Dipartimento di Matematica *"G. Peano"*, Università di Torino
Dipartimento di Matematica *"G.L. Lagrange"*, Politecnico di Torino

Ph. D. Thesis
Ph. D. Program in Pure and Applied Mathematics (35th cycle)

Dynamics and stability in Celestial Mechanics: from galactic billiards to Nekhoroshev estimates

Ph. D. Candidate: Irene De Blasi

Supervisors:
Prof. Susanna Terracini, Supervisor
Prof. Vivina Barutello, Co-Supervisor
Prof. Alessandra Celletti, Co-Supervisor

Referees:
Prof. Rafael de la Llave, Georgia Institute of Technology
Prof. Tere M-Seara, Universitat Politècnica de Catalunya

To my extended family

And the stars look very different today

David Bowie

Acknowledgements

This is most likely the last time I will write acknowledgments, so I'd better do it right.

As I usually say, despite a pandemic breaking out just at the beginning of my PhD course, the last three years have been the best of my academic - and maybe personal - life, and the merit of this goes to all the people mentioned below - and many others that I will guiltily forget.

The first and most important acknowledgment is of course for my supervisor, Prof. Susanna Terracini. It has been a true privilege and a honor to move my first steps in research under her guidance, learning *on the field* how to pass from studying other people's results to try and produce something new. Her brilliant intuitions have always amazed me, and her attitude towards Mathematics proves how one can really *enjoy* the research work. Honestly, I could not have hoped for anything more.

Many thanks to Prof. Vivina Barutello, with whom I have worked side by side during the past year, and co-supervisor of this work. Our daily collaboration taught me the value - and the pleasure - of writing and doing research together. I thank her also for being a precious confidant in such a delicate phase of my life, for comprehending patiently fears and doubts and always having the right answer to them.

Prof. Alessandra Celletti has first been my Master's thesis supervisor, and now a co-supervisor of my PhD's. With her I got a first taste of the meaning of research work, and I am overly grateful to have had the opportunity and the great privilege to continue working together.

Thanks to Prof. Christos Efthymiopoulos, for supervising my Master's thesis as well and for the fruitful and extremely interesting collaboration of the past three years. I have no doubt about the value of our conversations in deepening my understanding of many dynamical phenomena.

Going back in time, a special mention goes to Prof. Anna Maria Cherubini, my first supervisor and the first one to pronounce the words *Dynamical Systems* and *Celestial Mechanics* in my presence. It represented a true turning point for me, together with the advice and all the chats. I will do my best to deserve her trust.

To conclude, I would like to thank my family, for all the help, the love and the stubbornness in believing in me and in the path I chose.

My mother, without whose support I could never have gone so far. It's easy to go ahead when you know who is backing you. Thank you, from the bottom of my heart, for being such an extraordinary ally.

My sister Alessia, for the *surprise dinners* and all the creative ways she found to make me feel at home wherever I was. Go on without any doubt, because you *are* doing great.

My grandparents, the true pillars of this family, for always supporting me in every way they can, and for their commitment in understanding what *Celestial Mechanics* means.

The whole tribe of aunts, uncles and cousins I always return to: you know how much I cherish the time spent with you all, possibly together.

The members of this extended family that wrote about the 30% of this work - the *erased* part, unfortunately for you - with their paws on the keyboard.

Last but not least, Simone, my *partner in crime* in this adventure for a very long time. I do not know what to say that has not already been said just between ourselves. I cannot wait to know what is ahead of us.

As I had the opportunity to say recently, I really enjoyed to write this thesis, and I hope that whoever will read this work enjoys it as well.

Introduction

This thesis summarizes the results obtained as a part of the candidate's PhD research work described in the papers [1–6], and deals with the dynamical study of models of interest for Celestial Mechanics, using both analytical and numerical investigations. Given the heterogeneity of the subjects treated in this work, the detailed description of the specific analytical settings, as well as of the final outcomes, is left to the initial section of every chapter. We rather devote the current Introduction to a general illustration of the models and the approaches adopted, focusing our attention only on the main results. The variety of the themes treated, as well as of the methodologies used, is one of the peculiarities of Celestial Mechanics, and in general of Dynamical Systems, which make these fields such an attractive choice for those who decide to delve into them. Indeed, a wide range of techniques and approaches has been developed for this purpose.

The overall work is divided into two parts, each one based on a different type of dynamical system. In the first part, a particular class of billiards, designated as *galactic*, is taken into consideration (see Figure 1). Special emphasys is given to the case of the *galactic refraction billiard*, although the case of the *Kepler reflective billiard* (see for example [7, 8]) is treated as well in Chapter 3. In few words (see also Figure 1), the Kepler billiard can be described as a bounded domain inside which a massive body sits. If the energy of an inner particle is positive, it moves along Keplerian hyperbolæ, which, hitting the boundary, are reflected back. The major difference with refraction case is that, in this second model, the particle is no longer constrained in the interior of the boundary, and can exit following a specific refraction rule. The word *galactic* for such models emphasizes, of course, their relation with Celestial Mechanics: in the case of the reflective billiard, this link is correlated with the presence of a gravitational center. On the other hand, the refractive case has been inspired by the physical model of an elliptic galaxy having a central core, such as a black hole, in its center. This model has already been introduced, although with a different formalism, in [9], where, through both an analytical and a numerical approach, its chaoticity has been inferred from estimates of the mean Lyapunov exponent (for which it is worthwhile to mention [10] and the references within). While the reflective case has already been
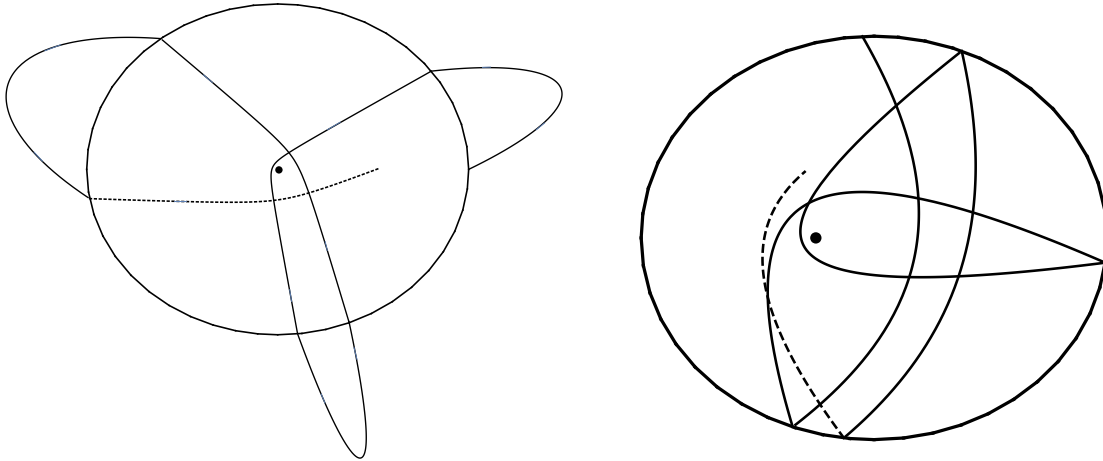
Fig. 1 Left: example of galactic refraction billiard. Right: galactic reflection billiard.

introduced and studied (for example, in [7, 8], where the problem of integrability is addressed), the refractive one presents some substantial novelties. In particular, in this case the particle is subject to the action of a discontinuous potential: indeed, taken a closed and regular domain $D \in \mathbb{R}^2$, we suppose that a harmonic oscillator-type potential acts in $\mathbb{R}^2 \setminus \overline{D}$, while a central mass μ dominates the dynamics in D. As a consequence, the resulting trajectories are concatenations of elliptic arcs outside D and, as long as the inner energy of the particle is positive, hyperbolic ones in D. On the boundary ∂D, a refraction Snell's law holds, deflecting the velocity vector while preserving the inner or outer energy. The reasons for requiring this type of junction rule can be retrieved by studying the problem from different perspectives: from a physical point of view, it represents a generalization of the classical Snell's law for light rays and can be obtained as a limit behavior when, taken two constant potentials and an interface of size ϵ which continuously connects them, we send $\epsilon \to 0$. On the other hand, from a variational point of view it is the result of a critical point argument, which translates, choosing suitable coordinates, in the conservativity of the first return map associated to the complete outer-inner dynamics. We stress that this kind of model belongs to the general family of refractive billiards, which can be constructed whenever a particle moves under the influence of a discontinuous potential and the transition is governed by some generalization of Snell's law (see also [11] and (∗∗) below), and that most of the techniques that we will describe can be generalized by modifying the potential. Similar models which take advantage of the relation between the rules of the motion of the light and the behavior of particles under a gravitational influence (the so-called *optical-mechanical analogy*) appear in more applied contexts, as in [12, 13]. Here, the metamaterials, that is, materials whose refraction index can be opportunely engineered, are used in order to have the light rays mimic the trajectories of a particle moving according to a gravitational force field. Other models where the two different potentials are coupled and the corresponding trajectories are patched are

the *Sun-shadow dynamics* (see [14]), where the pure Kepler's and the Stark's problem are paired, and the *inverse magnetic billiard* (cfr. [15]), which presents very strong analogies with our model from the point of view of the analytical setting. Here, straight lines in the interior of a closed billiard are concatenated with circular trajectories with proper Larmor radii, depending on a constant outer magnetic field. However, in both these models the transition between the two regimes does not deflect the velocity vector, as it happens in ours. The formalism used, as well as, where possible, the adopted techniques (for example, the construction of a first return map), are mainly inspired by the ample literature on classical Birkhoff billiards, where the particle is free to move along straight lines in the inner region and, hitting the boundary, is reflected back. Although this subject has been widely studied since the beginning of the XXth century (cfr. [16], and, for an accurate survey, the book [17]), at present there is a number of open problems (just to cite some examples, see [18]) and recent relevant advances (such as [19–21]). As for the study of the refractive case, to the best of our knowledge this has never been carried on. It is a worthwhile effort to begin its analysis with some basic results, such as the search for equilibrium trajectories and the derivation of their stability (which, as we will see, will require an accurate analysis). We will then pass on to the application of powerful analytical tools, deriving for example from KAM and Aubry-Mather theories (see [22], [23, 24] and [25]), and conclude with the study of the (possible) chaoticity of both the refraction and the Keplerian billiard.

This first part is organized into three main chapters:

- Chapter 1 starts by introducing the analytical model for the refractive galactic billiard. In particular, given a closed and regular domain $D \in \mathbb{R}^2$, we will study the motion of a particle subject to a discontinuous potential of the type

$$V(z) = \begin{cases} V_E(z) = \mathcal{E} + \dfrac{\omega^2}{2}\|z\|^2 & z \notin D \\ V_I(z) = \mathcal{E} + h + \dfrac{\mu}{\|z\|} & z \in D \end{cases}, \qquad (*)$$

where $\mathcal{E}, h, \omega$ and μ are positive constants. We study the zero-energy trajectories associated to this potential, supposing that, whenever an inner or outer orbit reaches the boundary, its velocity is refracted by following Snell's law

$$\sqrt{V_I(\tilde{z})} \sin \alpha_I = \sqrt{V_E(\tilde{z})} \sin \alpha_E, \qquad (**)$$

where $\tilde{z}$ is the transition point and α_I, α_E are respectively the angles that the inner and outer trajectories form with the normal unit vector to ∂D in $\tilde{z}$ (see Figure 2). It is clear that the shape of the boundary ∂D influences heavily the overall dynamics, even determining whether it is globally well defined or not. Of course, this influence extends on the form and properties of the associated first
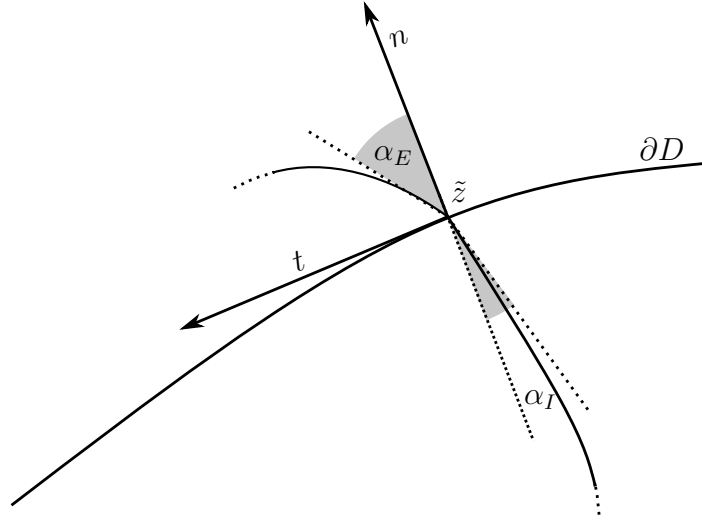
Fig. 2 Snell's law between inner and outer arcs. Here, t and n are respectively the tangent and the outward-pointing normal unit vectors to ∂D in $\tilde{z}$.

return map. The main purpose of this chapter is to study the local complete outer-inner dynamics around the so-called *homothetic periodic solutions*, obtained in the direction of a point $\bar{z} \in \partial D$ whenever $\bar{z} \perp \partial D$ and the half-line starting from the origin in the direction of $\bar{z}$ intersects ∂D only once (in the following sections of this work, and in particular in Chapter 3, these points will be denoted as *central configurations*). Indeed, in such cases the radial arcs in the direction of $\bar{z}$ are invariant for both the outer and inner [1] dynamics, and are not deflected by Snell's law as $\alpha_I = \alpha_E = 0$. We will show that, although, in general, the complete dynamics may not be well defined for every initial condition on ∂D, locally around the homothetic orbits it is possible to construct a well-defined and differentiable first return map (we stress that the construction of a first return map to study the continuous dynamics of the billiard is a quite classical technique deriving from the theory of Birkhoff billiards, see for example [17]). We can then study the linear stability of the equilibrium homothetic trajectories, defining a stability criterion based on the sign of the discriminant Δ of the characteristic polynomial associated to the Jacobian matrix of the first return map centered in the equilibrium itself. The quantity Δ depends on the physical parameters $\mathcal{E}, h, \omega, \mu$ and on the geometric features of the boundary ∂D up to the second order (in particular, on its curvature in the central configuration): it is then possible to study the stability of every homothetic trajectory under changes of the physical parameters, as well as of the boundary's geometry, searching for some kind of bifurcation phenomena. This is done explicitly in the case of a centered elliptic domain (namely, ∂D is an ellipse with center in the origin), for which asymptotic estimates are provided as well. In this case, the analytical results

---

[1]To treat the singularity at the origin, a regularization technique, such as the Levi-Civita one (see [26]) is in order.

are also substantiated with numerical computations, with the construction of a suitable Poincaré map. The study of the elliptic case is finally concluded with the analysis of non-homothetic *two-periodic brake orbits*, where two homothetic outer arcs are connected with an inner Keplerian hyperbola: under some conditions involving the ellipse's eccentricity and the inner parameters h and μ, we will be able to prove the existence of such kind of orbits by means of a *shooting method.*

- Chapter 2 is devoted to the analysis of the refractive dynamics for *close-to-circle* domains, whose boundaries can be parametrized in complex notation with smooth curves of the type

$$\gamma_\epsilon : \mathbb{R}_{/2\pi\mathbb{Z}} \to \mathbb{R}^2, \quad \gamma_\epsilon(\xi) = (1 + \epsilon f(\xi; \epsilon)) \, e^{i\xi},$$

where $f(\xi; \epsilon)$ is a suitable smooth function on $\mathbb{R}_{/2\pi\mathbb{Z}} \times [-C, C]$ and ξ is the polar angle. When $\epsilon = 0$, the domain ∂D is a circle, and the corresponding first return map is *globally well defined* and *completely integrable*: in particular, choosing suitable action angle variables (ξ, I), it can be expressed as a shift in ξ, while I is preserved under iterates, and the *rotation number* (see [25]) associated to every orbit is represented by the shift itself. On the other hand, when $\epsilon \neq 0$ and f is a generic smooth function, it is not possible anymore to give an explicit expression for the first return map associated to the perturbed domain, and neither establish any general facts upon its good definition. Nevertheless, is ϵ is small and f is regular enough, it is possible to construct invariant sets of initial conditions for which the perturbed dynamics, though not explicitly determined, is well defined and conservative. To do this, a sharper analysis of the variational properties of our map is in order, and the critical point argument that leads to the Snell's law assumes a crucial role. Whenever the perturbed dynamics is well defined and area-preserving, abstract theorems coming from the Aubry-Mather theory can be used to prove the existence of orbits with prescribed rotation numbers, including periodic ones. As a by-product of the invariance of perturbed orbit obtained from KAM theorem, a result concerning the existence of caustics (see for example [27, 28, 17] for the classical billiard case, and again [15] for the inverse magnetic billiard) is obtained.

- Chapter 3 will consider a wider class of domain shapes, focusing on the construction of a symbolic dynamics (see [29, 30] for a systematic dissertation on the subject) locally around the homothetic equilibrium trajectories. The overall reasoning follows the trails of [31, 32], where the existence of a symbolic dynamics is shown respectively for the classical N-centre and the anisotropic N-centre problem by discerning again between an outer and an inner dynamics. In our case, the inner dynamics is simpler, since the presence of a single Keplerian centre instead of a cluster of N massive bodies allows us to obtain more explicit

results; on the other hand, the presence of a completely different outer dynamics, and especially the replacement of a C^1 junction law with the Snell's refraction law, requires the application of new techniques, as for example the use of the Poincaré-Miranda fixed point theorem (see [33]). In our case, the existence of a symbolic dynamics for the refractive model is proved provided the energy jump h between the inner and outer potential is large enough and under very general hypotheses on D. In particular, it will be required that the function $\|\cdot\|$ restricted to ∂D admits at least two central configurations $\bar{P}, \hat{P}$ such that

- they are strict maxima or minima for the function;
- they are not antipodal, in the sense that the origin does not belong to the straight segment between $\bar{P}$ and $\hat{P}$.

An example of such kind of domain is represented by the centered non-circular ellipse already presented and analyzed in Chapter 1. We will denote as *admissible* the domains that satisfy such assumption. As we will see in the chapter, the non-antipodality property for the central configuration is a necessary condition for ensure that the inner dynamics is well defined, that is, that there exists a unique inner Keplerian hyperbola, satisfying suitable topological constraints (the so-called *(TnT) property*), connecting every pair of points of ∂D sufficiently close to the homothetics; indeed, to connect two point with a unique arc of this type it is necessary that they belong to neighborhoods of non-antipodal homothetics. The alphabet of our symbolic dynamics will then be given by the non-degenerate central configurations of ∂D which admit at least another non-antipodal and non-degenerate central configuration. Analogous results can be obtained, applying a somehow simplified reasoning, to the case of the Keplerian reflective billiard, under the same admissibility hypotheses.

The chapter continues going through the consequences of the existence of a symbolic dynamics for admissible domains, combining it with the stability analysis performed in Chapter 1, to investigate the possible presence of a chaotic regime in both our billiard models. In particular, under stricter assumptions on the nondegeneration of our central configurations, by the joint application of the results obtained in these Chapters 1 and 3, one can infer that, for large enough energies, every homothetic trajectory used as a letter for our symbolic dynamics is a saddle equilibrium point; as a consequence, one can construct infinitely-many heteroclinic connections between every pair of them. The presence of multiple heteroclinics, along with the existence of a symbolic dynamics, is a strong indicator of complex behavior. That goes in the direction of proving that our systems are chaotic, although not being enough. The final step can be performed by adapting a classical argument by Kozlov (see [34]), in order to show that, if h is large enough, then there are no analytic first integrals associated to our system.

Again, the reflective and refractive case go side by side in this analysis, suggesting that, under the admissibility conditions described in Chapter 3 along with the nondegeneration of at least one central configuration and provided that h is large enough, both the billiards are analytically non-integrable. This conclusion, when applied to the case of a centered elliptic domain, can lead to different consequences depending on the case: for the refractive billiard, this represent an analytical justification for what has already been observed, numerically, in the simulations presented in Chapter 1. As for the reflective billiard, this result represents an interesting complement to [7], where the integrability of the focused elliptic billiard (namely, with an elliptic boundary with one of the two foci in the origin) is proved; note that a focused elliptic billiard does not satisfy the admissibility property, having only two central configurations which are antipodal.

The second part of this work is focused on providing stability estimates for geocentric satellites and orbiting bodies, taking into consideration reliable models, within a Hamiltonian setting. Knowing the long-term behavior of the bodies orbiting around Earth is a crucial goal in Celestial Mechanics, from both a theoretical and a practical point of view, especially in view of estimating the orbital survival times of operating satellites or space debris (see for example [35–37]). In particular, this last theme has been an important subject of recent studies and source of concern for scientists and space agencies: at present, more than 30000 bodies have been recorded and are constantly monitored, and, based on the European Space Agency simulations, the actual number of orbiting objects larger than 1 cm in size is likely over one million [38]. As the altitude strongly influences the behavior of a satellite, it is convenient to distinguish the geocentric orbits into three main subgroups: LEO (Low Earth Orbit), up to 2000 km from the surface of Earth, MEO (Medium Earth Orbit), between 2000 and 30000 km, and GEO (Geosynchronous Earth Orbit), about 42164 km from Earth's center, where the geostationary orbits lie (see [39–43] for a thorough study of the dynamics of satellites in different regimes). While in LEO the dynamics is highly influenced by the atmospheric drag (cfr. [44]), in MEO and beyond the dissipation due to the atmosphere is negligible, and one can consider a model that takes into consideration only the gravitational effects of Earth, Sun and Moon, leading to a conservative system. In this region, the shape of the Earth must be considered to find its actual gravitational potential (the so-called *geopotential*, see [45] where it is derived and expressed in terms of spherical harmonics), and the lunar and solar gravitational attractions act as a third-body perturbations. The study presented in this work is mainly focused on object in MEO, and takes into consideration, sometimes underlining analogies and differences, two models:

viii

- the so-called J_2-*model*, where only the effects of the geopotential, truncated up to order $(2,0)$, are taken into account. Physically, this means that the oblateness of Earth is considered;

- the *secular geolunisolar model*, which includes, in addition to the J_2 effects, the gravitational attractions of the Sun and the Moon. We mention [46–49] and the references therein to show the relevance of such model to describe all the important dynamical behaviors occurring in the MEO region. Here, *secular* stands for *averaged* over short-period terms.

The mathematical tools employed to perform our stability analysis are based on perturbative methods (see for example [50]), and, more precisely, on *normal form theory* (see [51] for a detailed description) and the celebrated *Nekhoroshev theorem* (see [52, 53]), here presented in its nonresonant version, as stated by Pöschel in [54]. In practice, we will describe and apply two methods, based on the above theoretical tools, ensuring that, as long as the considered time period is within suitable computed bounds, the orbital elements of the body do not undergo large variations that could produce a drastic change in the overall orbit. We refer to these procedures as *semi-analytical*, in the sense that, while rigorous analytical methods are followed, the coefficients of the Hamiltonian functions that we will use are computed numerically.

In general, starting from a generic Hamiltonian $\mathcal{H}(\mathbf{I}, \mathbf{u})$, expressed in action-angle coordinates $(\mathbf{I}, \mathbf{u}) \in \mathbb{R}^n \times \mathbb{T}^n$ (n being the number of degrees of freedom of the system), we refer to *normal form* as a new Hamiltonian function, obtained through a series of canonical transformations, of the form

$$\widetilde{\mathcal{H}}(\mathbf{I}, \mathbf{u}) = h_0\left(\mathbf{I}, \mathbf{u}\right) + h_1\left(\mathbf{I}, \mathbf{u}\right),$$

where h_0 enjoys suitable properties, depending on the model, and h_1, the *remainder*, is *small* with respect to h_0. As an example, if h_0 depends only on the actions, then by means of Hamilton's equation one can observe that $I_1, \ldots, I_n$ are *quasi-integrals* of the motions, in the sense that their variation depends only on the small remainder.

This is for example the starting point of Nekhoroshev theorem, which, under suitable non-degeneracy hypotheses on h_0, ensures the stability of the action variables, in the sense specified before, for times which are exponentially long in the inverse of the remainder's size. In the original paper, the non-degeneracy hypothesis required by Nekhoroshev is the so-called *steepness* condition, a geometric assumption rather complex to verify in practice. Nevertheless, there are sufficient conditions implying steepness, such as *convexity, quasi-convexity* and *three-jet non-degeneracy* (see [55]), which are simpler to check; recent studies [56, 57] extend the class of sufficient conditions to more general properties. Applications of Nekhoroshev theorem to systems of interest for Celestial Mechanics can be found in [58] in the case of the three body problem and in [59] for the Trojan asteroids.
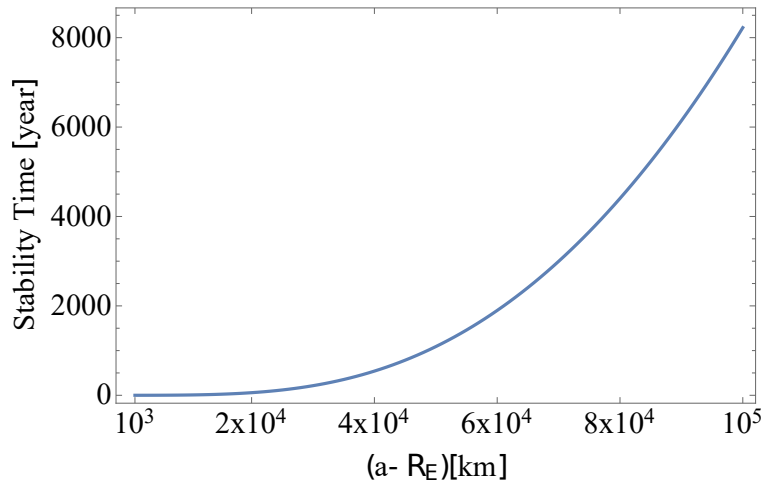
Fig. 3 Stability time in the J_2 model for the semimajor axis a for different initial values. Here the allowed excursion is $\Delta a = 0.1 R_E$, where R_E is the Earth's radius. Figure taken from [4].

This second part is again organized into two chapters, which describe the reasonings and the results obtained by applying two different methods to study the stability of our satellites:

- Chapter 4 takes into considerations both the J_2 and the geolunisolar models, following the trails of [60]. In general, the reasoning adopted to obtain stability estimates will be as follows: given the Hamiltonian $\mathcal{H}(\mathbf{I}, \mathbf{u})$, in our case, either the J_2 or the geolunisolar one, we will search for a normal form such that the term h_0 admits an integral of motion (which will depend on the model; in general, let us call it A). At this point, given a suitable functional norm $\| \cdot \|_D$ over a bounded domain D in the variables, by the mean value theorem one has that

$$ t \le T_{stab} = \frac{\Delta A}{\|h_1\|_D} \quad \Longrightarrow \quad |A(t)| \le \Delta A. $$

We obtain then a lower bound for the stability time, up to which the variation of the quasi-integral A is controlled.

The above reasoning is employed for the J_2 model to estimate the stability of the satellite's semimajor axis, while, in the secular geolunisolar system, it is applied to the *Lidov-Kozai* quasi integral $A = \sqrt{1 - e^2}\,(1 - \cos i)$ (see [61]). Figure 3 and Table 1 show the stability time, in years, obtained by means of the above analysis.

As a final result presented in the chapter, the three conditions of convexity, quasi-convexity and three-jet non-degeneracy, sufficient to guarantee the steepness of h_0, are checked on our two models. From the numerical investigation, it emerges that, while the J_2 normalized Hamiltonian is only three-jet non-degenerate, the geolunisolar one is quasi-convex: this is a nontrivial fact, translating in the

Altitude	Stability time
3000 km	$4.61551 \cdot 10^{13}$
20000 km	$2.20144 \cdot 10^{12}$
35790 km	$3.51266 \cdot 10^{10}$
50000 km	$1.07263 \cdot 10^{8}$
100000 km	$3.36609 \cdot 10^{4}$

Table 1 Stability time for different altitudes in the domain $(e, i) \in D = [0, 0.1] \times [0, 0.1]$ for the secular geolunisolar model.

evidence that the addition of the Moon and the Sun to our model removes the degeneracy from the J_2 model.

- Chapter 5 resumes the algorithm and the results of the application of Nekhoroshev theorem in its nonresonant formulation to the secular geolunisolar model. This version of the theorem, which can be applied to domains in the variables which are *far* from the resonances, does not require any particular assumption on the nondegeneracy of h_0, and is again based on the *smallness* of the remainder term. In particular, whenever the size of h_1 is less than a certain threshold value (depending on h_0 and the size of the action's domain), we are again able to bound the variation of the actions for a certain interval of time, which turns out to be exponentially long with respect to the inverse of the remainder's size itself; it is therefore crucial to control the growth of the term h_1 during the normalization process. While the study in Chapter 4 was focused on small inclinations and eccentricities, in this case a wider domain, including the orbital parameters of most of the operating satellites, is considered; as for the semimajor axis, relevant results have been obtained for MEO distances between 10000 and 20000 km. Figure 4 presents the stability time obtained in such cases, and requires some words to be correctly interpreted. The white regions in every graphic represent the values of eccentricity and inclination for which the nonresonant Nekhoroshev theorem (at least, following the presented algorithm) does not apply because the remainder is too large; the color scale in the remaining region denotes the obtained stability time. As one can see, the domain of applicability of the theorem tends to shrink when a grows, while the stability times diminish: a heuristic explanation for this phenomenon will be provided at the end of the chapter. Moreover, as we will see, the *inclination-dependent* resonances of the system (see for example [62]) play a crucial role, leading to the presence of domains in the orbital parameters for which the nonresonant hypothesis is not satisfied. In such regimes, it is worth to use Nekhoroshev theorem in its complete formulation, performing a systematic and accurate study of the geometry of the resonances that characterize of the system.
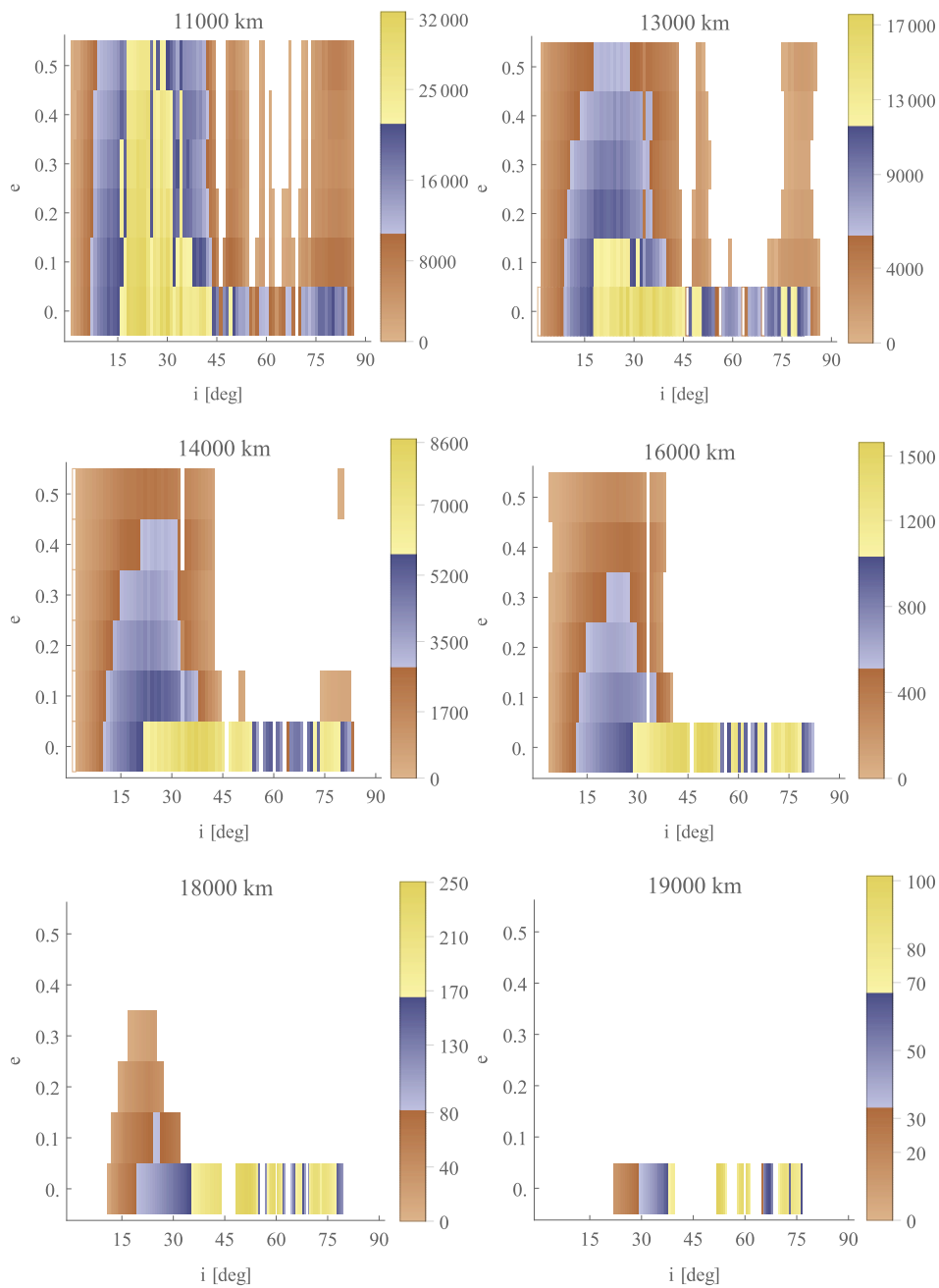
Fig. 4 Stability times (in years) computed by applying a semi-analytical algorithm based on the nonresonant Nekhoroshev theorem. On the top side of every graphic the corresponding value of the semimajor axis is specified. Figure taken from [5]

Now that the models, as well as the problems related, have been described, it is clear that, although both of them come from Celestial Mechanics, they present fundamental differences in the construction and in the methods of investigation. Nevertheless, there is a *fil rouge* that connects the research presented in this whole work. First of all, we are dealing with models that can be considered as particular examples of $n-$body problem, where the gravitational interaction of the particle with many different massive objects has been considered. In particular, in the refractive galactic billiard the influence of the central mass is counterbalanced by the gravitational interaction of our test particle with the elliptic galaxy's mass distribution, which translates in a harmonic oscillator-type potential. As for the geocentric motion described in the second part, we have again to take into account the gravitational field generated by a body with a proper volume, namely, the oblate Earth, as well as the Sun's and Moon's point-mass attractions.

Furthermore, in both models the dynamics of the particle is highly influenced by its distance from the center of the corresponding reference frame, and different regimes can be detected: in the refractive billiard the separation between them (harmonic and hyperbolic motion) is of course neat; in the satellite's case, it is less evident and definitely smooth, but it is clear from the results obtained that, increasing the altitude, we pass from a regime in which the Earth's attraction dominates to a condition where the effects of Sun and Moon are more and more relevant.

As for the problems addressed, the issue of stability plays a central role in both parts, representing the main argument of study in the second one. On the other hand, the motion around particular equilibrium trajectories, for which an accurate stability analysis has been provided, is crucial in Part I, and in particular in Chapters 1 and 3.

Contents

Part I

Dynamical results on galactic billiards

Chapter 1

Stability and bifurcations in galactic refraction billiards

1.1   Introduction and description of the model

According with Bertrand's theorem, among all central forces with bounded trajectories, there are only two cases with the property that all orbits are also periodic: the attractive inverse-square gravitational force and the linear elastic restoring one governed by Hooke's law.

In this first part we consider a dynamical system of physical interest, where such two forces act in two complementary regions of the space; a Keplerian attractive center sits in the inner region, while a harmonic oscillator is acting in the outer one. In addition, the two regions are separated by an interface Σ, where a Snell's law of ray refraction holds. Hence trajectories concatenate arcs of Keplerian hyperbolæ with harmonic ellipses, with a refraction at the boundary. When the interface also has a radial symmetry, then the system is integrable; in this chapter we will study the effect of symmetry-breaking on the stability and bifurcation of periodic orbits. Chapter 2 will be devoted to the analysis, in terms of KAM and Mather theories, of systems with close to circular interfaces.

Our first motivation comes from an elliptical galaxy model with a central core, of interest in Celestial Mechanics [9], which deals with the dynamics of a point-mass particle P moving in a galaxy with a harmonic biaxial core, in whose center there is a Black Hole. As known, Black Holes appear when, caused by gravitation collapse, the mass densities of celestial bodies exceeds some critical value, and act as attractors of both matter and light. Following the relativistic equivalence between energy and matter, the critical behaviour in the presence of Black Holes has been the recent object of investigation related with optical properties of metamaterials [12]. In this framework, light behaves in space as in an optical medium having an effective refraction index which incorporates the gravitational field and may have a discontinuity accounting for

the inhomogeneity of the material itself. Therefore, our type of systems are of interests in view of possible applications in engineering artificial optical devices that control, slow and trap light in a controlled manner [63].

Going back to our model in Celestial Mechanics, supposing the axes of the galaxy's mass distribution being orthogonal, we can then use a planar reference frame whose x and y axes are the galaxy's ones, while the BH is at its origin. In the described reference frame, we denote with $z \in \mathbb{R}^2$ the particle's coordinates. Here, the plane $\mathbb{R}^2$ is divided into two regions, according to whether the gravitational effects of the galaxy's mass distribution or of the BH dominate. The BH's domain of influence is set to be a generic regular domain $\mathbf{0} \in D \subset \mathbb{R}^2$, and the particle moves on the plane under the influence of inner and external potentials

$$V(z) = \begin{cases} V_I(z) = \mathcal{E} + h + \dfrac{\mu}{\|z\|} & \text{if } z \in D \\[2mm] V_E(z) = \mathcal{E} - \dfrac{\omega^2}{2}\|z\|^2 & \text{if } z \notin D, \end{cases} \tag{1.1.1}$$

with $\mathcal{E}, \mu, \omega > 0$ and $\mathcal{E} + h > 0$, while the behaviour of the particle's trajectory while it reaches the boundary $\partial D = \Sigma$ is ruled by a generalization of Snell's law (i.e. the conservation of the tangential component of the velocity through the interface, see Section 1.2). The motion of P will take place inside the *Hill's region*

$$\mathcal{H} = \{z \in \mathbb{R}^2 \quad | \quad V_E(z) \geq 0\};$$

for computational reasons, we impose $2\mathcal{E} > \omega^2$ to ensure that the circle of radius 1 is contained in $\mathcal{H}$.
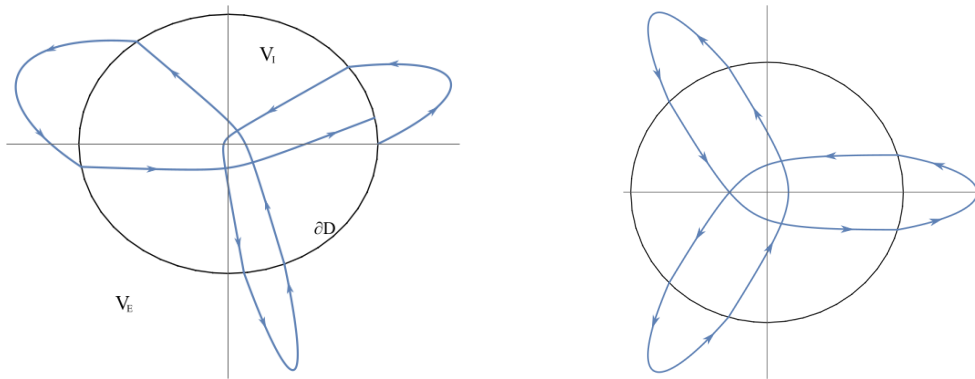
Fig. 1.1 Left: trajectory for the general case. The inner and outer arcs are connected by a refraction Snell's law. Right: a period three orbit for an elliptic domain with eccentricity $e = 0.3$ and physical parameters $\mathcal{E} = 2.5$, $\omega = \sqrt{2}$, $h = 0.1$ and $\mu = 1$.

Our aim is to study the trajectories of zero energy of the system whose potential is defined as in (1.1.1), in relation with the geometry of the boundary ∂D, taking $\mathcal{E}, h, \mu$ and ω as parameters. The study of this kind of orbits is performed by means of the *broken geodesics method* (see [64]), where the inner and outer dynamics are considered

separately and then the corresponding arcs are connected on ∂D. Though usually ∂D will be elliptical shaped, most of our results just involve its geometrical features, namely its tangent and curvature. If for example ∂D intersects orthogonally both coordinate axes, as we will explain in a moment there will be two collision homothetic periodic solutions in the horizontal and vertical directions. Taking advantage of Levi-Civita regularisation ([26]), we may indeed assume motions to be extended after a collision with the gravity center by complete reflection. We are concerned with the stability (dynamical and structural) of such periodic trajectories and their bifurcations in dependence of the system parameters. We shall focus in particular on bifurcations of period-two brake trajectories, where the term *brake* refers to orbits admitting a point with vanishing velocity. In the case of an elliptic domain we will be able to describe the full picture, in dependence of the physical parameters.

Although the presence of periodic orbits depends in general on the global geometry of D, as well as on the physical parameters $\mathcal{E}, h, \mu, \omega$, there is a class of them whose existence is guaranteed by particular local conditions on ∂D: this is the case with the *homothetic orbits*, namely, of the form $z(t) = \varphi(t)\mathbf{v}$, where $\mathbf{v} \in \mathbb{R}$ is a given configuration vector and $\varphi(t) : [0, \infty) \to \mathbb{R}$ is a scalar function. Let us suppose that ∂D is a curve of class C^2, and, with an abuse of notation, identify any point $\mathbf{q} \in \partial D$ with its position vector $\mathbf{q} - \mathbf{0}$. Now take $\mathbf{p} \in \partial D$, and suppose that it satisfies the two conditions[1]

(i) $\mathbf{p} \perp \partial D$,

(ii) the ray starting from $\mathbf{0}$ in the direction of $\mathbf{p}$ does not intersect   (1.1.2)

            ∂D more than once.

In this case, the system admits a collision homothetic orbit in the direction of $\mathbf{p}$, which we denote by $\bar{z}_p(t)$. Condition (i) is necessary to assure that the orbit is not deflected by Snell's law when crossing the interface ∂D. As a consequence, conditions (1.1.2) not only imply the existence of the homothetic orbits, but also the existence and uniqueness of inner and outer arcs in some neighbourhoods, as well as the good definition of the refraction law in its vicinity (see Section 1.2). Under the hypotheses (1.1.2), it makes then sense to study the linear stability of $\bar{z}_p$ under the regularised flow by considering its Jacobian matrix $\mathcal{M}$ centerd in $\bar{z}_p$. Taking $\Delta(\bar{\xi})$ as the discriminant of its characteristic polynomial, the following Theorem provides a full characterisation of stability in terms of the physical parameters and the local properties of ∂D in $\mathbf{p}$.

Theorem 1.1.1. *Let us suppose $\partial D = \gamma(I)$, with $\gamma \in C^2(I)$, $\gamma : \xi \in I \subset \mathbb{R} \mapsto \gamma(\xi) \in \mathbb{R}^2$. Take $\bar{\xi} \in I$ such that, setting $\mathbf{p} = \gamma(\bar{\xi})$ and identifying it with its position vector, it satisfies (1.1.2). Denote $k(\bar{\xi})$ the curvature of γ at $\gamma(\bar{\xi})$, and denote with $\bar{z}_p$ the homothetic orbit in the direction of $\mathbf{p}$. Let the inner and outer potentials be defined as in (1.1.1). Therefore*

- *if $\Delta(\bar{\xi}) > 0$, then $\bar{z}_p$ is linearly unstable;*

---

[1]In the following chapters, we will refer to such points as *central configurations* of our problem.

- *if $\Delta(\bar{\xi}) < 0$, then $\bar{z}_p$ is linearly stable,*

where we have denoted:

$$\Delta(\bar{\xi})(\mathcal{E}, h, \mu, \omega; \gamma) = ABCD$$

$$A = \frac{16}{\mathcal{E}^2 \mu^2} \left(\sqrt{V_I(\gamma(\bar{\xi}))} - \sqrt{V_E(\gamma(\bar{\xi}))} \right) \left(\|\gamma(\bar{\xi})\| k(\bar{\xi}) - 1 \right),$$

$$B = \mathcal{E} - \left(\|\gamma(\bar{\xi})\| k(\bar{\xi}) - 1 \right) \left(\sqrt{V_I(\gamma(\bar{\xi}))} - \sqrt{V_E(\gamma(\bar{\xi}))} \right) \sqrt{V_E(\gamma(\bar{\xi}))},$$

$$C = -\mu \sqrt{V_E(\gamma(\bar{\xi}))} + 2\|\gamma(\bar{\xi})\| B \sqrt{V_I(\gamma(\bar{\xi}))},$$

$$D = \mu + 2\|\gamma(\bar{\xi})\| \left(\|\gamma(\bar{\xi})\| k(\bar{\xi}) - 1 \right) \sqrt{V_I(\gamma(\bar{\xi}))} \left(\sqrt{V_I(\gamma(\bar{\xi}))} - \sqrt{V_E(\gamma(\bar{\xi}))} \right).$$

When ∂D is an ellipse, the stability of the four homothetic orbits, which are parallel to the coordinate axes, can be studied explicitly in terms of the physical parameters of the problem and the eccentricity $0 \leq e < 1$ of the ellipse. By symmetry, only the homothetic orbits intersecting the positive directions of the axes, which we denote with $\bar{z}_0$ and $\bar{z}_{\pi/2}$, are considered.

Corollary 1.1.2. *If ∂D is an ellipse, explicit expressions for $\Delta(0)$ and $\Delta(\pi/2)$ are provided in* (1.6.3)*, leading to the complete description, in terms of the physical parameters and of the eccentricity, of all stability regimes.*

As the expressions of the $\Delta(0)$ and $\Delta(\pi/2)$, though explicit, include the many different parameters in a rather intricated formula, the general study of their sign can be arduous and we shall perform it numerically in general and analytically in some specific regimes; indeed, an asymptotic analysis for $e \to 0$ can be done, leading to a rather simple stability criterion for small eccentricities. In particular, we have the following result for small eccentricities.

Corollary 1.1.3. *If* $\dfrac{\sqrt{\mathcal{E} + h + \mu}}{\mu} < \dfrac{\sqrt{2\mathcal{E} - \omega^2}}{2\sqrt{2\mathcal{E}}}$*, then, for small eccentricities, $\bar{z}_0$ is stable and $\bar{z}_{\pi/2}$ is unstable. Symmetrically, if* $\dfrac{\sqrt{\mathcal{E} + h + \mu}}{\mu} > \dfrac{\sqrt{2\mathcal{E} - \omega^2}}{2\sqrt{2\mathcal{E}}}$*, then, for small eccentricities, $\bar{z}_0$ is unstable and $\bar{z}_{\pi/2}$ is stable.*

A similar asymptotic analysis, which holds for arbitrary eccentricities, can be performed for high values of h or μ and $\mathcal{E}$ (see Proposition 1.6.1): fixing all the parameters but h (resp. μ), if h (resp. μ) is large enough, both $\bar{z}_0$ and $\bar{z}_{\pi/2}$ are unstable homothetic orbits. In such cases, with the additional hypothesis of a good definition of the dynamics on the whole ellipse, we can infer the existence of an intermediate non-homothetic stable periodic orbit with exactly two distinct crossings of ∂D. Furthermore, if $\mathcal{E}$ is large enough, one has that $\bar{z}_{\pi/2}$ is unstable, while the stability of $\bar{z}_0$ is determined by the value of μ, in the sense that there is a threshold value $\bar{\mu}(\omega, h, e)$ such that if $\mu < \bar{\mu}$ $\bar{z}_0$ is unstable, while if $\mu > \bar{\mu}$ its stability is reversed.

As the eigenvalues of the Jacobian matrix $\mathcal{M}$ depend smoothly on the physical parameters of the problem, bifurcation phenomena (see for example [65]) occur whenever a variation of any of $\mathcal{E}, h, \mu$ or ω determines a change in the sign of Δ: Section 1.7 provides concrete examples of such transitions.

The second class of periodic orbits on which this chapter is focused is represented by the two-periodic brake orbits, where two homothetic outer arcs are connected by an inner Keplerian hyperbola.

Theorem 1.1.4. *Suppose that ∂D is an ellipse with eccentricity $e \in (0, 1/\sqrt{2})$, and consider $\mathcal{E} > 0, \omega > 0$ such that $2\mathcal{E} > \omega^2$. Then there are $\bar{\mu} = \bar{\mu}(\mathcal{E}, \omega, e)$ and $\bar{h} = \bar{h}(\mathcal{E}, \omega, e, \mu)$ such that, if $\mu > \bar{\mu}$ and $h > \bar{h}$, then the dynamics admits at least two nontrivial brake orbits of period two.*

Nontrivial here stands for non homothetic. The existence of this type of periodic orbits on the ellipse is a significant fact, which distinguishes the strictly elliptic case, namely, with $e \neq 0$ from the circular case: we have indeed that, while in the latter there are infinitely many homothetic orbits, there is no possibility to have a nontrivial two periodic brake trajectory.

As in the case of Theorem 1.1.1, also Theorem 1.1.4 admits an extension for general curves which share with the ellipse a common behaviour near to the homothetic orbits up to the second order and a particular type of global convexity property with respect to the hyperbolæ. In particular, we shall define a class of boundaries γ for which the inner arcs are globally well defined.

Definition 1.1.5. *We say that the domain D is **convex for hyperbolæ for fixed $h, \mathcal{E}$ and μ** if every Keplerian hyperbola with energy $\mathcal{E} + h$ and central mass μ intersects ∂D at most in two points.* [2]
*The domain D is **convex for hyperbolæ** if the previous condition holds for every positive $\mathcal{E}, h$ and μ.*

As we will see in Remark 1.6.5, Theorem 1.1.4 remains true whenever a domain D:

- is convex for hyperbolæ;

- is everywhere transverse to the radial direction, in the sense that there are not rays starting form the origin and tangent to ∂D;

- $\partial D = \gamma(I)$, $\gamma \in C^2(I)$ and $\gamma(\xi) = \gamma_0(\xi) + \gamma_1(\xi)$, where γ_0 parametrises the ellipse and γ_1 has the same symmetry of the ellipse and is such that

$$\gamma_1(k\pi/2) = \dot{\gamma}_1(k\pi/2) = \ddot{\gamma}_1(k\pi/2) = (0,0) \quad \text{for } k = 0, 1, 2, 3.$$

---

[2]We recall that a Keplerian hyperbola is the solution of the Cauchy problem

$$\begin{cases} z''(s) = -\mu z(s)/\|z(s)\|^3, \quad \|z'(s)\|^2/2 = \mathcal{E} + h + \mu/\|z(s)\| \\ z(0) = p_0, \quad z'(0) = v_0 \end{cases}$$

for some initial conditions $p_0, v_0 \in \mathbb{R}^2$.

In the case of the ellipse, our analytical study is enriched by a numerical investigation, presented in Section 1.7, where the behaviour of the dynamics in different cases of interest is described. Of special interest is the evidence, under particular circumstances, of diffusive orbits, even for very small eccentricities (i.e., near to the circular case, which is integrable), which is a strong sign of chaotic behaviour.

The figures of the current chapter are taken from [1].

1.1.1   Analogies and differences with Birkhoff billiards

The model investigated in this work falls into the category of billiards and it is worth to focus our attention on the possible analogies and the fundamental differences with the classical Birkhoff case (see [16]). First of all, the rays are curved by the gravitational force inside D; moreover, reflection at the boundary is replaced by an excursion in the outer region in between two refractions. Similarly to billiards, our model can be described by an area preserving map of the cylinder, but, as we shall show in Chapter 2, the twist condition (see [25]) may be violated even in the simplest case of a circular domain. It should be noted that refraction imposes a new constraint, because the interface can only be crossed outwards when the inner arc is transverse enough to the boundary (cfr Section 1.2).

There is a wide literature on Birkhoff billiards, with recent relevant advances (see the book [17] and papers [19–21, 66]), including some cases of composite billiard with reflections and refractions [11], also in the case of a periodic inhomogenous lattice [67]. Special mention should be paid to the work on inverse magnetic billiards, where the trajectories of a charged particle in this setting are straight lines concatenated with circular arcs of a given Larmor radius [15, 15]. Let us add that, compared with the cases quoted above, additional difficulties arise because the corresponding return map is not globally well defined and from the singularity of the Kepler potential.

As a concrete example on how the considered model presents intrinsic analogies with classical billiards, as well as important differences, let us consider the case of an elliptic domain and take again Corollary 1.1.3. In classical elliptic billiards of every eccentricity (see for example Section 4 in [17]), the straight orbit segment corresponding to the major axis determines always a saddle point for the associated billiard map, while the one coinciding with the minor axis is a center. A similar behaviour can be observed in our refractive model when the domain is an ellipse with small eccentricity, but with a fundamental difference: while this model admits the same equilibrium orbits of the classical billiard, their stability depend on the value of the physical parameter $\mathcal{E}, h, \omega, \mu$.

1.2 Preliminaries: Jacobi distance and Snell's law

As already pointed out in the Introduction, the Snell's law used to rule the junction between outer and inner arcs relies on a variational (critical point) argument. To state it properly, it is necessary to give some basic definitions and results, which are treated in more details in [31]. As we will see in Chapter 2, the below discussion is intimately related to the good definition of our first return map as well, which can be retrieved, as in the classical Birkhoff case, by taking into account a suitable *generating function.*

1.2.1 Jacobi length and distance

Let us start by taking a generic fixed-ends problem of the type

$$\begin{cases} z''(s) = \nabla V(z(s)) & s \in [0,T] \\ \dfrac{1}{2}\|z'(s)\|^2 - V(z(s)) = 0 & s \in [0,T] \\ z(0) = z_0, z(T) = z_1 \end{cases} \tag{1.2.1}$$

with z_0, z_1 in a suitable subset of $\Omega \subset \mathbb{R}^2$, and $V(z)$ a generic potential of class $C^1(\Omega)$ such that $V(z) > 0$ almost everywhere, The reasoning described in the current section, although treated in general, will be then applied to the potentials V_E and V_I: as for the inner case, the Levi-Civita regularisation technique (described in details in Section 1.3) makes us retrieve the regularity of the solutions of the corresponding fixed-ends problem, allowing to proceed with the subsequent demonstrations also for V_I, at least in a regularised sense.

Let us take $z_0, z_1 \in \Omega$ such that the following assumption holds.

Assumption 1.2.1. *Let us assume that, given the problem*

$$\begin{cases} z''(s) = \nabla V(z(s)) & s \in [0,T] \\ \dfrac{1}{2}\|z'(s)\|^2 - V(z(s)) = 0 & s \in [0,T] \\ z(s) \in \Omega & s \in [0,T] \\ z(0) = z_0, z(T) = z_1 \end{cases} \tag{1.2.2}$$

for some $T > 0$, there exists a solution $\bar{z}(\cdot) \doteq z(\cdot; z_0, z_1) \in C^2([0,T])$ (possibly in a regularized sense). This solution is supposed to be unique in a suitable homotopy class.

With reference to the outer and inner potential, suitable conditions on the endpoints for which the existence and uniqueness hold are guaranteed, case by case, in Sections 1.3, 2.3.1, 2.4.1 and 3.2. The definition of a suitable homotopy class will be used, in Chapters 2 and 3, to ensure the uniqueness of the inner arc with fixed endpoints.

Definition 1.2.2. *Given z_0, z_1 satisfying Assumption 1.2.1, define the* Jacobi length *of the path $\bar{z}$ as the quantity*

$$L\left(\bar{z}\right) \doteq \int_0^T \left\|\bar{z}'(s)\right\| \sqrt{V\left(\bar{z}(s)\right)} ds$$

and the Jacobi distance *between z_0 and z_1 as*

$$d\left(z_0, z_1\right) \doteq L\left(\bar{z}\left(\cdot; z_0, z_1\right)\right).$$

Remark 1.2.3. *In view of Definition 1.2.2, we observe that:*

- *whenever the arc connecting z_0 and z_1 is unique, the Jacobi distance $d\left(z_0, z_1\right)$ is well defined and differentiable;*

- *the functions d is not a proper distance: as a matter of fact, when for example $V = V_E$ and $z_0 = z_1$, $d\left(z_0, z_1\right)$ represents the non-zero Jacobi length of the homothetic outer arc. An analogous observation can be carried out when $V = V_I$ and again the endpoints coincide (see Section 1.3 for more details);*

- *the Jacobi length is invariant under reparametrizations of the path $\bar{z}$.*

The Jacobi distance, and in particular its partial derivatives with respect to the endpoints, represents the starting point to give a variational characterization of the Snell's law.

Although in the previous discussion the Jacobi length has been defined only for the solution of Problem (1.2.2), the corresponding quantity can be computed for any path contained in the set

$$H_{z_0 z_1} \doteq \left\{ z \in H^1\left([0,T], \mathbb{R}^2\right) \text{ for some } T > 0 \,\middle|\, \begin{array}{l} z(0) = z_0, z(T) = z_1, \\ V(z(s)) \geq 0 \text{ for all } s \in [0,T] \end{array} \right\},$$

namely, for every curve in H^1 that connects z_0 to z_1 lying in the Hill's region associated to the potential V.

There is a strong relation between solutions of Problem (1.2.2) and critical points of L in $H_{z_0 z_1}$, which passes, naturally, by the resolution of the Euler-Lagrange equations.

Lemma 1.2.4. *A path $z \in H_{P_1 P_2}$ is a critical point of $L(\cdot)$ if and only if it is a solution of the Euler-Lagrange equations*

$$\frac{d}{ds}\left(\sqrt{V(z(s))}\frac{z'(s)}{\|z'(s)\|}\right) - \frac{\|z'(s)\|}{2\sqrt{V(z(s))}}\nabla V(z(s)) = 0 \quad a.e. \ in \ [0,T]. \qquad (1.2.3)$$

Proof. One has that z is a critical point for the Jacobi length if and only if for every $v \in H_0^1([0,T])$ it results $dL(z)[v] = 0$. This is equivalent to require that the following

chain of equalities holds:

$$
0 = \frac{d}{d\epsilon}L(z + \epsilon v)_{|\epsilon=0} = \left(\frac{d}{d\epsilon} \int_0^T \|z'(s) + \epsilon v'(s)\| \sqrt{V(z(s) + \epsilon v(s))} \, ds \right)_{|\epsilon=0}
$$

$$
= \int_0^T \frac{\sqrt{V(z(s))}}{\|z'(s)\|} z'(s) \cdot v'(s) + \frac{\|z'(s)\|}{2\sqrt{V(z(s))}} \nabla V(z(s)) \cdot v(s) \, ds
$$

$$
= \int_0^T \left[-\frac{d}{ds}\left(\sqrt{V(z(s))} \frac{z'(s)}{\|z'(s)\|} \right) + \frac{\|z'(s)\|}{2\sqrt{V(z(s))}} \nabla V(z(s)) \right] \cdot v(s) \, ds,
$$

where in the last equation an integration by parts has been employed. As the identity must be true for every $v \in H_0^1([0,T])$, one has that the Euler-Lagrange equations (1.2.3) must hold for almost every $s \in [0,T]$. □

We stress that if $u \in C^1([0,T]) \cap H_{P_1 P_2}$ is such that $V(z(s)) > 0$ and $\|z'(s)\| > 0$ for every $s \in [0,T]$, by continuity one can infer that Eq. (1.2.3) holds everywhere in $[0,T]$. Making use of the previous result, it is possible to find a connection between solutions of problem (1.2.2) and critical points of $L(\cdot)$.

Lemma 1.2.5. *Let $\bar{z} \in H_{P_1 P_2}$ be a solution of Problem (1.2.2) such that $V(\bar{z}(s)) > 0$ for every $s \in [0,T]$. Then $\bar{z}$ is also a critical point of the Jacobi length L.*

Proof. Let us start by observing that, if $\bar{z}$ is a solution of problem (1.2.2) such that $V(\bar{z}(s)) > 0$ in $[0,T]$, then $\bar{z} \in C^2([0,T]) \cap H_{P_1 P_2}([0,T])$ and $\|\bar{z}'(s)\| > 0$ for every $s \in [0,T]$. We will then prove that $\bar{z}$ is a critical point for $L(\cdot)$ by verifying that it solves the Euler-Lagrange equations (1.2.3) for every $s \in [0,T]$. As a matter of fact, one has, for every $s \in [0,T]$,

$$
\frac{d}{ds}\left(\sqrt{V(\bar{z}(s))} \frac{\bar{z}'(s)}{\|\bar{z}'(s)\|} \right)
$$

$$
= \frac{\nabla V(\bar{z}(s)) \cdot \bar{z}'(s)}{2\sqrt{V(\bar{z}(s))}} \frac{\bar{z}'(s)}{\|\bar{z}'(s)\|} + \frac{\sqrt{V(\bar{z}(s))}}{\|\bar{z}'(s)\|} \bar{z}''(s) - \sqrt{V(\bar{z}(s))} \frac{\bar{z}'(s) \cdot \bar{z}''(s)}{\|\bar{z}'(s)\|^3} \bar{z}'(s),
$$

and the Euler-Lagrange equations follow from the first two lines in Eq. (1.2.2). The conclusion follows from Lemma 1.2.4. □

Lemmas 1.2.4 and 1.2.5 are fundamental to compute the derivatives of the Jacobi distance with respect to variations of the endpoints z_0 and z_1: this represents the first step to give a variational justification to our Snell's law.

1.2.2 Snell's law.

Let us now pass from the case of a general potential V to our model, considering the outer and inner potential V_E and V_I, respectively in $\mathbb{R}^2 \setminus D$ and $\overline{D}$. With reference

to Definition 1.2.2, we will denote with L_E, d_E (resp. L_I, d_I) the Jacobi length and distance corresponding to the outer (resp. inner) potential. As our arcs start and end always on the boundary of our billiard, from this moment on our endpoints will belong to ∂D, and will be such that Assumption 1.2.1 is verified.

Recalling that ∂D is parametrised by a regular curve γ (see Theorem 1.1.1), it is convenient to express the distances as a function of the curve's parameter rather than of the endpoints in the plane. Take then $\xi_1^E, \xi_2^E, \xi_1^I, \xi_2^I \in [0, L]$ such that the corresponding pair of points $\gamma\left(\xi_1^E\right), \gamma\left(\xi_2^E\right)$ and $\gamma\left(\xi_1^I\right), \gamma\left(\xi_2^I\right)$ satisfy Assumption 1.2.1 respectively for the outer and inner problem (with an abuse of notation, in this case we will say that $\xi_1^E, \xi_2^E, \xi_1^I$ and ξ_2^I satisfy Assumption 1.2.1 as well). We can then define the functions

$$S_E(\xi_1^E, \xi_2^E) \doteq d_E(\gamma(\xi_1^E), \gamma(\xi_2^E)), \quad S_I(\xi_1^I, \xi_2^I) \doteq d_I(\gamma(\xi_1^I), \gamma(\xi_2^I)). \tag{1.2.4}$$

Let us recall that the distances d_E and d_I are infinitely-many differentiable as functions of the endpoints in every set in which they are well defined; by the chain rule, this implies that the lengths S_E and S_I inherit the regularity of the curve γ. In our case, since this curve is supposed to be at least of class $C^2([0, L])$, the two Jacobi lengths have the same regularity in every subset of $[0, L] \times [0, L]$ in which the inner or outer dynamics are well defined.

Denoting with ∂_a and ∂_b the partial derivatives respectively with respect to the first and second variable, one has

$$\begin{aligned}
\partial_a S_E \left(\xi_1^E, \xi_2^E\right) &= \nabla_{P_1} d_E \left(\gamma\left(\xi_1^E\right), \gamma\left(\xi_2^E\right)\right) \cdot \dot{\gamma}\left(\xi_1^E\right) \\
\partial_b S_E \left(\xi_1^E, \xi_2^E\right) &= \nabla_{P_2} d_E \left(\gamma\left(\xi_1^E\right), \gamma\left(\xi_2^E\right)\right) \cdot \dot{\gamma}\left(\xi_2^E\right)
\end{aligned} \tag{1.2.5}$$

(and similarly for S_I), where ∇_{P_1} and ∇_{P_2} are the gradients with respect to the first and the second point.

Lemma 1.2.6. *Let $\xi_1^E, \xi_2^E, \xi_1^I, \xi_2^I \in [0, L]$ and $h > 0$ such that Assumption 1.2.1 is verified. Then*

$$\begin{aligned}
\partial_a S_E(\xi_1^E, \xi_2^E) &= -\sqrt{V_E(\gamma(\xi_1^E))} \frac{z_E'(0; \gamma(\xi_1^E), \gamma(\xi_2^E))}{\|z_E'(0; \gamma(\xi_1^E), \gamma(\xi_2^E))\|} \cdot \dot{\gamma}(\xi_1^E) \\
\partial_b S_E(\xi_1^E, \xi_2^E) &= \sqrt{V_E(\gamma(\xi_2^E))} \frac{z_E'(T_E; \gamma(\xi_1^E), \gamma(\xi_2^E))}{\|z_E'(T_E; \gamma(\xi_1^E), \gamma(\xi_2^E))\|} \cdot \dot{\gamma}(\xi_2^E) \\
\partial_a S_I(\xi_1^I, \xi_2^I) &= -\sqrt{V_I(\gamma(\xi_1^I))} \frac{z_I'(0; \gamma(\xi_1^I), \gamma(\xi_2^I))}{\|z_I'(0; \gamma(\xi_1^I), \gamma(\xi_2^I))\|} \cdot \dot{\gamma}(\xi_1^I) \\
\partial_b S_I(\xi_1^I, \xi_2^I) &= \sqrt{V_I(\gamma(\xi_2^I))} \frac{z_I'(T_I; \gamma(\xi_1^I), \gamma(\xi_2^I))}{\|z_I'(T_I; \gamma(\xi_1^I), \gamma(\xi_2^I))\|} \cdot \dot{\gamma}(\xi_2^I).
\end{aligned} \tag{1.2.6}$$

Proof. Let us observe that the partial derivatives in Eq. (1.2.5) can be expressed as directional derivatives of the (inner or outer) Jacobi length in the direction of $\dot{\gamma}(\xi)$ for suitable $\xi \in [0, L]$. Taking for example $\partial_a S_E(\xi_1^E, \xi_2^E)$ (analogous expressions hold for

the other derivatives enlisted in Eq. (1.2.6)), one has

$$\partial_a S_E(\xi_1^E, \xi_2^E) = \partial_{1,\dot{\gamma}(\xi_1^E)} d_E\left(\gamma\left(\xi_1^E\right), \gamma\left(\xi_2^E\right)\right) = \partial_{1,\dot{\gamma}(\xi_1^E)} L_E\left(z_E\left(\cdot; \gamma\left(\xi_1^E\right), \gamma\left(\xi_2^E\right)\right)\right),$$

where, in general, $\partial_{1,\sigma}$ denotes the directional derivative with respect to variations of the first endpoint in the direction of the unit vector σ.

To find the explicit expression of $\partial_a S_E$, let us start by assuming that $\xi_1^E \neq \xi_2^E$. It is straightforward to verify that $z_E(\cdot) \doteq z_E\left(\cdot; \gamma\left(\xi_1^E\right), \gamma\left(\xi_2^E\right)\right)$ is a classical solution of the associated Bolza problem such that $V_E\left(z_E(s)\right) > 0$ for every $s \in [0, T_E]$: by Lemma 1.2.4, the outer arc is a solution of the associated Euler-Lagrange equations for every $s \in [0, T_E]$. Let us now consider the new parametrization $s \mapsto t(s)$ given by

$$\begin{cases} \dfrac{dt}{ds} = \dfrac{\sqrt{2} V_E(z_E(s))}{L_E} \\ t(0) = 0 \end{cases} \tag{1.2.7}$$

where $L_E \doteq L_E(z_E) \in \mathbb{R}$ as in Definition 1.2.2 with $V = V_E$, and define $\tilde{z}(t) \doteq z_E(s(t))$. Defining the new time derivative as $\dot{} \doteq \dfrac{d}{dt}$, it holds that $t \in [0, 1]$ and that

$$\forall t \in [0, 1] \quad \|\dot{\tilde{z}}(t)\| \sqrt{V_E\left(\tilde{z}(t)\right)} = L_E.$$

Moreover, by the invariance of the Jacobi length under reparametrizations, one has that

$$S_E\left(\xi_1^E, \xi_2^E\right) = L_E(z_E) = L_E\left(\tilde{z}\right),$$

and then, by (1.2.5), $\partial_a S_E(\xi_1^E, \xi_2^E) = \partial_{1,\dot{\gamma}(\xi_1^E)} L_E\left(\tilde{z}\right)$. Starting by (1.2.3), one can prove that the reparametrised curve $\tilde{z}$ satisfies the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{z}}\right) = \frac{\partial \mathcal{L}}{\partial z},$$

where $\mathcal{L} \doteq \|\dot{z}(t)\|^2 V_E(z(t)) = L_E^2$, namely,

$$\frac{d}{dt}\left(2 V_E\left(\tilde{z}(t)\right) \dot{\tilde{z}}(t)\right) = \|\dot{\tilde{z}}(t)\|^2 \nabla V_E\left(\tilde{z}(t)\right). \tag{1.2.8}$$

Let us now compute $\partial_{1,\dot{\gamma}(\xi_1^E)} L_E\left(\tilde{z}\right)$: differentiating $\mathcal{L}$ with respect to the first endpoint, one gets

$$2 L_E \partial_{1,\dot{\gamma}(\xi_1^E)} L_E = \partial_{1,\dot{\gamma}(\xi_1^E)} \mathcal{L}$$

$$= \partial_{1,\dot{\gamma}(\xi_1^E)} \int_0^1 \|\dot{\tilde{z}}(t)\|^2 \sqrt{V_E\left(\tilde{z}(t)\right)} dt$$

$$= \int_0^1 \partial_{1,\dot{\gamma}(\xi_1^E)} \left(\|\dot{\tilde{z}}(t)\|^2 \sqrt{V_E\left(\tilde{z}(t)\right)}\right) dt$$

$$= \int_0^1 \left(2 V_E\left(\tilde{z}(t)\right) \dot{\tilde{z}}(t) \cdot \partial_{1,\dot{\gamma}(\xi_1^E)} \dot{\tilde{z}}(t) + \|\dot{\tilde{z}}(t)\|^2 \nabla V_E\left(\tilde{z}(t)\right) \cdot \partial_{1,\dot{\gamma}(\xi_1^E)} \tilde{z}(t)\right) dt. \tag{1.2.9}$$

Moreover, multiplying (1.2.8) by $\partial_{1,\dot{\gamma}(\xi_1^E)}\tilde{z}(t)$ and integrating the result in $[0,1]$, one obtains

$$\int_0^1 \frac{d}{dt}\left(2V_E\left(\tilde{z}(t)\right)\dot{\tilde{z}}(t)\right)\cdot\partial_{1,\dot{\gamma}(\xi_1^E)}\tilde{z}(t) = \int_0^1 \|\dot{\tilde{z}}(t)\|^2\nabla V_E\left(\tilde{z}(t)\right)\partial_{1,\dot{\gamma}(\xi_1^E)}\tilde{z}(t)$$

$$\implies \int_0^1 2V_E\left(\tilde{z}(t)\right)\dot{\tilde{z}}(t)\cdot\partial_{1,\dot{\gamma}(\xi_1^E)}\dot{\tilde{z}}(t) + \|\dot{\tilde{z}}(t)\|^2\nabla V_E\left(\tilde{z}(t)\right)\cdot\partial_{1,\dot{\gamma}(\xi_1^E)}\tilde{z}(t) \qquad (1.2.10)$$

$$= -2V_E\left(\tilde{z}(0)\right)\dot{\tilde{z}}(0)\cdot\dot{\gamma}\left(\xi_1^E\right),$$

where the second equation is obtained by integrating by part and observing that $\partial_{1,\dot{\gamma}(\xi_1^E)}\tilde{z}(0) = \dot{\gamma}(\xi_1^E)$ and $\partial_{1,\dot{\gamma}(\xi_1^E)}\tilde{z}(1) = 0$. Comparing now (1.2.10) and (1.2.9) and recalling the expression of L_E, one gets the final expression

$$\partial_{1,\dot{\gamma}(\xi_1^E)}L_E\left(\tilde{z}\right) = -\frac{V_E\left(\tilde{z}(0)\right)\dot{\tilde{z}}(0)}{L_E}\cdot\dot{\gamma}\left(\xi_1^E\right) = -\sqrt{V_E\left(\tilde{z}(0)\right)}\frac{\dot{\tilde{z}}(0)}{\|\dot{\tilde{z}}(0)\|}\cdot\dot{\gamma}\left(\xi_1^E\right).$$

Returning now to the time parameter s, one obtains

$$\partial_a S_E(\xi_1^E,\xi_2^E) = -\sqrt{V_E(\gamma(\xi_1^E))}\frac{z_E'(0;\gamma(\xi_1^E),\gamma(\xi_2^E))}{\|z_E'(0;\gamma(\xi_1^E),\gamma(\xi_2^E))\|}\cdot\dot{\gamma}(\xi_1^E).$$

The same identity can be extended to the case $\xi_1^E = \xi_2^E$ by observing that $V_E\left(z_E(s;\gamma(\xi_1^E),\gamma(\xi_1^E))\right) > 0$ almost everywhere in $[0,T_E]$ and by taking into account the differentiable dependence of $z_E\left(\cdot;\gamma(\xi_1^E),\gamma(\xi_2^E)\right)$ with respect to variations of the endpoints.
In the inner case one can use the same reasonings, keeping in mind that, whenever a collision occurs, one can consider the corresponding regularized system. $\qquad \square$

For the sake of completeness, let us focus on the time parameter t used in the proof of Lemma 1.2.6, and defined in Eq. (1.2.7). In general, it is called *geodesic time*, and corresponds to the unique time parametrization such that the quantity $L = \|\frac{d}{dt}z(t)\|\sqrt{V\left(z(t)\right)}$ is constant along z. We will refer as *kinetic time* as the usual time parameter s for which Problem (1.2.2) is solved.
We are now ready to validate the refraction Snell's law

$$\sqrt{V_E(\gamma(\xi))}\sin\alpha_E = \sqrt{V_I(\gamma(\xi))}\sin\alpha_I, \qquad (1.2.11)$$

where:

- $\gamma(\xi)$ is the transition point between the outer and inner region (or viceversa);

- α_E, α_I are the angles of the two arcs with respect to the outward-pointing normal unit vector to γ in ξ.

The variational argument justifying (1.2.11) can be explicited by means of the partial derivatives of S_E and S_I.
Let us take $\xi_E,\xi_I \in [0,L]$ such that there exists a concatenation outer-inner arc starting

from $\gamma(\xi_E)$ and arriving in $\gamma(\xi_I)$. More precisely, this means that there exists $\xi \in [0, L]$ such that the pair ξ_E, ξ and ξ, ξ_I verify Assumption 1.2.1.

We can then consider the total Jacobi length of the concatenation composed by these arcs, denoted, following the notation introduced in Eq. (1.2.4), by $S_E(\xi_E, \xi) + S_I(\xi, \xi_I)$.

Definition 1.2.7. *We say that the concatenation between $\gamma(\xi_E)$ and $\gamma(\xi_I)$ of transition point $\gamma(\bar{\xi})$, with $\bar{\xi} \in [0, L]$, satisfies the Snell's law if $\bar{\xi}$ is a critical point[3] for the total Jacobi length, namely, if*

$$\frac{d}{d\xi} \left(S_E(\xi_E, \xi) + S_I(\xi, \xi_I) \right)_{|_{\xi = \bar{\xi}}} = 0. \tag{1.2.12}$$

An analogous condition can be established for a concatenation starting with an inner arc.

Note that condition (1.2.12) is equivalent to require that

$$\partial_b S_E(\xi_E, \bar{\xi}) + \partial_a S_I(\bar{\xi}, \xi_I) = 0, \tag{1.2.13}$$

which, in view of Lemma 1.2.6, can be rephrased as

$$\sqrt{V_E(\gamma(\bar{\xi}))} \frac{z'_E(T_E; \gamma(\xi_E), \gamma(\bar{\xi}))}{\|z'_E(T_E; \gamma(\xi_E), \gamma(\bar{\xi}))\|} \cdot \dot{\gamma}(\bar{\xi}) = \sqrt{V_I(\gamma(\bar{\xi}))} \frac{z'_I(0; \gamma(\bar{\xi}), \gamma(\xi_I))}{\|z'_I(0; \gamma(\bar{\xi}), \gamma(\xi_I))\|} \cdot \dot{\gamma}(\bar{\xi}). \tag{1.2.14}$$

Eq. (1.2.14) can be easily translated into Eq. (1.2.11), and can be interpreted as a conservation law for the tangential component of the trajectory's velocity vector across the interface. Let us observe that Eq. (1.2.11) has an evident correlation with the classical Snell's law for straight light rays, which can be derived again from a variational minimization problem (known as *Fermat's principle*). In this sense, our refraction law can be seen as a generalization for generic potentials and curved geodesics of this classical Snell's law. In particular, while in the classical case the geodesic arcs are always minimizers for the Jacobi length, in our case the solutions of the Bolza problems are only critical points of the corresponding lengths $L_E(\cdot)$ and $L_I(\cdot)$: this justifies the use of a criticality condition rather than a minimality one (which, in any case, can be retrieved working locally around the transition point).

Let us now return to Eq. (1.2.12), observing that, if $\xi_E, \bar{\xi}$ and ξ_I are such that

$$\frac{d}{d\bar{\xi}} \left(\partial_b S_E(\xi_E, \bar{\xi}) + \partial_a S_I(\bar{\xi}, \xi_I) \right) \neq 0, \tag{1.2.15}$$

---

[3]This criticality argument can be replaced in a minimality argument if we consider a more general definition for the inner and outer Jacobi distances, whose variables can be points not necessarily lying on ∂D, and restrict our analysis to a strongly convex neighbourhood for both the inner and outer Jacobi metric (see [68]). In such case, indeed, the geodesic arcs connecting any two points are minimizers for the Jacobi length.

then by the implicit function theorem it is possible to express locally the transition point as a function of the endpoints, that is, $\bar{\xi} = \bar{\xi}\left(\xi_E, \xi_I\right)^4$. In such case, we can then express the total Jacobi length of the complete concatenation from $\gamma\left(\xi_E\right)$ to $\gamma\left(\xi_E\right)$ only as a function of the endpoints as

$$S\left(\xi_E, \xi_I\right) \doteq S_E\left(\xi_E, \bar{\xi}\left(\xi_E, \xi_I\right)\right) + S_I\left(\bar{\xi}\left(\xi_E, \xi_I\right), \xi_I\right).$$

The function $S\left(\xi_E, \xi_I\right)$, in accordance with the case of classical Birkhoff billiards (see [17, 25]), can be defined as the *generating function* of our model, and will be used in Chapter 2 to derive, at least implicitly, the associated first return map in the case of the close-to-circle refractive billiards.

Let us conclude this section by analysing the good definition of our refraction rule: as already mentioned, for the Jacobi distances (outer or inner) to be well defined it is necessary to take endpoints for which the existence and uniqueness of the (outer or inner) arc are satisfied, and this problem will be addressed case by case in the next chapters. On the other hand, this is not enough to guarantee that Snell's law can be always verified: as $\forall \bar{z} \in \mathbb{R}^2 \backslash \{0\}$ one has that $V_E(\bar{z}) < V_I(\bar{z})$, the equation

$$\alpha_I = \arcsin\left(\sqrt{\frac{V_E(\bar{z})}{V_I(\bar{z})}} \sin \alpha_E\right)$$

is always solvable in the domain $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Viceversa, the crossing from the interior to the exterior of the domain may encounter an obstruction: indeed, the equation

$$\alpha_E = \arcsin\left(\sqrt{\frac{V_I(\bar{z})}{V_E(\bar{z})}} \sin \alpha_I\right)$$

admits a solution if and only if $\left|\sqrt{\frac{V_I(\bar{z})}{V_E(\bar{z})}} \sin \alpha_I\right| \leq 1$: in order to guarantee the solvability, we define a *critical angle*, depending on $\bar{z}$ and, as parameters, on the proper quantities of the problem $\mathcal{E}, h, \mu, \omega$, that is

$$\alpha_{I,crit} = \arcsin\left(\sqrt{\frac{V_E(\bar{z})}{V_I(\bar{z})}}\right).$$

In this way, the passage from the inside to the outside of the domain D takes place as long as $\alpha_I \in [-\alpha_{I,crit}, \alpha_{I,crit}]$ (we stress that, for $|\alpha_I| = \alpha_{I,crit}$, the refracted outer arc turns out to be tangent to ∂D). From a dynamical point of view, this means that the, for Snell's law to happen, every incoming inner arc must be *transverse enough* to the boundary ∂D.

---

[4]The nondegeneracy condition (1.2.15), as well as the domain of good definition of $\bar{\xi}\left(\xi_E, \xi_I\right)$, will be the subject of a thorough analysis, in the context of the close-to-circle billiards, in Chapter 2.

1.3   Local existence of inner and outer arcs: a transversality approach

This section is devoted to state existence of outer and inner arcs, close to the homothetic ones, which will be next used to apply a broken geodesics technique. We shall use a classical transversality approach, reminiscent to the one in [69], to which we refer for a more detailed exposition. It is worthwhile stressing that, when dealing with the inner dynamics, we shall take advantage of Levi-Civita regularising transformation.

From now on, we will always suppose that D is contained in the outer potential's Hill's region

$$\mathcal{H} = \left\{ p \in \mathbb{R}^2 \ \middle| \ \mathcal{E} - \frac{\omega^2}{2}\|p\|^2 > 0 \right\},$$

and assume that its boundary ∂D is parametrised by $\gamma : I \mapsto \mathbb{R}^2$, with $\gamma \in C^2$. Moreover, as already said in Section 1.1 and as we intend to assume from this moment on, we identify any point in $\mathbb{R}^2$, and in particular of the curve, with its corresponding position vector with respect to $\mathbf{0}$. We focus on the points of γ which satisfy a local transversality property, as well as a local star-convexity, namely:

$$\bar{\xi} \in I \text{ such that}: \quad (i) \quad \gamma(\bar{\xi}) \nparallel \dot{\gamma}(\bar{\xi})$$
$$(ii) \quad \text{the ray starting from } \mathbf{0} \text{ in the direction of } \gamma(\bar{\xi}) \text{ intersects}$$
$$\partial D \text{ only once.}$$
$$(1.3.1)$$

As in this chapter our main interest lies in the local study of the trajectories around the homothetic solutions, defined as in Section 1.1, characterised as in (1.1.2) and whose directions are in a subset of the ones identified by (1.3.1), for the moment we restrict our analysis to a neighbourhood of $\bar{\xi}$: condition (1.3.1), along with the regularity of γ, assures indeed the existence of an open interval $I' \subset I$ such that $\bar{\xi} \in I'$ and

$$\forall \xi \in I' \ \gamma(\xi) \nparallel \dot{\gamma}(\xi). \qquad (1.3.2)$$

Furthermore, possibly taking a smaller I', we can suppose that condition (1.3.1(ii)) holds for every $\xi \in I'$. The local transversality property of $\gamma(I')$ with respect to the radial directions and its star-convexity with reference to the origin will be the main ingredients to guarantee the existence of the inner and outer arcs in a neighbourhood of a homothetic solution.

Theorem 1.3.1. *Suppose that the domain's boundary ∂D is a regular curve parametrised by $\gamma : I \to \mathbb{R}^2$ and suppose that $\bar{\xi} \in I$ satisfies (1.3.1). Then there are $\epsilon_\alpha(\bar{\xi}) > 0$ and $\epsilon_{\xi_0}(\bar{\xi}) > 0$ such that for every $\xi_0 \in I$, $\alpha \in [-\pi/2, \pi/2]$ with $|\bar{\xi} - \xi_0| < \epsilon_{\xi_0}(\bar{\xi})$ and $|\alpha| < \epsilon_\alpha(\bar{\xi})$, there exist $T > 0, \xi_1 \in I$ such that the problem (in complex notation*

$\gamma(\bar{\xi}) = \|\gamma(\bar{\xi})\| e^{i\bar{\theta}})$

$$\begin{cases} z''(s) = -\omega^2 z(s) \\ \dfrac{1}{2} \|z'(s)\|^2 - \mathcal{E} + \dfrac{\omega^2}{2} \|z(s)\|^2 = 0 \\ z(0) = \gamma(\xi_0), \, z'(0) = v_0 e^{i(\bar{\theta} + \alpha)}, \end{cases}$$

with $v_0 = \sqrt{2\mathcal{E} - \omega^2 \|\gamma(\xi_0)\|^2}$, admits the unique solution $z(s; \xi_0, \alpha)$ and $z(T; \xi_0, \alpha) = \gamma(\xi_1) \in \partial D$.
Moreover, for every $s \in (0, T)$ one has $y(s; \xi_0, \alpha) \notin \bar{D}$.

The proof of the above result is based on the following Lemma, and the subsequent Remark. We will postpone to Chapter 2 the explicit proof of an analogous result, obtained taking into consideration the Bolza fixed-ends problem instead of the Cauchy one.

Lemma 1.3.2. *Suppose that $\bar{\xi} \in I$ satisfies condition (1.3.1). Then there exist $\delta_0(\bar{\xi})$, $\delta_1(\bar{\xi})$, $\rho(\bar{\xi}) > 0$ such that for every $\xi_0 \in I$, $\dot{\theta}_0 \in \mathbb{R}$ with $|\bar{\xi} - \xi_0| < \delta_0(\bar{\xi})$ and $|\dot{\theta}_0| < \rho(\bar{\xi})$, defined the unit vectors (in exponential notation) $\hat{x}_1 = \gamma(\bar{\xi})/\|\gamma(\bar{\xi})\|$ and $\hat{x}_2 = i\hat{x}_1$, there exist $T > 0$ and $\xi_1 \in I$ such that the problem*

$$\begin{cases} z''(s) = -\omega^2 y(s) \\ \dfrac{1}{2} \|z'(s)\|^2 - \mathcal{E} + \dfrac{\omega^2}{2} \|z(s)\|^2 = 0 \\ z(0) = \gamma(\xi_0), \, z'(0) = \dot{r}_0 \hat{x}_1 + \dot{\theta}_0 \hat{x}_2, \end{cases}$$

with $\dot{r}_0 = \dot{r}_0(\dot{\theta}_0) = \sqrt{2\mathcal{E} - \omega^2 \|\gamma(\xi_0)\| - \dot{\theta}_0^2}$, admits the unique solution $z(s; \xi_0, \dot{\theta}_0)$ such that $z(T; \xi_0, \dot{\theta}_0) = \gamma(\xi_1) \in \partial D$. Moreover, $|\bar{\xi} - \xi_1| < \delta_1(\bar{\xi})$.

The proof relies on a transversality argument, standard in detecting one side Poincaré sections, based upon regularity of solutions of Cauchy's problems and the implicit function theorem (see, e.g. the similar construction in [69]).

Remark 1.3.3. *The validity of condition (1.3.1(ii)) in a neighbourhood of $\bar{\xi}$ entails that the point $\gamma(\xi_1)$ defined as in Lemma 1.3.2 is such that, for every $s \in (0, T(\xi_0, \dot{\theta}_0))$, $z(s; \xi_0, \dot{\theta}_0) \notin \bar{D}$, that is, there are no other intersections of ∂D and the arc $z([0, T(\xi_0, \dot{\theta}_0)], \xi_0, \dot{\theta}_0)$ other than $\gamma(\xi_0)$ ad $\gamma(\xi_1)$. This is in fact a consequence the continuous dependence on the initial conditions, for which, if $(\xi_0, \dot{\theta}_0)$ are sufficiently close to $(\bar{\xi}, 0)$, then $z(\cdot; \xi_0, \dot{\theta}_0)$ is arbitrarily close to $z(\cdot; \bar{\xi}, 0)$ in the C^0 topology.*

The above Lemma states the existence of a local Poincaré section in the energy manifold of $\partial D \times \mathbb{R}^2$ for the outer dynamics in a neighbourhood of the initial condition of a radial brake orbit (namely, the direction of the velocity vector coincides with the radial one) and under some local conditions on ∂D. The condition $|\dot{\theta}_0| < \rho(\bar{\xi}) \equiv \rho$ can be rephrased by considering the angle $\alpha \in [-\pi/2, \pi/2]$ between the initial velocity $y'(0) = \dot{r}_0 \hat{x}_1 + \dot{\theta}_0 \hat{x}_2$ and the radial unit vector $\hat{x}_1$ of $\gamma(\bar{\xi})$ (notice that α and $\dot{\theta}_0$ have

always the same sign). In particular, one has that

$$\tan \alpha = \frac{\dot{\theta}_0}{\dot{r}_0} = \frac{\dot{\theta}_0}{\sqrt{2\mathcal{E} - \omega^2 \|\gamma(\xi_0)\|^2 - \dot{\theta}_0^2}} \Leftrightarrow \dot{\theta}_0 = \tan \alpha \sqrt{\frac{2\mathcal{E} - \omega^2 \|\gamma(\xi_0)\|^2}{1 + \tan \alpha^2}} = f(\alpha, \xi_0).$$

As $f(\alpha, \xi_0)$ is continuous and $f(0, \bar{\xi}) = 0$, there exist $\epsilon_\alpha, \epsilon_{\xi_0} > 0$ such that $\epsilon_{\xi_0} < \delta_0(\bar{\xi})$ and, if $|\alpha| < \epsilon$ and $|\bar{\xi} - \xi_0| < \epsilon_{\xi_0}$, then $|\dot{\theta}_0| < \rho$. Taking together Lemma 1.3.2 and Remark 1.3.3, one can eventually state Theorem 1.3.1.

In the case of the inner arcs, we turn to the problem

$$\begin{cases} z''(s) = -\dfrac{\mu}{\|z(s)\|^3} z(s), & s \in [0, S], \\ \dfrac{1}{2} \|z'(s)\|^2 - \mathcal{E} - h - \dfrac{\mu}{\|z(s)\|} = 0, & s \in [0, S], \\ z(0) = z_0, z'(0) = \mathbf{v}_0; \end{cases} \tag{1.3.3}$$

for some $S > 0$ and some initial conditions $z_0 \in \partial D$ and $\mathbf{v}_0$ pointing inward the domain D, and denote with $z(s; z_0, \mathbf{v}_0)$ its solution (with an abuse of notation, in the following the initial velocity will be defined either by its angle with the radial direction or its orthogonal component to the latter). As the Keplerian orbits with positive energy are unbounded (see for example [70]) and D is bounded, for every initial condition for which the arc enters in D there is $\tilde{S} > 0$ such that the it encounters ∂D again in a point which we call z_1. We search for constraints for $z_0, \mathbf{v}_0$ such that the velocity vector at this point, denoted with $z'(\tilde{S}; z_0, \mathbf{v}_0)$, is transverse to ∂D in the sense that $z'(\tilde{S}; z_0, \mathbf{v}_0)$ and the tangent vector to ∂D in z_1 are not parallel.

The singularity at the origin of the inner potential can be treated by means of the Levi-Civita regularisation technique (see [26]), which consists in a change both in the temporal parameter and the spatial coordinates, in order to remove the singularity of Kepler-type potentials. In particular, the following Proposition holds.

Proposition 1.3.4. *Problem (1.3.3) is conjugated, via a suitable set of transformations called Levi-Civita transformations, to the problem*

$$\begin{cases} w''(\tau) = \Omega^2 w(\tau), & \tau \in [0, T], \\ \dfrac{1}{2} \|w'(\tau)\|^2 - E - \dfrac{\Omega^2}{2} \|w(\tau)\|^2 = 0, & \tau \in [0, T], \\ w(0) = w_0, w'(0) = \dot{w}_0 \end{cases} \tag{1.3.4}$$

for suitable $w_0, \dot{w}_0, T$, $\Omega = \sqrt{2\mathcal{E} + 2h}$, $E = \mu$.

Proof of Proposition 1.3.4. Setting $r(s) = \|z(s)\|$, consider the reparametrisation $s = s(\tilde{\tau})$ such that

$$\frac{d}{ds} = \frac{1}{r(s(\tilde{\tau}))} \frac{d}{d\tilde{\tau}} \Rightarrow \frac{d^2}{ds^2} = -\frac{1}{r(s(\tilde{\tau}))^3} \frac{d}{d\tilde{\tau}} + \frac{1}{r(s(\tilde{\tau}))^2} \frac{d^2}{d\tilde{\tau}^2}.$$

Denoting, with an abuse of notation, $' = d/d\tilde{\tau}$, the first and second equations in (1.3.3) become

$$r(\tilde{\tau})z''(\tilde{\tau}) - r'(\tilde{\tau})z'(\tilde{\tau}) + \mu z(\tilde{\tau}) = 0, \quad \frac{1}{2r(\tilde{\tau})^2}\|z'(\tilde{\tau})\|^2 - \mathcal{E} - h - \frac{\mu}{r(\tilde{\tau})} = 0.$$

Identifying now $\mathbb{R}^2$ and $\mathbb{C}$, let us consider a new spatial coordinate $w \in \mathbb{C}$ such that $z(\tilde{\tau}) = w^2(\tilde{\tau})$: we have then

$$2r(\tilde{\tau})w(\tilde{\tau})w''(\tilde{\tau}) - w(\tilde{\tau})^2(\|w'(\tilde{\tau})\|^2 - \mu) = 0, \quad 2\|w'(\tilde{\tau})\|^2 - \mu = r(\tilde{\tau})(\mathcal{E} + h)$$
$$\Rightarrow 2w''(\tilde{\tau}) = w(\tilde{\tau})(\mathcal{E} + h).$$

Finally, considering the new time variable $\tau = \tilde{\tau}/2$ (again, with an abuse of notation, $' = d/d\tau$), one obtains the final Cauchy problem

$$\begin{cases} w''(\tau) = 2(\mathcal{E} + h)w(\tau), & \tau \in [-T, T] \\ \frac{1}{2}\|w'(\tau)\|^2 - (\mathcal{E} + h)\|w(\tau)\|^2 = \mu & \tau \in [-T, T] \\ w(-T) = w_0, w'(-T) = \dot{w}_0 \end{cases}$$
$$= \begin{cases} w''(\tau) = \Omega^2 w(\tau), & \tau \in [-T, T] \\ \frac{1}{2}\|w'(\tau)\|^2 - \frac{\Omega^2}{2}\|w(\tau)\|^2 = E & \tau \in [-T, T], \\ w(-T) = w_0, w'(-T) = \dot{w}_0 \end{cases}$$

for some $T > 0$ and suitable initial conditions $w_0, \dot{w}_0$ (we will return to the determination of w_0 and $\dot{w}_0$ in Proposition 1.3.10). The solutions of (1.3.3) can be then seen, in a suitable parametrisation, as complex squares of solutions of a harmonic repulsor with fixed ends boundary conditions, energy equal to $E = \mu$ and frequency $\Omega = \sqrt{2(\mathcal{E} + h)}$. □

We will refer to the time variable τ as the Levi-Civita time, and to the new reference system as the Levi-Civita plane.

By means of this regularisation, which is of independent interest and will be used again in Section 1.5, one can infer the local existence and transversality of solutions of Problem (1.3.3) in the vicinity of the homothetic orbits, in terms of both initial positions and velocities.

Theorem 1.3.5. *Let us suppose that ∂D is a closed curve of class C^2, $\mathbf{0} \in D$, parametrised by $\gamma(\xi) : I \to \mathbb{R}^2$, and suppose that there is $\bar{\xi} \in I$ such that condition 1.3.1 is satisfied. Then there exist $\lambda_{\xi_0}(\bar{\xi}) > 0$ and $0 < \lambda_\alpha(\bar{\xi}) < \pi/2$ such that for every $\xi_0 \in [\bar{\xi} - \lambda_{\xi_0}, \bar{\xi} + \lambda_{\xi_0}]$ and $\alpha \in [-\lambda_\alpha, \lambda_\alpha]$ there are $T > 0$ and $\xi_1 \in I$ such that the problem in complex notation*

$$\begin{cases} z''(s) = -\frac{\mu}{\|z(s)\|^3}z(s), \quad \frac{1}{2}\|z'(s)\|^2 = \mathcal{E} + h + \frac{\mu}{\|z(s)\|}, \quad s \in [0, S] \\ z(0) = \gamma(\xi_0) := \rho(\xi_0)e^{i\theta(\xi_0)}, z'(0) = -\sqrt{2}\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi_0)}}e^{i(\theta(\bar{\xi})+\alpha)} \end{cases} \tag{1.3.5}$$

admits the unique solution $z(s; \xi_0, \alpha)$. Moreover, $z(T; \xi_0, \alpha) = \gamma(\xi_1) \in \partial D$, and $z'(T; \xi_0, \alpha) \equiv z'(T)$ is not tangent to ∂D. More precisely, there exists $0 < \sigma < \pi/2$, depending on $\bar{\xi}$, such that, if $\beta \in [-\pi/2, \pi/2]$ is such that

$$z'(T) = \sqrt{2}\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi_1)}} e^{i(\theta(\xi_1) + \beta)}, \text{ we have } \beta \in [-\sigma, \sigma].$$

Remark 1.3.6. *As in the case of the outer dynamics, taking possibly a smaller λ_α one can assure that $z((0, T); \xi_0, \alpha) \in D$, namely, that $z(s; \xi_0, \alpha)$ does not intersect ∂D for $s \in (0, T)$. This is in fact guaranteed by the validity of condition (1.3.1(ii)), the continuous dependence of problem (1.3.5) on the initial conditions and the fact that $0 \in D$.*

To prove Theorem 1.3.5 we rely again on some preliminary results, taking into account the transformation of the billiard's boundary passing from the physical to the Levi-Civita plane. Like in the outer problem, suppose that ∂D is a closed curve of class C^2 parametrised by $\gamma(\xi) : I \to \mathbb{R}^2$: passing to the Levi-Civita plane, γ is transformed according to the same rule $w^2 = z$. As the complex square determines a double covering of $\mathbb{C}$, it is clear that every arc $z(\tau)$ in the physical plane corresponds to two arcs $w(\tau)$ in the Levi-Civita plane, depending on the choice of w_0, which is such that $w_0^2 = z_0$, and a suitable transformed velocity $\dot{w}_0$. In the following, we will work with the Levi-Civita variables, taking respectively for w_0 the negative determination of the square root of z_0 and for w_1 the positive determination of the square root of z_1, namely, in polar coordinates,

$$z_0 = \|z_0\|e^{i\theta_0} \Rightarrow w_0 = -\sqrt{\|z_0\|}e^{i\frac{\theta_0}{2}}, \quad z_1 = \|z_1\|e^{i\theta_1} \Rightarrow w_1 = \sqrt{\|z_1\|}e^{i\frac{\theta_1}{2}}. \quad (1.3.6)$$

The transformed boundary follows the same rules, and is defined in two neighbourhoods of w_0 and w_1. More precisely, let us suppose that $\bar{\xi}$ satisfies condition (1.3.1) and, additionally, $\gamma(\bar{\xi})$ points in the direction of $e_1 = (1, 0)$: for the sake of simplicity, we will focus on this particular value of $\bar{\xi}$, as for every $\bar{\xi}' \in I$ satisfying (1.3.1) we can consider the rotated basis (e_1', e_2') such that $\bar{\xi}'$ has the properties of $\bar{\xi}$.

Definition 1.3.7. *Letting $\bar{\xi}$ be defined above, there exists $\bar{\epsilon} > 0$ such that, if $\gamma(\xi)$ is expressed in polar coordinates, namely, $\gamma(\xi) = \rho(\xi)e^{i\theta(\xi)}$, the curves*

$$\phi_+(\xi) : (\bar{\xi} - \bar{\epsilon}, \bar{\xi} + \bar{\epsilon}) \to \mathbb{C}, \quad \phi_+(\xi) = \sqrt{\rho(\xi)}e^{i\theta(\xi)/2},$$

$$\phi_-(\xi) : (\bar{\xi} - \bar{\epsilon}, \bar{\xi} + \bar{\epsilon}) \to \mathbb{C}, \quad \phi_-(\xi) = -\sqrt{\rho(\xi)}e^{i\theta(\xi)/2} = \sqrt{\rho(\xi)}e^{i(\theta(\xi)/2 + \pi)}$$

are well defined in the Levi-Civita plane and represent the local transform of γ.

As an immediate consequence of the conformality of the map $w \mapsto w^2$ we have the following

Lemma 1.3.8. *The transformed curves $\phi_\pm(\xi)$ preserve the angle between the radial and the tangent direction of $\gamma(\xi)$. In particular, if condition 1.3.1 holds for $\gamma(\xi)$, then it holds for $\phi_\pm(\xi)$ with $\xi \in (\bar{\xi} - \bar{\epsilon}, \bar{\xi} + \bar{\epsilon})$, possibly reducing $\bar{\epsilon}$.*

Let us focus on the transformed inner arc in the Levi-Civita plane: the next Proposition states the existence, under suitable hypotheses on the initial conditions, of a solution of Problem (1.3.4) which has the desired transversality properties.

Proposition 1.3.9. *If condition 1.3.1 holds for $\gamma(\bar\xi)$, then there are $\tilde\lambda > 0, 0 < \tilde\epsilon < \bar\epsilon$ such that, for every $\xi_0 \in [\bar\xi - \tilde\epsilon, \bar\xi + \tilde\epsilon], \dot\theta_0 \in [-\tilde\lambda, \tilde\lambda]$ there are $T > 0, \xi_1 \in I$ such that the Cauchy problem*

$$\begin{cases} w''(\tau) = \Omega^2 w(\tau), & \tau \in [-T, T] \\ \dfrac{1}{2}\|w'(\tau)\|^2 - \dfrac{\Omega^2}{2}\|w(\tau)\|^2 = E & \tau \in [-T, T], \\ w(0) = \phi_-(\xi_0), w'(0) = \dot r_0 e_1 + \dot\theta_0 e_2 \end{cases}$$

with $\dot r_0 = \sqrt{2E + \Omega^2 \|\phi_-(\xi_0)\|^2 - \dot\theta_0^2}$ admits the unique solution $w(\tau; \xi_0, \dot\theta_0)$. Moreover, $w(T; \xi_0, \dot\theta_0) = \phi_+(\xi_1)$. In addition, $w'(T(\xi_0, \dot\theta_0); \xi_0, \dot\theta_0) \nparallel \dot\phi_+(\xi_1(\xi_0, \dot\theta_0))$, namely, the regularised arc is not tangent to ∂D in $\phi_+(\bar\xi)$.

The proof is again rather standard and relies on a transversality argument for the regularised flow. Moreover, continuity of the regularised flow with respect to the initial conditions and angle preserving of the complex square map entail the desired transversality property.

Let us notice that the smallness condition on the velocity's orthogonal component $\dot\theta_0$ can be given also in terms of the angle between the radial direction and the initial velocity vector. As in the outer case, we can consider the angle $\alpha \in [-\pi/2, \pi/2]$ between $w'(0) = \dot r_0 e_1 + \dot\theta_0 e_2$ and $\phi_-(\bar\xi)$ and have

$$\tan \alpha = \frac{\dot\theta_0}{\dot r_0} = \frac{\dot\theta_0}{\sqrt{2E + \Omega^2\|\phi_-(\xi_0)\|^2 - \dot\theta_0^2}} \Leftrightarrow \dot\theta_0 = \tan\alpha\sqrt{\frac{2E + \Omega^2\|\phi_-(\xi_0)\|^2}{1 + \tan\alpha^2}} = g(\alpha, \xi_0).$$

As $g(\alpha, \xi_0)$ is continuous and $g(0, \bar\xi) = 0$, there exist $\lambda_\alpha > 0$ and $\lambda_{\xi_0} > 0$ such that $\lambda_{\xi_0} < \epsilon$ and, if $|\alpha| < \lambda_\alpha$ and $|\bar\xi - \xi_0| < \lambda_{\xi_0}$, then $|\dot\theta_0| < \tilde\lambda$.

Proposition 1.3.10. *Let us consider $\bar\xi \in I$ such that $\gamma(\bar\xi) = \rho e_1, \rho > 0$, and suppose that condition 1.3.1 holds. Let $\bar\epsilon > 0$ such that the curves $\phi_\pm : [-\bar\epsilon + \bar\xi, \bar\epsilon + \bar\xi]$ are well defined, and choose $\xi_0 \in [-\bar\epsilon + \bar\xi, \bar\epsilon + \bar\xi]$ and $\beta \in [-\dfrac{\pi}{2}, \dfrac{\pi}{2}]$. Then the system (in polar coordinates)*

$$\begin{cases} z''(s) = -\dfrac{\mu}{\|z(s)\|^3} z(s), & s \in [0, S], \\ \dfrac{1}{2}\|z'(s)\|^2 - \mathcal{E} - h - \dfrac{\mu}{\|z(s)\|} = 0, & s \in [0, S], \\ z(0) = \gamma(\xi_0) = \rho(\xi_0)e^{i\theta(\xi_0)}, z'(0) = \sqrt{2}\sqrt{\mathcal{E} + h + \dfrac{\mu}{\rho(\xi_0)}}e^{i(\theta(\xi_0)+\beta)} \end{cases}$$

is conjugated, in the Levi-Civita plane, and considering $\tau = \tau(s)$ the Levi-Civita time, to the problem

$$\begin{cases} w''(\tau) = \Omega^2 w(\tau), & \tau \in [-T, T] \\ \dfrac{1}{2}\|w'(\tau)\|^2 - \dfrac{\Omega^2}{2}\|w(\tau)\|^2 = E & \tau \in [-T, T], \\ w(0) = -\sqrt{\rho(\xi_0)}e^{i\theta(\xi_0)/2}, w'(0) = -\sqrt{2E + \Omega^2\rho(\xi_0)}e^{i(\theta(\xi_0)/2+\beta)} \end{cases}$$

for a suitable $T > 0$. In other words, the angles between the original initial conditions and the transformed ones, namely, $\widehat{z(0), \dot{z}(0)}$ and $\widehat{w(0), w'(0)}$, are equal.

Proof. From the definition of ϕ_-, we have that

$$w(0) = \phi_-(\xi_0) = -\sqrt{\rho(\xi_0)}e^{i(\theta(\xi_0))} = \sqrt{\rho(\xi_0)}e^{i(\theta(\xi_0)+\pi)}.$$

To compute $w'(0) = dw(0)/d\tau$, we go through the following Levi-Civita transformations:

- $\dfrac{d}{ds} = \dfrac{1}{\|z(\tilde{\tau}(s))\|}\dfrac{d}{d\tilde{\tau}}$, then, for $s = 0$,

$$\sqrt{2}\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi_0)}}e^{i\theta(\xi_0)} = \frac{d}{ds}z(0) = \frac{1}{\rho(\xi_0)}\frac{d}{d\tilde{\tau}}z(0);$$

- $z = w^2$, then

$$\sqrt{2}\rho(\xi_0)\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi_0)}}e^{i\theta(\xi_0)} = \frac{d}{d\tilde{\tau}}z(0) = 2w(0)\frac{d}{d\tilde{\tau}}w(0);$$

- $\tau = \tilde{\tau}/2$, then

$$\sqrt{2}\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi_0)}}e^{i(\theta(\xi_0)+\beta)} = w(0)\frac{d}{d\tau}w(0)$$

$$\implies w'(0) = -\sqrt{2E + \Omega^2\rho(\xi_0)}e^{i(\theta(\xi_0)/2+\beta)}$$

$\square$

The angle between the initial conditions in the physical plane is then preserved after the passage in the Levi-Civita reference frame: this assures that, if we fix some "smallness" condition on the angle α between the initial velocity and the direction of $\gamma(\bar{\xi})$ in the original reference frame, they will hold also in the Levi-Civita plane. This allows, along with the previous results, to state Theorem 1.3.5.

Notation 1.3.11. In the following, we will refer to the outer differential equation along with the energy conservation law with (HS_E), while (HS_I) and (HS_{LC}) will denote the inner differential problem respectively in the physical and in the Levi-Civita variables, with their own energy conservation conditions. More precisely, we will write $(HS_E)[z]$ if z satisfies the outer differential equation with zero energy, and use the analogous notation for (HS_I) and (HS_{LC}).

1.4   First return map

As in the previous Sections, let us suppose that the boundary of the regular domain D defined in Section 1.1 can be parametrised by a regular closed curve $\gamma : I \to \mathbb{R}^2$. Given some initial conditions $z_0^{(I)}, v_0^{(I)}, z_0^{(E)}$ and $v_0^{(E)}$, let us consider the solutions $z_I(s)$ and $z_E(s)$ of the two systems

$$\begin{cases} (HS_I)[z(s)] & s \in [0, T_I] \\ z_I(0) = z_0^{(I)}, z_I'(0) = v_0^{(I)} \end{cases} \qquad \begin{cases} (HS_E)[z(s)] & s \in [0, T_E] \\ z_E(0) = z_0^{(E)}, z_E'(0) = v_0^{(E)} \end{cases}$$

$$(1.4.1)$$

for some $T_I, T_E > 0$. Fixed $z_0 \in \partial D$, $v_0 \in \mathbb{R}^2$ such that it points towards the exterior of D, we want to describe (supposing that it exists) the trajectory obtained by the juxtaposition of an outer arc z_E and the subsequent inner arc z_I, namely $z_{EI}(s)$ defined by

$$z_{EI}(s) = \begin{cases} z_E(s) & s \in [0, T_E) \\ z_I(s) & s \in [T_E, T_E + T_I) \\ z_E^{(1)}(s) & s = T_E + T_I, \end{cases}$$

$$(1.4.2)$$

where the branches z_E, z_I and $z_E^{(1)}$ are solution either of the outer or the inner problem and are connected by following the Snell's rule. In particular, we require $z_E(T_E) = z_I(T_E)$ and $z_I(T_E + T_I) = z_E^{(1)}(T_E + T_I)$, and, using the notation

$$z_E(T_E) = z_I(T_E) = z_1, \quad z_I(T_E + T_I) = z_E^{(1)}(T_E + T_I) = z_2$$

$$v_1 = \frac{z_E'(T_E)}{\|z_E'(T_E)\|}, \quad v_1' = \frac{z_I'(T_E)}{\|z_I'(T_E)\|}, \quad v_2 = \frac{z_I'(T_E + T_I)}{\|z_I'(T_E + T_I)\|} \quad v_2' = \frac{z_E'^{(1)}(T_E + T_I)}{\|z_E'^{(1)}(T_E + T_I)\|}$$

we demand

$$\begin{cases} (HS_E)[z_E(s)] & s \in [0, T_E] \\ z_E(s) \notin D, z_E(T_E) \in \partial D & s \in (0, T_E) \\ z_E(0) = z_0, z_E'(0) = v_0 \end{cases}$$

$$(1.4.3)$$

$$\begin{cases} (HS_I)[z_I(s)] & s \in [T_E, T_E + T_I] \\ z_I(s) \in D, z_I(T_E + T_I) \in \partial D & s \in (T_E, T_E + T_I) \\ \sqrt{V_E(z_1)}v_1 \cdot e_1 = \sqrt{V_I(z_1)}v_1' \cdot e_1 \end{cases}$$

$$(1.4.4)$$

$$\begin{cases} (HS_E)[z_E^{(1)}(s)] & s \in [T_E + T_I, T_E + T_I + \tilde{T}] \\ z_E'^{(1)}(s) \notin D, & s \in (T_E + T_I, T_E + T_I + \tilde{T}] \\ \sqrt{V_I(z_2)}v_2 \cdot e_2 = \sqrt{V_E(z_2)}v_2' \cdot e_2, \end{cases}$$

for some $T_E, T_I, \tilde{T} > 0$ and where e_1 and e_2 are the unit vectors tangent to ∂D respectively in z_1 and z_2.

1.4.1   Local first return map

We wish to construct the iteration map which expresses
$(z_1, v_1) \doteq (z_{EI}(T_E + T_I), z'_{EI}(T_E + T_I))$ as a function of (z_0, v_0) in a suitable set of coordinates.

Let us suppose that the point $z_{EI}(0) = \gamma(\xi) \in \partial D$ is the starting point of the outer branch of $z_{EI}(t)$: then, denoting with $t(\xi)$ and $n(\xi)$ respectively the tangent and the outward-pointing normal unit vectors of γ in ξ, the initial velocity v can be expressed as $v = \sqrt{2V_E(\gamma(\xi))}(\cos\alpha \, n(\xi) + \sin\alpha \, t(\xi))$, where $\alpha \in [-\pi/2, \pi/2]$ is the angle between v and $n(\xi)$, positive if $v \cdot t(\xi) \geq 0$ and negative otherwise. Then, once ξ is fixed, the vector v is completely determined by α. We can then consider the map

$$F : B \subset (I \times [-\pi/2, \pi/2]) \to I \times [-\pi/2, \pi/2],$$
$$(\xi_0, \alpha_0) \mapsto (\xi_1, \alpha_1) = (\xi_1(\xi_0, \alpha_0), \alpha_1(\xi_0, \alpha_0)),$$

where the pair (ξ_1, α_1) completely determines $(z_1, v_1$. The determination of the domain of F, denoted with B, is a nontrivial problem, whose main issues are discussed in Remark 1.4.2.

Although F is not explicitly defined, taking together the properties of the solutions of Problem (1.4.1) and Snell's law (1.2.11), under some suitable hypotheses on ∂D, one can characterize one particular class of fixed points of F, deriving from one-periodic homothetic solutions of (1.4.1):

Remark 1.4.1. *Initial conditions* $z_{EI}(0) = \gamma(\bar{\xi})$, $z'_{EI}(0) = \sqrt{2V_E(\bar{\xi})}\gamma(\bar{\xi})/\|\gamma(\bar{\xi})\|$ *correspond to a homothetic solution of Problem (1.4.1) if and only if*

$$\gamma(\bar{\xi}) \perp \dot{\gamma}(\bar{\xi})$$
and the segment $t\gamma(\bar{\xi})$, $t \in [0, \infty)$, *does not intersect* ∂D *for* $t \neq 1$.

$$(1.4.5)$$

Therefore, if condition (1.4.5) holds, the pair $(\bar{\xi}, 0)$ *is a fixed point for* F, *which we call* homothetic.

Remark 1.4.2. *The conditions for F to be globally defined on $I \times [-\pi/2, \pi/2]$ are essentially two:*

(i) the existence and uniqueness of the outer and inner arcs for any initial conditions;

(ii) the good definition of the refraction rule for every incoming arc: according to the reasonings in Section 1.2, it is equivalent to require that for every inner arc, if

we denote with β_1 the angle between $z_I'(T_I + T_E)$ and the inward-pointing normal vector to ∂D in $\gamma(\xi_1)$, we shall have $|\beta_1| < \beta_{crit} = \arcsin(\sqrt{V_E(\gamma(\xi_1))/V_I(\gamma(\xi_1))})$.

In other words, condition (ii) is verified if the inner Keplerian arcs z_I are transverse enough *to the domain's boundary ∂D. This obstruction to the good definition of F may be circumvented by considering a suitable prolonging according to which, whenever $|\beta_1| > \beta_{crit}$, the test particle returns back in the interior of the domain with an angle $\beta = \beta_1$. This extension, corresponding to the so-called* total reflection, *is somehow suggested by physical intuition and by analogy with the classical Birkhoff billiards, as well as the traditional Snell's law for light rays. However, as this approach leads to technical difficulties due to the passage to the tangent case for $|\beta_1| = \beta_{crit}$ and the consequent loss in regularity, in this work such extension is not considered, and, by definition, the map F will be well defined only when $|\beta_1| < \beta_{crit}$ for every inner arc* [5]. *Theorems 1.3.1 and 1.3.5 provide sufficient conditions for (i) to be satisfied, as well as proving the existence of a small neighbourhood of the homothetic initial conditions for which the inner arc is arbitrarily transverse to ∂D. As a consequence, even though the global definition of the first return map F can not be assured without additional requirement on γ, the hypotheses of the existence theorems guarantee that* the map is locally well defined near to the homothetic solutions.

As we will see in some specific cases, there are particular conditions on which the homothetic solutions are not the only one-periodic solutions of Problem (1.4.1). On the other hand, the study of the stability of this particular class of points allows us to derive important informations on the behaviour of F.

1.5 Stability analysis of the homothetic fixed points of F

1.5.1 The Jacobian matrix of F

Without loss of generality, let us assume that, in complex notation, $\bar{\xi} \in I$ is such that $\gamma(\bar{\xi}) = \|\gamma(\bar{\xi})\|e^{i\bar{\xi}}$ and $\dot{\gamma}(\bar{\xi}) = \|\dot{\gamma}(\bar{\xi})\|ie^{i\bar{\xi}}$. Then the point $\bar{p} = (\bar{\xi}, 0)$ is a fixed point for F, whose stability properties can be deduced from the spectral properties of the Jacobian matrix

$$DF((\bar{\xi}, 0)) = \begin{pmatrix} \dfrac{\partial \xi_1}{\partial \xi_0}_{|\bar{p}} & \dfrac{\partial \xi_1}{\partial \alpha_0}_{|\bar{p}} \\ \dfrac{\partial \alpha_1}{\partial \xi_0}_{|\bar{p}} & \dfrac{\partial \alpha_1}{\partial \alpha_0}_{|\bar{p}} \end{pmatrix} \tag{1.5.1}$$

which can be derived through the implicit function theorem, even though F is not explicitly determined.

Let us consider a generic potential $V(z)$ and, once fixed $z_0, z_1 \in \mathbb{R}^2$, consider the

---

[5]As we will see in Sections 2.3.1 and 2.4.1, this is true if D is a disk or a perturbation of the latter, provided $\alpha_0 \neq \pm\pi/2$.

function $z(s) = z(s; z_0, z_1)$ which solves the fixed end problem (1.2.2). As already seen in Section 1.2, $z(t; z_0, z_1) = z(s(t); z_0, z_1)$ is a critical point for the Jacobi length $L(y(t))$ with endpoints $y(0) = z_0$ $y(1) = z_1$. We denote $L(z(t; z_0, z_1))$ the value of this length. Note that neither the outer nor the inner arcs are global minimizers of the Jacobi length with fixed ends. Indeed, it can be proved (though is not relevant in this context) that the inner arc is a local minimizer while the outer one has Morse index one (cfr [71, 72]). We recall from Section 1.2 that:

- t is the geodesic time, whose relation with the kinetic time s is discussed in Section 1.2;

- $L(y(t)) = \int_0^1 \|\dot{y}(t)\| \sqrt{V(y(t))} dt;$

- $L = L(z(t; z_0, z_1)) = \|\dot{z}(t)\| \sqrt{V(z(t))} = const.$

If we consider a generic unit vector e, recalling and generalizing (1.2.5) the directional derivatives of $d(z_0, z_1)$ with respect to the first or second variable, denoted respectively with v and w, can be written as

$$\partial_{e,v} d(z_0, z_1) = \nabla_{z_0} d(z_0, z_1) \cdot e, \quad \partial_{e,w} d(z_0, z_1) = \nabla_{z_1} d(z_0, z_1) \cdot e$$
$$\nabla_{z_0} d(z_0, z_1) = -\sqrt{V(z(0))} \frac{\dot{z}(0)}{\|\dot{z}(0)\|}, \quad \nabla_{z_1} d(z_0, z_1) = \sqrt{V(z(1))} \frac{\dot{z}(1)}{\|\dot{z}(1)\|}. \tag{1.5.2}$$

Note that the derivatives are expressed in terms of geodesic time. Let us now define the generating function

$$S(\xi_0, \xi_1) = d(\gamma(\xi_0), \gamma(\xi_1)),$$

and define the tangent unit vectors $e_0 = \dot{\gamma}(\xi_0)/\|\dot{\gamma}(\xi_0)\|$ and $e_1 = \dot{\gamma}(\xi_1)/\|\dot{\gamma}(\xi_1)\|$. From (1.2.5),

$$\partial_{\xi_0} S(\xi_0, \xi_1) = \nabla_{z_0} d(\gamma(\xi_0), \gamma(\xi_1)) \cdot e_0,$$
$$\partial_{\xi_1} S(\xi_0, \xi_1) = \nabla_{z_1} d(\gamma(\xi_0), \gamma(\xi_1)) \cdot e_1. \tag{1.5.3}$$

Turning to the trajectory $z_{EI}(s)$ which describes a complete cycle exterior-interior, we can use the previous formulas to describe some geometric properties of the latter. Referring to (1.4.1) and further equations, define:

- $\xi_0, \tilde{\xi}, \xi_1 \in I$ such that $\gamma(\xi_0) = z_0$, $\gamma(\tilde{\xi}) = z_1$, $\gamma(\xi_1) = z_2$;

- α_0 the angle between v_0 with $n(\xi_0)$;

- β_0, β_1 respectively the angles of $z'_E(T_E)$ and $z'_I(T_E)$ with $n(\tilde{\xi})$;

- α'_1, α_1 respectively the angles of $z'_I(T_E + T_I)$ and $z'^{(1)}_E(T_E + T_I)$ with $n(\xi_1)$.

Then, from (1.5.2), (1.5.3) and Snell's law, one finds the relations

$$
\begin{aligned}
-\sqrt{V_E(\gamma(\xi_0))}\sin\alpha_0 &= \partial_{\xi_0} S_E(\xi_0, \tilde\xi), \\
\sqrt{V_E(\gamma(\tilde\xi))}\sin\beta_0 &= \partial_{\xi_1} S_E(\xi_0, \tilde\xi), \\
-\sqrt{V_I(\gamma(\tilde\xi))}\sin\beta_1 &= \partial_{\xi_0} S_I(\tilde\xi, \xi_1), \\
\sqrt{V_I(\gamma(\xi_1))}\sin\alpha'_1 &= \partial_{\xi_1} S(\tilde\xi, \xi_1), \\
\sqrt{V_E(\gamma(\tilde\xi))}\sin\alpha_0 &= \sqrt{V_I(\gamma(\tilde\xi))}\sin\beta_0 \\
\sqrt{V_I(\gamma(\xi_1))}\sin\alpha'_1 &= \sqrt{V_E(\gamma(\xi_1))}\sin\alpha_1,
\end{aligned}
\tag{1.5.4}
$$

where S_E and S_I refer respectively to d_E and d_I. Removing β_0, β_1 and α'_1 from (1.5.4), one obtains

$$
\begin{aligned}
\partial_{\xi_0} S_E(\xi_0, \tilde\xi) + \sqrt{V_E(\gamma(\xi_0))}\sin\alpha_0 &= 0, \\
\partial_{\xi_1} S_E(\xi_0, \tilde\xi) + \partial_{\xi_0} S_I(\tilde\xi, \xi_1) &= 0, \\
\partial_{\xi_1} S_I(\tilde\xi, \xi_1) - \sqrt{V_E(\gamma(\xi_1))}\sin\alpha_1 &= 0.
\end{aligned}
\tag{1.5.5}
$$

We can then define the function

$$
\Phi = [\eta_\xi^-, \eta_\xi^+] \times [\eta_\alpha^-, \eta_\alpha^+] \times [\eta_\xi^-, \eta_\xi^+] \times [\eta_\xi^-, \eta_\xi^+] \times [\eta_\alpha^-, \eta_\alpha^+] \to \mathbb{R}^3,
$$

$$
(\xi_0, \alpha_0, \tilde\xi, \xi_1, \alpha_1) \mapsto
\begin{pmatrix}
\Phi_1(\xi_0, \alpha_0, \tilde\xi, \xi_1, \alpha_1) \\
\Phi_2(\xi_0, \alpha_0, \tilde\xi, \xi_1, \alpha_1) \\
\Phi_3(\xi_0, \alpha_0, \tilde\xi, \xi_1, \alpha_1)
\end{pmatrix}
=
\begin{pmatrix}
\dfrac{\partial_{\xi_0} S_E(\xi_0, \tilde\xi)}{\sqrt{V_E(\gamma(\xi_0))}} + \sin\alpha_0 \\
\partial_{\xi_1} S_E(\xi_0, \tilde\xi) + \partial_{\xi_0} S_I(\tilde\xi, \xi_1) \\
\sin\alpha_1 - \dfrac{\partial_{\xi_1} S_I(\tilde\xi, \xi_1)}{\sqrt{V_E(\gamma(\xi_1))}}
\end{pmatrix},
$$

where $[\eta_\xi^-, \eta_\xi^+]$ and $[\eta_\alpha^-, \eta_\alpha^+]$ are neighbourhoods respectively of $\bar\xi$ and 0 such that the inner and outer dynamics are well defined (we remark that the existence of such neighbourhoods is assured by Theorems 1.3.1 and 1.3.5).
If $\xi_0, \tilde\xi$ and ξ_1 define respectively the initial, junction and final point of $z_{EI}(s)$, and α_0, α_1 are the angles of the initial and final velocity vectors of $z_{EI}(s)$ with the direction normal to ∂D in the initial and final points, then, from (1.5.5), $\Phi((\xi_0, \alpha_0, \tilde\xi, \xi_1, \alpha_1)) = 0$. The point $\bar q$ which describes the homothetic solution defined in Remark 1.4.1, which we call $\hat z_0(s)$, is given by $\bar q = (\bar\xi, 0, \bar\xi, \bar\xi, 0)$: clearly, $\Phi(\bar q) = 0$.
Under the hypothesis of nonsingularity of the matrix

$$
D_{(\tilde\xi, \xi_1, \alpha_1)}\Phi(\bar q) =
\begin{pmatrix}
\dfrac{\partial\Phi_1}{\partial\tilde\xi}\Big|_{\bar q} & \dfrac{\partial\Phi_1}{\partial\xi_1}\Big|_{\bar q} & \dfrac{\partial\Phi_1}{\partial\alpha_1}\Big|_{\bar q} \\
\dfrac{\partial\Phi_2}{\partial\tilde\xi}\Big|_{\bar q} & \dfrac{\partial\Phi_2}{\partial\xi_1}\Big|_{\bar q} & \dfrac{\partial\Phi_2}{\partial\alpha_1}\Big|_{\bar q} \\
\dfrac{\partial\Phi_3}{\partial\tilde\xi}\Big|_{\bar q} & \dfrac{\partial\Phi_3}{\partial\xi_1}\Big|_{\bar q} & \dfrac{\partial\Phi_3}{\partial\alpha_1}\Big|_{\bar q}
\end{pmatrix},
\tag{1.5.6}
$$

which we will prove in Section 1.5.4, the implicit function theorem guarantees the existence of a function $\Psi : I_1 \times J_1 \to I_2 \times I_3 \times J_2$, $(\xi_0, \alpha_0) \mapsto (\tilde{\xi}(\xi_0, \alpha_0), \xi_1(\xi_0, \alpha_0), \alpha_1(\xi_0, \alpha_0))$, where I_1, I_2, I_3 and J_1, J_2 are suitable neighbourhoods respectively of $\bar{\xi}$ and 0, such that

$$\forall (\xi_0, \alpha_0) \in I_1 \times J_1 \quad \Phi((\xi_0, \alpha_0, \Psi((\xi_0, \alpha_0)))) = 0.$$

Moreover, defined

$$D_{(\xi_0,\alpha_0)}\Phi(\bar{q}) = \begin{pmatrix} \dfrac{\partial \Phi_1}{\partial \xi_0}_{|\bar{q}} & \dfrac{\partial \Phi_1}{\partial \alpha_0}_{|\bar{q}} \\[2mm] \dfrac{\partial \Phi_2}{\partial \xi_0}_{|\bar{q}} & \dfrac{\partial \Phi_2}{\partial \alpha_0}_{|\bar{q}} \\[2mm] \dfrac{\partial \Phi_3}{\partial \xi_0}_{|\bar{q}} & \dfrac{\partial \Phi_3}{\partial \alpha_0}_{|\bar{q}} \end{pmatrix}, \quad D_{(\xi_0,\alpha_0)}\Psi(\bar{p}) = \begin{pmatrix} \dfrac{\partial \tilde{\xi}}{\partial \xi_0}_{|(\bar{\xi},0)} & \dfrac{\partial \tilde{\xi}}{\partial \alpha_0}_{|(\bar{\xi},0)} \\[2mm] \dfrac{\partial \xi_1}{\partial \xi_0}_{|(\bar{\xi},0)} & \dfrac{\partial \xi_1}{\partial \alpha_0}_{|(\bar{\xi},0)} \\[2mm] \dfrac{\partial \alpha_1}{\partial \xi_0}_{|(\bar{\xi},0)} & \dfrac{\partial \alpha_1}{\partial \alpha_0}_{|(\bar{\xi},0)} \end{pmatrix},$$

one has that

$$D_{(\xi_0,\alpha_0)}\Psi(\bar{p}) = -(D_{(\tilde{\xi},\xi_1,\alpha_1)}\Phi(\bar{q}))^{-1} D_{(\xi_0,\alpha_0)}\Phi(\bar{q}).$$

Recalling (1.5.1), we see that $DF(\bar{p})$ is composed by the last two rows of $D_{(\xi_0,\alpha_0)}\Psi(\bar{p})$. To compute (1.5.6), the second derivatives of $S_E(\xi_0, \tilde{\xi})$ and $S_I(\tilde{\xi}, \xi_1)$ computed in $\bar{p}$ are needed.

1.5.2 Outer dynamics: computation of the derivatives of $S_E(\xi_0, \tilde{\xi})$

Let us define $\hat{z}_E(s) = z_E(s; \gamma(\bar{\xi}), \gamma(\bar{\xi}))$ the homothetic solution of problem (1.4.3) (without loss of generality, suppose that it is defined in $[-T, T]$ for some $T > 0$ to be determined). Recalling that, from the initial assumptions on $\bar{\xi}$, $\gamma(\bar{\xi}) = \|\gamma(\bar{\xi})\| e^{i\bar{\xi}}$ and $\dot{\gamma}(\bar{\xi}) = \|\dot{\gamma}(\bar{\xi})\| i e^{i\bar{\xi}}$, we have that $\hat{z}_E(s) = x_0^E(s) e^{i\bar{\xi}}$, where $x_0^E(s) : [-T, T] \to \mathbb{R}$ is a solution of the one-dimensional fixed-end problem

$$\begin{cases} x_0^{E''}(s) = -\omega^2 x_0^E(s), & s \in [-T, T] \\[2mm] \dfrac{1}{2}|x_0^{E'}(s)|^2 + \dfrac{\omega^2}{2}|x_0^E(s)|^2 - \mathcal{E} = 0, & s \in [-T, T] \\[2mm] x_0^E(-T) = x_0^E(T) = \|\gamma(\bar{\xi})\|. \end{cases}$$

Then we have

$$\hat{z}_E(s) = \frac{\sqrt{2\mathcal{E}}}{\omega} \cos(\omega s) e^{i\bar{\xi}}, \quad T = \frac{1}{\omega} \arccos\left(\frac{\omega\|\gamma(\bar{\xi})\|}{\sqrt{2\mathcal{E}}}\right),$$

$$\hat{z}_E'(-T) = -\hat{z}_E'(T) = \sqrt{2\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2} e^{i\bar{\xi}};$$

$$\tag{1.5.7}$$

taking into account (1.5.2) and the relations

$$\frac{d}{dt} = \frac{L}{\sqrt{2}V(z(t(s)))}\frac{d}{ds} \Rightarrow \begin{cases} \sqrt{V_E(z_0^E(0))}\dfrac{\dot{z}(0)}{\|\dot{z}(0)\|} = \dfrac{1}{\sqrt{2}}z'(-T) \\ \sqrt{V_E(z_0^E(1))}\dfrac{\dot{z}(1)}{\|\dot{z}(1)\|} = \dfrac{1}{\sqrt{2}}z'(T) \end{cases}$$

one has

$$\partial_{\xi_0}S_E(\bar{\xi},\bar{\xi}) = \nabla_{z_0}d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\cdot\dot{\gamma}(\bar{\xi}) = -\frac{\sqrt{2\mathcal{E}-\omega^2\|\gamma(\bar{\xi})\|^2}}{\sqrt{2}}e^{i\bar{\xi}}\cdot\dot{\gamma}(\bar{\xi}) = 0, \tag{1.5.8}$$

$$\partial_{\xi_1}S_E(\bar{\xi},\bar{\xi}) = 0.$$

As for the second derivatives, we have

$$\begin{aligned} \partial_{\xi_0}^2 S_E(\bar{\xi},\bar{\xi}) &= \nabla_{z_0}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi})\cdot\dot{\gamma}(\bar{\xi}) + \nabla_{z_0}d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\cdot\ddot{\gamma}(\bar{\xi}) \\ \partial_{\xi_1}^2 S_E(\bar{\xi},\bar{\xi}) &= \nabla_{z_1}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi})\cdot\dot{\gamma}(\bar{\xi}) + \nabla_{z_1}d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\cdot\ddot{\gamma}(\bar{\xi}) \\ \partial_{\xi_0,\xi_1}^2 S_E(\bar{\xi},\bar{\xi}) &= \nabla_{z_0,z_1}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi})\cdot\dot{\gamma}(\bar{\xi}) \\ \partial_{\xi_1,\xi_0}^2 S_E(\bar{\xi},\bar{\xi}) &= \nabla_{z_1,z_0}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi})\cdot\dot{\gamma}(\bar{\xi}), \end{aligned} \tag{1.5.9}$$

where, defining $\bar{e} = ie^{i\bar{\xi}}$,

$$\nabla_{z_0}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi}) = \|\dot{\gamma}(\bar{\xi})\|\partial_{\bar{e},v}\left(\partial_{\bar{e},v}d_E\right)(\gamma(\bar{\xi}),\gamma(\bar{\xi})) = \|\dot{\gamma}(\bar{\xi})\|\partial_{\bar{e},v}\left(-\frac{1}{\sqrt{2}}\dot{z}_E'(-T)\right)$$

$$= -\frac{\|\dot{\gamma}(\bar{\xi})\|}{\sqrt{2}}\frac{d}{ds}\left(\partial_{\bar{e},v}\hat{z}_E\right)(-T), \tag{1.5.10}$$

and, similarly,

$$\nabla_{z_1}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi}) = \frac{\|\dot{\gamma}(\bar{\xi})\|}{\sqrt{2}}\frac{d}{ds}\left(\partial_{\bar{e},w}\hat{z}_E\right)(T),$$

$$\nabla_{z_0,z_1}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi}) = -\frac{\|\dot{\gamma}(\bar{\xi})\|}{\sqrt{2}}\frac{d}{ds}\left(\partial_{\bar{e},w}\hat{z}_E\right)(-T), \tag{1.5.11}$$

$$\nabla_{z_1,z_0}^2 d_E(\gamma(\bar{\xi}),\gamma(\bar{\xi}))\dot{\gamma}(\bar{\xi}) = \frac{\|\dot{\gamma}(\bar{\xi})\|}{\sqrt{2}}\frac{d}{ds}\left(\partial_{\bar{e},v}\hat{z}_E\right)(T).$$

The functions $\partial_{\bar{e},v}\hat{z}_E(s)$ and $\partial_{\bar{e},w}\hat{z}_E(s)$ are the first-order variations of $\hat{z}_E(s)$ with respect to the variation respectively of its first and second endpoint along the unit vector $\bar{e}$, which is orthogonal to $\hat{z}_E(s)$. If we define $\tilde{f}_0(s) = \partial_{\bar{e},v}\hat{z}_E(s)$ and $\tilde{f}_1(s) = \partial_{\bar{e},w}\hat{z}_E(s)$, we have that $\tilde{f}_0(s) = f_0(s)\bar{e}$ and $\tilde{f}_1(s) = f_1(s)\bar{e}$, with $f_0, f_1 : [-T,T] \to \mathbb{R}$ to be determined. Consider $z(t) = \hat{z}_E(t) + \tilde{f}_0(t)$ the geodesics obtained by varying the first endpoint of $\hat{z}_E(t)$ in the direction of $\dot{\gamma}(\bar{\xi})$ expressed with respect to the geodesic time t: it solves

the Euler-Lagrange equation with $\mathcal{L} = \|\dot{z}(t)\|^2 V_E(z(t))$, namely,

$$0 = -\frac{d}{dt}\left(2\dot{z}(t)V(z(t))\right) + \|\dot{z}(t)\|^2 \nabla V_E(z(t)) =$$

$$= -\frac{d}{dt}\left(2(\dot{\hat{z}}E(t) + \dot{\tilde{f}}_0(t))V(\hat{z}_E(t) + \tilde{f}_0(t))\right) + \|\dot{\hat{z}}_E(t) + \dot{\tilde{f}}_0(t)\|^2 \nabla V_E(\hat{z}_E(t) + \tilde{f}_0(t)) =$$

$$= -2\frac{d}{dt}\left(\dot{\tilde{f}}_0(t)V(\hat{z}_E(t))\right) + \|\dot{z}_0(t)\|^2 \nabla^2 V_E(\hat{z}_E(t))\tilde{f}_0(t) =$$

$$= -2\frac{\|\dot{\hat{z}}_E(t(s))\|}{\sqrt{2V_E(\hat{z}_E(t(s)))}}\frac{d}{ds}\left(V(\hat{z}_E(t(s)))\frac{\|\dot{\hat{z}}_E(t(s))\|}{\sqrt{2V_E(\hat{z}_E(t(s)))}}\frac{d}{ds}\tilde{f}_0(t(s))\right) +$$

$$+ \|\dot{z}_0(t(s))\|^2 \nabla^2 V_E(\hat{z}_E(t(s)))\tilde{f}_0(t(s))$$

$$\Rightarrow \tilde{f}_0''(s) - \nabla^2 V_E(\hat{z}_E(s))\tilde{f}_0(s) = 0,$$

where we took only the first-order terms and used the transformation rules between d/dt and d/ds, the conservation of $L = \|\dot{\hat{z}}_E(t(s))\|\sqrt{V_E(\hat{z}_E(t(s)))}$ and the Euler-Lagrange equation for $\hat{z}_E(t)$.

Since $\tilde{f}_0(-T) = \bar{e}$ and $\tilde{f}_0(T) = \mathbf{0}$, $f_0(s)$ solves the one-dimensional system

$$\begin{cases} f_0''(s) = -\omega^2 f_0(s), & s \in [-T, T] \\ f_0(-T) = 1, f_0(T) = 0, \end{cases}$$

namely, recalling the definition of T in (1.5.7),

$$\tilde{f}_0(s) = \frac{1}{2}\left(\frac{\sqrt{2\mathcal{E}}}{\omega\|\gamma(\bar{\xi})\|}\cos(\omega s) - \frac{\sqrt{2\mathcal{E}}}{\sqrt{2\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2}}\sin(\omega s)\right)\bar{e}.$$

With the same reasoning and taking into account that $\tilde{f}_1(-T) = \mathbf{0}$ and $\tilde{f}_1(T) = \bar{e}$, we have that $\tilde{f}_1(s) = \tilde{f}_0(-s)$, and then we can finally find

$$\begin{aligned}
\frac{d}{ds}\partial_{\bar{e},v}z_0^E(-T) &= \frac{\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2}{\|\gamma(\bar{\xi})\|\sqrt{2\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2}}\bar{e}, \\
\frac{d}{ds}\partial_{\bar{e},v}z_0^E(T) &= -\frac{\mathcal{E}}{\|\gamma(\bar{\xi})\|\sqrt{2\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2}}\bar{e}, \\
\frac{d}{ds}\partial_{\bar{e},w}z_0^E(-T) &= \frac{\mathcal{E}}{\|\gamma(\bar{\xi})\|\sqrt{2\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2}}\bar{e}, \\
\frac{d}{ds}\partial_{\bar{e},w}z_0^E(T) &= -\frac{\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2}{\|\gamma(\bar{\xi})\|\sqrt{2\mathcal{E} - \omega^2\|\gamma(\bar{\xi})\|^2}}\bar{e}.
\end{aligned} \tag{1.5.12}$$

Taking together (1.5.9), (1.5.10), (1.5.11) and (1.5.12), one can find the analytical expressions of the second derivatives of $S_E(\xi_0, \xi_1)$, computed for $\xi_0 = \xi_1 = \bar{\xi}$:

$$\partial_{\xi_0}^2 S_E(\bar{\xi}, \bar{\xi}) = \partial_{\xi_1}^2 S(\bar{\xi}, \bar{\xi}) = -\frac{\|\dot{\gamma}(\bar{\xi})\|^2}{2\|\gamma(\bar{\xi})\|} \frac{\mathcal{E} - \omega^2 \|\gamma(\bar{\xi})\|^2}{\sqrt{V_E(\gamma(\bar{\xi}))}} - \sqrt{V_E(\gamma(\bar{\xi}))} e^{i\bar{\xi}} \cdot \ddot{\gamma}(\bar{\xi}),$$

$$\partial_{\xi_0,\xi_1}^2 S_E(\bar{\xi}, \bar{\xi}) = \partial_{\xi_1,\xi_0}^2 S(\bar{\xi}, \bar{\xi}) = -\frac{\|\dot{\gamma}\|^2}{2\|\gamma(\bar{\xi})\|} \frac{\mathcal{E}}{\sqrt{V_E(\gamma(\bar{\xi}))}}. \tag{1.5.13}$$

If γ is parametrised by arc length, $\|\dot{\gamma}(\bar{\xi})\| = 1$ and $\ddot{\gamma}(\bar{\xi}) = -k(\bar{\xi})n(\bar{\xi}) = -k(\bar{\xi})e^{i\bar{\xi}}$, where $k(\bar{\xi})$ is the curvature of γ in $\bar{\xi}$: Eqs.(1.5.13) simplify then in

$$\partial_{\xi_0}^2 S_E(\gamma(\bar{\xi}), \gamma(\bar{\xi})) = \partial_{\xi_1}^2 S_E(\gamma(\bar{\xi}), \gamma(\bar{\xi}))$$

$$= \frac{\mathcal{E}}{2\|\gamma(\bar{\xi})\|\sqrt{V_E(\gamma(\bar{\xi}))}} + \frac{\sqrt{V_E(\gamma(\bar{\xi}))}}{\|\gamma(\bar{\xi})\|} \left(\|\gamma(\bar{\xi})\|k(\bar{\xi}) - 1 \right),$$

$$\partial_{\xi_0,\xi_1}^2 S_E(\gamma(\bar{\xi}), \gamma(\bar{\xi})) = \partial_{\xi_1,\xi_0}^2 S_E(\gamma(\bar{\xi}), \gamma(\bar{\xi})) = -\frac{\mathcal{E}}{2\|\gamma(\bar{\xi})\|\sqrt{V_E(\gamma(\bar{\xi}))}}. \tag{1.5.14}$$

Eq. (1.5.14) highlights that the second term in $\partial_{\xi_0}^2 S(\gamma(\bar{\xi}), \gamma(\bar{\xi}))$ represents a perturbation of the homogeneous second derivative with respect to the circular case, where $\left(\|\gamma(\bar{\xi})\|k(\bar{\xi}) - 1 \right) = 0$ for every $\bar{\xi} \in I$.

1.5.3 Inner dynamics: computation of the derivatives of $S_I(\tilde{\xi}, \xi_1)$

With reference to the Notation 1.3.11, from Proposition 1.3.4 we know that the fixed ends inner problem

$$\begin{cases} (HS_I)[z(s)] & s \in [0, T_I], \\ z_I(0) = z_0^I, z_I(T_I) = z_1^I. \end{cases} \tag{1.5.15}$$

is conjugated, by means of the Levi-Civita transformations, to the regularised problem

$$\begin{cases} (HS_{LC})[w(\tau)] & \tau \in [-T, T], \\ w(-T) = w_0, w(T) = w_1 \end{cases}$$

where $\Omega^2 = 2(\mathcal{E} + h)$, $E = \mu$, $w_0^2 = z_0^I$, $w_1^2 = z_1^I$ and $\tau = \tau(s)$ such that $\frac{d\tau}{ds} = \frac{1}{2\|z(s)\|}$. In the following, we will work with the Levi-Civita variables, taking respectively for w_0 the negative determination of the square root of z_0^I and for w_1 the positive determination of the square root of z_1^I, namely, in polar coordinates,

$$z_0^I = \|z_0^I\|e^{i\theta_0} \Rightarrow w_0 = -\sqrt{\|z_0^I\|}e^{i\frac{\theta_0}{2}}, \quad z_1^I = \|z_1^I\|e^{i\theta_1} \Rightarrow w_1 = \sqrt{\|z_1^I\|}e^{i\frac{\theta_1}{2}}. \tag{1.5.16}$$

To compute the derivatives of $S_I(\gamma(\xi_0), \gamma(\xi_1))$, define then

$$L_I(z(t)) = \int_0^1 \|\dot{z}(t)\| \sqrt{V_I(z(t))} dt = \int_0^1 \|\dot{z}(t)\| \sqrt{\mathcal{E} + h + \frac{\mu}{\|z(t)\|}} dt,$$

where $t \in [0, 1]$ is the usual geodesic time. Passing to the Levi-Civita plane:

$$L_I(z) = 2 \int_0^1 \|\dot{w}(t)\| \sqrt{2(\mathcal{E} + h)\frac{\|w(t)\|^2}{2} + \mu}\, dt$$

$$= 2 \int_0^1 \|\dot{w}(t)\| \sqrt{\frac{\Omega^2}{2}\|w(t)\|^2 + E}\, dt = 2\tilde{L}_I(w)$$

According to the choice for the initial and final point of $\omega(\tau)$ defined in (1.5.16), in the Levi-Civita plane the function $S(\xi_0, \xi_1)$ can be written as

$$S_I(\xi_0, \xi_1) = d_I(\gamma(\xi_0), \gamma(\xi_1)) = 2\tilde{d}_I(\phi_-(\xi_0), \phi_+(\xi_1)) = 2\tilde{S}_I(\xi_0, \xi_1),$$

where $\tilde{d}_I$ is the distance associated to $\tilde{L}_I$ and $\phi_-(\xi), \phi_+(\xi)$ are as in Definition 1.3.7. We can then compute the transformed of $\gamma(\bar{\xi})$, $\dot{\gamma}(\bar{\xi})$, $\ddot{\gamma}(\bar{\xi})$, seen both as initial and ending point of our arc. Without loss of generalization, let us suppose that $\gamma(\bar{\xi}) = \|\gamma(\bar{\xi})\|(1, 0)$ and $\dot{\gamma}(\bar{\xi}) = \|\dot{\gamma}(\bar{\xi})\|(0, 1)$: using the relation $\phi_\pm(\xi)^2 = \gamma(\xi)$, one has

$$\phi_-(\bar{\xi}) = \sqrt{\|\gamma(\bar{\xi})\|}(-1, 0), \quad \phi_+(\bar{\xi}) = \sqrt{\|\gamma(\bar{\xi})\|}(1, 0)$$

$$\dot{\phi}_-(\bar{\xi}) = \frac{\|\dot{\gamma}(\bar{\xi})\|}{2\sqrt{\|\gamma(\bar{\xi})\|}}(0, -1) = \frac{\|\dot{\gamma}(\bar{\xi})\|}{2\sqrt{\|\gamma(\bar{\xi})\|}}t_-(\bar{\xi}),$$

$$\dot{\phi}_+(\bar{\xi}) = \frac{\|\dot{\gamma}(\bar{\xi})\|}{2\sqrt{\|\gamma(\bar{\xi})\|}}(0, 1) = \frac{\|\dot{\gamma}(\bar{\xi})\|}{2\sqrt{\|\gamma(\bar{\xi})\|}}t_+(\bar{\xi})$$

where $t_-(\bar{\xi}) = (0, -1)$ and $t_+(\bar{\xi}) = (0, 1)$, and $\ddot{\phi}_\pm(\bar{\xi})$ satisfy the equations $\ddot{\gamma}(\bar{\xi}) = 2(\dot{\phi}_\pm^2(\bar{\xi}) + \phi_\pm(\bar{\xi})\ddot{\phi}_\pm(\bar{\xi}))$.

In order to compute the derivatives of $S_I(\bar{\xi}, \bar{\xi})$, we can use the same techniques used in Section 1.5.2 for the outer dynamics, taking into account that, in the Levi-Civita plane, the starting and final point are different.

Let us start with the derivation of the homothetic equilibrium orbit: in the physical plane, it corresponds to the ejection-collision solution $\hat{z}_I(s)$ of the fixed-end problem

$$\begin{cases} (HS_I)[z(s)], & s \in [0, T_I], \\ z(0) = z(T_I) = \gamma(\bar{\xi}), \end{cases}$$

which corresponds, in the Levi-Civita variables, to the solution $w_0(\tau)$ of the problem

$$\begin{cases} (HS_{LC})[w(\tau)], & \tau \in [-T, T], \\ w(-T) = \phi_-(\bar{\xi}), w(T) = \phi_+(\bar{\xi}), \end{cases}$$

from which one obtains

$$w_0(\tau) = \frac{\sqrt{2E}}{\Omega} \sinh(\Omega\tau)(1,0), \quad T = \frac{1}{\Omega} \operatorname{arcsinh}\left(\frac{\Omega\sqrt{\|\gamma(\bar{\xi})\|}}{\sqrt{2E}}\right)$$

$$\Rightarrow w_0'(-T) = w_0'(T) = \sqrt{2E + \Omega^2 \|\gamma(\bar{\xi})\|}(1,0).$$

Proceeding as in (1.5.8), one has then

$$\partial_{\xi_0} S_I(\bar{\xi}, \bar{\xi}) = -\frac{2}{\sqrt{2}} w_0'(-T) \cdot \dot{\phi}_-(\bar{\xi}) = 0, \quad \partial_{\xi_1} S_I(\bar{\xi}, \bar{\xi}) = \frac{2}{\sqrt{2}} w_0'(T) \cdot \dot{\phi}_+(\bar{\xi}) = 0.$$

As for the second derivatives, taking into account that $\phi_\pm$ are not parametrised by arc length:

$$
\begin{aligned}
\partial_{\xi_0}^2 S_I(\bar{\xi}, \bar{\xi}) \quad &= 2\partial_{\xi_0}^2 \tilde{S}_I(\bar{\xi}, \bar{\xi}) \\
&= 2\nabla_{w_0}^2 \tilde{d}_I(\phi_-(\bar{\xi}), \phi_+(\bar{\xi}))\dot{\phi}_-(\bar{\xi}) \cdot \dot{\phi}_-(\bar{\xi}) + 2\nabla_{w_0}\tilde{d}_I(\phi_-(\bar{\xi}), \phi_+(\bar{\xi})) \cdot \ddot{\phi}_-(\bar{\xi}) \\
&= \frac{\|\dot{\gamma}(\bar{\xi})\|^2}{2\|\gamma(\bar{\xi})\|} \nabla_{w_0}^2 \tilde{d}_I(\phi_-(\bar{\xi}), \phi_+(\bar{\xi}))t_-(\bar{\xi}) \cdot t_-(\bar{\xi}) \\
&\quad + 2\nabla_{w_0}\tilde{d}_I(\phi_-(\bar{\xi}), \phi_+(\bar{\xi})) \cdot \ddot{\phi}_-(\bar{\xi}) = \\
&= -\frac{\|\dot{\gamma}(\bar{\xi})\|^2}{2\sqrt{2}\|\gamma(\bar{\xi})\|} \frac{d}{d\tau}\left(\partial_{t_-(\bar{\xi}),v} w_0\right)(-T) \cdot t_-(\bar{\xi}) - \sqrt{2}w_0'(-T) \cdot \ddot{\phi}_-(\bar{\xi}), \\
\partial_{\xi_1}^2 S_I(\bar{\xi}, \bar{\xi}) \quad &= \frac{\|\dot{\gamma}(\bar{\xi})\|^2}{2\sqrt{2}\|\gamma(\bar{\xi})\|} \frac{d}{d\tau}\left(\partial_{t_+(\bar{\xi}),w} w_0\right)(T) \cdot t_+(\bar{\xi}) + \sqrt{2}w_0'(T) \cdot \ddot{\phi}_+(\bar{\xi}) \\
\partial_{\xi_0,\xi_1}^2 S_I(\bar{\xi}, \bar{\xi}) \quad &= -\frac{\|\dot{\gamma}(\bar{\xi})\|^2}{2\sqrt{2}\|\gamma(\bar{\xi})\|} \frac{d}{d\tau}\left(\partial_{t_+(\bar{\xi}),w} w_0\right)(-T) \cdot t_-(\bar{\xi}), \\
\partial_{\xi_1,\xi_0}^2 S_I(\bar{\xi}, \bar{\xi}) \quad &= \frac{|\dot{\gamma}(\bar{\xi})|^2}{2\sqrt{2}|\gamma(\bar{\xi})|} \frac{d}{d\tau}\left(\partial_{t_-(\bar{\xi}),v} w_0\right)(T) \cdot t_+(\bar{\xi}).
\end{aligned}
$$

$$(1.5.17)$$

We can compute the variations $\partial_{t_-(\bar{\xi}),v} w_0(\tau)$, $\partial_{t_+(\bar{\xi}),w} w_0(\tau)$ as in Section 1.5.2: by imposing $\partial_{t_-(\bar{\xi}),v} w_0(\tau) = \tilde{g}_0(\tau) = g_0(\tau)t_-(\bar{\xi})$ and $\partial_{t_+(\bar{\xi}),w} w_0(\tau) = \tilde{g}_1(\tau) = g_1(\tau)t_+(\bar{\xi})$ we have that $g_0(\tau)$ and $g_1(\tau)$ are solutions of the two one-dimensional systems

$$
\begin{cases} g_0''(\tau) = \Omega^2 g_0(\tau), \tau \in [-T, T] \\ g_0(-T) = 1, g_0(T) = 0, \end{cases}
\quad
\begin{cases} g_1''(\tau) = \Omega^2 g_1(\tau), \tau \in [-T, T] \\ g_1(-T) = 0, g_1(T) = 1, \end{cases}
$$

and then we obtain

$$\frac{d}{d\tau}(\partial_{t_-(\bar{\xi}),v}w_0)(-T) = -\frac{E + \Omega^2\|\gamma(\bar{\xi})\|}{\sqrt{\|\gamma(\bar{\xi})\|}\sqrt{2E + \Omega^2\|\gamma(\bar{\xi})\|}}t_-(\bar{\xi}),$$

$$\frac{d}{d\tau}(\partial_{t_-(\bar{\xi}),v}w_0)(T) = -\frac{E}{\sqrt{\|\gamma(\bar{\xi})\|}\sqrt{2E + \Omega^2\|\gamma(\bar{\xi})\|}}t_-(\bar{\xi}),$$

$$\frac{d}{d\tau}(\partial_{t_+(\bar{\xi}),w}w_0)(-T) = \frac{E}{\sqrt{\|\gamma(\bar{\xi})\|}\sqrt{2E + \Omega^2\|\gamma(\bar{\xi})\|}}t_+(\bar{\xi}),$$

$$\frac{d}{d\tau}(\partial_{t_+(\bar{\xi}),w}w_0)(T) = \frac{E + \Omega^2\|\gamma(\bar{\xi})\|}{\sqrt{\|\gamma(\bar{\xi})\|}\sqrt{2E + \Omega^2\|\gamma(\bar{\xi})\|}}t_+(\bar{\xi}).$$

Then, taking into account (1.5.17) and recalling that $E = \mu$, $\Omega^2 = \mathcal{E} + h$, we finally obtain (recall that we are assuming that $\gamma(\bar{\xi}) \parallel (1,0)$)

$$\partial^2_{\xi_0}S_I(\bar{\xi},\bar{\xi}) \quad = \frac{\|\dot{\gamma}(\bar{\xi})\|^2}{4\|\gamma(\bar{\xi})\|^2}\frac{\mu + 2(\mathcal{E} + h)\|\gamma(\bar{\xi})\|}{\sqrt{V_I(\gamma(\bar{\xi}))}} - 2\sqrt{\|\gamma(\bar{\xi})\|}\sqrt{V_I(\gamma(\bar{\xi}))}(1,0)\cdot\ddot{\phi}_-(\bar{\xi}),$$

$$\partial^2_{\xi_1}S_I(\bar{\xi},\bar{\xi}) \quad = \frac{\|\dot{\gamma}(\bar{\xi})\|^2}{4\|\gamma(\bar{\xi})\|^2}\frac{\mu + 2(\mathcal{E} + h)\|\gamma(\bar{\xi})\|}{\sqrt{V_I(\gamma(\bar{\xi}))}} + 2\sqrt{\|\gamma(\bar{\xi})\|}\sqrt{V_I(\gamma(\bar{\xi}))}(1,0)\cdot\ddot{\phi}_+(\bar{\xi}),$$

$$\partial^2_{\xi_0,\xi_1}S_I(\bar{\xi},\bar{\xi}) = \partial^2_{\xi_1,\xi_0}S_I(\bar{\xi},\bar{\xi}) = \frac{\|\dot{\gamma}(\bar{\xi})\|^2}{4\|\gamma(\bar{\xi})\|^2}\frac{\mu}{\sqrt{V_I(\gamma(\bar{\xi}))}}.$$

$$(1.5.18)$$

If $\gamma(\xi)$ is parametrised by arc length, $\|\dot{\gamma}(\bar{\xi})\| = 1$ and $\ddot{\gamma}(\bar{\xi})k(\bar{\xi}) = (-1,0)$, then

$$\ddot{\phi}_-(\bar{\xi}) = \frac{1}{2\sqrt{\|\gamma(\bar{\xi})\|}}\left(k(\bar{\xi}) - \frac{1}{2\|\gamma(\bar{\xi})\|}\right)(1,0) = -\ddot{\phi}_+,$$

and Eqs.(1.5.18) simplify as

$$\partial^2_{\xi_0}S_I(\bar{\xi},\bar{\xi}) = \partial^2_{\xi_1}S_I(\bar{\xi},\bar{\xi}) = -\frac{\mu}{4\|\gamma(\bar{\xi})\|^2\sqrt{V_I(\gamma(\bar{\xi}))}} - \sqrt{V_I(\gamma(\bar{\xi}))}\left(k(\bar{\xi}) - \frac{1}{\|\gamma(\bar{\xi})\|}\right),$$

$$\partial^2_{\xi_0,\xi_1}S_I(\bar{\xi},\bar{\xi}) = \partial^2_{\xi_1,\xi_0}S_I(\bar{\xi},\bar{\xi}) = \frac{\mu}{4\|\gamma(\bar{\xi})\|^2\sqrt{V_I(\gamma(\bar{\xi}))}}$$

$$(1.5.19)$$

1.5.4 Stability properties of $(\bar{\xi}, 0)$

Let us now suppose that $\gamma(\bar{\xi})$ is parametrised by arc length (the general case can be treated in the same way, taking into account the explicit expression of $\ddot{\gamma}(\bar{\xi})$): taking

together (1.5.14) and (1.5.19), one can see that they can be written in the form

$$\partial^2_{\xi_0} S_E(\bar{\xi}, \bar{\xi}) = \partial^2_{\xi_1} S_E(\bar{\xi}, \bar{\xi}) = E_0 + \varepsilon_E,$$

$$\partial^2_{\xi_0, \xi_1} S_E(\bar{\xi}, \bar{\xi}) = \partial^2_{\xi_1, \xi_0} S_E(\bar{\xi}, \bar{\xi}) = -E_0,$$

$$E_0 = \frac{\mathcal{E}}{2\|\gamma(\bar{\xi})\|\sqrt{V_E(\gamma(\bar{\xi}))}}, \qquad \varepsilon_E = (\|\gamma(\bar{\xi})\| k(\bar{\xi}) - 1) \frac{\sqrt{V_E(\gamma(\bar{\xi}))}}{\|\gamma(\bar{\xi})\|},$$

$$\partial^2_{\xi_0} S_I(\bar{\xi}, \bar{\xi}) = \partial^2_{\xi_1} S_I(\bar{\xi}, \bar{\xi}) = I_0 + \varepsilon_I,$$ (1.5.20)

$$\partial^2_{\xi_0, \xi_1} S_I(\bar{\xi}, \bar{\xi}) = \partial^2_{\xi_1, \xi_0} S_I(\bar{\xi}, \bar{\xi}) = -I_0,$$

$$I_0 = -\frac{\mu}{4\|\gamma(\bar{\xi})\|^2 \sqrt{V_I(\gamma(\bar{\xi}))}}, \qquad \varepsilon_I = -\left(k(\bar{\xi}) - \frac{1}{\|\gamma(\bar{\xi})\|}\right)\sqrt{V_I(\gamma(\bar{\xi}))},$$

The terms $\varepsilon_{E\backslash I}$ can be seen as the perturbations induced to the second derivatives when the domain's boundary ∂D is not a circle. Turning to the matrices defined in Section 1.5.1, we have that

$$D_{(\tilde{\xi}, \xi_1, \alpha_1)}\Phi(\bar{q}) = \begin{pmatrix} \dfrac{\partial^2_{\xi_0, \xi_1} S_E(\bar{\xi}, \bar{\xi})}{\sqrt{V_E(\gamma(\bar{\xi}))}} & 0 & 0 \\ \partial^2_{\xi_1} S_E(\bar{\xi}, \bar{\xi}) + \partial^2_{\xi_0} S_I(\bar{\xi}, \bar{\xi}) & \partial^2_{\xi_0, \xi_1} S_I(\bar{\xi}, \bar{\xi}) & 0 \\ -\dfrac{\partial_{\xi_0, \xi_1} S_I(\bar{\xi}, \bar{\xi})}{\sqrt{V_E(\gamma(\bar{\xi}))}} & -\dfrac{\partial^2_{\xi_1} S_I(\bar{\xi}, \bar{\xi})}{\sqrt{V_E(\gamma(\bar{\xi}))}} & 1 \end{pmatrix},$$

whose determinant is given by

$$\det\left(D_{(\tilde{\xi}, \xi_1, \alpha_1)}\Phi(\bar{q})\right) = -\frac{\mathcal{E}\mu}{8\|\gamma(\bar{\xi})\|^3 \sqrt{V_I(\gamma(\bar{\xi}))} V_E(\gamma(\bar{\xi}))} < 0$$

in the Hill's region $\mathcal{H}$. The implicit function theorem can be then applied and we have that there exist I_1, I_2, I_2, J_1, J_2 neighbourhoods respectively of $\bar{\xi}$ and 0 and there is a function $\Psi : I_1 \times J_1 \to I_2 \times I_3 \times J_2$ such that for every $(\xi_0, \alpha_0) \in I_1 \times J_1$ one has $\Phi(\xi_0, \alpha_0, \Psi(\xi_0, \alpha_0)) = 0$. Moreover

$$D_{(\xi_0, \alpha_0)}\Psi(\bar{p}) = -\left(D_{(\tilde{\xi}, \xi_1, \alpha_1)}\Phi(\bar{q})\right)^{-1} D_{(\xi_0, \alpha_0)}\Phi(\bar{q}).$$

The function $F : (\xi_0, \alpha_0) \mapsto (\xi_1(\xi_0, \alpha_0), \alpha_1(\xi_0, \alpha_0))$ is given by the last two components of Ψ, then $DF(\bar{p})$ is composed by the last two rows of $D_{(\xi_0, \alpha_0)}\Psi(\bar{p})$. Direct computations show that

$$DF(\bar{p}) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$ (1.5.21)

where

$$A_{11} = 1 + \frac{2\varepsilon_E + \varepsilon_I}{I_0} + \frac{\varepsilon_E(\varepsilon_E + \varepsilon_I + I_0)}{E_0 I_0},$$

$$A_{12} = \sqrt{V_E(\gamma(\bar{\xi}))} \left(\frac{1}{I_0} + \frac{1}{E_0} \right) + \sqrt{V_E(\gamma(\bar{\xi}))} \frac{\varepsilon_E + \varepsilon_I}{E_0 I_0},$$

$$A_{21} = \frac{2\varepsilon_E(\varepsilon_I + I_0) + \varepsilon_I(\varepsilon_I + 2I_0)}{I_0 \sqrt{V_E(\gamma(\bar{\xi}))}} + \frac{\varepsilon_E[\varepsilon_E(\varepsilon_I + I_0) + \varepsilon_I(\varepsilon_I + 2I_0)]}{E_0 I_0 \sqrt{V_E(\gamma(\bar{\xi}))}}$$

$$A_{22} = 1 + \frac{\varepsilon_E}{E_0} + \frac{\varepsilon_I(2I_0 + \varepsilon_I + E_0 + \varepsilon_E)}{E_0 I_0}$$

The stability properties of the equilibrium in $(\bar{\xi}, 0)$ can be studied looking at the eigenvalues of $DF(\bar{p})$, see [65]: let us denote them with λ_1 and λ_2. Direct computations show that $\det(DF(\bar{p})) = 1$: this is a completely general fact, as the map F describes an area-preserving system, and, from an algebraic point of view, implies that $\lambda_1 \lambda_2 = 1$. Therefore we can have two cases:

- $\lambda_1, \lambda_2 \in \mathbb{R} \Rightarrow \lambda_1 = 1/\lambda_2$: if $\lambda_1 \neq \pm 1$, then $(\bar{\xi}, 0)$ is an unstable saddle;

- $\lambda_1, \lambda_2 \in \mathbb{C}/\mathbb{R} \Rightarrow \lambda_1 = \bar{\lambda}_2$ and $\lambda_1, \lambda_2 \in \mathbb{S}^1$: then $(\bar{\xi}, 0)$ is a stable center.

We can distinguish between the two cases by considering the characteristic polynomial of $DF((\bar{p})$.

Remark 1.5.1. *Denoted by* $p(\lambda) = \lambda^2 + b\lambda + 1$ *the characteristic polynomial of* $DF(\bar{\xi}, 0)$, *let* $\Delta = b^2 - 4$ *its discriminant. Then*

- *if* $\Delta > 0 \Rightarrow (\bar{\xi}, 0)$ *is a saddle for F;*

- *if* $\Delta < 0 \Rightarrow (\bar{\xi}, 0)$ *is a center for F;*

The value of Δ with respect to the physical quantities of the problem can be directly computed: it results that

$$\Delta = ABCD,$$

where

$$A = \frac{16}{\mathcal{E}^2 \mu^2} \left(\sqrt{V_I(\gamma(\bar{\xi})} - \sqrt{V_E(\gamma(\bar{\xi}))} \right) \left(\|\gamma(\bar{\xi})\| k(\bar{\xi}) - 1 \right),$$

$$B = \mathcal{E} - \left(\|\gamma(\bar{\xi})\| k(\bar{\xi}) - 1 \right) \left(\sqrt{V_I(\gamma(\bar{\xi}))} - \sqrt{V_E(\gamma(\bar{\xi}))} \right) \sqrt{V_E(\gamma(\bar{\xi}))},$$

$$C = -\mu \sqrt{V_E(\gamma(\bar{\xi}))} + 2\|\gamma(\bar{\xi})\| B \sqrt{V_I(\gamma(\bar{\xi}))},$$

$$D = \mu + 2\|\gamma(\bar{\xi})\| \left(\|\gamma(\bar{\xi})\| k(\bar{\xi}) - 1 \right) \sqrt{V_I(\gamma(\bar{\xi}))} \left(\sqrt{V_I(\gamma(\bar{\xi}))} - \sqrt{V_E(\gamma(\bar{\xi}))} \right).$$

1.6 A direct investigation: elliptic domains

When the expression of $\gamma(\xi)$ is given, the general theory developed in Section 1.5 can be used to study the effective stability of the fixed points of the map F. In this Section we investigate the existence and stability of equilibrium orbits for our dynamical system when D is an elliptic domain. Let us suppose that ∂D is an ellipse with semimajor axis $a = 1$ and eccentricity $0 \le e < 1$. Denoted by $b = a\sqrt{1 - e^2} = \sqrt{1 - e^2}$ the semiminor axis, one can parametrise ∂D as

$$\gamma(\xi) = (\cos \xi, b \sin \xi), \ \xi \in [0, 2\pi],$$

which can be written as

$$\gamma(\theta) = (1 + f(e, \theta))e^{i\theta},$$
$$f(e, \theta) = \frac{\sqrt{1 - e^2}}{(1 - e^2 \cos^2 \theta)} - 1, \quad \theta \in [0, 2\pi], \quad e \in [0, 1).$$

From direct computations and from Remark 1.4.1, one has that the orbit with initial conditions $z(0) = \gamma(\bar{\xi})$, $z'(0) = \sqrt{V_E(z(0))}z(0)/\|z(0)\|$ is a homothetic equilibrium orbit if and only if $\bar{\xi} = k\pi/2$, $k \in \{0, 1, 2, 3, 4\}$: due to the symmetry of the problem, we can restrict our study to the two cases $\bar{\xi}_0 = 0$ and $\bar{\xi}_1 = \pi/2$. We have that

$$\gamma(0) = (1, 0), \qquad \dot{\gamma}(0) = (0, b), \qquad \ddot{\gamma}(0) = (-1, 0)$$
$$\gamma(\pi/2) = (0, b), \quad \dot{\gamma}(\pi/2) = (-1, 0), \quad \ddot{\gamma}(\pi/2) = (0, -b) :$$

The stability properties of the F-fixed points $(\bar{\xi}_0, 0)$ and $(\bar{\xi}_1, 0)$ can be deduced as in Section 1.5: in particular, from Eqs.(1.5.13) and (1.5.17), one obtains

$$\partial_{\xi_0}^2 S_I(\bar{\xi}_0, 0) = \partial_{\xi_1}^2 S_I(\bar{\xi}_0, 0) = I_0^{(0)} + \varepsilon_I^{(0)}, \quad \partial_{\xi_0,\xi_1}^2 S_I(\bar{\xi}_0, 0) = \partial_{\xi_1,\xi_0}^2 S_I(\bar{\xi}_0, 0) = -I_0^{(0)},$$

$$I_0^{(0)} = -\frac{(1 - e^2)\mu}{4\sqrt{V_I(\gamma(\bar{\xi}_0))}}, \quad \varepsilon_I^{(0)} = -e^2\sqrt{V_I(\gamma(\bar{\xi}_0))},$$

$$\partial_{\xi_0}^2 S_E(\bar{\xi}_0, 0) = \partial_{\xi_1}^2 S_E(\bar{\xi}_0, 0) = E_0^0 + \varepsilon_E^{(0)}, \quad \partial_{\xi_0,\xi_1}^2 S_E(\bar{\xi}_0, 0) = \partial_{\xi_1}^2 S_E(\bar{\xi}_0, 0) = -E_0^{(0)},$$

$$E_0^{(0)} = \frac{(1 - e^2)\mathcal{E}}{2\sqrt{V_E(\gamma(\bar{\xi}_0))}}, \quad \varepsilon_E^{(0)} = e^2\sqrt{V_E(\gamma(\bar{\xi}_0))},$$

$$(1.6.1)$$

$$\partial_{\xi_0}^2 S_I(\bar{\xi}_1,0) = \partial_{\xi_1}^2 S_I(\bar{\xi}_1,0) = I_0^{(1)} + \varepsilon_I^{(1)}, \quad \partial_{\xi_0,\xi_1}^2 S_I(\bar{\xi}_1,0) = \partial_{\xi_1}^2 S_I(\bar{\xi}_1,0) = -I_0^{(1)},$$

$$I_0^{(1)} = -\frac{\mu}{4(1-e^2)\sqrt{V_I(\gamma(\bar{\xi}_1))}}, \quad \varepsilon_I^{(1)} = \frac{e^2}{\sqrt{1-e^2}}\sqrt{V_I(\gamma(\bar{\xi}_1))},$$

$$\partial_{\xi_0}^2 S_E(\bar{\xi}_1,0) = \partial_{\xi_1}^2 S_E(\bar{\xi}_1,0) = E_0^{(1)} + \varepsilon_E^{(1)}, \quad \partial_{\xi_0,\xi_1}^2 S_E(\bar{\xi}_1,0) = \partial_{\xi_1}^2 S_E(\bar{\xi}_1,0) = -E_0^{(1)},$$

$$E_0^{(1)} = \frac{\mathcal{E}}{2\sqrt{1-e^2}\sqrt{V_E(\gamma(\bar{\xi}_1))}}, \quad \varepsilon_E^{(1)} = -\frac{e^2}{\sqrt{1-e^2}}\sqrt{V_E(\gamma(\bar{\xi}_1))},$$

$$(1.6.2)$$

and then we have

$$\Delta^{(0)} = A^{(0)}B^{(0)}C^{(0)}D^{(0)}, \qquad\qquad\qquad \Delta^{(1)} = A^{(1)}B^{(1)}C^{(1)}D^{(1)},$$

$$A^{(0)} = -\frac{16}{\mathcal{E}^2\mu^2}\frac{e^2}{1-e^2}\left(\sqrt{V_E(\gamma(\bar{\xi}_0))} - \sqrt{V_I(\gamma(\bar{\xi}_0))}\right)$$

$$A^{(1)} = \frac{16e^2}{\mathcal{E}^2\mu^2}\left(\sqrt{V_E(\gamma(\bar{\xi}_1))} - \sqrt{V_I(\gamma(\bar{\xi}_1))}\right)$$

$$B^{(0)} = \mu + 2\frac{e^2}{1-e^2}\sqrt{V_I(\gamma(\bar{\xi}_0))}\left(\sqrt{V_I(\gamma(\bar{\xi}_0))} - \sqrt{V_E(\gamma(\bar{\xi}_0))}\right)$$

$$B^{(1)} = \mu - 2e^2\sqrt{1-e^2}\sqrt{V_I(\gamma(\bar{\xi}_1))}\left(\sqrt{V_I(\gamma(\bar{\xi}_1))} - \sqrt{V_E(\gamma(\bar{\xi}_1))}\right)$$

$$C^{(0)} = \mathcal{E} + \frac{e^2}{1-e^2}\sqrt{V_E(\gamma(\bar{\xi}_0))}\left(\sqrt{V_E(\gamma(\bar{\xi}_0))} - \sqrt{V_I(\gamma(\bar{\xi}_0))}\right)$$

$$C^{(1)} = \mathcal{E} - e^2\sqrt{V_E(\gamma(\bar{\xi}_1))}\left(\sqrt{V_E(\gamma(\bar{\xi}_1))} - \sqrt{V_I(\gamma(\bar{\xi}_1))}\right)$$

$$D^{(0)} = -\mu\sqrt{V_E(\gamma(\bar{\xi}_0))} + 2\sqrt{V_I(\gamma(\bar{\xi}_0))}C^{(0)}$$

$$D^{(1)} = -\mu\sqrt{V_E(\gamma(\bar{\xi}_1))} + 2\sqrt{1-e^2}\sqrt{V_I(\gamma(\bar{\xi}_1))}C^{(1)}$$

$$(1.6.3)$$

1.6.1 Asymptotic behaviours

It is convenient to start the study of the elliptic case by investigating some of the properties of the first return map on a circular domain. When $e = 0$, for every $\bar{\xi} \in [0, 2\pi]$ the pair $(\bar{\xi}, 0)$ is a homothetic fixed point for F, with

$$DF(\bar{\xi},0) = \begin{pmatrix} 1 & \sqrt{\mathcal{E} - \frac{\omega^2}{2}}\left(\frac{2}{\mathcal{E}}\sqrt{\mathcal{E} - \frac{\omega^2}{2}} - \frac{4}{\mu}\sqrt{\mathcal{E} + h + \mu}\right) \\ 0 & 1 \end{pmatrix}.$$

This is consistent with the expression of F for a circular domain: when D is a disk of radius 1, from the central symmetry of both the domain and the inner and outer potentials one has that F is a rigid translation of the form

$$F_{circ}(\xi_0, \alpha_0) = (\xi_0 + \theta(\alpha_0), \alpha_0),$$

where the explicit expression of $\theta(\alpha)$ will be provided in Chapter 2. As a consequence, $\Delta_{circ} = 0$ for every homothetic point $(\bar{\xi}, 0)$. The circular case represents then a degenerate case for the study of the linear stability of the homothetic points; nevertheless, the possibility to compute the explicit expression of F_{circ} allows to study directly the map: considering the phase space (ξ, α), one has that the set $[0, 2\pi] \times \{0\}$ is the invariant set containing all the homothetic points, and that all the orbits of F_{circ} lie on the invariant lines $[0, 2\pi] \times \{\bar{\alpha}\}$, where the value of $\theta(\bar{\alpha})$ determines their nature. The systematical study of the circular case, in a more convenient variational setting, is one of the subject of Chapter 2.

Let us suppose that $e > 0$ and small. Recalling that $b = \sqrt{1 - e^2}$, the expression of $\Delta^{(0)}$ and $\Delta^{(1)}$ in Eqs.(1.6.3) can be expanded in Taylor series around $e = 0$, obtaining, from direct computations,

$$\Delta^{(0)} = f_2 e^2 + f_4 e^4 + \mathcal{O}(e^6), \quad \Delta^{(1)} = g_2 e^2 + g_4 e^4 + \mathcal{O}(e^6), \quad (1.6.4)$$

$$f_2 = -g_2 = -\frac{4\left(\sqrt{\mathcal{E} - \omega^2/2} - \sqrt{\mathcal{E} + h + \mu}\right)\left(2\mathcal{E}\sqrt{\mathcal{E} + h + \mu} - \mu\sqrt{\mathcal{E} - \omega^2/2}\right)}{\mu\mathcal{E}}. \quad (1.6.5)$$

Hence, when e is sufficiently small, the sign of $\Delta^{(0)}$ and $\Delta^{(1)}$ is determined by the quantity $\left(2\mathcal{E}\sqrt{\mathcal{E} + h + \mu} - \mu\sqrt{\mathcal{E} - \omega^2/2}\right)$.

Let us now suppose to fix the parameters related to the external dynamics, namely, $\mathcal{E}$ and ω, and to let vary μ and h. If e is small enough, $\Delta^{(0)}$ and $\Delta^{(1)}$ have opposite sign; in particular:

- if $\dfrac{\sqrt{\mathcal{E} + h + \mu}}{\mu} < \dfrac{\sqrt{2\mathcal{E} - \omega^2}}{2\sqrt{2\mathcal{E}}}$, $\Delta^{(0)} < 0$ and $\Delta^{(1)} > 0$. Then, from Remark 1.5.1, one has that $(0, 0)$ is a stable center and $(\pi/2, 0)$ is an unstable saddle for F;

- if $\dfrac{\sqrt{\mathcal{E} + h + \mu}}{\mu} > \dfrac{\sqrt{2\mathcal{E} - \omega^2}}{2\sqrt{2\mathcal{E}}}$, for the same reasoning $(0, 0)$ is a saddle and $(\pi/2, 0)$ is a center.

Fixing $\mathcal{E}$ and ω, one has also:

$$\lim_{h \to \infty} \Delta^{(0)} = \lim_{\mu \to \infty} \Delta^{(0)} = \lim_{h \to \infty} \Delta^{(1)} = \lim_{\mu \to \infty} \Delta^{(1)} = \infty.$$

As a final investigation on the asymptotical behaviour of $\Delta^{(0)}$ and $\Delta^{(1)}$, let us suppose to fix the physical parameters related to the inner dynamics and analyse the sign of the discriminants for $\mathcal{E} \to \infty$. From direct computations, one has

$$\ell_0 = \lim_{\mathcal{E} \to \infty} \Delta^{(0)} = \frac{(b^2 - 1)(2h + 2\mu + \omega^2)(2(b^2 - 1)h - 2\mu + (b^2 - 1)\omega^2)}{b^4 \mu^2}$$

$$\ell_1 = \lim_{\mathcal{E} \to \infty} \Delta^{(1)} = \frac{b(b^2 - 1)(2bh + 2\mu + b^3\omega^2)(2(b^2 - 1)h + b(2\mu + b(b^2 - 1)\omega^2))}{\mu^2}.$$

In particular, it results $\ell_0 > 0$ for every fixed $0 < b < 1$ and $h, \mu, \omega > 0$ and

$$\ell_1 > 0 \iff 0 < b < 1 \text{ and } 0 < \mu < \bar{\bar{\mu}} = \frac{(b^2 - 1)(2h + b^2\omega^2)}{2b}.$$

Taking together the above considerations, one can give some general results, which hold for small eccentricity or for high values of h, μ or $\mathcal{E}$.

Proposition 1.6.1. *For every* $\mathcal{E}, \omega > 0$ *with* $\omega^2 > 2\mathcal{E}$ *we have:*

I) for every fixed $h, \mu > 0$:

Ia) if $\dfrac{\sqrt{\mathcal{E} + h + \mu}}{\mu} < \dfrac{\sqrt{2\mathcal{E} - \omega^2}}{2\sqrt{2\mathcal{E}}}$, *then there is* $\bar{e} \in (0,1)$ *such that, for every* $e \in (o, \bar{e})$: $\bar{z}_0$ *is stable and* $\bar{z}_{\pi/2}$ *is unstable;*

Ib) if $\dfrac{\sqrt{\mathcal{E} + h + \mu}}{\mu} > \dfrac{\sqrt{2\mathcal{E} - \omega^2}}{2\sqrt{2\mathcal{E}}}$, *then there is* $\bar{e} \in (0,1)$ *such that, for every* $e \in (o, \bar{e})$: $\bar{z}_0$ *is unstable and* $\bar{z}_{\pi/2}$ *is stable.;*

II) for all fixed $e \in (0,1)$, $h > 0$, *there is* $\bar{\mu} > 0$ *such that for every* $\mu > \bar{\mu}$ *the homothetic fixed points* $(0,0)$ *and* $(\pi/2, 0)$ *are saddles;*

III) for all fixed $e \in (0,1)$, $\mu > 0$, *there is* $\bar{h} > 0$ *such that for every* $h > \bar{h}$ *the homothetic fixed points* $(0,0)$ *and* $(\pi/2, 0)$ *are saddles.*

For all fixed $e \in (0,1)$, $h, \omega > 0$ *there are* $\bar{\mathcal{E}} > 0$ *and* $\bar{\bar{\mu}} > 0$ *such that, if* $\mathcal{E} > \bar{\mathcal{E}}$:

IVa) if $\mu > \bar{\bar{\mu}}$, $(0,0)$ *is a saddle and* $(\pi/2, 0)$ *is a center;*

IVb) if $0 < \mu < \bar{\bar{\mu}}$, $(0,0)$ *and* $(\pi/2, 0)$ *are both saddles.*

With the additional hypothesis that the F is well defined on the whole ellipse, in cases (II) and (III), as well as (IVb), there must be at least a stable fixed point with $\xi_0 \in (0, \pi/2)$; *hence, by symmetry, F admits at least 4 stable period one non-homothetic fixed points.*

Proposition 1.6.1 (I) provides an approximated relation between h and μ through which one can find two regimes in the parameters such that, for e sufficiently small, the stability of the homothetic fixed points can be easily deduced. In particular, there is a curve which, for e sufficiently small, divides the two cases (Ia) and (Ib). As all the involved quantities are positive, one has

$$\frac{\sqrt{\mathcal{E} + h + \mu}}{\mu} > \frac{\sqrt{2\mathcal{E} - \omega^2}}{2\sqrt{2\mathcal{E}}} \iff h > \frac{2\mathcal{E} - \omega^2}{8\mathcal{E}^2}\mu^2 - \mu - \mathcal{E} = p(\mu).$$

When $\mathcal{E}$ and ω are fixed, as well as e small enough, the 2-degree polynomial $p(\mu)$ describes then a parabola on the plane (h, μ) such that, for fixed μ, if $h >> p(\mu)$, then we are in case (Ib); on the contrary, if $h << p(\mu)$, the case (Ia) is verified.
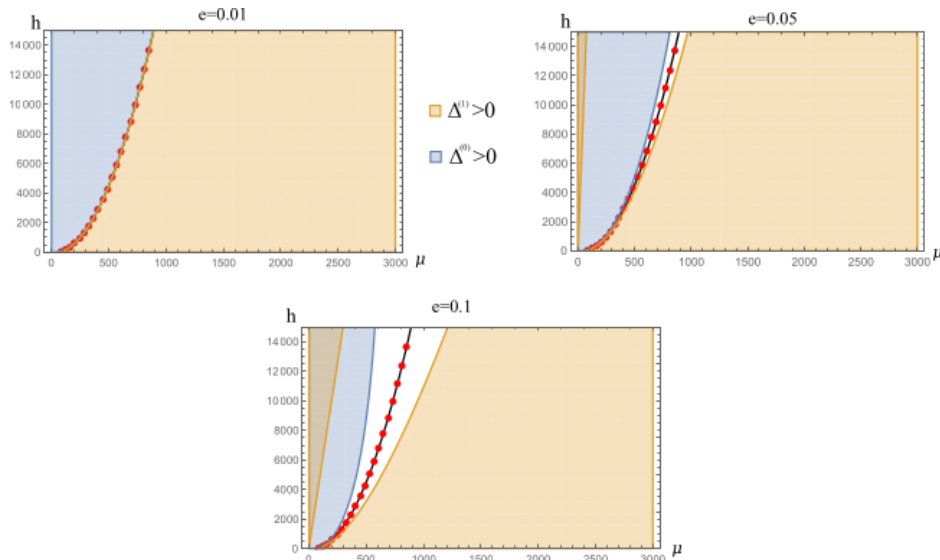
Fig. 1.2 Sign of $\Delta^{(0)}$ and $\Delta^{(1)}$ in the $(\mu, h)-$plane for $\mathcal{E} = 10$ and $\omega = 2$. The red dotted curve represents the parabola $h = p(\mu)$.

We stress that this behaviour holds only for e small enough for f_2 and g_2 to be the dominant terms in the expansions of $\Delta^{(0)}$ and $\Delta^{(1)}$. Moreover, one can verify that $f_4, g_4 = \mathcal{O}(h^2\sqrt{\mu})$, and that all the further terms of the Taylor expansion are of the order of positive powers of μ and h: as a consequence, for every $e > 0$, eventually the two parameters would be too large to use the above approximation.

Figure 1.2 gives a comparison between the parabola $p(\mu)$ (red dots) and the effective curves of change of sign for $\Delta^{(0)}$ and $\Delta^{(1)}$ in the (μ, h) plane for $\mathcal{E} = 10$, $\omega = 2$ and increasing eccentricities. As one can see, for very small eccentricities the approximation given by $p(\mu)$ is very good even for extremely high values of μ and h; on the other hand, the increase of the eccentricity and of the two inner parameters made this approximation worse.

Moving to moderate and high eccentricities, the behaviour of the signs of $\Delta^{(0)}$ and $\Delta^{(1)}$ becomes more complex: to give an example of this, Figure 1.3 shows the sign of both the discriminants as functions of h and μ and for fixed $\mathcal{E}$, ω and eccentricity of the ellipse. It is present a reminiscence of the original parabola $p(\mu)$, which tends to widen for increasing eccentricity, while other sign-changing curves, deriving by the influence of the higher order terms in (1.6.4), are present.

1.6.2   Arising of 2-periodic brake orbits

As already seen in Section 1.6.1, the existence of non homothetic $1-$ periodic points of F can be deduced by the signs of $\Delta^{(0)}$ and $\Delta^{(1)}$, namely, by the stability properties of the homothetic points. On the other hand, other analytical techniques can be
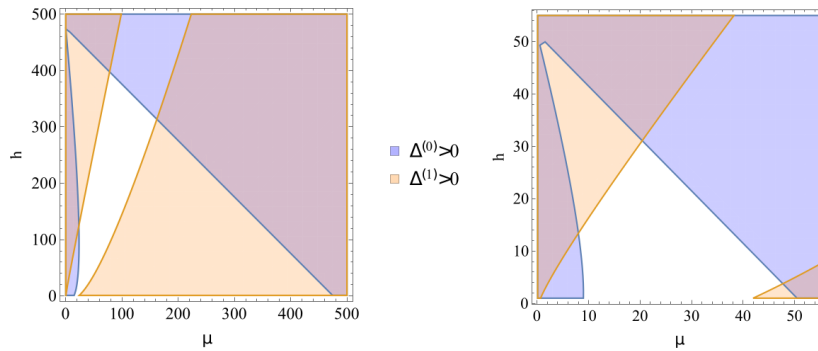
Fig. 1.3 Sign of $\Delta^{(0)}$ and $\Delta^{(1)}$ in the $(\mu, h)-$plane for $\mathcal{E} = 2.5$, $\omega = \sqrt{2}$ and $e = 0.3$ (left) or $e = 0.5$ (right).

used to assure the existence of particular periodic orbits with period greater than 1. It is the case of the so-called non homothetic *brake orbits*, namely, 2-periodic orbits with homothetic outer arcs (see Figure 1.4), whose existence can be proved for suitable regimes of the parameters through an application of the shooting method. The
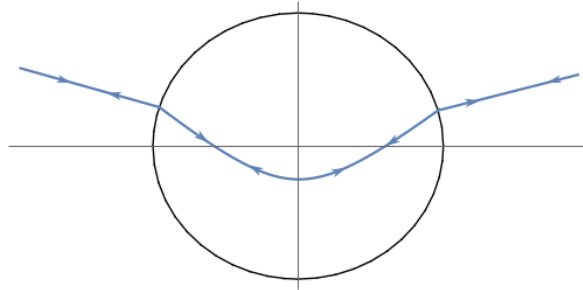
Fig. 1.4 Example of 2-periodic brake orbit: the homothetic outer arcs are connected by an inner hyperbola.

existence of brake orbits is equivalent to the existence of non-homothetic zeros for the *Free Fall* map, which quantifies the scattering with respect to the radial direction of the trajectory after entering the domain. Given $\theta \in [0, 2\pi]$, consider the homothetic outer arc with initial points (p_0, v_0), where p_0 is the intersection between the ellipse and the radial straight line of inclination θ, while v_0 is the outward-pointing radial vector in the direction of θ and such that $\|v_0\|^2/2 - V_E(p_0) = 0$: if we denote, as in the previous Sections, with (p_1, v_1) the position and velocity vectors after two consecutive outer and inner crossings (with the respective refractions), the free fall map $\theta \mapsto \delta(\theta)$ returns the angle δ between v_1 and p_1 (see Figure 1.5).

 If we consider general domains whose boundary intersects orthogonally the axes, as in the case of the ellipse, Theorems 1.3.1 and 1.3.5 guarantee that the Free Fall map is well defined in suitable neighbourhoods of the homothetic orbits in the horizontal and vertical directions. Nevertheless, as the construction of $\delta(\theta)$ only involves the refraction rule and the inner dynamics, under suitable hypotheses on ∂D one can assure that it
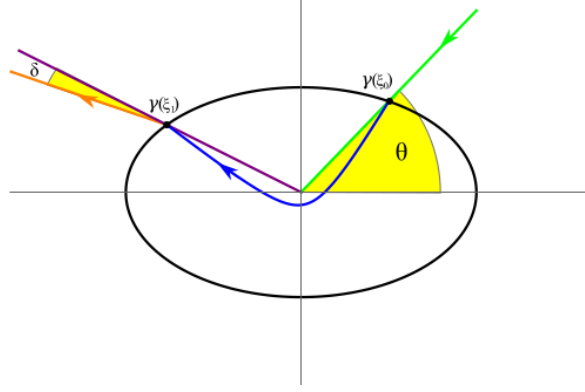
Fig. 1.5 Free fall map on the ellipse.

is well defined globally on $[0, 2\pi]$; in particular, it is sufficient to require

(I) the good definition of the inner dynamics globally on ∂D;

(II) a global transversality property of ∂D with respect to the radial directions holds:

$$\forall \xi \in I \quad \gamma(\xi) \nparallel \dot{\gamma}(\xi).$$

(1.6.6)

When ∂D satisfies the above properties, one can continuously extend $\delta(\theta)$ even in the case that the first return map F is not well defined (see Section 1.2): it suffices to impose $\delta(\theta) = \pi/2$ whenever the inner angle β_1 is greater or equal to $\beta_{crit} = \arcsin(\sqrt{V_E(p_1)/V_I(p_1)})$ and $\delta(\theta) = -\pi/2$ when $\beta_1 \leq -\beta_{crit}$. As a consequence, the function δ results to be a continuous function of $\theta \in [0, 2\pi]$, differentiable whenever $|\beta_1| < \beta_{crit}$, as in neighbourhoods of homothetic solutions. Moreover, condition (II) assures that whenever $\delta(\theta) = 0$ the Free Fall map is well-defined, since the refracted outer arc is not tangent to ∂D.

While the geometrical implications of condition (II) are rather immediate, condition (I) is implied by taking a particular class of domains characterized by a convexity property with respect to the hyperbolæ. In particular, we shall give the following definition.

Definition 1.6.2. *We say that the domain D is **convex for hyperbolæ for fixed h, $\mathcal{E}$ and μ** if every Keplerian hyperbola with energy $\mathcal{E} + h$ and central mass μ intersects ∂D at most in two points.*
*The domain D is **convex for hyperbolæ** if the previous condition holds for every positive $\mathcal{E}, h$ and μ.*

The connection between the Free Fall map and the brake orbits is straightforward: $(\cos\bar{\theta}, \sin\bar{\theta})$ is the direction of a 2-periodic brake orbit if and only if $\delta(\bar{\theta}) = 0$ and, denoting with $\xi_{\bar{\theta}}$ the parameter in I such that $\gamma(\xi_{\bar{\theta}})$ has polar angle $\bar{\theta}$, $\gamma(\xi_{\bar{\theta}}) \nparallel \dot{\gamma}(\xi_{\bar{\theta}})$. Let us remark that, by the properties of the ellipse, one has that $\delta(k\pi/2) = 0$ for $k = 0, 1, 2, 3$, and that condition (II) is trivially true. The following Proposition shows that, when the eccentricity is small enough the elliptical domains are also convex by hyperbolæ, leading to the conclusion that, in these cases, the Free Fall map is globally well defined.
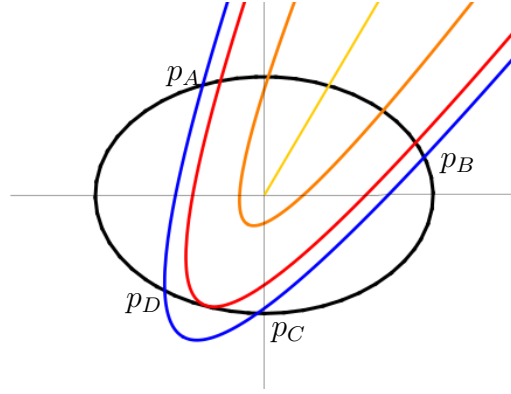
Fig. 1.6 Nested family of Keplerian hyperbolæ $\mathcal{F}'$ for fixed $\mathcal{E}, h, \mu$ and fixed transverse axis. As ℓ increases, the hyperbolæ move from the inner ones (orange) to the outer ones (blue).

Proposition 1.6.3. *If D is an ellipse parametrised by $(\cos(\xi), b\sin(\xi))$ with eccentricity $e \in [0, 1/\sqrt{2})$, then it is convex by hyperbolæ.*

Proof. Let us start by fixing the ellipse's eccentricity e and the parameters $\mathcal{E}, h$ and μ and by taking the associated family $\mathcal{F}$ of hyperbolæ, which is continuous with respect to variations of the angular momentum and rotations of the axes. Denoting with ℓ the absolute value of the angular momentum of a Keplerian hyperbola $\mathcal{K}$ in such a family and with r_p its minimal distance from the origin, one has that (see e.g. [70])

$$r_p = \frac{\ell^2}{\mu\left(1 + \sqrt{1 + \frac{2(\mathcal{E}+h)\ell^2}{\mu^2}}\right)},$$

which is continuous and strictly increasing for $\ell \geq 0$. The distance at the pericenter is then 0 when $\ell = 0$ (homothetic orbit) and varies continuously with ℓ. Moreover, since for the ellipse Theorem 1.3.5 is true for every $\xi \in [0, 2\pi]$, for ℓ small enough the hyperbolæ of $\mathcal{F}$ intersect ∂D exactly twice. Let us now fix a direction in $\mathbb{R}^2$, and consider only the hyperbolæ in $\mathcal{F}$ whose transverse axis is in the chosen direction, denoting them with $\mathcal{F}'$. As the eccentricity of a Keplerian hyperbolæ, whose expression is

$$e_{hyp} = \sqrt{1 + \frac{2(\mathcal{E} + h)\ell^2}{\mu^2}},$$

is strictly increasing in ℓ, such hyperbolæ are nested as in Figure 1.6. Let us suppose that there exists a Keplerian hyperbola in $\mathcal{F}'$ which intersects ∂D in four points p_A, p_B, p_C and p_D. For the previous considerations on the continuous dependence and monotonicity of r_p and e_{hyp} on ℓ, there exists $\bar{\ell}$ such that the corresponding hyperbola $\bar{\mathcal{K}}$ in $\mathcal{F}'$ is tangent from inside to ∂D in a point p_T; define $r_T = \|p_T\|$. This implies that, denoted with $k_{hyp}(p)$ and $k_{ell}(p)$ respectively the curvatures with respect to the inward-pointing normal vector of $\bar{\mathcal{K}}$ and of ∂D in a point p, we obtain

$$k_{hyp}(p_T) \geq k_{ell}(p_T),$$

which is a necessary condition for the family $\mathcal{F}'$ to admit a hyperbola which intersects ∂D four times. The ellipse's curvature is always bounded from below by b, while one can compute $k_{hyp}(p_T)$ by parametrising $\bar{\mathcal{K}}(s)$ by the kinetic time and recalling that, for

a generic curve $r(t)$, the curvature is given by

$$k_r(t) = \frac{\|r'(t) \wedge r''(t)\|}{\|r'(t)\|^3}.$$

Observing that $p_T \in \partial D$, which implies $r_T \geq b$, one obtains that, if $\bar{s}$ is such that $\bar{\mathcal{K}}(\bar{s}) = p_T$,

$$k_{hyp}(p_T) = \frac{\|\bar{\mathcal{K}}'(\bar{s}) \wedge \bar{\mathcal{K}}''(\bar{s})\|}{\|\bar{\mathcal{K}}'(\bar{s})\|^3} \leq \frac{\|\bar{\mathcal{K}}''(\bar{s})\|}{\|\bar{\mathcal{K}}'(\bar{s})\|^2} = \frac{\mu}{2r_T(r_T(\mathcal{E} + h) + \mu)} \leq \frac{\mu}{2b(b(\mathcal{E} + h) + \mu)}$$

Taking together the bounds obtained for $k_{ell}(p_T)$ and $k_{hyp}(p_T)$, one can find a necessary condition for the family $\mathcal{F}'$ to admit hyperbolæ with four intersection points with ∂D, given by

$$\frac{\mu}{2b(b(\mathcal{E} + h) + \mu)} \geq b.$$

It is then sufficient to require

$$\frac{\mu}{2b(b(\mathcal{E} + h) + \mu)} < b \iff 2b^2 \left(\frac{\mathcal{E} + h}{\mu} b + 1 \right) > 1 \qquad (1.6.7)$$

to ensure that $\mathcal{F}'$ does not admit hyperbolæ of such kind. It is straightforward to observe that (1.6.7) is trivially satisfied for every $\mathcal{E} + h > 0$ ad $\mu > 0$ whenever $2b^2 > 1$, namely $e \in [0, 1/\sqrt{2})$. This reasoning can be repeated for every fixed direction for the axis and for every $\mathcal{E}, h, \mu > 0$; in particular, it holds also when two of the four points of the original hyperbola (the blue one in Figure 1.6) coincide: it is in fact trivially true when $p_C = p_D$, and, if $p_A = p_D$ or $p_B = p_C$, one can take a lower ℓ to retrieve the original case. Then the convexity for hyperbolæ is proved whenever $e \in [0, 1/\sqrt{2})$.

$\square$

Although deriving the explicit expression of $\delta(\theta)$ goes beyond the extent of this study, the values of its derivatives computed at the homothetic points, which can be found by making use of Eqs.(1.6.1, 1.6.2), along with the global good definition of the Free Fall map, provide a sufficient condition for the existence of the brake orbits in the elliptic case.

Theorem 1.6.4. *For every $\mathcal{E} > 0$, $\omega > 0$ such that $\omega^2 > 2\mathcal{E}$, $e \in (0, 1/\sqrt{2})$ there are $\bar{h} > 0$, $\bar{\mu} > 0$ such that, if $h > \bar{h}$ and $\mu > \bar{\mu}$, the first return map F admits at least four $2-$periodic brake orbits.*

Proof. By symmetry, it is sufficient to prove that, for $\mathcal{E}, \omega, b$ satisfying the hypotheses of the Theorem, there are $\bar{h} > 0$ and $\bar{\mu} > 0$ such that, if $h > \bar{h}$ and $\mu > \bar{\mu}$, then $\exists \bar{\theta} \in (0, \pi/2)$ such that $\delta(\bar{\theta}) = 0$. To this end, we want to find a regime for the parameters such that $\delta'(0) > 0$ and $\delta'(\pi/2) > 0$.
Recall the definitions of (ξ_0, α_0), (ξ_1, α_1) used in Section 1.4, suppose to work in a neighbourhood of $\theta = \pi/2$, and consider the $6-$dimensional variable

$$\boldsymbol{q} = (\theta, \xi_0, \alpha_0, \xi_1, \alpha_1, \delta) \in (0, \pi) \times (0, \pi) \times \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \times (0, \pi) \times \left[-\frac{\pi}{2}, \frac{\pi}{2} \right] \times [0, 2\pi]$$

From elementary geometric considerations and recalling the refraction relations, one has that, defined $\bar{\boldsymbol{q}} = (\pi/2, \pi/2, 0, \pi/2, 0, 0)$, $\Phi(\bar{\boldsymbol{q}}) = 0$, where

$$
\Phi(q) = \begin{pmatrix}
\dfrac{1}{b}\cot\xi_0 - \cot\theta \\
\alpha_0 + \theta - \operatorname{arccot}(b\cot\xi_0) \\
\partial_a S_I(\xi_0, \xi_1) + \sqrt{V_E(\gamma(\xi_0))}\sin\alpha_0 \\
\partial_b S_I(\xi_0, \xi_1) - \sqrt{V_E(\gamma(\xi_1))}\sin\alpha_1 \\
\delta + \alpha_1 + \operatorname{arccot}(b\cot\xi_1) - \operatorname{arccot}\left(\dfrac{\cot\xi_1}{b}\right),
\end{pmatrix}
$$

is defined in a suitable neighbourhood of $\bar{\boldsymbol{q}}$. As a consequence,

$$
M = D_{(\xi_0, \alpha_0, \xi_1, \alpha_1, \delta)}\Phi(\bar{\boldsymbol{q}})
$$

$$
= \begin{pmatrix}
-\dfrac{1}{b} & 0 & 0 & 0 & 0 \\
-b & 1 & 0 & 0 & 0 \\
\partial_a^2 S_I(\bar{\xi}_1, \bar{\xi}_1) & \sqrt{V_E(\gamma(\bar{\xi}_1))} & \partial_{ab} S_I(\bar{\xi}_1, \bar{\xi}_1) & 0 & 0 \\
\partial_{ab} S_I(\bar{\xi}_1, \bar{\xi}_1) & 0 & \partial_b^2 S_I(\bar{\xi}_1, \bar{\xi}_1) & -\sqrt{V_E(\gamma(\bar{\xi}_1))} & 0 \\
0 & 0 & b - \dfrac{1}{b} & 1 & 1
\end{pmatrix}
$$

$$
\Rightarrow \det(M) = \frac{\mu}{4b^3\sqrt{\mathcal{E} + h + \mu/b}}\sqrt{\mathcal{E} - \frac{\omega^2}{2}b^2} > 0.
$$

Applying then the implicit function theorem, $\delta'(\pi/2)$ can be computed as the last component of the vector

$$
-M^{-1}D_\theta\Phi(\bar{\boldsymbol{q}}) = -M^{-1}\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},
$$

obtaining

$$
\delta'(\pi/2) = -\frac{\sqrt{V_E(\gamma(\bar{\xi}_1))}b^2 + \epsilon_I^{(1)}b - \sqrt{V_E(\gamma(\bar{\xi}_1))}}{bI_0^{(1)}\sqrt{V_E(\gamma(\bar{\xi}_1))}}
$$
$$
\times \left(\sqrt{V_E(\gamma(\bar{\xi}_1))}b^2 + (2I_0^{(1)} + \epsilon_I^{(1)})b - \sqrt{V_E(\gamma(\bar{\xi}_1))}\right) \tag{1.6.8}
$$

With the same reasoning and taking $\theta \in (-\pi/2, \pi/2)$, one gets

$$
\delta'(0) = \frac{\sqrt{V_E(\gamma(\bar{\xi}_0))} + \epsilon_I^{(0)}b - \sqrt{V_E(\gamma(\bar{\xi}_0))}b^2}{bI_0^{(0)}\sqrt{V_E(\gamma(\bar{\xi}_0))}}
$$
$$
\times \left(\sqrt{V_E(\gamma(\bar{\xi}_0))}b^2 - (2I_0^{(0)} + \epsilon_I^{(0)})b - \sqrt{V_E(\gamma(\bar{\xi}_0))}\right)
$$

Taking $b = \sqrt{1 - e^2}$, direct computations show that, if

$$\mu > \bar{\mu} = \frac{e^2(1 - e^2)^{3/2}(2\mathcal{E} - e^2\omega^2)}{(2e^2 - 1)^2}$$

and

$$h > \bar{h} = \frac{1}{4}\left(-2\mathcal{E} - \frac{(4e^2 - 2)\mu}{e^2\sqrt{1 - e^2}} - (1 - e^2)\omega^2\right)$$
$$+ \sqrt{\frac{(2\mathcal{E} - (1 - e^2)\omega^2)(4\mu - e^2\sqrt{1 - e^2}((1 - e^2)\omega^2 - 2\mathcal{E}))}{e^2\sqrt{1 - e^2}}}$$

then $\delta'(0) > 0$ and $\delta'(\pi/2) > 0$, and the statement is proved. $\qquad\square$

Remark 1.6.5. *Notice that Theorem 1.6.4 can be extended to general domains D with boundary ∂D parametrized by a curve $\tilde{\gamma}$, provided that:*

- *conditions (1.6.6) hold;*

- *γ shares the symmetry properties of the ellipse;*

- *in the vicinity of the intersections between the coordinate axes and ∂D, $\tilde{\gamma}$ and $\gamma(\xi) = (\cos\xi, b\sin\xi)$ are equal up to the second order, namely:*

$$\tilde{\gamma}(\xi) = (\cos\xi + f(\xi), b\sin\xi + g(\xi)),$$
$$f(\bar{\xi}_{0\backslash 1}) = g(\bar{\xi}_{0\backslash 1}) = f'(\bar{\xi}_{0\backslash 1}) = g'(\bar{\xi}_{0\backslash 1}) = f''(\bar{\xi}_{0\backslash 1}) = g''(\bar{\xi}_{0\backslash 1}) = 0.$$

As a matter of fact, one has that, locally around $\pi/2$ (the reasoning for 0 is the same), the vector $\boldsymbol{q}$ defined as in the Theorem satisfies the relation $\tilde{\Phi}(\boldsymbol{q})$, with

$$\tilde{\Phi}(\boldsymbol{q}) = \begin{pmatrix} \dfrac{\cos\xi_0 + f(\xi_0)}{b\sin\xi_0 + g(\xi_0)} - \cot\theta \\ \alpha_0 + \theta - \operatorname{arccot}\left(\dfrac{b\cos\xi_0 + g'(\xi_0)}{\sin\xi_0 - f'(\xi_0)}\right) \\ \partial_a S_I(\xi_0, \xi_1) + \sqrt{V_E(\gamma(\xi_0))}\sin\alpha_0 \\ \partial_b S_I(\xi_0, \xi_1) - \sqrt{V_E(\gamma(\xi_1))}\sin\alpha_1 \\ \delta + \alpha_1 + \operatorname{arccot}\left(\dfrac{b\cos\xi_1 + g'(\xi_1)}{\sin\xi_1 - f'(\xi_1)}\right) - \operatorname{arccot}\left(\dfrac{\cos\xi_1 + f(\xi_1)}{b\sin\xi_1 + g(\xi_1)}\right) \end{pmatrix}, \qquad (1.6.9)$$

whose derivatives with respect to all the variables, computed in $\bar{\boldsymbol{q}}$, are the same as in the Theorem.

Example 1.6.6. To make the reasoning quantitative, let us now consider the case $\mathcal{E} = 2.5, \omega = \sqrt{2}, \mu = 2$ and $e = 0.1$. Figure 1.7 shows the signs of $\Delta^{(0)}$ and $\Delta^{(1)}$ as a function of h. One can see that, while $(0,0)$ is always an unstable saddle, there is a bifurcation value of h for which $(\pi/2, 0)$ changes its stability, whose value is precisely $h_{bif} = 109.091$. Figure 1.8 shows the transition of $(\pi/2, 0)$ from center to saddle, with the concurrent arising of a two periodic orbit. With reference to Theorem 1.6.4, we have in this case $\bar{\mu} \simeq 0.0511$, while $\bar{h} = h_{bif}$: the threshold value for the existence of the 2-periodic brake orbits is then equal to the one for the change of stability of the
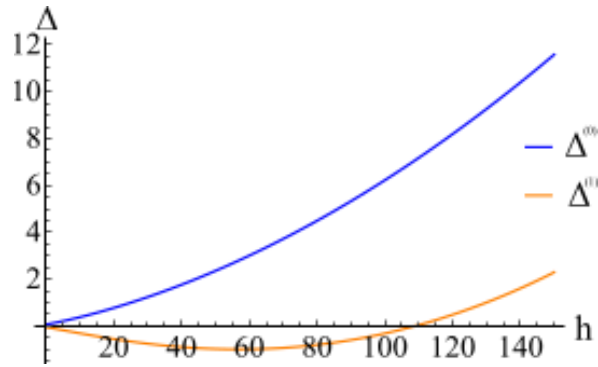
Fig. 1.7 Values of $\Delta^{(0)}$ and $\Delta^{(1)}$ as a function of h, with $\mathcal{E} = 2.5, \omega = \sqrt{2}, \mu = 2, e = 0.1$.
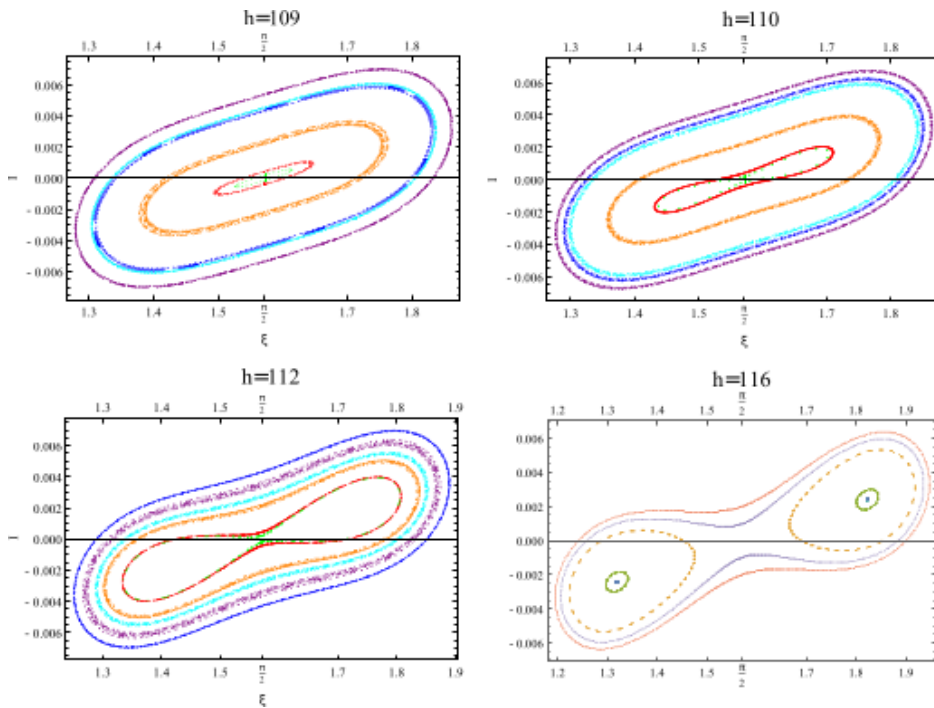
Fig. 1.8 Orbits of F in a neighbourhood of the homothetic fixed point $(\pi/2, 0)$ for $\mathcal{E} = 2.5, \omega = \sqrt{2}, \mu = 2, e = 0.1$ and different values of h. The transition of the fixed point from center to saddle is evident. Bottom-Right: the 2-periodic fixed point is detected as the 2-points blue orbit.
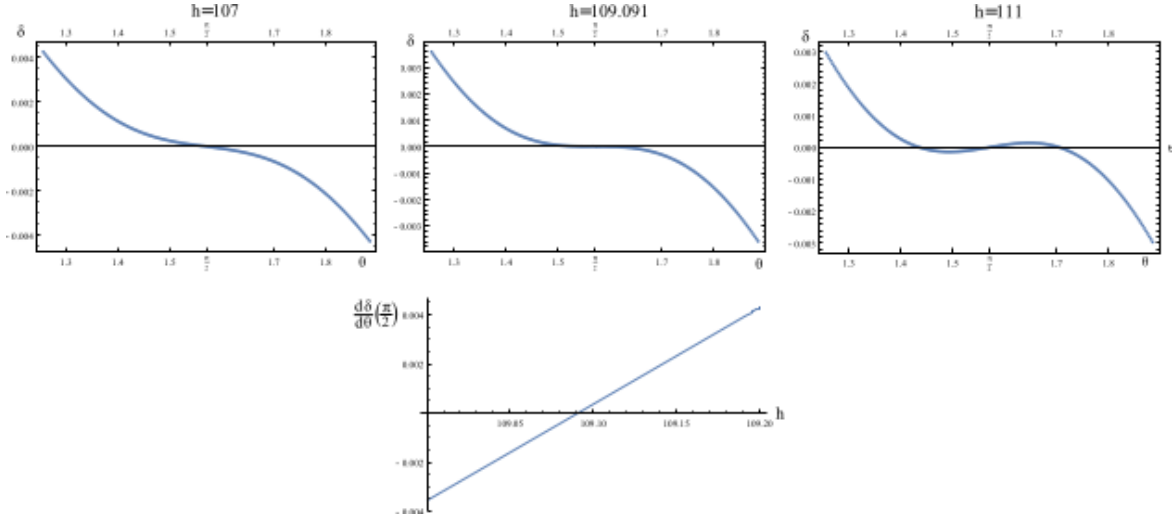
Fig. 1.9 Plot of the free fall map (top) and its derivative in $\pi/2$ (bottom) as a function of h. The other parameters are $\mathcal{E} = 2.5, \omega = \sqrt{2}, \mu = 2, e = 0.1$.

homothetic equilibrium point, underlying the concurrence of the two phenomena.

The direct study of the Free Fall map corroborates these findings. As a matter of fact, Figure 1.9 shows the plots of $\delta(\theta)$ in a neighbourhood of $\theta = \pi/2$ for different values of h (before h_{bif}, at h_{bif} and after it), along with the value of $\delta'(\pi/2)$ as a function of h: as one can see, before the bifurcation value the free fall map is strictly decreasing, while for $h = h_{bif}$ it has an inflection point with zero derivative at $\pi/2$. After the bifurcation value, two zeros, corresponding precisely to the brake orbits values of θ, appear.

1.7   Numerical simulations

As already pointed out in Section 1.6.2, the validity of the analytical investigations can be corroborated by a direct comparison with the plots of the map F in specific cases, which highlights the variety of the behaviours of the dynamics for different values of the involved parameters.

This Section aims to gather cases of interest for the dynamics, underlying the effective role of the bifurcations in the change of stability and the subsequent arising or disappearance of new periodic points for F, as well as the potential presence of diffusive orbits, that represents a strong signal of chaoticity.

All the below simulations are performed by considering D as an ellipse centerd in the origin, with semiaxes $a = 1$ and $b = \sqrt{1 - e^2}$, for different values of e. The routine is implemented in *Mathematica*, and involves the numerical integration for the outer problem in its original form and, in order to avoid the numerical instability due to the presence of the possible singularity, of the inner problem in its regularised formulation.

Figure 1.10 shows the transition of the map through different stability regimes as the parameters modify the sign of $\Delta^{(0)}$ and $\Delta^{(1)}$. The changes of stability of $(0,0)$ between **(b)** and **(c)** and of $(\pi/2, 0)$ between **(c)** and **(d)** are consistent with the plot
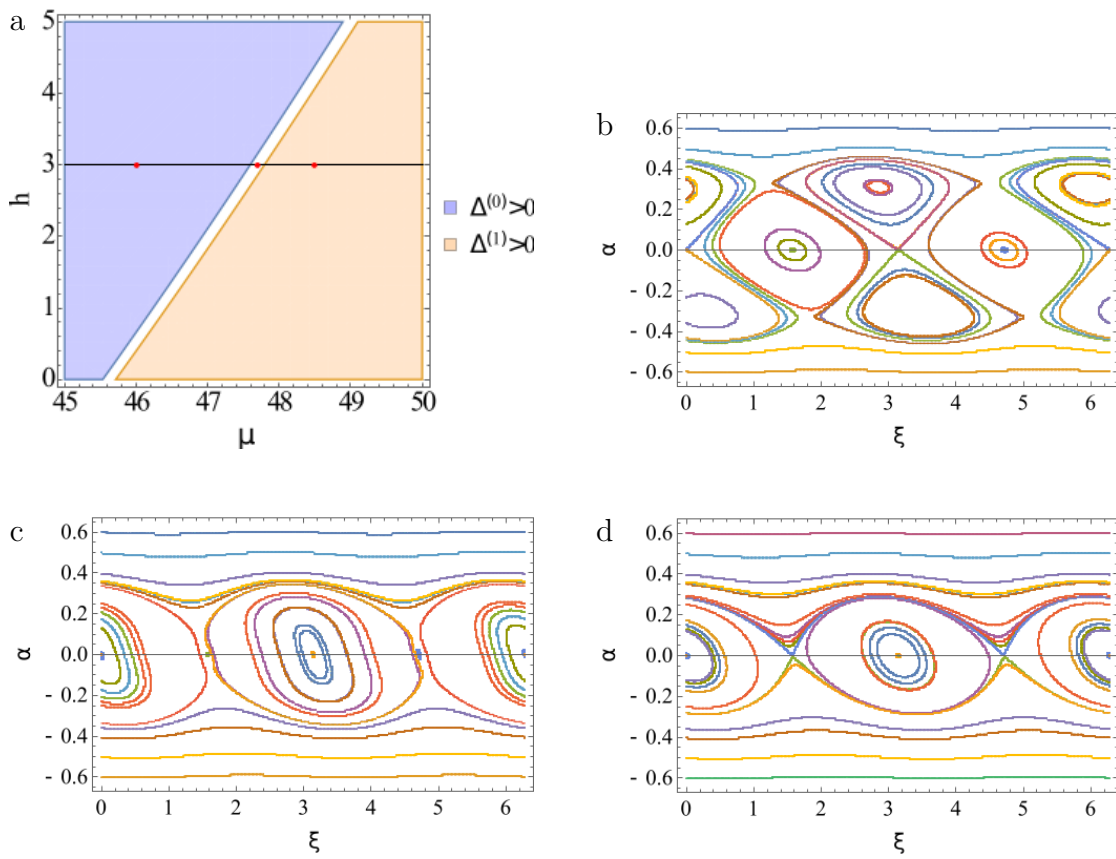
Fig. 1.10 Bifurcations and plot of F for a small eccentricity in a neighbourhood of the axis $\alpha = 0$. **(a)** Sign of $\Delta^{(0)}$ and $\Delta^{(1)}$ for $\mathcal{E} = 9$, $\omega = 1$, $e = 0.03$ as a function of h and μ. The red dots correspond to $h = 3$, $\mu = 46$ **(b)**, $h = 3$, $\mu = 47.7$ **(c)** and $h = 3$, $\mu = 48.5$ **(d)**.

of the discriminants sketched in **(a)**. In this case, where the eccentricity is small and the parameters are not much different from each other, the maps results to be regular also in the vicinity of the fixed points. We observe that in all the considered cases, for high values of α the map results essentially in a rotation on the ellipse, with small oscillations in α. A noticeable fact is represented by the overall number of respectively stable and unstable equilibria in each regime, which is the same even in the case of generation of non-trivial fixed points for $\alpha \neq 0$:

- in the case **(b)** the saddle nature of $(0,0)$ and $(0,\pi)$ give rise to four non-homothetic stable fixed points, whose presence are balanced by four non-homothetic saddles in the vicinity of $(\pi/2, 0)$ and $(3\pi/2, 0)$;

- in the case **(c)**, all the homothetic fixed points result to be stable; although the stable equilibrium points generated by the saddles in $(0,0)$ and $(\pi,0)$ disappear with their change of stability, the saddles near to $(\pi,0)$ and $(3\pi/2, 0)$ still remain, leading to four stable and four unstable points;

- in the case **(d)**, the stability of the homothetic points is balanced, and no other equilibrium points are detected.

This non-trivial fact is coherent with the results one can obtain by applying the theory of the topological degree to the study of the stability of the fixed points in a discrete dynamical system (cfr. [73]), although the rigorous application of such theory would require the good definition and non-degeneration of F on the whole ellipse.

In view of the approximation given in Section 1.6.1, it is reasonable to think that for small eccentricities and small values of the physical parameters the dynamics induced by F does not differ much to the one sketched in Figure 1.10. Nevertheless, when de ellipse becomes more eccentric or the parameters differ much from each others, a variety of behaviours can manifest, including the presence of diffusive orbits, that are strong indicators of chaos.

Figure 1.11 shows the transition of F for $e = 0.05$, $\mathcal{E} = 20$, $\omega = 1$, $\mu = 0.13$ and $h = 1$ **(a)**, $h = 10$ **(b)**, $h = 40$ **(c)**, namely, for e very small but with a high difference in magnitude between h and μ. In the considered regime, direct computations assure that $\Delta^{(0)} > 0$ and $\Delta^{(1)} < 0$, leading to the conclusion $(\pi/2, 0)$ is a center and $(0,0)$ is an unstable saddle. For increasing values of h, the saddle orbits around $(0,0)$ tend to diffuse, leading finally a chaotic cloud which surrounds the two stability islands. As in the case of Figure 1.10, the chaotic region is bounded by invariant curves which induce oscillating rotations on the ellipse. Furthermore, periodic orbits of period 4 **(b)** and 3 **(c)** are detectable.

The other factor which can induce chaotic behaviour is the increasing eccentricity of the domain.
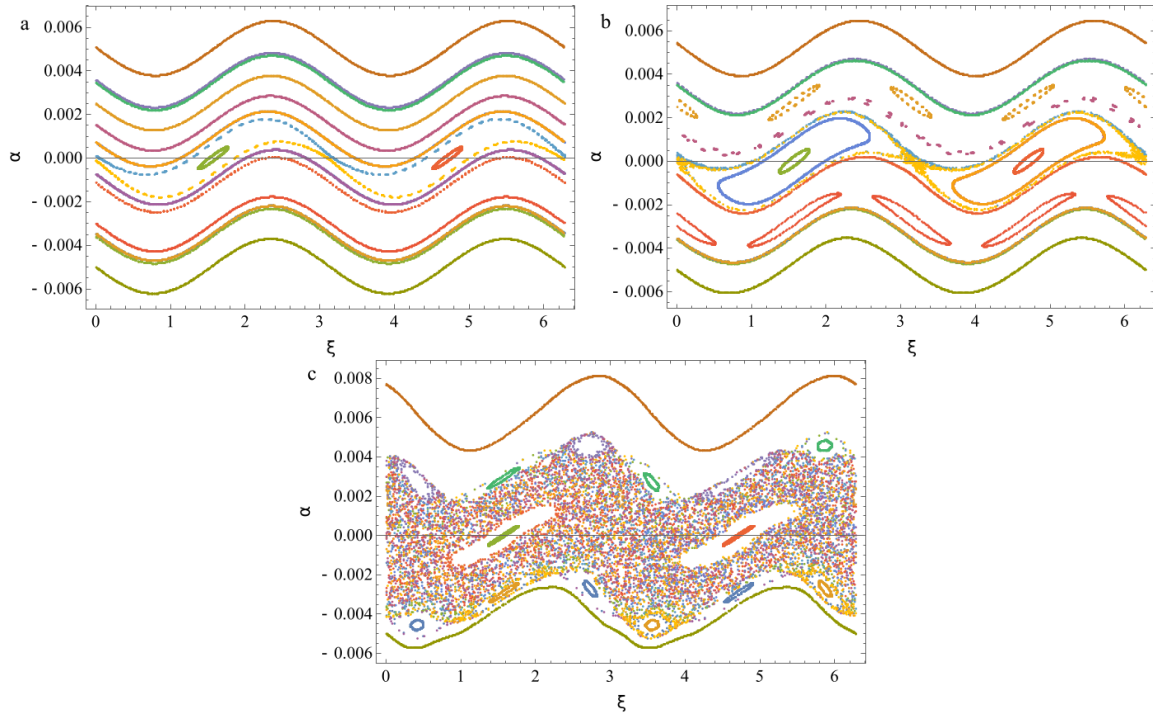
Fig. 1.11 Plots of F for $e = 0.05$, $\mathcal{E} = 20, \omega = 1, \mu = 0.13$ and $h = 1$ **(a)**, $h = 10$ **(b)**, $h = 40$ **(c)**. The chaotic behaviour around $(0,0)$ and $(\pi, 0)$ is evident even for very small eccentricities.
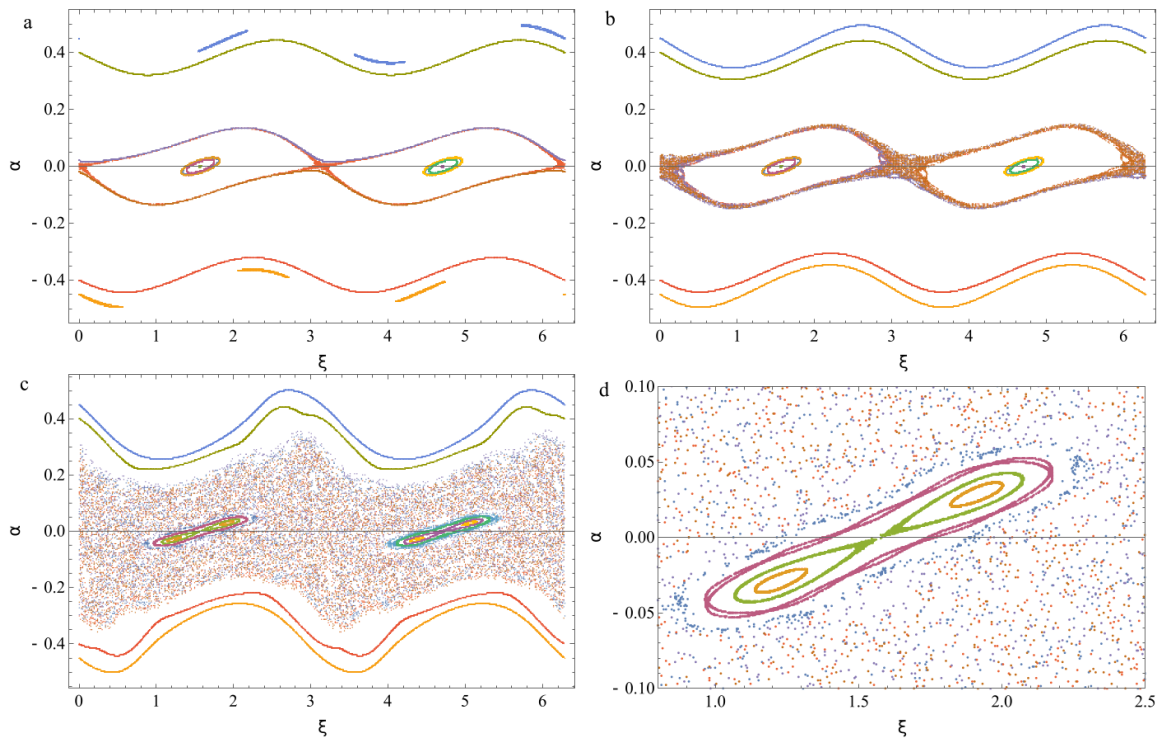
Fig. 1.12 Plots of F for $e = 0.3$, $\mathcal{E} = 2.5, \omega = \sqrt{2}, \mu = 1$ and $h = 0.1$ **(a)**, $h = 1$ **(b)**, $h = 7$ **(c)**. **(d)**: refining of **(c)** in a neighbourhood of $(\pi/2, 0)$.

Figure 1.12 illustrates how for moderate values of the eccentricity the system could have diffusive orbits around the unstable fixed points, even for small values of the physical parameters. This case is analogous to Figure 1.8, where the transition of $(\pi/2, 0)$ from center to saddle produce two 2-periodic brake orbits.

Chapter 2

KAM and Aubry-Mather theories for the close-to-circle refractive case

2.1 Introduction and statement of the results

In this chapter we will continue the analysis started in Chapter 1, addressing the general issue of periodic and quasi-periodic orbits (see Figure 2.1) and associated caustics when the domain is a perturbation of the circle, taking advantage of both KAM and Aubry-Mather theories (see for example [22–24]).

In order to tackle the problem, we shall recall the definition of the first[1] return map at the interface ∂D, after two consecutive (outer and inner) excursions, working at the zero energy level for the potential V as in (1.1.1). Here, we shall exploit the Lagrangian structure of the problem, building (locally) a generating function as the sum of an inner and outer contribution. Here comes a first problem, as these can not be globally defined. In addition, major difficulties arise from the return map not being globally defined, from the singularity of the attraction center and from a lack of twist condition

---

[1] The phrase *first return* may be confusing, since it is in fact a second crossing of the boundary; however, such apparent ambiguity will be clarified in Section 2.2.
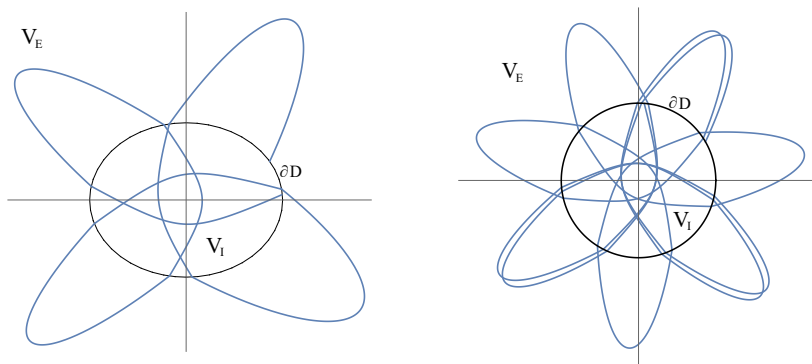
Fig. 2.1 Examples of trajectories for $\mathcal{E} = 2.5$, $\omega = 1$, $h = 2$ and $\mu = 2$. Left: general trajectory for an elliptic domain with eccentricity $e = 0.6$. Right: quasi-periodic trajectory for a circular domain.

(cfr. [25]), as it can be in shown that, in certain regimes of the parameters, even the completely integrable circular case admits at least a twist change (see Remark 2.3.8). Here are our main results.

Theorem 2.1.1 (Circular domains). *When D is the unit circle, there are action-angle coordinates $(\xi, I) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times (-I_c, I_c)$, where ξ is the polar angle, such that we can express the first return map as a shift*

$$
\begin{aligned}
\mathcal{F} : \mathbb{R}_{/2\pi\mathbb{Z}} \times (-I_c, I_c) &\to \mathbb{R}_{/2\pi\mathbb{Z}} \times (-I_c, I_c), \\
(\xi_0, I_0) &\mapsto (\xi_1, I_1) = (\xi_0 + \bar{\theta}(I_0), I_0),
\end{aligned}
\tag{2.1.1}
$$

where $\bar{\theta}(I) = f(I) + g(I)$ and

$$
f(I) = \begin{cases}
\operatorname{arccot}\left(\dfrac{\mathcal{E} - 2I^2}{I\sqrt{4\mathcal{E} - 2(2I^2 + \omega^2)}}\right) & \text{if } I \in (0, I_c) \\
0 & \text{if } I = 0 \\
\operatorname{arccot}\left(\dfrac{\mathcal{E} - 2I^2}{I\sqrt{4\mathcal{E} - 2(2I^2 + \omega^2)}}\right) - \pi & \text{if } I \in (-I_c, 0)
\end{cases}
$$

and

$$
g(I) = \begin{cases}
2\arccos\left(\dfrac{2I^2 - \mu}{\sqrt{4(\mathcal{E} + h)I^2 + \mu^2}}\right) - 2\pi & \text{if } I \in (0, I_c) \\
0 & \text{if } I = 0 \\
-2\arccos\left(\dfrac{2I^2 - \mu}{\sqrt{4(\mathcal{E} + h)I^2 + \mu^2}}\right) + 2\pi & \text{if } I \in (-I_c, 0)
\end{cases}
$$

are real analytic functions in $(-I_c, I_c)$. For every $I \in (-I_c, I_c)$, except for a finite number (at most ten) of points, there holds $\bar{\theta}'(I) \neq 0$.

The action I can be expressed as a function of the variables (ξ, α) already introduced in Chapter 1. The critical value I_c corresponds to the action associated with the *total reflection* of the trajectory at the boundary, i.e. when the outgoing refracted trajectory becomes tangent to the boundary (see Section 1.2) and is given in (2.2.10). In the circular case, the *rotation number* (see [25]) of the orbit is $\rho_I = \bar{\theta}(I) = f(I) + g(I)$. Depending whether this value is rational with 2π or not, the corresponding orbits are periodic or quasi-periodic. In both cases they determine an invariant curve and a pair of caustics, that is, smooth closed curves such that every trajectory which starts tangent to remains tangent after every passage in and out the domain D. Let us point out that caustics come in pair, respectively in D and in its complement. The proof of this Theorem is performed in Section 2.3.

Next we consider a perturbation D_ε of the domain whose boundary $\partial D_\varepsilon = \gamma_\epsilon \left(\mathbb{R}_{/2\pi\mathbb{Z}} \right)$ is given by a radial deformation of the circle of the form

$$\gamma_\epsilon : \mathbb{R}_{/2\pi\mathbb{Z}} \to \mathbb{R}^2 \quad \gamma_\epsilon(\xi) = (1 + \varepsilon f(\xi; \varepsilon)) \, e^{i\xi}, \tag{2.1.2}$$

where $f(\xi; \varepsilon)$ is a suitably smooth function of $\mathbb{R}_{/2\pi\mathbb{Z}} \times [-C, C]$ and ξ is the polar angle. Then, the first return map on the perturbed boundary can be extended as $\mathcal{F}(\xi, I; \varepsilon)$ (see definition 2.4.18), where (ξ, I) are canonical variables defined in suitable neighbourhoods of such invariant curves. The following theorem resumes the results stated in Theorems 2.4.12, here considering a single invariant curve, and 2.4.21.

Theorem 2.1.2 (Persistence of invariant curves (KAM)). *Let $f \in \mathcal{C}^k$, with $k > 5$. Let us suppose that $\bar{\theta}'(I_0) \neq 0$, and assume $\rho_0 = \bar{\theta}(I_0)$ has a Diophantine ratio with 2π (see Definition 2.4.8). Then there exists $\bar{\varepsilon}_{\rho_0}$ such that for every $\varepsilon \in \mathbb{R}$, $|\varepsilon| < \bar{\varepsilon}_{\rho_0}$ the map $\mathcal{F}(\xi, I; \varepsilon)$ admits a closed invariant curve of class C^1 with rotation numbers ρ_0. Each of these invariant curves generates a pair of regular caustics.*

Two invariant curves with Diophantine rotation numbers border an invariant region for the map $\mathcal{F}(\xi, I; \varepsilon)$, subject to the application of Poincaré-Birkhoff theorem and Aubry-Mather theory (see [23, 24, 22]). As the map is area-preserving, we only need to verify the twist condition. This is a nontrivial issue, as the function $\bar{\theta}(I)$ may indeed change its monotonicity. This fact poses some technical difficulties but also gives rise to a richer phenomenology. We have the following result.

Theorem 2.1.3 (Existence of Aubry-Mather invariant sets). *Let $\bar{\rho}_- < \bar{\rho}_+$ be Diophantine rotation numbers, such that there are no critical values of the function $\bar{\theta}$ in the range $[\bar{\rho}_-, \bar{\rho}_+]$. Then there exists $\bar{\varepsilon} > 0$ such that for every $\varepsilon \in \mathbb{R}$ with $|\varepsilon| < \bar{\varepsilon}$ and for every $\rho \in [\bar{\rho}_-, \bar{\rho}_+]$ the map $\mathcal{F}(\xi, I; \varepsilon)$ admits at least one orbit with rotation number ρ. When $\rho = 2\pi \dfrac{m}{n}$ then for ε sufficiently small there are at least 2 (m, n)-orbits, namely, such that, denoted with $\{(\xi_k, I_k)\}_{k \in \mathbb{N}} = \left\{ \mathcal{F}^k(\xi_0, I_0) \right\}_{k \in \mathbb{N}}$ the orbit generated by the initial point (ξ_0, I_0), one has*

$$\forall k \in \mathbb{N} \quad (\xi_{k+n}, I_{k+n}) = (\xi_k + 2\pi m, I_k) \equiv_{2\pi} (\xi_k, I_k).$$

This statement can be easily deduced from Theorem 2.4.20; as we shall see there, however, the actual number of solutions can be larger, depending on the number of monotonicity changes of the twist $\bar{\theta}(I)$.

To conclude this preamble, let us remark that, although the computations are not explicitly performed here, with means of the same analytical tools and techniques other variations of the considered model, more similar to the one described in [9], can be investigated.

As an example, let us consider $\epsilon \in \mathbb{R}$ and a non-isotropic perturbation of the outer potential given by

$$\tilde{V}_E(z) = \mathcal{E} - \frac{\omega^2}{2} x^2 - \frac{(\omega + \epsilon)^2}{2} y^2,$$

where $z = (x, y) \in \mathbb{R}^2$. Taking, if necessary, $h \in \mathbb{R}$ instead of $h > 0$, one can consider the dynamics induced by the potential

$$\tilde{V}(z) = \begin{cases} V_I(z) & \text{if } z \in D \\ \tilde{V}_E(z) & \text{if } z \notin D \end{cases} \tag{2.1.3}$$

when

$$\partial D = \{ z \in \mathbb{R}^2 \mid V_I(z) = \tilde{V}_E(z) \};$$

In this case, the potential is continuous and the refraction reduces to a conservation of both the position and velocity on ∂D, leading to a C^1 junction between inner and outer arcs. For $\varepsilon \in \mathbb{R}$ small, this system be considered again as a small perturbation of the circular case.

Let us conclude this section by underlining, as already done in Chapter 1, the connections between our refractive case and the classical Birkhoff case related to the methods used to study the respective dynamics: in this chapter, the fundamental argument to prove Theorem 2.1.3 is the extension to small perturbations of the existence results obtained for a circular domain, which is completely integrable. Analogous reasonings are used for example in [74] and [20] to prove a local version of Birkhoff conjecture starting from small deformations respectively of a circular and an elliptic domain, both sharing the complete integrability property in the classic case.
The figures of the current chapter are taken from [2].

2.2   First return map

As in Chapter 1, we will use the first return map to describe the dynamics in the close-to-circle case as well. On the other hand, in the framework presented in this chapter we will use a *variational* approach, starting from the definition of a suitable generating function, as it is usually done in the case of classical Birkhoff billiards (see [17]). In this framework, the variational formulation of Snell's law (1.2.12) turns out to be crucial for the generating function to be well defined.

2.2.1   Variational approach

Choosing the right set of canonical variables, the first return map already defined in Section 1.4 can be expressed in a variational form, which allows to take advantage of a wide range of powerful theoretical tools, coming from KAM and Aubry-Mather theories, which allow to tackle the case of small perturbations of a circular domain D_0 (the latter is discussed in Section 2.3, while the perturbative approach is the subject of

Section 2.4). Following e.g. [25], we can consider the *generating function*

$$S : \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathbb{R}_{/2\pi\mathbb{Z}} \to \mathbb{R},$$
$$S(\xi_0, \xi_1) = S_E(\xi_0, \tilde{\xi}) + S_I(\tilde{\xi}, \xi_1) = d_E(\gamma(\xi_0), \gamma(\tilde{\xi})) + d_I(\gamma(\tilde{\xi}), \gamma(\xi_1)), \tag{2.2.1}$$

where d_E and d_I are defined as in Section 1.2 and, according to Snell's law, the intermediate point $\tilde{p} = \gamma(\tilde{\xi})$ is such that $\tilde{\xi}$ is a critical point for the function $f(\xi_0, \xi, \xi_1) = d_E(\gamma(\xi_0), \gamma(\xi)) + d_I(\gamma(\xi), \gamma(\xi_1))$ with ξ_0 and ξ_1 fixed. In other words, $\tilde{\xi}$ is a solution of

$$\partial_b S_E(\xi_0, \tilde{\xi}) + \partial_a S_I(\tilde{\xi}, \xi_1) = 0, \tag{2.2.2}$$

where ∂_a and ∂_b denote respectively the partial derivatives with respect to the first and second variable. In general, one can not guarantee the global good definition of the generating function $S(\xi_0, \xi_1)$, which depends strongly on the specific geometry of the domain and on the values of the physical parameters $\mathcal{E}, h, \mu, \omega$. On the other hand, the local definition of $S(\xi_0, \xi_1)$ needs a nondegeneracy condition: under the assumption of existence and uniqueness of the inner and outer arcs, which ensures the differentiability of $d_E(p_0, p_1)$ and $d_I(p_0, p_1)$ separately, let us consider $(\xi_0, \tilde{\xi}, \xi_1) \in \left(\mathbb{R}_{/2\pi\mathbb{Z}}\right)^3$ such that (2.2.2) is satisfied. If

$$\partial_{\tilde{\xi}}(\partial_b S_E(\xi_0, \tilde{\xi}) + \partial_a S_I(\tilde{\xi}, \xi_1)) = \partial_b^2 S_E(\xi_0, \tilde{\xi}) + \partial_a^2 S_I(\tilde{\xi}, \xi_1) \neq 0, \tag{2.2.3}$$

then locally around ξ_0, ξ_1 one can express $\tilde{\xi} = \tilde{\xi}(\xi_0, \xi_1)$ as a function of the endpoints. Moreover, one has

$$\partial_{\xi_0}\tilde{\xi}(\xi_0, \xi_1) = -\frac{\partial_{ab} S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1))}{\partial_b^2 S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1)) + \partial_a^2 S_I(\tilde{\xi}(\xi_0, \xi_1), \xi_1)},$$
$$\partial_{\xi_1}\tilde{\xi}(\xi_0, \xi_1) = -\frac{\partial_{ab} S_I(\tilde{\xi}(\xi_0, \xi_1), \xi_1)}{\partial_b^2 S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1)) + \partial_a^2 S_I(\tilde{\xi}(\xi_0, \xi_1), \xi_1)}. \tag{2.2.4}$$

If (2.2.3) holds, the generating function is well defined locally around ξ_0 and ξ_1, and one can define the *canonical actions* associated to the system and to the above coordinates by the relations

$$I_0 = -\partial_{\xi_0} S(\xi_0, \xi_1), \quad I_1 = \partial_{\xi_1} S(\xi_0, \xi_1); \tag{2.2.5}$$

note that, when $S(\xi_0, \xi_1)$ is well defined, the same is true also for the actions as functions of the angles ξ_0 and ξ_1.

In order to define a first return map in the new canonical action-angle variables, one needs to express ξ_1 and I_1 as functions of ξ_0 and I_0: this is possible if a second nondegeneracy condition holds. Let us consider $\xi_0, \xi_1, \tilde{\xi}(\xi_0, \xi_1)$ such that (2.2.3) holds, and define I_0 as in (2.2.5): if

$$\partial_{\xi_1}(I_0 + \partial_{\xi_0} S(\xi_0, \xi_1)) \neq 0, \tag{2.2.6}$$

one can find $\xi_1 = \xi_1(\xi_0, I_0)$ as a function of the initial action-angle variables. In particular, making use of (2.2.2) and (2.2.4), condition (2.2.6) translates in

$$\frac{\partial_{ab}S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1))\partial_{ab}S_I(\tilde{\xi}(\xi_0, \xi_1), \xi_1)}{\partial_b^2 S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1)) + \partial_a^2 S_I(\tilde{\xi}(\xi_0, \xi_1), \xi_1)} \neq 0, \tag{2.2.7}$$

which is well defined in view of (2.2.3). If (2.2.7) holds, one can then find two neighbourhoods $[\xi_0 - \lambda_{\xi_0}, \xi_0 + \lambda_{\xi_0}]$ and $[I_0 - \lambda_{I_0}, I_0 + \lambda_{I_0}]$ such that the new *local first return map*

$$\mathcal{F} : [\xi_0 - \lambda_{\xi_0}, \xi_0 + \lambda_{\xi_0}] \times [I_0 - \lambda_{I_0}, I_0 + \lambda_{I_0}] \to \mathbb{R}_{/2\pi\mathbb{Z}} \to \mathbb{R},$$
$$(\xi_0, I_0) \mapsto (\xi_1(\xi_0, I_0), I_1(\xi_0, I_0)), \tag{2.2.8}$$

where $I_1(\xi_0, I_0) = \partial_{\xi_1}S(\xi_0, \xi_1(\xi_0, I_0))$, is well defined.

The switch from a *Lagrangian* approach adopted by using the original first return map F and an *Hamiltonian* one involving the generating function and the canonical action-angle variables is crucial as it will allow, in Section 2.4, to take advantage of the results coming from KAM and Aubry-Mather theories to derive the existence of orbits with prescribed rotation numbers in the case of small perturbations of a circular domain (see Section 2.4).

In the circular case, which will be analyzed in details in Section 2.3, both the potentials and the domain are centrally symmetric: a consequence of the subsequent invariance under rotations is that the nondegeneracy conditions (2.2.3) and (2.2.7) will result in fact equivalent, and S, denoted in this case with S_0, will be well defined almost everywhere.

As a final remark, one can observe that Snell's law, along with (2.2.5) and (2.2.2), provides a general relation between the actions I_0, I_1 and the angles α_0, α_1 defined in Section 1.4: in particular,

$$I_0(\xi_0, \xi_1) = -\partial_{\xi_0}(S_E(\xi_0, \tilde{\xi}) + S_I(\tilde{\xi}, \xi_1))) = -\partial_a S_E(\xi_0, \tilde{\xi}) =$$
$$= \frac{1}{\sqrt{2}}z'_E(0) \cdot \frac{\dot{\gamma}(\xi_0)}{\|\dot{\gamma}(\xi_0)\|} = \sqrt{V_E(\gamma(\xi_0))}\sin\alpha_0$$
$$I_1(\xi_0, \xi_1) = \partial_{\xi_1}(S_E(\xi_0, \tilde{\xi}) + S_I(\tilde{\xi}, \xi_1))) = \partial_b S_I(\tilde{\xi}, \xi_1) =$$
$$= \frac{1}{\sqrt{2}}z'_I(T_E + T_I) \cdot \frac{\dot{\gamma}(\xi_1)}{\|\dot{\gamma}(\xi_1)\|} = \sqrt{V_I(\gamma(\xi_0))}\sin\alpha'_1 = \sqrt{V_E(\gamma(\xi_1))}\sin\alpha_1.$$
$$\tag{2.2.9}$$

Eqs.(2.2.9) provides a natural boundary for the values that the actions can assume, which, in a inhomogenous case, depend on ξ_0 and ξ_1: in particular

$$|I_0| \leq \sqrt{V_E(\gamma(\xi_0))} \text{ and } |I_1| \leq \sqrt{V_E(\gamma(\xi_1))},$$

where the equalities correspond to the tangent case $\alpha_{0\backslash 1} = \pm\pi/2$.

In the case of a circular domain, the bound on the actions is uniform and given by

$$I_0, I_1 \in \left[-\sqrt{\mathcal{E} - \frac{\omega^2}{2}}, \sqrt{\mathcal{E} - \frac{\omega^2}{2}} \right] = \left[-I_c, I_c \right]. \tag{2.2.10}$$

For reasons which will be clear in Section 2.4, related to the good definition of the map $\mathcal{F}$, we will not consider the tangent case, restricting ourselves to the open interval $(-I_c, I_c)$.

2.2.2 Good definition of d_E and d_I: preliminary results

In this section we deal with the good definition of the Jacobi distances d_E and d_I, which, as already pointed out in Section 1.2, Chapter 1, is heavily based on the existence and uniqueness of the solutions of the problems (recall the notation introduced in 1.3.11)

$$\begin{cases} (HS_E)[z_E(s)] & s \in [0, T_E] \\ z_E(s) \notin D & s \in (0, T_E) \\ z_E(0) = p_0^E, \ z_E(T_E) = p_1^E \end{cases} \quad \begin{cases} (HS_I)[z_I(s)] & s \in [0, T_I] \\ z_I(s) \in D & s \in (0, T_I) \\ z_I(0) = p_0^I, \ z_I(T_I) = p_1^I \end{cases} \tag{2.2.11}$$

for some $T_E, T_I > 0$ and fixed $p_0^E, p_1^E, p_0^I, p_1^I \in \partial D$. Unlike Chapter 1, where the Cauchy problems with prescribed initial conditions are considered, here the inner and outer fixed-ends problems are taken into account: our aim is to find conditions for which problems (2.2.11) admit unique solutions. Depending on the specific nature of the considered problem (if for example we are dealing with close-to-circle domains, as in the case of the current chapter, or with general domains, as in Chapter 3), these conditions could be related to the billiard shape, the endpoints, the physical parameters, as well as the geometric properties of the solutions themselves. Here we will present some auxiliary results, which will be used both in Chapters 2 and 3 to prove the final results in the respective frameworks.

The outer and inner cases will be treated in different ways, in particular:

- outer arcs will connect points near to the same central configuration (as defined in 1.1.2) and will be obtained via perturbative methods;

- inner ones connect not antipodal points near possibly different central configurations and will act as *transfer orbits* between possibly disjoint regions of ∂D; these arcs are obtained by purely geometric arguments.

We can say that $\bar{\xi} \in I$ is a $\lceil$central configuration$\rceil$ if the corresponding point $\gamma\left(\bar{\xi}\right)$ satisfies the conditions in 1.1.2.

Let us start with the outer dynamics, observing that the following general result holds for both Chapter 2 and 3.

Theorem 2.2.1 (Local existence of the outer arcs)**.** *Let $\bar{\xi} \in I$ be a central configuration. Then there exists $\delta_{E,\bar{\xi}} > 0$ such that for every $\xi_1, \xi_2 \in (\bar{\xi} - \delta_{E,\bar{\xi}}, \bar{\xi} + \delta_{E,\bar{\xi}})$ there is a unique solution $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2)) : [0, T] \to \mathbb{R}^2$ of the fixed-ends problem*

$$\begin{cases} (HS_E)\,[z(s)] & s \in [0, T] \\ z(s) \notin D & s \in (0, T) \\ z(0) = \gamma(\xi_1), \ z(T) = \gamma(\xi_2) \end{cases} \tag{2.2.12}$$

for some $T \doteq T(\xi_1, \xi_2) > 0$. Moreover the solution $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2))$ is transversal to the boundary ∂D at the endpoints, namely,

$$z_E(0; \gamma(\xi_1), \gamma(\xi_2)) \not\Vert \dot{\gamma}(\xi_1) \quad and \quad z_E(T; \gamma(\xi_1), \gamma(\xi_2)) \not\Vert \dot{\gamma}(\xi_2).$$

Proof. The proof of the theorem relies again on a perturbative strategy starting from the homothetic solution. Let us then start with the case $\xi_1 = \xi_2 = \bar{\xi}$: consider the Cauchy problem

$$\begin{cases} z''(s) = -\omega^2 z(s) \\ z(0) = \gamma(\bar{\xi}), \ z'(0) = v_0 \doteq \sqrt{2\left(\mathcal{E} - \omega^2 \|\gamma(\bar{\xi})\|^2\right)} \dfrac{\gamma(\bar{\xi})}{\|\gamma(\bar{\xi})\|}. \end{cases}$$

Its solution is the outer homothetic arc in the direction of $\gamma(\bar{\xi})$

$$z_E(s; \gamma(\bar{\xi}), \gamma(\bar{\xi})) = \cos(\omega s)\gamma(\bar{\xi}) + \frac{\sin(\omega s)}{\omega} v_0.$$

Setting $T = 2\pi/\omega$, one has that $z_E(T; \gamma(\bar{\xi}), \gamma(\bar{\xi})) = \gamma(\bar{\xi})$, and, by the choice of v_0,

$$\|z_E'(s; \gamma(\bar{\xi}), \gamma(\bar{\xi}))\|^2 + \omega^2 \|z_E(s, \gamma(\bar{\xi}), \gamma(\bar{\xi}))\|^2 = 2\mathcal{E} \quad \text{for every} \quad s \in [0, T],$$

so that the energy conservation law is satisfied. Moreover, from the second condition in (1.1.2), it results that $z_E(s; \gamma(\bar{\xi}), \gamma(\bar{\xi})) \notin \bar{D}$ for every $s \in (0, T)$; summing up the previous considerations, one can conclude that $z_E(\cdot; \gamma(\bar{\xi}), \gamma(\bar{\xi}))$ is a solution of (2.2.12) when $\xi_1 = \xi_2 = \bar{\xi}$.

Let us now pass to the general case; by the regularity of γ and the hypotheses on $\bar{\xi}$, there exists $\delta > 0$ such that for every $\xi \in (\bar{\xi} - \delta, \bar{\xi} + \delta)$:

- $\gamma(\xi) \not\Vert \dot{\gamma}(\xi)$;

- the half-line starting from 0 in the direction of $\gamma(\xi)$ intersects ∂D only once.

This means that, denoting with S the sector between the half-lines in the direction of $\gamma(\bar{\xi} - \delta)$ and $\gamma(\bar{\xi} + \delta)$, one has that it intersects ∂D exactly in $\gamma((\bar{\xi} - \delta, \bar{\xi} + \delta))$.

By the differentiable dependence of the Bolza problem with respect to variations of the endpoints, along with the regularity of γ, there exists $\delta_{E,\bar{\xi}} \in (0, \delta)$ such that for every $\xi_1, \xi_2 \in (\bar{\xi} - \delta_{E,\bar{\xi}}, \bar{\xi} + \delta_{E,\bar{\xi}})$ there is $T = T(\xi_1, \xi_2) > 0$ and a unique $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2))$

solution of

$$\begin{cases} (HS_E)[z(s)] \\ z(0) = \gamma(\xi_1), \, z(T) = \gamma(\xi_2). \end{cases}$$

Moreover, $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2))$ converges in C^1-norm to the homothetic solution $z_E(\cdot; \gamma(\bar{\xi}), \gamma(\bar{\xi}))$ when $(\xi_1, \xi_2) \to (\bar{\xi}, \bar{\xi})$: possibly taking a smaller $\delta_{E,\bar{\xi}}$, one has then that for every $\xi_1, \xi_2 \in (\bar{\xi} - \delta_{E,\bar{\xi}}, \bar{\xi} + \delta_{E,\bar{\xi}})$ the corresponding solution $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2))$ is completely contained in the sector S, and then $z_E(s; \gamma(\xi_1), \gamma(\xi_2)) \notin \bar{D}$ for every $s \in (0, T)$. The same convergence property also ensures that, for $\delta_{E,\bar{\xi}}$ small enough, $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2))$ is transversal to the boundary ∂D at the endpoints. $\qquad\square$

As for the inner dynamics, the perturbative approach around the homothetic arcs is not sufficient, as our aim is to connect points close to possibly different central configurations (we stress that this was not necessary in Chapter 1, where only the local dynamics in the vicinity of the homothetic trajectories has been considered). We will rather use a geometric approach, based on classical results in Celestial Mechanics (see [70, 75]) and somehow similar to the ones obtained, for different cases, in [71, 72]. A further difficulty is related to the uniqueness of the inner arc, which is obtained by requiring an additional topological condition on the solution z_I of (2.2.11).

Definition 2.2.2. *Let $P_1, P_2 \in \mathbb{R}^2 \setminus \{0\}$ be not antipodal points and $\alpha : [0, A] \to \mathbb{R}^2 \setminus \{0\}$ be a continuous curve such that $\alpha(0) = P_1$ and $\alpha(A) = P_2$. We say that α is **topologically non-trivial** (TnT) if $\alpha([0, A])$ is not homotopic to the line segment connecting P_1 and P_2 in the punctured plane $\mathbb{R}^2 \setminus \{0\}$.*

We stress thatm in Chapter 1, such condition was not necessary, since the Cauchy problem, instead of the fixed-ends one, has been considered. On the other hand, in that case the inner arcs obtained by perturbing the collision-ejection homothetic arc enjoy the *(TnT)* property by contruction.

Let us start recalling some basic facts on Keplerian hyperbolæ. In general, the solutions for the Kepler problem with central mass μ at the origin and fixed energy $H > 0$

$$\begin{cases} z''(s) = -\mu \dfrac{z(s)}{\|z(s)\|^3} & s \in \mathbb{R} \\ \dfrac{1}{2}\|z'(s)\|^2 - \dfrac{\mu}{\|z(s)\|} = H & s \in \mathbb{R} \end{cases} \tag{2.2.13}$$

are branches of hyperbola with a focus at the origin. In particular (see [70]):

- the hyperbola's semimajor axis a is given by $a = \dfrac{\mu}{2H}$;

- the eccentricity e is

$$e = \sqrt{1 + \frac{2Hk^2}{\mu^2}},$$

    where k is the norm of the angular momentum, which is constant along the solutions;
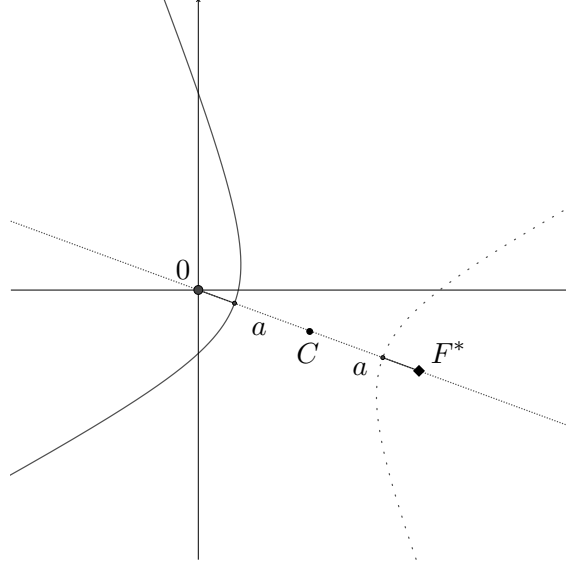
Fig. 2.2 Keplerian hyperbola with a focus in the origin, center C and second focus F^*. The semimajor axis a is the distance between C and the pericenter of the arc.

- the selected branch is always the closest to the origin.

If we call F^* the second focus of the hyperbola (see Figure 2.2), one has that $\|F^*\| = 2ae$, and, by the analytic definition of the conic, any point P of the hyperbolic Keplerian arc satisfies

$$\|\overline{PF^*}\| - \|\overline{P0}\| = 2a. \tag{2.2.14}$$

These geometric considerations allow us to state an existence theorem for Keplerian arcs connecting two distinct points of $\mathbb{R}^2 \setminus \{0\}$.

Theorem 2.2.3 (Global existence of inner arcs)**.** *Let $P_1, P_2 \in \mathbb{R}^2 \setminus \{0\}$, $P_1 \neq P_2$, be two distinct points such that either:*

(NC) they are not collinear with the origin, or

(C) the origin is contained in the segment $\overline{P_1 P_2}$.

Then for every $H, \mu > 0$ there are exactly two Keplerian arcs with energy H and central mass μ connecting P_1 to P_2. In particular, the two arcs are contained in the opposite half-planes generated by the line connecting the two points. Moreover, if condition (NC) is verified, then exactly one of the two arcs is (TnT).

Proof. We prove the theorem by showing that for every P_1, P_2 satisfying the hypotheses, there are exactly two possible positions for the second focus F^*. Let us call $r_i = \|P_i\|$, $i = 1, 2$, and assume, without loss of generality, that $r_1 \geq r_2$; then from (Eq. 2.2.14) we have that $\|P_i F^*\| = r_i + 2a$, $i = 1, 2$, which means that F^* belongs to both the circles centered in P_i with radius $R_i = r_i + 2a$ for $i = 1, 2$. By a simple result from elementary geometry, one has that the two circumferences have two distinct intersection points if and only if

$$R_1 - R_2 < \overline{P_1 P_2} < R_1 + R_2,$$
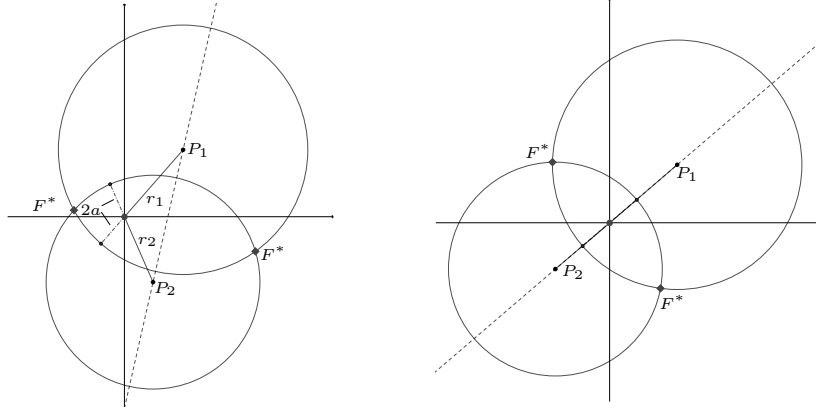
Fig. 2.3 Possible positions of the virtual focus F^* in the non collinear (left) and collinear (right) case.

which, recalling the definition of the radii, is equivalent to

$$r_1 - r_2 < \overline{P_1 P_2} < r_1 + r_2 + 4a.$$

As for any $H > 0$ the semimajor axis a is positive, the above condition is satisfied if and only if condition (NC) or (C) hold (see Figure 2.3): in this case, the two circumferences have exactly two intersection points, corresponding to the only two possible positions for the second focus F^*. This proves the existence of the two hyperbolæ.

As for the topological *(TnT)* characterization, let us observe that if the points P_1 and P_2 are distinct and not collinear with the origin, it is possible to divide the plane $\mathbb{R}^2$ into two half-planes, separated by the line connecting the two points (see Figure 2.3), and such that only one of them contains the origin. We will call $\mathcal{H}_0(P_1, P_2; H)$ the Keplerian branch of hyperbola lying in the same half-plane with the origin and $\mathcal{H}_1(P_1, P_2; H)$ the other one. It is then straightforward to observe that the branch $\mathcal{H}_1(P_1, P_2; H)$ is homotopic to the line segment connecting P_1 and P_2, while $\mathcal{H}_0(P_1, P_2; H)$ has the *(TnT)* property.

$\square$

Theorem 2.2.3 will be the starting point for proving the existence and uniqueness of a *(TnT)* inner arc in both Chapter 2 and 3. Nevertheless, let us observe that, unlike Theorem 2.2.1, it does not give any information on whether our arc lies inside of D or not. This additional condition, necessary for the inner dynamics to be well defined, will be proved separately in the two chapters; in particular:

- in Chapter 2, constraints on the boundary ∂D are required, as it must be sufficiently close to a circle;

- in Chapter 3, where a very general class of domains is considered, the result will be achieved by requiring that the energy jump h from the outer to the inner region is large enough.

For this reason, we will leave to Sections 2.3.1, 2.4.1 and 3.2 the analogous of Theorem 2.2.1 for the inner dynamics.
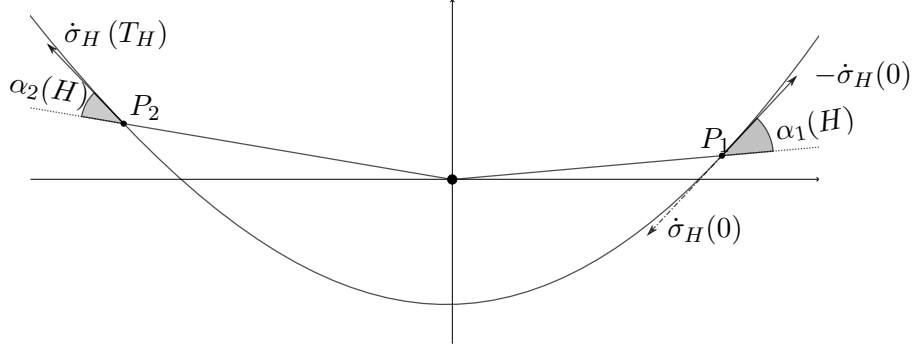
Fig. 2.4 Representation of the angles $\alpha_1(H)$ and $\alpha_2(H)$ between the radial directions defined by the points P_1 and P_2 and the tangent vectors $-\dot{\sigma}_H(0)$ and $\dot{\sigma}_H(T_H)$ to the Keplerian hyperbolic arc $\mathcal{H}_0(P_1, P_2; H)$ (see Eq. (2.2.15) and the definitions above.)

Let us now focus our attention on the *transversality* properties of our hyperbola, which, as already pointed out in Section 1.2, are fundamental to guarantee the good definition of Snell's law. Fix then P_1, P_2 and μ, and define the families

$$\mathcal{F}_i \equiv \mathcal{F}_i(P_1, P_2) \doteq \{\mathcal{H}_i(P_1, P_2; H), \ H > 0\} \quad i = 0, 1.$$

Focusing on $\mathcal{F}_0$, we parametrize its elements. For any $H > 0$, let σ_H be a smooth parametrization of the branch $\mathcal{H}_0(P_1, P_2; H)$ such that $\sigma_H(0) = P_1$ and $\sigma_H(T_H) = P_2$ for some $T_H > 0$. Moreover, let us introduce the following angles (see Figure 2.4):

$$\alpha_1(H) \equiv \alpha_1(P_1, P_2; H) \doteq \angle\,(P_1, -\dot{\sigma}_H(0))\,, \quad \alpha_2(H) \equiv \alpha_2(P_1, P_2; H) \doteq \angle\,(P_2, \dot{\sigma}_H(T_H))\,.$$
$$(2.2.15)$$

Proposition 2.2.4 provides an asymptotic estimate for the angles $\alpha_1(H)$ and $\alpha_2(H)$ as $H \to \infty$.

Proposition 2.2.4. *Let $\mu > 0$ and $P_1, P_2 \in \mathbb{R}^2 \setminus \{0\}$, $P_1 \neq P_2$, such that (NC) is satisfied, and for every $H > 0$ consider $\alpha_1(H)$ and $\alpha_2(H)$ as defined in (2.2.15). Then*

$$\lim_{H \to \infty} \alpha_i(H) = 0, \quad i = 1, 2.$$

Proof. Le us start recalling that the semimajor axis a of the hyperbola which generates the branch $\mathcal{H}_0(P_1, P_2; H)$ is given by $a(H) = \mu/2H$; hence, $a(H) \to 0$ when $H \to \infty$. Referring to the left side of Figure 2.3, one has that the intersection point between the two circumferences centered at P_1 and P_2 and with radii R_1 and R_2 lying in the same side of the origin tends to 0. Then, denoting with $F^*(H)$ the second focus of the considered hyperbola, we have

$$\lim_{H \to \infty} R^*(H) = 0, \quad R^*(H) \doteq \|F^*(H)\|. \qquad (2.2.16)$$

Define now $k(H)$ as the norm of the angular momentum for the corresponding Keplerian orbit along $\mathcal{H}_0(P_1, P_2; H)$. From the beginning of this Section, we have that

$$R^*(H) = 2ae = \frac{\mu}{H}\sqrt{1 + \frac{2Hk^2(H)}{\mu^2}},$$

hence

$$k(H) = \sqrt{\frac{(HR^*(H))^2 - \mu^2}{2H}}.$$

We choose now as a parametrization of $\mathcal{H}_0(P_1, P_2; H)$ the corresponding solution of the Kepler problem (2.2.13); we still call $\sigma_H(t)$, $t \in \mathbb{R}$, such a parametrization assuming that $\sigma_H(0) = P_1$ and $\sigma_H(T_H) = P_2$. By the conservation of the angular momentum we have

$$k(H) = \|\sigma_H(0)\| \, \|\dot{\sigma}_H(0)\| \, |\sin\alpha_1(H)|$$
$$= \|\sigma_H(T_H)\| \, \|\dot{\sigma}_H(T_H)\| \, |\sin\alpha_2(H)|,$$

while, by the energy conservation law in Eq. (2.2.13) and recalling that $r_i = \|P_i\|$, $i = 1, 2$, it results

$$\|\dot{\sigma}_H(0)\| = \sqrt{2}\sqrt{\frac{\mu}{r_1} + H}, \quad \|\dot{\sigma}_H(T_H)\| = \sqrt{2}\sqrt{\frac{\mu}{r_2} + H};$$

from which one obtains that, for $i = 1, 2$,

$$|\sin\alpha_i(H)| = \frac{1}{2\sqrt{r_i}}\sqrt{\frac{(HR^*(H))^2 - \mu^2}{r_i H^2 + \mu H}},$$

which, in view of (2.2.16), tends to zero when $H \to \infty$. $\qquad\qquad\square$

More precise estimates on the convergence rate of the angles α_1 to 0 will be given, in the specific case of the circular domain, in Section 2.3. Moreover, this result will be of crucial importance in Chapter 3, where more general domains are considered.

Remark 2.2.5. *As pointed out in [75, pp. 273-274], the branch $\mathcal{H}_0(P_1, P_2; H)$ as $H \to \infty$ converges to the two straight-line segments from P_1 to 0 and from 0 to P_2.*
With analogous computations coming from elementary geometry one can prove that the branch $\mathcal{H}_1(P_1, P_2; H)$ converges to the line connecting P_1 to P_2.
In this setting, the meaning of convergence is purely geometric and it is due to the convergence of the second foci to some fixed positions (since a tends to 0). We will translate this concept of convergence within a dynamical setting in Section 3.2 and further.

We stress that more precise results on the rate of convergence of the fixed-end hyperbolæ for $H \to \infty$ are provided in [76].

Remark 2.2.6. *When P_1 and P_2 satisfy condition (C), the two branches of hyperbolæ found in Theorem 2.2.3 lose the topological classification into $\mathcal{H}_0$ and $\mathcal{H}_1$, because the origin belongs to the segment connecting P_1 and P_2. Nevertheless, the geometric asymptotic behaviour described in Proposition 2.2.4 and Remark 2.2.5 continues to hold: in particular, both branches converge to the line connecting P_1 and P_2; hence, the tangent directions in P_1 and P_2 tend in both cases to the directions of P_1 and P_2.*

Remark 2.2.7. *When $P_1 = P_2$, trough the Levi-Civita regularization one can prove that the inner problem*

$$\begin{cases} (HS_I)[z(s)] \\ z(0) = z(T) = P_1 \end{cases} \qquad (2.2.17)$$
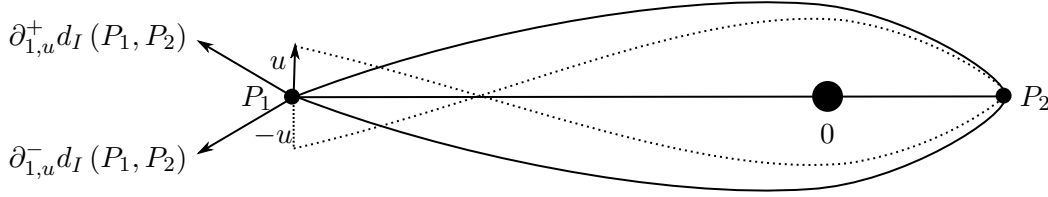
Fig. 2.5 Antipodal case: the derivative of $d_I(P_1, P_2)$ is not well defined in P_1.

admits a unique collision-ejection solution $z(\cdot; P_1, P_1; H)$ (which coincides to the segment connecting the origin to P_1) for some $T > 0$, in the sense that it solves problem (2.2.17) in $[0, T] \setminus \{T/2\}$. Although this collision-ejection solution can not be considered as a classical solution of the Kepler problem (one can refer to it as a regularised *solution of the latter), one can prove that it is the limit of the arc $\mathcal{H}_0(P_1, P_2; H)$, whenever $P_2 \to P_1$ (see Section 1.3 for the derivation and the detailed study of the regularised problem). Moreover, as $z(\cdot; P_1, P_1; H)$ is parallel to the straight half-line starting from 0 and containing P_1, the angles $\alpha_1(H)$ and $\alpha_2(H)$ defined in Eq. (2.2.15) are identically zero, and then Proposition 2.2.4 trivially holds also in the collisional case.*

Let us conclude this section with few words about the antipodal case. As already pointed out in Remark 2.2.6, when the origin is contained in the segment connecting P_1 to P_2 the *(TnT)* topological characterization does not make sense anymore. This is not only a technical difficulty, but has fundamental consequences on the differentiability of the inner distance d_I. As a matter of fact, let us suppose to start with two antipodal points P_1 and P_2; by Theorem 2.2.3, we know that there are exactly two Keplerian arcs connecting them, and it is easy to prove that they are symmetric with respect to the segment $\overline{P_1 P_2}$ and have the same Jacobi length: the distance $d_I(P_1, P_2)$ is then still well defined. As for its derivatives, from Section 1.2, and in particular Eq. (1.5.2) we know that they are related to the tangent vector to the chosen arc in the endpoints. Let us now apply a variation to one of the two points, say P_1, in the direction of a unit vector u transverse to P_1. Depending on the sign of our variation (either it points towards u or $-u$), the corresponding *(TnT)* arcs will not belong to the same homotopy class in the punctured plane $\mathbb{R}^2 \setminus \{0\}$, as they are deformations of one of the two distinct inner antipodal arcs (see Figure 2.5). As a consequence, directional derivative of d_I with respect to variations of the first endpoints in the direction of u is not well defined in P_1.

2.3   The unperturbed case: circular domain

Let us suppose now that D is a disk of radius 1, and denote it with D_0: in this particular case, both the potentials and the domain are centrally symmetric, and, as a consequence, the system is integrable. In particular, it is possible to find the explicit expression of the first return map in action-angle variables, which in this case is denoted with $\mathcal{F}_0$, as it will be done in Section 2.3.3. From now and until the end of the chap-

ter, we will assume that $\mathcal{E} > \omega^2$ to ensure that D_0 is contained in the outer Hill's region.

2.3.1   Good definition of d_E and d_I in the circular case

Let us now adapt the results already obtained in general in Section 2.2.2 when the domain D is a circle. As we will see, the explicit expression of ∂D_0, as well as its central symmetry, allows to obtain precise estimates on the quantities involved in Theorem 2.2.1 and its analogous for the inner case. Let us start with the outer case..

Theorem 2.3.1. *For every $p_0, p_1 \in \partial D_0$ such that $\|p_0 - p_1\| < 2$, there exists $T > 0$ and a unique $z(s; p_0, p_1) : [0, T] \to \mathbb{R}^2$ solution of the fixed ends problem*

$$\begin{cases} (HS_E)[z(s)] & s \in [0, T] \\ \|z(s)\| > 1 & s \in (0, T) \\ z(0) = p_0, \quad z(T) = p_1. \end{cases} \tag{2.3.1}$$

Moreover, $z(s; p_0, p_1)$ is of class C^1 with respect to variations of the endpoints.

This result can be verified using the same reasoning of Theorem 2.2.1 and observing some straightforward geometric features of the elliptic arcs connecting every pair of points on $\partial D0$. Nevertheless, it is worth to propose another proof, with more details, whose by-products will be useful in Section 2.3.2.

Proof. Fix $p_0 = e^{i\theta_0} \in \partial D_0$, and, given $\alpha \in (-\pi/2, \pi/2)$, consider the Cauchy problem

$$\begin{cases} z''(s) = -\omega^2 z(s), \\ z(0) = p_0, z'(0) = v_0 = \sqrt{2\mathcal{E} - \omega^2} e^{i\theta_0 + \alpha}, \end{cases} \tag{2.3.2}$$

whose solution $z(s; p_0, v_0)$ is an ellipse whose parameters depend on the initial conditions and can be decoupled as

$$z(s) = (x(s), y(s)) = \left(p_{0,x} \cos \omega s + \frac{v_{0,x}}{\omega} \sin \omega s, p_{0,y} \cos \omega s + \frac{v_{0,y}}{\omega} \sin \omega s \right) \tag{2.3.3}$$

Since $\alpha \in (-\pi/2, \pi/2)$, the orbit is exterior to D in a neigborhood of $s = 0$. Let $s_1 > 0$ the first positive instant for which $z(s_1; p_0, v_0) \in \partial D_0$ again, and define $p_1 = e^{i\theta_1} = p_1(p_0, \alpha) = z(s_1; p_0, v_0)$. As the system is invariant under rotations, the shift θ_E from θ_0 to θ_1 and s_1 depend only on the direction of v_0 with respect to the radial direction, i.e. on α. We can then fix $p_0 = \bar{p}_0 = (1, 0)$, and we have $\theta_E(\alpha) \doteq \theta_1$. The solution $z(s; \bar{p}_0, v_0)$ simplifies as

$$z(t) = (x(s), y(s)) = \left(\cos \omega s + \frac{v_x}{\omega} \sin \omega s, \frac{v_y}{\omega} \sin \omega s \right), \tag{2.3.4}$$

from which one has

$$r^2(s) = x^2(s) + y^2(s) = \frac{\omega^2 - \mathcal{E}}{\omega^2} \cos(2\omega s) + \frac{v_x}{\omega} \sin(2\omega s) + \frac{\mathcal{E}}{\omega^2} = A \cos(2\omega s + \bar{\alpha}) + \frac{\mathcal{E}}{\omega^2},$$
$$(2.3.5)$$

with $A \in \mathbb{R}$ and $\bar{\alpha} \in [0, 2\pi)$ such that

$$\left\{ A \cos \bar{\alpha} = \frac{\omega^2 - \mathcal{E}}{\omega^2}, \quad A \sin \bar{\alpha} = -\frac{v_x}{\omega}, \right. \qquad (2.3.6)$$

and, since $v_x > 0$,

$$\cot \bar{\alpha} = \frac{\mathcal{E} - \omega^2}{\omega v_x} \Rightarrow \bar{\alpha} = \operatorname{arccot}\left(\frac{\mathcal{E} - \omega^2}{\omega v_x}\right) \in \left(0, \frac{\pi}{2}\right),$$

$$\cos \bar{\alpha} = \frac{\mathcal{E} - \omega^2}{\sqrt{\omega^2 v_x^2 + (\mathcal{E} - \omega^2)^2}}, \quad \sin \bar{\alpha} = \frac{\omega v_x}{\sqrt{\omega^2 v_x^2 + (\mathcal{E} - \omega^2)^2}} \qquad (2.3.7)$$

$$A = -\frac{v_x}{\omega \sin \bar{\alpha}} = -\frac{\sqrt{\omega^2 v_x^2 + (\mathcal{E} - \omega^2)^2}}{\omega^2} < 0.$$

The time $s_1 > 0$ is such that $\rho(s_1) = 1$ and is given by $s_1 = (\pi - \bar{\alpha})/\omega$: if $y(s_1) \neq 0$ (namely, $\alpha \neq 0$), the polar angle θ_1 of the point p_1 is given by

$$\theta_1 = \begin{cases} \operatorname{arccot}\left(\dfrac{x(s_1)}{y(s_1)}\right) & \text{if } \alpha > 0, \\[4mm] \operatorname{arccot}\left(\dfrac{x(s_1)}{y(s_1)}\right) - \pi & \text{if } \alpha < 0, \end{cases} \qquad (2.3.8)$$

where we took into account that, for $\alpha < 0$, $\bar{\theta}_E \in [\pi, 2\pi]$, then one has to take the second determination of arccot.
Direct computations of the homothetic solution (corresponding to $\alpha = 0$) and equation (2.3.7), along with the definition of θ_E, lead finally to

$$\theta_E(\alpha) = \begin{cases} \theta_E^+(\alpha) = \operatorname{arccot}\left(\dfrac{\omega^2}{(2\mathcal{E} - \omega^2)\sin(2\alpha)} + \cot(2\alpha)\right) & \text{if } \alpha > 0, \\[4mm] 0 & \text{if } \alpha = 0, \\[4mm] \theta_E^-(\alpha) = \operatorname{arccot}\left(\dfrac{\omega^2}{(2\mathcal{E} - \omega^2)\sin(2\alpha)} + \cot(2\alpha)\right) - \pi & \text{if } \alpha < 0. \end{cases} \qquad (2.3.9)$$

If $\mathcal{E} > \omega^2$, the function $\theta_E(\alpha)$ is of class C^1 in $(-\pi/2, \pi/2)$ and assumes all the values in $(-\pi, \pi)$. Moreover,

$$\frac{d\theta_E}{d\alpha}(\alpha) = \frac{(2\mathcal{E} - \omega^2)(2\mathcal{E} - \omega^2 + \omega^2 \cos(2\alpha))}{2\mathcal{E}(\mathcal{E} - 2\omega^2) - (2\mathcal{E} - \omega^2)\omega^2 \cos(2\alpha)} > 0 \text{ for all } \alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right). \qquad (2.3.10)$$

From the inverse function theorem, there exist a unique function $\alpha : (-\pi, \pi) \to (-\pi/2, \pi/2)$, $\theta_1 \mapsto \alpha(\theta_1)$ such that for every θ_1 we have

$$p_1 = e^{i\theta_1} = z\left(s_1(\alpha(\theta_1)); \bar{p}_0, \sqrt{2\mathcal{E} - \omega^2} e^{i\alpha(\theta_1)}\right). \qquad (2.3.11)$$

Moreover, $\alpha(\theta_1) \in C^1(-\pi, \pi)$.

Fixing now $p_0, p_1 \in D_0$ such that $|p_0 - p_1| < 2$, we have that $|\theta_0 - \theta_1| < \pi$, then problem

$$
\begin{cases}
z''(s) = -\omega^2 z(s), \\
z(0) = p_0, \, z'(0) = v_0 = \sqrt{2\mathcal{E} - \dfrac{\omega^2}{2}} e^{i(\theta_0 + \alpha(\theta_1 - \theta_0))},
\end{cases}
\tag{2.3.12}
$$

admits the unique solution $z(s; p_0, p_1)$. If we define $T = s_1$ as above, we have that $z(T; p_0, p_1) = p_1$ and $\|z(s)\| > 1$ for every $s \in (0, T)$, while the energy conservation law is ensured by the choice of v_0. Moreover, by the differentiable dependence on the initial conditions of the Cauchy problem and the fact that $\alpha(\theta_1)$ is of class C^1, one can conclude that $z(s; p_0, p_1)$ is differentiable as a function of its endpoints.

$$\square$$

As for the inner dynamics, we can state the below theorem, which refers to the *(TnT)* topological characterization introduced in Definition 2.2.2.

Theorem 2.3.2. *For every $p_0, p_1 \in \partial D_0$, $\|p_0 - p_1\| < 2$, there is a unique $T > 0$ and a unique solution $z(s; p_0, p_1)$ of*

$$
\begin{cases}
(HS_I)[z(s)] & s \in [0, T] \\
\|z(s)\| < 1 & s \in (0, T) \\
z(0) = p_0, \; z(T) = p_1
\end{cases}
\tag{2.3.13}
$$

such that $z(s; p_0, p_1)$ is of class C^1 with respect to p_0 and p_1 and:

- *if $p_0 = p_1$, $z(s; p_0, p_0)$ is an ejection-collision solution;*

- *if $p_0 \neq p_1$, $z(s; p_0, p_1)$ is a classical solution of (2.3.13) which satisfies the* (TnT) *characterization. If $p_1 \to p_0$, $z(s; p_0, p_1)$ tends to the ejection-collision solution $z(s; p_0, p_0)$.*

Moreover, there is $0 < C < 1$ such that, for every p_0, p_1 as above,

$$
- p_0 \cdot \frac{z_1'(0; p_0, p_1)}{\|z_1'(0; p_0, p_1)\|} > C \quad and \quad p_1 \cdot \frac{z_1'(T; p_0, p_1)}{\|z_1'(T; p_0, p_1)\|} > C.
\tag{2.3.14}
$$

The quantity C depends on the physical parameters $\mathcal{E} + h, \mu$ and on $\|p_0 - p_1\|$. In particular, it tends to 1 when $\mathcal{E} \to \infty$ or $\|p_0 - p_1\| \to 0$

Proof. If $p_1 \neq p_2$, the first part of the proof is a straightforward consequence of Theorem 2.2.3: as a matter of fact, $\|p_0 - p_1\| < 2$ is equivalent to require that condition (NC) is satisfied. Moreover, by the monotonicity properties of the function $\|z(s)\|$ when z is a Keplerian hyperbola, one can infer also that $\|z(s)\| < 1$ for every $s \in (0, T)$. When instead $p_1 = p_2$, one can refer to Remark 2.2.7 to gain the proof.

As for the inequalities in (2.3.14), for $\mathcal{E} + h$ large enough they are a natural consequence of the asymptotic result in Theorem 2.2.3. On the other hand, using again the peculiar properties of a circular domain, we can extend this result to any positive value of the

energy, providing more precise estimates. First of all, let us observe that, whenever $p_1 = p_2$, the claims are trivially true, as the ejection-collision solution is always parallel to the radial direction. Let us then assume that $p_1 \neq p_2$, and, as the system is invariant under rotation, suppose $p_0 = e^{-i\beta}$ and $p_1 = e^{i\beta}$, $\beta \in (0, \pi/2)$ (note that, if $\beta = \pi/2$, then p_1 and p_2 are antipodal). Let us call $x_0 \in (0,1)$ the $x-$coordinate of both the points, to have $p_1 = \left(x_0, -\sqrt{1 - x_0^2}\right)$ and $p_2 = \left(x_0, \sqrt{1 - x_0^2}\right)$. From direct computations, it results that the unique *(TnT)* arc which solves the inner fixed-ends problem with endpoints p_1 and p_2 is parametrized by the equation

$$(x - ae)^2(e^2 - 1) - y^2 = a^2(e^2 - 1), \quad x \le a(e - 1)$$

$$a = \frac{\mu}{2(\mathcal{E} + h)}, \quad e = \frac{x_0 + \sqrt{4a^2 + 4a + x_0^2}}{2a},$$

whose positive branch can be parametrized as $y(x) = \sqrt{e^2 - 1}\sqrt{(x + a)^2 - a^2}$. Writing $p_1 = p_1(x_0) = (x_0, y(x_0))$ and $v_1 = (1, \partial_x y(x_0))$, we can express the cosine of the angle between p_1 and v_1 as a function

$$c(x_0) \equiv p_1 \cdot \frac{v_1}{|v_1|} = e\sqrt{\frac{2a + x_0(x_0 + \sqrt{4a^2 + 4a + x_0^2})}{2 + 4a}}, \tag{2.3.15}$$

which is strictly increasing for $x_0 \in [0, 1]$ and such that $c(0) = \sqrt{\dfrac{1 + a}{1 + 2a}} \doteq C < 1$ and $c(1) = 1$: this prove the claim for p_1 and $s = T$. The same estimate holds for $-p_0$ and $s = 0$, taking into account that for $s = 0$ the hyperbola points inward the domain D_0. $\qquad\square$

Remark 2.3.3. *The value of C depends on the physical parameters of the problem: in particular, with reference to the proof of Theorem 2.3.2, one has*

$$c(0) = \sqrt{\frac{1 + a}{1 + 2a}} = \sqrt{\frac{\mathcal{E} + h + \mu/2}{\mathcal{E} + h + \mu}},$$

which tends to 1 when $\mathcal{E} + h \to \infty$. As a consequence, one can control the transversality of $z_1(s; p_0, p_1)$ by acting on the value of the total inner energy. This fact is of particular importance in view of Remark 1.4.2, since for the first return $\mathcal{F}_0$ to be well defined one needs that the angle $\beta_1 = \angle(p_1, v_1)$ is such that

$$|\sin \beta_1| \le \sqrt{\frac{V_E(p_1)}{V_I(p_1)}} = \sqrt{\frac{\mathcal{E} - \omega^2/2}{\mathcal{E} + h + \mu}}. \tag{2.3.16}$$

If $\mathcal{E} + h$ is such that $\sqrt{1 - C^2} < \sqrt{(\mathcal{E} - \omega^2/2)/(\mathcal{E} + h + \mu)}$, and this is true if $\sqrt{\mu/2} < \sqrt{\mathcal{E} - \omega^2/2}$, Eq.(2.3.16) is satisfied.
Moreover, since for $p_0 \to p_1$ the inner arc tends to the ejection-collision solution, the lower bound C can be controlled also by choosing the endpoint to be close enough.

The estimates given by (2.3.14) are crucial to ensure that the inner arcs are transversal to ∂D_0 as much as needed: this is necessary for the first return map to be

well defined, since, according to Remark 1.4.2, it is clear that these arcs can not be tangent to the interface.

2.3.2 Study of the map $\mathcal{F}_0$

Once the good definition and differentiability of the distances $d_I(p_0, p_1)$ and $d_E(p_0, p_1)$ is ensured, one can eventually consider the generating function introduced in Section 2.2.1 and given by

$$S_0(\xi_0, \xi_1) = S_{E,0}(\xi_0, \tilde{\xi}) + S_{I,0}(\tilde{\xi}, \xi_1) = d_E(\gamma_0(\xi_0), \gamma_0(\tilde{\xi})) + d_I(\gamma_0(\tilde{\xi}), \gamma_0(\xi_1)),$$

where $\gamma_0 : \mathbb{R}_{/2\pi\mathbb{Z}} \to \mathbb{R}^2$ denotes the circle of radius 1, and investigate the associated nondegeneracy conditions (2.2.3) and (2.2.7). Although the complete analysis on its good definition will be done after the derivation of the explicit formulation of the associated first return map $\mathcal{F}_0$ in Section 2.3.3, the central symmetry of the circular case allows to give some preliminary informations. As both the outer and inner systems are invariant under rotations, the associated generating functions can be expressed as univariate functions depending on the angle spanned by the arc; more precisely,

$$S_0(\xi_0, \xi_1) = \tilde{S}_0(\xi_1 - \xi_0) = \tilde{S}_{E,0}(\tilde{\xi} - \xi_0) + \tilde{S}_{I,0}(\xi_1 - \tilde{\xi}).$$

Going through the same analysis described in general in Section 2.2.1, the intermediate coordinate $\tilde{\xi}$ is implicitly determined as a function of ξ_0 and ξ_1 by the relation

$$\partial_{\tilde{\xi}}(\tilde{S}_{E,0}(\tilde{\xi} - \xi_0) + \tilde{S}_{I,0}(\xi - \tilde{\xi})) = 0$$

if (2.2.3) is verified. In the circular case, the latter translates in

$$\tilde{S}''_{E,0}(\tilde{\xi} - \xi_0) + \tilde{S}''_{I,0}(\xi_1 - \tilde{\xi}) \neq 0. \tag{2.3.17}$$

If (2.3.17) holds, one has

$$\partial_{\xi_0}\tilde{\xi} = \frac{\tilde{S}''_{E,0}(\tilde{\xi} - \xi_0)}{\tilde{S}''_{E,0}(\tilde{\xi} - \xi_0) + \tilde{S}''_{I,0}(\xi_1 - \tilde{\xi})}, \quad \partial_{\xi_1}\tilde{\xi} = \frac{\tilde{S}''_{I,0}(\xi_1 - \tilde{\xi})}{\tilde{S}''_{E,0}(\tilde{\xi} - \xi_0) + \tilde{S}''_{I,0}(\xi_1 - \tilde{\xi})},$$

and the canonical actions are defined by

$$\begin{aligned} I_0 &= -\partial_{\xi_0}\tilde{S}_0(\xi_1 - \xi_0) = \tilde{S}'_{E,0}(\tilde{\xi}(\xi_0, \xi_1) - \xi_0), \\ I_1 &= \partial_{\xi_1}\tilde{S}_0(\xi_1 - \xi_0) = \tilde{S}'_{I,0}(\xi_1 - \tilde{\xi}(\xi_0, \xi_1)). \end{aligned} \tag{2.3.18}$$

The first equation in (2.3.18) defines implicitly $\xi_1 = \xi_1(\xi_0, I_0)$, and, as a consequence, the map $\mathcal{F}_0$, if (2.2.7) holds, that is, if

$$\partial_{\xi_1}(I_0 - \tilde{S}'_{E,0}(\tilde{\xi}(\xi_0, \xi_1) - \xi_0)) = -\frac{\tilde{S}''_{E,0}(\tilde{\xi}(\xi_0, \xi_1) - \xi_0)\tilde{S}''_{I,0}(\xi_1 - \tilde{\xi}(\xi_0, \xi_1))}{\tilde{S}''_{E,0}(\tilde{\xi}(\xi_0, \xi_1) - \xi_0) + \tilde{S}''_{I,0}(\xi_1 - \tilde{\xi}(\xi_0, \xi_1))} \neq 0. \tag{2.3.19}$$

As a final remark, note that the validity of conditions (2.3.17) and (2.3.19) are strongly related to the *twist condition* (see [77, 25, 23]) associated to the map $\mathcal{F}_0$, defined as $\partial_{I_0}\xi_1 \neq 0$. As a matter of fact, one has

$$\partial_{I_0}\xi_1 = \frac{\tilde{S}''_{E,0}(\tilde{\xi}(\xi_1 - \xi_0) - \xi_0) + \tilde{S}''_{I,0}(\xi_1 - \tilde{\xi}(\xi_1 - \xi_0))}{\tilde{S}''_{E,0}(\tilde{\xi}(\xi_0,\xi_1) - \xi_0)\tilde{S}''_{I,0}(\xi_1 - \tilde{\xi}(\xi_1 - \xi_0))}$$

and one can then say that, if $\mathcal{F}_0$ is given, the twist condition is equivalent to require the nondegenerations (2.3.17) and (2.3.19) to be true.

2.3.3 Explicit formulation of $\mathcal{F}_0$

When the domain D is circular, the first return map $\mathcal{F}_0 : (\xi_0, \alpha_0) \mapsto (\xi_1, \alpha_1)$ can be explicitly determined: in this case, the nondegeneracy given through (2.3.17) and (2.3.19) can be investigated in the equivalent form given by the twist condition. The boundary ∂D_0 can be parametrized as $\gamma(\xi) = (\cos\xi, \sin\xi)$, with $\xi \in \mathbb{R}_{/2\pi\mathbb{Z}}$, and the symmetry properties of the potentials V_E and V_I and the isotropy of the Snell's law on a circular domain imply that $\mathcal{F}_0$ is of the form

$$\mathcal{F}_0(\xi_0, \alpha_0) = (\xi_1(\xi_0, \alpha_0), \alpha_1(\xi_0, \alpha_0)) = (\xi_0 + \bar{\theta}(\alpha_0), \alpha_0); \tag{2.3.20}$$

in other words, the first return map on the circle reduces to a conservation of the velocity variable α and a shift in the angle ξ of a suitable quantity $\bar{\theta}$ which depends only on the physical parameters of the problem and on α_0. The Jacobian matrix $D\mathcal{F}_0((\xi_0, \alpha_0))$ can be then expressed for every pair $(\xi_0, \alpha_0) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times (-\pi/2, \pi/2)$ as

$$D\mathcal{F}_0(\xi_0, \alpha_0) = \begin{pmatrix} 1 & \dfrac{\partial\bar{\theta}}{\partial\alpha_0}(\alpha_0) \\ 0 & 1 \end{pmatrix}.$$

From the above considerations, we have that $\bar{\theta}(\alpha_0) = \bar{\theta}_E(\alpha_0) + \bar{\theta}_I(\alpha_0)$, where $\bar{\theta}_E(\alpha_0)$ and $\bar{\theta}_I(\alpha_0)$ represent the excursions in the angles due respectively to the outer and the inner arcs of the orbit $z_{EI}(s)$.

Outer shift The outer shift has been already computed as an additional result in the proof of Theorem 2.3.1, and is equal to

$$\theta_E(\alpha) = \begin{cases} \theta_E^+(\alpha) = \operatorname{arccot}\left(\dfrac{\omega^2}{(2\mathcal{E} - \omega^2)\sin(2\alpha)} + \cot(2\alpha)\right) & \text{if } \alpha > 0, \\ 0 & \text{if } \alpha = 0, \\ \theta_E^-(\alpha) = \operatorname{arccot}\left(\dfrac{\omega^2}{(2\mathcal{E} - \omega^2)\sin(2\alpha)} + \cot(2\alpha)\right) - \pi & \text{if } \alpha < 0. \end{cases} \tag{2.3.21}$$

Inner shift   As the system is invariant under rotations, without loss of generality let $p_0 = (1, 0)$ be the initial point of the inner orbit and, denoted by v_0 its initial velocity, let $\beta_0 \in (-\pi/2, \pi/2)$ the angle between v_0 and the inward-pointing radial unit vector, namely, $-p_0$; then $v_0 = \sqrt{2(\mathcal{E} + h) + 2\mu}(-\cos\beta_0, \sin\beta_0) = (v_x, v_y)$. The inner Cauchy problem is then given by

$$\begin{cases} z''(s) = -\dfrac{\mu}{\|z(s)\|^3} z(s), & s \in [0, T] \\ z(0) = p_0, z'(0) = v_0. \end{cases} \tag{2.3.22}$$

Unlike the outer case, for the inner Keplerian orbit it is not possible to decouple the $2-$dimensional system into two one-dimensional systems in the variables (x, y); we shall rely on other classical techniques (see [70]) which require the passage in polar coordinates. Consider the functions $r(s) \in \mathbb{R}^+$ and $\theta(s) \in \mathbb{R}_{/2\pi\mathbb{Z}}$ such that $z(s) = r(s)e^{i\theta(s)}$. From the conservation of the angular momentum, we have that

$$r(s)^2 \theta'(s) = const = k = \|p_0\| \|v_0\| \sin\beta_0 = \sqrt{2\mathcal{E} + 2h + 2\mu} \sin\beta_0, \tag{2.3.23}$$

while the energy conservation law implies

$$\mathcal{E} + h = \frac{1}{2}\left(r'(s)^2 + r(s)^2\theta'(s)^2\right) - \frac{\mu}{r(s)} \Rightarrow r'(s) = -\sqrt{2(\mathcal{E} + h) - \frac{k^2}{r^2(s)} + \frac{2\mu}{r(s)}}, \tag{2.3.24}$$

where the sign depend by the fact that, according to the chosen initial conditions, $r(s)$ is decreasing. Taking together (2.3.23) and (2.3.24), one has then

$$d\theta = -\frac{k}{r^2\sqrt{2(\mathcal{E} + h) - \frac{k^2}{r^2} + \frac{2\mu}{r}}} dr. \tag{2.3.25}$$

The classical results for the two-body problem ensure that, for positive energies, the Kepler problem is unbounded, and $r(t)$ reaches its unique minimum r_p at a time $s_p > 0$. The value of r_p is given by (see [70])

$$r_p = \frac{k^2}{\mu}\left(1 + \sqrt{1 + \frac{2(\mathcal{E} + h)k^2}{\mu^2}}\right)^{-1}. \tag{2.3.26}$$

If we denote with θ_p the polar angle of the pericenter and consider the initial conditions given by (2.3.22), taking into account the symmetry of $r(s)$ with respect to s_p, we have that the inner shift angle is given by

$$\bar{\theta}_I = 2\theta_p, \tag{2.3.27}$$

where θ_p can be obtained by integration from (2.3.25):

$$\theta_p = \int_0^{\theta_p} d\theta = -k \int_1^{r_p} \frac{dr}{r^2 \sqrt{2(\mathcal{E} + h) - \frac{k^2}{r^2} + \frac{2\mu}{r}}} = \frac{k}{|k|} \int_1^{\frac{1}{r_p}} \frac{du}{\sqrt{\frac{2(\mathcal{E}+h)}{k^2} - u^2 + \frac{2\mu}{k^2}u}}$$

setting $x = u - \mu/k^2$, $x_0 = 1 - \mu/k^2$ and $x_1 = \frac{\mu}{k^2}\sqrt{1 + \frac{2(\mathcal{E} + h)k^2}{\mu^2}}$:

$$\theta_p = \frac{k}{|k|} \int_{x_0}^{x_1} \frac{dx}{\sqrt{\frac{2(\mathcal{E}+h)}{k^2} + \frac{\mu^2}{k^4} - x^2}}.$$

Finally, defining $y = x \left(\frac{2(\mathcal{E} + h)}{k^2} + \frac{\mu^2}{k^4}\right)^{-1/2}$, $y_0 = (k^2 - \mu)\left(2(\mathcal{E} + h)k^2 + \mu^2\right)^{-1/2}$ and $y_1 = 1$:

$$\theta_p = \frac{k}{|k|} \int_{y_0}^{y_1} \frac{dy}{\sqrt{1 - y^2}} = \frac{k}{|k|} \arccos y_0 = \frac{k}{|k|} \arccos\left(\frac{k^2 - \mu}{\sqrt{2(\mathcal{E} + h)k^2 + \mu^2}}\right). \qquad (2.3.28)$$

Casting together (2.3.23), (2.3.27) and (2.3.28), one obtains

$$\bar{\theta}_I(\beta_0) = \begin{cases} \bar{\theta}_I^+(\beta_0) = 2\arccos\left(\frac{(2\mathcal{E} + 2h + 2\mu)\sin{\beta_0}^2 - \mu}{\sqrt{4(\mathcal{E} + h)(\mathcal{E} + h + \mu)\sin{\beta_0}^2 + \mu^2}}\right) - 2\pi & \text{if } \beta_0 \geq 0 \\[4mm] \bar{\theta}_I^-(\beta_0) = -2\arccos\left(\frac{(2\mathcal{E} + 2h + 2\mu)\sin{\beta_0}^2 - \mu}{\sqrt{4(\mathcal{E} + h)(\mathcal{E} + h + \mu)\sin{\beta_0}^2 + \mu^2}}\right) + 2\pi & \text{if } \beta_0 < 0, \end{cases}$$

(2.3.29)

where the shift is such that $\bar{\theta}_I(\beta_0) \in (-\pi, \pi)$ for every $\beta_0 \in (-\pi/2, \pi/2)$. Note that

$$\lim_{\beta_0 \to 0^+} \bar{\theta}_I^+(\beta_0) = 0 = \lim_{\beta_0 \to 0^-} \bar{\theta}_I^-(\beta_0),$$

$$\lim_{\beta_0 \to 0^+} \frac{d}{d\beta_0}\bar{\theta}_I^+(\beta_0) = -\frac{4(\mathcal{E} + h + \mu)}{\mu} = \lim_{\beta_0 \to 0^-} \frac{d}{d\beta_0}\bar{\theta}_I^-(\beta_0)$$

and then $\bar{\theta}_I \in C^1(-\pi/2, \pi/2)$.

Total shift and properties of the overall trajectories The total shift angle $\bar{\theta}(\alpha_0)$ is computed by taking the sum of the outer and the inner shifts and taking into account the transition laws for the velocities across the interface ∂D_0. In particular, if $\tilde{\alpha}$ and $\tilde{\beta}$ denote respectively the angles wit the normal unit vector of the outer and the inner velocities of an orbit crossing the interface in a point $\tilde{p} \in \partial D_0$, from Snell's law one has

$$\sqrt{\mathcal{E} - \frac{\omega^2}{2}\|\tilde{p}\|^2} \sin \tilde{\alpha} = \sqrt{\mathcal{E} + h + \frac{\mu}{\|\tilde{p}\|}} \sin \tilde{\beta};$$

in the particular case of a circular domain, the Snell's law is uniform over all the points of ∂D_0, and the initial and final angles with the radial direction are equal for every branch of the orbit. Performing in (2.3.29) the substitution $\sin \beta_0 = \sqrt{(2\mathcal{E} - \omega^2)/(2(\mathcal{E} + h + \mu))} \sin \alpha_0$, one obtains the total shift

$$\bar{\theta}(\alpha_0) = \begin{cases} \bar{\theta}_E^+(\alpha_0) + \bar{\theta}_I^+(\alpha_0) & \text{if } \alpha_0 > 0 \\ 0 & \text{if } \alpha_0 = 0 \\ \bar{\theta}_E^-(\alpha_0) + \bar{\theta}_I^-(\alpha_0) & \text{if } \alpha_0 < 0, \end{cases}$$

where $\bar{\theta}_E^+(\alpha_0)$ and $\bar{\theta}_E^-(\alpha_0)$ are given by (2.3.21) and

$$\bar{\theta}_I^+(\alpha_0) = 2 \arccos \left(\frac{(2\mathcal{E} - \omega^2)\sin {\alpha_0}^2 - \mu}{\sqrt{2(\mathcal{E} + h)(2\mathcal{E} - \omega^2)\sin {\alpha_0}^2 + \mu^2}} \right) = -\bar{\theta}_I^-(\alpha_0).$$

The map is continuous and differentiable with respect to α_0, and

$$\frac{d}{d\alpha_0}\bar{\theta}(0) = \lim_{\alpha_0 \to 0^+} \frac{d}{d\alpha_0}\bar{\theta}(\alpha_0) = \lim_{\alpha_0 \to 0^-} \frac{d}{d\alpha_0}\bar{\theta}(\alpha_0) =$$
$$= \frac{2\mathcal{E} - \omega^2}{\mathcal{E}} - \frac{2\sqrt{2}\sqrt{\mathcal{E} + h + \mu}\sqrt{2\mathcal{E} - \omega^2}}{\mu}.$$

Passing to the canonical coordinates (ξ, I), the axisymmetry of the potentials and the isotropy of Snell's law on the circle translates in the conservation of the quantity I both in the endpoints and the transition point $\tilde{\xi}$. The first claim is a straightforward consequence of Eq.(2.3.20), while to prove the conservation of the action across the intermediate point one needs to consider the actions I^E and I^I associated to $S_{E,0}$ and $S_{I,0}$ separately:

$$\begin{aligned} I_1^E(\xi_0, \xi_1) &= \partial_b S_{E,0}(\xi_0, \tilde{\xi}) = \sqrt{V_E(\gamma(\tilde{\xi}))} \sin \beta_0 = \sqrt{V_I(\gamma(\tilde{\xi}))} \sin \beta_1 \\ I_0^I(\xi_0, \xi_1) &= -\partial_a S_{I,0}(\tilde{\xi}, \xi_1) = \sqrt{V_I(\gamma(\tilde{\xi}))} \sin \beta_1. \end{aligned} \qquad (2.3.30)$$

Since on the circle $\alpha_0 = \beta_0$ and $\beta_1 = \alpha_1'$, we have that for every $\xi_0, \xi_1 \in \mathbb{R}_{/2\pi\mathbb{Z}}$

$$I_0(\xi_0, \xi_1) = I_1^E(\xi_0, \xi_1) = I_0^I(\xi_0, \xi_1) = I_1(\xi_0, \xi_1) \equiv I(\xi_0, \xi_1). \qquad (2.3.31)$$

Moreover, from (2.2.9) one has that in the circular case the global domain of definition of the actions does not depend on the points ξ_0, ξ_1, that is

$$I_0, I_1 \in \left(-\sqrt{\mathcal{E} - \frac{\omega^2}{2}}, \sqrt{\mathcal{E} - \frac{\omega^2}{2}} \right) = (-I_c, I_c) = \mathcal{I}.$$

Taking into account Eq.(2.3.20), the definitions of $\bar{\theta}_E$ and $\bar{\theta}_I$ and the relations (2.2.9), (2.3.31), in the new set of canonical coordinates $(\xi, I) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I}$ we can express the

first return map as

$$\mathcal{F}_0 : \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I} \to \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I},$$
$$(\xi_0, I_0) \mapsto (\xi_1, I_1) = (\xi_0 + \bar{\theta}(I_0), I_0) = (\xi_0 + f(I_0) + g(I_0), I_0), \tag{2.3.32}$$

where

$$f(I) = \begin{cases} \operatorname{arccot}\left(\dfrac{\mathcal{E} - 2I^2}{I\sqrt{4\mathcal{E} - 2(2I^2 + \omega^2)}}\right) & \text{if } I \in (0, I_c) \\[2ex] 0 & \text{if } I = 0 \\[2ex] \operatorname{arccot}\left(\dfrac{\mathcal{E} - 2I^2}{I\sqrt{4\mathcal{E} - 2(2I^2 + \omega^2)}}\right) - \pi & \text{if } I \in (-I_c, 0) \end{cases}$$

and

$$g(I) = \begin{cases} 2\arccos\left(\dfrac{2I^2 - \mu}{\sqrt{4(\mathcal{E} + h)I^2 + \mu^2}}\right) - 2\pi & \text{if } I \in (0, I_c) \\[2ex] 0 & \text{if } I = 0 \\[2ex] -2\arccos\left(\dfrac{2I^2 - \mu}{\sqrt{4(\mathcal{E} + h)I^2 + \mu^2}}\right) + 2\pi & \text{if } I \in (-I_c, 0) \end{cases}$$

are C^1 functions in $\mathcal{I}$.

Remark 2.3.4. *Direct computations show that for every $\mathcal{E} > \omega^2, h > 0, \mu > 0$ and for every $I \in \mathcal{I}$ one has $f'(I) > 0$ and $g'(I) < 0$: the outer and inner shifts are then invertible in $\mathcal{I}$, and one can define the inverse functions $\tilde{f}(\theta) = f^{-1}(I)_{|I=I(\theta)}$ and $\tilde{g}(\theta) = g^{-1}(I)_{|I=I(\theta)}$. From the regularity of both f and g, we have that $\tilde{f}$ and $\tilde{g}$ are of class C^1 in the respective domains. In particular,*

$$f\,(\mathcal{I}) = (-\pi, \pi), \quad g\,(\mathcal{I}) = (-\bar{\theta}, \bar{\theta}),$$

$$\bar{\theta} = 2\pi - 2\arccos\left(\frac{2\mathcal{E} - \omega^2 - \mu}{\sqrt{2(\mathcal{E} + h)(2\mathcal{E} - \omega^2) + \mu^2}}\right). \tag{2.3.33}$$

Moreover,

$$\tilde{\xi} = \xi_0 + f(I) \Leftrightarrow I = \tilde{f}(\tilde{\xi} - \xi_0) \equiv I_1^E(\xi_0, \tilde{\xi}),$$
$$\xi_1 = \tilde{\xi} + g(I) \Leftrightarrow I = \tilde{g}(\xi_1 - \tilde{\xi}) \equiv I_0^I(\tilde{\xi}, \xi_1). \tag{2.3.34}$$

Lemma 2.3.5. *For every $I \in \mathcal{I}$, except for a finite number if points, $f'(I) + g'(I) \neq 0$.*

Proof. Direct computations lead to

$$f'(I) + g'(I) = \frac{\sqrt{2}(2\mathcal{E}^2 - (\mathcal{E} + 2I^2)\omega^2)}{\sqrt{2\mathcal{E} - \omega^2 - 2I^2}(\mathcal{E}^2 - 2\omega^2 I^2)} - \frac{8(\mathcal{E} + h)I^2 + 4(\mathcal{E} + h)\mu + 4\mu^2}{\sqrt{\mathcal{E} + h + \mu - I^2}(4(\mathcal{E} + h)I^2 + \mu^2)}$$
$$= \frac{A(I^2)}{B(I^2)} - \frac{C(I^2)}{D(I^2)}. \tag{2.3.35}$$

Since for every $I \in \mathcal{I}$ we have that $A(I^2), B(I^2), C(I^2), D(I^2) > 0$,

$$f'(I) + g'(I) = 0 \Leftrightarrow X = I^2 \in \left[0, I_c^2\right) \text{ is a solution of } p(x) = 0,$$

where $p(x) = A^2(x)D^2(x) - B^2(x)C^2(x)$. As $p(x)$ is a real polynomial of degree 5 in X, one can have at most ten values of $I \in \mathcal{I}$ such that $f'(I) + g'(I) = 0$. $\qquad\square$

We define $\bar{\mathcal{I}} = \{I \in \mathcal{I} \mid f'(I) + g'(I) = 0\}$ as the set of the critical points of the function $f + g$.

Proposition 2.3.6. *The generating function $S_0(\xi_0, \xi_1)$ is well defined in $\mathbb{R}_{/2\pi\mathbb{Z}} \times \mathbb{R}_{/2\pi\mathbb{Z}}$ except for a finite number of pairs (ξ_0, ξ_1) in the quotient space $(\mathbb{R}_{/2\pi\mathbb{Z}} \times \mathbb{R}_{/2\pi\mathbb{Z}})/\sim$, where $(\xi_0, \xi_1) \sim (\xi_0', \xi_1') \Leftrightarrow \xi_1 - \xi_0 = \xi_1' - \xi_0'$.*

Proof. For $S_0(\xi_0, \xi_1) = S_{E,0}(\xi_0, \tilde{\xi}) + S_{I,0}(\tilde{\xi}, \xi_1)$ to be well defined, one needs to verify condition (2.2.3). From the definition of the actions

$$\partial_{\tilde{\xi}}(\partial_b S_{E,0}(\xi_0, \tilde{\xi}) + \partial_a S_{I,0}(\tilde{\xi}, \xi_1)) = \partial_{\tilde{\xi}}(I_1^E(\xi_0, \tilde{\xi}) - I_0^I(\tilde{\xi}, \xi_1)) = \partial_{\tilde{\xi}} I_1^E(\xi_0, \tilde{\xi}) - \partial_{\tilde{\xi}} I_0^I(\tilde{\xi}, \xi_1)$$

$$= \partial_{\tilde{\xi}} \tilde{f}(\tilde{\xi} - \xi_0) - \partial_{\tilde{\xi}} \tilde{g}(\xi_1 - \tilde{\xi}) = \left(\frac{1}{f'(I)} + \frac{1}{g'(I)}\right)_{I=I(\xi_0, \xi_1)}$$

$$= \left(\frac{f'(I) + g'(I)}{f'(I)g'(I)}\right)_{I=I(\xi_0, \xi_1)}$$

which is zero if and only if $\xi_1 - \xi_0 \in (f+g)(\bar{\mathcal{I}})$. $\qquad\square$

Proposition 2.3.7. *For every $(\xi_0, I_0) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times \left(\mathcal{I}\setminus\bar{\mathcal{I}}\right)$ the first return map $\mathcal{F}_0$*

1. *is area-preserving;*

2. *satisfies the twist condition*
$$\frac{\partial \xi_1}{\partial I_0}(\xi_0, I_0) \neq 0.$$

Proof. The area-preserving property of $\mathcal{F}_0$ is a direct consequence of the variational formulation of the problem: when $\tilde{\xi}$ is well defined, we have (expressing $\xi_1 = \xi_1(\xi_0, I_0)$)

$$\partial_{\xi_0}\xi_1 = -\frac{\partial_a^2 S_0(\xi_0, \xi_1)}{\partial_{ab} S_0(\xi_0, \xi_1)}, \qquad\qquad \partial_{I_0}\xi_1 = -\frac{1}{\partial_{ab} S_0(\xi_0, \xi_1)}$$

$$\partial_{\xi_0} I_1 = \partial_{ab} S_0(\xi_0, \xi_1) - \frac{\partial_b^2 S_0(\xi_0, \xi_1)\partial_a^2 S_0(\xi_0, \xi_1)}{\partial_{ab} S_0(\xi_0, \xi_1)}, \qquad \partial_{I_0} I_1 = -\frac{\partial_b^2 S_0(\xi_0, \xi_1)}{\partial_{ab} S_0(\xi_0, \xi_1)},$$

where, from (2.3.34),

$$\partial_{ab} S(\xi_0, \xi_1) = \partial_{\xi_0 \xi_1} S(\xi_0, \xi_1) = \partial_{ab} S_I(\tilde{\xi}, \xi_1)\partial_{\xi_0}\tilde{\xi} = -\frac{\tilde{g}'(\xi_1 - \tilde{\xi})\tilde{f}'(\tilde{\xi} - \xi_0)}{\tilde{f}'(\tilde{\xi} - \xi_0) + \tilde{g}'(\xi_1 - \tilde{\xi})}$$

is well defined and different from zero for every $(\xi_0, I_0) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I}\setminus\bar{\mathcal{I}}$. Whenever $\mathcal{F}_0$ is well defined, the determinant of its Jacobian matrix is

$$\det\left(D_{(\xi_0, I_0)}\mathcal{F}_0\right) = \partial_{\xi_0}\xi_1\partial_{I_0}I_1 - \partial_{I_0}\xi_1\partial_{\xi_0}I_1 = 1,$$

thus $\mathcal{F}_0$ is area-preserving.

As for the twist condition, we have $\partial_{I_0}\xi_1 = f'(I_0) + g'(I_0)$, which is nonzero whenever $I_0 \notin \bar{\mathcal{I}}$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Summarizing the previous results, we can then conclude that the set $\mathcal{I}\backslash\bar{\mathcal{I}}$ is the finite union of open intervals (at most eleven, but possibly the whole $(-I_c, I_c)$ if $\bar{\mathcal{I}} = \emptyset^2$), in which $\mathcal{F}_0$ is well defined, area-preserving and satisfies the twist condition with constant sign.

Remark 2.3.8. *Locally around $\pm I_c$ and 0 the sign of $\partial_{I_0}\xi_1$ can be easily determined as a function of the physical parameters $\mathcal{E}, \omega, h, \mu$: as a matter of fact, one has*

$$\lim_{I\to I_c^-} \partial_{I_0}\xi_1 = \lim_{I\to -I_c^+} \partial_{I_0}\xi_1 = +\infty$$

and

$$\partial_{I_0}\xi_1(\xi_0, 0) = \frac{2\sqrt{\mathcal{E} - \frac{\omega^2}{2}}}{\mathcal{E}} - \frac{4\sqrt{\mathcal{E} + h + \mu}}{\mu},$$

then, for every $\mathcal{E} > \omega^2, h > 0, \mu > 0$

- *$\exists \bar{I} \in (0, I_c)$ such that for every $I \in \mathcal{I}$ with $|I| > \bar{I}$ it results $\partial_{I_0}\xi_1 > 0$;*

- *if $\dfrac{2\sqrt{\mathcal{E} - \frac{\omega^2}{2}}}{\mathcal{E}} > \dfrac{4\sqrt{\mathcal{E} + h + \mu}}{\mu}$ (resp. $\dfrac{2\sqrt{\mathcal{E} - \frac{\omega^2}{2}}}{\mathcal{E}} < \dfrac{4\sqrt{\mathcal{E} + h + \mu}}{\mu}$), $\exists \bar{\bar{I}} \in (0, I_c)$ such that for every $I \in (-I_c, I_c)$ with $|I| < \bar{\bar{I}}$ one has $\partial_{I_0}\xi_1 > 0$ (resp. $\partial_{I_0}\xi_1 > 0$);*

- *additionally, if $\dfrac{2\sqrt{\mathcal{E} - \omega^2/2}}{\mathcal{E}} < \dfrac{4\sqrt{\mathcal{E} + h + \mu}}{\mu}$, the derivative $\partial_{I_0}\xi_1$ admits at least a change of sign, which corresponds to a change of twist for the map $\mathcal{F}_0$.*

2.3.4 Periodic solutions on the circle

Once the general properties of $\mathcal{F}_0$ on the circle are defined, we can pass to the study of its orbits. To this end, given $(\xi_0, I_0) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I}\backslash\tilde{\mathcal{I}}$, let us define the orbit of (ξ_0, I_0) as the sequence of the iterates $\{(\xi_k, I_k)\}_{k\in\mathbb{Z}} = \{\mathcal{F}_0^k(\xi_0, I_0)\}_{|k\in\mathbb{Z}}$.

Definition 2.3.9. *The rotation number[3] associated to (ξ_0, I_0) through $\mathcal{F}_0$ is given by*

$$\rho(\xi_0, I_0) = \lim_{k\to\infty} \frac{\xi_k - \xi_0}{k}. \qquad\qquad (2.3.36)$$

In the circular case, one can easily see that for every (ξ_0, I_0) for which $\mathcal{F}_0$ is well defined one has $\rho(\xi_0, I_0) = \bar{\theta}(I_0)$.

---

[2]Numerical investigations shows that this case is consistent, in the sense that there are values of the parameters $\mathcal{E}, h, \mu$ and ω such that the sign of $f' + g'$ is constant (for example $\mathcal{E} = 2.5, \omega = 2, \mu = 2$ and $h = 2$).

[3]With an abuse of notation, in Section 2.4.2 we will use the same definition to identify the rotation number of the *lift* of a map of the annulus $\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b]$, that is, its periodic extension to $\mathbb{R} \times [a, b]$.

As the action I_0 is preserved on the circle, we have that, taking into account the phase space $(\xi, I) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I}$, the straight lines $\mathbb{R}_{/2\pi\mathbb{Z}} \times \{I_0\}$ are invariant for the dynamics induced by $\mathcal{F}_0$. We can then distinguish between two types of orbits:

- if $\bar{\theta}(I_0)/2\pi = p/q \in \mathbb{Q}$, then

$$(\xi_q, I_q) = (\xi_0 + 2\pi p, I_0) \equiv_{2\pi} (\xi_0, I_0);$$

  in this case, we say that the point (ξ_0, I_0) and the associated orbit are *(p,q)-periodic*;

- if $\bar{\theta}(I_0)/2\pi \notin \mathbb{Q}$, then for all $\xi_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}$ the orbit with initial point (ξ_0, I_0) is dense in $\mathbb{R}_{/2\pi\mathbb{Z}} \times \{I_0\}$.

A particular class of fixed points for $\mathcal{F}_0$ is given by the ejection-collision solutions, which form an invariant line of periodic points of period one defined on $\mathbb{R}_{/2\pi\mathbb{Z}} \times \{0\}$. Taking advantage of the continuity of the function $f + g$ on $\mathcal{I}$, one can state the following existence result.

Proposition 2.3.10. *Given $C = \bar{\theta} - \pi$, where $\bar{\theta}$, as in (2.3.33), depends only on the physical parameters $\mathcal{E}, h, \mu, \omega$, for every $\rho \in (-C, C)$ there are two values $I^{\pm} \in (-I_c, I_c)$ of the actions such that, for every $\xi_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}, \rho(\xi_0, I_0^{\pm}) = \rho$.*
In particular, for every $p, q \in \mathbb{Z}$ such that $-C < 2\pi p/q < C$, there are $I_{\pm}^{(p,q)} \in (-I_c, I_c)$ such that for every $\xi_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}$ the points $(\xi_0, I_+^{(p,q)})$ and $(\xi_0, I_-^{(p,q)})$ are (p,q)-periodic.

In the circular case, the existence of two orbits of all the rotation numbers is a simple consequence of the continuity of the total shift function. As it will be analysed in Section 2.4, a deformation of the boundary ∂D_0 breaks the symmetry of the system: in general, the first return map will be not integrable anymore and more sophisticated tools should be used to retrieve, at least partially, analogous existence results. In this framework, the persistence of the twist condition under small perturbations of the boundary will play a crucial role, and this is the reason why, although not immediately used, this nondegeneracy condition has been investigated in the circular case.

Under particular assumptions on the physical parameters, one can prove the existence of a second type of fixed points different from the ones which correspond to the ejection-collision solutions:

Proposition 2.3.11. *Fixed $\mathcal{E} > \omega^2 > 0$, let us define*

$$\bar{\mu} = \frac{4\mathcal{E} + \sqrt{8\mathcal{E}^3(4\mathcal{E} - \omega^2)}}{2\mathcal{E} - \omega^2} > 2\mathcal{E} - \omega^2, \quad \bar{h} = \frac{2\mathcal{E} - \omega^2}{8\mathcal{E}^2}\mu^2 - (\mathcal{E} + \mu).$$

If $(\mu > \bar{\mu} \ \wedge \ h > \bar{h})$ or $(2\mathcal{E} - \omega^2 < \mu \leq \bar{\mu} \ \wedge \ h > 0)$ there is $\bar{I}^{(1)} \in (0, I_c)$ such that for every $\xi_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}$ the points $(\xi_0, \bar{I}^{(1)})$ and $(\xi_0, -\bar{I}^{(1)})$ are non-homothetic fixed points of $\mathcal{F}$.
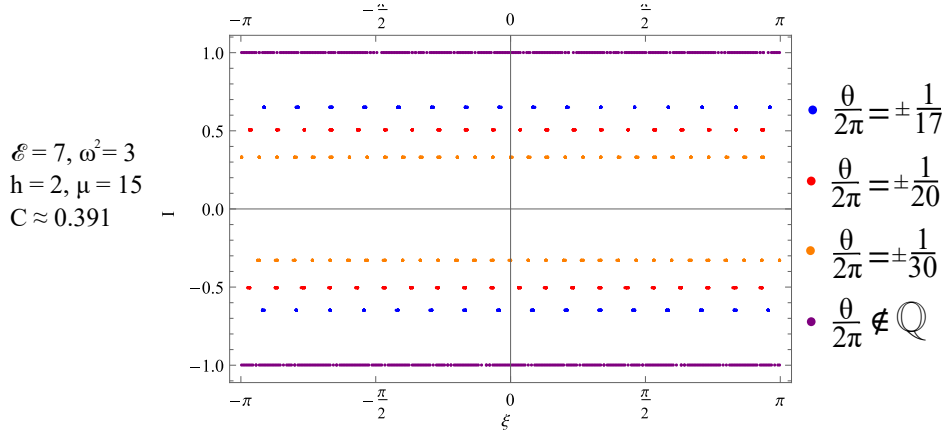
Fig. 2.6 Examples of periodic and non-periodic orbits on the circle in the phase space $(\xi, I) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I}$.

Proof. Recalling that $f(0) + g(0) = 0$, $\lim\limits_{I \to I_c^-} f(I) + g(I) = \bar{\theta} + \pi$ and Eq.(2.3.35), from direct computations one has that

- fixed $\mathcal{E} > \omega^2 > 0$ and $\mu > 0$,

$$f'(0) + g'(0) < 0 \iff h > \bar{h};$$

- fixed $\mathcal{E} > \omega^2 > 0$,

$$\bar{h} > 0 \Leftrightarrow \mu > \bar{\mu} \quad \text{and} \quad \bar{\theta} + \pi > 0 \Leftrightarrow \mu > 2\mathcal{E} - \omega^2.$$

If $(\mu > \bar{\mu} \,\wedge\, h > \bar{h})$ or $(2\mathcal{E} - \omega^2 < \mu \leq \bar{\mu} \,\wedge\, h > 0)$, we have then that $f'(0) + g'(0) < 0$ and $\lim\limits_{I \to I_c^-} f(I) + g(I) > 0$: as a consequence, there exists $\bar{I}^{(1)} > 0$ such that $f(\bar{I}^{(1)}) + g(\bar{I}^{(1)}) = 0 = f(-\bar{I}^{(1)}) + g(-\bar{I}^{(1)})$, and then for every $\xi_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}$, $(\xi_0, \pm\bar{I}^{(1)})$ are fixed points of $\mathcal{F}$. Given that $\bar{I}^{(1)} \neq 0$, these points are not homothetic (see Figure 2.7). $\qquad\square$

2.3.5   Caustics for the unperturbed case

A question of great interest in the study of billiards is the one of caustics, which plays a key role in the determining the regions of the plane where the orbits can access. A caustic is a smooth closed curve Γ such that every trajectory which is tangent to Γ in a point remains tangent to the latter after every passage in and out the domain D. The issue of the existence of caustics in standard billiards ([27, 17]) and its variants ([15]) has been widely studied; in particular, in the framework of a standard convex billiard D, Lazutkin used the KAM approach to prove that, if ∂D is sufficiently smooth (of class C^{553} in the original paper [27], later improved to C^6 by Douady in [28]), then there exists a discontinuous family of caustics in a small neighbourhood of ∂D.
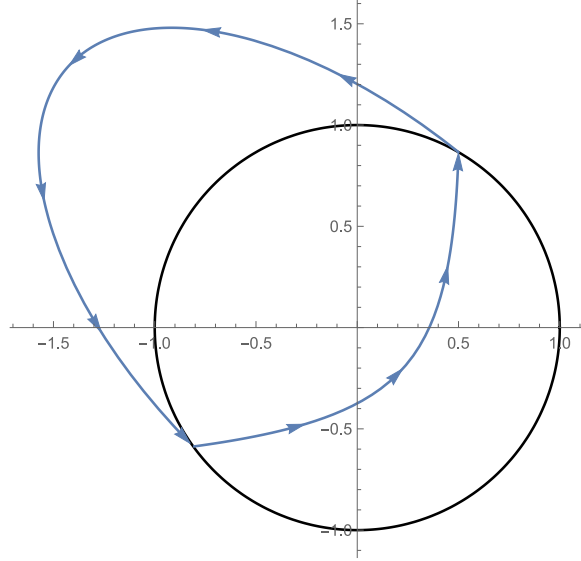
Fig. 2.7 Example of a non-homothetic fixed point for $\mathcal{F}$, with $\mathcal{E} = 7, \omega^2 = 3, h = 2, \mu = 15$. In this case, with reference to Proposition 2.3.11, $\bar{\mu} = 41.6287$.

The aim of this Section is to extend the concept of caustic to our refractive model in the circular case: in view of the presence of two distinct dynamics inside and outside the domain D, one shall search for two of such curves, which can be studied separately. Moreover, by the central symmetry typical of the circular case, it is reasonable to foresee that the inner and outer caustics are circles of suitable radii depending on the action I_0.

Theorem 2.3.12. *For every* $\mathcal{E}, h, \omega, \mu > 0$, $\mathcal{E} > \omega^2$, *given* $I_0 \in (-I_c, I_c)$:

- *the exterior caustic* $\Gamma_E(\zeta; I_0)$ *is given by the locus of the apocenters of the outer ellipses, namely,*

$$\Gamma_E : [0, 2\pi] \to \mathbb{R}^2, \quad \Gamma_E(\zeta; I_0) = R_E(\cos \zeta, \sin \zeta),$$

$$R_E = \frac{\sqrt{\mathcal{E} + \sqrt{\mathcal{E}^2 - 2I_0^2\omega^2}}}{\omega};$$

- *the interior caustic* $\Gamma_I(\zeta; I_0)$ *is the locus of the pericenters of the inner Keplerian hyperbolæ. In particular,*

$$\Gamma_I : [0, 2\pi] \to \mathbb{R}^2 \quad \Gamma_I(\zeta; I_0) = R_I(\cos \zeta, \sin \zeta),$$
$$R_I = \frac{p}{1 + e},$$

where

$$p = \frac{2I_0^2}{\mu}, \quad e = \sqrt{1 + \frac{4I_0^2(\mathcal{E} + h)}{\mu^2}}.$$

In general, following [17] and [78], one shall give the following characterization for the caustics: take an orbit for our dynamical system and suppose that one of its

(interior or exterior) branches is implicitly defined through the relation

$$G(x, y; \xi) = 0, \tag{2.3.37}$$

where $G : \mathbb{R}^2 \to \mathbb{R}$ is of class C^2 in all the variables and ξ acts as a parameter (for example, it could denote the polar angle of the initial point of the branch, its pericenter or apocenter). The caustic Γ can be then seen as th envelope of the family of curves defined by (2.3.37) varying $\xi \in [0, 2\pi]$, that is, the set of points $(x_0(\xi), y_0(\xi))$ satisfying

$$\begin{cases} G(x, y; \xi) = 0 \\ \partial_\xi G(x, y; \xi) = 0 \end{cases}. \tag{2.3.38}$$

By means of the implicit function theorem, it is straightforward that if

$$\nabla_{(x,y)} G(x, y; \xi) \nparallel \nabla_{(x,y)} \partial_\xi G(x, y; \xi) \text{ on the solutions of } (2.3.38), \tag{2.3.39}$$

then (2.3.38) defines a regular curve $\Gamma(\xi) = (x_0(\xi), y_0(\xi))$.
The proof of Theorem 2.3.12 relies on the evaluation of (2.3.38) in the particular cases of the inner and outer dynamics: in the case of circular domains, the solutions of such system can be computed explicitly.

Outer caustic Given $p_0 = (p_x, p_y) = e^{i\xi_0}$, $v_0 = (v_x, v_y) \in \mathbb{R}^2$, $\|v_0\| = \sqrt{2\mathcal{E} - \omega^2}$, from the proof of Theorem 2.3.1, one has that the solution of

$$\begin{cases} (HS_E)[z(s)] & s \in [0, T_E] \\ z(0) = p_0, \; z'(0) = v_0 \end{cases} \tag{2.3.40}$$

can be parametrized as

$$(x(s), y(s)) = \left(p_x \cos(\omega s) + \frac{v_x}{\omega} \sin(\omega s), p_y \cos(\omega s) + \frac{v_y}{\omega} \sin(\omega s) \right).$$

If, as in Section 2.2, $\alpha \in (-\pi/2, \pi/2)$ denotes the angle between p_0 and v_0, recalling the definition of canonical action (2.2.9) one has

$$v_0 = \sqrt{2\mathcal{E} - \omega^2}(\cos \alpha \; p_0 + \sin \alpha \; t_0) = \sqrt{2\mathcal{E} - \omega^2 - 2I_0^2} \; p_0 + \sqrt{2} I_0 \; t_0,$$

were $t_0 = ie^{i\xi}$ is the tangent unit vector to ∂D_0 in p_0. Since in the circular case the action I_0 is constant along the orbits, it can be treated as a parameter in $\mathcal{I}$. Additionally, consider the non-homothetic case, that is, suppose $I_0 \neq 0$ (the case $I_0 = 0$ can be easily analysed separately, leading to the same result).

Taking the function $r^2(s) = x^2(s) + y^2(s)$, by direct computations one has

$$a^2 = \max_{s \in \left[0, \frac{2\pi}{\omega}\right]} r^2(s) = \frac{\mathcal{E} + \sqrt{\mathcal{E}^2 - 2I_0^2 \omega^2}}{\omega^2} > 0$$

$$b^2 = \min_{s \in \left[0, \frac{2\pi}{\omega}\right]} r^2(s) = \frac{\mathcal{E} - \sqrt{\mathcal{E}^2 - 2I_0^2 \omega^2}}{\omega^2} > 0.$$

In the reference frame whose axes coincide with the ellipse's ones, denoted with $R(O, x'', y'')$, the outer arc can be then implicitly defined as a segment of the conic

$$G_{E,0}(x'', y'') = \frac{x''^2}{a^2} + \frac{y''^2}{b^2} - 1 = 0.$$

Denoting by ζ the polar angle of one of the apocenter points, all the solutions of (2.3.40) with $\angle(p_0, v_0) = \alpha$ are then implicitly defined by

$$G_E(x, y; \zeta) = \frac{(x \cos \zeta + y \sin \zeta)^2}{a^2} + \frac{(y \cos \zeta - x \sin \zeta)^2}{b^2} - 1 = 0$$

where $\zeta \in [0, 2\pi]$ is treated as a parameter.
As

$$\partial_\zeta G_E(x, y; \zeta) = \frac{\sqrt{\mathcal{E} - 2I_0^2 \omega^2}}{I_0^2} \left((x^2 - y^2) \sin(2\zeta) - 2xy \cos(2\zeta) \right),$$

the explicit formulation of (2.3.38) for the outer arcs is then, for $\zeta \neq k\dfrac{\pi}{2}$, $k = 0, 1, 2, 3$

$$\begin{cases} \dfrac{(x \cos \zeta + y \sin \zeta)^2}{a^2} + \dfrac{(y \cos \zeta - x \sin \zeta)^2}{b^2} - 1 = 0 \\ \dfrac{\sqrt{\mathcal{E}^2 - 2I_0^2 \omega^2}}{I^2} \sin(2\zeta) \, (x + y \cot \zeta) \, (x - y \tan \zeta) \,; \end{cases} \qquad (2.3.41)$$

note that, in the degenerate cases $\zeta = k\dfrac{\pi}{2}$, from $\partial_\zeta G_E(x, y; \zeta) = 0$ one obtains $x = 0$ or $y = 0$.

The solutions of (2.3.41) are the ellipse's apocenters and pericenters: since the outer arc of the considered dynamical systems involves only the first apocenter, the only admissible solution of (2.3.41) is given by

$$(\bar{x}(\zeta), \bar{y}(\zeta)) = \frac{\sqrt{\mathcal{E} + \sqrt{\mathcal{E}^2 - 2I_0^2 \omega^2}}}{\omega} (\cos \zeta, \sin \zeta),$$

which describes, for $\zeta \in [0, 2\pi]$, the caustic $\Gamma_E(\zeta; I_0)$ as the circle of radius R_E of Theorem 2.3.12.

Although the caustics for the circular domain are completely determined, let us investigate the nondegeneracy condition (2.3.39), which will be generalized for small

perturbations of D_0 in Section 2.4.3. By direct computations, one has

$$\nabla_{(x,y)}G_E(x,y;\zeta)_{|(\bar{x},\bar{y})} = 2a\frac{\sqrt{\mathcal{E}^2 - 2I_0^2\omega^2}}{I_0^2}(-\cos\zeta, -\sin\zeta),$$

$$\nabla_{(x,y)}\partial_\zeta G_E(x,y;\zeta)_{|(\bar{x},\bar{y})} = 2a\frac{\sqrt{\mathcal{E}^2 - 2I_0^2\omega^2}}{I_0^2}(\sin\zeta, -\cos\zeta),$$

leading to

$$\nabla_{(x,y)}G_E(x,y;\zeta)_{|(\bar{x},\bar{y})} \perp \nabla_{(x,y)}\partial_\zeta G_E(x,y;\zeta)_{|(\bar{x},\bar{y})}. \tag{2.3.42}$$

Inner caustic Let us now consider the inner problem

$$\begin{cases} (HS_I)[z(s)] & s \in [0, T_I] \\ z(0) = p_0, \ z'(0) = v_0, \end{cases} \tag{2.3.43}$$

and denote with $\alpha = (\pi/2, 3\pi/2)$ the angle between p_0 and v_0. Given that

$$k = \|p_0 \wedge v_0\| = \|p_0\|\|v_0\|\sin\alpha = \sqrt{2}I_0,$$

one has that the polar equation of the Keplerian inner arc is

$$r = \frac{p}{1 + e\cos f},$$

with

$$p = \frac{2I_0^2}{\mu}, \quad e = \frac{\sqrt{\mu^2 + 4I_0^2(\mathcal{E} + h)}}{\mu}.$$

Choosing the reference frame $R(O, x'', y'')$ where the pericenter is on the positive branch of the x-axis, the inner Keplerian arc is expressed by

$$G_I(x'', y'') = (e^2 - 1)x''^2 - y''^2 - 2pex'' + p^2 = 0, \quad x'' \le \frac{p}{e+1}, \tag{2.3.44}$$

where the inequality condition expresses the choice of the branch of the hyperbola with the concavity in the direction of the central mass.

As in the outer case, denoting with ζ the polar angle of the pericenter, one has that all the Keplerian hyperbolæ with central mass μ, energy $\mathcal{E} + h$ and angular momentum $k = \sqrt{2}I_0$ are given by

$$G_I(x, y; \zeta) = (e^2 - 1)(x\cos\zeta + y\sin\zeta)^2 - (y\cos\zeta - x\sin\zeta)^2 - 2pe(x\cos\zeta + y\sin\zeta) + p^2$$
$$= 0$$

along with

$$x\cos\zeta + y\sin\zeta \le \frac{p}{e+1},$$

with $\zeta \in [0, 2\pi]$. The system (2.3.38) in the inner case becomes then

$$\begin{cases} (e^2 - 1)(x \cos \zeta + y \sin \zeta)^2 - (y \cos \zeta - x \sin \zeta)^2 - 2pe(x \cos \zeta + y \sin \zeta) + p^2 = 0 \\ 2e\,(y \cos \zeta + x \sin \zeta)\,(p - e\,x \cos \zeta + e\,y \sin \zeta) = 0 \end{cases}$$

$$(2.3.45)$$

with the additional condition $(x \cos \zeta + y \sin \zeta) \leq p/(1 + e)$. Problem (2.3.43) admits the unique solution

$$(\bar{x}(\zeta), \bar{y}(\zeta)) = \frac{p}{1 + e}(\cos \zeta, \sin \zeta)$$

which corresponds to the position of the pericenter of the corresponding Keplerian arc, taking ζ as a parameter. The inner caustic $\Gamma_I(\zeta, I_0)$ is then expressed by a circle with radius R_I as in Theorem 2.3.12.

It is straightforward to verify that the nondegeneracy condition (2.3.39) is verified: from

$$\nabla_{(x,y)} G_I(x, y; \zeta)_{|(\bar{x},\bar{y})} = -2p(\cos \zeta, \sin \zeta), \quad \nabla_{(x,y)} \partial_\zeta G_I(x, y; \zeta)_{|(\bar{x},\bar{y})} = \frac{2ep}{1 + e}(\sin \zeta, -\cos \zeta),$$

one has

$$\nabla_{(x,y)} G_I(x, y; \zeta)_{|(\bar{x},\bar{y})} \perp \nabla_{(x,y)} \partial_\zeta G_I(x, y; \zeta)_{|(\bar{x},\bar{y})}.$$

2.4 Perturbations of the circle

Many of the results obtained in the circular case, although significant in themselves, can be generalized to non-circular smooth domains, provided they are close enough to D_0 in a way which will be specified soon. This extension can be performed by means of classical perturbation theory, as well as of more sophisticated results such as KAM and Aubry-Mather theorems (see [25, 22, 77, 24, 23]).

To this end, let us consider a class of domains D_ε whose boundary $\partial D_\varepsilon = \gamma_\epsilon\left(\mathbb{R}_{/2\pi\mathbb{Z}}\right)$ is given by a radial deformation of the circle of the form

$$\gamma_\epsilon : \mathbb{R}_{/2\pi\mathbb{Z}} \to \mathbb{R}^2 \quad \gamma_\epsilon(\xi) = (1 + \varepsilon f(\xi; \varepsilon))\, e^{i\xi}, \tag{2.4.1}$$

where $f(\xi; \varepsilon)$ is a smooth function of $\mathbb{R}_{/2\pi\mathbb{Z}} \times [-C_\varepsilon, C_\varepsilon]$, with $C_\varepsilon > 0$ arbitrarily large; note that, from the choice of the parametrization of γ_ϵ, the variable ξ still represents the polar angle of ξ.

This section aims to analyze the generating function S_ε, with particular emphasis to its good definition and nondegeneracy properties, and the associated first return map $\mathcal{F}_\varepsilon$, whose orbits, when possible, will be studied in terms of their rotation numbers.

2.4.1 Global existence of the outer and inner arcs for the perturbed dynamics

As for the circular case (see Section 2.3.1), the generating function associated to the so-called *perturbed dynamics*, that is, the dynamics induced by the potential (1.1.1) inside and outside the perturbed domain D_ε, is given by

$$S(\xi_0, \xi_1; \varepsilon) = d_E(\gamma_\epsilon(\xi_0), \gamma_\epsilon(\tilde{\xi})) + d_I(\gamma_\epsilon(\tilde{\xi}), \gamma_\epsilon(\xi_1)),$$

where $\gamma_\epsilon(\tilde{\xi})$ is the passage point of $z_{EI}(s)$, as defined in Section 1.4, through ∂D_ε. A preliminary passage to discuss the good definition of S_ε as a whole is to ensure that the functions $d_E(p_0, p_1)$ and $d_I(p_0, p_1)$ are differentiable as functions of $p_0, p_1 \in \partial D_\varepsilon$, namely, that the inner and outer dynamics admit a unique geodesic arc joining p_0 and p_1.

In view of the results of Section 2.3.1, this follows from the continuous dependence of the solutions of the fixed ends problems (2.3.1) and (2.3.13) with respect to p_0 and p_1. To fix the notation, let us denote with $z_{E \backslash I}(s; p_0, p_1; 0)$ the respective solutions in the unperturbed circular case, where the last variable refers to $\varepsilon = 0$.

Remark 2.4.1. *Focusing on the outer problem, from the continuous dependence on p_0 and p_1 of the solution $z_E(s; p_0, p_1; 0)$ defined in Theorem 2.3.1, along with the invariance of the system under rotations, there exists $\rho_E > 0$ such that for every $p_0, p_1, \tilde{p}_0, \tilde{p}_1$ satisfying $\|p_0 - p_1\| < 2$ and $\|\tilde{p}_0 - p_0\|, \|\tilde{p}_1 - p_1\| < \rho_E$ one finds $T > 0$ and a unique solution $z_E(s; \tilde{p}_0, \tilde{p}_1)$ of the problem*

$$\begin{cases} (HS_E)[z(s)] & s \in [0, T] \\ z(0) = \tilde{p}_0, \quad z(T) = \tilde{p}_1. \end{cases} \tag{2.4.2}$$

For computational reasons, we require $\rho_E < 1$, ad set

$$\mathcal{S}_{\rho_E} = \bigcup_{p_0 \in \partial D_0} B_{\rho_E}(p_0) = \{p \in \mathbb{R}^2 \mid dist(p, \partial D_0) < \rho_E\}. \tag{2.4.3}$$

Proposition 2.4.2. *There exists $\delta > 0$ such that for every $\tilde{p}_0, \tilde{p}_1 \in \mathcal{S}_{\rho_E}$ with $\|\tilde{p}_0 - \tilde{p}_1\| < \delta$ there is $T > 0$ and a unique $z_E(s; \tilde{p}_0, \tilde{p}_1)$ solution of the fixed-end problem*

$$\begin{cases} (HS_E)[z(s)] & s \in [0, T] \\ z(0) = \tilde{p}_0, \quad z(T) = \tilde{p}_1. \end{cases} \tag{2.4.4}$$

Proof. It is sufficient to set $\delta < 2(1 - \rho_E)$. Denoting in polar coordinates $\tilde{p}_0 = r_0 e^{i\theta_0}$ and $\tilde{p}_1 = r_1 e^{i\theta_1}$, consider $p_0 = e^{i\theta 0}$ and $p_1 = e^{i\theta_1}$: we have then $\|\tilde{p}_0 - p_0\|, \|\tilde{p}_1 - p_1\| < \rho_E$, and

$$\|p_0 - p_1\| \leq \|\tilde{p}_0 - p_0\| + \|\tilde{p}_0 - \tilde{p}_1\| + \|\tilde{p}_1 - p_1\| < \delta + 2\rho_E < 2, \tag{2.4.5}$$

then, by Remark 2.4.1, the thesis is proved. □

Theorem 2.4.3. *There are $\bar{\delta}_E > 0, \bar{\epsilon}_E > 0$ such that for every $\varepsilon \in \mathbb{R}$ and for every $\xi_0, \xi_1 \in \mathbb{R}_{/2\pi\mathbb{Z}}$ with $|\xi_0 - \xi_1| < \bar{\delta}_E$ and $|\varepsilon| < \bar{\varepsilon}_E$ there is $T > 0$ and a unique function $z_E(s; \gamma_\epsilon(\xi_0), \gamma_\epsilon(\xi_1)) \equiv z_E(s; \xi_0, \xi_1; \varepsilon)$ which is a classical solution of*

$$\begin{cases} (HS_E)[z(s)] & s \in [0, T] \\ z(0) = \gamma_\epsilon(\xi_0), \quad z(T) = \gamma_\epsilon(\xi_1). \end{cases} \tag{2.4.6}$$

Proof. The claim is true if $\bar{\delta}_E < \delta/2$ and $\bar{\epsilon}_E < \min\{1/\left(\|f\|_\infty + \|\partial_\xi f\|_\infty\right), \rho_E/\|f\|_\infty\}$. For, fixed $\varepsilon \in \mathbb{R}$ such that $|\varepsilon| < \bar{\varepsilon}_E$, one has that for every $\xi \in \mathbb{R}_{/2\pi\mathbb{Z}}$

$$\|\dot{\gamma}_\varepsilon(\xi)\| = |\varepsilon \partial_\xi f(\xi, \varepsilon) e^{i\xi} + (1 + \varepsilon f(\xi, \varepsilon)) \, i e^{i\xi}| \leq 1 + |\varepsilon| \left(\|\partial_\xi f\|_\infty + \|f\|_\infty\right) < 2.$$

If $\xi_0, \xi_1 \in \mathbb{R}_{/2\pi\mathbb{Z}}$ are such that $|\xi_0 - \xi_1| < \bar{\delta}_E < \delta/2$, defining $\tilde{p}_0 = \gamma_\epsilon(\xi_0)$, $\tilde{p}_1 = \gamma_\epsilon(\xi_1)$, $p_0 = e^{i\xi_0}$ and $p_1 = e^{i\xi_1}$:

$$\begin{aligned} \|\tilde{p}_0 - p_0\| &= |\epsilon| \, |f(\xi_0)| \leq |\epsilon| \, \|f\|_\infty < \rho_E, \\ \|\tilde{p}_1 - p_1\| &< \rho_E, \\ \|\tilde{p}_0 - \tilde{p}_1\| &< \|\dot{\gamma}_\varepsilon\|_\infty |\xi_0 - \xi_1| < 2|\xi_0 - \xi_1| < \delta, \end{aligned} \tag{2.4.7}$$

then the hypotheses of Proposition 2.4.2 hold and the claim is true.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We stress that Theorem 2.4.3 can be proved exactly as in its general analogous in Section 2.2.2 as well. Nevertheless, relying again on the properties of the map on the circle, with the proof proposed we obtained more precise estimates on the quantities $\bar{\delta}_E$ and $\bar{\epsilon}_E$, including their relation with the perturbing function f.

Passing to the inner dynamics, let us recall that, by the choice of the parametrization of γ_ϵ, the parameter ξ still represents the polar angle of the corresponding point $\gamma_\epsilon(\xi)$. This means that if we want two endpoints $\gamma_\epsilon(\xi_1)$ and $\gamma_\epsilon(\xi_2)$ such that they are not antipodal, it is sufficient to require that they are sufficiently close to each others.

Theorem 2.4.4. *There exist $\bar{\varepsilon}_I > 0$, $\bar{\delta}_I > 0$ and $C > 0$ such that for every $\varepsilon \in \mathbb{R}$, $\xi_0, \xi_1 \in \mathbb{R}_{/2\pi\mathbb{Z}}$ satisfying $|\xi_0 - \xi_0| < \bar{\delta}_I$ and $|\varepsilon| < \bar{\varepsilon}_I$ there exists a unique $T(\xi_0, \xi_1) \equiv T > 0$ and a unique solution $z_I(s; \xi_0, \xi_1; \varepsilon)$ of*

$$\begin{cases} (HS_I)[z(s)] & s \in [0, T] \\ z(s) \in D_\varepsilon & s \in (0, T) \\ z(0) = \gamma_\epsilon(\xi_0), \; z(T) = \gamma_\epsilon(\xi_1) \end{cases} \tag{2.4.8}$$

with the following properties:

- *if $\xi_0 \neq \xi_1$, then $z(s; \xi_0, \xi_1; \varepsilon)$ is a classical solution of (2.4.8) which satisfies the* (TnT) *topological condition;*

- *if $\xi_0 = \xi_1$, $z(s; \xi_0, \xi_0; \varepsilon)$ is an ejection-collision solution.*

In any case, $z(s; \xi_0, \xi_1; \varepsilon)$ is of class C^1 with respect to variations of ξ_0 and ξ_1 and, if we define

$$\beta_0 = \angle(\dot{\gamma}_\epsilon(\xi_0), z'(0; \xi_0, \xi_1; \varepsilon)), \quad \beta_1 = \angle(\dot{\gamma}_\epsilon(\xi_1), z'(T; \xi_0, \xi_1; \varepsilon)), \tag{2.4.9}$$

one has that $|\beta_0| > \bar{C}$ and $|\beta_1| > \bar{C}$ for some $\bar{C} > 0$.

Proof. Given $\tilde{p}_0 \doteq \gamma_\epsilon(\xi_0)$ and $\tilde{p}_1 \doteq \gamma_\epsilon(\xi_1)$, the existence of a unique *(TnT)* solution of

$$\begin{cases} (HS_I)[z(s)] & s \in [0, T] \\ z(0) = \tilde{p}_0, \ z(T) = \tilde{p}_1 \end{cases}$$

is guaranteed by Theorem 2.2.3 as long as $|\xi_0 - \xi_1| < \pi$. Moreover, from the regularity of γ_ϵ with respect to ξ and of $z(s; \tilde{p}_0, \tilde{p}_1)$ with respect to the endpoints, we have that $z(s; \xi_0, \xi_1; \varepsilon)$ is of class C^1 in the variables ξ_0 and ξ_1.

To prove the transversality properties of $z_I(s; \xi_0, \xi_1; \varepsilon)$, let us observe that by Theorem 2.3.2 and the differentiability of $z(s) \equiv z(s; \xi_0, \xi_1; \varepsilon)$, as well as by the invariance under rotations of the system, if ϵ is small enough (let us say, $|\epsilon| < \epsilon_1$ for a certain $\epsilon_1 > 0$) there is $C_1 > 0$, possibly lower than C, such that for every ξ_0, ξ_1 satisfying the existence hypotheses, defined $\tilde{p}_0$ and $\tilde{p}_1$ as above, one has

$$1 \geq -\frac{\tilde{p}_0}{|\tilde{p}_0|} \cdot \frac{z'(0)}{|z'(0)|} > C_1, \ \ 1 \geq \frac{\tilde{p}_1}{|\tilde{p}_1|} \cdot \frac{z'(T)}{|z'(T)|} > C_1. \tag{2.4.10}$$

Let us now consider $\alpha_1 = \angle(\tilde{p}_1, z'(T))$ (if $\alpha_0 = \angle(-\tilde{p}_0, z'(0))$ we proceed analogously): setting $C_2 = \arccos(C_1) \in (0, \pi/2)$, we have $|\alpha_1| < C_2$, and, taking $\beta_1 = \angle(\dot{\gamma}_\epsilon(\xi_1), z'(T))$, one has

$$\beta_1 = \angle(\dot{\gamma}_\epsilon(\xi_1), \tilde{p}_1^\perp) + \angle(\tilde{p}_1^\perp, z'(T)), \tag{2.4.11}$$

where $\tilde{p}_1^\perp = i\tilde{p}_1$. Then we have that

$$|\beta_1| \geq \left| \angle(\tilde{p}_1^\perp, z'(T)) \right| - \left| \angle(\dot{\gamma}_\epsilon(\xi_1), \tilde{p}_1^\perp) \right|,$$
$$\left| \angle(\tilde{p}_1^\perp, z'(T)) \right| > \frac{\pi}{2} - C_2 \equiv C_3 \in \left(0, \frac{\pi}{2}\right). \tag{2.4.12}$$

To estimate $\angle(\dot{\gamma}_\epsilon(\xi_1), \tilde{p}_1^\perp)$, let us observe that

$$\left| \sin\left(\angle(\dot{\gamma}_\epsilon(\xi_1), \tilde{p}_1^\perp)\right) \right| = \frac{\|\dot{\gamma}_\epsilon(\xi_1) \wedge \tilde{p}_1^\perp\|}{\|\dot{\gamma}_\epsilon(\xi_1)\| \|\tilde{p}_1^\perp\|}, \tag{2.4.13}$$

where $\tilde{p}_1^\perp / \|\tilde{p}_1^\perp\| = ie^{i\xi_1}$ and $\dot{\gamma}_\epsilon(\xi_1) = \varepsilon \partial_\xi f(\xi_1, \varepsilon) e^{i\xi_1} + (1 + \varepsilon f(\xi_1, \varepsilon))ie^{i\xi_1}$.
If $|\varepsilon| < \varepsilon_1' \doteq \left\{ \epsilon_1, \dfrac{1}{2\|f\|_\infty} \right\}$:

$$\|\dot{\gamma}_\epsilon(\xi_1)\| = \sqrt{(\varepsilon \partial_\xi f(\xi_1, \varepsilon))^2 + (1 + \varepsilon f(\xi_1, \varepsilon))^2} \geq 1 - |\varepsilon| \, \|f\|_\infty > \frac{1}{2}$$
$$\Rightarrow \left| \sin\left(\angle(\dot{\gamma}_\epsilon(\xi_1), \tilde{p}_1^\perp)\right) \right| = \frac{|\varepsilon| \, |\partial_\xi f(\xi_1, \varepsilon)|}{\|\dot{\gamma}_\epsilon(\xi_1)\|} < 2|\varepsilon| |\partial_\xi f(\xi_1, \varepsilon)|. \tag{2.4.14}$$

If we consider $C_4 > 0$ with $0 < \arcsin(C_4) < C_3$, setting $\varepsilon < \min\left\{\varepsilon_1', \dfrac{C_4}{2\|\partial_\xi f\|_\infty}\right\} = \bar{\varepsilon}_I$:

$$\left| \sin\left(\angle(\dot{\gamma}_\epsilon(\xi_1), \tilde{p}_1^\perp)\right)\right| < C_4 \Rightarrow |\beta_1| > C_3 - \arcsin(C_4) \equiv \bar{C} > 0. \qquad (2.4.15)$$

Recalling the definitions which lead to $\bar{\varepsilon}$ and $\bar{C}$, it is clear that they do not depend on ξ_0 nor ξ_1.

Finally, the condition $z(s) \in D_\varepsilon$ for $s \in (0, T)$ follows from the smallness of ε (possibly reducing $\bar{\epsilon}_I$) and the transversality of $z(s)$ with respect to the perturbed domain ∂D_ε, which ensures that $z(s)$ do not intersect twice the domain's boundary in a neighbourhood of $z(T)$. $\qquad\square$

Remark 2.4.5. *Using the same transversality argument described in details for the inner dynamics, one can prove that, if ε is small enough and ξ_0, ξ_1 sufficiently close, the solution $z_E(s; \xi_0, \xi_1; \varepsilon)$ of (2.4.6), whose existence is ensured by Theorem 2.4.3, is such that $z_E(s; \xi_0, \xi_1; \varepsilon) \notin \bar{D}_\varepsilon$ for $s \in (0, T)$.*

2.4.2   Invariant sets for $\mathcal{F}_\varepsilon$

The good definition of the distances $d_E(p_0, p_1)$ and $d_I(p_0, p_1)$ for $p_0, p_1 \in \partial D_\varepsilon$ allows to consider the associated generating function

$$S(\xi_0, \xi_1; \varepsilon) = S_E(\xi_0, \tilde{\xi}; \varepsilon) + S_I(\tilde{\xi}, \xi_1; \varepsilon) = d_E(\gamma_\epsilon(\xi_0), \gamma_\epsilon(\tilde{\xi})) + d_I(\gamma_\epsilon(\tilde{\xi}), \gamma_\epsilon(\xi_1)). \quad (2.4.16)$$

When well defined, $S(\xi_0, \xi_1; \varepsilon)$ has the same regularity of $f(\xi, \varepsilon)$ as a function of both the angle variables ξ_0, ξ_1 and the perturbative parameter ε.

This section aims to prove that, under suitable assumptions, the results proved for the circle regarding the twist condition and the existence of invariant sets with prescribed rotation numbers (see Section 2.3.4) can be extended to the perturbed dynamics as described in (2.4.1).

Recalling the notation of Section 2.3, $\bar{\mathcal{I}}$ is the finite set in $\mathcal{I} = (-I_c, I_c)$ for which $\mathcal{F}_0$ is not well defined, with $I_c = \sqrt{\mathcal{E} - \omega^2/2}$. To highlight the dependence on ε, from now on we will use the notation $\mathcal{F}_0(\xi_0, I_0) \equiv \mathcal{F}(\xi_0, I_0; 0)$.

Proposition 2.4.6. *Let $[a, b] \subset \mathcal{I}\backslash\bar{\mathcal{I}}$, and suppose that, in (2.4.1), $f \in C^k\left(\mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I}\right)$ with $k \geq 2$. Then there exists $\bar{\varepsilon} > 0$ such that for every $\varepsilon \in \mathbb{R}$, $|\varepsilon| < \bar{\varepsilon}$, the perturbed first return map*

$$\mathcal{F}(\xi_0, I_0; \varepsilon) = (\xi_1(\xi_0, I_0; \varepsilon), I_1(\xi_0, I_0; \varepsilon))$$

is well defined and of class $C^{k-2}(\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b])$. Moreover, $\mathcal{F}(\cdot, \cdot, \varepsilon)$ is area-preserving and twist.

Proof. Let us consider $[a, b] \subset \mathcal{I}\backslash\bar{\mathcal{I}}$, and, with reference to (2.3.32), define

$$K = \left\{(\xi_0, \xi_0 + \bar{\theta}(I_0)) \mid \xi_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}, I_0 \in [a, b]\right\}:$$

in view of Propositions 2.3.6 and 2.3.7, the generating function $S(\xi_0, \xi_1; 0)$ is well defined and infinitely many differentiable in K, and the same holds for $\mathcal{F}(\xi_0, I_0; 0)$ in $\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b]$. Moreover, K is a compact subset of the torus $\mathbb{R}_{/2\pi\mathbb{Z}} \times \mathbb{R}_{/2\pi\mathbb{Z}}$. In particular, one has that the quantity

$$\partial_b^2 S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1; 0); 0) + \partial_a^2 S_I(\tilde{\xi}(\xi_0, \xi_1; 0), \xi_1; 0), \qquad (2.4.17)$$

with $\tilde{\xi}(\xi_0, \xi_1; 0)$ such that $\partial_b S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1; 0); 0) + \partial_a S_I(\tilde{\xi}(\xi_0, \xi_1; 0), \xi_1; 0) = 0$, is different from 0 and has always the same sign for $(\xi_0, \xi_1) \in K$.

Let us now fix $(\bar{\xi}_0, \bar{\xi}_1) \in K$: by the implicit function theorem, there are two neighbourhoods $A_{\bar{\xi}_0}, A_{\bar{\xi}_1}$ respectively of $\bar{\xi}_0$ and $\bar{\xi}_1$, a quantity $\bar{\varepsilon}(\bar{\xi}_0, \bar{\xi}_1) > 0$ and a unique function $\tilde{\xi}(\xi_0, \xi_1; \varepsilon)$, defined in $A_{\bar{\xi}_0} \times A_{\bar{\xi}_1} \times [-\bar{\varepsilon}(\bar{\xi}_0, \bar{\xi}_1), \bar{\varepsilon}(\bar{\xi}_0, \bar{\xi}_1)]$, such that the refraction law

$$\partial_b S_E(\xi_0, \tilde{\xi}(\xi_0, \xi_1; \varepsilon); \varepsilon) + \partial_a S_I(\tilde{\xi}(\xi_0, \xi_1; \varepsilon), \xi_1; \varepsilon) = 0$$

holds also in the perturbed case. Moreover, the function $\tilde{\xi}$ is of class C^{k-1} in all its variables. As a consequence, the generating function $S(\xi_0, \xi_1; \varepsilon)$ is well defined in $A_{\bar{\xi}_0} \times A_{\bar{\xi}_1} \times [-\bar{\varepsilon}(\bar{\xi}_0, \bar{\xi}_1), \bar{\varepsilon}(\bar{\xi}_0, \bar{\xi}_1)]$. Varying $(\bar{\xi}_0, \bar{\xi}_1) \in K$, the family

$$\left\{ A_{\bar{\xi}_0} \times A_{\bar{\xi}_1} \mid (\bar{\xi}_0, \bar{\xi}_1) \in K \right\}$$

is a covering of K such that, if $(A_{\bar{\xi}_0} \times A_{\bar{\xi}_1}) \cap (A_{\bar{\xi}_0'} \times A_{\bar{\xi}_1'}) \neq \emptyset$, then $\tilde{\xi}(\xi_0, \xi_1; \varepsilon)$ coincide in the intersection. Since K is compact, there exists a finite sequence $(\bar{\xi}_0^{(i)}, \bar{\xi}_1^{(i)})_{i=1}^N$ such that

$$K \subset \bigcup_{i=1}^N A_{\bar{\xi}_0^{(i)}} \times A_{\bar{\xi}_1^{(i)}}.$$

Setting $\bar{\varepsilon}' = \min_{i=1,\dots,N} \bar{\varepsilon}(\bar{\xi}_0^{(i)}, \bar{\xi}_1^{(i)})$, one has that for every $(\xi_0, \xi_1) \in K$ and every $\varepsilon \in \mathbb{R}$ such that $|\varepsilon| < \bar{\varepsilon}'$, the perturbed generating function $S(\xi_0, \xi_1; \varepsilon)$ is well defined and of class C^{k-1}. In such set one can define the canonical actions

$$I_0(\xi_0, \xi_1; \varepsilon) = -\partial_{\xi_0} S(\xi_0, \xi_1; \varepsilon), \quad I_1(\xi_0, \xi_1; \varepsilon) = \partial_{\xi_1} S(\xi_0, \xi_1; \varepsilon),$$

and, by the definition of K, one has that for every $(\xi_0, \xi_1) \in K$, $I_0(\xi_0, \xi_1; 0) \in [a, b]$. Fixing $\bar{\xi}_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}$ and $\bar{I}_0 \in [a, b]$, set $\bar{\xi}_1 = \xi_1(\bar{\xi}_0, \bar{I}_0; 0)$: from the proof of Proposition 2.3.7, one has that $\partial_{\xi_1} \left(\bar{I}_0 + \partial_{\xi_0} S(\bar{\xi}_0, \bar{\xi}_1; 0) \right) \neq 0$, and then, varying $(\bar{\xi}_0, \bar{I}_0)$ in the compact rectangle $[0, 2\pi] \times [a, b]$, one can apply the same reasoning used before to find $0 < \bar{\varepsilon} < \bar{\varepsilon}'$ such that for every $\xi_0 \in [0, 2\pi]$, $I_0 \in [a, b]$ and $|\varepsilon| < \bar{\varepsilon}$ the function $\xi_1(\xi_0, I_0; \varepsilon)$ is well defined and of class C^{k-2}. Extending $\xi_1(\xi_0, I_0; \varepsilon)$ by periodicity for $\xi_0 \in \mathbb{R}_{/2\pi\mathbb{Z}}$, one has then that, for every $\varepsilon > 0$, $|\varepsilon| < \bar{\varepsilon}$, the perturbed first return map

$$\mathcal{F}(\xi_0, I_0; \varepsilon) = (\xi_1(\xi_0, I_0; \varepsilon), I_1(\xi_0, I_0; \varepsilon)),$$

where $I_1(\xi_0, I_0; \varepsilon) = I_1(\xi_0, \xi_1(\xi_0, I_0; \varepsilon); \varepsilon)$, satisfies the claim in terms of good definition and regularity. The area-preserving property is a straightforward consequence of the existence of the perturbed generating function, while the twist property depends on

the existence of $\xi_1(\xi_0, \xi_1; \varepsilon)$, since

$$\frac{\partial \xi_1}{\partial I_0} = \left(\frac{\partial I_0}{\partial \xi_1}\right)^{-1} = -\frac{1}{\partial_{\xi_0 \xi_1} S(\xi_0, \xi_1; \varepsilon)}.$$

□

Remark 2.4.7. *Proposition 2.4.6 remains valid if we ask weaker regularity hypotheses on $f(\xi, \varepsilon)$. In particular, if f is of class C^k in ξ and is continuous, along with all its k $\xi-$derivatives, in ε, one can find $\bar{\varepsilon} > 0$ such that, for $|\varepsilon| < \bar{\varepsilon}$, the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ is of class C^{k-2} in $\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b]$ and continuous in ε, and the same holds for all its $k-2$ derivatives.*

The map $\mathcal{F}(\xi_0, I_0; \varepsilon)$, whose existence under suitable conditions and for subsets of $\mathbb{R}_{/2\pi\mathbb{Z}} \times \mathcal{I}$ is ensured by Proposition 2.4.6, can be expressed in the form

$$\mathcal{F}(\xi_0, I_0; \varepsilon) = \begin{cases} \xi_1 = \xi_0 + \bar{\theta}(I_0) + F(\xi_0, I_0; \varepsilon) \\ I_1 = I_0 + G(\xi_0, I_0; \varepsilon) \end{cases} \tag{2.4.18}$$

where F and G are of class C^{k-2} in all the variables and

$$\|F\|_{C^{k-2}} \xrightarrow{\varepsilon \to 0} 0, \quad \|G\|_{C^{k-2}} \xrightarrow{\varepsilon \to 0} 0.$$

We can now prove the existence of particular orbits with prescribed rotation number for $\mathcal{F}_\varepsilon$. We will make use of *KAM Theorem* in the finitely differentiable version of Moser (cfr [22]); before stating the Theorem, let us now give some preliminary definitions.

Definitions 2.4.8. *Let $s \geq 1$ and $f(\xi, I)$ of class C^s in $\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b]$. The s-th derivative norm of f is given by*

$$|f|_s = \sup \left|\left(\frac{\partial}{\partial I}\right)^{m_1} \left(\frac{\partial}{\partial \xi}\right)^{m_2} f(\xi, I)\right|, \quad m_1 + m_2 \leq s.$$

Let us now consider $\mathcal{F}(\xi_0, I_0) = (\xi_1(\xi_0, I_0), I_1(\xi_0, I_0))$ a given map on the annulus $\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b]$. We say that $\mathcal{F}$ has the intersection property if for any closed curve α near the circle, that is, of the form

$$\alpha(\xi_1) = (\xi_1, f(\xi_1))$$

where f is $2\pi-$periodic and with f' small, one has

$$supp(\alpha) \cap supp\,(\mathcal{F}(\alpha)) \neq \emptyset.$$

If $\mathcal{F}$ exact symplectic, the intersection property is straightforwardly verified (see [25, Chapter 1]).
Finally, given $\sigma > 0$, define

$$\mathcal{D}(\sigma) = \left\{\omega \in \mathbb{R} \mid \forall n, m \in \mathbb{Z}, \ n > 0, \ |n\omega - m2\pi| \geq \sigma n^{-3/2}\right\}$$

the set of Diophantine numbers with respect to the constant $\sigma/2\pi$ and the exponent $5/2$. According to [25], given

$$\mathcal{D} = \bigcup_{\sigma > 0} \mathcal{D}(\sigma),$$

one has that $\mathcal{D}$ is dense in $[0,1]$, and by extension in every closed interval of $\mathbb{R}$. As a consequence, for every $[c,d] \subset \mathbb{R}$ and every $\rho \in (c,d) \cap \mathcal{D}$ there exists $\sigma_\rho > 0$ such that

$$\forall \sigma < \sigma_\rho \quad \rho \in (c + \sigma, d - \sigma) \cap D(\sigma).$$

Theorem 2.4.9. *(KAM Theorem, [22]) Let $a, b \in \mathbb{R}$ such that $0 < a < b$ and $b - a \geq 1$, and let*

$$\mathcal{F}_0(\xi_0, I_0) = \begin{cases} \xi_1 = \xi_0 + \bar{\theta}(I_0) \\ I_1 = I_0 \end{cases}$$

be a map on the annulus $\mathbb{R}_{/2\pi\mathbb{Z}} \times [a,b]$; suppose that there is $c_0 \geq 1$ such that

$$c_0^{-1} \leq \frac{\partial \bar{\theta}}{\partial I_0}(I_0) \leq c_0.$$

Moreover, let

$$\mathcal{F}(\xi_0, I_0) = \begin{cases} \xi_1 = \xi_0 + \bar{\theta}(I_0) + F(\xi_0, I_0) \\ I_1 = I_0 + G(\xi_0, I_0) \end{cases}$$

a perturbation of $\mathcal{F}_0$ that satisfies the intersection property.
Fixed $\sigma > 0$ and $s \geq 1$, there are $\delta_0 = \delta_0(c_0, \sigma, s) > 0$ and an integer $l = l(s) > 0$ such that, if

1. *$|F|_0 + |G|_0 < \delta_0$,*

2. *F and G are of class $C^l(\mathbb{R}_{/2\pi\mathbb{Z}} \times [a,b])$ and $|\bar{\theta}|_l + |F|_l + |G|_l < c_0$,*

then $\mathcal{F}$ admits a closed invariant curve

$$\begin{cases} \xi = u + p(u) \\ I = \bar{I} + q(u), \end{cases} \tag{2.4.19}$$

with $\bar{I} \in [a,b]$, which induces a mapping

$$u_1 = u_0 + \bar{\theta}(\bar{I}) \tag{2.4.20}$$

and such that p and q are $2\pi-$periodic functions in the parameter u with s continuous derivatives and

$$|p|_s + |q|_s < \sigma. \tag{2.4.21}$$

Moreover, for every $\omega \in \left(\bar{\theta}(a) + \sigma, \bar{\theta}(b) - \sigma\right) \cap \mathcal{D}(\sigma)$ there exists an invariant curve of the form (2.4.19) with rotation number $\bar{\theta}(\bar{I}) = \omega$.

Remark 2.4.10. *The rotation number of the invariant curve (2.4.19) can be derived from the mapping (2.4.20) as follows: let us take $u_0 \in \mathbb{R}$ and consider the sequence $\{u_n\}_{n \in \mathbb{N}}$ produced by (2.4.20), which is trivially given by $u_n = u_0 + n\bar{\theta}(\bar{I})$. The orbit of*

$\mathcal{F}$ generated by (2.4.20) and lying in the invariant curve (2.4.19) is then $\{(\xi_n, I_n)\}_{n \in \mathbb{N}}$, with

$$\begin{cases} \xi_n = u_0 + n\bar{\theta}(\bar{I}) + p\left(u_0 + n\bar{\theta}(\bar{I})\right) \\ I_n = \bar{I} + q\left(u_0 + n\bar{\theta}(\bar{I})\right). \end{cases} \tag{2.4.22}$$

From the definition (2.3.36) and given that p is bounded, one can easily compute the rotation number associated to the initial condition $(\xi_0, I_0) = (\xi(u_0), I(u_0))$ through the map $\mathcal{F}$ as

$$\rho(\xi_0, I_0) = \lim_{n \to \infty} \frac{\xi_n - \xi_0}{n} = \lim_{n \to \infty} \bar{\theta}(\bar{I}) + \frac{p(u_0 + n\bar{\theta}(\bar{I})) - p(u_0)}{n} = \bar{\theta}(\bar{I}).$$

Remark 2.4.11. *Although in the original paper [22] for $s = 1$ the minimal number of continuous derivatives required for the application of Theorem (2.4.9) is $l = 333$, Rüssman and Hermann reduced this number to $l = 5$ and then to $l > 3$ (see [79] and [80]). Here, the authors require a more restrictive Diophantine hypothesis, requiring that the rotation numbers belong to the set*

$$\tilde{\mathcal{D}}_2 = \bigcup_{\sigma > 0} \mathcal{D}_2(\sigma) = \bigcup_{\sigma > 0} \left\{ \omega \in \mathbb{R} \mid \forall n, m \in \mathbb{Z}, \ n > 0, \ \left| \frac{\omega}{2\pi} - \frac{m}{n} \right| \geq \frac{\sigma}{n^2} \right\}.$$

On the other hand, it is a known fact that $\tilde{\mathcal{D}}_2$ enjoys the same density properties already claimed for $\mathcal{D}$. For this reason, and in view of Proposition 2.4.6, we require the normal perturbation $f(\xi, \varepsilon)$ to be of class $C^k\left(\mathbb{R}_{/2\pi\mathbb{Z}} \times [-\bar{\varepsilon}, \bar{\varepsilon}]\right)$, with $k > 5$: as a consequence, $\mathcal{F}(\xi_0, I_0; \varepsilon) \in C^{k'}\left(\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b] \times [-\bar{\varepsilon}, \bar{\varepsilon}]\right)$, with $k' > 3$, and the invariant curves, if existing, are of class $C^1(\mathbb{R}_{/2\pi\mathbb{Z}})$.

Theorem 2.4.12. *Let us suppose that $\bar{\theta}'(I_0) > 0$ in $[a, b]$, and take $\rho_0, \rho_1 \in \left(\bar{\theta}(a), \bar{\theta}(b)\right) \cap \tilde{\mathcal{D}}_2$. Then there exists $\bar{\varepsilon}_{\rho_0\rho_1}$ such that for every $\varepsilon \in \mathbb{R}$, $|\varepsilon| < \bar{\varepsilon}_{\rho_0\rho_1}$ the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ defined in (2.4.18) admits two closed invariant curves of class C^1 with rotation numbers ρ_0 and ρ_1.*

Proof. To verify the hypotheses of Theorem 2.4.9, let us choose $C > (b - a)^{-1}$ such that $\rho_0' = \rho_0/C, \rho_1' = \rho_1/C \in \mathcal{D}$ (such C exists for the density of $\tilde{\mathcal{D}}_2$ in $\mathbb{R}$), and consider the canonical change of coordinates

$$\left\{ \xi' = \frac{\xi}{C}, \quad I' = C \, I. \right.$$

Expressing $\mathcal{F}_0$ and $\mathcal{F}_\varepsilon$ in the new variables, one obtains the rescaled problem

$$\tilde{\mathcal{F}}_0(\xi_0', I_0') = \begin{cases} \xi_1' = \xi_0' + \Theta(I_0') \\ I_1' = I_0' \end{cases} \qquad \tilde{\mathcal{F}}(\xi_0', I_0'; \varepsilon) = \begin{cases} \xi_1' = \xi_0' + \Theta(I_0') + \tilde{F}(\xi_0', I_0'; \varepsilon) \\ I_1' = I_0' + \tilde{G}(\xi_0', I_0'; \varepsilon), \end{cases}$$

where $I' = CI \in [a', b'] = C[a, b]$, $b' - a' > 1$, and

$$\Theta(I_0') = \frac{\bar{\theta}\left(\frac{I_0'}{C}\right)}{C} = \frac{\bar{\theta}(I_0)}{C}, \quad \tilde{F}(\xi_0', I_0'; \varepsilon) = \frac{1}{C}F\left(C\xi_0', \frac{I_0'}{C}; \varepsilon\right) = \frac{1}{C}F(\xi_0, I_0; \varepsilon),$$

$$\tilde{G}(\xi_0', I_0'; \varepsilon) = C\, G\left(C\xi_0', \frac{I_0'}{C}; \varepsilon\right) = C\, G(\xi_0, I_0; \varepsilon)$$

are defined for $\xi_0' \in \mathbb{R}_{/2\pi\mathbb{Z}}$, $I_0' \in [a', b']$ and $|\varepsilon| < \bar{\varepsilon}$ (see Proposition 2.4.6). Note that the monotonicity and convergence properties

$$\partial_{I_0'}\Theta\left(I_0'\right) > 0, \quad \|\tilde{F}\|_{C^{k-2}} \xrightarrow{\varepsilon \to 0} 0, \quad \|\tilde{G}\|_{C^{k-2}} \xrightarrow{\varepsilon \to 0} 0 \qquad (2.4.23)$$

hold also for the rescaled map, as well as the area-preserving property. Moreover, as $\Theta(a') = \bar{\theta}(a)/C$ and $\Theta(b') = \bar{\theta}(b)/C$, one finds $\sigma > 0$ such that $\rho_0', \rho_1' \in (\Theta(a') + \sigma, \Theta(b') - \sigma) \cap \tilde{\mathcal{D}}_2(\sigma)$.

Since $\tilde{\mathcal{F}}$ is exact symplectic, it satisfies the intersection property, and, given that $\Theta \in C^1([a', b'])$, there is $c_0 > 1$ such that

$$\forall I_0' \in [a', b'] \quad c_0^{-1} \le \Theta'(I_0') \le c_0.$$

Fixed $s = 1$, let us consider $l = l(s)$ as in Theorem 2.4.9, and, eventually taking a higher c_0, suppose $c_0 > |\Theta|_l$. By Theorem 2.4.9, there exists $\delta_0 = \delta_0(c_0, \sigma, s) > 0$ such that, if (1) and (2) hold for $\tilde{\mathcal{F}}$, then the existence of the two invariant orbits for the rescaled problem is ensured. From (2.4.23), one can choose $0 < \bar{\varepsilon}_{\rho_0\rho_1} < \bar{\varepsilon}$, such that for every $\varepsilon \in [-\bar{\varepsilon}_{\rho_0\rho_1}, \bar{\varepsilon}_{\rho_0\rho_1}]$

$$|\tilde{F}|_0 + |\tilde{G}|_0 < \delta_0 \quad \text{and} \quad |\tilde{F}|_l + |\tilde{G}|_l < c_0 - |\Theta|_l,$$

then the hypotheses of Theorem 2.4.9 hold and the invariant curves obtained for $\tilde{\mathcal{F}}$ can be reparametrized to be invariant curves for $\mathcal{F}$. In particular, such curves have rotation number ρ_0 and ρ_1: for example, let us consider the invariant curve for the rescaled problem with rotation number ρ_0', which, in view of Theorem 2.4.9, can be expressed as

$$\begin{cases} \xi_0' = u' + \tilde{p}(u') \\ I_0' = \bar{I}' + \tilde{q}(u') \end{cases} \quad \text{with mapping } u_1' = u_0' + \Theta(\bar{I}') = u_0' + \rho_0'.$$

Returning to the original coordinates and setting $u = C\, u'$, one gets the rescaled invariant curve

$$\begin{cases} \xi_0 = u + p(u) \\ I_0 = \bar{I} + q(u) \end{cases} \quad \text{with mapping } u_1 = u_0 + \bar{\theta}(\bar{I}) = u_0 + \rho_0,$$

with $p(u) = C\tilde{p}(u/C)$ and $q(u) = C^{-1}\tilde{q}(u/C)$. $\qquad\qquad\qquad \square$

In the phase space (ξ, I), the curves obtained in Theorem 2.4.12 can be identified as the graphs of functions of the form $I_\rho(\xi; \varepsilon) \in C^1(\mathbb{R}_{/2\pi\mathbb{Z}})$: fixing $\varepsilon \in (-\bar{\varepsilon}_{\rho_0\rho_1}, \bar{\varepsilon}_{\rho_0\rho_1})$, let us consider for example the closed invariant curve of $\mathcal{F}$ of rotation number ρ_0, which
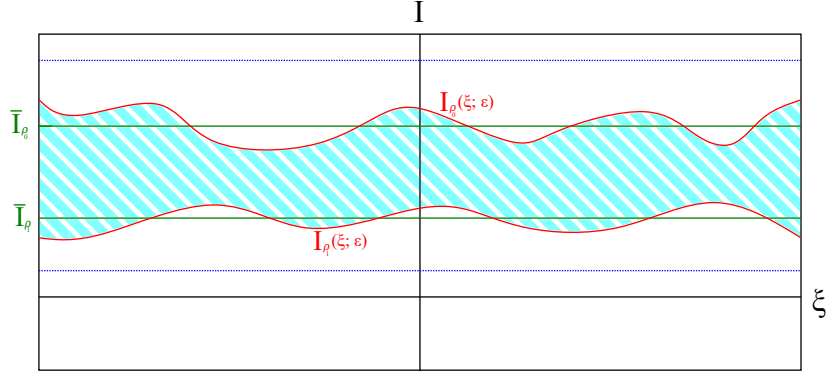
Fig. 2.8 Sketch of the perturbed dynamics in the region described by Proposition 2.4.6 and Theorem 2.4.12 in the phase plane (ξ, I). Red: the invariant curves of Diophantine rotation numbers ρ_0 and ρ_1, which are deformations of the unperturbed invariant straight lines $I = \bar{I}_{\rho_0}$, $I = \bar{I}_{\rho_1}$ (green) such that $\bar{\theta}\left(\bar{I}_{\rho_0}\right) = \rho_0$ and $\bar{\theta}\left(\bar{I}_{\rho_1}\right) = \rho_1$. In the striped region the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ is area-preserving and twist. The blue dashed lines denote two singular action values for the unperturbed dynamics (i.e. $I \in \bar{\mathcal{I}}$).

can be expressed, according to (2.4.19) and (2.4.20), as

$$\begin{cases} \xi_{\rho_0}(u; \varepsilon) = u + p(u; \varepsilon) \\ I_{\rho_0}(u; \varepsilon) = \bar{I} + q(u; \varepsilon) \end{cases} \quad \text{with} \quad \bar{\theta}(\bar{I}) = \rho_0. \tag{2.4.24}$$

From the boundedness of p asserted in (2.4.21), if σ is small enough (e.g. $\sigma < 1$) the quantity

$$\partial_u \xi(u; \varepsilon) = 1 + \partial_u p(u; \varepsilon) \tag{2.4.25}$$

is always positive: one can then invert the first equation in (2.4.24) obtaining $u(\xi)$, which is differentiable. As a consequence, one can parametrize the curve (2.4.24) as the graph of the C^1 function

$$I_{\rho_0} : \mathbb{R}_{/2\pi\mathbb{Z}} \to \mathbb{R}, \quad I_{\rho_0}(\xi; \varepsilon) = I_{\rho_0}(u(\xi); \varepsilon). \tag{2.4.26}$$

Remark 2.4.13. *Taking σ sufficiently small and a suitable $\bar{\varepsilon}_{\rho_0\rho_1}$, one can find invariant curves of $\mathcal{F}$ which are arbitrarily close to the unperturbed orbits $\mathbb{R}_{/2\pi\mathbb{Z}} \times \{\bar{I}\}$ in the plane (ξ, I). Then, as $\varepsilon \to 0$, the functions $I_\rho(\xi; \varepsilon)$ which define the invariant curves in the perturbed phase space tend in norm $C^1(\mathbb{R}_{/2\pi\mathbb{Z}})$ to the constant functions $\bar{I}_\rho$ with $\bar{\theta}(\bar{I}_\rho) = \rho$.*
Moreover, Theorem 2.4.9 can be extended to negative twist maps, leading to the existence result for invariant curves as stated in Theorem 2.1.2

As sketched in Figure 2.8, the presence of two invariant curves with irrational rotation number leads to a *confinement* in the dynamics of $\mathcal{F}(\xi_0, I_0; \varepsilon)$, where, in view of Proposition (2.4.6), the map is area-preserving and twist. More precisely, the set

$$\mathcal{A} = \left\{ (\xi, I) \in \mathbb{R}_{/2\pi\mathbb{Z}} \times \mathbb{R} \mid I_{\rho_0}(\xi; \varepsilon) \leq I \leq I_{\rho_1}(\xi; \varepsilon) \right\} \tag{2.4.27}$$

is invariant under $\mathcal{F}_\varepsilon$, as well as its boundaries.

In the unperturbed dynamics, the existence of periodic orbits of any rotation number in a suitable interval is the simple consequence of the continuity of the total shift $\bar\theta(I)$; when $\varepsilon \neq 0$, one can not take advantage of the explicit formulation of the perturbed map, then this strategy is no longer suitable. Nevertheless, the broad properties of the map, such as its area-preserving one and the existence of the invariant curves ensured by Theorem 2.4.12, enable the use of more sophisticated topological results, where the existence of orbits with prescribed rotation number is proved under more general assumptions: this is the case of the Poincaré-Birkhoff theorem, here presented in the version of [25].

Theorem 2.4.14 (Poincaré-Birkhoff). *Let $\mathcal{F}$ an area preserving map on the annulus $\mathbb{R}_{/2\pi\mathbb{Z}} \times [c,d]$ which preserves the boundaries and $\mathcal{F}_\mathbb{R}$ its lift over $\mathbb{R} \times [c,d]$. Suppose that $\mathcal{F}_\mathbb{R}$ satisfies the boundary twist condition, that is, the restrictions of $\mathcal{F}_\mathbb{R}$ to each boundary component $u_- = \mathbb{R} \times \{c\}$ and $u_+ = \mathbb{R} \times \{d\}$ have rotation numbers $\rho_\pm$ with $\rho_- < \rho_+$ (the case $\rho_+ < \rho_-$ is analogous). If $2\pi\dfrac{m}{n} \in [\rho_-, \rho_+]$ and m,n are coprime, then $\mathcal{F}$ has at least two $(m,n)-$orbits.*

Remark 2.4.15. *Theorem 2.4.14 can be extended to area-preserving maps which preserve invariant strips in $\mathbb{R}_{/2\pi\mathbb{Z}} \times \mathbb{R}$ whose boundaries are fixed by $\mathcal{F}$ and are graphs of C^1 functions over the $\xi-$axis. Let us take the set $\mathcal{A}$ defined in 2.4.27 and consider $\Phi_0(\xi;\varepsilon), \Phi_1(\xi;\varepsilon), \bar{I}_0, \bar{I}_1$ such that*

$$\partial_\xi \Phi_0(\xi;\varepsilon) = I_{\rho_0}(\xi;\varepsilon), \quad \partial_\xi \Phi_1(\xi;\varepsilon) = I_{\rho_1}(\xi;\varepsilon), \quad \bar\theta(\bar{I}_{\rho_0}) = \rho_0, \quad \bar\theta(\bar{I}_{\rho_1}) = \rho_1,$$

choosing Φ_1 and Φ_0 such that $\Phi_1(\xi;0) = \bar{I}_1\xi$ and $\Phi_0(\xi;0) = \bar{I}_0\xi$. For ε sufficiently small, consider the quantity

$$A(\varepsilon) = \int_0^{2\pi} I_{\rho_1}(\xi;\varepsilon) - I_{\rho_0}(\xi;\varepsilon)d\xi = \Phi_1(2\pi;\varepsilon) - \Phi_0(2\pi;\varepsilon) - (\Phi_1(0;\varepsilon) - \Phi_0(0;\varepsilon)) > 0,$$

and, noted that $A(0) = \left(\bar{I}_1 - \bar{I}_0\right)2\pi$, define the change of coordinates

$$\Psi(\xi, I;\varepsilon) = \begin{cases} \xi' = \dfrac{2\pi}{A(\varepsilon)}\left(\Phi_1(\xi;\varepsilon) - \Phi_0(\xi;\varepsilon)\right) \\ I' = \dfrac{A(\varepsilon)}{2\pi}\left(\dfrac{I - I_{\rho_0}(\xi;\varepsilon)}{I_{\rho_1}(\xi;\varepsilon) - I_{\rho_0}(\xi;\varepsilon)} + \dfrac{\bar{I}_0}{\bar{I}_1 - \bar{I}_0}\right). \end{cases}$$

From direct computations, one has that:

(i) *$\Psi(\xi, I;\varepsilon)$ is C^1 in all its variables and for every fixed $\varepsilon \in \mathbb{R}$ $\det\left(D_{(\xi;I)}\Psi(\xi, I;\varepsilon)\right) = 1$: hence, Ψ defines a canonical change of variables;*

(ii) *for $\varepsilon = 0$, $\Psi(\xi, I;0) = Id$;*

(iii) Ψ maps the horizontal boundaries of $\mathcal{A}$, that is, $\{(\xi, I_{\rho_0}(\xi; \varepsilon)) \mid \xi \in \mathbb{R}_{/2\pi\mathbb{Z}}\}$ and $\{(\xi, I_{\rho_1}(\xi; \varepsilon)) \mid \xi \in \mathbb{R}_{/2\pi\mathbb{Z}}\}$ respectively into the straight lines

$$I = I_0' = \frac{A(\varepsilon)}{2\pi} \frac{\bar{I}_0}{\bar{I}_1 - \bar{I}_0} > 0 \quad and \quad I = I_1' = \frac{A(\varepsilon)}{2\pi} \frac{\bar{I}_1}{\bar{I}_1 - \bar{I}_0} > I_0';$$

(iv) for every $\xi \in \mathbb{R}_{/2\pi\mathbb{Z}}$ and every fixed ε, one has

$$\Phi_1(\xi + 2\pi; \varepsilon) - \Phi_0(\xi + 2\pi; \varepsilon) = \Phi_1(\xi; \varepsilon) - \Phi_0(\xi; \varepsilon) + A(\varepsilon),$$

and then

$$\xi'(\xi + 2\pi) = \frac{2\pi}{A(\varepsilon)} \left(\Phi_1(\xi + 2\pi; \varepsilon) - \Phi_0(\xi + 2\pi; \varepsilon) \right) =$$

$$= \frac{2\pi}{A(\varepsilon)} \left(\Phi_1(\xi; \varepsilon) - \Phi_0(\xi; \varepsilon) + A(\varepsilon) \right) = \xi'(\xi) + 2\pi;$$

(v) ξ' is strictly increasing in ξ, while I' is $2\pi-$periodic in ξ.

Globally, Ψ maps the $\mathcal{F}_\varepsilon$invariant set $\mathcal{A}$ into the straight line $\mathcal{B} = \mathbb{R}_{/2\pi\mathbb{Z}} \times [I_0', I_1']$ preserving the orientation and the boundaries. One can then consider the map $\bar{\mathcal{F}}_\varepsilon : \mathcal{B} \to \mathcal{B}$ such that $\bar{\mathcal{F}}_\varepsilon \circ \Psi = \Psi \circ \mathcal{F}_\varepsilon$, namely, such that $\bar{\mathcal{F}}_\varepsilon = \Psi \circ \mathcal{F}_\varepsilon \circ \Psi^{-1}$. It can be proved that $\bar{\mathcal{F}}_\varepsilon$ preserves the rotation number of the corresponding orbits of $\mathcal{F}$: for a $(m, n)-$periodic orbit, it is a simple consequence of (iv), as, taken $\{(\xi_k, I_k)\}_{k\in\mathbb{N}}$ $(m, n)-$periodic for $\mathcal{F}_\varepsilon$ and defined for every $k \in \mathbb{N}$ $(\xi_k', I_k') = \Psi(\xi_k, I_k)$, one has

$$(\xi_{k+n}', I_{k+n}') = \Psi(\xi_{k+n}, I_{k+n}) = \Psi(\xi_k + 2\pi m, I_k) = (\xi_k' + 2\pi m, I_k').$$

Let us now take a $\mathcal{F}_\varepsilon-$orbit with rotation number $\rho \in \mathbb{R}$ parametrized, according to Mather's definition in [24], by $(\xi_k, I_k) = (\psi_1(t_k), \psi_2(t_k))$ such that

$$t_{k+1} = t_k + \rho, \quad \mathcal{F}_\varepsilon(\xi_k, I_k) = (\xi_{k+1}, I_{k+1}) = (\psi_1(t_k + \rho), \psi_2(t_k + \rho)),$$
$$(\psi_1(t + 2\pi), \psi_2(t + 2\pi)) = (\psi_1(t) + 2\pi, \psi_2(t)),$$

with $\psi_1 : \mathbb{R} \to \mathbb{R}$ a weakly order preserving map (not necessarily continuous). Now, setting $(\xi_k', I_k') = \Psi(\xi_k, I_k) = \Psi(\psi_1(t_k), \psi_2(t_k))$, one has

$$\bar{\mathcal{F}}_\varepsilon(\xi_k', I_k') = \Psi \circ F(\xi_k, I_k) = \Psi(\xi_{k+1}, I_{k+1}) = (\xi_{k+1}', I_{k+1}'),$$

and, defined $(\tilde{\psi}_1(t), \tilde{\psi}_2(t)) = \Psi(\psi_1(t), \psi_2(t))$,

$$(\tilde{\psi}_1(t + 2\pi), \tilde{\psi}_2(t + 2\pi)) = (\tilde{\psi}_1(t) + 2\pi, \tilde{\psi}_2(t)),$$

leading to the conclusion that the $\bar{\mathcal{F}}_\varepsilon-$orbit $\{(\xi_k', I_k')\}_{k\in\mathbb{N}}$ has rotation number ρ. Note that for ε sufficiently small $\bar{\mathcal{F}}_\varepsilon$ satisfied the hypotheses of Theorem 2.4.14, as for $\varepsilon = 0$ the identity map is trivially twist. As the preservation of the rotation number holds also for Ψ^{-1}, given a twist map on invariant sets of the type $\mathcal{A}$ which preserves the horizontal boundaries one can pass to the strip $\mathcal{B}$ and use Theorem 2.4.14 to prove the existence of $(m, n)-$periodic orbits for $\bar{\mathcal{F}}_\varepsilon$; returning then to the map $\mathcal{F}_\varepsilon$, this translates to the existence of $(m, n)-$periodic orbits for the original map.

Making use of Theorem 2.4.14, one can prove, under suitable conditions on the perturbation, the existence of periodic orbits for the dynamics induced by the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ with $\varepsilon \neq 0$. We recall that in Section 2.3 we denoted with $\mathcal{I} \backslash \bar{\mathcal{I}}$ the set of well definition of the unperturbed map $\mathcal{F}(\xi_0, I_0; 0)$ and we proved that it is the finite union of open intervals in $\mathbb{R}$. In particular, the set of the singular points $\bar{\mathcal{I}}$ is composed by the critical points of the C^1 function $\bar{\theta}(I)$ (see Proposition 2.3.7), and one can set

$$\mathcal{I} \backslash \bar{\mathcal{I}} = \bigcup_{i=1}^{N} A_i$$

with $N > 0$ (possibly $N = 1$) and A_i open intervals in $\mathcal{I}$. In the following, to ensure the good definition of the perturbed map in a compact set, a finite union of closed intervals in $\mathcal{I} \backslash \bar{\mathcal{I}}$ will be fixed, : in particular, we fix $a_i, b_i \in \mathcal{I}$ such that $\forall i \in \{1, \ldots, N\}$

$$[a_i, b_i] \subset A_i \text{ and, if } \bar{\theta}_- = \min_i \left\{ \bar{\theta}(a_i), \bar{\theta}(b_i) \right\}, \ \bar{\theta}_+ = \max_i \left\{ \bar{\theta}(a_i), \bar{\theta}(b_i) \right\},$$

$$\bar{\theta}_-^{(i)} = \min\{\bar{\theta}(a_i), \bar{\theta}(b_i)\}, \text{ and } \bar{\theta}_+^{(i)} = \max\{\bar{\theta}(a_i), \bar{\theta}(b_i)\}, \text{ one has} \qquad (2.4.28)$$

$$\bigcup_{i=1}^{N} \left[\bar{\theta}_-^{(i)}, \bar{\theta}_+^{(i)} \right] = [\bar{\theta}_-, \bar{\theta}_+].$$

Note that, by the continuity of $\bar{\theta}$, such sets $\{a_i\}_{i=1}^N$, $\{b_i\}_{i=1}^N$ exist.

Proposition 2.4.16. *Let $a_i, b_i \in \mathcal{I}$ as in (2.4.28), and fix $\rho_{\pm}^{(i)} \in \tilde{\mathcal{D}}_2$ such that for every $i = 1, \ldots, N$ one has $\bar{\theta}_-^{(i)} < \rho_-^{(i)} < \rho_+^{(i)} < \bar{\theta}_+^{(i)}$. Then there exists $\bar{\varepsilon} > 0$ such that for every $\varepsilon \in \mathbb{R}$, $|\varepsilon| < \bar{\varepsilon}$, and for every $m, n \in \mathbb{Z}$ coprime, $n > 0$, with $2\pi \dfrac{m}{n} \in (\rho_-^{(i)}, \rho_+^{(i)})$ for some $i \in \{1, \ldots, N\}$, the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ admits at least $2k$ (m,n)-orbits, where k is the number of the pairs $(\rho_-^{(i)}, \rho_+^{(i)})$ such that $\rho_-^{(i)} < 2\pi \dfrac{m}{n} < \rho_+^{(i)}$.*

Proof. According to Theorem 2.4.12, for every pair $\rho_-^{(i)}, \rho_+^{(i)}$ there is $\bar{\varepsilon}_{\rho_{\pm}}^{(i)}$ such that for every $|\varepsilon| < \bar{\varepsilon}_{\rho_{\pm}}^{(i)}$ the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ admits two orbits of rotation numbers $\rho_-^{(i)}$ and $\rho_+^{(i)}$. Moreover, the perturbed map is area-preserving and twist between these two orbits. Setting $\bar{\varepsilon} = \min_{i \in \{1, \ldots, N\}}$, one has that if ε is such that $|\varepsilon| < \bar{\varepsilon}$ all the orbits of rotation numbers $\rho_{\pm}^{(i)}$ are preserved and the perturbed map in between is well defined and area-preserving.

Fixing $\varepsilon \in \mathbb{R}$, $|\varepsilon| < \bar{\varepsilon}$, if m, n are such that $2\pi \dfrac{m}{n} \in \left(\rho_-^{(i)}, \rho_+^{(i)} \right)$, then by Theorem 2.4.14 the perturbed map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ admits at least 2 (m,n)-orbits. As this reasoning can be repeated whenever a pair $(\rho_-^{(i)}, \rho_+^{(i)})$ is such that $\rho_-^{(i)} < 2\pi \dfrac{m}{n} < \rho_+^{(i)}$, the claim is true. $\qquad \square$

Proposition 2.4.16 claims the existence of a unique threshold value of ε under which the presence of periodic orbits of prescribed rotation numbers in a certain set is guaranteed. Another slightly different approach is proposed in Proposition 2.4.17, where, fixed m, n such that $2\pi \dfrac{m}{n}$ lies in a suitable interval which does not depend on

prefixed boundary rotation numbers $\rho_\pm^{(i)}$, one can find a threshold $\bar{\varepsilon}_{mn}$, depending on m, n, such that for every $|\varepsilon| < \bar{\varepsilon}_{mn}$ the presence of the corresponding (m, n)-orbit is ensured.

Proposition 2.4.17. *Given* $\{a_i\}_{i=1}^N$, $\{b_i\}_{i=1}^N$, $\{\bar{\theta}_+^{(i)}\}_{i=1}^N$, $\{\bar{\theta}_-^{(i)}\}_{i=1}^N$ *as in (2.4.28), let* $m, n \in \mathbb{Z}$ *coprime,* $n > 0$, *such that* $2\pi\dfrac{m}{n} \in \left(\bar{\theta}_-^{(i)}, \bar{\theta}_+^{(i)}\right)$ *for some* $i \in \{1, \ldots, N\}$. *Then* $\exists \bar{\varepsilon}_{mn} > 0$ *such that for every* $\varepsilon \in \mathbb{R}$, $|\varepsilon| < \bar{\varepsilon}_{mn}$ *the map* $\mathcal{F}(\xi_0, I_0; \varepsilon)$ *admits at least* $2k$ (m, n)-*orbits, where* k *is the number of intervals* (a_i, b_i) *such that* $2\pi\dfrac{m}{n}$ *is between* $\bar{\theta}_-^{(i)}$ *and* $\bar{\theta}_+^{(i)}$.

Proof. By the density of $\tilde{\mathcal{D}}_2$ in every bounded interval, one can find $\rho_\pm \in \tilde{\mathcal{D}}_2$ with $\bar{\theta}_-^{(i)} < \rho_- < 2\pi\dfrac{m}{n} < \rho_+ < \bar{\theta}_+^{(i)}$. From Theorem 2.4.12, one can find $\bar{\varepsilon}_{mn} = \bar{\varepsilon}_{\rho_\pm}$ such that, if ε is such that $|\varepsilon| < \bar{\varepsilon}_{mn}$, the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ admits two orbit with rotation numbers $\rho_\pm$, and it is area-preserving between them. Applying again Theorem 2.4.14, the claim follows. $\square$

Extending the discussion beyond periodic orbits, one may search for more general class of invariant sets. KAM theory allowed us to claim the persistence of orbits with Diophantine rotation numbers within certain ranges, while Poincaré-Birkhoff theorem extended the existence result to periodic number with 2π-rational numbers between them. The Aubry-Mather theory allows to move further, providing the existence of orbits of the perturbed map of every prescribed rotation number in suitable subsets of $\mathbb{R}$.

Theorem 2.4.18 (Aubry-Mather on the compact annulus). *Let* $\mathcal{F}$ *an area and orientation- preserving twist homeomorphism of the annulus* $\mathbb{R}_{/2\pi\mathbb{Z}} \times [a, b]$ *which preserves* $\mathbb{R}_{/2\pi\mathbb{Z}} \times \{a\}$ *and* $\mathbb{R}_{/2\pi\mathbb{Z}} \times \{b\}$, *and define* ρ_a *and* ρ_b *as the rotation numbers of the two boundary components. Then for every* $\rho \in [\rho_a, \rho_b]$ *there exists at least an orbit for* $\mathcal{F}$ *with rotation number* ρ. *In particular:*

- *if* $\rho = m/n \in \mathbb{Q}$, *such orbit is periodic of period* n;

- *if* $\rho \notin \mathbb{Q}$, *the orbit rotates either on a closed continuous curve or on a Cantor set.*

In any case, the orbits with the same rotation number belong to a common invariant set Γ_ρ, *called Mather set, which is a subset of the graph of a Lipschitz-continuous function over the* $\xi-$*axis.*

We refer to [23, 77, 24] for the definition of Mather set and for a thorough discussion on the Aubry-Mather theory.

Remark 2.4.19. *As in the case of Poincaré-Birkhoff Theorem 2.4.14, with the same reasoning also Aubry-Mather Theorem can be extended to maps on invariant sets of the type* $\mathcal{A}$ *defined in 2.4.27.*

Making use of the same arguments used in the proofs of Propositions 2.4.16 and 2.4.17, one can take advantage of Theorem 2.4.18 to state these existence results in a more general way.

Theorem 2.4.20. *Let $\{a_i\}_{i=1}^N$, $\{b_i\}_{i=1}^N$, $\{\bar{\theta}_-^{(i)}\}_{i=1}^N$, $\{\bar{\theta}_+\}_{i=1}^N$ as in (2.4.28). Then:*

- *letting $\rho_\pm^{(i)} \in \mathcal{D}$ as in Proposition 2.4.16, there exists $\bar{\varepsilon} > 0$ such that for every $\varepsilon \in \mathbb{R}$ with $|\varepsilon| < \bar{\varepsilon}$ and for every $\rho \in \mathbb{R}$ such that $\rho \in [\bar{\rho}_-^{(i)}, \bar{\rho}_+^{(i)}]$ for some $i \in \{1, \dots, N\}$ the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ admits k orbits with rotation number ρ, with k defined as in Proposition 2.4.16;*

- *for every $\rho \in \mathbb{R}$ such that $\rho \in [\bar{\theta}_-^{(i)}, \bar{\theta}_+^{(i)}]$ for some $i \in \{1, \dots, N\}$ there is $\bar{\varepsilon}_\rho > 0$ such that for every ε with $|\varepsilon| < \bar{\varepsilon}_\rho$ the map $\mathcal{F}(\xi_0, I_0; \varepsilon)$ admits k orbits with rotation number ρ, where k is defined as in Proposition 2.4.17.*

In both cases, if $\rho = 2\pi \dfrac{m}{n}$ then for ε sufficiently small there are at least $2k$ (m, n)-orbits, where k is defined suitably according to the cases.

2.4.3 Caustics for the perturbed case

The persistence of invariant curves with Diophantine rotation numbers ensured by the KAM theorem has important consequences for the existence of caustics in the perturbed dynamics. As a matter of fact, for such invariant tori (which are dense in the phase space) it is possible to find, although not explicitly, the inner and outer caustics also for small perturbations of the circular domain D_0.

Theorem 2.4.21. *Let $\xi_0 \in [0, 2\pi]$, $I_0 \in \mathcal{I}\backslash\bar{\mathcal{I}}$ such that $\theta(I_0) \in \tilde{\mathcal{D}}_2$. Then there exists $\bar{\varepsilon} > 0$ such that for every $|\varepsilon| < \bar{\varepsilon}$ there are $\Gamma_E(\xi; \varepsilon, \theta(I_0))$, $\Gamma_I(\xi; \varepsilon, \theta(I_0))$ respectively outer and inner caustics related to the perturbed orbit of rotation number $\theta(I_0)$.*

The proof of Theorem 2.4.21 relies on showing that, for ε small enough, system (2.3.38) evaluated both for the outer and inner dynamics admits a unique solution for each $\xi \in [0, 2\pi]$, which defines a regular and closed curve. To prove that, it is worthwhile to derive the form of $G_{E\backslash I}(x, y; \xi)$ for a perturbed domain.

Outer dynamics Let us consider $\xi \in [0, 2\pi]$, $p_0 = \gamma_\varepsilon(\xi)$ and $v_0 \in \mathbb{R}^2$ such that $\|v_0\| = \sqrt{2V_E(p_0)}$ and $\alpha = \angle(p_0, v_0) \in (-\pi/2, \pi/2)$.
To fix the notation, recall the definition ^{4} of $\gamma_\varepsilon(\xi) = (1 + \varepsilon f(\xi, \varepsilon))e^{i\xi} = \rho(\xi; \varepsilon)e^{i\xi}$: as the perturbation of the circle is only in the normal direction, the curve's parameter ξ still represents the polar angle of the point $\gamma_\varepsilon(\xi)$. We want to find the Cartesian equation of the outer elliptic arc of initial conditions p_0 and v_0.

---

^{4}Note that, in this section, the quantity ρ represents the radial disttance of our points, and has nothing to do with the rotation number of a trajectory.

Following the same reasoning of Appendix 2.3.1 and denoting with $(x(s), y(s))$ the parametrization of such ellipse, its maximal and minimal distances from the origin can be then computed as

$$
\begin{aligned}
a^2 &= \max_{s\in[0,2\pi/\omega]} r^2(s) = A + \frac{\mathcal{E}}{\omega^2} = \frac{\mathcal{E} + \sqrt{(\mathcal{E} - \omega^2\|p_0\|^2)^2 + \omega^2(p_0 \cdot v_0)^2}}{\omega^2} \\
b^2 &= \min_{s\in[0,2\pi/\omega]} r^2(s) = -A + \frac{\mathcal{E}}{\omega^2} = \frac{\mathcal{E} - \sqrt{(\mathcal{E} - \omega^2\|p_0\|^2)^2 + \omega^2(p_0 \cdot v_0)^2}}{\omega^2} :
\end{aligned}
\tag{2.4.29}
$$

in the reference frame $R(O, x'', y'')$ whose axes coincide with the ellipse's ones the latter is then implicitly defined by the equation

$$
\frac{x''^2}{a^2} + \frac{y''^2}{b^2} = 1.
$$

Let us now search for the angle $\bar{\beta}$ such that the rotated ellipse

$$
\frac{(x'\cos\bar{\beta} + y'\sin\bar{\beta})^2}{a^2} + \frac{(y'\cos\bar{\beta} - x'\sin\bar{\beta})^2}{b^2} = 1
$$

passes from $p_0 = \|p_0\|(1,0)$ in $R(O, x', y')$: one has to solve the equation

$$
\frac{\|p_0\|^2 \cos^2\bar{\beta}}{a^2} + \frac{\|p_0\|^2 \sin^2\bar{\beta}}{b^2} - 1 = 0 \Rightarrow \sin^2\bar{\beta} = \frac{b^2}{a^2 - b^2}\left(\frac{a^2}{\|p_0\|^2} - 1\right) \geq 0.
$$

Denoting by (v'_x, v'_y) the components of v_0 in $R(O, x', y')$, one has that

$$
\sin\bar{\beta} = \begin{cases} -\dfrac{b}{\sqrt{a^2 - b^2}}\sqrt{\dfrac{a^2}{|p_0|^2} - 1} & \text{if } v'_y < 0 \\[2ex] \dfrac{b}{\sqrt{a^2 - b^2}}\sqrt{\dfrac{a^2}{|p_0|^2} - 1} & \text{if } v'_y > 0 \end{cases} \Rightarrow \cos\bar{\beta} = \frac{a}{\sqrt{a^2 - b^2}}\sqrt{1 - \frac{b^2}{\|p_0\|^2}}.
$$

$$\tag{2.4.30}$$

Returning to the original frame $R(O, x, y)$, one can then retrieve the Cartesian equation of the outer arc as

$$
G_E(x, y; \xi, \varepsilon) = \frac{\left(x\cos(\xi + \bar{\beta}) + y\sin(\xi + \bar{\beta})\right)^2}{a^2} + \frac{\left((y\cos(\xi + \bar{\beta}) - x\sin(\xi + \bar{\beta})\right)^2}{b^2} - 1 = 0.
$$

$$\tag{2.4.31}$$

Note that, although not explicitly indicated, the quantities a, b and $\bar{\beta}$ depend on ϵ. Once obtained the general expression for an ellipse of initial conditions p_0 and v_0, we shall return to the framework of our perturbed problem. Let us then consider $I_0 \in \mathcal{I}\backslash\bar{\mathcal{I}}$ such that $\theta(I_0)$ is Diophantine: from Theorem 2.4.9 there exists $\bar{\varepsilon}^{(1)} > 0$ such that, if $|\varepsilon| < \bar{\varepsilon}^{(1)}$, we can define $I(\xi; \varepsilon)$ invariant curve in the plane (ξ, I) for the perturbed map $\mathcal{F}_\varepsilon$ such that $I(\xi; 0) \equiv I_0$ and with rotation number $\theta(I_0)$. Moreover, $I(\xi; \varepsilon)$ is continuous in ε and differentiable in ξ, with $\partial_\xi I(\xi; \varepsilon)$ continuous in ε: as a consequence, since $\theta(I_0) \in \tilde{\mathcal{D}}_2$ implies $I_0 \neq 0$, possibly reducing $\bar{\varepsilon}^{(1)}$ one can assume that $I(\xi; \varepsilon)$ has

always the same sign of I_0.

For the caustic of the orbit associated to $(\xi, I(\xi; \varepsilon))$ to be well defined, it is necessary that the system

$$\begin{cases} G_E(x, y; \xi, \varepsilon) = 0 \\ \partial_\xi G_E(x, y; \xi, \varepsilon) = 0 \end{cases} \tag{2.4.32}$$

defines implicitly a unique curve $\Gamma_E(\xi; \varepsilon)$ for $\xi \in [0, 2\pi]$, that is, that x and y can be expressed as functions of (ξ, ε) globally defined for $\xi \in [0, 2\pi]$. As already pointed out in Section 2.3.5, from the implicit function theorem the local existence of $\Gamma_E(\xi; \varepsilon)$ is then ensured by requiring the nondegeneracy condition

$$\nabla_{(x,y)} G_E(x, y; \xi, \varepsilon) \nparallel \nabla_{(x,y)} \partial_\xi G_E(x, y; \xi, \varepsilon) \tag{2.4.33}$$

on the solutions of (2.4.32).

Lemma 2.4.22. *If $I_0 \in \mathcal{I} \backslash \bar{\mathcal{I}}$ is such that $\theta(I_0)$ is Diophantine, then there is $\bar{\varepsilon}^{(2)} > 0$ such that, if $|\varepsilon| < \bar{\varepsilon}^{(2)}$, then $G_E(x, y; \xi, \varepsilon)$ is continuous in ε, differentiable in ξ and such that $\partial_\xi G_E(x, y; \xi, \varepsilon)$ is continuous in ε.*

Proof. Recalling (2.4.31), the proof of the Lemma relies on showing that all the quantities involved in the definition of $G_E(x, y; \xi, \varepsilon)$, namely, $a^{-2}, b^{-2}, \cos \bar{\beta}$ and $\sin \bar{\beta}$ are continuous in ε, differentiable in ξ and with derivative continuous in ε, provided the latter is small enough.

Starting from a^{-2} and b^{-2}, from (2.4.29) it is clear that the expression of $p_0 \cdot v_0$ as a function of ξ and ε is needed. Recalling the definition (2.4.1), denoted with $t(\xi; \varepsilon)$ and $n_e(\xi; \varepsilon)$ the tangent and the outward-pointing normal unit vectors to γ_ε in p_0, one has that

$$v_0 = \sqrt{2\mathcal{E} - \omega^2 \rho(\xi; \varepsilon)^2} \left(\cos \alpha \, n_e(\xi; \varepsilon) + \sin \alpha \, t(\xi; \varepsilon) \right).$$

Expliciting $\cos \alpha, \sin \alpha, t(\xi; \varepsilon), n_e(\xi; \varepsilon)$ and setting for simplicity $\rho \equiv \rho(\xi; \varepsilon)$, $\rho' \equiv d\rho(\xi; \varepsilon)/d\xi$ and $I(\xi; \varepsilon) \equiv I$, one obtains

$$v_0 = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

$$= \frac{1}{\sqrt{\rho^2 + \rho'^2}} \begin{pmatrix} \sqrt{2}I \left(\rho' \cos \xi - \rho \sin \xi \right) + \sqrt{2\mathcal{E} - \omega^2 \rho^2 - 2I^2} \left(\rho' \sin \xi + \rho \cos \xi \right) \\ \sqrt{2}I \left(\rho' \sin \xi + \rho \cos \xi \right) + \sqrt{2\mathcal{E} - \omega^2 \rho^2 - 2I^2} \left(\rho \sin \xi - \rho' \cos \xi \right) \end{pmatrix}$$

$$\Rightarrow p_0 \cdot v_0 = \frac{\rho(\xi; \varepsilon)}{\sqrt{\rho^2(\xi; \varepsilon) + \rho'^2(\xi; \varepsilon)}}$$

$$\times \left(\rho(\xi; \varepsilon) \sqrt{2\mathcal{E} - \omega^2 \rho^2(\xi; \varepsilon) - 2I^2(\xi; \varepsilon)} + \sqrt{2}I(\xi, \varepsilon) \rho'(\xi; \varepsilon) \right),$$

$$\tag{2.4.34}$$

which has the desired continuity and differentiability properties provided ε is small enough. This implies that a^2 and b^2 have the same properties. Moreover, it is trivial that $a^2 > 0$ and, since for $\varepsilon = 0$

$$b^2_{|\varepsilon=0} = \frac{\mathcal{E} - \sqrt{\mathcal{E}^2 - 2\omega^2 I_0^2}}{\omega^2} > 0,$$

by the continuity of b with respect to ε we have also $b^2 > 0$ for ε small enough. Applying the same reasoning, we can infer $\sqrt{a^2 - b^2} > 0$.

Going back to (2.4.30), $\cos \bar{\beta}$ is then continuous and differentiable, and the same conclusion holds for $\sin \bar{\beta}$ if one can ensure that v_y' has the same sign for all the points of the orbit $(\xi, I(\xi; \varepsilon))$. From (2.4.34), in the plane $R(O, x', y')$ one has

$$v_y' = \frac{1}{\sqrt{\rho^2(\xi; \varepsilon) + \rho'^2(\xi; \varepsilon)}} \left(\rho(\xi; \varepsilon) \sqrt{2} I(\xi; \varepsilon) - \rho'(\xi; \varepsilon) \sqrt{2\mathcal{E} - \omega^2 \rho^2(\xi; \varepsilon) - 2I^2(\xi; \varepsilon)} \right),$$

which for $\varepsilon = 0$ translates in

$$v_{y|\varepsilon=0}' = \sqrt{2} I_0 \neq 0.$$

Taking again advantage of the continuity of v_y' with respect ε, we can finally ensure that for ε small enough the thesis is proved. $\hspace{4cm}\square$

Proposition 2.4.23. *If $I_0 \in \mathcal{I} \backslash \bar{\mathcal{I}}$ is such that $\theta(I_0) \in \tilde{\mathcal{D}}_2$, then there exists $\bar{\varepsilon}_E$ such that for $|\varepsilon| < \bar{\varepsilon}_E$ the caustic $\Gamma_E(\xi; \varepsilon, \theta(I_0))$ is globally well defined.*

Proof. As the nondegeneracy condition (2.4.33) holds for $\varepsilon = 0$ (cfr. (2.3.42)), from Lemma 2.4.22, for every $\bar{\xi} \in [0, 2\pi]$ there exists $\bar{\varepsilon}^{(2)}(\bar{\xi})$ such that for every $|\varepsilon| < \bar{\varepsilon}^{(2)}(\bar{\xi})$ condition (2.4.33) is satisfied. By the implicit function theorem, there are $\lambda_\xi(\bar{\xi}), \lambda_\varepsilon(\bar{\xi}) > 0$ such that the curve $(x(\xi; \varepsilon), y(\xi, \varepsilon))$ solution of (2.4.32) is well defined in $R(\bar{\xi}) = \left(\bar{\xi} - \lambda_\xi(\bar{\xi}), \bar{\xi} + \lambda_\xi(\bar{\xi}) \right) \times \left(-\lambda_\varepsilon(\bar{\xi}), \lambda_\varepsilon(\bar{\xi}) \right)$. For the uniqueness of the solution, if $\bar{\xi}_1$ and $\bar{\xi}_2$ are such that $R(\bar{\xi}_1) \cap R(\bar{\xi}_2) \neq \emptyset$, the curve coincides in such intersection. As $[0, 2\pi]$ is compact, it is possible to find $N > 0$, $\{\bar{\xi}_1, \dots \bar{\xi}_N\} \subset [0, 2\pi]$ such that

$$[0, 2\pi] \subset \bigcup_{i=1}^{N} \left(\bar{\xi}_i - \lambda_\xi(\bar{\xi}_i), \bar{\xi}_i + \lambda_\xi(\bar{\xi}_i) \right),$$

then, setting

$$\bar{\varepsilon}_E = \min_{i \in \{1, \dots, N\}} \lambda_\varepsilon(\bar{\xi}_i),$$

for every $\varepsilon > 0$ such that $|\varepsilon| < \bar{\varepsilon}_E$ the curve $\Gamma_E(\xi; \varepsilon, I_0) = (x(\xi; \varepsilon), y(\xi; \varepsilon))$ is globally well defined in $[0, 2\pi]$. $\hspace{2cm}\square$

Inner caustics Following the same reasoning applied for the outer caustic, let us consider the inner problem

$$\begin{cases} z''(s) = -\dfrac{\mu}{\|z(s)\|^3} z(s), & s \in [0, T_I] \\[2mm] \dfrac{1}{2} \|z'(s)\|^2 - \mathcal{E} - h - \dfrac{\mu}{\|z(s)\|} = 0 & s \in [0, T_I] \\[2mm] z(0) = p_0, \quad z'(0) = v_0 \end{cases}$$

by fixing $p_0 = \|p_0\| e^{i\xi}$, $v_0 = \sqrt{2(\mathcal{E} + h + \mu/\|p_0\|)} e^{i\theta_v}$ such that $\theta_v - \xi \in (\pi/2, 3\pi/2)$. This last assumption, which is done to guarantee that the hyperbola points inward a circle of radius $\|p_0\|$, can be ensured for ε small enough and suitable bounds on $I(\xi)$. Rotating again the reference frame $R(O, x, y)$ by an angle $-\xi$, we obtain $R(O, x', y')$

such that $p_0 = \|p_0\|(1,0)$.

Recalling (2.3.44), in the reference frame $R(O, x'', y'')$ were the hyperbola's pericenter lies on the positive half of the x-axis, its Cartesian equation is given by:

$$(e^2 - 1)x''^2 - y''^2 - 2pex'' + p^2 = 0 \text{ with } x \leq \frac{p}{e+1},$$

where

$$p = \frac{k^2}{\mu}, e = \frac{\sqrt{\mu^2 + 2(\mathcal{E} + h)k^2}}{\mu}, \quad k = \|p_0 \wedge v_0\|.$$

To find the corresponding equation in the reference frame $R(O, x', y')$, one can search again for the angle $\bar{\delta}$ such that the arc defined by

$$(e^2 - 1)(x' \cos \bar{\delta} + y' \sin \bar{\delta})^2 - (y' \cos \bar{\delta} - x' \sin \bar{\delta})^2 - 2ep(x' \cos \bar{\delta} + y' \sin \bar{\delta}) + p^2 = 0,$$
$$x' \cos \bar{\delta} + y' \sin \bar{\delta} \leq \frac{p}{e+1}$$

$$(2.4.35)$$

passes from $p_0 = \|p_0\|(1,0)$. Solving (2.4.35) with $x' = \|p_0\|$ and $y' = 0$, one obtains

$$\cos \bar{\delta} = \frac{p - \|p_0\|}{e\|p_0\|},$$

which is in $[-1, 1]$ if we take non-degenerate hyperbolæ. Referring to v'_y as the vertical component of v_0 in $R(O, x', y')$, one has then

$$\sin \bar{\delta} = \begin{cases} \dfrac{(e^2 - 1)\|p_0\|^2 + 2p\|p_0\| - p^2}{e\|p_0\|} & \text{if } v'_y > 0 \\[3mm] -\dfrac{(e^2 - 1)\|p_0\|^2 + 2p\|p_0\| - p^2}{e\|p_0\|} & \text{if } v'_y < 0 \end{cases}.$$

Returning to the original reference frame $R(O, x, y)$, one obtains then the Cartesian equation for the inner Keplerian arc

$$G_I(x, y; \xi\varepsilon) = (e^2 - 1)(x \cos(\bar{\delta} + \xi) + y \sin(\bar{\delta} + \xi))^2 - (y \cos(\bar{\delta} + \xi) - x \sin(\bar{\delta} + \xi))^2$$
$$- 2pe(x \cos(\bar{\delta} + \xi) + y \sin(\bar{\delta} + \xi)) + p^2 = 0$$
$$x \cos(\bar{\delta} + \xi) + y \sin(\bar{\delta} + \xi) \leq \frac{p}{e+1}.$$

$$(2.4.36)$$

Note that, with reference to the polar angles ξ and θ_v, the angle $\bar{\delta}$ can be also expressed as

$$\bar{\delta} = \frac{\sin(\theta_v - \xi)}{|\sin(\theta_v - \xi)|} \arccos\left(\frac{p - \|p_0\|}{e\|p_0\|}\right).$$

$$(2.4.37)$$

As in the case of the outer dynamics, the global good definition of the inner caustic $\Gamma_I(\xi; \varepsilon, \theta(I_0))$ depends on proving that $G_I(x, y; \xi, \varepsilon)$ differentiable in ξ and that both G_I and $\partial_\xi G_I$ are continuous in ε.

Lemma 2.4.24. *If $I_0 \in \mathcal{I}\backslash\bar{\mathcal{I}}$ is such that $\theta(I_0) \in \tilde{\mathcal{D}}_2$, then there is $\bar{\varepsilon}^{(3)} > 0$ such that, if $|\varepsilon| < \bar{\varepsilon}^{(3)}$, then $G_I(x, y; \xi, \varepsilon)$ is continuous in ε, differentiable in ξ and such that $\partial_\xi G(x, y; \xi, \varepsilon)$ is continuous in ε.*

Proof. As in the case of Lemma 2.4.22, one needs to prove the desired regularity properties on the quantities p and e, as well as $\sin\bar{\delta}$ and $\cos\bar{\delta}$. As all these quantities depend on $k = \|p_0 \wedge v_0\|$, let us find the expression of the angular momentum as a function of ξ and I. As already done in Section 2.2, let us now denote with α the angle between v_0 and the *inward*-pointing normal unit vector to γ_ε in p_0, which we indicate with $n_i(\xi)$; then referring to (2.2.9) and using the same notation of Lemma 2.4.22, we have

$$
v_0 = \sqrt{2\left(\mathcal{E} + h + \frac{\mu}{\rho}\right)} \left(\sin\alpha\, t(\xi) + \cos\alpha\, n_i(\xi)\right) =
$$

$$
= \frac{1}{\sqrt{\rho^2 + \rho'^2}} \begin{pmatrix} \sqrt{2}I\left(\rho'\cos\xi - \rho\sin\xi\right) - \sqrt{2(\mathcal{E} + h + \mu/\rho - I^2)}\left(\rho'\sin\xi + \rho\cos\xi\right) \\ \sqrt{2}I\left(\rho'\sin\xi + \rho\cos\xi\right) + \sqrt{2(\mathcal{E} + h + \mu/\rho - I^2)}\left(\rho'\cos\xi - \rho\sin\xi\right). \end{pmatrix}
$$

And, since $p_0 = \rho e^{i\xi}$,

$$
k = \|p_0 \wedge v_0\| = \frac{\sqrt{2}\rho(\xi; \varepsilon)}{\sqrt{\rho^2(\xi; \varepsilon) + \rho'^2(\xi; \varepsilon)}}
$$

$$
\times \left(I(\xi; \varepsilon)\rho(\xi; \varepsilon) + \rho'(\xi; \varepsilon)\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi; \varepsilon)} - I^2(\xi; \varepsilon)}\right)
$$

$$
p = \frac{k^2}{\mu} = \frac{2\rho^2(\xi; \varepsilon)}{\mu(\rho^2(\xi; \varepsilon) + \rho'^2(\xi; \varepsilon))}
$$

$$
\times \left(I(\xi; \varepsilon)\rho(\xi; \varepsilon) + \rho'(\xi; \varepsilon)\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi; \varepsilon)} - I^2(\xi; \varepsilon)}\right)^2
$$

$$
e = \sqrt{1 + \frac{4(\mathcal{E} + h)\rho^2(\xi; \varepsilon)}{(\rho^2(\xi; \varepsilon) + \rho'^2(\xi; \varepsilon))\mu^2}}
$$

$$
\times \left(I(\xi; \varepsilon)\rho(\xi; \varepsilon) + \rho'(\xi; \varepsilon)\sqrt{\mathcal{E} + h + \mu/\rho(\xi; \varepsilon) - I^2(\xi; \varepsilon)}\right)
$$

The regularity of p and e is then ensured whenever $\rho^2(\xi; \varepsilon) + \rho'^2(\xi; \varepsilon) \neq 0$, which is true for ε small enough. As for $\sin\bar{\delta}$ and $\cos\bar{\delta}$, from (2.4.37) one can infer that the requested regularity is ensured if $\sin(\theta_v - \xi)$ has always the same sign on the orbit $(\xi, I(\xi; \varepsilon))$. As in the case of the outer orbit, this is a consequence of the continuity of $\rho(\xi; \varepsilon)$, $\rho'(\xi; \varepsilon)$ and $I(\xi; \varepsilon)$ with respect to ε. Denoting with θ_{n_i} the polar angle of

$n_i(\xi)$, from the definition of α one has $\theta_v - \xi = \theta_{n_i} - \xi - \alpha$, and then

$$
\begin{aligned}
\sin(\theta_v - \xi) &= \sin(\theta_{n_i} - \xi)\cos\alpha + \cos(\theta_{n_i} - \xi)\sin\alpha \\
&= \frac{1}{\rho(\xi;\varepsilon)\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi;\varepsilon)}}} \\
&\qquad \times \left(\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi;\varepsilon)} - I^2(\xi;\varepsilon)}\|n_I(\xi) \wedge \gamma_\varepsilon(\xi)\| - I(\xi;\varepsilon)\gamma_\varepsilon(\xi)\cdot n_i(\xi) \right) \\
&= \frac{\rho'(\xi;\varepsilon)\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi;\varepsilon)} - I^2(\xi;\varepsilon)} + \rho(\xi;\varepsilon)I(\xi;\varepsilon)}{\sqrt{\rho^2(\xi;\varepsilon) + \rho'^2(\xi;\varepsilon)}\sqrt{\mathcal{E} + h + \frac{\mu}{\rho(\xi;\varepsilon)}}}.
\end{aligned}
$$

For $\varepsilon = 0$, $\sin(\theta_v - \xi)_{|\varepsilon=0} = I_0/\sqrt{\mathcal{E} + h + \mu} \neq 0$, then, if ε is small enough, $\sin(\theta_v - \xi)$ has always the same sign of I_0, and $\sin\bar\delta$, $\cos\bar\delta$ are differentiable in ξ and continuous in ε, with derivative continuous in ε. $\square$

Making use of Lemma 2.4.24 and following the same reasoning used in the proof of Proposition 2.4.23, it is possible to prove the existence of a well-defined inner caustic $\Gamma_I(\xi;\varepsilon,\theta(I_0))$ related to the invariant curve for the map $\mathcal{F}_\varepsilon$ with rotation number $\theta(I_0)$.

Proposition 2.4.25. *If $I_0 \in \mathcal{I}\backslash\bar{\mathcal{I}}$ is such that $\theta(I_0) \in \tilde{\mathcal{D}}_2$, then there exists $\bar\varepsilon_I$ such that for $|\varepsilon| < \bar\varepsilon_I$ the caustic $\Gamma_I(\xi;\varepsilon,\theta(I_0))$ is globally well defined.*

Chapter 3

Symbolic dynamics and analytic non-integrability for galactic billiards

3.1 Introduction

In this chapter we prove the existence of a symbolic dynamics for a class of mechanical refraction billiards (but our technique covers also also reflection ones), which symbols encode the geometry of the associated motions (for an extensive discussion about symbolic dynamics, we refer to [29, 30]). As for the reflection model (known as *Kepler billiard* as well), it refers to the mechanical billiard in which a particle moves under the influence of a fixed gravitational center and reflects elastically at the boundary of the inner region; in this case, trajectories concatenate arcs of hyperbolæ joined by boundary reflections.

Both the reflective and the refractive situations correspond to some (possibly ill-defined) area-preserving map in the cylinder. It is worth to note that reflection billiards have been extensively investigated with and without internal potentials (as general reference we quote the monographs [81, 17]).

As in the previous chapters, the configurations associated to the homothetic equilibrium trajectories play a prominent role in proving our results. For the readers' convenience, here we recall their definition, along with the one of the corresponding solutions.

Definition 3.1.1. *A* central configuration *for our dynamics is a point $\bar{P} \in \partial D$ such that*

- *$\bar{P}$ is a constrained critical point for the distance function $\|\cdot\|_{\partial D}$ to 0[1], that is, the position vector $\bar{P}$ is orthogonal to the boundary ∂D at $\bar{P}$;*

- *the half-line connecting the origin to $\bar{P}$ intersects ∂D only at $\bar{P}$.*

---
[1] We denote with $\|\cdot\|_{\partial D}$ the distance function $P \mapsto \|P\|$ restricted to the domain's boundary ∂D.
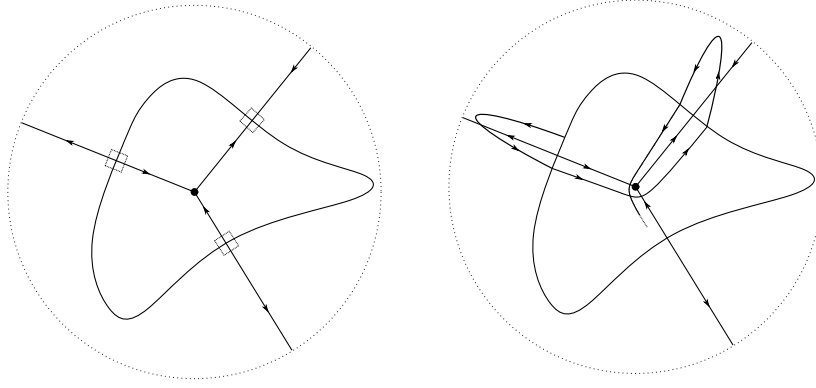
Fig. 3.1 Left: homothetic trajectories for the complete dynamics; the dashed circle denotes the boundary of the Hill's region for the outer potential. Right: trajectories for the complete dynamics in the vicinity of homothetic arcs.

A homothetic collision-ejection trajectory *is a one dimensional solution for the complete dynamics of the type* $z(t) = \lambda(t)w$ *with* $\lambda \colon \mathbb{R} \to [0, +\infty)$ *and* $w \in \mathbb{R}^2$ *(see Figure 3.1, left).*

In order to construct our symbolic dynamics, more restrictive assumptions on our central configurations need to be made:

Definition 3.1.2. *Suppose that* ∂D *admits a* C^1 *parametrization. A central configuration* $\bar{P}$ *is said to be* admissible *if*

- *there exists a second central configurations,* $\hat{P}$, *such that the origin does not belong to the segment connecting* $\bar{P}$ *and* $\hat{P}$ *(we say then that* $\bar{P}$ *and* $\hat{P}$ *are* not antipodally directed, or shortly, antipodal*), and*

- *both* $\hat{P}$ *and* $\bar{P}$ *are strict local maxima or minima for the function* $\|\cdot\|_{\partial D}$.

The domain D *is called* admissible *if its boundary contains at least two admissible central configurations denoted by* $\bar{P}_i$, $i \in \mathcal{I}$, $\#\mathcal{I} \geq 2$.

As we already observed in Section 2.2.2, the non antipodality property is crucial to guarantee the uniqueness of a *(TnT)* inner arc (recall Definition 2.2.2) starting from arbitrary endpoints. Note that admissible domains are generic in $\mathcal{C}^1$.

Now, assume that the domain D is admissible: we will use the indices $i \in \mathcal{I}$ as the symbols of our alphabet. The corresponding words will be composed as bi-infinite sequences of symbols in $\mathcal{I}$, with the following further admissibility requirement.

Definition 3.1.3. *We define the set of* admissible words *for our symbolic dynamics as*

$$\mathcal{L} \doteq \left\{ \underline{\ell} = (\ell_i)_i \in \mathcal{I}^{\mathbb{Z}} \;\middle|\; \begin{array}{l} \textit{for every } i \in \mathbb{Z}, \textit{ the symbols } \ell_i \textit{ and } \ell_{i+1} \textit{ do not} \\ \textit{correspond to antipodally directed central configurations} \end{array} \right\}.$$

Definition 3.1.4. *Given mutually disjoint neighbourhoods* $\mathcal{N}_i \subseteq \partial D$ *of the central configurations* $\bar{P}_i$, *we say that a trajectory realizes a word* $\underline{\ell} \in \mathcal{I}^{\mathbb{Z}}$ *if it visits the*
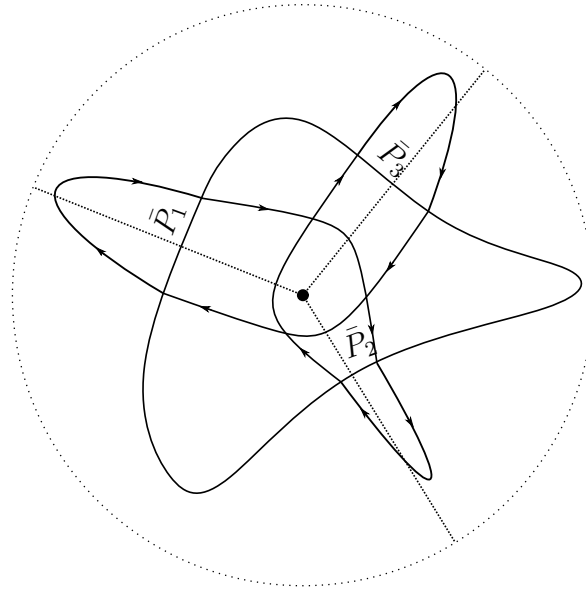
Fig. 3.2 Example of orbit which realizes the periodic word $\underline{\ell} = (\ldots, 1, 2, 3, 1, 2, 3, \ldots)$. The orbit visits the neighbourhoods of the three central configurations $\bar{P}_i$, $i = 1, 2, 3$, in the order prescribed by $\underline{\ell}$.

neighbourhoods $\mathcal{N}_i$, $i \in \mathcal{I}$, in the order imposed by $\underline{\ell}$ (see Figure 3.2). For refraction trajectories, this means that there are two consecutive crossings of ∂D in-out-in in each $\mathcal{N}_i$, while for the case of reflection trajectories there is a reflection point in N_i.

Given these premises, we are in a position to state our main result.

Theorem 3.1.5. *Let D be an admissible domain. For any sufficiently large internal energy h there exists a subset X of the initial conditions-set, a first return map $\mathcal{F}$ ad a continuous surjective map $\pi : X \to \mathcal{L}$ such that the diagram*

$$
\begin{array}{ccc}
X & \xrightarrow{\mathcal{F}} & X \\
\downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\
\mathcal{L} & \xrightarrow{\sigma_r} & \mathcal{L}
\end{array}
$$

commutes, where σ_r is the Bernoulli right-shift. In other words, for large enough h, our refraction (resp. reflection) billiard model admits a symbolic dynamics.

A particular case included in Theorem 3.1.5, which is by the way the key of its proof, regards the existence of periodic orbits, which is the content of the following corollary.

Corollary 3.1.6. *If h is large enough, for every periodic admissible word $\underline{\ell} \in \mathcal{L}$ there exists a periodic trajectory $z_{\underline{\ell}}$ which realizes $\underline{\ell}$, in the sense described by Definition 3.1.4.*

Under some restriction on the words in $\mathcal{L}$, the corresponding symbolic dynamics is collision-free. In order to do that we define the set of bi-infinite symmetric words $\mathcal{L}_s \subset \mathcal{L}$ admitting a symmetry axis and state the following corollary.

Corollary 3.1.7. *Replacing $\mathcal{L}$ with $\widetilde{\mathcal{L}} \doteq \mathcal{L} \setminus \mathcal{L}_s$ in the diagram of Theorem 3.1.5, we obtain that the symbolic dynamics is not collisional, in the sense that any trajectory corresponding to a word $\underline{\ell} \in \widetilde{\mathcal{L}}$ does not have any collisional inner arc.*

To prove Theorem 3.1.5 we shall use a broken geodesic method, reminiscent of the one used in [69, 32], together with a shadowing lemma based on Poincaré-Miranda's fixed point theorem (see [33]). We will perform the proof for planar domains, but it can be easily extended in any dimension. The technique of the proofs, in essence, is the same for refraction and reflection billiards, the latter being somewhat simpler. This is why, throughout this chapter, we shall enter here more in the details of the refraction case.

We stress that, strictly speaking, the existence of a symbolic dynamics is not equivalent to claim that our model is chaotic; as a matter of fact, according to Devaney [29], to have a chaotic dynamical system also the *injectivity* of the map π is required. In our case, the possible injectivity of π may be obstructed by the lack of uniqueness of the critical point we given by our critical point argument (that is, Poincaré-Miranda theorem does not guarantee the uniqueness of the fixed point). Nevertheless, the presence of a symbolic dynamics is a strong indicator of chaos, which is coherent with the numerical simulations presented in [1], where one can see that, for h large enough, the refraction model with an elliptic interface presents diffusive orbits typical of chaotic behaviours. The gap between symbolic dynamics and topological chaos in the general setting will be filled, in full details, in [3].

An intermediate result, regarding the analytic non-integrability of our billiard can be obtained by requiring further assumption on a subset (possibly composed by a single element) of the central configurations of the admissible domain.

Definition 3.1.8. *Let D an admissible domain, and let P_i, $i \in \mathcal{I}$ the set of its admissible configurations. We say that a central configuration P_j is nondegenerate if it is a* nondegenerate *critical point for $\| \cdot \|_{\partial D}$.*

Adapting a classical argument by Kozlov ([34]) we can link the presence of a symbolic dynamics with the non analytic integrability of the system under the assumption that at least one of the central configurations is nondegenerate.

Theorem 3.1.9. *Let D be an admissible domain, and assume that it admits at least one nondegenerate central configuration. If h is large enough, then there are not non-constant analytic first integrals of the motion.*

This result is heavily based on two facts: first of all, every nondegenerate homothetic orbit is a saddle hyperbolic equilibrium if h is large enough; on the other hand, whenever a central configuration is a saddle, by means of the symbolic dynamics it is possible to construct a one-side infinite trajectory starting from an arbitrary point of (a suitable subset of) ∂D and intersecting the stable or unstable manifold of the saddle itself.

Furthermore, requiring that at least two central configurations are non-degenerate, it is possible to extend this result to find infinite complete heteroclinic connections between them: this fact is relevant, since multiple heteroclinics are usually indicators of a complex dynamics (see for instance [82–84]).

Theorem 3.1.9 implies, for example, that elliptic one-center reflection billiards can not be analytically integrable at high energies, when the gravitational center coincides with the center of the ellipses. This negatively complements the recent results about integrability of the one center reflection elliptical billiard with the center occupying one of the foci by Takeuchi and Zhao [85, 7, 86]. Note that an ellipse with focus at the center is not an admissible domain, while it becomes admissible when moving the gravitational center at the center of the ellipses. The interested reader can compare our result with the well established theory of integrability of the gravitational n-centre problem ([87–92]). More on integrability at high energies of the n-center problem can be found in [93, 94].

3.2 Existence of a local dynamics for general domains

In this first section we will study separately the outer and inner dynamics proving two local existence results, given by Theorems 3.2.3 and 3.2.7; such results are heavily based on Theorems 2.2.3 and 2.2.1, presented in Section 2.2.2. We begin recalling some geometric properties of the domain of the refractive billiard, focusing our attention on what will be crucial in the following. Given our bounded open domain $D \subset \mathbb{R}^2$ containing the origin, let us take $\gamma : [0, L] \to \mathbb{R}^2$, $\gamma \in C^1([0, L])$, such that $\partial D = \gamma([0, L])$ and $\|\dot\gamma(\xi)\| = 1$ for every $\xi \in [0, L]^2$.

Since γ is a closed smooth curve, and in particular $\gamma(0) = \gamma(L)$, by Weierstrass Theorem, along with the regularity of γ, the C^1-real-valued function $\xi \mapsto \|\gamma(\xi)\|$ admits at least two stationary points; for any of such points $\bar\xi$, it holds

$$\frac{d}{d\xi}\|\gamma(\xi)\|_{\xi=\bar\xi} = \frac{\gamma(\bar\xi) \cdot \dot\gamma(\bar\xi)}{\|\gamma(\bar\xi)\|} = 0 \quad \implies \quad \gamma(\bar\xi) \perp \dot\gamma(\bar\xi). \tag{3.2.1}$$

The following definitions have a central role in the construction of both the local inner and outer dynamics. Although they have been already presented in Chapter 1, let us recall them for the sake of clarity.

Definition 3.2.1. *We say that the domain D satisfies the* local star-convexity assumption *with respect to $\xi \in [0, L]$ if the half-line connecting the origin to $\gamma(\xi)$ intersects ∂D only in $\gamma(\xi)$.*

---

[2]Note that, in the current chapter, the regularity hypothesis on γ has been weakened, from C^2 to C^1.

Definition 3.2.2. *With an abuse of notation, we say that the parameter $\bar{\xi} \in [0, L]$ is a central configuration if the corresponding point $\gamma(\bar{\xi})$ is a central configuration in the sense of Definition 3.1.2.*

The local existence of outer arcs connecting point near a central configuration follows exactly as in Theorem 2.2.1, here repeated for completeness. Recall that the notation used to denote the inner and puter problems has been introduced in 1.3.11.

Theorem 3.2.3 (Local existence of the outer arcs)**.** *Let $\bar{\xi} \in [0, L]$ be a central configuration. Then there exists $\delta_{E,\bar{\xi}} > 0$ such that for every $\xi_1, \xi_2 \in (\bar{\xi} - \delta_{E,\bar{\xi}}, \bar{\xi} + \delta_{E,\bar{\xi}})$ there is a unique solution $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2)) : [0, T] \to \mathbb{R}^2$ of the fixed-ends problem*

$$\begin{cases} (HS_E)[z(s)] & s \in [0, T] \\ z(s) \notin D & s \in (0, T) \\ z(0) = \gamma(\xi_1), \ z(T) = \gamma(\xi_2) \end{cases} \qquad (3.2.2)$$

for some $T \doteq T(\xi_1, \xi_2) > 0$. Moreover the solution $z_E(\cdot; \gamma(\xi_1), \gamma(\xi_2))$ is transversal to the boundary ∂D at the endpoints, namely,

$$z_E(0; \gamma(\xi_1), \gamma(\xi_2)) \nparallel \dot{\gamma}(\xi_1) \quad and \quad z_E(T; \gamma(\xi_1), \gamma(\xi_2)) \nparallel \dot{\gamma}(\xi_2).$$

Let us outline that the outer dynamics is *local* in the sense that an outer arc connects two points on ∂D belonging to a neighbourhood of the same point $\gamma(\bar{\xi})$.

Concerning the inner dynamics, our aim is to connect two points in neighbourhoods of possibly different images of critical points of $\|\gamma(\cdot)\|$. Hence, also the inner dynamics is local; nevertheless, as we will see in the proof of Theorem 3.2.7, this local structure is not necessary to show the existence of the Keplerian arcs themselves, but it is required to ensure that they are completely contained in the domain D.

In order to proceed with the construction of the inner arcs we give the following definitions, necessary to address the non-antipodality property to neighbourhoods of a central configuration in a general domain.

Definition 3.2.4. *We say that D is an* admissible domain *if there are $N \geq 2$ central configurations for D, $\bar{\xi}_1, \ldots, \bar{\xi}_N$, such that for every $i = 1, \ldots, N$, it holds $\#\mathrm{NA}(i) \geq 2$, where*

$$\mathrm{NA}(i) \doteq \left\{ j \in \{1, \ldots, N\} : \gamma(\bar{\xi}_i) \text{ and } \gamma(\bar{\xi}_j) \text{ are not antipodal} \right\};$$

in this case every central configuration $\bar{\xi}_i$, $i = 1, \ldots, N$ is termed admissible.

We observe that $i \in \mathrm{NA}(i)$ for every i, hence $\bar{\xi}_i$ is admissible $\gamma(\bar{\xi}_i)$ admits a distinct not antipodal $\gamma(\bar{\xi}_j)$.

Definition 3.2.5. *To ease the notation, for every $\delta > 0$, we define the following sets*

$$\mathcal{U}(\delta) \doteq \bigcup_{i=1,\ldots,N} (\bar{\xi}_i - \delta, \bar{\xi}_i + \delta),$$

$$\mathcal{U}_{\mathrm{NA}}(\delta; i) \doteq \bigcup_{j \in \mathrm{NA}(i)} (\bar{\xi}_j - \delta, \bar{\xi}_j + \delta),$$

$$\mathcal{P}_{\mathrm{NA}}(\delta; i) \doteq (\bar{\xi}_i - \delta, \bar{\xi}_i + \delta) \times \mathcal{U}_{\mathrm{NA}}(\delta; i).$$

Definition 3.2.6. *Let $\bar{\delta} > 0$ be such that*

- *for every $i \in \{1, \ldots, N\}$ and for every $(\xi_1, \xi_2) \in \mathcal{P}_{\mathrm{NA}}(\delta; i)$ the corresponding images $\gamma(\xi_1)$ and $\gamma(\xi_2)$ are not antipodal;*

- *the domain D satisfies the local star-convexity assumption with respect to any $\xi \in \mathcal{U}(\bar{\delta})$.*

Note that, if D is an admissible domain, the quantity $\bar{\delta}$ always exists by construction. We also point out that Theorem 3.2.3 ensures the existence of an outer dynamics when we restrict to the components of the set $\mathcal{U}(\delta_E)$ with

$$\delta_E = \min\{\bar{\delta}, \delta_{E, \bar{\xi}_1}, \ldots, \delta_{E, \bar{\xi}_N}\}. \tag{3.2.3}$$

The existence of a local inner dynamics is heavily based on the general results described in Section 2.2.2, where the problem of finding a Keplerian arc connecting two given points is dealt in a more *geometric*, rather than *dynamical*, framework.
We are now ready to present the analogous of Theorem 3.2.3 for the inner dynamics. In the following, the domain D is always supposed to be admissible; recall that the definition of (TnT) arcs has been introduced in Definition 2.2.2.

Theorem 3.2.7 (Existence and transversality of the inner arcs)**.** *Let $N \geq 2$ and $\bar{\xi}_1, \ldots, \bar{\xi}_N \in [0, L]$ be admissible central configurations for D. Then:*

(i) there exist $h_0 > 0$ and $\delta_I > 0$ such that for every $i = 1, \ldots, N$, $(\xi_1, \xi_2) \in \mathcal{P}_{NA}(\delta_I; i)$, and for every $h > h_0$ there exists a unique $T \doteq T(\xi_1, \xi_2; h) > 0$ such that:

- *if $\xi_1 \neq \xi_2$, there exists a unique inner arc, $z_I(\cdot; \gamma(\xi_1), \gamma(\xi_2); h) \in C^2([0, T])$, connecting $\gamma(\xi_1)$ and $\gamma(\xi_2)$ and satisfying (TnT), which solves the fixed-ends problem*

$$\begin{cases} (HS_I)[z(s)] & s \in [0, T] \\ z(s) \in D & s \in (0, T) \\ z(0) = \gamma(\xi_1), \ z(T) = \gamma(\xi_2) \end{cases} \tag{3.2.4}$$

- *if $\xi_1 = \xi_2$, there exists a unique inner arc $z_I(\cdot; \gamma(\xi_1), \gamma(\xi_2); h) \colon [0, T] \to \mathbb{R}^2$ which is a collision-ejection solution of the fixed-ends problem (3.2.4) in $[0, T] \setminus \{T/2\}$.*
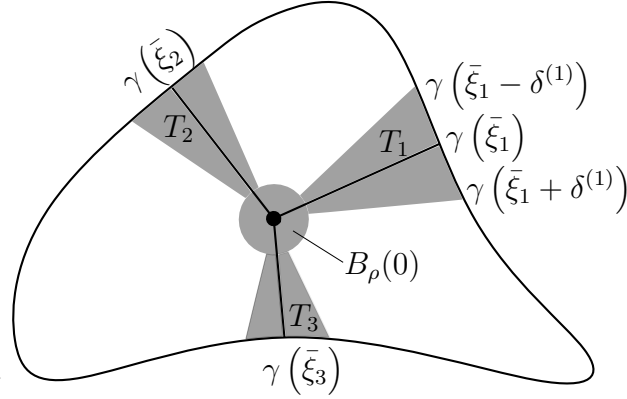
Fig. 3.3 The open set A introduced in the proof of Theorem 3.2.7. A is the union of the circular neighbourhood of the origin $B_\rho(0)$ and of the sectors T_i, $i = 1, 2, 3$.

(ii) For every $\epsilon > 0$ there exists $h_1 \doteq h_1(\epsilon) > h_0$ such that for every $i = 1, \ldots, N$, $(\xi_1, \xi_2) \in \mathcal{P}_{NA}(\delta_I; i)$, ad every $h > h_1$, defined the angles

$$\alpha_1 \doteq \angle\left(\gamma(\xi_1), -z_I'(0; \gamma(\xi_1), \gamma(\xi_2); h)\right),$$
$$\alpha_2 \doteq \angle\left(\gamma(\xi_2), z_I'(T; \gamma(\xi_1), \gamma(\xi_2); h)\right),$$

one has $|\alpha_1| < \epsilon$ and $|\alpha_2| < \epsilon$.

Proof. Let us start with claim *(i)* when $\xi_1 \neq \xi_2$. Recalling the Definition 3.2.6, by Theorem 2.2.3 one has that taking $(\xi_1, \xi_2) \in \overline{\mathcal{P}_{NA}(\bar{\delta}/2; i)} \subsetneq \mathcal{P}_{NA}(\bar{\delta}; i)$ and $h > 0$, there exists $T = T(\gamma(\xi_1), \gamma(\xi_2); h) > 0$ and a unique $z_I(\cdot; \gamma(\xi_1), \gamma(\xi_2); h)$ solution of the Bolza problem

$$\begin{cases} (HS_I)[z(s)] & s \in [0, T] \\ z(0) = \gamma(\xi_1),\ z(T) = \gamma(\xi_2) \end{cases} \tag{3.2.5}$$

which is *(TnT)*. We now have to find a condition ensuring that $z_I(s; \gamma(\xi_1), \gamma(\xi_2); h) \in D$ for every $s \in (0, T)$. In general, given $P_1, P_2 \in \mathbb{R}^2 \setminus \{0\}$ not antipodal and distinct, let us define $c(P_1 0 P_2)$ as the union of the two straight-line segments from P_1 to 0 and from 0 to P_2. Considering now the arc $\mathcal{H}_0(\gamma(\xi_1), \gamma(\xi_2); \mathcal{E} + h)$ parametrized by $z_I(\cdot; \gamma(\xi_1), \gamma(\xi_2); h)$, the convergence property stated in Remark 2.2.5 can be rephrased as

$$\lim_{h \to \infty} dist\left(\mathcal{H}_0(P_1, P_2; h), c(P_1 0 P_2)\right) = 0. \tag{3.2.6}$$

As $0 \in D$ and since D is open, there exists $\rho > 0$ such that $B_\rho(0) \subsetneq D$; moreover, for every $i = 1, \ldots, N$, one can define the conic set (see Figure 3.3)

$$T_i = \left\{ \lambda\gamma(\xi) \mid \lambda \in (0, 1),\ \xi \in \left(\bar{\xi}_i - \bar{\delta}, \bar{\xi}_i + \bar{\delta}\right) \right\},$$

and the open set

$$A = \bigcup_{i=1,\ldots,N} T_i \cup B_\rho(0).$$

Fixed $i \in \{1, \ldots, N\}$, by the convergence in (3.2.6), for every $(\xi_1, \xi_2) \in \overline{\mathcal{P}_{NA}(\bar{\delta}/2; i)}$, there exists $\bar{h}_i(\xi_1, \xi_2) > 0$ such that for every $h > \bar{h}_i(\xi_1, \xi_2)$

$$z_I\left((0, T); \gamma(\xi_1), \gamma(\xi_2); h\right) \subset A \subset D. \tag{3.2.7}$$
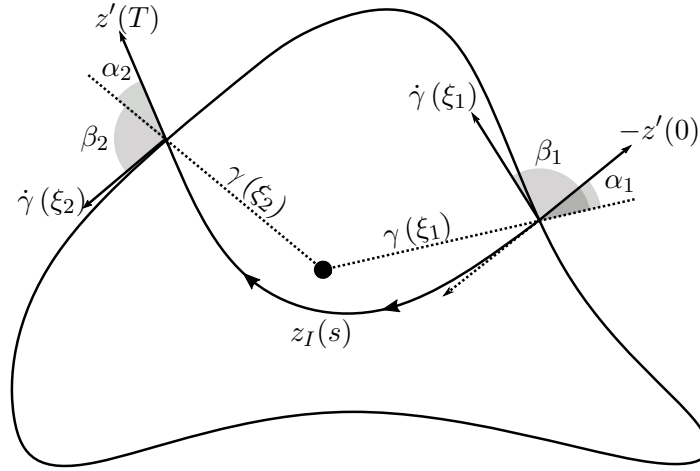
Fig. 3.4 Angles of the inner arc $z_I(\cdot) \equiv z_I(\cdot; \gamma(\xi_1), \gamma(\xi_2); h)$ with respect to the initial and final radial and tangent directions. In particular, α_1 and α_2 follow the same definition of Theorem 3.2.7 and $\beta_k = \angle(\dot\gamma(\xi_k), \gamma(\xi_k))$, $k = 1, 2$.

By the differentiable dependence of the Bolza problem (3.2.5) with respect to the parameter h and to the endpoints, and the regularity of γ, for every i the function $(\xi_1, \xi_2) \mapsto \bar{h}_i(\xi_1, \xi_2)$ is continuous on the compact set $\overline{\mathcal{P}_{NA}(\bar\delta/2; i)}$; as a consequence, it admits a maximum value $\bar{h}_{i,M}$. Hence, for every $h \geq \bar{h}_{i,M}$, condition (3.2.7) is satisfied for every $(\xi_1, \xi_2) \in \overline{\mathcal{P}_{NA}(\bar\delta/2; i)}$. Setting now $h_0 \doteq \max\{\bar{h}_{1,M}, \ldots, \bar{h}_{N,M}\}$ and $\delta_I = \bar\delta/2$, one has that for every $i = 1, \ldots, N$, $(\xi_1, \xi_2) \in \mathcal{P}_{NA}(\delta_I; i)$ and every $h > h_0$ the unique *(TnT)* solution of Pb. (3.2.5) satisfies also Pb. (3.2.4).

In the case $\xi_1 = \xi_2$, the existence of the collision-ejection solution $z_I(\cdot; \gamma(\xi_1), \gamma(\xi_2), h)$ such that $z_I(0; \gamma(\xi_1), \gamma(\xi_2); h) = z_I(T; \gamma(\xi_1), \gamma(\xi_2); h) = \gamma(\xi_1)$ is guaranteed by Remark 2.2.7.

The proof of claim *(ii)* follows the same reasoning: if $\xi_1 \neq \xi_2$, once fixed $\epsilon > 0$, recalling Proposition 2.2.4 one has that for every $i = 1, \ldots, N$ and $(\xi_1, \xi_2) \in \overline{\mathcal{P}_{NA}(\delta_I; i)}$ there exists $\underline{h}_i(\xi_1, \xi_2; \epsilon)$ such that for every $h > \underline{h}_i(\xi_1, \xi_2; \epsilon)$ it results $|\alpha_1| < \epsilon$ and $|\alpha_2| < \epsilon$. The function $(\xi_1, \xi_2) \mapsto \underline{h}_i(\xi_1, \xi_2; \epsilon)$ is again continuous on the compact set $\overline{\mathcal{P}_{NA}(\delta_I; i)}$ and achieves its maximal value $h_{i,1}$; hence, letting

$$h_1 \doteq \max\{h_0, h_{1,1}, \ldots, h_{N,1}\}$$

the claim is proved.

In the case $\xi_1 = \xi_2$, one can refer again to Remark 2.2.7 to ensure that $\alpha_1 = \alpha_2 = 0$. $\qquad\square$

Note that if we define

$$\delta \doteq \frac{1}{2}\min\{\delta_E, \delta_I\}, \tag{3.2.8}$$

both Theorems 3.2.3 and 3.2.7 hold simultaneously.

We recall that, casting together Theorems 3.2.3 and 3.2.7 and definition 3.2.2, we can conclude that if $\bar\xi$ is a central configuration then the complete dynamics admits an homothetic trajectory in the direction of $\gamma(\bar\xi)$.

To conclude the exploration of the geometric properties of the inner arcs, in the next result we show that every arc defined in Theorem 3.2.7 is arbitrarily transversal to the boundary ∂D, possibly reducing δ_I and increasing h_0.

Proposition 3.2.8. *For every $\epsilon \in (0, \pi/2)$ there exist $\delta_I^\epsilon \in (0, \delta_I]$ and $\bar{h}^\epsilon > h_0$ such that for every $i = 1, \ldots, N$, $(\xi_1, \xi_2) \in \overline{\mathcal{P}_{NA}(\delta_I^\epsilon; i)}$, and every $h > \bar{h}^\epsilon$*

$$
\left| \angle(\dot{\gamma}(\xi_1), -z_I'(0; \gamma(\xi_1), \gamma(\xi_2); h)) - \frac{\pi}{2} \right| < \epsilon,
$$

$$
\left| \angle(\dot{\gamma}(\xi_2), z_I'(T(\xi_1, \xi_2); \gamma(\xi_1), \gamma(\xi_2); h)) - \frac{\pi}{2} \right| < \epsilon. \tag{3.2.9}
$$

Proof. For every $i = 1, \ldots, N$, the geometric properties of the admissible central configuration $\bar{\xi}_i$ imply that

$$
\cos\left(\angle\left(\dot{\gamma}(\bar{\xi}_i), \gamma(\bar{\xi}_i)\right)\right) = \dot{\gamma}(\bar{\xi}_i) \cdot \frac{\gamma(\bar{\xi}_i)}{\|\gamma(\bar{\xi}_i)\|} = 0;
$$

hence, by the regularity of γ, for every $\epsilon > 0$, there exists $\delta_I^\epsilon \in (0, \delta_I]$ such that

$$
\left| \angle\left(\dot{\gamma}(\xi), \gamma(\xi)\right) - \frac{\pi}{2} \right| < \frac{\epsilon}{2}, \tag{3.2.10}
$$

for every $\xi \in \overline{\mathcal{U}(\delta_I^\epsilon)}$. Moreover, from claim (ii) of Theorem 3.2.7 one has that there exists $\bar{h}^\epsilon$ such that for every $i = 1, \ldots, N$, $(\xi_1, \xi_2) \in \overline{\mathcal{P}_{NA}(\delta_I^\epsilon; i)}$, and for every $h > \bar{h}^\epsilon$

$$
\left| \angle(\gamma(\xi_1), -z_I'(0; \gamma(\xi_1), \gamma(\xi_2); h)) \right| < \frac{\epsilon}{2} \quad \text{and} \quad \left| \angle(\gamma(\xi_2), z_I'(T; \gamma(\xi_1), \gamma(\xi_2); h)) \right| < \frac{\epsilon}{2}. \tag{3.2.11}
$$

Taking together Eqs. (3.2.10) and (3.2.11), one can then conclude that, for every $i = 1, \ldots, N$, $(\xi_1, \xi_2) \in \overline{\mathcal{P}_{NA}(\delta_I^\epsilon; i)}$, and every h greater than $\bar{h}^\epsilon$,

$$
\left| \angle\left(\dot{\gamma}(\xi_1), -z_I'(0; \gamma(\xi_1), \gamma(\xi_2); h)\right) - \frac{\pi}{2} \right|
$$

$$
\leq \quad \left| \angle\left(\dot{\gamma}(\xi_1), \gamma(\xi_1)\right) \right| + \left| \angle(\gamma(\xi_1), -z_I'(0; \gamma(\xi_1), \gamma(\xi_2); h)) - \frac{\pi}{2} \right| < \epsilon,
$$

and the same reasoning applies to $\angle(\dot{\gamma}(\xi_2), z_I'(T; \gamma(\xi_1), \gamma(\xi_2); h))$. $\qquad \square$

Remark 3.2.9. *Theorems 3.2.3 and 3.2.7 ensure the existence of the outer and inner dynamics separately, provided that the endpoints' parameters ξ_1 and ξ_2 are in suitable neighbourhoods of $\bar{\xi}_i$, $i = 1, \ldots, N$. Nevertheless, this is not sufficient to ensure the good definition of the* complete *dynamics, that is, a concatenation between outer and inner arcs satisfying the Snell's law.*
In particular, as observed in Section 1.2, one has that the refraction exterior-interior is always possible, while the converse, interior-exterior, can take place if and only if the inner arc is sufficiently transverse to the boundary. Hence, in order to prove the existence of a complete dynamics, we should find conditions to have uniform transversality properties of the inner arcs. On the other hand, this is not really necessary to our purposes: we will prove a posteriori *that the particular concatenations*

of outer and inner arcs that realize a symbolic dynamics for our problem are admissible trajectories. The idea is that the transversality of an inner arc is indirectly deduced from the validity of the variational formulation of the Snell's law (see Eq. (1.2.13) in Section 1.2), along with the transversality of the subsequent outer arc.

3.3 Symbolic dynamics

The current section is devoted to the construction of a symbolic dynamics using as building blocks the outer and inner arcs whose existence and properties have been proved in Section 3.2. The final result is Theorem 3.1.5. Before starting, some technical results connected to the geometric properties of the inner and outer arcs, along with their variational consequences in terms of Jacobi distances, are stated. From now on, we refer to the definitions already given in Section 1.2.

3.3.1 Estimates on angles

As substantiated in Section 3.2, to ensure the local good definition of the outer and inner dynamics in neighbourhoods of $\gamma(\bar{\xi}_i)$, $i = 1, \ldots, N$, the only hypothesis required on the points $\bar{\xi}_i$ is that they are admissible central configurations, and, in particular, each $\bar{\xi}_i$ must be a critical point for $\|\gamma(\xi)\|$. In the following, the nature of such critical points is crucial to relate the geometry of the domain D to the variational properties of the inner and outer arcs. We then make additional assumption on D.

Assumption 3.3.1. *Let us suppose that there exists a subset $\mathcal{I} \subseteq \{1, \ldots, N\}$ such that:*

- *$\#\mathcal{I} \geq 2$;*

- *for any $i \in \mathcal{I}$, the critical point $\bar{\xi}_i$ is a local strict minimum or maximum for $\|\gamma(\xi)\|$, so that we can write*
$$\mathcal{I} = \mathcal{I}_m \cup \mathcal{I}_M,$$
where if $i \in \mathcal{I}_m$ (resp. $\mathcal{I}_M$) then $\bar{\xi}_i$ is a minimum (resp. maximum). Let us note that one of these two sets can be empty;

- *the set $\mathcal{I}$ corresponds to a set of admissible central configurations for D (see Definition 3.2.4).*

Although Theorem 3.1.5 can be proved under the assumption $\gamma \in C^1$ and satisfying Assumption 3.3.1, for the sake of simplicity we will present a proof based on stricter hypotheses, specified in Assumption 3.3.2; a way more general case will be exposed in [3].

Assumption 3.3.2. *Let us suppose $\gamma \in C^2([0, L])$. With reference to Eq. (3.2.8), let us suppose that δ is such that:*

(δ1) *the function $\|\gamma(\xi)\|$ is respectively strictly convex/concave in $[\bar{\xi}_i - \delta, \bar{\xi}_i + \delta]$ for every $i \in \mathcal{I}$, depending on the nature of the critical point $\bar{\xi}_i$;*

(δ2) *the angle between the radial directions and the tangent vectors to γ is in a small neighbourhood of $\pi/2$, namely, there exists $A_\delta \in (0, \pi/2)$ such that for every $i \in \mathcal{I}$ and every $\xi \in [\bar{\xi}_i - \delta, \bar{\xi}_i + \delta]$ one has*

$$0 < \frac{\pi}{2} - A_\delta < \angle\left(\gamma(\xi), \dot{\gamma}(\xi)\right) < \frac{\pi}{2} + A_\delta < \pi.$$

Let us start our study by considering a geometric property holding when the critical point $\bar{\xi}_i$ is a strict local minimum. From now on, the quantity $\delta > 0$ is the one introduced in Assumption 3.3.2, h_0 is the energy bound introduced in Theorem 3.2.7 and the set $\mathcal{U}_{NA}(\delta; i)$ is the one introduced in Definition 3.2.5.

Lemma 3.3.3. *There exist $\bar{h}_m > h_0$ and $C_m > 0$ such that for every $i \in \mathcal{I}_m$, for every $\xi \in \mathcal{U}_{NA}(\delta; i)$ and every $h > \bar{h}_m$ one has*

$$- \frac{z_I'\left(0; \gamma(\bar{\xi}_i + \delta), \gamma(\xi); h\right)}{\|z_I'\left(0; \gamma(\bar{\xi}_i + \delta), \gamma(\xi); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i + \delta) > C_m,$$

$$- \frac{z_I'\left(0; \gamma(\bar{\xi}_i - \delta), \gamma(\xi); h\right)}{\|z_I'\left(0; \gamma(\bar{\xi}_i - \delta), \gamma(\xi); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i - \delta) < -C_m,$$

$$\frac{z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i + \delta); h\right)}{\|z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i + \delta); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i + \delta) > C_m,$$

$$\frac{z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i - \delta); h\right)}{\|z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i + \delta); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i - \delta) < -C_m.$$

where $z_I(\cdot; \cdot, \cdot; h)$ is the unique inner arc connecting the specified pair of points with energy $\mathcal{E} + h$, whose existence is guaranteed in Theorem 3.2.7.

Proof. We will prove only the last two inequalities, as the first two can be derived straightforwardly from them using the time reversibility and the uniqueness of the solutions claimed in Theorem 3.2.7. Let now $h > h_0$. By the definition of scalar product, for every $i \in \mathcal{I}_m$ and $\xi \in \mathcal{U}_{NA}(\delta; i)$, it holds

$$\frac{z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i \pm \delta); h\right)}{\|z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i \pm \delta); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i \pm \delta) = \cos\left(\angle\left(z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i \pm \delta); h\right), \dot{\gamma}(\bar{\xi}_i \pm \delta)\right)\right),$$

(3.3.1)

where the last angle can be splitted as a sum of the form

$$\angle\left(z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i \pm \delta); h\right), \dot{\gamma}(\bar{\xi}_i \pm \delta)\right) = \alpha_\pm(\xi, \bar{\xi}_i, h) + \beta_\pm(\bar{\xi}_i), \qquad (3.3.2)$$

where

$$\alpha_\pm(\xi, \bar{\xi}_i, h) = \angle\left(z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i \pm \delta); h\right), \gamma(\bar{\xi}_i \pm \delta)\right), \quad \text{and}$$

$$\beta_\pm(\bar{\xi}_i) = \angle\left(\gamma(\bar{\xi}_i \pm \delta), \dot{\gamma}(\bar{\xi}_i \pm \delta)\right).$$

We observe that

$$\cos\beta_\pm(\bar{\xi}_i) = \frac{\gamma(\bar{\xi}_i \pm \delta)}{\|\gamma(\bar{\xi}_i \pm \delta)\|} \cdot \dot{\gamma}(\bar{\xi}_i \pm \delta) = \frac{d}{d\xi}\|\gamma(\xi)\|_{|\xi = \bar{\xi}_i \pm \delta}.$$

Since $\bar{\xi}_i$ is a strict minimum, the function $\|\gamma(\xi)\|$ is strictly convex in $[\bar{\xi}_i - \delta, \bar{\xi}_i + \delta]$, and its derivative $\dfrac{d\|\gamma(\xi)\|}{d\xi}$ is strictly increasing in the same interval. Furthermore, the extremality of $\bar{\xi}_i$ ensures that $\dfrac{d}{d\xi}\|\gamma(\xi)\|_{|\xi = \bar{\xi}_i} = 0$, thus, there exist $\tilde{A} \in (0,1)$ and $A \in (0, A_\delta)$ (where A_δ has been introduced in Assumption 3.3.2) such that for every $i \in \mathcal{I}_m$

$$\cos\beta_+(\bar{\xi}_i) > \tilde{A}, \quad \cos\beta_-(\bar{\xi}_i) < -\tilde{A}, \tag{3.3.3}$$

and, still by virtue of Assumption 3.3.2,

$$0 < \frac{\pi}{2} - A_\delta < \beta_+(\bar{\xi}_i) < \frac{\pi}{2} - A < \pi, \quad 0 < \frac{\pi}{2} + A < \beta_-(\bar{\xi}_i) < \frac{\pi}{2} + A_\delta < \pi.$$

Let us now consider the angles $\alpha_\pm(\xi, \bar{\xi}_i, h)$; fix $\bar{\epsilon} > 0$ such that $\bar{\epsilon} < \min\{A, \pi/2 - A_\delta\}$. By Theorem 3.2.7, there exists $\bar{h}_m \equiv h_1(\bar{\epsilon})$ such that for every $i \in \mathcal{I}_m$, for every $\xi \in \overline{\mathcal{U}_{NA}(\delta; i)}$, and $h > \bar{h}_m$ one has $|\alpha_\pm(\xi, \bar{\xi}_i, h)| < \bar{\epsilon}$. Recalling Eq. (3.3.2), one can then conclude that

$$0 < \frac{\pi}{2} - A_\delta - \bar{\epsilon} < \angle\left(z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i + \delta) <; h\right), \dot{\gamma}(\bar{\xi}_i + \delta)\right) < \frac{\pi}{2} - A + \bar{\epsilon} < \pi$$

$$0 < \frac{\pi}{2} + A - \bar{\epsilon} < \angle\left(z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i - \delta); h\right), \dot{\gamma}(\bar{\xi}_i - \delta)\right) < \frac{\pi}{2} + A_\delta + \bar{\epsilon} < \pi,$$

from which one has

$$\cos\left(\angle\left(z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i + \delta); h\right), \dot{\gamma}(\bar{\xi}_i + \delta)\right)\right) > \sin\left(A - \bar{\epsilon}\right) \doteq C_m \in (0,1)$$

$$\cos\left(\angle\left(z_I'\left(T; \gamma(\xi), \gamma(\bar{\xi}_i - \delta); h\right), \dot{\gamma}(\bar{\xi}_i - \delta)\right)\right) < -\sin\left(A - \bar{\epsilon}\right) = -C_m.$$

Recalling Eq. (3.3.1), the last inequalities conclude the proof. $\qquad\square$

The geometric interpretation of Lemma 3.3.3 is the following: if δ is small enough and $\bar{\xi}_i$ is a strict minimum for $\|\gamma(\xi)\|$, for sufficiently large energies every inner arc connecting $\gamma(\bar{\xi}_i - \delta)$ with a non-antipodal point $\gamma(\xi)$ forms with $\dot{\gamma}(\bar{\xi}_i - \delta)$ an angle strictly greater than $\pi/2$. Conversely, the Keplerian arcs starting from or arriving in $\gamma(\bar{\xi}_i + \delta)$ form with $\dot{\gamma}(\bar{\xi}_i + \delta)$ an angle strictly smaller than $\pi/2$ (see Figure 3.5, left).
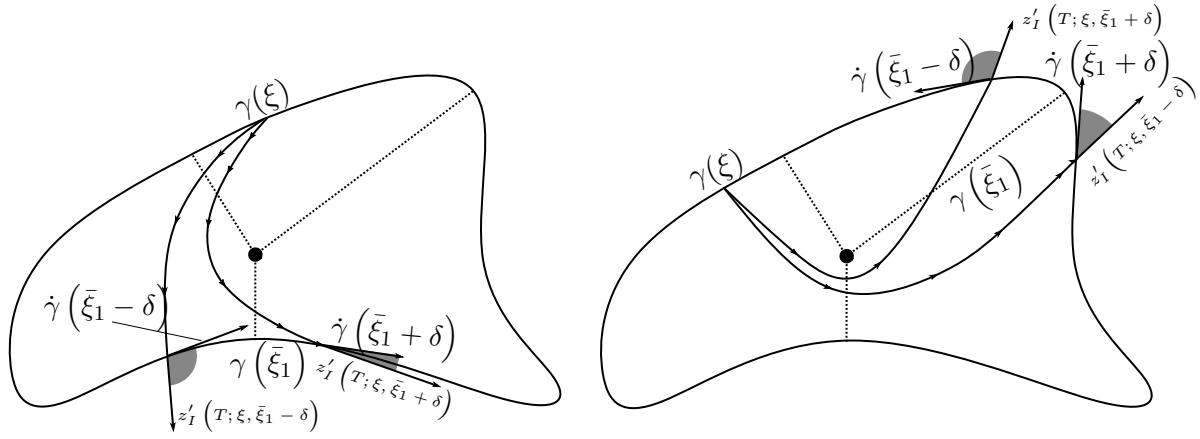
Fig. 3.5 Estimates on the angles proved in Lemmas 3.3.3 and 3.3.4 for the strict minimum case (left) and the strict maximum case (right). While in the first case the angle between the inner arc and tangent vector to γ is obtuse in $\bar{\xi}_i - \delta$ and acute in $\bar{\xi}_i + \delta$; the estimates are reversed in the second case. Here, for the sake of brevity, $z_I(t; \xi_1, \xi_2) \doteq z_I(t; \gamma(\xi_1), \gamma(\xi_2); h)$.

When the critical point $\bar{\xi}_i$ is a strict local maximum we make use of the same reasoning of the previous lemma, taking now into account the strict concavity of $\|\gamma(\xi)\|$ around such point. Again, the geometric interpretation of the maximal case can be found in Figure 3.5, right, and it is completely specular with respect to the minimal case.

Lemma 3.3.4. *There exist $\bar{h}_M > 0$ and $C_M > 0$ such that for every $i \in \mathcal{I}_M$, for every $\xi \in \overline{\mathcal{U}_{NA}(\delta; i)}$ and $h > \bar{h}_M$ one has*

$$- \frac{z'_I\left(0; \gamma(\bar{\xi}_i + \delta), \gamma(\xi); h\right)}{\|z'_I\left(0; \gamma(\bar{\xi}_i + \delta), \gamma(\xi); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i + \delta) < -C_M,$$

$$- \frac{z'_I\left(0; \gamma(\bar{\xi}_i - \delta), \gamma(\xi); h\right)}{\|z'_I\left(0; \gamma(\bar{\xi}_i - \delta), \gamma(\xi); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i - \delta) > C_M,$$

$$\frac{z'_I\left(T; \gamma(\xi), \gamma(\bar{\xi}_i + \delta); h\right)}{\|z'_I\left(T; \gamma(\xi), \gamma(\bar{\xi}_i + \delta); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i + \delta) < -C_M,$$

$$\frac{z'_I\left(T; \gamma(\xi), \gamma(\bar{\xi}_i - \delta); h\right)}{\|z'_I\left(T; \gamma(\xi), \gamma(\bar{\xi}_i - \delta); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i - \delta) > C_M.$$

where $z_I(\cdot; \cdot, \cdot; h)$ is the unique inner arc connecting the specified pair of points with energy $\mathcal{E} + h$ whose existence is guaranteed in Theorem 3.2.7.

Corollary 3.3.5. *Taking $\bar{h} = \max\left\{\bar{h}_m, \bar{h}_M\right\}$ and $C = \min\{C_m, C_M\}$, one can unify Lemmas 3.3.3 and 3.3.4 to obtain estimates holding for every $i \in \mathcal{I}$.*

Taking into account the definition of the distances $S_E(\xi_1, \xi_2)$ and $S_I(\xi_1, \xi_2; h)^3$, as well as the expressions of their derivatives, provided in Section 1.2, Lemmas 3.3.3

---

[3]Here, the dependence of S_I is explicited since, in view of Theorem 3.2.7, this parameter influences the good definition of the inner arc itself.

and 3.3.4 can be translated into the variational formulation given in the following result, where again strict minima and maxima behave symmetrically. Recall that, given suitable ξ_1, ξ_2, ξ_3 in $[0, L]$ and $h > h_0$, the quantity $S_E(\xi_1, \xi_2) + S_I(\xi_2, \xi_3; h)$ is the total Jacobi length of the concatenation of the outer and inner arcs connecting $\gamma(\xi_1)$ to $\gamma(\xi_2)$ and $\gamma(\xi_2)$ to $\gamma(\xi_3)$. Symmetrically, $S_I(\xi_1, \xi_2; h) + S_E(\xi_2, \xi_3)$ is the total Jacobi length of the concatenation of the inner arc from $\gamma(\xi_1)$ to $\gamma(\xi_2)$ and the outer one from $\gamma(\xi_2)$ to $\gamma(\xi_3)$. Moreover, $\partial_a S_{E/I}$ and $\partial_b S_{E/I}$ denote the partial derivatives of the lengths with respect to their first and second variable: hence $\partial_b S_{E/I} + \partial_a S_{I/E}$ are the partial derivatives of the total Jacobi lengths with respect to the *transition point*.

Lemma 3.3.6. *There exists $\bar{h}_1 > 0$ such that*

- *for every $i \in \mathcal{I}_m$, for every $\xi_E \in [\bar{\xi}_i - \delta, \bar{\xi}_i + \delta]$, $\xi_I \in \overline{\mathcal{U}_{NA}(\delta; i)}$ and every $h > \bar{h}_1$:*

$$
\begin{aligned}
\partial_b S_E(\xi_E, \bar{\xi}_i + \delta) + \partial_a S_I(\bar{\xi}_i + \delta, \xi_I) &> 0 \\
\partial_b S_E(\xi_E, \bar{\xi}_i - \delta) + \partial_a S_I(\bar{\xi}_i - \delta, \xi_I) &< 0 \\
\partial_b S_I(\xi_I, \bar{\xi}_i + \delta) + \partial_a S_E(\bar{\xi}_i + \delta, \xi_E) &> 0 \\
\partial_b S_I(\xi_I, \bar{\xi}_i - \delta) + \partial_a S_E(\bar{\xi}_i - \delta, \xi_E) &< 0
\end{aligned}
\tag{3.3.4}
$$

- *for every $i \in \mathcal{I}_M$, the above result holds with the inverted inequalities.*

Proof. As the other cases are completely symmetric, we will prove only the first two inequalities of Eq. (3.3.4). With reference to Section 1.2, and in particular to Eq. (1.2.6), one has that if $i \in \mathcal{I}_m$

$$
\begin{aligned}
|\partial_b S_E(\xi_E, \bar{\xi}_i + \delta)| &= \sqrt{V_E\left(\gamma(\bar{\xi}_i + \delta)\right)} \left| \frac{z_E'\left(T; \gamma(\xi_E), \gamma(\bar{\xi}_i + \delta)\right)}{\|z_E'\left(T; \gamma(\xi_E), \gamma(\bar{\xi}_i + \delta)\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i + \delta) \right| \\
&= \sqrt{V_E\left(\gamma(\bar{\xi}_i + \delta)\right)} \left| \cos\left(\angle\left(z_E'\left(T; \gamma(\xi_E), \gamma(\bar{\xi}_i + \delta)\right), \dot{\gamma}(\bar{\xi}_i + \delta) \right) \right) \right| \\
&\leq \sqrt{V_E\left(\gamma(\bar{\xi}_i + \delta)\right)},
\end{aligned}
$$

while, making use of Lemma 3.3.3 and Corollary 3.3.5, one obtains that there exist $C > 0$ and $\bar{h} > 0$ such that for every $h > \bar{h}$,

$$
\begin{aligned}
\partial_a S_I(\bar{\xi}_i + \delta, \xi_I; h) &= -\sqrt{V_I\left(\gamma(\bar{\xi}_i + \delta)\right)} \frac{z_I'\left(0; \gamma(\bar{\xi}_i + \delta), \gamma(\xi_I); h\right)}{\|z_I'\left(0; \gamma(\bar{\xi}_i + \delta), \gamma(\xi_I); h\right)\|} \cdot \dot{\gamma}(\bar{\xi}_i + \delta) \\
&> \sqrt{V_I\left(\gamma(\bar{\xi}_i + \delta)\right)} C.
\end{aligned}
$$

Taking together the two expressions, one obtains the chain of inequalities

$$
\begin{aligned}
\partial_b S_E(\xi_E, \bar{\xi}_i + \delta) + \partial_a S_I(\bar{\xi}_i + \delta, \xi_I; h) &> -\sqrt{V_E\left(\gamma(\bar{\xi}_i + \delta)\right)} + \sqrt{V_I\left(\gamma(\bar{\xi}_i + \delta)\right)} C \\
&= -\sqrt{\mathcal{E} - \frac{\omega^2}{2}\|\gamma(\bar{\xi}_i + \delta)\|^2} + \sqrt{\mathcal{E} + h + \frac{\mu}{\|\gamma(\bar{\xi}_i + \delta)\|}} C > -\sqrt{\mathcal{E}} + \sqrt{h}\, C. \quad (3.3.5)
\end{aligned}
$$

Taking then $\bar{h}_1 = \max\left\{\mathcal{E}/C^2, \bar{h}\right\}$ one has that for every $h > \bar{h}_1$ the quantity $\partial_b S_E(\xi_E, \bar{\xi}_i + \delta) + \partial_a S_I(\bar{\xi}_i + \delta, \xi_I)$ is strictly positive.

As for the second inequality of the claim, using again the same reasonings, Lemma 3.3.3 and Corollary 3.3.5, one obtains $\partial_b S_E(\xi_E, \bar{\xi}_i - \delta) + \partial_a S_I(\bar{\xi}_i - \delta, \xi_I) < \sqrt{\mathcal{E}} - \sqrt{h}\,C$, which is negative for $h > \bar{h}_1$.

Carrying on the same computations one can conclude the proof also when $i \in \mathcal{I}_M$. $\qquad \square$

Note that, adapting the estimates presented in [76] in a way more general framework, one can obtain analogous results. On the other hand, for the readers' convenience here the computations are described, explicitly, in our specific case.

Remark 3.3.7. *If $\gamma \in C^2$, in view of the strict convexity or concavity around any $\bar{\xi}_i$, $i \in \mathcal{I}$, it is straightforward that results analogous to Lemmas 3.3.3, 3.3.4 and 3.3.6 can be obtained for any $\tilde{\delta} \in (0, \delta)$. Taking into account this additional consideration, one can provide the variational interpretation of Lemma 3.3.6 in the C^2-case. For large enough energies, the minimal critical points $\bar{\xi}_i$ with $i \in \mathcal{I}_m$ are also strict local minima for the total Jacobi length of the arc concatenation connecting $\gamma(\xi_E)$ to $\gamma(\xi_I)$, either starting with an inner or an outer arc. Conversely, the maximal points $\bar{\xi}_i$, $i \in \mathcal{I}_M$, are also strict maxima for the total Jacobi length.*

Remark 3.3.8. *From the estimates* (3.3.5) *one can observe that, as far as h is large, the dominant quantity in the derivatives of the total Jacobi length is the one related to the inner arc. In general, this fact holds also for the total length itself and not only for its derivatives.*

3.3.2   Existence of suitable periodic trajectories

The present section is devoted to the statement and proof of a key result, which will allow, in Section 3.3.3, to prove the existence of a symbolic dynamics for our model. In particular, we will prove that, for every finite (admissible, in a sense that will be specified in a moment) sequence of suitable symbols, there exists a periodic trajectory which realizes it in the sense described in Definition 3.1.4. The first step is to define our *alphabet*, as well as the rules to build admissible words.

For every $i \in \mathcal{I}$ as in Assumption 3.3.1, define

$$I^{(i)} = (\bar{\xi}_i - \delta, \bar{\xi}_i + \delta),$$

with δ as in Assumption 3.3.2. Let $n \in \mathbb{N}$, $n \geq 1$, and define the set $\mathbb{L}_n \subset \mathcal{I}^n$ of the admissible words of length n as

$$\mathbb{L}_n = \left\{\underline{\ell} = (\ell_0, \ldots, \ell_{n-1}) : \ell_i \in \mathcal{I} \text{ and } \forall i = 0, \ldots, n-1 \quad \ell_{(i+1) \bmod n} \in NA\left(\ell_i\right)\right\},$$

where the sets $NA\left(\ell_i\right)$ have been introduced in Definition 3.2.4. By Assumption 3.3.1, for every $n \in \mathbb{N} \setminus \{0\}$ the set $\mathbb{L}_n$ is not empty. Taking the union of such sets over $n \in \mathbb{N}$,

we obtain the set of the admissible finite words

$$\mathbb{L} = \bigcup_{n \in \mathbb{N} \setminus \{0\}} \mathbb{L}_n. \tag{3.3.6}$$

For any fixed $\underline{\ell} \in \mathbb{L}$ we define the $(2n+1)-$dimensional open rectangle

$$\mathbb{U}_{\underline{\ell}} = (I_0 \times I_0) \times (I_1 \times I_1) \times \cdots \times (I_{n-1} \times I_{n-1}) \times I_0, \tag{3.3.7}$$

where for all $k = 0, \ldots, n-1$ one has $I_k \doteq I^{(\ell_k)} = (\bar{\xi}_{\ell_k} - \delta, \bar{\xi}_{\ell_k} + \delta)$. The rectangle $\mathbb{U}_{\underline{\ell}}$ is the domain of (the parameters of) the transition points between the inner and outer arcs of the periodic solutions we are searching for. To this purpose, define the closed set

$$\mathbb{S}_{\underline{\ell}} = \left\{ \underline{\xi} = (\xi_0, \ldots, \xi_{2n}) \in \overline{\mathbb{U}}_{\underline{\ell}} : \xi_0 = \xi_{2n} \right\}, \tag{3.3.8}$$

where the identification between the first and last points takes into account the periodicity of the searched trajectory.

Let us point out that for every $j = 0, \ldots, n-1$ the points ξ_{2j} and ξ_{2j+1} belong to the same interval I_j, while the points ξ_{2j+1} and ξ_{2j+2} belong to possibly different intervals I_j and $I_{(j+1) \bmod n}$ and correspond to non-antipodal points in $\mathbb{R}^2$. Since our complete dynamics starts with an outer arc, for every $j = 0, \ldots, n-1$ the points $\gamma(\xi_{2j})$ and $\gamma(\xi_{2j+1})$ must be connected by an outer arc, while we will search for inner arcs connecting $\gamma(\xi_{2j+1})$ and $\gamma(\xi_{2j+2})$. We observe that this justifies also the claim of Theorem 3.2.3, where we proved the existence of the outer dynamics only for endpoints belonging to the same neighbourhood of a generic central configuration $\gamma(\bar{\xi}_i)$. In particular, in view of Theorems 3.2.3 and 3.2.7, for every $j = 0, \ldots, n-1$ the arcs

$$z_E(\cdot; \gamma(\xi_{2j}), \gamma(\xi_{2j+1})), \quad z_I(\cdot; \gamma(\xi_{2j+1}), \gamma(\xi_{2j+2}); h) \tag{3.3.9}$$

exist and are unique, as far as h is sufficiently large. Moreover, we recall that the uniqueness of z_I is related to the topological characterization *(TnT)* introduced in Definition 2.2.2.

Our purpose is to determine periodic trajectories for the complete dynamics as critical points of the functional $W_{\underline{\ell}}(\cdot; h) : \mathbb{S}_{\underline{\ell}} \to \mathbb{R}$ defined as

$$W_{\underline{\ell}}(\underline{\xi}; h) = \sum_{j=0}^{n-1} S_E(\xi_{2j}, \xi_{2j+1}) + \sum_{j=0}^{n-1} S_I(\xi_{2j+1}, \xi_{2j+2}; h). \tag{3.3.10}$$

The relation between sequences in $\mathbb{S}_{\underline{\ell}}$ and periodic trajectories for the complete dynamics can be built as follows: given $\underline{\xi} \in \mathbb{S}_{\underline{\ell}}$, the corresponding periodic orbit $z(\cdot)$ is the concatenation of the arcs in Eq. (3.3.9), which is unique in view of Theorems 3.2.3 and 3.2.7. More precisely, one can give the following definition.

Definition 3.3.9. *Given* $n \in \mathbb{N}$, $\underline{\ell} \in \mathbb{L}_n$, $\underline{\xi} \in \mathbb{S}_{\underline{\ell}}$ *and* $h > h_0$, *let us consider the unique arcs enlisted in Eq.* (3.3.9) *which connect pairs of subsequent points; with reference to Theorems 3.2.3 and 3.2.7, let us define*[4] *for every* $j = 0, \ldots, n-1$

$$T_E^{(j)} \doteq T_E\left(\gamma(\xi_{2j}), \gamma(\xi_{2j+1})\right), \quad T_I^{(j)} \doteq T_I\left(\gamma(\xi_{2j+1}), \gamma(\xi_{2j+2}); h\right).$$

and the partial sums $T^{(j)} \doteq \sum_{k=0}^{j} T_E^{(k)} + T_I^{(k)}$. *Setting* $T^{(-1)} \doteq 0$, *consider the concatenation* $z\left(\cdot; \underline{\xi}; h\right) : \left[0, T^{(n-1)}\right] \to \mathbb{R}^2$ *where, for every* $j = 0, \ldots, n-1$,

$$
\begin{cases}
z\left(s; \underline{\xi}; h\right) \doteq z_E\left(s - T^{(j-1)}; \gamma\left(\xi_{2j}\right), \gamma\left(\xi_{2j+1}\right)\right) & s \in \left[T^{(j-1)}, T^{(j-1)} + T_E^{(j)}\right] \\[2ex]
z\left(s; \underline{\xi}; h\right) \doteq z_I\left(s - T^{(j-1)} - T_E^{(j)}; \gamma\left(\xi_{2j+1}\right), \gamma\left(\xi_{2j+2}\right); h\right) & s \in \left[T^{(j-1)} + T_E^{(j)}, T^{(j)}\right]
\end{cases}.
$$

The function $z\left(\cdot; \underline{\xi}; h\right)$ *is trivially continuous, and, since* $z\left(0; \underline{\xi}; h\right) = z\left(T^{(n-1)}; \underline{\xi}; h\right)$, *it is also periodic. We can then extend it by periodicity and, with an abuse of notation, suppose* $z\left(\cdot; \underline{\xi}; h\right) : \mathbb{R} \to \mathbb{R}^2$.

We stress that, except for very special cases, concatenations are not C^1, as, at every transition point, the Snell's law determines a deflection of the velocity vector.

We can give a necessary and sufficient condition for concatenations introduced in Definition 3.3.9 to be an admissible trajectories for the complete dynamics: the Snell's law must be satisfied at every transition point $P_i \doteq \gamma(\xi_i)$, $i = 0, \ldots, 2n$. More precisely, for every $j = 0, \ldots, n-1$,

$$
\begin{aligned}
\sqrt{V_E\left(P_{2j+1}\right)} \; & \frac{z_E'\left(T_E^{(j)}; P_{2j}, P_{2j+1}\right)}{\left\|z_E'\left(T_E^{(j)}; P_{2j}, P_{2j+1}\right)\right\|} \cdot \dot{\gamma}(\xi_{2j+1}) \\
& = \sqrt{V_I\left(P_{2j+1}\right)} \frac{z_I'\left(0; P_{2j+1}, P_{2j+2}; h\right)}{\left\|z_I'\left(0; P_{2j+1}, P_{2j+2}; h\right)\right\|} \cdot \dot{\gamma}(\xi_{2j+1}), \\[2ex]
\sqrt{V_I\left(P_{2j+2}\right)} \; & \frac{z_I'\left(T_I^{(j)}; P_{2j+1}, P_{2j+2}; h\right)}{\left\|z_I'\left(T_I^{(j)}; P_{2j+1}, P_{2j+2}; h\right)\right\|} \cdot \dot{\gamma}(\xi_{2j+2}) \\
& = \sqrt{V_E\left(P_{2j+2}\right)} \frac{z_E'\left(0; P_{2j+2}, P_{2j+3}\right)}{\left\|z_E'\left(0; P_{2j+2}, P_{2j+3}\right)\right\|} \cdot \dot{\gamma}(\xi_{2j+2}),
\end{aligned}
\tag{3.3.11}
$$

where, with an abuse of notation, $P_{2n+1} = P_1$.

Remark 3.3.10. *We stress that, fixed* h *and* $\underline{\xi}$, *the concatenation* $z\left(\cdot; \underline{\xi}; h\right)$ *is uniquely determined: this means that the validity of conditions* (3.3.11) *depends only on the transition points determined by* $\underline{\xi}$ *and on the inner energy.*

---

[4] Here the T_E and T_I denote the final time respectively of the outer and inner arcs defined in Theorems 3.2.3 and 3.2.7.

With reference to Section 1.2, the functional $W_{\underline{\ell}}(\cdot; h)$ introduced in Eq. (3.3.10) can be interpreted as the *total Jacobi length* of the trajectory having transition points in $P_0, \ldots, P_{2n-1}$. Actually, there exists a core relation between trajectories for the complete dynamics and critical points for the total Jacobi length as a function of the parameters ξ of γ. Let us in fact suppose that, fixed $\underline{\ell} \in \mathbb{L}$ and $h > h_0$, the functional $W_{\underline{\ell}}(\cdot; h)$ admits a critical point in $\mathbb{S}_{\underline{\ell}}$, namely, $\exists \hat{\underline{\xi}} = \left(\hat{\xi}_0, \ldots, \hat{\xi}_{2n}\right) \in \mathbb{S}_{\underline{\ell}}$ such that $\nabla W_{\underline{\ell}}\left(\hat{\underline{\xi}}; h\right) = \underline{0}$. Then, for every $i = 0, \ldots, n-1$

$$
\begin{aligned}
\frac{\partial W_{\underline{\ell}}}{\partial \xi_{2i}}\left(\hat{\underline{\xi}}; h\right) &= \partial_b S_I\left(\hat{\xi}_{(2i-1) \bmod 2n}, \hat{\xi}_{2i}; h\right) + \partial_a S_E\left(\hat{\xi}_{2i}, \hat{\xi}_{2i+1}\right) = 0 \\
\frac{\partial W_{\underline{\ell}}}{\partial \xi_{2i+1}}\left(\hat{\underline{\xi}}; h\right) &= \partial_b S_E\left(\hat{\xi}_{2i}, \hat{\xi}_{2i+1}\right) + \partial_a S_I\left(\hat{\xi}_{2i+1}, \hat{\xi}_{2i+2}; h\right) = 0.
\end{aligned}
\tag{3.3.12}
$$

With reference to Definition 3.3.9, consider now the concatenation $z\left(\cdot; \hat{\underline{\xi}}; h\right)$. Comparing Eq. (3.3.12) with the expressions of $\partial_{a/b} S_{I/E}$ computed in Eq. (1.2.6), one can conclude that the stationarity of $\hat{\underline{\xi}}$ is equivalent to state that z satisfies the Snell's law at every transition point: this means that *the concatenation is an admissible periodic trajectory for the complete dynamics if and only if* $\nabla W_{\underline{\ell}}\left(\hat{\underline{\xi}}; h\right) = \underline{0}$. The above reasoning justifies why the search for critical points of $W_{\underline{\ell}}(\cdot; h)$ and for periodic trajectories are equivalent. Provided that h is large enough, the existence of a critical point of $W_{\underline{\ell}}(\cdot; h)$ is a straightforward consequence of the following classical result.

Theorem 3.3.11 (Poincaré-Miranda Theorem, [33])**.** *Let $F_1, \ldots, F_d$ d-functions in the variables $(x_1, \ldots, x_d)$ continuous on the d-dimensional hypercube*

$$
R = \{(x_1, \ldots, x_d) \mid |x_k| \leq L \text{ for every } k = 1, \ldots, d\}
$$

and such that for every $k = 1, \ldots, d$

$$
\begin{cases}
F_k\left(x_1, \ldots, x_d\right)_{|x_k = -L} \geq 0, \\
F_k\left(x_1, \ldots, x_d\right)_{|x_k = L} \leq 0.
\end{cases}
\tag{3.3.13}
$$

Then there exists at least a solution in R of

$$
F_k(x_1, \ldots, x_d) = 0 \quad \text{for every } k = 1, \ldots, d.
\tag{3.3.14}
$$

Proposition 3.3.12. *Given $\underline{\ell} \in \mathbb{L}$, the total Jacobi length $W_{\underline{\ell}}(\cdot; h)$ admits a critical point $\hat{\underline{\xi}}$ in $S_{\underline{\ell}}^{\circ}$ provided $h > \bar{h}_1$, where $\bar{h}_1$ is introduced in Lemma 3.3.6.*

Proof. This proof relies on a direct application of Poincaré-Miranda Theorem; to ease the notation and the overall proof, some considerations can be made:

- if there is $k \in 1, \ldots, d$ such that the inequalities Eq. (3.3.13) are satisfied with the opposite signs, taking $-F_k$ instead of F_k the Theorem remains valid. As a matter of fact, the key hypothesis is that any function $F_{\bar{j}}$, is continuous and changes its sign on the boundaries of $\bar{j}$-th edge $\left\{(x_1 \ldots, x_d) \in R \mid |x_{\bar{j}}| \leq L\right\}$;

- with the change of variables $y_k = x_k - c_k$, $k = 1, \ldots, d$, one can prove that the Theorem is valid also in hypercubes centered at any point $(c_1, \ldots, c_d) \in \mathbb{R}^d$;

- if the inequalities Eq. (3.3.13) are strict, the solution of Eq. (3.3.14) must lie in the interior of R.

Given these premises, let us fix $\underline{\ell} \in \mathbb{L}$ and set $n = |\underline{\ell}|$. Recalling the definitions of $\mathbb{U}_{\underline{\ell}}$ and $\mathbb{S}_{\underline{\ell}}$ given in Eqs. (3.3.7) and (3.3.8) and of the corresponding intervals I_i, $i = 0, \ldots, n-1$, we know that, if $\underline{\xi} = (\xi_0, \ldots, \xi_{2n}) \in \mathbb{S}_{\underline{\ell}}$, then $\xi_{2i}, \xi_{2i+1} \in \bar{I}_i = \left[\bar{\xi}_{\ell_i} - \delta, \bar{\xi}_{\ell_i} + \delta \right]$, where δ is defined as in Assumption 3.3.2. Let us now set $d = 2n$ and $R = \Pi_{i=0}^{n-1} \left(\bar{I}_i \times \bar{I}_i \right)$. For every $i = 0, \ldots, n-1$ and every $\underline{\xi} \in \mathbb{S}_{\underline{\ell}}$, define

$$F_{2i}\left(\underline{\xi}\right) = \partial_b S_I \left(\xi_{2i-1}, \xi_{2i}; h\right) + \partial_a S_E \left(\xi_{2i}, \xi_{2i+1}\right),$$
$$F_{2i+1}\left(\underline{\xi}\right) = \partial_b S_E \left(\xi_{2i}, \xi_{2i+1}\right) + \partial_a S_I \left(\xi_{2i+1}, \xi_{2i+2}; h\right),$$

where $\xi_{-1} = \xi_{2n-1}$. Computing the above functions on the hypercube's edges, one has that for every $i = 0, \ldots, n-1$,

$$\begin{aligned} F_{2i}\left(\underline{\xi}\right)_{|\xi_{2i}=\bar{\xi}_{\ell_i}\pm\delta} &= \partial_b S_I \left(\xi_{2i-1}, \bar{\xi}_{\ell_i} \pm \delta\right) + \partial_a S_E \left(\bar{\xi}_{\ell_i} \pm \delta, \xi_{2i+1}\right), \\ F_{2i+1}\left(\underline{\xi}\right)_{|\xi_{2i+1}=\bar{\xi}_{\ell_i}\pm\delta} &= \partial_b S_E \left(\xi_{2i}, \bar{\xi}_{\ell_i} \pm \delta\right) + \partial_a S_I \left(\bar{\xi}_{\ell_i} \pm \delta, \xi_{2i+2}\right). \end{aligned} \tag{3.3.15}$$

If $h > \bar{h}_1$, Lemma 3.3.6 ensures that for every $i = 0, \ldots, n-1$

$$F_{2i}\left(\underline{\xi}\right)_{|\xi_{2i}=\bar{\xi}_{\ell_i}-\delta} \cdot F_{2i}\left(\underline{\xi}\right)_{|\xi_{2i}=\bar{\xi}_{\ell_i}+\delta} < 0, \quad F_{2i+1}\left(\underline{\xi}\right)_{|\xi_{2i+1}=\bar{\xi}_{\ell_i}-\delta} \cdot F_{2i+1}\left(\underline{\xi}\right)_{|\xi_{2i+1}=\bar{\xi}_{\ell_i}+\delta} < 0$$

one can then apply Poincaré-Miranda Theorem to obtain a solution $\hat{\underline{\xi}} \in \mathbb{S}_{\underline{\ell}}^{\circ}$ of system (3.3.14). By the choice of the functions F_k and from direct computations, it results that $\nabla W_{\underline{\ell}}\left(\hat{\underline{\xi}}; h\right) = 0$. $\qquad\square$

The main result of this section now follows straightforwardly.

Theorem 3.3.13. *Given $\underline{\ell} \in \mathbb{L}$ and $h > \bar{h}_1$, there exists $\hat{\underline{\xi}} \in \mathbb{S}_{\underline{\ell}}^{\circ}$ and a periodic trajectory $z\left(\cdot; \hat{\underline{\xi}}; h\right)$ for the complete dynamics which realizes the word $\underline{\ell}$, in the sense that it connects $\gamma\left(\hat{\xi}_0\right), \ldots, \gamma\left(\hat{\xi}_{2n-1}\right)$.*

3.3.3 Construction of the symbolic dynamics

We are finally ready to prove the existence of a symbolic dynamics for our refractive billiard model; in particular, we will construct a surjective and continuous application between a suitable set of initial conditions of trajectories and admissible bi-infinite words. To this end let us define the energy shell for the external dynamics

$$\Xi \doteq \left\{ (\xi, v) \colon \xi \in [0, L], v \in \mathbb{R}^2, \frac{1}{2}\|v\|^2 - V_E(\gamma(\xi)) = 0 \right\}.$$

We now introduce the sets of initial conditions in Ξ for which:

- the parameter ξ belongs to $\overline{I^{(r)}}$, for a fixed $r \in \mathcal{I}$, where $I^{(r)}$ has been introduced at the beginning of Section 3.3.2,

- the velocity vector points respectively outward or inward the domain D.

More precisely, we define

$$\Xi_r^+ \doteq \left\{ (\xi, v) \in \Xi \colon \xi \in \overline{I^{(r)}} \text{ and } \langle n(\xi), v \rangle > 0 \right\}$$
$$\Xi_r^- \doteq \left\{ (\xi, v) \in \Xi \colon \xi \in \overline{I^{(r)}} \text{ and } \langle n(\xi), v \rangle < 0 \right\},$$

where $n(\xi)$ is the outward-pointing normal unit vector to γ in $\gamma(\xi)$.

Since in our model a crossing through γ implies a refraction of the trajectory, it is convenient to define analytically a *refraction map* which of course depends on the parameter ξ and the energy jump h. From now on we assume $h > \bar{h}_1$, where $\bar{h}_1$ is the threshold value introduced in Lemma 3.3.6 and used in Proposition 3.3.12.

Definition 3.3.14. *Fixed $\xi \in [0, L]$ and $h > 0$, we define the sets*

$$B_E^\xi \doteq \left\{ v \in \mathbb{R}^2 \colon \|v\|^2 = 2V_E(\gamma(\xi)) \right\} \qquad B_I^\xi \doteq \left\{ v \in \mathbb{R}^2 \colon \|v\|^2 = 2V_I(\gamma(\xi)) \right\}$$

and the refraction map

$$R_{EI}(\cdot; \xi, h) \colon B_E^\xi \to B_I^\xi$$
$$v = a\,t(\xi) + b\,n(\xi) \mapsto R_{EI}(v; \xi, h) = a\,t(\xi) + sgn(b)\sqrt{2V_I(\gamma(\xi)) - a^2}\,n(\xi)$$

where we recall that $t(\xi)$ is the tangent unit vector to γ in $\gamma(\xi)$.

Remark 3.3.15. *Some words are due to understand the previous definition. First of all, we observe that for every $\xi \in [0, L]$ the vectors $(t(\xi), n(\xi))$ form a orthonormal basis of the plane. Moreover, since $v \in B_E^\xi$ (namely, $a^2 + b^2 = 2V_E(\gamma(\xi))$) and $V_I > V_E$ everywhere in the punctured plane, the quantity $2V_I(\gamma(\xi)) - a^2$ is always positive, so that the refraction map is well defined in its domain. Furthermore, it is clear that the normal component preserves the sign, so that the image of an outward (resp. inward)-pointing vector is still outward (resp. inward)-pointing. It can also be easily proved that given a vector $v \in B_E^\xi$, its image $R_{EI}(v; \xi, h)$ is the unique vector with which v satisfies the refraction Snell's law (1.2.11).*

The refraction map R_{EI} is clearly injective; if we introduce the set

$$\tilde{B}_I^\xi \doteq \left\{ v \in \mathbb{R}^2 \colon \|v\|^2 = 2V_I(\gamma(\xi)) \text{ and } \langle v, t(\xi) \rangle \leq \sqrt{2V_E(\gamma(\xi))} \right\}$$

then $R_{EI}(\cdot; \xi, h) \colon B_E^\xi \to \tilde{B}_I^\xi$ is invertible.

Remark 3.3.16. *As already pointed out in Section 1.2, the Snell's Law (1.2.11) depends on the point $z \in \partial D$, where the inequality $V_I(z) > V_E(z)$ is always satisfied. Hence, the*

transition from inside to outside needs that

$$|\alpha_I| \leq \arcsin\left(\frac{\sqrt{V_E(z)}}{\sqrt{V_I(z)}}\right) = \alpha_{crit}.$$

The restriction of the map $R_{EI}(\cdot; \xi, h)$ to the set $\tilde{B}_I^\xi$ corresponds to require that the angle between an internal velocity and $n(\xi)$ is less than the critical value α_{crit}.

Let us now fix $(\xi, v) \in \Xi_r^+$ and follow step by step the trajectory of the complete dynamics starting from the initial condition $(\gamma(\xi), v)$. Assumptions on the allowed initial conditions will become more and more restrictive as the dynamics proceeds in order to obtain a final set

$$X \subseteq \bigcup_{r \in \mathcal{I}} \Xi_r^+ \qquad (3.3.16)$$

for which it is possible to construct a symbolic dynamics.

First of all, let us consider the flow $\Phi_E^s(\gamma(\xi), v)$ generated by the Cauchy problem associated to the outer potential, that is

$$\begin{cases} (HS_E)[z(s)] \\ z(0) = \gamma(\xi), \ z'(0) = v. \end{cases} \qquad (3.3.17)$$

As customary dealing with this kind of systems, we consider the projections of such flow onto the configuration and the velocity space respectively:

$$\Pi_z \Phi_E^s(\cdot, \cdot), \qquad \Pi_v \Phi_E^s(\cdot, \cdot).$$

Let us now consider the outer arc with initial condition $(\gamma(\xi), v)$. Then we can define the set

$$\mathbb{T}^-(\xi, v) \doteq \left\{ s_1 > 0 \ \middle| \ \begin{array}{l} \Phi_E^{s_1}(\gamma(\xi), v) = (\gamma(\xi_1), v_1) \text{ for some } (\xi_1, v_1) \in \Xi_r^- \\ \Pi_z \Phi_E^s(\gamma(\xi), v) \notin \bar{D} \text{ for every } s \in (0, s_1) \end{array} \right\}. \qquad (3.3.18)$$

which contains at most one element. In view of Theorem 3.2.3, there holds

$$\left\{ (\xi, v) \in \Xi_r^+ : \mathbb{T}^-(\xi, v) \neq \emptyset \right\} \neq \emptyset \quad \text{for all } r \in \mathcal{I}.$$

Let us now suppose that $\mathbb{T}^-(\xi, v) = \{s_1\} \neq \emptyset$ and call

$$(\gamma(\xi_1), v_1) \doteq \Phi_E^{s_1}(\gamma(\xi), v).$$

To proceed with the inner dynamics we need to refract our arc: let us then consider the initial condition $(\gamma(\xi_1), v_1') \doteq (\gamma(\xi_1), R_{EI}(v_1; \xi_1, h))$ to start with an inner arc. Exactly
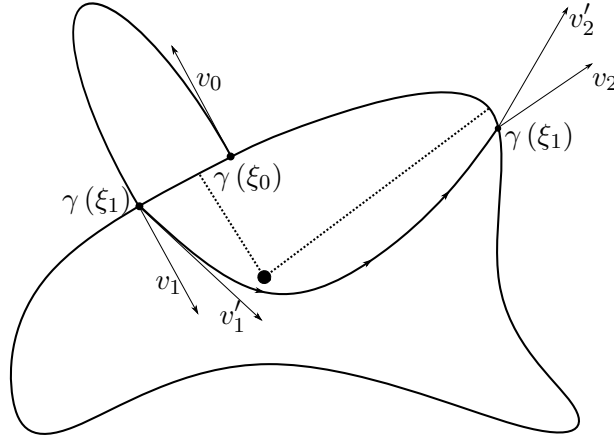
Fig. 3.6 An example of concatenation starting from a point $(\gamma(\xi), v)$, with $(\xi, v) \in A$.

as for the outer dynamics, we consider the flow associated to the inner problem

$$
\begin{cases}
(HS_I)[z(s)] \\
z(0) = \gamma(\xi_1), \ z'(0) = v_1',
\end{cases}
\tag{3.3.19}
$$

and the set

$$
\mathbb{T}^+(\xi, v) \doteq \left\{ s_2 > 0 \ \middle| \ \begin{array}{l} \Phi_I^{s_2}(\gamma(\xi_1), v_1') = (\gamma(\xi_2), v_2') \text{ for some } (\xi_2, v_2') \in \bigcup_{r' \in NA(r)} \Xi_{r'}^+ \\[2mm] v_2' \in \tilde{B}_I^{\xi_2} \\[2mm] \Pi_z \Phi_I^s(\gamma(\xi_1), v_1') \in D \text{ for every } s \in (0, s_2) \end{array} \right\},
\tag{3.3.20}
$$

where $NA(r)$ has been introduced in Definition 3.2.4. Once more $\mathbb{T}^+(\xi, v)$ has at most one element and we can extend its definition to every pair $(\xi, v) \in \Xi_r^+$ by requiring that

$$
\mathbb{T}^-(\xi, v) = \emptyset \quad \implies \quad \mathbb{T}^+(\xi, v) = \emptyset.
$$

In view of Theorem 3.3.13, there holds again that

$$
\left\{ (\xi, v) \in \Xi_r^+ : \mathbb{T}^+(\xi, v) \neq \emptyset \right\} \neq \emptyset \quad \text{for all } r \in \mathcal{I};
$$

indeed, it is sufficient to consider a word $\underline{\ell}$ of length at least 2 with the first element equal to r and take the initial condition of the corresponding trajectory.

Let now $(\xi, v) \in \Xi_r^+$ be such that $\mathbb{T}^+(\xi, v) \neq \emptyset$ and let (ξ_2, v_2') as in Eq. (3.3.20). As $v_2' \in \tilde{B}_I^{\xi_2}$ we are allowed to define

$$
v_2 \doteq R_{EI}^{-1}(v_2'; \xi_2, h)
\tag{3.3.21}
$$

so that $(\xi_2, v_2) \in \Xi_{r'}^+$ and we have the initial condition for a second outer arc.

The following non-empty set contains pairs (ξ, v) which satisfy all assumptions made up to now (see Figure 3.6).

Definition 3.3.17. *Let us define the set $A \subset [0, L] \times \mathbb{R}^2$ as the set of pairs (ξ, v) such that*

1. *$(\xi, v) \in \Xi_r^+$ for some $r \in \mathcal{I}$;*

2. *there exist $s_1 \in \mathbb{T}^-(\xi, v)$ and $s_2 \in \mathbb{T}^+(\xi, v)$;*

The set A contains initial conditions for which it is possible to construct a complete concatenation outer-inner arc. It is then worthwhile to define the first return map on A as follows.

Definition 3.3.18. *Let us define the first return map[5]*

$$F \colon A \to F(A), \qquad F(\xi, v) \doteq (\xi_2, v_2) \tag{3.3.22}$$

where ξ_2 and v_2 are introduced respectively in (3.3.20) and (3.3.21).

The function F is a bijection by the uniqueness of the arcs claimed in Theorems 3.2.3 and 3.2.7. Now, by recurrence, we construct the set X as the set on which all the positive and negative iterates of the map F are defined. Let

$$X_1^+ \doteq A \quad \text{and} \quad X_1^- \doteq A \cap F(A) = X_1^+ \cap F(X_1^+)$$

and observe that $X_1^- \neq \emptyset$ (by Theorem 3.3.13); moreover both F and F^{-1} are well defined on X_1^-. Then we introduce the (still non empty) sets

$$X_2^+ \doteq X_1^- \cap F^{-1}(X_1^-) \quad \text{and} \quad X_2^- \doteq X_2^+ \cap F(X_2^+)$$

so that the maps F, F^{-1} and F^2 are well defined on both X_2^+ and X_2^-, and on X_2^- is also defined the map F^{-2}. Actually, for any positive integer $k \geq 2$ we introduce the non-empty sets

$$X_k^+ \doteq X_{k-1}^- \cap F^{-1}(X_{k-1}^-) \quad \text{and} \quad X_k^- \doteq X_k^+ \cap F(X_k^+)$$

so that on X_k^- the iterates F^j, with $j \in \mathbb{Z}$ and $|j| \leq k$, are well defined (in our notation $F^0 = \mathrm{Id}$). We are then ready to define the set of initial conditions in Ξ_r^+ that generate trajectories with an infinite number of transversal intersections with ∂D, namely

$$X \doteq \bigcap_{k \in \mathbb{N}} X_k^-.$$

Of course, the set X is non empty, and on X are defined the iterates F^j, for any integer $j \in \mathbb{Z}$. As the set X is invariant for the first return map F (see also the remark below), it is convenient to consider the restriction

$$\mathcal{F} \doteq F|_X$$

---

[5]Note that, in this chapter, the first return map is expressed in terms of the velocity instead of its angle with the normal vector. Of course, the definitions are completely equivalent.

Remark 3.3.19. *Let us observe that the set X is the set of initial conditions which generate trajectories for the complete dynamics that cross ∂D an infinite number of times. For this reason, the double-inclusion $F(X) = X$ holds. Indeed, for every $x \doteq (\xi, v) \in X$ it is straightforward to understand that both $F(x)$ and $F^{-1}(x)$ belong to X.*

Let us now consider the set of bi-infinite admissible words

$$\mathcal{L} \doteq \left\{ \underline{\ell} \in \mathcal{I}^{\mathbb{Z}} \colon \ell_{j+1} \in NA(\ell_j), \, \forall j \in \mathbb{Z} \right\} \tag{3.3.23}$$

endowed with the metric (here $\rho(i, j) = 0$ if $i = j$ and $\rho(i, j) = 1$ if $i \neq j$)

$$d(\underline{\ell}, \underline{m}) \doteq \sum_{k \in \mathbb{Z}} \frac{\rho(\ell_k, m_k)}{4^k}, \tag{3.3.24}$$

we refer to the book [30] for complete treatment on the subject. It is straightforward to prove that, with this metric, the subset of the periodic bi-infinite words

$$\mathbb{L}_P \doteq \{ \underline{\ell} \in \mathcal{L} \colon \underline{\ell} \text{ is periodic} \}$$

is dense in $\mathcal{L}$. Furthermore, we observe that the elements of $\mathbb{L}_P$ are the periodic extensions of the finite admissible words of the set $\mathbb{L}$ introduced in Eq. (3.3.6).

In order to project the set of initial conditions X into the set of admissible words $\mathcal{L}$ we introduce the map $\chi \colon X \to \mathcal{I}$ such that

$$\chi(\xi, v) = r \iff \xi \in \overline{I^{(r)}} \tag{3.3.25}$$

which associates to every (ξ, v) the index corresponding to the neighbourhood where ξ belongs.

Finally, we define the projection map

$$\pi \colon X \to \mathcal{L}, \quad (\xi, v) \mapsto \pi(\xi, v) = (\ell_j)_{j \in \mathbb{Z}} \quad \text{with } \ell_j = \chi(\mathcal{F}^j(\xi, v))$$

through which we are able to associate to every (ξ, v) the bi-infinite word realized by the trajectory of initial condition $(\gamma(\xi), v)$.
We can then consider the commutative diagram

$$\begin{array}{ccc} X & \xrightarrow{\mathcal{F}} & X \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ \mathcal{L} & \xrightarrow{\sigma_r} & \mathcal{L} \end{array}$$

where σ_r is the Bernoulli right-shift. In order to prove the existence of a symbolic dynamics for our model, we are left to show that the map π is a continuous surjection. The rest of this section is devoted to prove this result. Again, we recall that $\bar{h}_1$ is the threshold value for the energy jump h defined in Theorem 3.3.13.

Proposition 3.3.20. *If $h > \bar{h}_1$, the projection map π is surjective.*

Proof. Let us take a sequence $\underline{\ell} = (\ell_j)_{j \in \mathbb{Z}} \in \mathcal{L}$; our aim is to find $(\xi, v) \in X$ such that $\pi(\xi, v) = \underline{\ell}$, namely, the initial condition for a trajectory for the complete dynamics which realizes $\underline{\ell}$ in the sense of Theorem 3.3.13.

First of all, for every $n \in \mathbb{N}$ let us consider the truncated sequences

$$\underline{\ell}^{(n)} \doteq (\ell_{-n}, \ldots, \ell_n),$$

which are elements of $\mathbb{L}$ of length $2n + 1$. Since $h > \bar{h}_1$, by Theorem 3.3.13, for every $n \in \mathbb{N}$ there exists $\hat{\underline{\xi}}^{(n)} \in \mathbb{S}_{\underline{\ell}^{(n)}}$ such that the corresponding trajectory

$$z^{(n)}(\cdot) \doteq z\left(\cdot; \hat{\underline{\xi}}^{(n)}, h\right)$$

realizes the word $\underline{\ell}^{(n)}$. Since the trajectories $z^{(n)}(\cdot)$ are periodic, without loosing in generality and possibly with a time translation, we can assume that

$$z^{(n)}(0) \in \gamma(\overline{I^{(\ell_0)}}).$$

We now extend by periodicity $\hat{\underline{\xi}}^{(n)}$ to obtain a sequence in $\mathcal{L}$

$$\left(\underline{\xi}^{(n)}\right)_n, \qquad \underline{\xi}^{(n)} = \left(\xi_k^{(n)}\right)_{k \in \mathbb{Z}},$$

where, for every n, $\underline{\xi}^{(n)} \in \mathbb{L}_P$ is a $(4n + 2)$-periodic sequence. By construction,

$$\forall k \in \mathbb{Z}, \; \exists N_k > 0, \; \exists i_k \in \mathcal{I} \; : \; \xi_k^{(n)} \in \overline{I^{(i_k)}}, \; \forall n \geq N_k, \tag{3.3.26}$$

namely, fixed k, the points $\xi_k^{(n)}$ belong eventually to the same compact interval prescribed by the sequence $\underline{\ell}$.

We now start with the diagonal process that will imply the existence of a limit sequence of parameters in $[0, L]^{\mathbb{Z}}$. Let us fix $k = 0$ and consider the sequence $\left(\xi_0^{(n)}\right)_n$: by Eq. (3.3.26), the compactness of $\overline{I^{(\ell_0)}}$ (in this case $\ell_0 = i_0$) implies the existence of a convergent subsequence, that is

$$\exists \bar{\xi}_0 \in \overline{I^{(\ell_0)}}, \; \exists (\eta_{j,0})_j \subset \mathbb{N} \; : \; \eta_{j,0} \xrightarrow{j \to \infty} \infty \; \text{ and } \; \xi_0^{(\eta_{j,0})} \xrightarrow{j \to \infty} \bar{\xi}_0.$$

Now, going on with the same reasoning, for every integer $k \geq 1$ there exist $\bar{\xi}_{\pm k} \in \overline{I^{(i_{\pm k})}}$ and a subsequence $(\eta_{j,k})_j \subset (\eta_{j,k-1})_j$, $\eta_{j,k} \to \infty$, such that

$$\xi_i^{(\eta_{j,k})} \xrightarrow{j \to \infty} \bar{\xi}_i, \quad \text{for every } i \in \mathbb{Z} \text{ such that } |i| \leq k,$$

and in particular $\xi_{\pm k}^{(\eta_{j,k})} \xrightarrow{j \to \infty} \bar{\xi}_{\pm k}$.

We now define the diagonal index sequence

$$(a_n)_n \text{ such that } a_n \doteq \eta_{n,n}.$$

Since $a_n \to \infty$ and, for every k, it holds $(a_n)_n \subset (\eta_{j,k})_{j \in \mathbb{N}}$, we have that

$$\xi_i^{(a_n)} \xrightarrow{n \to \infty} \bar{\xi}_i, \quad \text{for every } i \in \mathbb{Z}.$$

Let us now define the limit sequence $\bar{\xi} \doteq (\bar{\xi}_i)_i$: by construction it results that $\bar{\xi}$ realizes the word ℓ, in the sense that

$$\bar{\xi}_{2j}, \bar{\xi}_{2j+1} \in \overline{I^{(\ell_j)}}, \quad \text{for every } j \in \mathbb{Z}.$$

In order to conclude the proof, we need to find the concatenation for the complete dynamics which realizes the limit sequence $\bar{\xi}$. To do that the admissibility of the word ℓ will be crucial. Indeed, since $h > \bar{h}_1$, Theorems 3.2.3 and 3.2.7 guarantee that for every $j \in \mathbb{Z}$ there exists a unique pair of arcs

$$z_E\left(\cdot; \gamma(\bar{\xi}_{2j}), \gamma(\bar{\xi}_{2j+1})\right) \quad \text{and} \quad z_I\left(\cdot; \gamma(\bar{\xi}_{2j+1}), \gamma(\bar{\xi}_{2j+2}); h\right).$$

We can then define the concatenation $\bar{z} \colon \mathbb{R} \to \mathbb{R}^2$ of these arcs as in Definition 3.3.9: our aim is now to verify that $\bar{z}$ is an admissible trajectory for the complete dynamics. By construction, $\bar{z}$ realizes the word ℓ, hence we are left to show that it satisfies the Snell's law at every transition point. This fact follows from the differentiable dependence of the arcs with respect to the endpoints; indeed, for every $j \in \mathbb{Z}$, we have the C^1 convergence

$$z_E\left(\cdot; \gamma\left(\xi_{2j}^{(a_n)}\right), \gamma\left(\xi_{2j+1}^{(a_n)}\right)\right) \to z_E\left(\cdot; \gamma(\bar{\xi}_{2j}), \gamma(\bar{\xi}_{2j+1})\right)$$

and

$$z_I\left(\cdot; \gamma\left(\xi_{2j+1}^{(a_n)}\right), \gamma\left(\xi_{2j+2}^{(a_n)}\right); h\right) \to z_I\left(\cdot; \gamma(\bar{\xi}_{2j+1}), \gamma(\bar{\xi}_{2j+2}); h\right).$$

Now the claim follows straightforwardly from the fact that the concatenations $\left(z^{(a_n)}\right)_n \subset \left(z^{(n)}\right)_n$ satisfy the Snell's law at every transition point.

Defining now $(\xi, v) \doteq (\bar{\xi}_0, \bar{z}'(0))$ we have that $\pi(\xi, v) = \ell$ and the surjectivity is proved. $\qquad \square$

At this point, the continuity of π is the last property to verify. Before proving it, let us start with a preliminary lemma, which ensures that the time intervals associated to every (outer or inner) arc are bounded from above uniformly with respect to the endpoints.

Lemma 3.3.21. *Let $h > \bar{h}_1$; then there exists $C > 0$ such that for every $j \in \mathcal{I}$, $\xi_0, \xi_1 \in \overline{I^{(j)}}$ and $\xi_2 \in \bigcup_{i \in NA(j)} \overline{I^{(i)}}$ it holds*

$$T_E(\xi_0, \xi_1) \leq C, \quad T_I(\xi_0, \xi_2; h) \leq C,$$

where T_E and T_I are as in Definition 3.3.9.

Proof. The proof follows easily by the continuous dependence of the outer and inner arcs with respect to variations of the endpoints and by the compactness of the intervals $\overline{I^{(j)}}$, $j \in \mathcal{I}$. $\qquad \square$

We are now ready to verify the continuity of the map π; we recall that the space of the admissible words $\mathcal{L}$ is endowed with the metric associated to the distance $d(\cdot, \cdot)$, defined in Eq. (3.3.24), while in X we will consider the usual Euclidean metric over $\mathbb{R}^3$.

Proposition 3.3.22. *If $h > \bar{h}_1$, the projection map π is continuous.*

Proof. Let us fix $(\xi_0, v_0) \in X$, and take $\epsilon > 0$. We are searching for $\delta > 0$ such that, if $(\xi, v) \in X$ satisfies $\|(\xi, v) - (\xi_0, v_0)\| < \delta$, then

$$d\left(\pi(\xi, v), \pi(\xi_0, v_0)\right) < \epsilon. \tag{3.3.27}$$

To this end, for every $k \in \mathbb{Z}$ it is convenient to define the $k - th$ projection map

$$\pi_k \colon X \to \mathcal{I}, \qquad \pi_k(\xi, v) \doteq \chi\left(\mathcal{F}^k(\xi, v)\right),$$

where χ is defined as in Eq. (3.3.25). Condition (3.3.27) translates in

$$\sum_{k \in \mathbb{Z}} \frac{\rho\left(\pi_k(\xi, v), \pi_k(\xi_0, v_0)\right)}{4^{|k|}} < \epsilon. \tag{3.3.28}$$

As the above series is always convergent, there exists $k_0 \in \mathbb{N}$ such that for every $(\xi, v) \in X$

$$\sum_{|k| \geq k_0} \frac{\rho\left(\pi_k(\xi, v), \pi_k(\xi_0, v_0)\right)}{4^{|k|}} \leq \sum_{|k| \geq k_0} \frac{1}{4^{|k|}} < \epsilon;$$

we will now prove that, if $\delta > 0$ is small enough and $\|(\xi, v) - (\xi_0, v_0)\| < \delta$, then

$$\sum_{|k| < k_0} \frac{\rho\left(\pi_k(\xi, v), \pi_k(\xi_0, v_0)\right)}{4^{|k|}} = 0, \tag{3.3.29}$$

and hence Eq. (3.3.27) is true. We observe that Eq. (3.3.29) is equivalent to require that, if $\underline{\ell} \doteq \pi(\xi, v)$ and $\underline{m} \doteq \pi(\xi_0, v_0)$, then $\ell_k = m_k$ for every $|k| < k_0$, namely, that the trajectories generated by (ξ, v) and (ξ_0, v_0) intersect the boundary ∂D in the same neighbourhoods in the first k_0 steps forward and backward. For δ small enough, this is a straightforward consequence of the continuous dependence on the initial data of both the inner and outer arc and the refraction law (along with its inverse): let us in fact define $z(\cdot)$ and $z_0(\cdot)$ as the trajectories for the complete dynamics such that $z(0) = \gamma(\xi)$, $z'(0) = v$, $z_0(0) = \gamma(\xi_0)$ and $z_0'(0) = v_0$. In view of Lemma 3.3.21, it is possible to find $a > 0$ such that both the trajectories cross ∂D at least $4k_0 + 3$ times within the time interval $[-a, a]$ (recall that, for every concatenation outer-inner arcs, we have exactly two intersections with ∂D): for every $\epsilon' > 0$, there exists then $\delta' > 0$ such that, if $\|(\xi, v) - (\xi_0, v_0)\| < \delta'$

$$\|z - z_0\|_{C^0([-a,a])} < \epsilon'.$$

As the intervals $\gamma\left(\overline{I^{(j)}}\right)$, $j \in \mathcal{I}$, are disjoint, for ϵ' small enough the two trajectories must cross the interface in the same neighbourhoods for $s \in [-a, a]$, and then, taking the corresponding δ', Eq.(3.3.29) holds. $\qquad \square$

Taking together Propositions 3.3.20 and 3.3.22, one can then prove Theorem 3.1.5 in the case of the refractive dynamics. As for the reflective case, all the previous considerations can be reproduced (actually, the absence of an outer dynamics leads to a simpler model), including the existence of a symbolic dynamics under the same admissibility hypotheses described in Assumption 3.3.1. This consideration concludes the proof of Theorem 3.1.5

3.4   Non-collisional dynamics

Going back to Theorem 3.2.7, we recall that the inner dynamics naturally includes collisional arcs. As a consequence, we can not *a priori* exclude the occurrence of collisions in the symbolic dynamics found by means of Theorem 3.1.5. Nevertheless, as we will prove in the present section, if we suitably restrict the set $\mathcal{L}$ we can construct a *non-collisional* symbolic dynamics.

Let us start dealing with periodic trajectories: let $\underline{\ell} \in \mathbb{L}$, as in Eq. (3.3.6), be an admissible word and, with reference to Theorem 3.3.13, $h > \bar{h}_1$. Following the reasoning of Section 3.3, one can then prove the existence of a sequence $\hat{\underline{\xi}} \in \mathbb{S}_{\underline{\ell}}^{\circ}$ and a corresponding trajectory $z\left(\cdot; \hat{\underline{\xi}}; h\right)$ which realizes the word $\underline{\ell}$, in the sense that the transition points $\gamma\left(\hat{\xi}_0\right), \ldots, \gamma\left(\hat{\xi}_{2|\underline{\ell}|-1}\right)$ belong to the segments of ∂D fixed by the sequence $\underline{\ell}$. It is possible that two or more subsequent elements of $\hat{\underline{\xi}}$ are equal: in view of Theorems 3.2.3 and 3.2.7, this implies the presence of homothetic outer arcs or collisional inner arcs. As we will see in this section, the presence of radial segments has strong consequences on the global structure of $z\left(\cdot; \hat{\underline{\xi}}; h\right)$, as it implies a reflection which involves every arc in the concatenation: this is possible only if $\underline{\ell}$ satisfies precise symmetry conditions.

Take now $\underline{\ell} = (\ell_0, \ldots, \ell_{n-1})$ and $h > \bar{h}_1$. As the trajectory $z\left(\cdot; \hat{\underline{\xi}}; h\right)$ is periodic, one can identify $\underline{\ell}$ with any of its shift, namely,

$$\tilde{\underline{\ell}} = \left(\tilde{\ell}_0, \ldots, \tilde{\ell}_{n-1}\right) = \left(\ell_{(k) \bmod n}, \ldots, \ell_{(n-1+k) \bmod n}\right) \tag{3.4.1}$$

for any $k \in \{1, \ldots, n-1\}$.

To describe in details the trajectory $z\left(\cdot; \hat{\underline{\xi}}; h\right)$, it is worth to keep trace not only of the transition points $\gamma\left(\hat{\underline{\xi}}_i\right)$, $i = 0, \ldots, 2n-1$, but also of the inner and outer velocity vectors at such points. To this aim, let us define the finite sequences $(v_i)_{i=0}^{2n}$, $(\tilde{v}_i)_{i=0}^{2n}$,
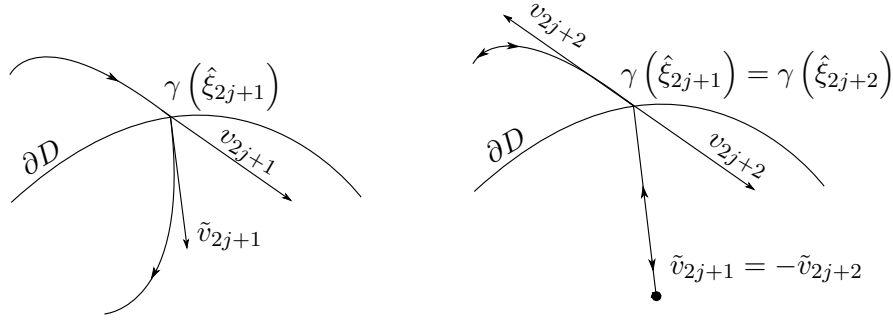
Fig. 3.7 Velocity vectors as defined in Eq. (3.4.2) in the non-collisional case (left) and in the collisional case (right). In the second case, the collisional inner arc forces a reflection in the adjacent outer arcs.

where, for every $j = 0, \ldots, n-1$ (see also Figure 3.7, left)[6],

$$v_{2j} = z'_E\left(0; \gamma\left(\hat{\xi}_{2j}\right), \gamma\left(\hat{\xi}_{2j+1}\right)\right), \qquad v_{2j+1} = z'_E\left(T; \gamma\left(\hat{\xi}_{2j}\right), \gamma\left(\hat{\xi}_{2j+1}\right)\right),$$

$$v_{2n} = v_0,$$

$$\tilde{v}_{2j+1} = z'_I\left(0; \gamma\left(\hat{\xi}_{2j+1}\right), \gamma\left(\hat{\xi}_{2j+2}\right); h\right), \quad \tilde{v}_{2j+2} = z'_I\left(T; \gamma\left(\hat{\xi}_{2j+1}\right), \gamma\left(\hat{\xi}_{2j+2}\right); h\right),$$

$$\tilde{v}_0 = \tilde{v}_{2n}.$$

$$(3.4.2)$$

We now provide, omitting the straightforward proof, necessary and sufficient conditions for the outer and inner arc to be respectively homothetic and collisional, depending on their endpoints (see Figure 3.7, right).

Lemma 3.4.1. *For any $j = 0, \ldots, n-1$:*

- $z_E\left(\cdot; \gamma\left(\hat{\xi}_{2j}\right), \gamma\left(\hat{\xi}_{2j+1}\right)\right)$ *is a homothetic outer arc if and only if $\hat{\xi}_{2j} = \hat{\xi}_{2j+1}$. In that case, one has also that $v_{2j} = -v_{2j+1}$ and the velocities are parallel to the vector $\gamma\left(\underline{\hat{\xi}}_{2j}\right)$;*

- $z_I\left(\cdot; \gamma\left(\hat{\xi}_{2j+1}\right), \gamma\left(\hat{\xi}_{2j+2}\right); h\right)$ *is a collision-ejection inner arc if and only if $\hat{\xi}_{2j+1} = \hat{\xi}_{2j+2}$. As in the outer case, whenever it happens one has that $\tilde{v}_{2j} = -\tilde{v}_{2j+1}$ and the velocities are parallel to the vector $\gamma\left(\underline{\hat{\xi}}_{2j+1}\right)$.*

The following Theorem underlines as the presence of collisional inner arcs can impact the overall structure of $z\left(\cdot; \underline{\hat{\xi}}; h\right)$.

Theorem 3.4.2. *Let $\underline{\ell} \in \mathbb{L}$ and $h > \bar{h}_1$; define $n \doteq |\underline{\ell}|$, and suppose $n > 1$. Let $\underline{\hat{\xi}} \in \mathbb{S}^\circ_{\underline{\ell}}$ and $z(\cdot) \doteq z\left(\cdot; \underline{\hat{\xi}}; h\right)$ as in Theorem 3.3.13, and suppose that the concatenation $z(\cdot)$ admits a collisional inner arc. Then:*

---

[6]From this moment on, for the sake of brevity and with an abuse of notation, we will denote with T the final time of an arc, omitting the dependence on the endpoints and without discerning between the inner and the outer case.

1. *if n is even, $z(\cdot)$ has another collisional arc and there exists a shift $\tilde{\underline{\ell}}$ which is symmetric with respect to the axis that separates $\tilde{\ell}_{n/2-1}$ from $\tilde{\ell}_{n/2}$;*

2. *if n is odd, then $z(\cdot)$ has a homothetic outer arc and there exists a shift $\tilde{\underline{\ell}}$ such that the finite sequence $(\ell_0, \dots, \ell_{n-1}, \ell_0)$ is symmetric with respect to the axis that separates $\tilde{\ell}_{(n+1)/2-1}$ from $\tilde{\ell}_{(n+1)/2}$.*

Before providing the proof of Theorem 3.4.2, let us give an explicit example, which will be useful to fix the notation and to clarify the reasoning used, in a completely general framework, below.

Example 3.4.3. Let $\underline{\ell} = (\ell_0, \ell_1, \ell_2, \ell_3) \in \mathbb{L}$ with $n = |\underline{\ell}| = 4$, and take $h > \bar{h}_1$: the sequence $\hat{\underline{\xi}}$ provided by Theorem 3.3.13 is then of the form $\hat{\underline{\xi}} = \left(\hat{\xi}_0, \dots, \hat{\xi}_8\right)$, with $\hat{\xi}_0 = \hat{\xi}_8$. Let us suppose that the corresponding concatenation $z\left(\cdot; \hat{\underline{\xi}}; h\right)$ has a collisional inner arc, and, in view of Eq. (3.4.1), suppose that this arc is given by $z_I\left(\cdot; \gamma\left(\hat{\xi}_3, \hat{\xi}_4; h\right)\right)$. From Lemma 3.4.1, this implies that $\hat{\xi}_3 = \hat{\xi}_4$ and $\tilde{v}_3 = -\tilde{v}_4$: by the Snell's law, one has then $v_3 = -v_4$. Recalling the definitions of v_j given in Eq. (3.4.2), we have then

$$z'_E\left(T; \gamma\left(\hat{\xi}_2\right), \gamma\left(\hat{\xi}_3\right)\right) = -z'_E\left(0; \gamma\left(\hat{\xi}_4\right), \gamma\left(\hat{\xi}_5\right)\right).$$

By the uniqueness of the solution of a Cauchy problem, one can then conclude that the arcs $z_E\left(\cdot; \gamma\left(\hat{\xi}_2\right), \gamma\left(\hat{\xi}_3\right)\right)$ and $z_E\left(\cdot; \gamma\left(\hat{\xi}_4\right), \gamma\left(\hat{\xi}_5\right)\right)$ parametrize the same curve but in the opposite sense. More precisely,

$$z_E\left(\cdot; \gamma\left(\hat{\xi}_2\right), \gamma\left(\hat{\xi}_3\right)\right) = z_E\left(T - \cdot; \gamma\left(\hat{\xi}_4\right), \gamma\left(\hat{\xi}_5\right)\right),$$

and, as a consequence, $\hat{\xi}_2 = \hat{\xi}_5$ and $v_2 = -v_5$. Using the same reasonings, possibly applied to the inner arc, one can prove that

$$\tilde{v}_2 = -\tilde{v}_5 \implies z_I\left(\cdot; \gamma\left(\hat{\xi}_1\right), \gamma\left(\hat{\xi}_2\right); h\right) = z_I\left(T - \cdot; \gamma\left(\hat{\xi}_5\right), \gamma\left(\hat{\xi}_6\right); h\right)$$

$$\implies \hat{\xi}_1 = \hat{\xi}_6 \text{ and } \tilde{v}_1 = -\tilde{v}_6$$

$$\implies z_E\left(\cdot; \gamma\left(\hat{\xi}_0\right), \gamma\left(\hat{\xi}_1\right)\right) = z_E\left(T - \cdot; \gamma\left(\hat{\xi}_6\right), \gamma\left(\hat{\xi}_7\right)\right) \implies \hat{\xi}_0 = \hat{\xi}_7 \text{ and } v_0 = -v_7.$$

By periodicity, one has also that $\hat{\xi}_0 = \hat{\xi}_8$: by means of Lemma 3.4.1 one can then conclude that the inner arc $z_I\left(\cdot; \gamma\left(\hat{\xi}_7\right), \gamma\left(\hat{\xi}_8\right); h\right)$ must be of collision-ejection. The concatenation $z\left(\cdot; \hat{\underline{\xi}}; h\right)$ bounces then between the collisional arcs described by $z_I\left(\cdot; \gamma\left(\hat{\xi}_3\right), \gamma\left(\hat{\xi}_4\right); h\right)$ and $z_I\left(\cdot; \gamma\left(\hat{\xi}_7\right), \gamma\left(\hat{\xi}_8\right); h\right)$ (see Figure 3.8).

Let us now analyze the actual structure of $\underline{\ell}$; recalling Eqs. (3.3.7) and (3.3.8), one has that $\hat{\xi}_0, \hat{\xi}_1 \in I^{(\ell_0)}$, $\hat{\xi}_2, \hat{\xi}_3 \in I^{(\ell_1)}$, $\hat{\xi}_4, \hat{\xi}_5 \in I^{(\ell_2)}$ and $\hat{\xi}_5, \hat{\xi}_6 \in I^{(\ell_3)}$. Hence, since $\hat{\xi}_3 = \hat{\xi}_4$ and $\hat{\xi}_5 = \hat{\xi}_1$, we can infer that $\ell_1 = \ell_2$ and $\ell_0 = \ell_3$, and then the finite sequence $\underline{\ell}$ is symmetric with respect to the axis which divides $\ell_1 = \ell_{4/2-1}$ from $\ell_2 = \ell_{4/2}$.

The reasoning used in the previous example can be generalised to find necessary conditions on $\underline{\ell}$ to have a collisional orbit: as claimed in Theorem 3.4.2 and underlined in its proof, proposed below, the presence of a collisional reflection arc always forces the
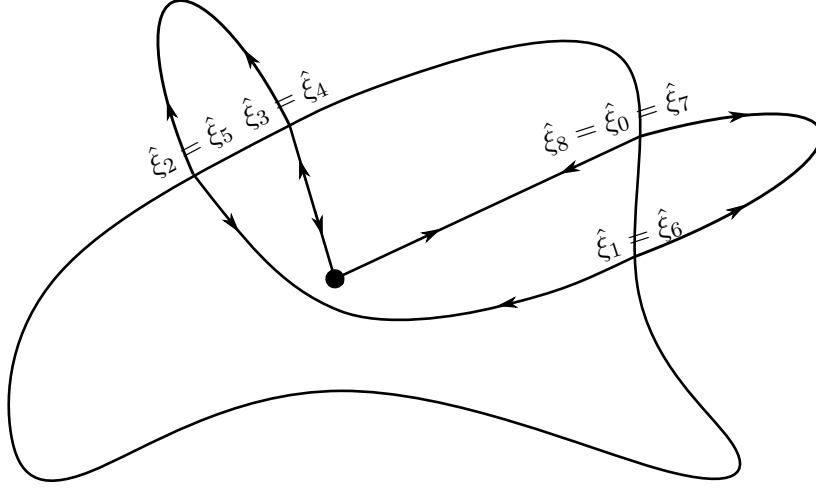
Fig. 3.8 Graphical representation of a possible trajectory with the structure described in Example 3.4.3 $\left(\text{to ease the notation, here } \hat{\xi}_j = \gamma\left(\hat{\xi}_j\right)\right)$. The presence of a collisional inner arc, along with the periodicity of the whole concatenation and the total number of transition points (determined by the length of $\underline{\ell}$), implies the existence of a second collisional arc.

complete trajectory to have another reflection, which can be either another collision-ejection arc or a homothetic outer arc depending, on the parity of $|\underline{\ell}|$.

Proof of Theorem 3.4.2. Let us start by assuming that n is even: in view of Eq. (3.4.1), we can suppose without loss of generality that the collisional arc is the one connecting $\gamma\left(\hat{\xi}_{n-1}\right)$ to $\gamma\left(\hat{\xi}_n\right)$, namely, $z_I\left(\cdot\,;\gamma\left(\hat{\xi}_{n-1}\right),\gamma\left(\hat{\xi}_n\right);h\right)$. Then, following the same reasonings used in Example 3.4.3, one has the following equalities:

$$\forall j = 0, \dots, n-1 \qquad \hat{\xi}_{n-(j+1)} = \hat{\xi}_{n+j}, \quad \tilde{v}_{n-(j+1)} = -\tilde{v}_{n+j}, \quad v_{n-(j+1)} = -v_{n+j},$$

$$\forall j = 0, \dots, n-2$$

$$\begin{cases} z_E\left(\cdot\,;\gamma\left(\hat{\xi}_{n-(j+2)}\right),\gamma\left(\hat{\xi}_{n-(j+1)}\right)\right) = z_E\left(T-\cdot\,;\gamma\left(\hat{\xi}_{n+j}\right),\gamma\left(\hat{\xi}_{n+j+1}\right)\right) & \text{if } j \text{ is even,} \\[2mm] z_I\left(\cdot\,;\gamma\left(\hat{\xi}_{n-(j+2)}\right),\gamma\left(\hat{\xi}_{n-(j+1)}\right);h\right) = z_I\left(T-\cdot\,;\gamma\left(\hat{\xi}_{n+j}\right),\gamma\left(\hat{\xi}_{n+j+1}\right);h\right) & \text{if } j \text{ is odd.} \end{cases}$$
$$(3.4.3)$$

In particular, taking $j = n-1$, one has $\hat{\xi}_0 = \hat{\xi}_{2n-1}$; by periodicity, $\hat{\xi}_{2n} = \hat{\xi}_{2n-1}$. Hence by Lemma 3.4.1, the inner arc $z_I\left(\cdot\,;\gamma\left(\hat{\xi}_{2n-1}\right),\gamma\left(\hat{\xi}_{2n}\right),h\right)$ must be collisional, and then the first claim is proved. Let us now focus on the structure of $\underline{\ell} = (\ell_0, \dots, \ell_{n-1})$. Since for every $k = 0, \dots, n-1$ one has $\hat{\xi}_{2k}, \hat{\xi}_{2k+1} \in I^{(\ell_k)}$, equalities in (3.4.3) imply that for every $j = 0, \dots, n/2 - 1$ it holds $\ell_{n/2-(j+1)} = \ell_{n/2+j}$, and then the sequence $(\ell_0, \dots, \ell_{n-1})$ is symmetric with respect to the axis which separates $\ell_{n/2-1}$ from $\ell_{n/2}$. Let us now suppose that $n > 1$ is odd and, without loss of generality, assume that the

arc $z_I\left(\cdot;\gamma\left(\hat{\xi}_n\right),\gamma\left(\hat{\xi}_{n+1}\right),h\right)$ is collisional. One can infer that

$$\forall j=0,\ldots,n-1 \qquad \hat{\xi}_{n-j}=\hat{\xi}_{n+j+1},\quad \tilde{v}_{n-j}=\tilde{v}_{n+j+1},\quad v_{n-j}=v_{n+j+1},$$
$$\forall j=0,\ldots,n-2$$
$$\begin{cases} z_E\left(\cdot;\gamma\left(\hat{\xi}_{n-(j+1)}\right),\gamma\left(\hat{\xi}_{n-j}\right)\right)=z_E\left(T-\cdot;\gamma\left(\hat{\xi}_{n+j+1}\right),\gamma\left(\hat{\xi}_{n+j+2}\right)\right) & \text{if } j \text{ is even} \\ z_I\left(\cdot;\gamma\left(\hat{\xi}_{n-(j+1)}\right),\gamma\left(\hat{\xi}_{n-j}\right);h\right)=z_I\left(T-\cdot;\gamma\left(\hat{\xi}_{n+j+1}\right),\gamma\left(\hat{\xi}_{n+j+2}\right);h\right) & \text{if } j \text{ is odd} \end{cases}$$
$$(3.4.4)$$

Again, taking $j=n-1$ and considering the periodicity of $\hat{\underline{\xi}}$, one has $\hat{\xi}_1=\hat{\xi}_{2n}=\hat{\xi}_0$, and this implies that the outer arc $z_E\left(\cdot;\gamma\left(\hat{\xi}_0\right),\gamma\left(\hat{\xi}_1\right)\right)$ is homothetic. Eqs. (3.4.4) imply also that for every $j=0,\ldots,(n+1)/2-2$ one has $\ell_{(n+1)/2-(j+1)}=\ell_{(n+1)/2+j}$, and then the finite sequence $(\ell_0,\ldots,\ell_{n-1},\ell_0)$ is symmetric with respect to the axis separating $\ell_{(n+1)/2-1}$ from $\ell_{(n+1)/2}$. $\qquad\square$

Remark 3.4.4. *The proof of Theorem 3.4.2 is particularly rich of further informations:*

1. *in both the described cases one can not have more than two radial (inner or outer) arcs; in particular, the concatenation segment between two radial arcs must contain $n-1$ non-radial intermediate arcs. This is necessary to have $|\ell|=n$;*

2. *with completely analogous reasonings, one can prove that, if $z\left(\cdot,\hat{\underline{\xi}},h\right)$ admits an outer homothetic arc, then*

   - *if $|\underline{\ell}|$ is even, there must be another outer homothetic arc;*
   - *if $|\underline{\ell}|$ is odd, there must be a collision-ejection inner arc.*

   In both cases, the sequence $\underline{\ell}$ must satisfy symmetry properties analogous to the ones described in Theorem 3.4.2;

3. *as a consequence, a periodic trajectory can have either zero or two radial arcs, between which it is reflected.*

Remark 3.4.5. *In the particular case $n=1$, one has that the concatenation $z\left(\cdot,\hat{\underline{\xi}},h\right)$ is collisional if an only if it is a homothetic orbit for the complete dynamics. In particular, this is possible if and only if, being $\underline{\ell}=(j)$, with $j\in\mathcal{I}$, one has $\hat{\xi}_0=\hat{\xi}_1=\hat{\xi}_2=\bar{\xi}_j$.*

Theorem 3.4.2 provides necessary conditions on a finite word $\underline{\ell}$ for the corresponding dynamics to be collisional. By periodicity, we can extend this result to periodic bi-infinite words in $\mathbb{L}_P$, and, actually, to any admissible word in $\mathcal{L}$. Indeed, if a concatenation z satisfying $\underline{\ell}\in\mathcal{L}$ has a collision, then using the same reasoning as in Theorem 3.4.2, we deduce the *symmetry* of the word $\underline{\ell}$. In particular, if $\underline{\ell}\in\mathbb{L}_P$ we are again in the case described in Remark 3.4.4.

On the other hand, the same results can be used in a *reversed* formulation to have *sufficient* conditions for the concatenation induced by $\underline{\ell}$ to be non-collisional. In particular, restricting the set of the admissible words to the ones which do not present the symmetry highlighted in Theorem 3.4.2, one can construct a *non-collisional symbolic dynamics* for our refractive billiard.
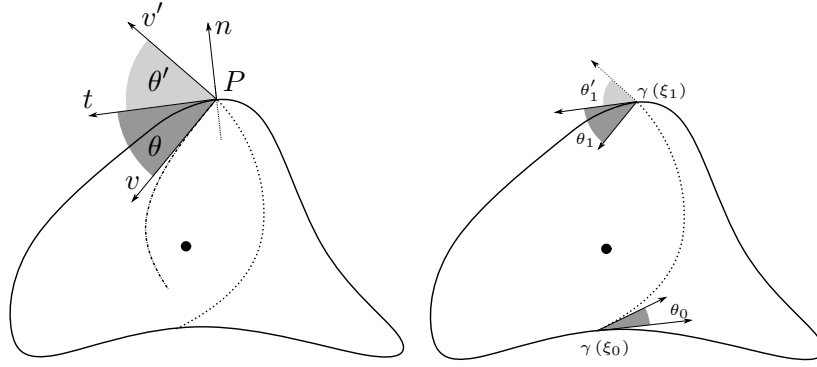
Fig. 3.9 Left: reflection law for the Keplerian billiard. Right: coordinates of the first return map $f : (\xi_0, \theta_0) \mapsto (\xi_1, \theta_1)$.

Corollary 3.4.6. *Define the set $\tilde{\mathcal{L}} \subsetneq \mathcal{L}$ as*

$$\tilde{\mathcal{L}} = \{\underline{\ell} \in \mathcal{L} \,|\, \underline{\ell} \text{ is not symmetric}\},$$

then the symbolic dynamics semi-conjugated to $\tilde{\mathcal{L}}$ is not collisional, in the sense any trajectory corresponding to a word $\underline{\ell} \in \tilde{\mathcal{L}}$ does not have any collisional inner arc.

3.5   Dynamical consequences: analytic non-integrability

This last section recollects results obtained along the whole study of the galactic refraction (and reflection) billiard, through Chapters 1, 2 and 3, to investigate the issue of the chaoticity and integrability of such model. In Section 1.7 we already observed that a centred ellipse could produce, given a certain choice of the physical parameters, diffusive chaotic orbits. More generally, the presence of a symbolic dynamics proved in Section 3.3 is a strong hint of complex behaviors. It is then reasonable to ask whether, as far as h is large enough, our system is non-integrable.

Let us start with some considerations related to the presence of infinitely many heteroclinic connections between different homothetic collision-ejection trajectories (on this topic see for example [82–84]). We will present results which hold both in the refractive and reflective case, assuming some further assumptions on the domain D, according to the following definition.

Definition 3.5.1. *An admissible domain D (according to Definition 3.2.4) is termed* non degenerate *if every central configuration $\bar{\xi}_i$, $i \in \mathcal{I}$, is* non degenerate *in the sense that*

$$\frac{d^2}{d\xi^2}\|\gamma(\xi)\|_{|\xi = \bar{\xi}_i} \neq 0.$$

Proposition 3.5.2. *Suppose D is admissible and that there exists a non-degenerate central configuration $\bar{\xi}_i$, for some $i \in \mathcal{I}$. Then if h is large enough, the homothetic trajectory in the direction of $\gamma\left(\bar{\xi}_i\right)$ is a hyperbolic saddle equilibrium.*

Proof. The proof of this result in the case of the refractive dynamics relies on asymptotic estimates based on 1.5.1. Here, in particular, a general domain is considered, and,

after the construction of a suitable first return map f, the stability of the homothetic trajectories is investigated in relation to the local geometric features of the boundary and the value of the physical parameters $\mathcal{E}, h, \omega, \mu$. The inspection is carried on by considering the sign of the discriminant Δ of the characteristic polynomial associated to the Jacobian matrix of f centered in an homothetic direction. By straightforward estimates, one can prove that, as long as the considered homothetic direction is nondegenerate, $\lim_{h \to \infty} \Delta > 0$, and then $\bar{\xi}$ is a saddle point.

Let us now pass to the reflective Kepler billiard. We refer again to the procedure used in [1] for the computation of Δ, presenting the explicit computation adapted to the reflective case. Let us start by giving the analytic formulation of the reflection law: consider an inner Keplerian arc that intersects the boundary in the point P and with velocity vector v' (see Figure 3.9, left). Decomposing v' with respect to the basis (t, n) given respectively by the tangent and the normal unit vectors to ∂D in P, one has that the reflected velocity v has the same tangent component of v', while the normal component is equal except for the opposite sign. Equivalently, if θ and θ' are the angles respectively of v and v' with t, one has $\theta = -\theta'$ and, as a consequence,

$$\cos \theta = \cos \theta'. \tag{3.5.1}$$

This relation will be used, in the computation of Δ, as replacement of the Snell's law. Let us now take $(P, v) \in \partial D \times B_r$ with $r = \sqrt{2}\sqrt{\mathcal{E} + h + \mu/\|P\|}$ as the initial condition of an inward-pointing Keplerian hyperbola: it is uniquely determined by the pair (ξ, θ), where $\gamma(\xi) = p$ and θ is defined as before. We want to construct (at least implicitly) the first return map that, given a pair (ξ_0, θ_0), returns the parameters (ξ_1, θ_1) corresponding to the position and velocity vector obtained after going through an inner arc and the subsequent reflection (see Figure 3.9, right). By virtue of Lemma 1.2.6 and Eq. (3.5.1), we can infer that the quantities $(\xi_0, \theta_0, \xi_1, \theta_1)$ have to satisfy the relation

$$\Phi(\xi_0, \theta_0, \xi_1, \theta_1) = \begin{pmatrix} \Phi_1(\xi_0, \theta_0, \xi_1, \theta_1) \\ \Phi_2(\xi_0, \theta_0, \xi_1, \theta_1) \end{pmatrix} = \begin{pmatrix} \partial_{\xi_0} S_I(\xi_0, \xi_1) + \sqrt{V_I(\gamma(\xi_0))} \cos \theta_0 \\ \partial_{\xi_1} S_I(\xi_0, \xi_1) - \sqrt{V_I(\gamma(\xi_1))} \cos \theta_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Let us now consider a central configuration $\bar{\xi}$: it is straightforward that the coordinates $(\bar{\xi}, \bar{\theta}) = (\bar{\xi}, \pi/2)$ correspond to the collision-ejection homothetic solution in the direction of $\gamma(\bar{\xi})$, which is reflected exactly in itself. This means that, denoted with $\underline{q} = (\bar{\xi}, \bar{\theta}, \bar{\xi}, \bar{\theta})$ the initial and final parameters related to an homothetic trajectory, one has that $\Phi(\underline{q}) = \underline{0}$. Let us now define the matrix

$$D_{(\xi_1, \theta_1)}(\underline{q}) = \begin{pmatrix} \dfrac{\partial \Phi_1}{\partial \xi_1}(\underline{q}) & \dfrac{\partial \Phi_1}{\partial \theta_1}(\underline{q}) \\ \dfrac{\partial \Phi_2}{\partial \xi_1}(\underline{q}) & \dfrac{\partial \Phi_2}{\partial \theta_1}(\underline{q}) \end{pmatrix} :$$

if $\det\left(D_{(\xi_1, \theta_1)}(\underline{q})\right) \neq 0$, then by the implicit function theorem $f \colon (\xi_0, \theta_0) \mapsto (\xi_1, \theta_1)$ is well defined in a neighbourhood of $(\bar{\xi}, \bar{\theta})$, and

$$Df(\bar{\xi}, \bar{\theta}) = -\left(D_{(\xi_1, \theta_1)}(\underline{q})\right)^{-1} D_{(\xi_0, \theta_0)}(\underline{q}).$$

From straightforward computations, one has that

$$\det\left(D_{(\xi_1,\theta_1)}\left(\underline{q}\right)\right) = \sqrt{V_I\left(\gamma\left(\bar{\xi}\right)\right)}\partial_{\xi_0\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right) = \frac{\mu}{4\|\gamma\left(\bar{\xi}\right)\|^2},$$

where the last equation is obtained by taking into consideration Eqs. (1.5.20). The implicit function theorem can then be applied to obtain

$$Df(\bar{\xi},\bar{\theta}) = \begin{pmatrix} -\dfrac{\partial^2_{\xi_0}S_I\left(\bar{\xi},\bar{\xi}\right)}{\partial_{\xi_0\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)} & \dfrac{\sqrt{V_I\left(\gamma\left(\bar{\xi}\right)\right)}}{\partial_{\xi_0\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)} \\ \dfrac{1}{\sqrt{V_I\left(\gamma\left(\bar{\xi}\right)\right)}}\left(\dfrac{\partial^2_{\xi_0}S_I\left(\bar{\xi},\bar{\xi}\right)\partial^2_{\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)}{\partial_{\xi_0\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)}\right) & -\dfrac{\partial^2_{\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)}{\partial_{\xi_0\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)} \end{pmatrix}$$

and then Δ, which is the discriminant of the characteristic polynomial associated to the above matrix, is given by

$$\Delta = \left(\frac{\partial^2_{\xi_0}S_I\left(\bar{\xi},\bar{\xi}\right)\partial^2_{\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)}{\partial_{\xi_0\xi_1}S_I\left(\bar{\xi},\bar{\xi}\right)}\right)^2 - 4$$

$$= \frac{32\|\gamma\left(\bar{\xi}\right)\|^4}{\mu^2}\left(V_I\left(\gamma\left(\bar{\xi}\right)\right)\right)^{3/2}\left(k\left(\bar{\xi}\right) - \frac{1}{\|\gamma\left(\bar{\xi}\right)\|}\right) \times$$

$$\times \left[\sqrt{V_I\left(\gamma\left(\bar{\xi}\right)\right)}\left(k\left(\bar{\xi}\right) - \frac{1}{\|\gamma\left(\bar{\xi}\right)\|}\right) + \frac{\mu}{2\|\gamma\left(\bar{\xi}\right)\|^2\sqrt{V_I\left(\gamma\left(\bar{\xi}\right)\right)}}\right]$$

where we used again Eqs. (1.5.20). The final result is then trivially obtained by observing that, if $\bar{\xi}$ is a non degenerate critical point for $\|\gamma\left(\cdot\right)\|$, then $\Delta > 0$ for h large enough. $\qquad\square$

For the readers' convenience, as these sets will be widely used in the following, let us recall that the construction of the set of initial conditions X is done in Section 3.3.3, and that the set $\mathcal{U}(\delta)$ is introduced in Definition 3.2.5.

Let us now present two lemmas, which will be the basis to prove the existence of heteroclinic connections between nondegenerate saddles. They are based on a construction which is analogous to the one presented in Section 3.3.2, except for the fact that, instead of periodic trajectories, we consider fixed-ends ones. For this reason, for any $n \geq 2$ consider the sets

$$\mathbb{L}'_n = \{\underline{\ell} = (\ell_0,\ldots,\ell_{n-1}) : \ell_i \in \mathcal{I} \text{ and } \forall i = 0,\ldots,n-2 \quad \ell_{i+1} \in NA\left(\ell_i\right)\},$$

$$\mathbb{U}'_{\underline{\ell}} = (I_0 \times I_0) \times (I_1 \times I_1) \times \cdots \times (I_{n-2} \times I_{n-2}) \times I_{n-1},$$

$$\mathbb{S}'_{\underline{\ell}} = \left\{\underline{\xi} = (\xi_0,\ldots,\xi_{2n-2}) \in \overline{\mathbb{U}}_{\underline{\ell}} : \xi_0 = \tilde{\xi}_a, \xi_{2n-2} = \tilde{\xi}_b\right\},$$

where ξ_a ad ξ_b are fixed respectively in I_0 and I_{n-1}. It is easy to prove that

$$\mathbb{L} = \bigcup_{n \in \mathbb{N} \setminus \{0\}} \mathbb{L}'_n,$$

with $\mathbb{L}$ defined as in Eq. (3.3.6).

Let us remark that, in this setting, we have $2n - 3$ free points $\xi_1, \ldots, \xi_{2n-3}$. As a consequence, given a sequence $\underline{\xi} \in \mathbb{S}'_{\underline{\ell}}$, the total Jacobi length takes the form

$$W'_{\underline{\ell}}(\underline{\xi}; h) = \sum_{j=0}^{n-2} S_E(\xi_{2j}, \xi_{2j+1}) + \sum_{j=0}^{n-2} S_I(\xi_{2j+1}, \xi_{2j+2}; h). \tag{3.5.2}$$

In this framework, Definition 3.3.9 and Eq. (3.3.11) are the same, with straightforward modifications in the indices. Again, sequences in $S'_{\underline{\ell}}$ such that $\nabla W'_{\underline{\ell}}(\underline{\xi}) = 0$ correspond to admissible trajectories for our fixed-ends dynamics which realize the word $\underline{\ell}$.

Lemma 3.5.3. *Let us fix $\underline{\ell} \in \mathbb{L}$, and set $n \doteq |\underline{\ell}|$. Fixed $\xi_a \in I_0$ and $\xi_b \in I_{n-1}$, there exists a trajectory which connects $\gamma(\xi_a)$ to $\gamma(\xi_b)$ realizing the word $\underline{\ell}$.*

Proof. The proof of this result is completely analogous of the one of Theorem 3.3.13, setting $d \doteq 2n - 3$,

$$R \doteq I_0 \times (I_1 \times I_1) \times \ldots \times (I_{n-2} \times I_{n-2}),$$
$$F_i(\underline{\xi}) \doteq \frac{\partial W'_{\underline{\ell}}}{\partial \xi_i}(\underline{\xi})$$

for every $i = 1, \ldots, 2n - 3$. $\qquad\qquad\square$

Following exactly the techniques used in Proposition 3.3.20 we can pass from fixed-ends finite trajectories to one-side infinite ones starting from a chosen point. Before doing that, it is worth to define the set of the admissible *infinite* words

$$\mathcal{L}^+ \doteq \left\{ (\ell_0, \ldots, \ell_n, \ldots) \in \mathcal{I}^{\mathbb{N}} : \ell_i + 1 \in NA(\ell_i) \ \forall i \in \mathbb{N} \right\}.$$

Lemma 3.5.4. *Let $\underline{\ell} \in \mathcal{L}^+$ and $\xi \in I^{(\ell_0)}$. The there exists $v \in \mathbb{R}^2$ such that $(\xi, v) \in X$ and $\pi(\xi, v) = \underline{\ell}$[7].*

The next result is a straightforward consequence of Lemma 3.5.4 and Proposition 3.5.2.

Corollary 3.5.5. *Suppose D is admissible and that there exists a non-degenerate central configuration $\bar{\xi}_i$, for some $i \in \mathcal{I}$. Then, if h is large enough, for every $\xi \in \mathcal{U}(\delta)$ there exist infinitely many half-heteroclinic connections tending forward (resp. backward) to the homothetic trajectory in the direction of $\gamma(\bar{\xi}_i)$.*

As far as more than one central configuration is non degenerate, Lemma 3.5.4 allows to construct heteroclinic connections between different saddle points.

---

[7]Here, the projection is intended only for the forward trajectory starting from $(\gamma(\xi), v)$.

Corollary 3.5.6. *If D is admissible and the energy jump h is large enough, then for every pair of non degenerate central configurations $\bar{\bar{\xi}}_i$, $\bar{\bar{\xi}}_j$, $i, j \in \mathcal{I}$, $i \neq j$, there exist infinitely many heteroclinic connections between the homothetic trajectories in the direction of $\gamma\left(\bar{\xi}_i\right)$ and $\gamma\left(\bar{\xi}_j\right)$ for both the refractive and the reflective dynamics.*

Proof. The proof is identical for both the cases: let us suppose that h is large enough such that both Theorem 3.1.5 and Proposition 3.5.2 hold. Now, taken $i \in \mathcal{I}$ and $j \in NA(i)$, call $\bar{z}_i$ and $\bar{z}_j$ the corresponding homothetic trajectories and consider the bi-infinite (non-periodic) word $\underline{\ell} = (\dots, i, i, i, [\dots], j, j, j, \dots)$, where $[\dots]$ denotes any word of finite length such that $\underline{\ell} \in \mathcal{L}$. By means of Theorem 3.3.13, find the sequence $(\hat{\xi}_k)_{k \in \mathbb{Z}}$ which realizes $\underline{\ell}$. By the hyperbolicity of both $\bar{z}_i$ and $\bar{z}_j$ as equilibrium trajectories, this sequence must belong to the unstable manifold of $\bar{z}_i$ as well as to the stable manifold of $\bar{z}_j$. $\qquad\square$

The last result of this section connects the presence of infinitely many half-heteroclinics to the analytic non-integrability of our dynamical systems. This result is obtained by adapting a classical argument by Kozlov ([34]) and makes use of Lemma 3.5.4 and Corollary 3.5.5.

Theorem 3.5.7. *Suppose D is an admissible domain and assume there exists a non-degenerate central configuration $\bar{\bar{\xi}}_i$, for some $i \in \mathcal{I}$. Then, if h is large enough, there are no non-constant analytic first integrals associated to the dynamics.*

Proof. On every initial condition $(\xi, v) \in X$, the first return map, its inverse and all their iterates are well defined. In particular, let us observe that, fixed $\xi \in [0, L]$, every outward-pointing velocity vector v starting from $\gamma(\xi)$ is uniquely determined by the angle $\alpha \in (-\pi/2, \pi/2)$ between v and the normal unit vector to γ at $\gamma(\xi)$. Using α as a new variable, the initial conditions corresponding to the homothetic arcs will be denoted with $(\bar{\hat{\xi}}_j, 0)$, $j \in \mathcal{I}$. By construction, and with a slight abuse of notation, there exists $\delta > 0$ and $\alpha_0 > 0$ such that

$$X \subset \mathcal{U}(\delta) \times (-\alpha_0, \alpha_0).$$

Let now $G : \mathcal{O} \to \mathbb{R}$, with $\mathcal{O}$ an open set containing $\mathcal{U}(\delta) \times (-\alpha_0, \alpha_0)$, be an analytic first integral and let $c \in \mathbb{R}$ be such that $G(\bar{\xi}_i, 0) = c$. If h is large enough, by Proposition 3.5.2, the stable and unstable manifolds of $(\bar{\xi}_i, 0)$ are contained in the same level set $\{G = c\}$.
Fix now $\hat{\xi} \in \mathcal{U}(\delta)$: Lemma 3.5.4 ensures that there exist infinite pairs $(\hat{\xi}, \alpha) \in \{\hat{\xi}\} \times (-\alpha_0, \alpha_0)$ that belong to the stable manifold of $(\bar{\xi}_i, 0)$. This means that the c-level of the analytic function $G(\hat{\xi}, \cdot) : (-\alpha_0, \alpha_0) \to \mathbb{R}$ admits an accumulation point. Hence, this function is constant. We conclude the proof by the arbitrarity of $\hat{\xi}$. $\qquad\square$

We stress that, with the obvious modifications on the quantity involved, the above results hold also in the reflective case.

If the domain D is nondegenerate as in Definition 3.5.1, making use of the estimates in [76], one can prove that the projection π that define the symbolic dynamics determined in Theorem 3.1.5 is actually injective, so that the dynamical system defined by $F_{|X}$ is *topologically chaotic*. The details of this proof, based on the uniqueness of the critical point provided by Poincaré-Miranda Theorem, will be described in [3].

Part II

Normal forms and Nekhoroshev theory for geocentric satellites

Chapter 4

Stability estimates for Earth's satellites through normal forms

4.1 Introduction

One of the major goals of Celestial Mechanics is the analysis of the stability of the dynamics of celestial bodies. Knowing the behavior in time of the orbital elements of an object allows one to predict its future, in particular whether it will cross the orbit of other celestial bodies and eventually undergo collisions. When applied to artificial spacecraft and space debris orbiting around the Earth, the question of the stability becomes of crucial importance, especially in view of the problem of estimating the orbital survival times of operating satellites or space debris. It is therefore crucial to devise methods that allow to study the orbital stability of objects moving around our planet.

In particular, we will not consider the complex dynamics of an artificial spacecraft, which should include the analysis of its shape, composition as well as its rotational motion. We will rather consider a point-mass body around the Earth, that we can identify with one of the several millions of space debris orbiting our planet. In fact, the proliferation of artificial satellites in the last decades has led to the generation of an enormous amount of space debris with different sizes, from meters down to microns, and at different altitudes. Space debris are remnants of non operational satellites or the result of break-up events, either collisions or explosions. Since the altitude determines the contribution of the different forces acting on the object (the gravitational attraction of the Earth, its geopotential perturbation, the influence of Sun and Moon, the Solar radiation pressure, etc.), it is convenient to make a distinction in terms of the altitude. To this end, the space in the surrounding of the Earth is commonly split into three main regions: LEO ('Low Earth Orbit') denotes the region up to about 2 000 km of altitude in which the Earth's attraction, the geopotential as well as the atmospheric drag are the terms which greatly affect the dynamics of an Earth's satellite; MEO ('Medium

Earth Orbit') refers to the region between $2\,000$ and $30\,000$ km in which the effects of Moon, Sun and Solar radiation pressure become important; GEO ('Geosynchronous Earth Orbit') refers to a thin (~ 200 km) zone around the geostationary orbits (at $42\,164$ km from Earth's center), where the satellites are in synchronous resonance with the 24-hour rotation of the Earth around its spin-axis.

The huge amount of objects (up to millions) in LEO, MEO, GEO needs a careful control of their orbits and the analysis of their dynamical stability ([39, 42, 41, 95, 43, 96]), also in view of devising appropriate mitigation measures (see, e.g., [35–37]). For objects in LEO, it is of crucial importance to evaluate the orbital lifetime, which is strongly affected by the atmospheric drag which provokes a decay of the orbits ([97, 44, 98–100]). On the other hand, the study of dynamics at MEO in the conservative regime has been subject of many works, including the development of analytical models (e.g., [101, 42, 102–104]), study of resonances (e.g., [105, 106, 95, 107–113]), as well as the dynamical chartography (stability maps, onset of chaos) of the MEO region (e.g., [114, 115, 46, 47, 116–121]).

In this work we focus on objects in MEO, GEO and beyond, thus not taking into account the dissipative effect of the atmosphere. Instead of using a propagation of the orbits to predict the stability time of the orbital elements, we propose a procedure based on analytical perturbative methods (see also [50]). More precisely, via a suitably defined sequence of canonical transformations, we construct a *normal form* of the Hamiltonian function, which enjoys the property that one or more of the Hamiltonian's Delaunay actions define *quasi-integrals of motion* (namely, integrals of the integrable part of the new Hamiltonian). Once the transformed Hamiltonian is obtained, the size of its *remainder* (which gives a control on the goodness of the approximation) can then be used to provide bounds on the time variations, and hence the stability time of the orbital elements (semimajor axis, eccentricity, inclination) as a function of the distance of the object from the Earth. We refer to this procedure as *semi-analytical*, which means that it uses an analytical method, precisely normal forms, whose coefficients are calculated numerically, namely with the aid of a computer.

We consider two different models to describe the motion of the debris around the Earth. The first model takes into account only the influence of the geopotential up to the term J_2 of its expansion in spherical harmonics (see [45]); we refer to this problem as the J_2 model and denote the corresponding Hamiltonian as $\mathcal{H}_{J_2}$, which results from truncating to a suitable power of the coordinates around reference values, and normalizing up to a suitable order, as described in Section 4.4.1. The second model, referred to as the secular 'geolunisolar' model (Hamiltonian $\mathcal{H}_{gls}$, truncated and normalized similarly to $\mathcal{H}_{J_2}$, see Sections 4.2.2 and 4.5.1), includes also the effects of the Moon and the Sun, placing, for simplicity, the Moon strictly on the ecliptic; this last restriction means to omit from the Hamiltonian terms corresponding to lunisolar resonances other

than the 'inclination-dependent' ones. The latter resonances, on the other hand, are those producing the most important effects as regards orbital stability (see [95, 42] for a review). Furthermore, instead of formally eliminating the fast angle via canonical transformations (as we do in the pure J_2 problem), in $\mathcal{H}_{gls}$ we just take the average of the Hamiltonian with respect to all fast angles, namely, the mean anomaly of the satellite as well as the fast angles of the Moon and Sun: this averaged model allows us to focus on the satellite's long-term dynamics (i.e. the *secular* one). The averaging is done in closed-form (see [122]) and leads to formulas equivalent to those described in [45]. Furthermore, we reduce this last Hamiltonian to action-angle variables around each forced equilibrium point, which corresponds to a non-zero inclination defining the so-called *Laplace plane* (see Section 4.2.2).

In summary, our stability estimates are obtained according to the procedure (i)-(iii) outlined below:

i) Within the J_2 model, we make a formal elimination in the Hamiltonian of the fast angle (mean longitude); as a consequence, we get the preservation of the conjugate action variable corresponding to the semimajor axis. This allows us to compute the stability time for the semimajor axis at different altitudes, yielding stability times that increase with the altitude.

ii) Using $\mathcal{H}_{gls}$, instead, the semimajor axis becomes a parameter (with *a priori* constant value), while we proceed to analyze the behaviour of eccentricity and inclination. The latter is obtained using a quasi-resonant normal form, which reflects the 1:1 near-resonance of the integrable part of $\mathcal{H}_{gls}$ between the frequencies of the longitude of the ascending node and the sum of the argument of perigee and the longitude of the ascending node (see Section 4.2.2). This means that, close to the Laplace plane, the quasi-preserved secular quantities cannot be defined neither as the eccentricity e nor the inclination i alone, but rather by the Kozai-Lidov combination $\mathcal{I} = 1 - \sqrt{1 - e^2} \cos i \approx (e^2 + i^2)/2$ (for e, i small). We then explore the dependence of the stability time of $\mathcal{I}$ on the altitude of the orbit. Our results show that the J_2 and lunisolar terms have an opposite effect on the time of stability as the distance from the Earth increases. As a by-product of this analysis, we also compute the so-called forced inclination (that is, the inclination of the Laplace plane), which corresponds to the shift of the secular equilibrium from a strictly equatorial orbit to an orbit with small positive initial inclination, an effect caused by the fact that the perturbing bodies (Moon and Sun) are in orbits inclined with respect to the Earth's equator.

iii) Finally, as a first step towards obtaining exponential stability estimates à la Nekhoroshev ([53]), we check whether some so-called 'steepness' conditions are satisfied for the integrable part of both Hamiltonians $\mathcal{H}_{J_2}$ and $\mathcal{H}_{gls}$, namely whether the integrable parts are convex, quasi-convex, or satisfy the three-jet condition (see [55] and references therein). The results show that the J_2 model is three-jet non-convex, while the contribution of the lunisolar terms removes the intrinsic degeneracy of the J_2 part and allows us to conclude that the geolunisolar model is quasi-convex. A detailed application of the non-resonant form of Nekhoroshev's theorem in the Hamiltonian $\mathcal{H}_{gls}$ is the subject of Chapter 5.

Summarizing, the previous strategy allows us to obtain three different stability results: one for the semimajor axis in the J_2 model, a second for the stability of the eccentricity and inclination in the geolunisolar model, and a third on the holding,

altogether, of necessary conditions for Nekhoroshev-type stability of the satellite orbits. All three results point towards the same direction, i.e. that, at least far from exact resonances, orbital stability can be ensured at MEO, GEO and beyond for quite long times ($10^4 - 10^6$ orbital periods, $10^2 - 10^4$ years). Besides these general numbers, one may remark that the calculation of the size of the remainder of the normal form actually provides an estimate of the rate of drift of the orbits in element space, an information required in orbital diffusion studies for defunct satellites and space debris. The figures of the current chapter are taken from [4].

4.2   The J_2 and geolunisolar models

Bodies orbiting around the Earth are primarily affected by the Keplerian attraction with our planet. However, for an accurate description of the dynamics it is mandatory to assume that the Earth is non-spherical. Beside the Earth, the satellite dynamics is subject to the gravitational influence of Sun and Moon. Section 4.2.1 describes the Hamiltonian model $\mathcal{H}_{J_2 kep}$, which includes the Earth's Keplerian term and the first non-trivial term in the expansion of the geopotential. Section 4.2.2 presents the Hamiltonian model $\mathcal{H}_{gls,sec}$, which includes J_2 and lunisolar terms, averaged over the fast angles.

4.2.1   The J_2 model

We consider a model describing the motion of a point-mass body, say a satellite S, under the effect of the Earth's gravitational attraction, including an approximation of the geopotential due to the oblateness of the Earth. Let $\underline{r} \equiv (x, y, z)$ be the position vector of S in a geocentric reference frame, with the plane (x, y) coinciding with the equatorial plane, and x pointing towards a fixed celestial point (e.g. the equinox). We consider the Hamiltonian describing the motion of S under the geopotential as the sum of two terms

$$\mathcal{H}_{J_2 kep} = \mathcal{H}_{kep} + \mathcal{V}_{J_2} \ , \tag{4.2.1}$$

where

$$\mathcal{H}_{kep} = \frac{p^2}{2} - \frac{\mu_E}{r} \tag{4.2.2}$$

is the Keplerian part ($r = \|\underline{r}\|$), and

$$\mathcal{V}_{J_2} = -J_2 \frac{\mu_E R_E^2}{r^3} \left(\frac{1}{2} - \frac{3z^2}{2r^2} \right) \ , \tag{4.2.3}$$

is the J_2 potential term, arising from expanding the geopotential in spherical harmonics and retaining only the largest coefficient (see, e.g., [45]). The constants are the Earth's mass parameter $\mu_E = \mathcal{G} M_E$ ($\mathcal{G}$ = Newton's constant, M_E = Earth's mass), R_E is the mean Earth's radius, and J_2 is a dimensionless coefficient describing the oblate shape of the Earth. Their numerical values are:

- $\mu_E = 1.52984 \times 10^9 \ R_E^3 / yr^2$;

- $R_E = 6378.14$ km;

- $J_2 = -1082.6261 \times 10^{-6}$.

The Hamiltonian (4.2.1) is expressed in Cartesian coordinates. However, by a standard procedure, it can be transformed to an expression in the following set of modified Delaunay canonical action-angle variables

$$
\begin{cases}
L = \sqrt{\mu_E a} \\
P = \sqrt{\mu_E a}(1 - \sqrt{1 - e^2}) \\
Q = \sqrt{\mu_E a}\sqrt{1 - e^2}(1 - \cos i)
\end{cases}
\qquad
\begin{cases}
\lambda = M + \omega + \Omega \\
p = -\omega - \Omega \\
q = -\Omega \ ,
\end{cases}
\tag{4.2.4}
$$

where $(a, e, i, M, \omega, \Omega)$ are the orbital elements of the satellite (semimajor axis, eccentricity, inclination, mean anomaly, argument of the perigee, longitude of the nodes). The passage is done by first expressing the Hamiltonian (4.2.1) in elements via the relations (see e.g. [123])

$$
\begin{aligned}
x &= \frac{1}{2}r(1 + \cos i)\cos(f + \omega + \Omega) + \frac{1}{2}r(1 - \cos i)\cos(f + \omega - \Omega) \\
y &= \frac{1}{2}r(1 + \cos i)\sin(f + \omega + \Omega) - \frac{1}{2}r(1 - \cos i)\sin(f + \omega - \Omega) \\
z &= r\sin i \sin(f + \omega),
\end{aligned}
\tag{4.2.5}
$$

where f is the true anomaly and r, $\cos f$ and $\sin f$ are given by the series

$$
r = a\left[1 + \frac{e^2}{2} - 2e\sum_{\nu=1}^{\infty}\frac{(J_{\nu-1}(\nu e) - J_{\nu+1}(\nu e))\cos(\nu M)}{2\nu}\right]
\tag{4.2.6}
$$

$$
\cos f = \frac{2(1 - e^2)}{e}\sum_{\nu=1}^{\infty}J_\nu(\nu e)\cos(\nu M) - e
$$

$$
\sin f = 2\sqrt{1 - e^2}\sum_{\nu=1}^{\infty}\frac{1}{2}(J_{\nu-1}(\nu e) - J_{\nu+1}(\nu e))\sin(\nu M) \quad .
$$

In the actual calculations, all series are truncated to order $N = 15$ in the eccentricity e. Finally, we pass from the elements $(a, e, i, M, \omega, \Omega)$ to the canonical variables (L, P, Q, λ, p, q) by inverting Eqs. (4.2.4).

To perform the high order normal form computations described in Section 4.4, using computer algebra, it is convenient that the dependence of the Hamiltonian on the action-angle variables be expressed as a trigonometric polynomial. To this end, we first make a shift transformation $L \to \delta L$ around a reference value a_*, with

$$
\delta L = L - L_* = \sqrt{\mu_E a} - \sqrt{\mu_E a_*} \ .
\tag{4.2.7}
$$

This means that the Hamiltonian found after expanding in powers of the quantity δL refers to the local dynamics of orbits with semimajor axis $a \approx a_*$. Every time when we change the reference value a_* (i.e. the 'altitude' or 'distance' of the orbit from the Earth's center), we then perform the Hamiltonian expansion anew around L_* and obtain the stability estimates corresponding to that reference value. One may also note that $P = \mathcal{O}(e^2/2)$ and $Q = \mathcal{O}(i^2/2)$, thus all three quantities δL, P and Q are small quantities for orbits not very far from the equator and not very far from circular. We

then expand $\mathcal{H}_{J_2 kep}(\delta L, P, Q, \lambda, p, q)$ in powers of $\sqrt{\delta L}$, $\sqrt{P}$, and $\sqrt{Q}$ up to the same order $N = 15$ as the original expansion in the eccentricity (this ensures missing no term in P, Q in the Hamiltonian up to the order N). After this change, the truncated Hamiltonian takes the form (apart from a constant):

$$
\begin{aligned}
\mathcal{H}^{\leq N}_{J_2 kep} &= n_* \delta L + \omega_{1*} P + \omega_{2*} Q + \sum_{\substack{s=1 \\ s \neq 2}}^{2N} \mathcal{Z}_s(\delta L, P, Q) \\
&+ \sum_{s=1}^{2N} \sum_{\substack{k_1, k_2, k_3 \in \mathbb{Z} \\ 0 < |k_1| + |k_2| + |k_3| \leq s}} \mathcal{P}_{s, k_1, k_2, k_3}(\delta L, P, Q) \cos(k_1 \lambda + k_2 p + k_3 q).
\end{aligned} \tag{4.2.8}
$$

The functions $\mathcal{Z}_s$ and $\mathcal{P}_{s, k_1, k_2, k_3}(\delta L, P, Q)$ are polynomials of degree $s/2$ in the action variables $(\delta L, P, Q)$. The frequency n_* is equal to:

$$
n_* = \sqrt{\frac{\mu_E}{a_*^3}} + J_2 \frac{3 \mu_E^{1/2} R_E^2}{a_*^{7/2}}, \tag{4.2.9}
$$

while ω_{1*} and ω_{2*} can be retrieved in Section 5.7 by setting $(i_* e_*) = (0, 0)$. The Hamiltonian (4.2.8) is the starting point for the stability estimates on the orbits' semimajor axes; one notices that $\omega_{1*}, \omega_{2*} = \mathcal{O}(J_2)$, a fact implying that both these frequencies are way smaller than $n_* \simeq (\mu_E / a_*^3)^{1/2}$ (third Kepler's law). Accordingly, for all orbits the angle λ circulates at a rate which is $\mathcal{O}(1/J_2)$ faster than the rate of circulation of the angles p, q. Hence, λ constitutes the 'fast angle' of the Hamiltonian $\mathcal{H}^{\leq N}_{J_2 kep}$. Its elimination through a suitable sequence of canonical transformations leads to the approximate constancy of the value of the semimajor axis, as detailed in Section 4.4.

4.2.2 The geolunisolar Hamiltonian

While stability estimates for the semimajor axis depend mostly on the Earth's J_2 term, the question of the long-term stability as regards secular variations in eccentricity and inclination requires considering the effects of the Lunar and Solar gravitational tides. Let us consider a celestial body B (either Moon or Sun) with mass M_b moving around the Earth and whose orbit is exterior to that of the satellite. Let $\underline{r} = (x, y, z)$ and $\underline{r}_b = (x_b, y_b, z_b)$ be the position vectors of S and B in a geocentric reference frame, with $r = |\underline{r}|$ and $r_b = |\underline{r}_b|$. The tidal disturbance caused by B on S is described by the potential

$$
\begin{aligned}
\mathcal{V}_b(\underline{r}, t) &= -\mu_b \left(\frac{1}{|\underline{r} - \underline{r}_b(t)|} - \frac{\underline{r} \cdot \underline{r}_b(t)}{r_b^3(t)} \right) \\
&= -\frac{\mu_b}{r_b(t)} + \frac{\mu_b}{2 r_b^3(t)} r^2 + \frac{3}{2} \frac{\mu_b (\underline{r} \cdot \underline{r}_b(t))^2}{r_b^5(t)} + \mathcal{O}\left(\left(\frac{r}{r_b} \right)^3 \right),
\end{aligned} \tag{4.2.10}
$$

where $\mu_b = \mathcal{G} M_b$. The first term $-\mu_b / r_b$ in the multipolar expansion (4.2.10) does not depend on the coordinates of S, therefore it can be omitted from the Hamiltonian of motion of S. Thus, the tidal (or 'third body') perturbation terms in the Hamiltonian

takes the form:

$$\mathcal{H}_{3B} = \mu_m \left(\frac{1}{2} \frac{r^2}{r_m^3(t)} - \frac{3}{2} \frac{(\underline{r} \cdot \underline{r}_m(t))^2}{r_m^5(t)} \right) + \mu_s \left(\frac{1}{2} \frac{r^2}{r_s^3(t)} - \frac{3}{2} \frac{(\underline{r} \cdot \underline{r}_s(t))^2}{r_s^5(t)} \right) + O_3 = \mathcal{H}_m + \mathcal{H}_s,$$
(4.2.11)

where $\mu_m, \underline{r}_m$ and $\mu_s, \underline{r}_s$ are the mass and geocentric position vectors of the Moon and Sun respectively, and O_3 denotes octupolar or higher order terms in the expansion of the third body potentials. The exact form of the term $\mathcal{H}_{3B}$ depends now on the model adopted for the geocentric orbits of the Sun and Moon. In the framework of this chapter, and of Chapter 5 as well, we adopt the following models for Sun and Moon:

1. we suppose that the Sun's geocentric orbit is an ellipse lying in the Earth's ecliptic plane (i.e., with inclination $i_{s0} = 23.43°$ with respect to the equatorial plane), $\Omega_s = 0°$, $a_s = 1.496 \cdot 10^8 \, km$ and $e_s = 0.0167$;

2. we assume the lunar orbit as elliptic and also lying on the ecliptic plane, with $a_m = 384748 \, km$, $e_m = 0.065$ and $i_{m0} = i_{s0}$. Note that this assumption ignores the precession of the Lunar node (with period $\simeq 18.6$ yr) associated with the inclination of the Moon's orbit with respect to the ecliptic (by $5°15'$). While the precession of the Lunar node is important near secular lunisolar resonances[1], it only has a minimal effect far from these resonances, as substantiated by numerical studies (e.g. [47], [124]). Thus, we ignore this effect in our present estimates.

Under the above approximations, the satellite Hamiltonian $\mathcal{H}_{J_2ls}$ takes the form

$$\mathcal{H}_{J_2ls} = \mathcal{H}_{J_2kep} + \mathcal{H}_{3B}.$$
(4.2.12)

This is a Hamiltonian depending on three degrees of freedom (the coordinates and momenta of the satellite) as well as on time (through the vectors $\underline{r}_m(t)$ and $\underline{r}_s(t)$). However, contrary to the case of the Hamiltonian $\mathcal{H}_{J_2kep}$, in which we are interested in establishing the long-term stability of the semimajor axis over short-period oscillations, here we are interested in the question of the stability of the eccentricity and inclination of the satellite over secular timescales. Thus, as customary (see [45], [125]), we average $\mathcal{H}_{J_2ls}$ with respect to the mean anomalies of the satellite, Moon and Sun. The averaging can be done in closed form (see, for example, [45]), and leads to:

$$\mathcal{H}_{J_2}^{(av)} = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{H}_{J_2kep} dM = \int_0^{2\pi} \mathcal{H}_{J_2kep} \frac{r^2}{a^2\sqrt{1-e^2}} df$$

$$\mathcal{H}_m^{(av)} = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \mathcal{V}_m dM dM_m = \int_0^{2\pi} \int_0^{2\pi} \mathcal{V}_m (1 - e\cos E) \frac{r_m^2}{a_m^2\sqrt{1-e_m^2}} dE df_m$$

$$\mathcal{H}_s^{(av)} = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \mathcal{V}_s dM dM_s = \int_0^{2\pi} \int_0^{2\pi} \mathcal{V}_s (1 - e\cos E) \frac{r_s^2}{a_s^2\sqrt{1-e_s^2}} dE df_s.$$

Here, f, E are the satellite's true and eccentric anomaly, while f_m, f_s are the Moon's and Sun's true anomaly along their geocentric orbits. The averaged J_2 term takes the

---

[1]By secular lunisolar resonances we mean resonances of the form $k_1\dot{\omega} + k_2\dot{\Omega} + k_3\dot{\Omega}_M = 0$, with $(k_1, k_2, k_3) \in \mathbb{Z}^3\backslash\{\underline{0}\}$, thus involving the rate of variation of the longitude of the ascending node of the Moon.

form (apart from constant terms):

$$\mathcal{H}_{J_2}^{(av)}(e,i,\omega,\Omega) = -J_2 \frac{\mu_E R_E^2}{a^3(1-e^2)^{3/2}}\left(\frac{1}{2}-\frac{3}{4}\sin^2 i\right). \tag{4.2.13}$$

The terms $\mathcal{H}_m^{(av)}(e,i,\omega,\Omega)$ and $\mathcal{H}_s^{(av)}(e,i,\omega,\Omega)$, instead, turn out to be identical to those given in equations (3.6) and (3.7) of [42], setting $i_M = 0$. Then, the Hamiltonian averaged over all short period terms, hereafter referred to as the *secular geolunisolar Hamiltonian*, takes the form

$$\mathcal{H}_{gls,sec}(e,i,\omega,\Omega) = \mathcal{H}_{J_2}^{(av)} + \mathcal{H}_s^{(av)} + \mathcal{H}_m^{(av)}, \tag{4.2.14}$$

which, in terms of the Delaunay modified variables, has two degrees of freedom.

Expansion around the forced inclination

As it was done in the case of the J_2 model (Eq. (4.2.8)), normal form computations for the Hamiltonian (4.2.14), expressed in Delaunay action-angle variables, require a polynomial expansion in the action variables around some preselected values. In the case of the secular geolunisolar Hamiltonian (4.2.14), a natural choice of the origin for such expansions is the *forced element* values: writing $\mathcal{H}_{gls,sec}$ as a function of the Delaunay variables, say, $\mathcal{H}_{gls,sec}(P,Q,p,q;a)$ (where the semimajor axis a is now a priori constant, hence, can be considered as a parameter), a forced equilibrium is defined as an equilibrium point of the secular Hamiltonian, i.e., a point $(Q^{(eq)},P^{(eq)},q^{(eq)},p^{(eq)})$ for which the following relations hold:

$$\left(\frac{\partial\mathcal{H}_{gls,sec}}{\partial P}\right)_{eq} = \left(\frac{\partial\mathcal{H}_{gls,sec}}{\partial Q}\right)_{eq} = \left(\frac{\partial\mathcal{H}_{gls,sec}}{\partial p}\right)_{eq} = \left(\frac{\partial\mathcal{H}_{gls,sec}}{\partial q}\right)_{eq} = 0, \tag{4.2.15}$$

where the subscript 'eq' denotes the condition $Q = Q^{(eq)}, P = P^{(eq)}, q = q^{(eq)}, p = p^{(eq)}$. In the case of the Hamiltonian (4.2.14), a forced equilibrium solution can be computed by writing first $\mathcal{H}_{gls,sec}$ in terms of the *Poincaré variables* as

$$\begin{cases} X_1 = \sqrt{2Q}\sin q, & X_2 = \sqrt{2P}\sin p, \\ Y_1 = \sqrt{2Q}\cos q, & Y_2 = \sqrt{2P}\cos p. \end{cases} \tag{4.2.16}$$

Expanding up to quadratic terms in the Poincaré variables, the truncated secular Hamiltonian has the form

$$\widetilde{\mathcal{H}}(Y_1,Y_2,X_1,X_2) = A_1 Y_1 + B_1(X_1^2 + Y_1^2) + B_2(X_2^2 + Y_2^2), \tag{4.2.17}$$

where the coefficients A_1, B_1 are given by:

$$A_1 = -\frac{3R_E^2 a^{7/4}\sin(2i_{s0})}{8(\mathcal{G}M_E)^{1/4}}\left(\frac{\mathcal{G}M_m}{a_m^3}+\frac{\mathcal{G}M_s}{a_s^3}\right),$$
$$B_1 = \frac{3}{4}\frac{\sqrt{\mathcal{G}M_E}R_E^2 J_2}{a^{7/2}} + \frac{3\mathcal{G}M_m(2-3\sin^2 i_{s0})}{16\sqrt{\frac{\mathcal{G}M_E}{a^3}}a_m^3} + \frac{3\mathcal{G}M_s(2-3\sin^2 i_{s0})}{16\sqrt{\frac{\mathcal{G}M_E}{a^3}}a_s^3}. \tag{4.2.18}$$

The Hamiltonian (4.2.17) corresponds to two decoupled harmonic oscillators in the variables (X_1, Y_1) and (X_2, Y_2). The second harmonic oscillator (corresponding to the action-angle pair (P, p), hence, to the orbit's eccentricity vector) has an equilibrium point at $\left(X_2^{(eq)}, Y_2^{(eq)} \right) = (0, 0)$, implying $P^{(eq)} = 0$ and any value $0 \leq p^{(eq)} < 2\pi$. This implies that the sub-manifold of circular orbits $e = 0$ (corresponding to $P = 0$) is invariant under the flow of the secular geolunisolar Hamiltonian. On the other hand, as regards the pair (X_1, Y_1), Hamilton's equations for the Hamiltonian (4.2.17) yield:

$$\begin{cases} \dot{X}_1 = A_1 + 2B_1 Y_1 \\ \dot{Y}_1 = -2B_1 X_1 \ . \end{cases} \tag{4.2.19}$$

For $i_{s0} \neq 0$, the equilibrium point of (4.2.19) is given by

$$X_1^{(eq)} = 0 \ , \quad Y_1^{(eq)} = -\frac{A_1}{2B_1} \neq 0.$$

Setting $Q^{(eq)} = \left(\left(X_1^{(eq)} \right)^2 + \left(Y_1^{(eq)} \right)^2 \right)/2$, $i_{eq} \simeq (2Q^{(eq)}/\sqrt{\mu_E a})^{1/2}$ (for $Q^{(eq)}$ small), we arrive at

$$i^{(eq)} \simeq -\frac{A_1}{2B_1} \frac{1}{(\mu_E a)^{1/4}}, \quad q^{(eq)} = -\Omega^{(eq)} = 0. \tag{4.2.20}$$

More accurate expressions for the *forced inclination* $i^{(eq)}$ can be obtained by introducing (4.2.20) along with the remaining equilibrium values in the derivatives of the full secular Hamiltonian (4.2.14) and finding the roots of Hamilton's equations. One can readily verify that $q^{(eq)} = 0$ at all orders, while $i^{(eq)}$ is subject to small corrections with respect to the expression (4.2.20). In physical terms, the forced inclination $i^{(eq)}$ defines the inclination of the *Laplace plane*: since the perturbing bodies (Moon and Sun) are in orbits inclined with respect to the equator, a satellite orbit can maintain its inclination constant when the latter has the value $i^{(eq)}$. Inspecting the form of the coefficients (4.2.18), we find that $i^{(eq)} \to 0$ as $a \to 0$, while it can be shown that $i^{(eq)} \to i_{s0}$ for values of a greater than the GEO one (see for example [126]), reflecting the fact that the Laplace plane tends to coincide with the equator for satellite orbits close to the Earth (as imposed by the oblateness of the Earth), while it tends to coincide with the ecliptic at large distances from the Earth (where the Lunar and Solar tides dominate).

Returning to the expansion of the secular geolunisolar Hamiltonian, making the shift transformation $\delta Y_1 = Y_1 - Y_1^{(eq)}$ allows us to express the Hamiltonian as a polynomial in the variables $(X_1, \delta Y_1)$ and (X_2, Y_2). The Hamiltonian $\mathcal{H}_{gls,sec}$ starts now with terms of second degree which we regroup in $\mathcal{H}_2$:

$$\mathcal{H}_2 = \frac{b_1 + \epsilon_1}{2} X_1^2 + \frac{b_1 + \epsilon_2}{2} \delta Y_1^2 + \frac{b_1 + \epsilon_3}{2} X_2^2 + \frac{b_1 + \epsilon_4}{2} Y_2^2, \tag{4.2.21}$$

where $b_1 = 2B_1$ and ϵ_1, ϵ_2, ϵ_3, ϵ_4 are corrections of order $\mathcal{O}(\mu_b a^{3/2}/(\mu_E^{1/2} a_b^3))$, with the index b referring to the Moon or Sun. All these corrections turn to be rather small, with relative size $\sim 10^{-3} B_{10}$ at semimajor axis $a \sim 10^4$ km, where

$$B_{10} = \frac{3}{4} \frac{\sqrt{\mathcal{G}M_E} R_E^2 J_2}{a^{7/2}}.$$

Thus, after a canonical rescaling $X_1 = c_{12}\tilde{X}_1$, $\delta Y_1 = \tilde{Y}_1/c_{12}$, $X_2 = c_{34}\tilde{X}_2$, $Y_2 = \tilde{Y}_2/c_{34}$, with $(c_{12})^4 = (b_1 + \epsilon_2)/(b_1 + \epsilon_1) = 1 + \mathcal{O}(\mu_b a^{3/2}/(B_{10}\mu_E^{1/2}a_b^3))$ and $(c_{34})^4 = (b_1+\epsilon_4)/(b_1+\epsilon_3) = 1+\mathcal{O}(\mu_b a^{3/2}/(B_{10}\mu_E^{1/2}a_b^3))$, the secular lunisolar Hamiltonian resumes the form:

$$
\begin{aligned}
\mathcal{H}_{gls,sec} =\ & \frac{\nu_1}{2}\left(\tilde{X}_1^2 + \tilde{Y}_1^2\right) + \frac{\nu_2}{2}\left(\tilde{X}_2^2 + \tilde{Y}_2^2\right) \\
& + \sum_{s=3}^{\infty} \sum_{\substack{k_1,k_2,l_1,l_2\in\mathbb{N}\\k_1+k_2+l_1+l_2=s}} h_{k_1,k_2,l_1,l_2}\tilde{X}_1^{k_1}\tilde{X}_2^{k_2}\tilde{Y}_1^{l_1}\tilde{Y}_2^{l_2}.
\end{aligned}
\tag{4.2.22}
$$

This is the typical form of a secular Hamiltonian, consisting of linear oscillators (with frequencies ν_1, ν_2) coupled with nonlinear terms. However, we have $\nu_1 = \nu_2 + \mathcal{O}(\mu_b a^{3/2}/\mu_E^{1/2}a_b^3)$, implying that the two frequencies are nearly equal

$$
\nu_1 \simeq \nu_2 \simeq \frac{3}{2}\frac{\sqrt{\mathcal{G}M_E}R_E^2 J_2}{a^{7/2}}\quad.
$$

This is a consequence of the axisymmetry of the J_2 model, implying that the secular frequencies $\dot{q} = -\dot{\Omega}$ and $\dot{p} = -\dot{\omega} - \dot{\Omega}$ are equal for nearly equatorial orbits in this model. As we will see in Section 4.5, this near-equality implies that with the present normal form estimates one cannot establish independently the long term stability of the eccentricity and the inclination, but only the long-term stability of the Kozai-Lidov integral $\mathcal{I} = \tilde{X}_1^2 + \tilde{Y}_1^2 + \tilde{X}_2^2 + \tilde{Y}_2^2$, which couples oscillations between the eccentricity and the proper inclination of the satellite.

Finally, the Hamiltonian (4.2.22) can be written in action-angle variables $\tilde{X}_1 = \sqrt{2I_1}\sin\phi_1$, $\tilde{Y}_1 = \sqrt{2I_1}\cos\phi_1$, $\tilde{X}_2 = \sqrt{2I_2}\sin\phi_2$, $\tilde{Y}_2 = \sqrt{2I_2}\cos\phi_2$ as

$$
\begin{aligned}
\mathcal{H}_{gls,sec} =\ & \nu_1 I_1 + \nu_2 I_2 \\
& + \sum_{s=3}^{\infty} \sum_{\substack{s_1,s_2\in\mathbb{N}\\s_1+s_2=s}} \sum_{\substack{k_1,k_2\in\mathbb{Z}\\|k_1|+|k_2|\leq s\\(|k|_1+|k|_2)\equiv s\,(\mathrm{mod}\,2)}} \tilde{h}_{s_1,s_2,k_1,k_2}I_1^{s_1/2}I_2^{s_2/2}\cos(k_1\phi_1 + k_2\phi_2).
\end{aligned}
\tag{4.2.23}
$$

The Hamiltonian (4.2.23) is the starting point for all normal form calculations in Section 4.5. For computational reasons, the expansion in (4.2.23) is truncated up to a maximal order $N = 15$, leading to the truncated form

$$
\begin{aligned}
\mathcal{H}_{gls,sec}^{\leq N}(I_1, I_2, \phi_1, \phi_2) =\ & \nu_1 I_1 + \nu_2 I_2 \\
& + \sum_{s=3}^{N} \sum_{\substack{s_1,s_2\in\mathbb{N}\\s_1+s_2=s}} \sum_{\substack{k_1,k_2\in\mathbb{Z}\\|k_1|+|k_2|\leq s\\(|k|_1+|k|_2)\equiv s\,(\mathrm{mod}\,2)}} \tilde{h}_{s_1,s_2,k_1,k_2}I_1^{s_1/2}I_2^{s_2/2}\cos(k_1\phi_1 + k_2\phi_2).
\end{aligned}
$$

4.3 Hamiltonian Normalization

In this Section we briefly recall some basic definitions related to normal form theory and its use in obtaining stability estimates based on the size of the normal form's *remainder*.

In Sections 4.4 and 4.5 we will discuss the particular normalizations implemented on the Hamiltonians (4.2.8) and (4.2.24) respectively.

4.3.1   Normal form and remainder

Consider a Hamiltonian function of the form

$$\mathcal{H}(\underline{A}, \underline{\varphi}) = Z_0(\underline{A}) + \mathcal{H}_1(\underline{A}, \underline{\varphi}) = \underline{\omega} \cdot \underline{A} + \mathcal{H}_1(\underline{A}, \underline{\varphi}), \qquad (4.3.1)$$

where ω_j are real constants, and $(\underline{A}, \underline{\varphi}) \in \mathbb{R}^n \times \mathbb{T}^n$ are action-angle variables. We assume that $\mathcal{H}_1$ is analytic in the complex domain $(\underline{A}, \underline{\varphi}) \in D_{\rho, \sigma}(U) = B_\rho U \times S_\sigma$ (or simply $D_{\rho, \sigma}$), where U is an open domain of $\mathbb{R}^n$, $B_\rho U$ is a complex neighborhood of U of size ρ:

$$B_\rho U = \{\underline{A} \in \mathbb{C}^n : \; dist(\underline{A}, U) < \rho\}, \qquad (4.3.2)$$

S_σ is the complex strip

$$S_\sigma = \{\underline{\varphi} \in \mathbb{C}^n : \; Re(\varphi_j) \in \mathbb{T}, \; |Im(\varphi_j)| < \sigma, \quad j = 1, \dots, n\} \qquad (4.3.3)$$

for $\rho, \sigma > 0$. On $D_{\rho, \sigma}(U)$ we define the norm of a function $f = f(\underline{A}, \underline{\varphi})$ as

$$\|f\|_{\rho, \sigma} = \sup_{(\underline{A}, \underline{\varphi}) \in D_{\rho, \sigma}} |f(\underline{A}, \underline{\varphi})|. \qquad (4.3.4)$$

The aim of normalization theory is to introduce a near to identity canonical transformation $\Phi : \; (\underline{A}, \underline{\varphi}) \to (\underline{A}', \underline{\varphi}')$, so that in the new variables $(\underline{A}', \underline{\varphi}')$ the Hamiltonian (4.3.1) takes the form

$$\mathcal{H}\left(\underline{A}(\underline{A}', \underline{\varphi}'), \underline{\varphi}(\underline{A}', \underline{\varphi}')\right) = Z(\underline{A}', \underline{\varphi}') + R(\underline{A}', \underline{\varphi}') \qquad (4.3.5)$$

with the following properties:
i) the transformation Φ is analytic in a domain $D_{\rho', \sigma'}(U)$ with $0 < \rho' < \rho$, $0 < \sigma' < \sigma$,
ii) the dynamics under $Z(\underline{A}', \underline{\varphi}')$, called the *normal form*, has some desired properties (see below), and
iii) under the norm definition (4.3.4) one has $\|R\|_{\rho', \sigma'} \ll \|Z\|_{\rho', \sigma'}$ implying that the function $R(\underline{A}', \underline{\varphi}')$, called the *remainder*, introduces only a small correction with respect to the flow under the normal form term $Z(\underline{A}', \underline{\varphi}')$.

Regarding point ii) above, see, e.g., [51] for a definition of the properties of the normal form term in various contexts of perturbation theory (e.g. in the Kolmogorov-Arnold-Moser or Nekhoroshev theories). Here we mention three cases of particular interest, pertinent to our context:

Case 1: Birkhoff normal form. The function Z is independent of the angles $\underline{\varphi}'$. This kind of normalization allows us to prove the near-constancy of the action variables $\underline{A}$.

Case 2: elimination of short-period terms. The real constants ω_j in (4.3.1), called the unperturbed frequencies, are divided in two groups, 'fast' $\{\omega_1, \dots, \omega_{K_f}\}$, $1 \leq K_f < n$, and 'slow' $\{\omega_{K_f+1}, \dots, \omega_n\}$, such that $\min\{|\omega_1|, \dots, |\omega_{K_f}|\} \gg \max\{|\omega_{K_f+1}|, \dots, |\omega_n|\}$. In this case, it turns convenient to introduce a normalizing transformation Φ such that

the normal form Z becomes independent of the 'fast angles' $\{\varphi'_1, \ldots, \varphi'_{K_f}\}$. Such is the case of the normal form encountered in Section 4.4, leading to estimates on the stability of the semimajor axis in the J_2 problem. The corresponding Hamiltonian is of the form (4.3.1), with $n = 3$, $A_1 = \delta L$, $A_2 = P$, $A_3 = Q$, $\varphi_1 = \lambda$, $\varphi_2 = p$, $\varphi_3 = q$.

Case 3: resonant normal form. The frequencies ω_j satisfy one or more quasi - commensurability relations of the form $\underline{m} \cdot \underline{\omega} \simeq 0$, with $\underline{m} \in \mathbb{Z}^n, |\underline{m}| \neq 0$. The maximum number of linearly independent and irreducible integer vectors $\underline{m}_l$, $1 \leq l \leq l_{max}$, yielding exact commensurabilities for a given set of frequencies ω_j, satisfies $0 \leq l_{max} \leq n$. Since $\mathcal{H}_1$ is analytic in $D_{\rho,\sigma}(U)$ and periodic in $\underline{\varphi}$, $\mathcal{H}_1$ admits the Fourier decomposition

$$\mathcal{H}_1(\underline{A}, \underline{\varphi}) = \sum_{\underline{k} \in \mathbb{Z}^n} h_{1,\underline{k}}(\underline{A}) e^{i\underline{k} \cdot \underline{\varphi}}, \tag{4.3.6}$$

where, according to Fourier theorem, the coefficients $|h_{1,\underline{k}}(\underline{A})|$ are bounded by exponentially decaying quantities $\mathcal{O}(e^{-|\underline{k}|\sigma})$. Then, it turns out that the appropriate normal form Z has the *resonant form*:

$$Z(\underline{A}', \underline{\varphi}') = \sum_{\underline{k} \in \mathcal{M}} \zeta_{\underline{k}}(\underline{A}') e^{i\underline{k} \cdot \underline{\varphi}'}, \tag{4.3.7}$$

for some Fourier coefficients $\zeta_{\underline{k}}(\underline{A}')$ and where

$$\mathcal{M} := \{\underline{k} \in \mathbb{Z}^n : \underline{k} \cdot \underline{m}_l = 0 \text{ for all } l = 1, \ldots, l_{max}\}$$

is the 'resonant module'. A normal form of the form (4.3.7) implies the existence of $n - l_{max}$ quasi-integrals of the form $I_i = \underline{K}_i \cdot \underline{A}$, $i = 1, \ldots, n - l_{max}$, where the vectors $\underline{K}_i$ satisfy the equations $\underline{K}_i \cdot \underline{m}_l = 0$ for all l with $1 \leq l \leq l_{max}$. The quantities I_i are called the *resonant integrals* of the Hamiltonian (4.3.7).

As an example, whenever $\nu_1 = \nu_2$ the secular geolunisolar Hamiltonian (4.2.24) admits a resonant normal form. We have $n = 2$, $l_{max} = 1$, $\underline{m}_1 = (1, -1)$, $A_1 = I_1$, $A_2 = I_2$, $\varphi_1 = \phi_1$, $\varphi_2 = \phi_2$. Therefore, the normal form contains terms independent of the angles or depending on the angles through trigonometric terms of the form $\cos(k(\phi'_1 - \phi'_2))$, $k = 1, 2, \ldots$. The associated resonant integral corresponds to the 'Kozai-Lidov' integral $\mathcal{I} = I_1 + I_2$ (see [61]).

Definition 4.3.1. *A r-th step Hamiltonian normalization process is a composition of near identity transformations*

$$\Phi^{(r)} = \Phi_r \circ \Phi_{r-1} \circ \ldots \circ \Phi_1 \tag{4.3.8}$$

mapping the initial action-angle variables to the r-th step normalized action-angle variables via the successive transformations $(\underline{A}^{(s)}, \underline{\phi}^{(s)}) = \Phi^{(s)}(\underline{A}^{(s-1)}, \underline{\phi}^{(s-1)})$, *$s = 1, 2, \ldots, r$, $(\underline{A}^{(0)}, \underline{\varphi}^{(0)}) \equiv (\underline{A}, \underline{\varphi})$, defined so that the compositions*

$$\Phi^{(s)} = \Phi_s \circ \Phi_{s-1} \circ \ldots \circ \Phi_1$$

for all $s = 1, \ldots, r$ are analytic and with inverse analytic within non-null domains $D_{\rho^{(s)}, \sigma^{(s)}} \neq \emptyset$, and the rth-step Hamiltonian takes the form:

$$\mathcal{H}^{(r)}\left(\underline{A}(\underline{A}^{(r)}, \underline{\varphi}^{(r)}), \underline{\varphi}(\underline{A}^{(r)}, \underline{\varphi}^{(r)})\right) = Z^{(r)}(\underline{A}^{(r)}, \underline{\varphi}^{(r)}) + R^{(r)}(\underline{A}^{(r)}, \underline{\varphi}^{(r)}) \tag{4.3.9}$$

with $\|R^{(r)}\|_{\rho^{(r)}, \sigma^{(r)}} \ll \|Z^{(r)}\|_{\rho^{(r)}, \sigma^{(r)}}$.

The semi-analytical estimates of stability that we will develop in the next Sections are based on defining a suitable r-step sequence of canonical transformations $\Phi_1, \Phi_2, \ldots, \Phi_r$ reducing the size of the remainder $\|R^{(r)}\|_{\rho^{(r)}, \sigma^{(r)}}$ as much as possible given the initial Hamiltonian model considered. The appropriate sequence is found using the method of *Lie series* (see Section 4.4). The obtained times of stability are of order $\|R^{(r_{opt})}\|_{\rho^{(r_{opt})}, \sigma^{(r_{opt})}}^{-1}$, where r_{opt} is the normalization order yielding the smallest possible remainder norm. The value of r_{opt} can be obtained via theoretical estimates (see [127]), but in practice, it is also limited by the maximum order in which our computer-algebra normal form calculations can proceed. Theoretical estimates imply that the size of the remainder norm is exponentially small in the inverse of the size of the perturbation $\|\mathcal{H}_1\|_{\rho, \sigma}$ in Eq. (4.3.1). For example, in the simplest case of the Birkhoff normal form, we have the following theorem (see [128] for full details).

Theorem 4.3.2. *Consider the Hamiltonian expressed in action-angle variables $\mathcal{H}(\underline{A}, \underline{\varphi}) = \underline{\omega} \cdot \underline{A} + f(\underline{A}, \underline{\varphi})$, where $\underline{\omega} \in \mathbb{R}^n$ satisfies the following Diophantine condition: there exist $\tau, \gamma > 0$ such that*

$$|\underline{k} \cdot \underline{\omega}| \geq \frac{\gamma}{|\underline{k}|^\tau} \qquad \forall \underline{k} \in \mathbb{Z}^n \backslash \{\underline{0}\} \tag{4.3.10}$$

and f is real analytic on $D_{\rho, \sigma}$ for some $\rho, \sigma > 0$. Consider two positive parameters $\delta < \rho/2$ and $\xi < \sigma/2$, and for $r \geq 1$, let

$$\epsilon_1^* = \frac{\gamma \delta \xi^{\tau+1}}{2^{n-\tau+4}\sqrt{(2\tau+2)!}\, \|f\|_{\rho, \sigma}} \ , \qquad \epsilon_r^* = \frac{\epsilon_1^*}{r^{\tau+2}} \ . \tag{4.3.11}$$

Then, for any

$$r < \left(\frac{\gamma \delta \xi^{\tau+1}}{2^{n-\tau+4}\sqrt{(2\tau+2)!}}\right)^{1/\tau+2} \frac{1}{\|f\|_{\rho, \sigma}^{1/\tau+2}} \ , \tag{4.3.12}$$

there exists a real analytic canonical transformation $\Phi : D_{\rho-2\delta, \sigma-2\xi} \mapsto D_{\rho, \sigma}$ such that the transformed Hamiltonian has the form

$$\mathcal{H} \circ \Phi = h(\underline{A}) + \sum_{s=1}^{r} Z_s(\underline{A}) + R^{(r+1)}(\underline{A}, \underline{\varphi}) \ , \tag{4.3.13}$$

where the remainder $R^{(r+1)}$ can be bounded as

$$\|R^{(r+1)}\|_{\rho-2\delta, \sigma-2\xi} \leq \frac{\|f\|_{\rho, \sigma}}{4r^{\tau+2}}\left(\frac{1}{\epsilon_r^*}\right)^r \frac{\epsilon_r^*}{\epsilon_r^* - 1} \ . \tag{4.3.14}$$

Casting together (4.3.11) and (4.3.14), one readily sees that the remainder grows more rapidly than any power of r, namely as $(r^{\tau+2})^{r-1}$. Consequently, this procedure

does not converge for $r \to \infty$. In any case, we remark that, as the threshold value for the normalization order r is proportional to the inverse of $\|f\|_{\rho,\sigma}^{1/(\tau+2)}$, if we manage to reduce the size of the initial remainder function, then we can increase the maximum value of r for which Theorem 4.3.2 is satisfied.

Similar estimates hold in the case of the resonant normal form constructions (see [127]). The behavior of the size of the remainder as a function of the normalization order r will be examined in detail in our semi-analytical computations in Sections 4.4 and 4.5 below.

4.3.2   Book-keeping and construction of the normal form

Both Hamiltonians (4.2.8) and (4.2.24) are of the form (4.3.1), therefore the above results on Hamiltonian normalization apply. In order to compute the composition of canonical transformations required in Eq. (4.3.8), we implement the method of *composition of Lie series*, after introducing a suitable *book-keeping* (see [51]) to separate terms in the Hamiltonian according to estimates of their order of smallness.

Definition 4.3.3. *Consider a 'book-keeping symbol' ϵ, with numerical value $\epsilon = 1$. A book-keeping rule is a splitting of the initial Hamiltonian $\mathcal{H}(\underline{A}, \underline{\varphi})$ in the form*

$$\mathcal{H}(\underline{A}, \underline{\varphi}) = \underline{\omega} \cdot \underline{A} + \sum_{s=1}^{\infty} \epsilon^s \mathcal{H}_s(\underline{A}, \underline{\varphi}). \tag{4.3.15}$$

Remark 4.3.4. *The splitting can in principle be arbitrary. However, the sequence of remainders $\|R^{(r)}\|_{\rho^{(r)},\sigma^{(r)}}$ found by Hamiltonian normalization behaves well, i.e. $\|R^{(s)}\|_{\rho^{(s)},\sigma^{(s)}} < \|R^{(s-1)}\|_{\rho^{(s-1)},\sigma^{(s-1)}}$ for $s = 1,\ldots,r_{opt}$ when the splitting (4.3.15) is done so as to reflect the order of smallness of different terms in the Hamiltonian. Roughly speaking, one must have $\|\mathcal{H}_s\|_{\rho,\sigma} = \mathcal{O}\left(\|\mathcal{H}_1\|_{\rho,\sigma}^s\right)$ (see [51]).*

Proposition 4.3.5. Lie series: *Let $\chi(\underline{A}, \underline{\varphi})$, called the Lie generating function, be a function analytic in the domain $D_{\rho,\sigma}(U)$, and $\mathcal{L}_\chi$ denote the Poisson bracket operator $\mathcal{L}_\chi \cdot = \{\cdot, \chi\}$. Given positive numbers $\delta < \rho$ and $\xi < \sigma$, assume that*

$$\min\left(\delta \left\|\frac{\partial \chi}{\partial q}\right\|_{\rho-\delta,\sigma-\xi}^{-1}, \xi \left\|\frac{\partial \chi}{\partial p}\right\|_{\rho-\delta,\sigma-\xi}^{-1}\right) > 1.$$

Then, the mapping

$$(\underline{A}', \underline{\varphi}') = \exp(\mathcal{L}_\chi)(\underline{A}, \underline{\varphi}) = \sum_{j=0}^{\infty} \frac{1}{j!} \mathcal{L}_\chi^j (\underline{A}, \underline{\varphi}) \tag{4.3.16}$$

is an analytic canonical transformation of the domain $D_{\rho-\delta,\sigma-\xi}(U)$ onto itself.

The proof consists in implementing Proposition 1 of [128] with $r = 1$.

Proposition 4.3.6. Exchange theorem: *Let f be a real analytic function $f : U \times \mathbb{T}^n \to \mathbb{R}$ extended to the domain $D_{\rho,\sigma}(U)$. The equality*

$$f(\underline{A}', \underline{\varphi}') = \left(\exp(\mathcal{L}_\chi) f(\underline{A}, \underline{\varphi})\right)_{\underline{A}=\underline{A}', \underline{\varphi}=\underline{\varphi}'} \tag{4.3.17}$$

holds, where $(\underline{A}', \underline{\varphi}')$ are given by the transformation (4.3.16) and $(\underline{A}', \underline{\varphi}') \in D_{\rho-\delta, \sigma-\xi}(U)$.

See [129] for the proof. In simple words, the exchange theorem implies that the result of a Lie series canonical transformation onto a function depending on $(\underline{A}, \underline{\varphi})$ can be found by implementing the sequence of Poisson brackets of the exponential operator $\exp(\mathcal{L}_\chi)$ directly on the function f, and substituting, after this operation, the arguments $(\underline{A}, \underline{\varphi})$ with $(\underline{A}', \underline{\varphi}')$.

Normal form algorithm: The above definitions allow us to establish an algorithm for the calculation of the sequence of canonical transformations (4.3.8) using Lie series. The algorithm is obtained recursively by defining the r-th step as follows. Assume the Hamiltonian after $r-1$ normalization steps, denoted by $\mathcal{H}^{(r-1)}$, is in normal form up to the book-keeping order $r-1$:

$$\mathcal{H}^{(r-1)} = Z_0 + \epsilon Z_1 + \ldots + \epsilon^{r-1} Z_{r-1} + \epsilon^r R_r^{(r-1)} + \epsilon^{r+1} R_{r+1}^{(r-1)} + \epsilon^{r+2} R_{r+2}^{(r-1)} + \ldots \quad (4.3.18)$$

Then, the r-th step Lie generating function χ_r and Hamiltonian $\mathcal{H}^{(r)}$ are computed as follows:

(i) split $R_r^{(r-1)}$ as $R_r^{(r-1)} = Z_r^{(r-1)} + h_r^{(r-1)}$, where $Z_r^{(r-1)}$ denotes the part of $R_r^{(r-1)}$ being in normal form;

(ii) compute χ_r as the solution of the homological equation

$$\{\underline{\omega} \cdot \underline{A}, \chi_r\} + \epsilon^r h_r^{(r-1)} = 0; \quad (4.3.19)$$

(iii) compute the r-th step normalized Hamiltonian as $\mathcal{H}^{(r)} = \exp(\mathcal{L}_{\chi_r})\mathcal{H}^{(r-1)}$. This yields the Hamiltonian

$$\mathcal{H}^{(r)} = Z_0 + \epsilon Z_1 + \ldots + \epsilon^{r-1} Z_{r-1} + \epsilon^r Z_r + \epsilon^{r+1} R_{r+1}^{(r)} + \epsilon^{r+2} R_{r+2}^{(r)} + \ldots \quad (4.3.20)$$

where $Z_r = Z_r^{(r-1)}$.

Remark 4.3.7. *In the above algorithm, the notation $f^{(r)}$ implies a function depending on the canonical variables $(\underline{A}^{(r)}, \underline{\varphi}^{(r)})$, which are connected to the original variables $(\underline{A}, \underline{\varphi})$ via the composition of Lie series transformations*

$$(\underline{A}, \underline{\varphi}) = \exp(\mathcal{L}_{\chi_r}) \circ \exp(\mathcal{L}_{\chi_{r-1}}) \circ \ldots \circ \exp(\mathcal{L}_{\chi_1})(\underline{A}^{(r)}, \underline{\varphi}^{(r)}). \quad (4.3.21)$$

For simplicity of notation, unless explicitly required in the sequel we do not write the superscripts in the canonical variables defined in every step, but only in the functions in which these variables are arguments of.

Remark 4.3.8. *In the computer-algebraic implementation of the normalization algorithm, all functions are truncated up to a maximum book-keeping order, specified by computational restrictions.*

Remark 4.3.9. *The solution of the homological equation (4.3.19) is trivial when the functions $h_r^{(r-1)}$ are written in the Fourier representation*

$$h_r^{(r-1)} = \sum_{\underline{k} \in \mathbb{Z}^n} \tilde{h}_{r,\underline{k}}^{(r-1)}(\underline{A}) \exp(i\underline{k} \cdot \underline{\varphi}),$$

which gives

$$\chi_r = \epsilon^r \sum_{\underline{k} \in \mathbb{Z}^n} \frac{i\tilde{h}_{r,\underline{k}}^{(r-1)}(\underline{A})}{\underline{k} \cdot \underline{\omega}} \exp(i\underline{k} \cdot \underline{\varphi}).$$

4.4 Stability of the semimajor axis in the J_2 model

We will now implement the Hamiltonian normalization discussed in Section 4.3 to eliminate the short period terms (depending on the mean longitude λ) in the Hamiltonian (4.2.8), leading to estimates on the long-term stability of the orbits' semimajor axis.

4.4.1 Normal form

We express the Hamiltonian function in the form (4.3.15), choosing the book-keeping power equal to $s-2$, where s is the index in the Hamiltonian expansion (4.2.8), that is, collecting together at book-keeping order s all polynomials $\mathcal{Z}_{s-2}$ and $\mathcal{P}_{s-2,k_1,k_2,k_3}$. Then

$$
\begin{aligned}
\mathcal{H}_{J_2}^{(0)}(\delta L, P, Q, \lambda, p, q) = \\
= \mathcal{H}_0(\delta L, P, Q) + \epsilon \mathcal{H}_1(\delta L, P, Q, \lambda, p, q) + \cdots + \epsilon^N \mathcal{H}_N(\delta L, P, Q, \lambda, p, q)
\end{aligned}
\tag{4.4.1}
$$

with $\mathcal{H}_0 = n_* \delta L + \omega_{1*} P + \omega_{2*} Q$. The truncation order (in eccentricity and inclination) is $N = 15$.

With reference to the algorithm of Subsection 4.3.2, normal form terms are specified as those non-depending on the mean longitude λ. That is, the Fourier harmonics $\cos(k_1\lambda + k_2 p + k_3 q)$ to survive in normal form are selected by the choice of resonant module (see Subsection 3.1) defined by the relation $l_{max} = 1$, $m_1 = (1, 0, 0)$. Following the choice of book-keeping as described above, the normal form and remainder computations where done using a program written by the authors in the symbolic package *Mathematica*©. The symbolic program performs $M = 12$ normalization steps, implementing steps (i) to (iii) of the normalization algorithm discussed at the end of Section 4.3. The operation $H^{(r)} = \exp(\mathcal{L}_{\chi_r}) H^{(r-1)}$ is truncated in book-keeping up to the order $N = 15$. The symbolic program performs this truncation automatically, since every term in both the Hamiltonians $H^{(r-1)}$ and the generating functions χ_r, $r = 1, \ldots, M$ carries the book-keeping symbol ϵ raised to some power. Mild memory requirements (of the order of 100MB) are required for the whole process. We also note that, for given book-keeping rule, this process takes automatically care of the minimum and maximum powers in the action variables, as well as the minimum and maximum Fourier harmonics associated with every book-keeping order (see [51] for a detailed description of this process).

After M normalization steps, the Hamiltonian takes the form

$$\mathcal{H}_{J_2} \equiv \mathcal{H}_{J_2}^{(M)}(\delta L, P, Q, \lambda, p, q) = \mathcal{H}_{J_2, sec}^{(M)}(\delta L, P, Q, p, q) + \mathcal{R}_{J_2}^{(M)}(\delta L, P, Q, \lambda, p, q) , \quad (4.4.2)$$

where (setting the book-keeping $\epsilon = 1$)

$$\mathcal{H}_{J_2, sec}^{(M)} = Z_0 + \ldots + Z_M,$$
$$\mathcal{R}_{J_2}^{(M)} = R_{M+1}^{(M)} + \ldots + R_N^{(M)}$$

with $Z_0 = \mathcal{H}_0$.

The term $\mathcal{H}_{J_2, sec}^{(M)}$ will be referred to as the 'secular Hamiltonian' (not depending on the fast angle λ). On the other hand, the remainder $\mathcal{R}_{J_2}^{(M)}$ quantifies the difference between the true evolution of all canonical variables and the one induced by $\mathcal{H}_{J_2, sec}^{(M)}$. Since in (4.4.2) we can only compute a truncated remainder, we probe numerically that the finite sum of the leading terms in the remainder (up to order N) yields a remainder norm close to the limiting one (which corresponds to the limit $N \to \infty$). To this end, we take as maximum normalization order $M = N - 3$, ensuring that at least the three first leading terms are included in the remainder (see [60]). Also, in estimating the size of the remainder through a suitable definition of the norm, we compute the *sup* norm on a closed and bounded domain $\mathcal{D} \subset \mathbb{R}^2$:

$$\|f\|_{\infty,\mathcal{D}} = \sup_{\substack{(e,i)\in\mathcal{D} \\ (\lambda,p,q)\in\mathbb{T}^3}} |f(e, i, \lambda, p, q)| . \quad (4.4.3)$$

The calculation of the sup norm in a fixed domain $(e, i) \in \mathcal{D}$, $(\lambda, p, q) \in \mathbb{T}^3$ can only be done approximately, by taking a grid of values for all variables involved in this domain, and computing the maximum of the absolute value of the function involved on this grid. Since this process can lead to quite strong fluctuations in the norm estimate, for practical purposes we substitute the sup norm calculation with one based on majorization: consider a function of the form

$$f(e, i, \lambda, p, q) = \sum_{k_1, k_2, k_3} f_{k_1, k_2, k_3}(e, i) \cos(k_1\lambda + k_2 p + k_3 q), \quad (4.4.4)$$

where the sum is over an arbitrary (finite) number of harmonics $(k_1, k_2, k_3) \in \mathbb{Z}^3$ with $|k_1| + |k_2| + |k_3| \neq 0$, and the functions $f_{k_1, k_2, k_3}(e, i)$ are sums of polynomials

$$f_{k_1, k_2, k_3}(e, i) = \sum_{s_1} \sum_{s_2} g_{k_1, k_2, k_3, s_1, s_2} P^{s_1/2}(e) Q^{s_2/2}(i) \quad (4.4.5)$$

over a finite set of integer pairs s_1, s_2. Define the domain $\mathcal{D}_*(e_*, i_*)$ in the action space (P, Q) via the relation: $\mathcal{D}_*(e_*, i_*) = \{0 \leq P \leq P(e_*), 0 \leq Q \leq Q(e_*, i_*)\}$. Then, one has:

$$\sup_{\substack{(e,i)\in\mathcal{D}_* \\ (\lambda,p,q)\in\mathbb{T}^3}} |f(e, i, \lambda, p, q)| \leq \|f\|_{\infty,\mathcal{D}_*} = \sum_{k_1, k_2, k_3} \sum_{s_1, s_2} |g_{k_1, k_2, k_3, s_1, s_2}| P^{s_1/2}(e_*) Q^{s_2/2}(e_*, i_*)$$

$$(4.4.6)$$

We can then use the norm (4.4.6) as an estimate substituting the sup norm in all actual calculations.

4.4.2 Numerical results: stability of the semimajor axis

Having fixed the procedure for the normal form and remainder computations, we proceed in deriving stability estimates based on the time variations of the value of the semimajor axis in the J_2 problem. Fixing a reference value a_* of the semimajor axis, we assume that, at the time $t = 0$, we have $L = L_* = \sqrt{\mu a_*}$, i.e. $\delta L = 0$. Our aim is to estimate the fluctuations of L as functions of the orbital parameters e and i.

The first question to settle is that, for every value of the reference parameter a_* we have to specify the range of values of the variables (e, i) for which the remainder $\mathcal{R}_{J_2}^{(M)}$ is small enough to represent only a perturbation with respect to the dynamics determined by the secular part. In applications, we compute the value of $\|\mathcal{R}_{J_2}^{(M)}\|_{\infty, \mathcal{D}_*}$ in the domain $(e, i) \in \mathcal{D} = [0, 0.15] \times [0, \pi/2]$, so that the inclination can take all possible values; the eccentricity is instead taken in a reasonable interval, where we can find almost all main Earth's satellites.

With reference to the Hamiltonian (4.4.2), if we consider the dynamics induced only by the secular part, we obtain that

$$\frac{d}{dt}\delta L = -\frac{\partial \mathcal{H}_{J_2, sec}^{(M)}}{\partial \lambda} = 0 \ ,$$

which implies that δL (hence L) is a constant of motion. We remind that δL is not the original Delaunay variable, but rather the one obtained after M normalization steps. If we denote by $\delta L^{(0)}$ the original variable, then we have

$$\delta L = \exp(-\mathcal{L}_{\chi^{(1)}}(\ldots(\exp(-\mathcal{L}_{\chi^{(M)}}(\delta L^{(0)}))))) \ . \tag{4.4.7}$$

To obtain $\delta L^{(0)}$ as a function of the new variable δL, we need to invert the transformation (4.4.7); we observe that

$$(\exp(\mathcal{L}_\chi))^{-1} = \exp(-\mathcal{L}_\chi) \ ,$$

implying

$$\delta L^{(0)} = \exp(\mathcal{L}_{\chi^{(M)}}(\ldots(\exp(\mathcal{L}_{\chi^{(1)}}(\delta L))))) \ .$$

Since we are dealing with *near-identity* canonical transformations, we realize that $\delta L^{(0)}$ is the sum of δL and short period (small) variations which do not affect its stability.

If we consider the full Hamiltonian in (4.4.2), then L is not constant anymore because of the dependence of $\mathcal{R}_{J_2}^{(M)}$ on λ. Using again Hamilton's equations, we see that

$$\frac{d}{dt}L = \frac{d}{dt}(\delta L + L_*) = \frac{d}{dt}\delta L = -\frac{\partial \mathcal{H}_{J_2}}{\partial \lambda} = -\frac{\partial \mathcal{R}_{J_2}^{(M)}}{\partial \lambda} \ .$$

Then, for every set of values, say $(e^*, i^*, \lambda^*, p^*, q^*) \in \mathcal{D} \times \mathbb{T}^3$, we obtain

$$\left|\frac{d}{dt}L(e^*, i^*, \lambda^*, p^*, q^*)\right| \leq \sup_{\substack{(e,i) \in \mathcal{D} \\ (\lambda, p, q) \in \mathbb{T}^3}} \left|\frac{d}{dt}L(e, i, \lambda, p, q)\right| \leq \left\|\frac{\partial \mathcal{R}_{J_2}^{(M)}}{\partial \lambda}\right\|_{\infty, \mathcal{D}_*} \ .$$

Let $L(e, i, \lambda, p, q; T)$ be the value at time $t = T$. To estimate its distance from the equilibrium point L_*, we can use the mean value theorem which gives

$$|L(e, i, \lambda, p, q; T) - L_*| \leq \|L(e, i, \lambda, p, q; T) - L_*\|_{\infty, \mathcal{D}_*} \leq \left\|\frac{dL}{dt}\right\|_{\infty, \mathcal{D}_*} T \ . \qquad (4.4.8)$$

Requiring that the right hand side of (4.4.8) is of order of unity, then the stability time T becomes order of $O\left(1/\|dL/dt\|_{\infty, \mathcal{D}_*}\right)$. Let us fix a constant value ΔL and suppose that we want to estimate the minimal time T_1 up to which the variation of $L(e, i, \lambda, p, q; T)$ stays bounded by ΔL:

$$\|L(e, i, \lambda, p, q; T) - L_*\|_{\infty, \mathcal{D}_*} \leq \Delta L \ .$$

Using (4.4.8) we obtain that T_1 is given by

$$T_1 \geq \frac{\Delta L}{\|dL/dt\|_{\infty, \mathcal{D}_*}} \ . \qquad (4.4.9)$$

Equation (4.4.9) can be used to derive the stability time of the semimajor axis a: recalling that, in general, $L = \sqrt{\mu a}$, one has that $\Delta L = \Delta a / 2\sqrt{\mu / a}$, which allows to obtain a lower bound for the stability time of a given by

$$T_2 = \frac{1}{2}\sqrt{\frac{\mu}{a_*}} \frac{\Delta a}{\|dL/dt\|_{\infty, \mathcal{D}_*}}. \qquad (4.4.10)$$

This estimate will be used in Section 4.5.3 to obtain results on the stability time at different altitudes; in particular, Δa is set to be equal to 0.1 R_E.

To check that the norm $\|\mathcal{R}_{J_2}^{(M)}\|_{\infty, \mathcal{D}_*}$ is *small* in the domain $\mathcal{D} = [0, 0.15] \times [0, \pi/2]$, we compute its value by taking a set of samples for the reference value of the semimajor axis a_*, that correspond to different distances from the Earth's center (the radius of the Earth is $R_E = 6378.14$ km). Precisely, we consider the following semimajor axes:

- $a_*^{(1)} = (42164 \ km)/R_E$: the reference value for GEO satellites;

- $a_*^{(2)} = (26560 \ km)/R_E$: the reference value for GPS satellites;

- $a_*^{(3)} = (8524.75 \ km)/R_E$: an intermediate value in terms of the altitude;

- $a_*^{(4)} = (7258.69 \ km)/R_E$: very close to the Earth's surface. We remark that in this case the results obtained are not very relevant from a practical point of view, because the effect of the atmosphere becomes important.

Table 4.1 shows the values of $\|\mathcal{R}_{J_2}^{(M)}\|_{\infty, \mathcal{D}_*}$ computed for the above values of a_* and for $J_2 = 1.084 \cdot 10^{-3}$, namely the real value of the coefficient for the Earth. As we can see, $\|\mathcal{R}_{J_2}^{(M)}\|_{\infty, \mathcal{D}_*}$ is typically very small for all values of a_*: this confirms that for the J_2 problem it is reasonable to take the domain in eccentricity and inclination as $\mathcal{D} = [0, 0.15] \times [0, \pi/2]$.

Table 4.2 provides the results for the estimate of $\|dL/dt\|_{\infty, \mathcal{D}_*}$, which show that, using (4.4.9) with ΔL equal for all the considered distances a_*, the stability time for L increases with the altitude. Figure 4.1 shows the logarithmic plot of $\|\mathcal{R}_{J_2}^{(M)}\|_{\infty, \mathcal{D}_*}$ in

Table 4.1 Estimates of $\|\mathcal{R}_{J_2}^{(M)}\|_{\infty,\mathcal{D}_*}$ for different values of a_* in the J_2 model.

Semimajor axis		a_*	$\|\mathcal{R}_{J_2}^{(M)}\|_{\infty,\mathcal{D}_*}$
42164	km	6.6107	$2.03365 \cdot 10^{-10}$
26560	km	4.16422	$1.78552 \cdot 10^{-9}$
8524.75	km	1.33656	$5.99718 \cdot 10^{-7}$
7258.69	km	1.13806	$1.45442 \cdot 10^{-6}$

Table 4.2 Estimates of $\|dL/dt\|_{\infty,\mathcal{D}_*}$ for different values of a_* in the J_2 model.

Semimajor axis		a_*	$\|dL/dt\|_{\infty,\mathcal{D}_*}$
42164	km	6.6107	$1.31051 \cdot 10^{-9}$
26560	km	4.16422	$1.04198 \cdot 10^{-8}$
8524.75	km	1.33656	$3.10705 \cdot 10^{-6}$
7258.69	km	1.13806	$7.41578 \cdot 10^{-6}$
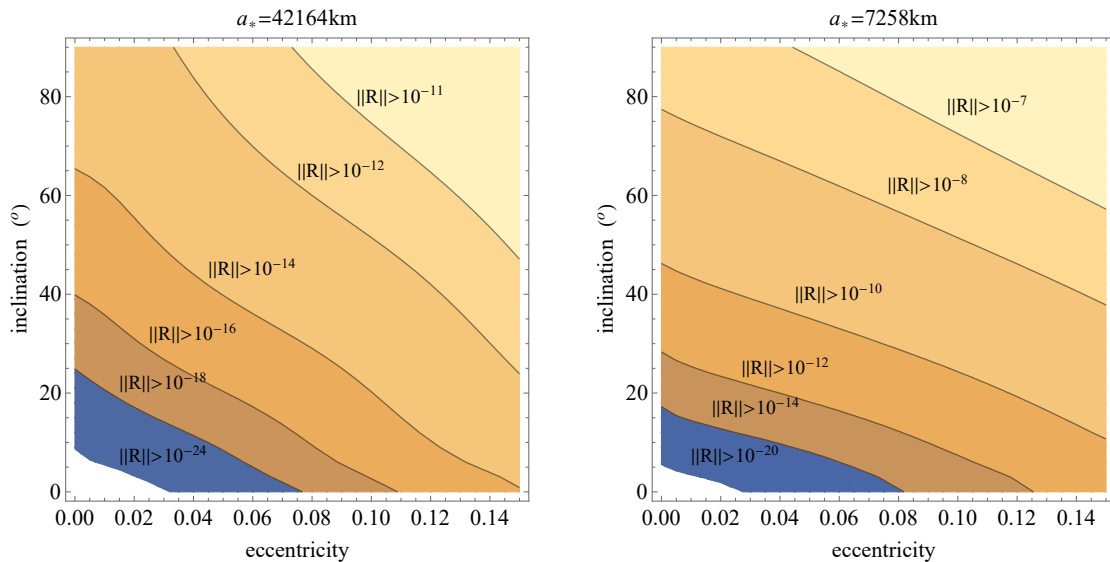
Fig. 4.1 Plots of $\|\mathcal{R}_{J_2}^{(M)}\|_{\infty,\mathcal{D}_*}$ for $(e,i) \in \mathcal{D}$: $a_* = a_*^{(1)}$ (left) and $a_* = a_*^{(4)}$ (right) in the J_2 model.
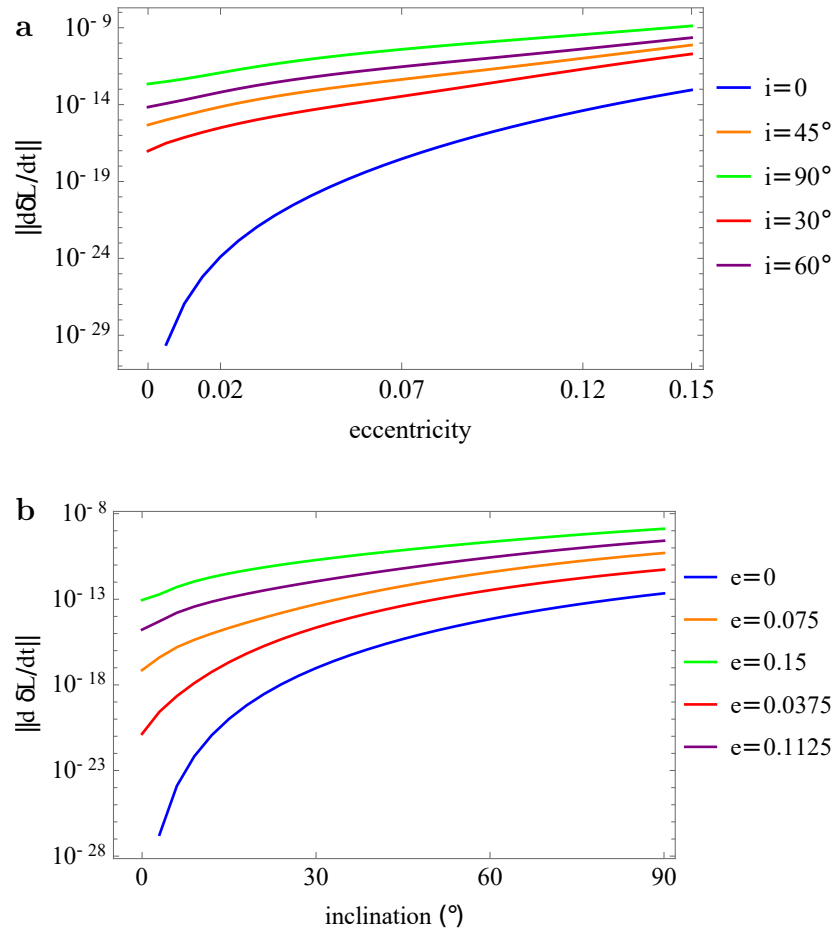
Fig. 4.2 Stability estimates for the GEO case (with $a = a_*^{(1)}$) in the J_2 model. **(a)** plot of $||dL/dt||_{\infty,\mathcal{D}_*}$ as a function of e for fixed values of i. **(b)** plot of $||dL/dt||_{\infty,\mathcal{D}_*}$ as a function of i for fixed values of e.
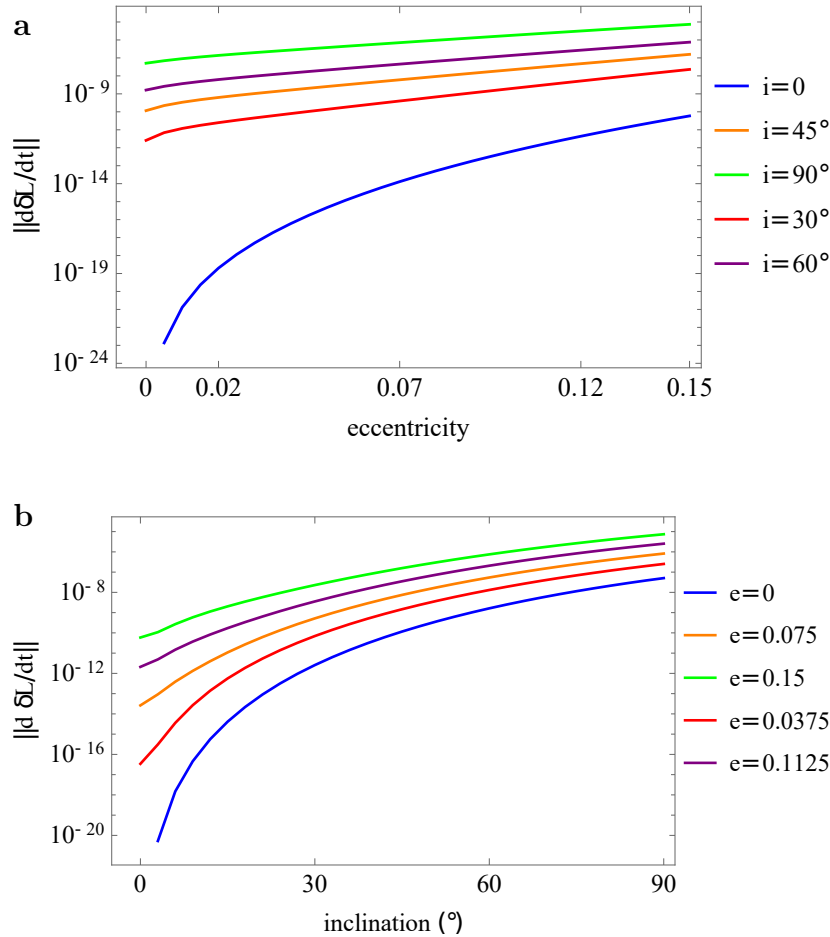
Fig. 4.3 Stability estimates for the near-Earth case (with $a = a_*^{(4)}$) in the J_2 model. **(a)** plot of $||dL/dt||_{\infty,\mathcal{D}_*}$ as a function of e for fixed values of i. **(b)** plot of $||dL/dt||_{\infty,\mathcal{D}_*}$ as a function of i for fixed values of e.

the limit cases $a_* = a_*^{(1)}$ and $a_* = a_*^{(4)}$. The plots show that the remainder decreases as one gets farther from the Earth and it becomes larger when increasing the eccentricity and inclination.

Figures 4.2 and 4.3 refer, respectively, to $a = a_*^{(1)}$ and $a = a_*^{(4)}$; the left plots provide the graph of $||dL/dt||_{\infty,\mathcal{D}_*}$ as a function of the eccentricity for fixed values of the inclination, while the right plots give the norm as a function of the inclination for fixed values of the eccentricity. We notice that the norms tend to decrease when the eccentricity and the inclination are smaller, although the effect is more evident in the GEO region than closer to the Earth.

We now examine how the stability time changes as a function of the semimajor axis a_*: in this case, we consider 1000 values for a_* uniformly distributed from $a_{in} = 1.15679$ (corresponding to an altitude of 1000 km, which we take as the first reference value, although in this region weak dissipative effects are possibly affecting the dynamics) to $a_f = 16.6786$ (corresponding to an altitude of $10^5\ km$), using Eq. (4.4.10) with $\Delta a = 0.1\ R_E$. Figure 4.4 confirms that the stability time increases with the altitude also in the case of the semimajor axis. Indeed, while for $a_* = a_{in}$ we can ensure the stability of the semimajor axis for a period of the order of one year, in the case $a_* = a_f$ we have a stability time of the order of 10^3 years. From an analytical point of view, this

behaviour of the stability time can be explained by the fact that, for higher distances, our model can be approximated by Kepler's problem in which the semimajor axis is constant.
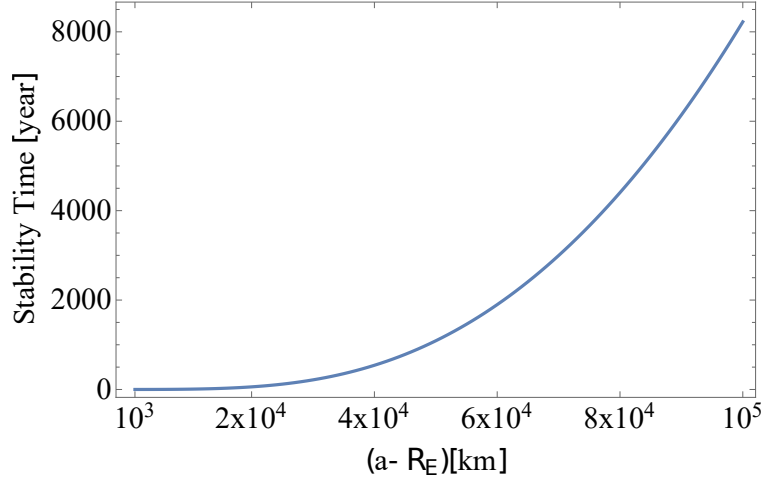

Fig. 4.4 Stability time in the J_2 model for $a \in [a_{in}, a_f]$ (see the text for the definition of a_{in}, a_f) allowing a variation of 0.1 R_E.

4.5   Secular stability in the geolunisolar model

Using the J_2 model, we have demonstrated how the stability of the semimajor axis can be established against *short-period* perturbations (depending on the satellite's mean anomaly). In this Section, we focus, instead, on the long-term variations in the eccentricity and inclination of the satellite's orbit, for orbits close to circular ($e < 0.1$ *rad*) and with small inclination ($|i| < 0.1$). One can easily verify that, within the geolunisolar problem (Hamiltonian $\mathcal{H}_{gls,sec}^{\leq N}$, see Eq. (4.2.24)), the phase-space manifold $e = 0$, corresponding to $I_2 = 0$, constitutes an invariant manifold of the flow, implying that circular orbits remain so for infinitely long times independently of their variations in inclination and longitude of the node. On the other hand, for e small, but non-zero, long-term variations of both the eccentricity and inclination can occur on timescales given by the inverse of the frequencies ν_1 and ν_2 (Eq. (4.2.22)). Since $\nu_1 \simeq \nu_2 \simeq \dfrac{3}{2} \dfrac{\sqrt{\mathcal{G}M_E} R_E^2 J_2}{a^{7/2}}$, the secular timescale is of order of $T_{sec} = \mathcal{O}\left((a/R_E)^2 J_2^{-1}\right) T_{short}$, where T_{short} is the characteristic time of the frequency associated to the fast angle. Since $J_2 \approx 10^{-3}$, the short and long periods are separated by three orders of magnitudes, a fact which justifies altogether the simple averaging over mean anomalies which leads to the model of departure $\mathcal{H}_{gls,sec}^{\leq N}$ for the analysis of the secular stability. On the other hand, the fact that $\nu_1 \simeq \nu_2$ implies that, near the equator (or, more precisely, for orbits near the Laplace plane, see Section 4.2.2), the eccentricity and inclination have coupled variations (the so-called 'Kozai-Lidov' mechanism). Thus, close to the Laplace plane, the term 'secular stability' cannot mean the long-term preservation of the eccentricity and inclination one independently of the other, but only the approximate preservation of the combination $\mathcal{I} \approx e^2 + i^2$ (see below for exact expressions) known as the Kozai-Lidov

integral. The normal form construction and remainder estimates in the present section reflect these basic properties of the dynamics.

4.5.1 Normal form

Starting with the model $\mathcal{H}_{gls,sec}^{\leq N}$ given in Eq. (4.2.24), the construction of the normal form proceeds with the algorithm described in Section 4.3 and the following settings:

i) The book-keeping rule (exponent s in Eq. (4.3.15)) is set as $s = s_1 + s_2 - 2$, where s_1 and s_2 are the exponents appearing in Eq. (4.2.24).
ii) The resonant module (Eq. (4.3.7), case 3 of Subsection 3.1) is set as:

$$\mathcal{M} := \{(k_1, k_2) \in \mathbb{Z}^2 : k_1 + k_2 = 0\}$$

where k_1, k_2 are the integers specifying each Fourier harmonic in Eq. (4.2.24).
iii) The maximum truncation order is set to $N = 15$, while the maximum normalization order is set to $M = 12$.

Here, as well, we use a symbolic program to perform all normalizations, which works in essentially the same way as described in Subsection 4.1 for the case of the normal form of the J_2 problem.

With the following settings, the Hamiltonian after r normalization steps, where r can take the values $r = 1, 2, ... M$, resumes the form:

$$\mathcal{H}_{gls,sec}^{(r)}(I_1, I_2, \phi_1, \phi_2) = \mathcal{Z}_{gls,sec}^{(r)}(I_1, I_2) + \mathcal{Z}_{gls,res}^{(r)}(I_1, I_2, \phi_1 - \phi_2) + \mathcal{R}_{gls}^{(r)}(I_1, I_2, \phi_1, \phi_2) \quad.$$
$$(4.5.1)$$

The term $\mathcal{Z}_{gls,sec}^{(r)}(I_1, I_2)$, hereafter called the *secular part*, contains all terms independent of the angles (corresponding to the choice $k_1 = k_2 = 0$ in the resonant module). The dynamics of this term implies separate preservation of the eccentricity and inclination (the latter around the Laplace plane). Instead, $\mathcal{Z}_{gls,res}^{(r)}(I_1, I_2, \phi_1 - \phi_2)$, called the *resonant part* of the normal form, collects all normal form terms depending on the resonant angle $\phi_1 - \phi_2$. Finally, $\mathcal{R}_{gls}^{(r)}(I_1, I_2, \phi_1, \phi_2)$ is the *remainder term*, which contains non-normalized terms of book-keeping orders $s = r + 1, \ldots, N$. After M normalization steps, we obtain the final geolunisolar Hamiltonian $\mathcal{H}_{gls} \equiv \mathcal{H}_{gls,sec}^{(M)}$.

We now look at the dynamics induced by the sum of secular and resonant parts:

$$\mathcal{H}_{norm}(I_1, I_2, \phi_1, \phi_2) = \mathcal{Z}_{gls,sec}^{(M)}(I_1, I_2) + \mathcal{Z}_{gls,res}^{(M)}(I_1, I_2, \phi_1 - \phi_2),$$

called, altogether, the *resonant normal form* $\mathcal{H}_{norm}$ (for simplicity, we drop the dependence on the normalization order r from the notation). The quantity $I_1 + I_2$ is a first integral for the dynamics induced by $\mathcal{H}_{norm}$, which implies that the vertical component of the angular momentum, which coincides with Θ, is preserved[2]. Given that L is

---

[2]For the J_2 model the preservation of the vertical component of the angular momentum is a direct consequence of the axisimmetry of the truncated geopotential. For the geolunisolar model, the addition of the external attractions breaks this symmetry. However, the preservation of this quantity turns to be still true for the Hamiltonian $\mathcal{H}_{norm}$.

constant, say $L = L_* = \sqrt{\mu_E a_*}$, the quantity

$$I_1 + I_2 = L_* - L_*\sqrt{1 - e^2}(1 - \cos i) \tag{4.5.2}$$

is a first integral and, as a consequence, the quantity

$$\mathcal{I}(e, i) = 1 - \sqrt{1 - e^2}(1 - \cos i) \tag{4.5.3}$$

is constant for the dynamics induced by the normal form. This means that e and i can change only in such a way that the value of $\mathcal{I}(e, i)$ remains constant.

The fact that the presence of resonant first integrals determines a *locking* in the values of e and i is at the basis of the so-called *Kozai-Lidov effect* ([61, 130]), which is common, in a wide range of resonant combinations, in many models of Celestial Mechanics.

4.5.2 Remainder and stability estimates

As already mentioned in Section 4.4.2, we need to guarantee that the remainder is *small* with respect to the normal part; we denote again by $\mathcal{D}$ the domain over which the norm $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}_*}$ is computed. For a function $f = f(e, i, \phi_1, \phi_2)$ of the form:

$$f(e, i, \phi_1, \phi_2) = \sum_{k_1,k_2} \sum_{s_1,s_2} f_{k_1,k_2,s_1,s_2} I_1(e, i)^{s_1/2} I_2(e, i)^{s_2/2}, \tag{4.5.4}$$

where the sums are over a finite number of terms, we have

$$\sup_{(e,i)\in\mathcal{D},(\phi_1,\phi_2)\in\mathbb{T}^2} |f(e, i, \phi_1, \phi_2) \leq \|f\|_{\infty,\mathcal{D}_*},$$

where, recalling the definition in (4.4.6), the norm of f is defined as

$$\|f\|_{\infty,\mathcal{D}_*} = \sum_{k_1,k_2} \sum_{s_1,s_2} |f_{k_1,k_2,s_1,s_2}| I_1(e_*, i_*)^{s_1/2} I_2(e_*, i_*)^{s_2/2}. \tag{4.5.5}$$

There exists an optimal value of M that minimizes the estimate of the remainder's norm, as shown in Section 4.5.3 for GEO orbits.

Since $I_1 + I_2$ is a first integral for $\mathcal{H}_{norm}$, we have that

$$\{I_1 + I_2, \mathcal{H}_{norm}\} = 0.$$

To evaluate the stability of $\mathcal{I}(e, i)$, we use the relation:

$$\frac{d}{dt}(I_1 + I_2) = \{I_1 + I_2, \mathcal{H}_{gls}\} = \{I_1 + I_2, \mathcal{R}_{gls}^{(M)}\};$$

then, for every $(e^*, i^*, \phi_1^*, \phi_2^*) \in \mathcal{D} \times \mathbb{T}^2$, we have the following estimate:

$$\left|\frac{d}{dt}(I_1 + I_2)(e^*, i^*, \phi_1^*, \phi_2^*)\right| \leq \sup_{\substack{(e,i)\in\mathcal{D} \\ (\phi_1,\phi_2)\in\mathbb{T}^2}} \left|\frac{d}{dt}(I_1 + I_2)(e, i, \phi_1, \phi_2)\right| \leq \|\{I_1 + I_2, \mathcal{R}_{gls}^{(M)}\}\|_{\infty,\mathcal{D}_*}.$$

Let us now consider an orbit with initial point $(I_{1,0}, I_{2,0})$ such that the corresponding eccentricity and inclination belong to $\mathcal{D}$; consider its evolution up to $t = T$. Using the mean value theorem, we have that

$$\|(I_1(T) + I_2(T)) - (I_{1,0} + I_{2,0})\| \leq \|\{I_1 + I_2, \mathcal{R}_{gls}^{(M)}\}\|_{\infty, \mathcal{D}_*} \, T. \tag{4.5.6}$$

Setting Γ to be the maximum value for the variation of $I_1 + I_2$ in time, let us denote by $\widetilde{T}$ the minimum time such that for every $T \leq \widetilde{T}$

$$\|(I_1(T) + I_2(T)) - (I_{1,0} + I_{2,0})\| \leq \Gamma \ .$$

From (4.5.6), we have

$$\widetilde{T} \geq \frac{\Gamma}{\|\{I_1 + I_2, \mathcal{R}_{gls}^{(M)}\}\|_{\infty, \mathcal{D}_*}};$$

then we can use the value of T as $T = \Gamma / \|\{I_1 + I_2, \mathcal{R}_{gls}^{(M)}\}\|_{\infty, \mathcal{D}_*}$, which gives an estimate for the stability time of $I_1 + I_2$ and, consequently, of $\mathcal{I}(e, i)$. The stability results for the quantity $\mathcal{I}$ can be translated in terms of the orbital elements (e, i) as follows: in view of (4.5.3), for small values of e and i we find

$$\mathcal{I} \simeq L_* \frac{e^2 + i^2}{2}, \tag{4.5.7}$$

hence, if we consider the variations of $\mathcal{I}$, e and i, they are connected by the relation

$$\frac{\Delta \mathcal{I}}{\mathcal{I}} \simeq 2 \frac{e \Delta e + i \Delta i}{e^2 + i^2}. \tag{4.5.8}$$

For the limit case of e or i fixed and small, one finds

$$\frac{\Delta \mathcal{I}}{\mathcal{I}} \simeq 2 \frac{\Delta e}{e} \simeq 2 \frac{\Delta i}{i}; \tag{4.5.9}$$

then the relative variation of $\mathcal{I}$ (and, as a consequence, of $I_1 + I_2$) is proportional to the relative variations of the orbital elements by a factor 2.

To make the stability results for the geolunisolar model consistent with the ones obtained in Section 4.4.2 for the J_2 model, in Section 4.5.3 we set

$$\Gamma = 0.05 \sqrt{\frac{\mu}{a_*}}, \tag{4.5.10}$$

namely, recalling that $\Delta L = \Delta a / 2 \sqrt{\mu / a}$ and $\Delta a = 0.1$, the maximal variation of $I_1 + I_2$ in the geolunisolar model is equal to the maximal variation allowed for the action L in the J_2 model.

4.5.3 Numerical results for the geolunisolar model

For the geolunisolar model, we take the domain $(e, i) \in \mathcal{D} = [0, 0.1] \times [0, 0.1]$ around the forced eccentricity (which is always zero) and the forced inclination (which depends on the chosen altitude).

Table 4.3 Estimate of $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}_*}$, with $M = 12$, in the geolunisolar model with $\mathcal{D} = [0,0.1] \times [0,0.1]$ for different altitudes.

Altitude	$\|R^{(M)}\|_{\infty,\mathcal{D}_*}$
3000 km	$2.13276 \cdot 10^{-14}$
20000 km	$3.22704 \cdot 10^{-13}$
35790 km	$1.3774 \cdot 10^{-11}$
50000 km	$5.39185 \cdot 10^{-9}$
100000 km	$1.80878 \cdot 10^{-5}$

Since the stability results strongly depend on the distance from the Earth, we select five different altitudes, that correspond to cases of interest for the satellite's problem:

- $h^{(1)} = 3000\ km$, above the atmosphere;

- $h^{(2)} = 20000\ km$, that is in MEO region;

- $h^{(3)} = 35786\ km$, the altitude of GEO orbits;

- $h^{(4)} = 50000\ km$, corresponding to far objects;

- $h^{(5)} = 100000\ km$, that corresponds to objects which are very far from the Earth's surface.

The value of the remainder's norm depends on the altitude of the orbit: in particular, we can state that the stability time decreases as the altitude increases.

Table 4.3 provides the value of $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}_*}$ as a function of the altitude, showing a significant worsening for altitudes after the GEO region.

Figure 4.5 shows the behaviour of the remainder's norm as a function of (e,i) in the bigger domain $\mathcal{D}' = [0,0.1] \times [0,\pi/2]$: as we can see, in all cases the domain $\mathcal{D}'$ is too large to ensure the smallness of $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}'_*}$. Moreover, the magnitude of $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}'_*}$ increases significantly with the altitude. We can easily notice that the value of $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}'_*}$ is strongly dependent on the inclination: using this fact, we can detect a value of i, denoted by i_{crit}, which is the minimum value of the inclination for which $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}'_*}$ is of the order of unity. Table 4.4 shows the computed values of i_{crit} (converted in degrees) for the considered altitudes: we can notice that the *smallness* domain shrinks substantially between 50000 km and 100000 km; in any case, we can see that for every value of the considered altitudes the domain $\mathcal{D} = [0,0.1] \times [0,0.1]$ is contained in the smallness domain of $\mathcal{R}_{gls}^{(M)}$.

As mentioned in Section 4.5, the remainder's norm depends on the normalization order M. Although the norm does not converge to zero if M tends to infinity, there is a value of M, called the *optimal normalization order*, say M_{opt}, for which the norm of the remainder is minimal. Typically, this optimal value is greater than the order of the Taylor expansions of the numerically computed functions, and the estimates for the remainder is so good that there is no reason to push further the order of the expansion;
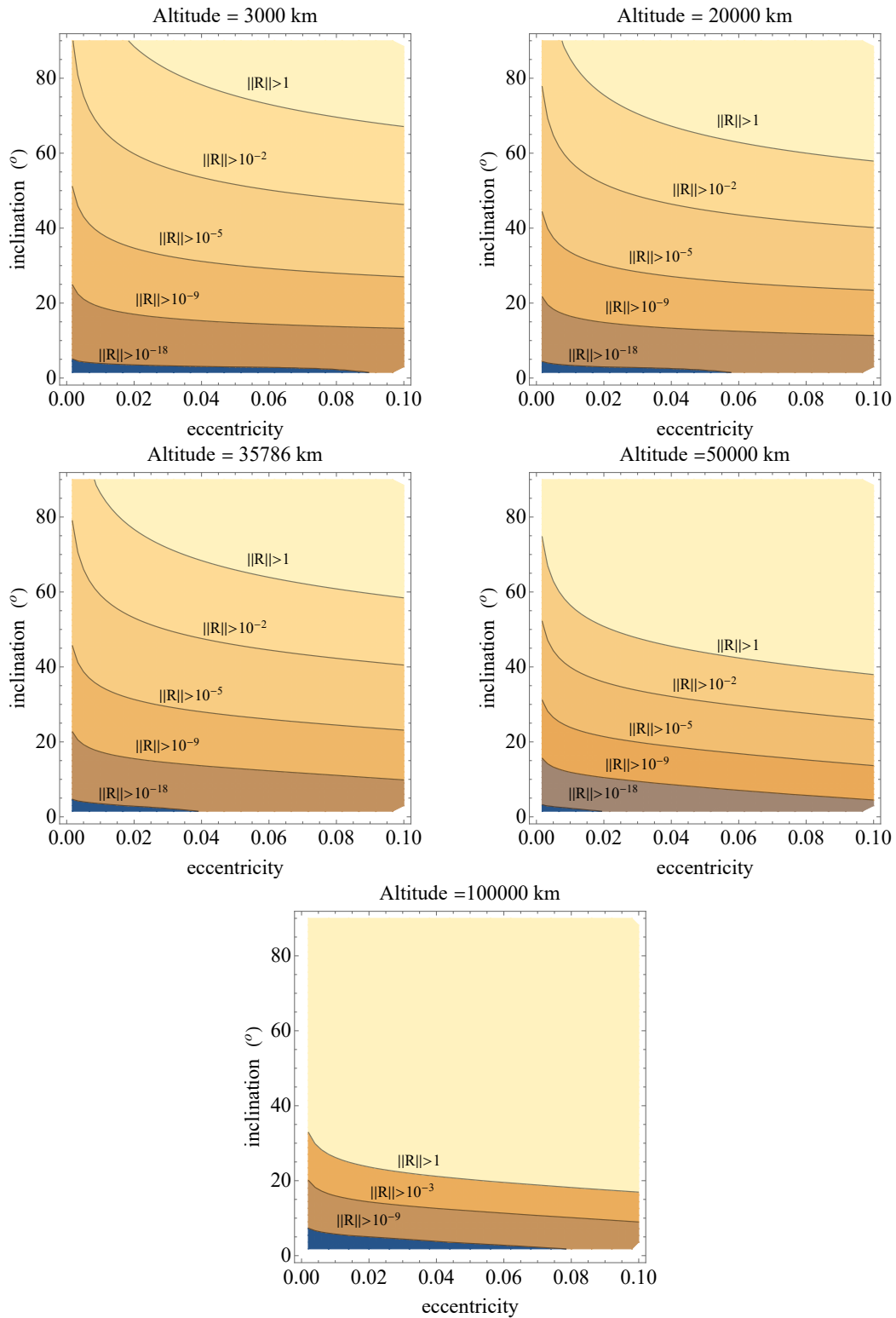
Fig. 4.5 Remainder's norm for the geolunisolar model in the domain $\mathcal{D}' = [0, 0.1] \times [0, \pi/2]$ for $h^{(i)}$, $i = 1, \ldots, 5$ (see the text for the definition of $h^{(i)}$).

Table 4.4 Value of i_{crit} as a function of different values of the altitude for the geolunisolar model.

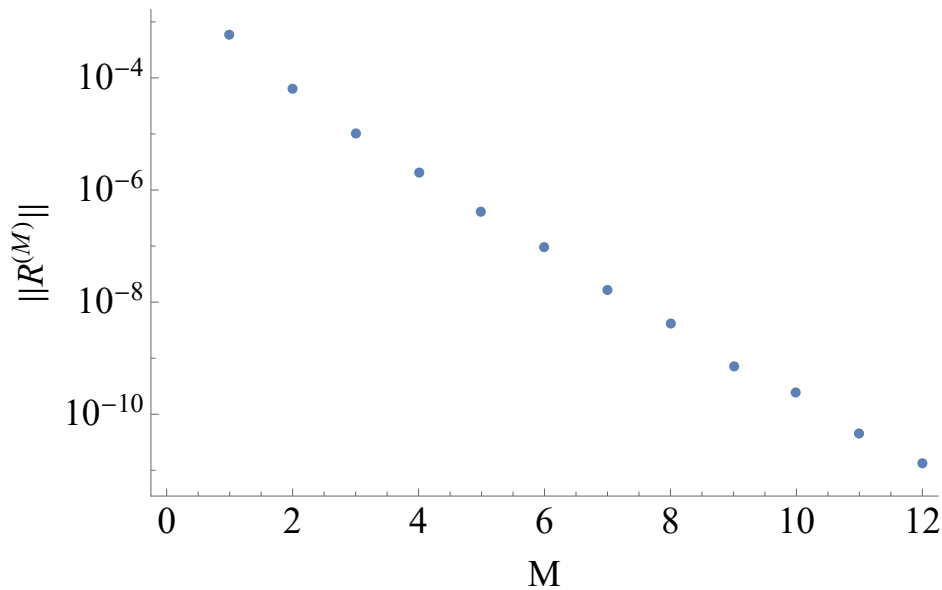| Altitude	$i_{crit}(deg)$
3000 km	67.1°
20000 km	57.89°
35790 km	58.43°
50000 km	37.95°
100000 km	16.93°

Fig. 4.6 Estimate of $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}_*}$ as a function of the normalization order M for $h = h^{(3)}$ (GEO distance) in the domain $\mathcal{D} = [0, 0.1] \times [0, 0.1]$ for the geolunisolar model.

for example, this is the case for the normalized Hamiltonian function which describes the geolunisolar problem computed for the GEO altitude.

As we can see from Figure 4.6, the optimal normalization order is greater than or equal to 12, that is the order at which we make our estimates.

Once obtained the smallness of $\|\mathcal{R}_{gls}^{(M)}\|_{\infty,\mathcal{D}_*}$ in $\mathcal{D}$, we proceed to compute the stability time for the quantity $\mathcal{I}(e,i) = 1 - \sqrt{1 - e^2}(1 - \cos i)$.

As we can see from Table 4.5, the stability times are extremely long: this fact depends on the model we considered, with the Lunar orbit in the ecliptic plane without precession effects. However, we can notice a relevant decrease in the stability time for distances greater than GEO. This behaviour is opposite to that of the J_2 model where the stability time was increasing with the altitude (see Figure 4.4). In fact, at low altitudes the J_2 model is strongly affected by the Keplerian part and the geopotential, while the geolunisolar model takes into account both the inner effect due to the Earth and the outer effect due to Moon and Sun.

As a final remark, to show the importance of taking the right domain in eccentricity and inclination, let us assume $h = h^{(5)}$ and consider the domain $(e, i) \in B = [0, 0.1] \times$

Table 4.5 Stability times in years for different altitudes in the domain $(e, i) \in \mathcal{D} = [0, 0.1] \times [0, 0.1]$ for the geolunisolar model.

Altitude	Stability time in $\mathcal{D}$
3000 km	$4.61551 \cdot 10^{13}$
20000 km	$2.20144 \cdot 10^{12}$
35790 km	$3.51266 \cdot 10^{10}$
50000 km	$1.07263 \cdot 10^{8}$
100000 km	$3.36609 \cdot 10^{4}$

$[0, 0.5]$, which is larger than the convergence domain $[0, 0.1] \times [0, i_{crit}]$ (see Table 4.4). If we compute the stability time in the enlarged domain B, we obtain just the value $T = 0.00085$ years.

4.6 Non-degeneracy conditions

In the previous sections, we examined the question of the long-term stability of the elements (a, e, i) in the case of the Earth's satellite orbits using a semi-analytical computation based on the size of the remainder of the Birkhoff normal form, computed as described in Sections 4.4 and 4.5. While providing stability times quite long with respect to any application of practical interests, such estimates cannot probe the question of the dependence of the optimal remainder on the small parameters of the problem (the value of J_2, as well as the values of (e, i) for non-resonant satellite orbits). Also, it was stressed before that we have no guarantee of the optimality of the estimates themselves with respect to the normalization order, which, in theory, should scale as a power of the inverse of the small parameters of the problem (see [51] for a tutorial introduction).

All such scalings can be examined, instead, in the framework of the outstanding theorem developed by Nekhoroshev ([53]). Under suitable assumptions, the theorem gives a confinement of the action variables for exponentially long times. In particular, the Hamiltonian must satisfy a non-degeneracy condition which, in the original formulation, is called *steepness* condition. The definition of the steepness condition is quite technical and typically not trivial to verify for a specific Hamiltonian system. However, there are some sufficient conditions which imply steepness, whose verification requires the resolution of algebraic equalities and inequalities. This motivates the introduction of the following definition (see [131, 54]).

Definition 4.6.1. *Consider the Hamiltonian $h = h(\underline{J})$ for $\underline{J} \in B$ where $B \subset \mathbb{R}^n$ is an open connected set. Denote by $\underline{\omega}(\underline{J})$ the gradient of h and by $\mathcal{Q}(\underline{J})$ its Hessian matrix. Then:*

1. $h(\underline{J})$ is convex in $\underline{J} \in B$ if

$$\forall \underline{u} \in \mathbb{R}^n \quad \mathcal{Q}(\underline{J})\underline{u} \cdot \underline{u} = 0 \Leftrightarrow \underline{u} = \underline{0};$$

2. $h(\underline{J})$ is quasi-convex in $\underline{J} \in B$ if $\underline{\omega}(\underline{J}) \neq 0$ and

$$\forall \underline{u} \in \mathbb{R}^n \qquad \begin{cases} \omega(\underline{J}) \cdot \underline{u} = 0 \\ \mathcal{Q}(\underline{J})\underline{u} \cdot \underline{u} = 0 \end{cases} \quad \Leftrightarrow \underline{u} = \underline{0};$$

3. $h(\underline{J})$ is three-jet non degenerate in $\underline{J} \in B$ if $\omega(\underline{J}) \neq 0$ and

$$\forall \underline{u} \in \mathbb{R}^n \qquad \begin{cases} \omega(\underline{J}) \cdot \underline{u} = 0 \\ \mathcal{Q}(\underline{J})\underline{u} \cdot \underline{u} = 0 \\ \displaystyle\sum_{i,j,k=1}^{n} \frac{\partial^3 h}{\partial J_i \partial J_j \partial J_k}(\underline{J})u_i u_j u_k = 0 \end{cases} \quad \Leftrightarrow \underline{u} = \underline{0}.$$

We remark that the convexity condition is equivalent to require that the Hessian matrix $\mathcal{Q}(\underline{J})$ is positive (or negative) definite in $\underline{J}$. We add also the following definition of isoenergetically non-degenerate which, for Hamiltonian systems with 2 degrees of freedom, implies quasi-convexity.

Definition 4.6.2. *The Hamiltonian $h = h(\underline{J})$ is called* isoenergetically non degenerate *in $\underline{J} \in B$ with $B \subset \mathbb{R}^n$ open, if*

$$\det \begin{pmatrix} \dfrac{\partial^2 h}{\partial \underline{J}^2}(\underline{J}) & \dfrac{\partial h(\underline{J})}{\partial \underline{J}} \\ \left(\dfrac{\partial h(\underline{J})}{\partial \underline{J}}\right)^T & \underline{0} \end{pmatrix} \neq 0.$$

One can prove (see [132]) that, for every Hamiltonian system with n degrees of freedom, quasi-convexity implies isoenergetically non-degeneracy: as a consequence, for two-dimensional Hamiltonian systems, the two conditions are equivalent.

4.6.1 Numerical verification of the non-degeneracy conditions

We now apply the above definitions to the Hamiltonian functions introduced in Section 4.2. We consider the following cases:

- the Hamiltonian function related to the J_2 problem $\mathcal{H}_{J_2}$, in form of Taylor expansion up to order 15 in eccentricity and inclination, normalized up to order 12 with respect to the fast angle λ; we denote the resulting Hamiltonian including the normalized part $\mathcal{H}_{J_2,sec}^{(M)}$ and the remainder $\mathcal{R}_{J_2}^{(M)}$ (see Eq. (4.4.2)), as

$$\mathcal{H}_{J_2}(\delta L, P, Q, \lambda, p, q) = \mathcal{H}_{J_2,sec}^{(M)}(\delta L, P, Q, p, q) + \mathcal{R}_{J_2}^{(M)}(\delta L, P, Q, \lambda, p, q) \ .$$

  Given the practical stability of the semimajor axis established in Section 4.4.2, in our computations we set $L = L_*$, i.e., $\delta L = 0$;

- the Hamiltonian function related to the geolunisolar problem $\mathcal{H}_{gls,sec}$ in (4.2.14), expanded around the forced values of inclination and eccentricity (see Section 4.2.2) up to order 15 in eccentricity and inclination, see (4.2.24). The Hamiltonian

$\mathcal{H}_{gls,sec}$ is averaged over the fast angle λ and put in resonant normal form with respect to the angles ϕ_1 and ϕ_2 up to order 12 in eccentricity and inclination. As a consequence, the resulting Hamiltonian $\mathcal{H}_{gls} = \mathcal{H}_{gls,sec}^{(M)}$, including the normalized part $\mathcal{Z}_{gls,sec}^{(M)}$, the resonant part $\mathcal{Z}_{gls,res}^{(M)}$ and the remainder $\mathcal{R}_{gls}^{(M)}$ (see Eq. (4.5.1)), has two degrees of freedom and it is the sum of three terms:

$$\mathcal{H}_{gls}(I_1, I_2, \phi_1, \phi_2) = \mathcal{H}_{gls,sec}(I_1, I_2) + \mathcal{H}_{gls,res}(I_1, I_2, \phi_1, \phi_2) + \mathcal{R}_{gls}(I_1, I_2, \phi_1, \phi_2),$$

where $\mathcal{H}_{gls,res}$ depends only on the quasi-resonant combination $\phi_1 - \phi_2$.

To analyze the non-degeneracy conditions, we write the Hamiltonian as the sum of two terms, namely an integrable Hamiltonian h and a perturbing function f. For the J_2-Hamiltonian, we set $h(P, Q)$ to contain all the terms of $\mathcal{H}_{J_2}$ that are independent on all angles, while the perturbing function f contains all other terms. For the geolunisolar case, we choose $h(I_1, I_2)$ to be the angle-independent part of the truncation up to order 2 of $\mathcal{H}_{gls}$: in this way, the Hessian matrix of h is independent of the actions, and the computations are easier[3].

Since the Hamiltonian functions depend on the parameter $L_* = \sqrt{\mu a_*}$, we select four reference values for the altitudes that correspond to distances of interest in satellite dynamics:

- 3000 km, for near-Earth objects;

- 20000 km, for distance of the order of MEO;

- 35790 km, for GEO orbits;

- 50000 km, for far objects.

For each of these values, we check the non-degeneracy conditions of convexity, quasi-convexity and three-jet, for both the case of the J_2-problem and the geolunisolar models in the domain[4] $(e, i) \in \mathcal{D} = [0, 0.1] \times [0, 0.1]$, which corresponds to a domain in the actions $\mathcal{D}'' = [0, P_{max}] \times [0, Q_{max}] \subset \mathbb{R}^2$, where P_{max}, Q_{max} correspond to $e = 0.1$, $i = 0.1$ and can be computed numerically.

Remark 4.6.3. *We notice that a Hamiltonian $h = h(P, Q)$ (or, equivalently, $h(I_1, I_2)$ in the geolunisolar case) is convex in $\mathcal{D}'' \in \mathbb{R}^2$, if the product of the eigenvalues of the Hessian matrix of h is greater than zero for every $(P, Q) \in \mathcal{D}''$. Moreover, $h(P, Q)$ is quasi-convex in $\mathcal{D}'' \in \mathbb{R}^2$, if for every $(P, Q) \in \mathcal{D}''$ the determinant of the matrix*

$$A = \begin{pmatrix} h_{11}(P,Q) & h_{12}(P,Q) & h_1(P,Q) \\ h_{21}(P,Q) & h_{22}(P,Q) & h_2(P,Q) \\ h_1(P,Q) & h_2(P,Q) & 0 \end{pmatrix} \tag{4.6.1}$$

is non zero.

---

[3]We made this particular choice after verifying that, in the chosen domain in the actions, there are no substantial differences between taking all the normalized terms up to order 12 or only the quadratic truncation.

[4]From now on, unless otherwise specified, the angles are expressed in *radians*.

Table 4.6 Values of $\lambda_1 \lambda_2$ for the J_2 model for different altitudes and $(P,Q) \in \mathcal{D}''$.

Altitudes	$\lambda_1 \lambda_2$ intervals
3000 km	$[-1.53606 \cdot 10^{-6}, -1.44844 \cdot 10^{-6}]$
20000 km	$[-3.9009 \cdot 10^{-10}, -3.67893 \cdot 10^{-10}]$
35790 km	$[-3.00586 \cdot 10^{-16}, -3.19895 \cdot 10^{-42}]$
50000 km	$[-4.3889 \cdot 10^{-32}, -1,30545 \cdot 10^{-47}]$

If the convexity and quasi-convexity tests fail, one can control the three-jet non-degeneracy condition, that we compute, again, numerically, checking that the system

$$\begin{cases} \underline{\omega}(P,Q) \cdot \underline{u} = 0 \\ (\partial^2 h(P,Q)\underline{u}) \cdot \underline{u} = 0 \\ ((\partial^3 h(P,Q)\underline{u})\underline{u}) \cdot \underline{u} = 0 \end{cases} \tag{4.6.2}$$

evaluated on a grid of values $(P,Q) \in \mathcal{D}''$ admits only the trivial solution $\underline{u} = (0,0,0)$. Since convexity implies quasi-convexity and quasi-convexity implies three-jet non-degeneracy, to identify which of the conditions is satisfied, we proceed in the following way:

- we begin with the convexity test on the product of the eigenvalues: if the product is positive for every value of $(P,Q) \in \mathcal{D}''$, then $h(P,Q)$ is convex;

- if the convexity test fails, we pass to the quasi-convexity condition, checking the criteria given in Definition 4.6.1 and Remark 4.6.3;

- if the quasi-convexity test fails, we check the three-jet non-degeneracy through the numerical test based on Definition 4.6.1.

4.6.2 Non-degeneracy of the J_2 Hamiltonian

We start from the convexity test; we denote by λ_1, λ_2 the eigenvalues of the Hessian matrix of h.

Table 4.6 gives the numerical values of $\lambda_1\lambda_2$ for different altitudes and (P,Q) in the domain $\mathcal{D}''$ (we recall that, since the values in the Hessian matrix depend on P and Q, we have an interval for $\lambda_1\lambda_2$ instead of a single value). As one can see, the product of the eigenvalues is always negative or zero within numerical precision level, leading to the conclusion that the Hamiltonian $\mathcal{H}_{J_2}$ is not convex in $\mathcal{D}''$ for the considered altitudes.

We can then pass to the quasi-convexity test. We consider the determinant of the matrix A defined in (4.6.1) for $(P,Q) \in \mathcal{D}''$. As we can see from Table 4.7, for every considered altitude the values of det A are equal to zero within the numerical precision level, leading to the conclusion that the J_2 Hamiltonian is not quasi-convex in $\mathcal{D}''$.

Table 4.7 Values of det A, with A as in (4.6.1) for the J_2 model for different altitudes and $(P, Q) \in \mathcal{D}''$.

Altitudes	det A intervals
3000 km	$[-1.99418 \cdot 10^{-10}, -1.82748 \cdot 10^{-10}]$
20000 km	$[-2.27575 \cdot 10^{-15}, -2.08271 \cdot 10^{-15}]$
35790 km	$[-1.30574 \cdot 10^{-17}, -1.19784 \cdot 10^{-17}]$
50000 km	$[-5.34517 \cdot 10^{-19}, -4.9035 \cdot 10^{-19}]$

Table 4.8 Values of λ_1 and λ_2 for different altitudes in the geolunisolar model.

Altitudes	λ_1	λ_2
3000 km	-11.6416	3.046
20000 km	-0.185307	0.0479062
35790 km	-0.0294666	0.00703346
50000 km	-0.0188881	0.00508227

The failure of the quasi-convexity for the J_2 problem is a relevant fact: as we will see in Section 4.6.3, the effects of the lunisolar attraction will eliminate such degeneracy, making the total Hamiltonian quasi-convex.

We conclude with the test on the three-jet non-degeneracy condition. To make the computations quantitative, we solved the system (4.6.2) for values (P_i, Q_j) on a mesh of 10000 points in $\mathcal{D}''$. For every pair of values (P_i, Q_j) the only solution of the system is the trivial one $\underline{u} = (0, 0)$, leading to conclude that the Hamiltonian of the J_2 model is three-jet non degenerate in $\mathcal{D}''$. This fact is not unexpected as for systems up to 3 degrees of freedom the three-jet condition is generically satisfied (see [133]).

4.6.3   Quasi-convexity of the geolunisolar Hamiltonian

As for the J_2 model, we start from the convexity test. In this case, the unperturbed Hamiltonian is a polynomial of degree 2 in the actions; then, the Hessian matrix of $h(I_1, I_2)$ does not depend on the values of I_1 and I_2, and the same holds for its eigenvalues. This makes the test on the convexity of the Hamiltonian easier.

Table 4.8 shows the values of λ_1 and λ_2 for different altitudes. As we can see, in every case the eigenvalues of the Hessian have opposite sign, showing that the geolunisolar unperturbed Hamiltonian is not convex in $\mathbb{R}^2$, and hence in $\mathcal{D}''$.

As for the quasi-convexity, we check whether the matrix A defined in (4.6.1) is nondegenerate for every value $(I_1, I_2) \in \mathcal{D}''$.

From Table 4.9 we can see that the determinant of A is strictly positive for every value of the selected altitudes and every $(I_1, I_2) \in \mathcal{D}''$. Hence, we conclude that the Hamiltonian for the geolunisolar case is quasi-convex. As observed at the end of Section 4.6.2, this fact is highly nontrivial, since it means that the lunisolar perturbation of the J_2 model removes the degeneracy.

Table 4.9 Values of $\det A$, with A in (4.6.1) in the geolunisolar case for different altitudes and $(I_1, I_2) \in \mathcal{D}''$.

Altitudes	$\det A$ intervals
3000 km	$[2206.82, 2335.21]$
20000 km	$[0.0271813, 0.0287145]$
35790 km	$[0.000309172, 0.000323288]$
50000 km	$[0.000113523, 0.000118622]$

Chapter 5

Nekhoroshev estimates for the orbital stability of Earth's satellites

5.1 Introduction

The aim of this chapter is to study the stability of a model for objects in MEO from an analytical point of view, providing exponential stability estimates using the celebrated *Nekhoroshev theorem* ([53]). We stress that, while the Nekhoroshev theorem is particularly relevant for systems with three or more degrees of freedom, which can be affected by the phenomenon known as *Arnold diffusion* ([134]), the applicability of the theorem in securing the long-term stability in *open domains* in the action space holds for systems of any number of degrees of freedom larger than or equal to two. As already pointed out in Section 4.6, the Nekhoroshev theorem was originally developed under a suitable non-degeneracy condition, called *steepness*, while later approaches (e.g. [135, 54]) focus on the important subcase of convex and quasi-convex Hamiltonians. As regards the applications, the theorem was proved useful in obtaining realistic estimates of the domains or times of practical stability of the orbits in a number of interesting problems in Celestial Mechanics. Among others, we mention the three-body problem ([136]) as well as the problem of the Trojan asteroids ([137], [138]).

In this chapter, we apply the Nekhoroshev theorem to a model approximating the (averaged over short period terms) dynamics of a small body around the Earth. As discussed below, this allows to obtain long-time stability estimates for realistic sets of parameters, at least for altitudes (values of the semi major axis) below $20\,000$ km.

Our main model, explained in detail in Section 5.2, is in principle the *geolunisolar model* already introduced in Chapter 4. Let us recall that it is ruled by a Hamiltonian function obtained as the sum of different contributions, namely, the geopotential J_2 term as well as the third-body perturbations on the small body by the Sun and Moon. We assume the spatial case of the small body's motion, while we approximate the Moon's and Sun's orbits as fixed Keplerian ellipses lying in the ecliptic plane. We argue (Section 5.2) that the Moon's precession of the nodes introduces only minimal effects as regards the problem of determining Nekhoroshev stability, due to the fact that the frequency of the precession is much smaller than any of the frequencies in the small body's motion. As a result, our point of departure is a Hamiltonian model obtained by a 3 degrees-of-freedom and time-dependent Hamiltonian function, which

depends quasi-periodically on time, since the (fast) frequencies of motion of the Sun and Moon are non-commensurable.

Now, similarly to what we already observed in in Chapter 4 and as will be discussed in Section 5.2, this model is still not convenient for the discussion of Nekhoroshev stability over secular timescales, because both short and long period effects are included in it. Working, however with closed-form perturbation theory (namely, without series expansions in eccentricity and inclination, see [139, 122] for a review), one can eliminate all short-period terms and arrive at an autonomous Hamiltonian with two degrees of freedom which is convenient for the description of the secular motions of the small body. As our basic model, we then adopt the one found after averaging (in closed-form) over the Earth's J_2 term and the Sun's and Moon's quadrupolar (P_2) terms. Several studies (see [46–48, 140] and references therein) have demonstrated the relevance of this model in capturing all important effects for the long-term dynamics at MEO. On the other hand, unlike Chapter 4, in Section 5.5 we will consider also a more complicated model including the Earth's J_3 and J_4 terms to first order, as well as J_2^2 terms. The latter are computed by implementing a *closed form* averaging through Deprit's elimination of the parallax technique ([141–143]). One finds (see the discussion in Section 5.5) that the relative importance of these terms over lunisolar terms decreases with altitude; yet, these terms provide relevant contributions to the Hamiltonian for the lowermost altitudes considered in our analysis (namely, with semimajor axis $a \approx 11000$ km).

After fixing the initial model, an important aspect of the work presented in this chapter concerns a number of preliminary operations performed on the initial Hamiltonian, which turn to be crucial to the purpose of bringing the Hamiltonian in a form allowing to explicitly demonstrate the fulfilment of the conditions for the holding of the Nekhoroshev theorem in the form given in [54]. These preliminary steps are explained in detail in Section 5.2 below, and can be summarized as follows:

(i) *Average over fast angles.* We start by averaging the Hamiltonian over the problem's fast angles, i.e., the mean anomalies of the small body's, Moon's and Sun's orbits. After this operation, the semi-major axis a of any orbit becomes a constant which can be used to label the altitude of each orbit. We stress that an analogous averaging procedure has been applied in Section 4.2.2.

(ii) *Expansion around reference values in the eccentricity and inclination.* The remaining elements (eccentricity e and inclination i), which can be mapped into the action variables of the problem, undergo 'secular' (slow) evolution under the averaged Hamiltonian. Our purpose is to characterize the stability of the orbits in the space (e, i) of the orbital elements. To this end, fixing a grid of reference values (e_*, i_*) in the plane (e, i) for each (constant) semi-major axis $a = a_*$, we perform a Taylor expansion of the averaged Hamiltonian around the points in action space associated with the reference point (e_*, i_*). This step is important, since the Taylor-expanded Hamiltonian can be easily manipulated in terms of normalizing canonical transformations necessary to perform with the aid of a computer-algebraic program (see below). This step represents a fundamental difference with respect to the expansion used in Chapter 4 to obtain the lunisolar Hamiltonian. As a matter of fact, while here the reference values (e_*, i_*) will range over a wide domain of eccentricities and inclinations, in the case already presented in Section 4.2.2 only *small* values of e and i have been considered.

Table 5.1 Inclination-dependent resonances of order ≤ 4 in the lunisolar model. The coefficients are such that $\alpha\dot{\omega} + \beta\dot{\Omega} = 0$, where ω is the argument of the perigee and Ω is the longitude of the ascending node ([112]).

α	β	$i(deg)$	α	β	$i(deg)$	α	β	$i(deg)$	α	β	$i(deg)$
1	0	63.43	0	1	90	1	1	46.37	1	-1	73.1
2	1	56	1	2	0	2	-1	69	-1	12	78
3	1	58.75	3	-1	67.33	1	-3	81.47			

(iii) *Preliminary normalization.* We perform a preliminary normalization of the averaged and Taylor-expanded Hamiltonian, aiming to eliminate some terms which, albeit reflecting a trivial dynamics (see Section 5.3), may artificially affect the estimates found by implementing Pöschel's version of the Nekhoroshev theorem. We argue below that this step is a consequence of the non-zero value of the inclination of the Laplace plane with respect to the Earth's equatorial plane. The value of such inclination has already been expressed in (4.2.20). A key result of this chapter is the use of normal form techniques to reduce the size, in the Hamiltonian, of all terms related to the Laplace plane (see Section 5.3); whenever convergent, this procedure is crucial to put the initial Hamiltonian in a form for which Nekhoroshev's nonresonant stability estimates can be produced, since it allows us to control the norm of the perturbing function under a suitable choice of the domain in the actions.

Now, following steps (i) to (iii) above, the procedure leads to a normalized 2-dimensional Hamiltonian expressed in suitably defined action-angle variables $(\mathbf{I}, \mathbf{u}) \in \mathbb{R}^2 \times \mathbb{T}^2$, of the form:

$$\mathcal{H}(\mathbf{I}, \mathbf{u}) = h_0(\mathbf{I}) + h_1(\mathbf{I}, \mathbf{u}) \ . \tag{5.1.1}$$

Using the Hamiltonian (5.1.1), we can derive stability results on the eccentricity and the inclination by implementing the estimates provided in [54, Proposition 1]. This proposition refers to the so-called *non-resonant* regime, i.e., when the fundamental frequencies deduced by the integrable part of the Hamiltonian, h_0, are subject to no resonance conditions. Under particular assumptions on the non-resonance condition for h_0, as well as on the smallness of the norm of h_1 in a suitable functional space and domain in the action variables (see Section 5.3.1), one can prove that the actions remain in a small neighborhood of their initial values for a period of time which is exponentially long with respect to the norm of h_1. We remark that Proposition 1 in [54] does not require any convexity assumption on the Hamiltonian. This assumption is relevant when analyzing resonant regimes (a thorough analysis of different non-degeneracy conditions such as convexity, quasi-convexity, 3-jet, etc., has been already presented in Chapter 4). However, we also stress that, despite our use of Pöschel's proposition in the non-resonant regime, the presence of resonances at MEO plays an important role also in our results, as becomes evident in the discussion of our results in Section 5.4. In fact, we find that our obtained stability domains typically exclude some zones around the so-called *inclination-dependent resonances* ([112]), i.e., resonances appearing for particular values of the inclination of the orbit, independently of the value of the semimajor axis or the eccentricity. This is because the series constructed in our preliminary normalization of the Hamiltonian are affected by small divisors related to the most-important of these resonances, given in Table 5.1. Also, the frequencies

associated with these divisors influence the determination of the so-called 'Fourier cut-off' (Section 5.3) which appears in the implementation of the Proposition 1 of [54].

As described in Section 5.4, the stability estimates obtained in this chapter depend strongly on the distance of the small body from the Earth's center: our results show that the domain of Nekhoroshev stability in the plane (e, i) has a large volume (limited only by narrow strips around resonances) at the distance of $10\,000$ km, while it shrinks to a near-zero volume beyond the distance of $20\,000$ km. We should stress that this result is partly due to the dynamics itself (the J_2 dynamics alone is integrable, but the third body perturbations increase in relative size as the distance from the Earth increases), but also probably due, in part, to our particular technique used to apply the Nekhoroshev theorem, i.e. including the processing of the Hamiltonian as described in steps (i)-(iii) above. We thus leave open the possibility that this latter constraint be relaxed with the use of a better technique. Also, our present treatment is simplified in that we ignore the periodic oscillation of the Moon's line of nodes (by an amplitude of $11.5°$ over a period of 18.6 yr) and inclination (by $\pm5°$) around the ecliptic of the Moon's orbit with respect to the Earth's equatorial plane. This oscillation introduces one more secular frequency to the problem; however, it substantially affects the orbits only for semi-major axes $a > 20\,000$ km, which is, anyway, beyond the domain of stability presently found even while ignoring this effect.
The figures of the current chapter are taken from [6].

5.2   Hamiltonian preparation

In this section, we provide details on the model (Section 5.2.1), on the corresponding secular Hamiltonian function averaged over fast angles (Section 5.2.2), the expansion around some reference values for the eccentricity and inclination (Section 5.2.3), and the preliminary normalization to remove specific terms (Section 5.2.4).

5.2.1   Model

Although the model used in this chapter is analogous to the geolunisolar one already introduced in Chapter 4, let us recall, for the readers' convenience, the basic steps of its construction. This will be useful as, starting from Section 5.2.3, the expansion we will consider will be different from the one used previously, as well as the normalization techniques carried on to apply Nekhoroshev theorem.
In particular, we consider the following approximation to the body's equations of motion:

$$\ddot{\mathbf{r}} = -\nabla V_E(\mathbf{r}) - \mu_\odot \left(\frac{\mathbf{r} - \mathbf{r}_\odot}{|\mathbf{r} - \mathbf{r}_\odot|^3} + \frac{\mathbf{r}_\odot}{|\mathbf{r}_\odot|^3} \right) - \mu_M \left(\frac{\mathbf{r} - \mathbf{r}_M}{|\mathbf{r} - \mathbf{r}_M|^3} + \frac{\mathbf{r}_M}{|\mathbf{r}_M|^3} \right) , \qquad (5.2.1)$$

where $V_E(\mathbf{r})$ approximates the geopotential via the relation

$$V_E(\mathbf{r}) = V_{kep}(\|\mathbf{r}\|) + V_{J_2}(\mathbf{r}) , \qquad (5.2.2)$$

where $V_{kep}(r) = -\dfrac{\mu_E}{r}$ and V_{J_2} in spherical coordinates (r, φ, ϕ) is given by

$$V_{J_2}(r, \varphi, \phi) = \frac{\mu_E J_2}{r} \left\{ \left(\frac{R_E}{r} \right)^2 \left(\frac{3}{2} \sin^2 \phi - \frac{1}{2} \right) \right\}{}^1. \qquad (5.2.3)$$

In the above formulas:

- $\mathcal{G}$ is the gravitational constant, $\mu_E = \mathcal{G} m_E$, $\mu_M = \mathcal{G} m_M$, $\mu_\odot = \mathcal{G} m_\odot$ with m_E, m_M, $m_\odot$ the masses of the Earth, Moon and Sun respectively.

- We adopt the value $J_2 = 1.082 \times 10^{-3}$ for the J_2 coefficient, and $R_E = 6\,400$ km for the Earth's equatorial radius.

- $\mathbf{r}$, $\mathbf{r}_\odot$ and $\mathbf{r}_M$ are, respectively, the geocentric position vectors of S, Sun and Moon with respect to an inertial reference frame.

The expressions of $\mathbf{r}_\odot$ and $\mathbf{r}_M$ depend on the assumptions on the orbits of Sun and Moon. Here, the geocentric orbit of the Sun is taken as a fixed ellipse with $a_\odot = 1.496 \times 10^8$ km, $e_\odot = 0.0167$ and $i_\odot = 23.44°$, while the geocentric orbit of the Moon is taken as a fixed ellipse with orbital parameters $a_M = 384\,748$ km, $e_M = 0.0554$ and $i_M = i_\odot$. As already pointed out, the last assumption implies that the only lunisolar resonances which affect the dynamics of the body are those whose location in the element space (a, e, i) depends only on the inclination (see [112]). More resonances, instead, appear when the effect of nodal precession (by a period of 18.6 years) of the Moon's orbit is taken into account. However, these resonances affect the dynamics only at altitudes exceeding the ones where we presently establish Nekhoroshev stability (see [47] and Section 5.4 below), thus they can be again ignored in the framework of our present study.

The Hamiltonian function which describes the motion of S can be expressed as the sum of three contributions:

$$\mathcal{H} = \mathcal{H}_E + \mathcal{H}_\odot + \mathcal{H}_M, \qquad (5.2.4)$$

where $\mathcal{H}_E = \mathbf{p}^2/2 + V_E(\mathbf{r})$ with $\mathbf{p} = \dot{\mathbf{r}}$, and $\mathcal{H}_\odot$ and $\mathcal{H}_M$ are the solar and lunar third-body perturbation terms. Considering the quadrupolar expansion of the third-body perturbation terms in the equations of motion (5.2.1), we find again the expansions in Eq. (4.2.10).

5.2.2 Average over fast angles - Secular Hamiltonian

As already pointed out in Section 4.2.2, the secular motion of the body S can be modeled by computing the average of (5.2.4) over all canonical angles associated to the fast motions of S, the Sun and the Moon. Note that the period of the Sun is only 'semi-fast' (one year, compared to secular periods of $\sim$ 10 yrs for the small body), and more detailed models can consider also the case of 'semi-secular' resonances, i.e., resonances in the case in which the equations of motion (and Hamiltonian) are not averaged with respect to the Sun's mean anomaly (see, for example, [42]).

---

[1]In the present section, as well as in Sections 5.3 and 5.4, we limit our analysis to the J_2-term, which is the dominant term of the Earth's potential at all altitudes; however, in Section 5.5 we will discuss the influence on our results by the terms J_3, J_4, and J_2^2 (obtained through a canonical transformation), which become relevant for the lowermost limit in altitude of the MEO domain.

Averaging with respect to all fast angles leads to the following, called hereafter, *secular Hamiltonian*, given by the sum of the averaged contributions of the Earth, Sun and Moon:

$$\mathcal{H}^{(sec)} = \mathcal{H}_E^{(av)} + \mathcal{H}_\odot^{(av)} + \mathcal{H}_M^{(av)} \ . \tag{5.2.5}$$

The function $\mathcal{H}^{(sec)} = \mathcal{H}^{(sec)}(G, \Theta, \omega, \Omega)$ is a two degrees of freedom Hamiltonian, which can be explicitly computed in terms of Delaunay canonical action-angle variables G, Θ (with conjugated angles ω, Ω), related to the orbital elements by the expressions (see, e.g., [70]):

$$G = \sqrt{\mu_E a(1 - e^2)}, \quad \Theta = \sqrt{\mu_E a(1 - e^2)} \cos i \ . \tag{5.2.6}$$

Since the averaged Hamiltonian does not depend on the mean anomaly M of S, the conjugated Delaunay action $L = \sqrt{\mu_E a}$, and hence the semi-major axis a, is a constant of motion of the Hamiltonian $\mathcal{H}^{(sec)}$. We set $L = L_*$, or, equivalently, $a = a_*$ when referring to trajectories whose semi-major axis has the reference value a_*.

Following the same procedure as in Section 4.2.2, one obtains the geopotential term

$$\mathcal{H}_E^{(av)} = -\frac{\mu_E^2}{2L^2} - J_2 \frac{\mu_E R_E^2}{a_*^3(1 - e^2)^{3/2}} \left(\frac{1}{2} - \frac{3}{4} \sin^2 i \right) \ . \tag{5.2.7}$$

We note that this procedure of *scissor averaging* yields a formula for $\mathcal{H}_E^{(av)}$, which is identical to the formula obtained at first order through a Lie canonical transformation, a procedure known as the *Delaunay normalization* (see [144, 139]). However, from a physical point of view, this implies that in all results described below, by *elements* are implied the ones found after eliminating from the Hamiltonian the short-period terms (see discussion in Section 5.5). The same property holds for the averaging of the terms $\mathcal{H}_\odot^{(av)}$, $\mathcal{H}_M^{(av)}$, which can be performed by a canonical transformation leading, to first order, to the same formula as using the scissor averaging integral

$$\mathcal{H}_\odot^{(av)} = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \left(-\frac{\mu_\odot}{r_\odot} - \frac{\mu_\odot}{2r_\odot^3} r^2 + \frac{3}{2} \frac{\mu_\odot (\mathbf{r} \cdot \mathbf{r}_\odot)^2}{r_\odot^5} \right) dM dM_\odot$$

(and analogously for $\mathcal{H}_M^{(av)}$); we recall that, in this case, it turns convenient to change the integration variables from M to u (eccentric anomaly of S) and from $M_\odot$ to $f_\odot$ (true anomaly of the Sun). We note that, up to quadrupolar terms, this yields the same result as considering the Moon and Sun in circular, instead of elliptic, orbits (in which case $M_\odot$, M_M would be equal to $f_\odot, f_M$), but replacing each third-body's semi-major axis a_b with the expression $a_b \to a_b(1 - e_b^2)^{1/2}$ (index b standing for Sun or Moon). This replacement accomplishes the first step in the Hamiltonian preparation.

5.2.3 Expansion around reference values (e_*, i_*)

After performing the above operations, the Hamiltonian $\mathcal{H}^{(sec)}$ becomes a function of the body's action-angle variables (G, ω), (Θ, Ω), while it depends also on the Delaunay action L, which however, does not affect the secular dynamics and can be carried on all subsequent expressions as a parameter (equal to L_*). We use, alternatively, a_* as the parameter appearing in the coefficients of all trigonometric terms in $\mathcal{H}^{(sec)}$. Furthermore, it turns convenient to express $\mathcal{H}^{(sec)}$ in terms of *modified Delaunay variables* instead

of the original Delaunay variables, as they have already been introduced in (4.2.4) and (4.2.7) (note that, in the following, Γ corresponds to the old P and $\tilde{\Theta}$ to the old Q). Starting now from the Hamiltonian $\mathcal{H}^{(sec)}(\Gamma, \tilde{\Theta}, p, q)$, our goal will be to examine Nekhoroshev stability in a covering of the action space in terms of local neighborhoods around a grid of reference values corresponding to a grid of element values (a_*, e_*, i_*) (see Section 5.3.2). This motivates to introduce the variables P and Q defined by

$$\begin{cases} P = \Gamma_* - \Gamma, \\ Q = \tilde{\Theta}_* - \tilde{\Theta} \ , \end{cases} \tag{5.2.8}$$

where Γ_* and $\tilde{\Theta}_*$ are the values corresponding to the orbital elements (e_*, i_*), and compute the Taylor expansion of $\mathcal{H}^{(sec)}$ in powers of the small quantities (Q, P), truncated at a maximum order N (we set $N = 12$)[2]. We then arrive at the following truncated secular Hamiltonian model

$$\mathcal{H}^{(sec,N)}(P, Q, p, q) = \sum_{j=1}^{N} g^{(j)}(P, Q, p, q) \ . \tag{5.2.9}$$

In the model (5.2.9) we have

$$g^{(1)}(P, Q) = \omega_1 P + \omega_2 Q \ . \tag{5.2.10}$$

For reasons that will become clear later, for $j \geq 2$ we split each of the functions $g^{(j)}(P, Q, p, q)$ as a sum depending only on the actions and a sum depending also on the angles:

$$g^{(j)}(P, Q, p, q) = \sum_{\substack{\mathbf{l} \in \mathbb{Z}^2 \\ |\mathbf{l}| = j}} a_{\mathbf{l}}^{(j)} P^{l_1} Q^{l_2} + \sum_{\substack{\mathbf{l}, \mathbf{k} \in \mathbb{Z}^2 \\ |\mathbf{l}| = j-2}} b_{\mathbf{l}, \mathbf{k}}^{(j)} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)} \ . \tag{5.2.11}$$

Expansion (5.2.11) is the analogous of Eq. (4.2.23) presented in Chapter 4, though the latter has been expressed as a trigonometric polynomial instead of the exponential expansion. Moreover, for reasons related to the particular normalization algorithm we will use in Section 5.2.4, the terms depending only on the action are separated from the time-dependent ones. This last splitting completes the second step in the Hamiltonian preparation. The explicit expressions of the quantities ω_1, ω_2, $a_{\mathbf{l}}$, $b_{\mathbf{l}, \mathbf{k}}$ for $j = 2$ are given in Section 5.7, in terms of the orbital elements of the satellite, Moon and Sun.

5.2.4   Preliminary normalization

It was already mentioned in Section 5.1 that the presence of the averaged lunisolar terms in (5.2.9) implies the existence of a secular equilibrium solution of Hamilton's equations of motion under the Hamiltonian $\mathcal{H}^{(sec)}$, corresponding to the values $e = 0, i = i^{(eq)}$ (see Eq.(4.2.20), where $i^{(eq)} = i^{(p)}$), where $i^{(p)}$ is called the inclination of the *Laplace plane* or *proper inclination*. It is easy to see that the non-zero value of the inclination

---

[2]This is actually the first difference between the geolunisolar Hamiltonian presented in Chapter 4 and the one constructed here. As a matter of fact, while in the current Hamiltonian (i_*, e_*) could be arbitrary, in Chapter 4 the expasion was done around $e = 0$ and the inclination proper of the Laplace plane.

of the Laplace plane is reflected into the Hamiltonian $\mathcal{H}^{(sec,N)}$ by the presence of purely trigonometric terms, i.e., terms with $|\mathbf{l}| = 0$. Such terms yield coefficients which are dominant with respect to the remaining terms in the Hamiltonian expansion. Furthermore, in the splitting of the Hamiltonian as $\mathcal{H} = h_0(\mathbf{I}) + h_1(\mathbf{I}, \mathbf{u})$, where $(\mathbf{I}, \mathbf{u})$ are action-angle variables, as required for the implementation of the Nekhoroshev theorem (see next section), the above terms generate terms with a dominant coefficient largely affecting the size of the perturbation $h_1(\mathbf{I}, \mathbf{u})$. In the present subsection, we implement a procedure for controlling the size of the terms (5.2.9) of the expansion, so that we obtain a Hamiltonian satisfying the norm bounds required for the implementation of the Nekhoroshev theorem.

More specifically, the aim of the normalization algorithm described below is to remove, up a certain order N_{norm} with respect to the expansion (5.2.9), the angle-dependent terms which are constant or linear in the actions: this leads to a Hamiltonian $\mathcal{H}^{(N_{norm})}$, in which the norm of the angle-dependent part decreases at least quadratically with the size of the domain A_{r_0} in which local action variables are defined.

The normalization procedure relies on the use of *Lie series*, already introduced in Section 4.4. In every normalization step, the transformed Hamiltonian is given by

$$\mathcal{H}^{(new)} = \exp^{(N)}(\mathcal{L}_\chi)\mathcal{H}^{(old)}, \tag{5.2.12}$$

where $\mathcal{L}_\chi f = \{f, \chi\}$ ($\{\cdot, \cdot\}$ denotes the *Poisson bracket*) and $\exp^{(N)}(\mathcal{L}_\chi)$ is defined by

$$\exp^{(N)}(\mathcal{L}_\chi)f = \sum_{s=0}^{N} \frac{1}{s!}\mathcal{L}_\chi^s f \ . \tag{5.2.13}$$

To illustrate the procedure, rename the initial Hamiltonian (5.2.9) as $\mathcal{H}^{(0)}$ (where superscripts denote how many normalization steps were performed). Then:

$$\mathcal{H}^{(0)}(P, Q, p, q) = \sum_{j=1}^{N} g^{(j,0)}(P, Q, p, q) \ , \tag{5.2.14}$$

where

$$g^{(1,0)}(P, Q) = \omega_1 P + \omega_2 Q$$

and, when $j \geq 2$,

$$g^{(j,0)}(P, Q, p, q) = \sum_{\substack{\mathbf{l} \in \mathbb{Z}^2 \\ |\mathbf{l}|=j}} a_{\mathbf{l}}^{(j,0)} P^{l_1} Q^{l_2} + \sum_{\substack{\mathbf{l}, \mathbf{k} \in \mathbb{Z}^2 \\ |\mathbf{l}|=j-2}} b_{\mathbf{l},\mathbf{k}}^{(j,0)} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)}. \tag{5.2.15}$$

The second term of the sum (5.2.14) takes the form

$$g^{(2,0)}(P, Q, p, q) = \sum_{\substack{\mathbf{l} \in \mathbb{Z}^2 \\ |\mathbf{l}|=2}} a_{\mathbf{l}}^{(2,0)} P^{l_1} Q^{l_2} + \sum_{\mathbf{k} \in \mathbb{Z}^2} b_{\mathbf{0},\mathbf{k}}^{(2,0)} e^{i(k_1 p + k_2 q)} \ . \tag{5.2.16}$$

The generating function $\chi^{(1)}$ eliminating the above terms has the form

$$\chi^{(1)}(P,Q,p,q) = \sum_{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2} x^{(1)}_{\mathbf{l},\mathbf{k}} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)} \; , \tag{5.2.17}$$

where the coefficients $x^{(1)}_{\mathbf{l},\mathbf{k}}$ are obtained as the solution of the homological equation

$$\{\omega_1 P + \omega_2 Q, \chi^{(1)}\} = -\sum_{\mathbf{k}\in\mathbb{Z}^2} b^{(2,0)}_{\mathbf{0},\mathbf{k}} e^{i(k_1 p + k_2 q)} \; , \tag{5.2.18}$$

namely

$$\chi^{(1)}(p,q) = -\sum_{\mathbf{k}\in\mathbb{Z}^2} \frac{b^{(2,0)}_{\mathbf{0},\mathbf{k}}}{i(\omega_1 k_1 + \omega_2 k_2)} e^{i(k_1 p + k_2 q)} \; . \tag{5.2.19}$$

The normalized Hamiltonian after the first step can be written as

$$\mathcal{H}^{(1)}(P,Q,p,q) = \omega_1 P + \omega_2 Q + Z^{(2,1)}(P,Q,p,q) + \sum_{j=3}^{N} g^{(j,1)}(P,Q,p,q) \; , \tag{5.2.20}$$

where

$$Z^{(2,1)} = g^{(2,0)} + \mathcal{L}_{\chi^{(1)}}(\omega_1 P + \omega_2 Q) = \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=2}} a^{(2,0)}_{\mathbf{l}} P^{l_1} Q^{l_2} \tag{5.2.21}$$

and

$$g^{(j,1)} = \sum_{s=0}^{j-1} \frac{1}{s!} \mathcal{L}^s_{\chi^{(1)}} g^{(j-s,1)}. \tag{5.2.22}$$

In general, since the generating function $\chi^{(1)}$ is constant in the actions, one can see that, if $f(P,Q,p,q)$ has polynomial order ℓ in the actions, then the order in the actions of the transformed function $\mathcal{L}_{\chi^{(1)}} f$ is $\ell - 1$. This means that all terms in $\mathcal{H}^{(1)}$ can be labeled through their polynomial orders in the actions: choosing the expansion order N to be odd and distinguishing the indices j with respect to their parity, we have, for $n = 1, \ldots, (N-1)/2$:

$$g^{(2n,1)}(P,Q,p,q) = \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=2n}} a^{(2n,1)}_{\mathbf{l}} P^{l_1} Q^{l_2} + \sum_{s=0}^{n-1} \sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=2s}} b^{(2n,1)}_{\mathbf{l},\mathbf{k}} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)} \quad (n \geq 2),$$

$$g^{(2n+1,1)}(P,Q,p,q) = \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=2n+1}} a^{(2n+1,1)}_{\mathbf{l}} P^{l_1} Q^{l_2} + \sum_{s=0}^{n-1} \sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=2s+1}} b^{(2n+1,1)}_{\mathbf{l},\mathbf{k}} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)}.$$

$$\tag{5.2.23}$$

After the classical normalization step, the function $Z^{(2,1)}(P,Q,p,q)$ does not contain angle-dependent terms which are constant or linear in the actions.

The second step focusses on the manipulation of the term

$$g^{(3,1)}(P,Q,p,q) = \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=3}} a^{(3,1)}_{\mathbf{l},\mathbf{k}} P^{l_1} Q^{l_2} + \sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=1}} b^{(3,1)}_{\mathbf{l},\mathbf{k}} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)} \; . \tag{5.2.24}$$

Precisely, the second normalization step aims to remove the second sum in $g^{(3,1)}$ which is angle-dependent and linear in the actions. The generating function $\chi^{(2)}$, given by (5.2.17) with a suitable change in the upper indexes, must satisfy the normal form equations

$$\{\omega_1 P + \omega_2 Q, \chi^{(2)}\} = -\sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=1}} b_{\mathbf{l},\mathbf{k}}^{(3,1)} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)} , \qquad (5.2.25)$$

which gives

$$\chi^{(2)}(P,Q,p,q) = -\sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=1}} \frac{b_{\mathbf{l},\mathbf{k}}^{(3,1)}}{i(\omega_1 k_1 + \omega_2 k_2)} P^{l_1} Q^{l_2} e^{i(k_2 p + k_2 q)} .$$

As a result, the generating function $\chi^{(2)}$ is linear in the actions, so that the operator $\mathcal{L}_{\chi^{(2)}} f$ preserves the polynomial degree in the actions of any generic function $f(P,Q,p,q)$.

The second-order transformed Hamiltonian $\mathcal{H}^{(2)}$ can be written as

$$\mathcal{H}^{(2)}(P,Q,p,q) = \omega_1 P + \omega_2 Q + \sum_{j=2}^{3} Z^{(j,2)}(P,Q) + \sum_{j=4}^{N} g^{(j,2)}(P,Q,p,q) , \qquad (5.2.26)$$

where, noticing that $g^{(0,2)} \equiv 0$, one obtains

$$Z^{(2,2)} = \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=2}} a_{\mathbf{l}}^{(2,2)} P^{l_1} Q^{l_2}, \quad Z^{(3,2)} = \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=3}} a_{\mathbf{l}}^{(3,2)} P^{l_1} Q^{l_2}, \quad g^{(j,2)} = \sum_{s=0}^{\lfloor \frac{j}{2} \rfloor} \frac{1}{s!} \mathcal{L}^s g^{(j-2s,2)} .$$

$$(5.2.27)$$

Taking into account the parities of the indexes j, one can obtain also for $g^{(j,2)}$ the analogous of (5.2.23).

We can now give the explicit formulas for the normalization steps for $r > 2$.

- The $r-$th normalization step allows one to transform the Hamiltonian

$$\mathcal{H}^{(r-1)}(P,Q,p,q) = \omega_1 P + \omega_2 Q + \sum_{j=2}^{r-1} Z^{(j,r-1)}(P,Q,p,q) + \sum_{j=r}^{N} g^{(j,r-1)}(P,Q,p,q)$$

$$(5.2.28)$$

  into

$$\mathcal{H}^{(r)}(P,Q,p,q) = \omega_1 P + \omega_2 Q + \sum_{j=2}^{r} Z^{(j,r)}(P,Q,p,q) + \sum_{j=r+1}^{N} g^{(j,r)}(P,Q,p,q),$$

$$(5.2.29)$$

with

$$
\begin{aligned}
Z^{(2,r)} &= \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=2}} a_{\mathbf{l}}^{(2,r)} P^{l_1} Q^{l_2}, \quad Z^{(3,r)} = \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=3}} a_{\mathbf{l}}^{(3,r)} P^{l_1} Q^{l_2}, \\
Z^{(j>3,r)} &= \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=j}} a_{\mathbf{l}}^{(j,r)} P^{l_1} Q^{l_2} + \sum_{s=2}^{j-2} \sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=s}} b_{\mathbf{l},\mathbf{k}}^{(j,r)} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)}, \\
g^{(j,r)} &= \sum_{\substack{\mathbf{l}\in\mathbb{Z}^2 \\ |\mathbf{l}|=j}} a_{\mathbf{l}}^{(j,r)} P^{l_1} Q^{l_2} + \sum_{s=0}^{j-2} \sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=s}} b_{\mathbf{l},\mathbf{k}}^{(j,r)} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)} .
\end{aligned}
\tag{5.2.30}
$$

By the above parity rules, which apply also for $r > 3$, both $Z^{(j,i)}$ and $g^{(j,i)}$ contain only the terms with s even if j is even and s odd if j is odd. Notice that, for $j > 3$, $Z^{(j,i)}$ can contain also angle-dependent terms, which are at least quadratic in the actions.

- The r-th order generating function can be expressed as

$$
\chi^{(r)}(P,Q,p,q) = - \sum_{\substack{\mathbf{l},\mathbf{k}\in\mathbb{Z}^2 \\ |\mathbf{l}|=0,1}} \frac{b_{\mathbf{l},\mathbf{k}}^{(r+1,r-1)}}{i(\omega_1 k_1 + \omega_2 k_2)} P^{l_1} Q^{l_2} e^{i(k_1 p + k_2 q)},
\tag{5.2.31}
$$

which contains only purely trigonometric terms (independent on the actions) if r is odd and only terms linear in the actions if r is even.

- After N_{norm} normalization steps, the final Hamiltonian is given by

$$
\begin{aligned}
\mathcal{H}^{(N_{norm})}(P,Q,p,q) = \omega_1 P + \omega_2 Q &+ \sum_{j=2}^{N_{norm}} Z^{(j,N_{norm})}(P,Q,p,q) \\
&+ \sum_{j=N_{norm}+1}^{N} g^{(j,N_{norm})}(P,Q,p,q).
\end{aligned}
\tag{5.2.32}
$$

From (5.2.30) it is clear that the functions $g^{(j,\ell)}$ might contain terms which are angle-dependent and constant or linear in the actions. As we will see later, the series are convergent in particular domains of the parameters. In that case, the normalization procedure succeeds to reduce the magnitude of all the terms in the perturbation to a size sufficiently small for the application of the Nekhoroshev theorem.

It is also important to observe that particular angle combinations in the angle-dependent part of the Hamiltonians can produce, if r is odd, constant terms both in actions and angles, which do not affect the dynamics; however, when r is even, the same combinations can produce terms which do not depend on the angles, but are linear in the actions. These terms represent a perturbation on the frequencies, which can have important effects on the applicability of Nekhoroshev theorem.

From the definition of the $r-$th order generating function (5.2.31), one can observe that the convergence of the normalization algorithm depends heavily on the presence of *resonances*, which produce small divisors of the type $\omega_1 k_1^{(res)} + \omega_2 k_2^{(res)} \approx 0$ for suitable

integers $k_1^{(res)}$, $k_2^{(res)}$. Section 5.4.2 provides numerical examples of how the presence of resonances can affect the convergence of the normalization procedure, along with effects on the variation of the initial frequencies.

5.3   Nekhoroshev stability estimates

We are now ready to recall the version of the Nekhoroshev theorem developed in [54] for frequencies satisfying a non-resonance condition (see Section 5.3.1). Based on this theorem, we developed an algorithm computing all quantities needed in order to check whether the necessary conditions for the holding of the theorem are fulfilled in the case of the Hamiltonian (5.2.32). The algorithm is presented in Section 5.3.2.

5.3.1   Theorem on exponential stability

Let us consider an $n-$dimensional quasi-integrable Hamiltonian of the form

$$\mathcal{H}(\mathbf{I}, \mathbf{u}) = h(\mathbf{I}) + f_\epsilon(\mathbf{I}, \mathbf{u}) \ ,$$

with h called the integrable part and f_ϵ the perturbing function, depending on a small real parameter $\epsilon > 0$. The Hamiltonian $\mathcal{H}$ is assumed real analytic in the domain $(\mathbf{I}, \mathbf{u}) \in A \times \mathbb{T}^n$ with $A \subseteq \mathbb{R}^n$ open and bounded. Besides, we assume that $\mathcal{H}$ can be extended analytically to the set D_{r_0, s_0} defined as

$$D_{r_0, s_0} = A_{r_0} \times \mathbb{T}^n_{s_0} \ , \tag{5.3.1}$$

where for $r_0, s_0 > 0$:

$$A_{r_0} = \{\mathbf{I} \in \mathbb{C}^n : dist(\mathbf{I}, A) < r_0\} \tag{5.3.2}$$

and

$$\mathbb{T}^n_{s_0} = \{\mathbf{u} \in \mathbb{C}^n : Re(u_j) \in \mathbb{T}, \ \max_{j=1,\dots,n} |Im(u_j)| < s_0\} \ .$$

Finally, we assume that there exists a positive constant M such that

$$\sup_{\mathbf{I} \in A_{r_0}} \|\mathcal{Q}(\mathbf{I})\|_o \leq M \ ,$$

where $\mathcal{Q}$ denotes the Hessian matrix associated to h and $\| \cdot \|_o$ denotes the operator norm induced by the Euclidean norm on $\mathbb{R}^n$.

For any analytic function

$$g(\mathbf{I}, \mathbf{u}) = \sum_{\mathbf{k} \in \mathbb{Z}^n} g_{\mathbf{k}}(\mathbf{I}) e^{i\mathbf{k} \cdot \mathbf{u}},$$

in D_{r_0, s_0}, we define its Cauchy norm as

$$|g|_{A, r_0, s_0} = \sup_{\mathbf{I} \in A_{r_0}} \sum_{\mathbf{k} \in \mathbb{Z}^n} |g_{\mathbf{k}}(\mathbf{I})| e^{|\mathbf{k}| s_0} \ , \tag{5.3.3}$$

where $|\mathbf{k}|$ is the ℓ^1-norm of $\mathbf{k} \in \mathbb{Z}^n$. Finally, let ϵ be such that

$$|f_\epsilon|_{A,r_0,s_0} \leq \epsilon \ . \tag{5.3.4}$$

The following Theorem provides a bound on the action variables for exponentially long times; we refer to [54] for the proof and further extensions. First we need the following definition.

Definition 5.3.1. *A set $D \subseteq A$ is said to be a completely α, K-nonresonant domain in A, if for every $\mathbf{k} \in \mathbb{Z}^n \backslash \{\mathbf{0}\}$, $|\mathbf{k}| \leq K$, and for every $\mathbf{I} \in D$*

$$|\mathbf{k} \cdot \boldsymbol{\omega}(\mathbf{I})| \geq \alpha > 0 \ , \tag{5.3.5}$$

where $\boldsymbol{\omega}(\mathbf{I}) = \partial_\mathbf{I} h(\mathbf{I})$.

Theorem 5.3.2 ([54]). *Let $D \subseteq A$ be a completely α, K-nonresonant domain. Let $a, b > 0$ such that $\dfrac{1}{a} + \dfrac{1}{b} = 1$. Let ϵ be as in (5.3.4) for some $r_0, s_0 > 0$. If the following inequalities are satisfied:*

$$\epsilon \leq \frac{1}{2^7 b} \frac{\alpha r}{K} = \epsilon^*, \qquad r \leq \min\left(\frac{\alpha}{aMK}, r_0\right) \ , \tag{5.3.6}$$

then, denoting by $||\cdot||$ the Euclidean norm in A, one has

$$\|\mathbf{I}(t) - \mathbf{I}_0\| \leq r \quad for \quad |t| \leq \frac{s_0 r}{5\epsilon} e^{K s_0/6} \tag{5.3.7}$$

for every orbit of the perturbed system with initial position $(\mathbf{I}_0, \mathbf{u}_0)$ in $D \times \mathbb{T}^n$.

5.3.2 Algorithm for the application of the theorem

To apply Theorem 5.3.2 to the final Hamiltonian $\mathcal{H}^{(N_{norm})}$ defined in (5.2.32), one has to compute all the quantities involved in the Theorem itself. This procedure gives rise to an explicit constructive algorithm to give stability estimates for every pair of reference values (e_*, i_*) in the uniform grid $[0, 0.5] \times [0, 89.5°]$ with step-size equal to 0.1 in eccentricity and 0.5° in inclination. Notice that the upper value of the grid in inclination is equal to 89.5° to avoid singularities.

First, we need to determine the greatest integer $\bar{K}$, to which we refer as the *cut-off value*, such that conditions (5.3.6) hold. From the definition of α in (5.3.5) and ϵ^* in (5.3.6), it is clear that ϵ^* decreases as K increases; then, provided that condition (5.3.6) holds for $K = 1$, the maximal value $\bar{K}$ exists. On the other hand, if (5.3.6) does not hold for $K = 1$, it continues to remain false for all $K > 1$.

From a computational point of view, the procedure is composed by the following steps, (S1), ..., (S8), performed for every pair (e_*, i_*) in the grid defined above; by trial and error, we fix the values of r_0, s_0, a, b. Their choice is arbitrary and can be tuned so to satisfy the conditions of the Theorem and to optimize the final estimates.

(S1) Taylor expansion up to order $N = 12$ in the expansion (5.2.9) around the actions (P_*, Q_*), corresponding to the Keplerian elements (e_*, i_*);

(S2) normalization up to order $N_{norm} = 6$, following the procedure described in Section 5.2.4, which leads to compute the normalized Hamiltonian $\mathcal{H}^{(N_{norm})}$;

(S3) splitting of the Hamiltonian $\mathcal{H}^{(N_{norm})}$ in the unperturbed part $h_0(P, Q)$, containing the terms of $\mathcal{H}^{(N_{norm})}$ which depend only on the actions, and the perturbing part $h_1(P, Q, p, q) = \mathcal{H}^{(N_{norm})}(P, Q, p, q) - h_0(P, Q)$; computation[3] of $\boldsymbol{\omega} = (\omega_1, \omega_2)$, with ω_1 and ω_2 coefficients respectively of P and Q in h_0;

(S4) definition of the real and complexified domains in the actions as in (5.3.1) and computation of the quantity

$$M = \sup_{(P,Q) \in A_{r_0}} ||\mathcal{Q}(P, Q)||_o \; ; \tag{5.3.8}$$

in particular, we define

$$A = [P_* - dP^{(max)}, P_* + dP^{(max)}] \times [Q_* - dQ^{(max)}, Q_* + dQ^{(max)}]$$

with $dP^{(max)} = dQ^{(max)} = 0.1$; we select $r_0 = s_0 = 0.1$ and, following [54], we take $a = 9/8$ and $b = 9$;

(S5) for every $K = 1, \ldots, 50$, computation of the quantities

$$\alpha_K = \min_{|\mathbf{l}| \leq K} \{\boldsymbol{\omega} \cdot \mathbf{l}\}, \qquad r_K = \min\left\{\frac{\alpha_K}{aMK}, r_0\right\}, \qquad \epsilon_K^* = \frac{1}{2^7 b} \frac{\alpha_K r_K}{K} \; ; \tag{5.3.9}$$

(S6) defining $\epsilon = |h_1|_{A, r_0, s_0}$, check of the condition $\epsilon \leq \epsilon_K^*$ for every $K = 1, \ldots, 50$;

(S7) if $\epsilon \leq \epsilon_1^*$, computation of $\bar{K}$, namely the greatest K such that $\epsilon \leq \epsilon_K^*$, and of the corresponding stability time

$$t = \frac{s_0 r_{\bar{K}}}{5\epsilon} e^{\bar{K} s_0/6} \; ; \tag{5.3.10}$$

(S8) if $\epsilon > \epsilon_1^*$, the conditions for the application of Theorem 5.3.2 are not satisfied. In this case, we impose $\bar{K} = 0$.

We remark that the order of the Taylor expansion $N = 12$, the order of normalization $N_{norm} = 6$, the iteration of K up to 50 are set on the basis of having a reasonable computational execution time on standard laptops.

5.4 Results

The current section is devoted to the presentation of the results of the application of Theorem 5.3.2 to the Hamiltonian model described in Section 5.2. This allows us to derive stability estimates as well as to discuss the convergence of the normalization procedure.

---

[3]With an abuse of notation, we continue to define the new frequencies, which could be modified by the normalization, with the symbols ω_1 and ω_2. When, in Section 5.4.2, it will be required to distinguish between the initial and the final frequencies, the latter will be denoted by $\tilde{\omega}_1$ and $\tilde{\omega}_2$.

Table 5.2 Inclination-dependent resonances which affect the stability in the lunisolar model. The coefficients α and β are such that $\alpha\dot{p} + \beta\dot{q} = 0$.

α	β	$i(deg)$	α	β	$i(deg)$	α	β	$i(deg)$	α	β	$i(deg)$
1	0	46.37	0	1	90	1	1	0	1	-1	63.4351
2	1	33.0156	-1	2	73.1484	-2	1	56.0646	-2	3	69.007
-4	3	60.0001	-4	1	51.5596	-1	3	78.4633	-4	5	66.422

5.4.1   Stability estimates

We apply the algorithm of Section 5.3.2 to probe the Nekhoroshev stability for satellites with semimajor axes between $11\,000$ km and $19\,000$ km under the model presented in Section 5.2. The results exposed below highlight the strong dependence of the stability conditions on the precise values of the elements (e, i). Of crucial role in this dependence is the location of the 'inclination-dependent' resonances (see Section 5.1). These satisfy a condition of the form $\alpha\dot{p} + \beta\dot{q} = 0$ for some coefficients α, $\beta \in \mathbb{Z}$.

Table 5.2 shows the values of the inclinations corresponding to each pair of coefficients (α, β). We find that these resonances determine regions where Theorem 5.3.2 cannot be applied. This can be exemplified with the help of Figure 5.1, showing (in blue) the region where the algorithm of Section 5.3.2 returns that the necessary conditions of Theorem 5.3.2 hold true. The algorithm provides an answer as a function of the chosen reference values i_* and e_* (for a fixed a_*). Figure 5.2 shows the Nekhoroshev stability times computed at every grid point (e_*, i_*) in the previous figure's blue region.

It is evident from Figure 5.1 that increasing the distance from the Earth's center causes a shrinking of the size of the domains of Nekhoroshev stability, as well as a fast decrease of the corresponding computed stability times. From the physical point of view, this tendency is evident and can be explained on the basis of the simple remark that the averaged Hamiltonian $\mathcal{H}_{kep} + \mathcal{H}_{J_2}$ without third-body perturbations is integrable (this Hamiltonian has in fact no dependence on the Delaunay angles). Since the overall relative size of third body perturbations increases with the altitude, these perturbations affect the stability more as a_* increases. At a formal level the effect of the semimajor axis on the estimates can be identified by an analysis of the convergence of the preliminary normalization algorithm (see Section 5.4.2 below).

On the other hand, also evident from Figure 5.1 is the strong role of resonances in affecting the stability properties of the system: in fact, around every of the resonances listed in Table 5.2 we observe, in the figure, the formation of a white zone, which indicates values (e_*, i_*) excluded from the Nekhoroshev stability as detected by our algorithm. As a general comment, the presence of the resonances acts at two different stages of the computation:

(i) it can affect the convergence of the classical normalization, producing an increase of the size of the perturbing function and a consequent failure of conditions (5.3.6);
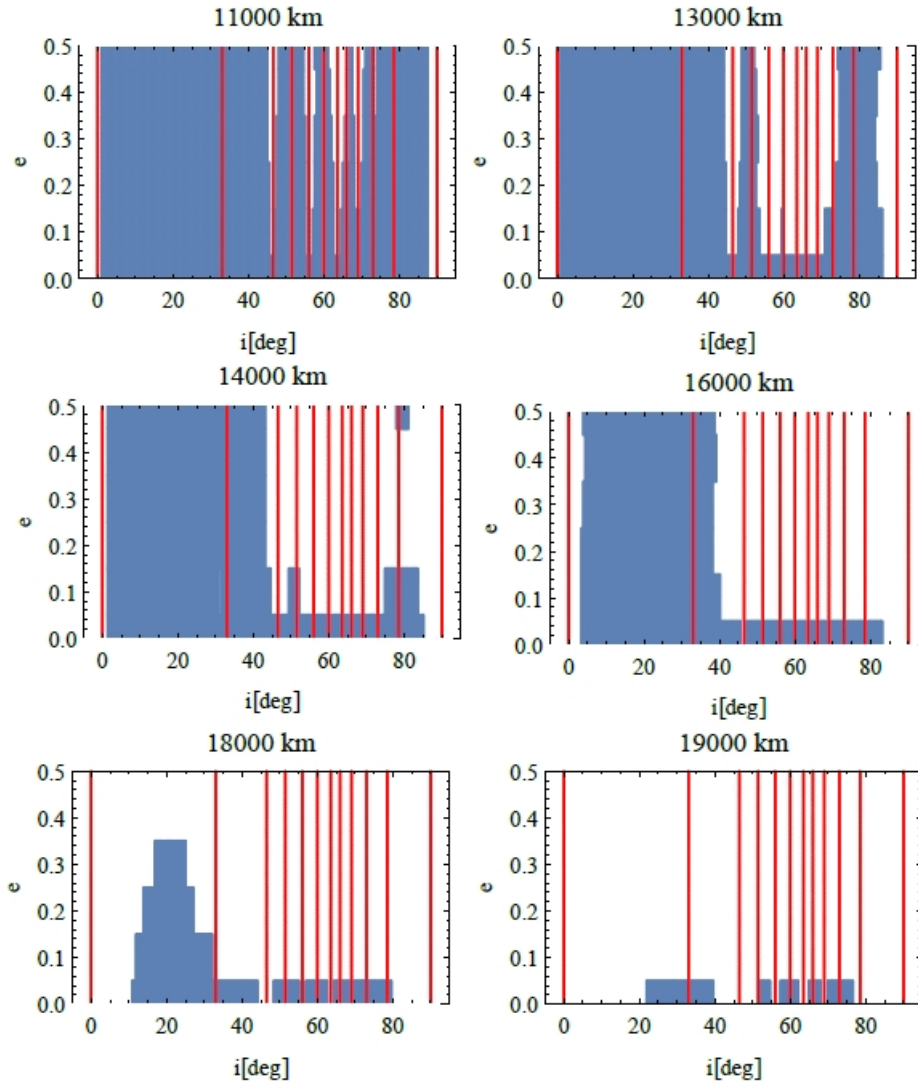
Fig. 5.1 Domains of applicability of Theorem 5.3.2 for different values of the altitude. The blue regions represent the values of (i_*, e_*) for which the Theorem can be applied, while the red lines define the values of the inclination which are associated with the most important resonances in the considered regions (see Table 5.2).

(ii) near the low-order resonant values of the inclination, the quantity α_K (see (5.3.9)) can be extremely small, even for low values of K. As a consequence, in the proximity of a resonance, the corresponding value of the quantity ϵ_K^* might not be small enough to satisfy (5.3.6).

At any rate, we stress that Theorem 5.3.2 used in this framework holds only for non-resonant domains in the phase space; therefore, by definition it cannot be used to probe the Nekhoroshev stability very close to resonances. We defer to a future study the question of the precise investigation of the conditions for Nekhoroshev stability inside resonances, by implementing a resonant form of the theorem, as first suggested in [53].
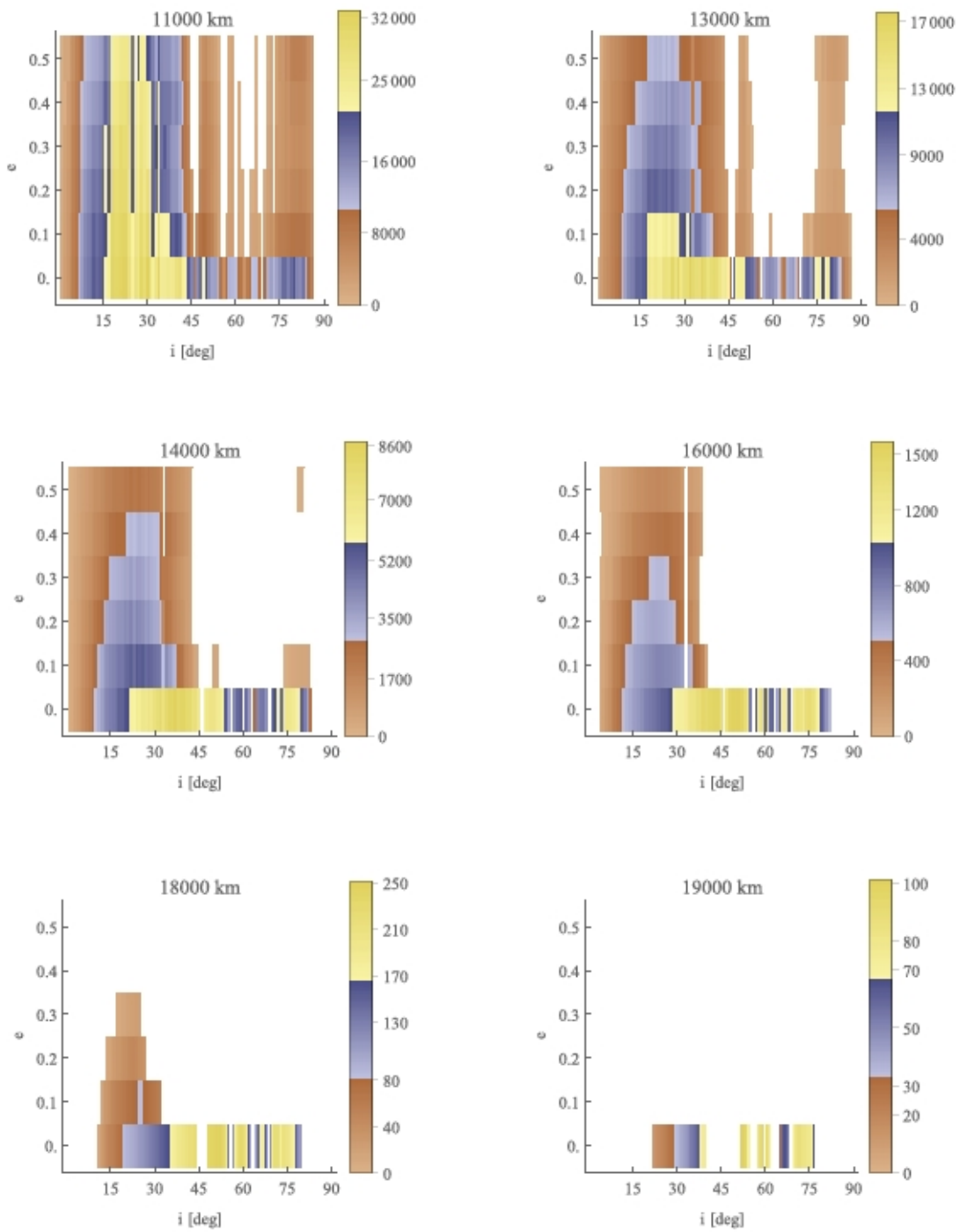
Fig. 5.2 Stability time (in years) computed for the values of (i_*, e_*) in the domain of applicability of Theorem 5.3.2.
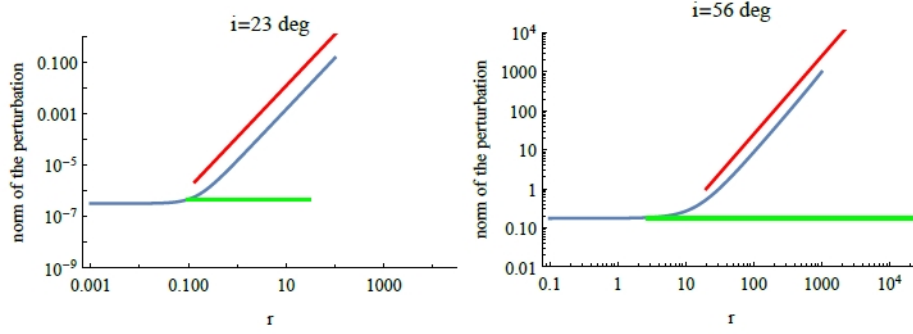
Fig. 5.3 Blue: plot in *LogLog* scale of $|f_\epsilon|_{A,r,s_0}$ for $a = 13\,000$ km, $e_* = 0.2$, and $i_* = 23°$ (left) and $i_* = 56°$ (right). For the computations, we selected $A = [P_* - r, P_* + r] \times [Q_* - r, Q_* + r]$, $r_0 = r$, $s_0 = 0.1$. The slope of the plot for high r is compared with that of a line with slope 2 (red); the value at the plateau (denoted with a green line) is compared with the value of the norm of the purely trigonometric part of f_ϵ with $s_0 = 0.1$.

5.4.2   Convergence of the preliminary normalization

As pointed out in Section 5.2.4, the aim of the preliminary normalization is to allow to control the norm of the perturbing function $|f_\epsilon|_{A,r_0,s_0}$ by reducing the size of the complexified action domain A_{r_0} (see (5.3.2)). In particular, the consequence of the removal of angle-dependent terms which are constant or linear in the actions is that, within certain values of the size of the domain A_{r_0}, the norm of the perturbation decreases quadratically with the actions.

Figure 5.3 shows the behaviour of $|f_\epsilon|_{A,r,s_0}$ for $a = 13\,000$ km, $e_* = 0.2$ and two selected values of i_*, as a function of the size of the action in the complexified domain A_r (the domain A is set to be a rectangle of width $2r$ around the central values P_* and Q_*). As expected, the value of $|f_\epsilon|_{A,r,s_0}$ decreases quadratically with r, until it reaches a plateau, whose value is the norm of the terms of f_ϵ which do not depend on the actions.

As already mentioned in Section 5.4.1, the convergence of the normalization presented in Section 5.2.4 for $\mathcal{H}^{(sec)}$ is crucial to control the size of the perturbing function h_1; such value plays a fundamental role in Theorem 5.3.2. A first study of the effect of the chosen value of the semimajor axis on the convergence can be performed by considering a simpler model to which a normalization procedure similar to the one implemented in Section 5.2.4 can be performed. The model is defined by the Hamiltonian

$$\widetilde{\mathcal{H}}^{(in)}(P, Q, p, q) = \omega_1 P + \omega_2 Q + \frac{c_2}{2}Q^2 + f_1 \cos q \; , \tag{5.4.1}$$

where the frequencies ω_1, ω_2 and the coefficient c_2 depend essentially only on the J_2 averaged Hamiltonian, while the coefficient f_1 depends on the lunar and solar third-body perturbation potentials, and it is proportional to the sinus of the inclination i_0 of the ecliptic. We will examine the effect of performing the preliminary normalization algorithm on the Hamiltonian $\widetilde{\mathcal{H}}^{(in)}$ so as to remove purely trigonometric terms. After N_{norm} normalization steps, the Hamiltonian takes the form:

$$\widetilde{\mathcal{H}}^{(fin)} = \omega_1 P + \omega_2 Q + \frac{c_2}{2}Q^2 + \sum_{\beta=1}^{N_{norm}} Z_i(P, Q, q) + \sum_{i=N_{norm}+1}^{\infty} R_i(P, Q, q) \; , \tag{5.4.2}$$

where the normalized parts $Z_i(P, Q, q)$ do not contain terms which depend only on the angle q (as well as linear terms in the actions multiplied by trigonometric terms). By an explicit computation of the Poisson brackets involved in the normalization, we readily find that $R_{N_{norm}+1}$ contains trigonometric terms with coefficients proportional to the quantity

$$2^a f_1 \left(\frac{c_2 f_1}{4\omega_2^2} \right)^{N_{norm}} , \tag{5.4.3}$$

where $a = 1, 2, 3$ depends on the value of N_{norm}. The convergence of the remainder through the steps of the normalization algorithm depends, then, on the value of the ratio $c_2 f_1 / 4\omega_2^2$; in particular, when this quantity is greater than 1, the normalization does not converge. Neglecting the lunar and solar contributions in ω_1, ω_2 and c_2, the coefficient $c_2 f_1 / 4\omega_2^2$ can be expressed in terms of the orbital elements of debris, Sun and Moon as

$$\frac{c_2 f_1}{4\omega_2^2} = \frac{1}{32} \frac{\sin 2i_0}{R_E^2 \mu_E J_2} \left(\frac{\mu_M}{(a_M(1 - e_M))^3} + \frac{\mu_\odot}{(a_\odot(1 - e_\odot))^3} \right) a^5 (2 + 3e_*^2)(1 - e_*^2)^{3/2} \tan i_* .$$

$$\tag{5.4.4}$$

As a consequence, it is clear that its size strongly depends on a and i_*: it grows sharply when a increases and when i_* approaches $90°$.

On the other hand, the coefficient f_1 is proportional to $\sin 2i_0$, that is, proportional to the (non-zero) inclination $i^{(p)}$ of the Laplace plane (see Eqs. (4.2.20) and further). Hence, the presence in the secular Hamiltonian of purely trigonometric terms is a manifestation of the presence in the model of a Laplace plane. Since $i^{(p)}$ increases with a and f_1 increases both with $i^{(p)}$ and i_*, this gives a first explanation of the loss of stability of the model as a and i_* increase.

As already mentioned in Section 5.4.1, the other important factor influencing the size of the remainder across the preliminary normalization process is the effect of resonances, which, due to Eq.(5.2.31), leads to the appearance, in the series terms, of small divisors. Of particular importance are the small divisors appearing in the series' purely trigonometric terms, whose size cannot be controlled by altering the size of the domain in the actions A_{r_0}.

Figure 5.4 shows the behaviour of the norm of the purely trigonometric part of the perturbation h_1 (with the notation (S3) of Section 5.3.2) as a function of the inclination for four different values of a and two different values of e. As one can see, the size of the trigonometric part reaches its peaks in correspondence of the resonant values of the inclination, as expected. We also notice that the number of resonances involved in the growth of the size of the trigonometric part increases with a and e.

As explained in Section 5.2.4, the normalization algorithm used here does not perform a re-tuning of the frequencies for every normalization step. This fact has important effects on the applicability of Theorem 5.3.2: when the normalization converges, the change between the original and the new frequencies is negligible with respect to their magnitude; on the other hand, when it does not converge, a large variation in the value of the frequencies occurs, with important consequences on the computation of α_K and, therefore, of the quantities involved in Theorem 5.3.2.

As an example, Figure 5.5 shows the variations of the frequencies as a function of the inclination for $a = 13\,000$ km and $e_* = 0.2$. Comparing Figures 5.4 and 5.5, it is
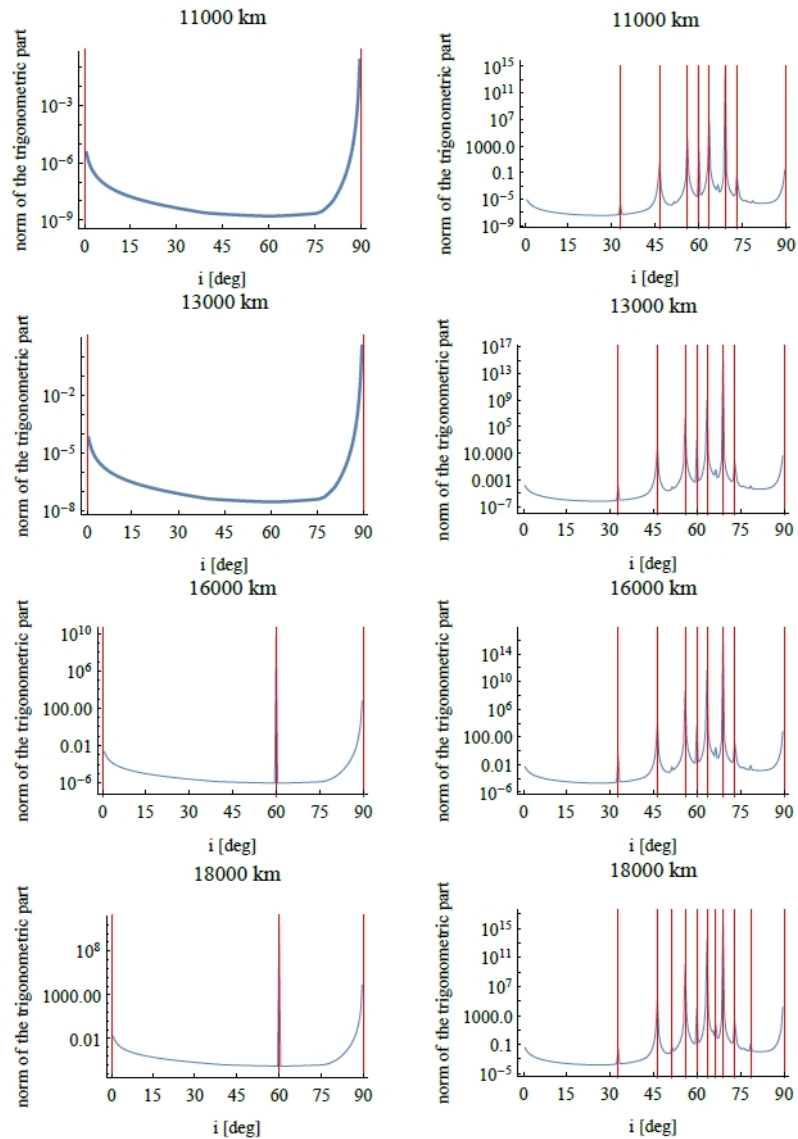
Fig. 5.4 Behaviour of the norm of the purely trigonometric part of h_1 as a function of the inclination i_* for different semimajor axes and eccentricities (left: $e = 0$, right: $e = 0.5$). The red lines represent the inclinations of the resonances (see Table 5.2).
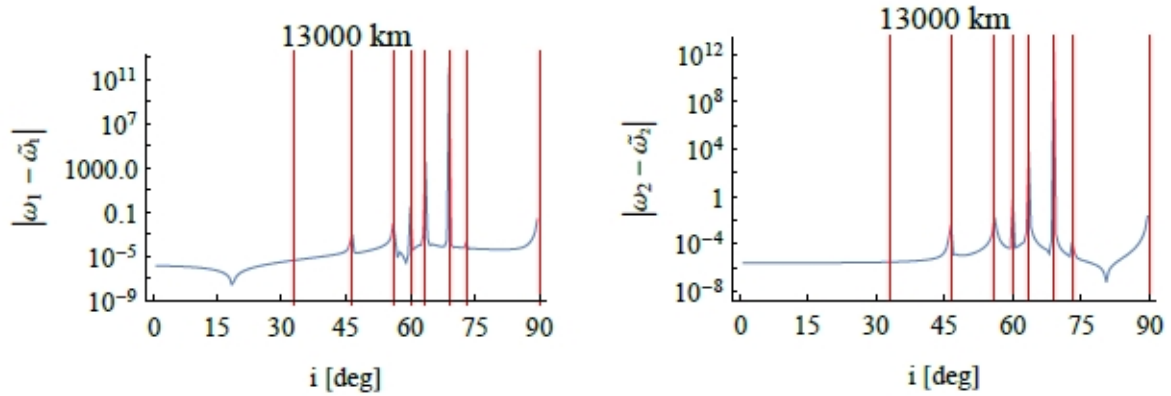
Fig. 5.5 Variation between the initial (ω_1 and ω_2) and the final ($\tilde{\omega}_1$ and $\tilde{\omega}_2$) frequencies as a function of the inclination i_* for $a = 13\,000$ km and $e_* = 0.2$. The red lines represent the values of i_* associated to the resonances which affect the convergence of the normalization algorithm (see Figure 5.4).

$i_*(deg)$	$N(i_*)$	$\bar{K}$	$i_*(deg)$	$N(i_*)$	$\bar{K}$	$i_*(deg)$	$N(i_*)$	$\bar{K}$	$i_*(deg)$	$N(i_*)$	$\bar{K}$
46.5	1	0	89.5	1	0	1	2	1	63.5	2	0
33	3	2	73	3	0	56	3	0	38	3	3
53	4	1	78.5	4	3	40.5	4	5	27	5	4
51.5	5	4	58.5	5	0	69	5	0	81.5	5	4
41.5	6	5	50.5	6	5	83.5	6	4			

Table 5.3 Comparison between the order $N(i_*)$ of the nearest resonance and the computed cut-off value $\bar{K}$, computed for $a = 13\,000$ km and $e = 0.1$.

clear that the resonances which affect the growth in size of the purely trigonometric part of h_1 and the variation of the frequencies are the same.

5.4.3 Behaviour of the cut-off value $\bar{K}$

Provided that the classical normalization converges, from the definition of the cut-off value $\bar{K}$ given in Section 5.3.2, one expects that exactly at a resonance, once denoted with $N(i_*) = |\alpha| + |\beta|$ its order, one has $\bar{K} = N(i_*) - 1$. Since in our analysis the inclinations are selected in a mesh of $[0, 89.5°]$ with step $0.5°$, the computations of the quantities involved in Theorem 5.3.2, including $\bar{K}$, are not performed exactly at resonance (with the exception of $i_* = 60°$, whose distance from the exact resonance is of the order of 10^{-3}): Table 5.3 shows the value of $\bar{K}$ computed for the points of the mesh which are near to the resonances up to order 6, with $a = 13\,000$ km and $e = 0.1$, along with the resonance order $N(i_*)$ of the nearest one. With the exception of the inclinations associated to resonances which affect the convergence of the classical normalization, the majority of the listed inclinations follows the expected rule $\bar{K} = N(i_*) - 1$, while some slight deviation is probably due to the numerical computation.
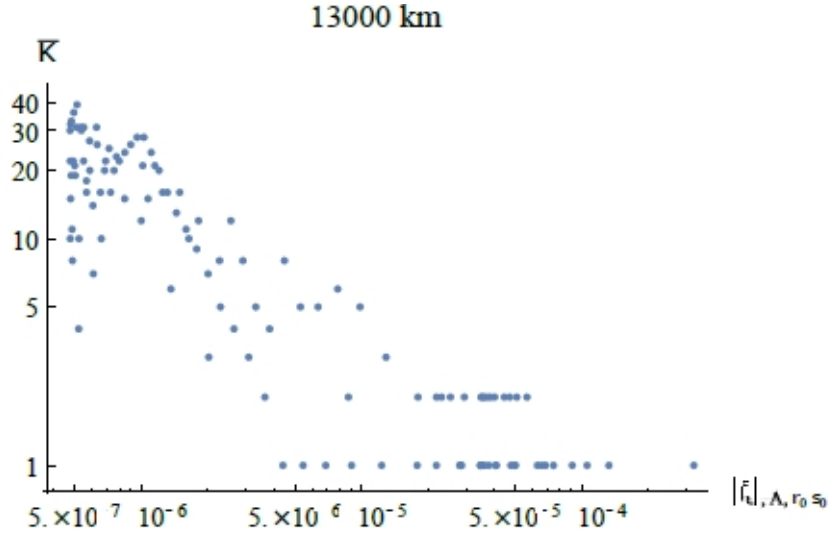
Fig. 5.6 Plot in *LogLog* scale of the points $\{|f_\epsilon|_{A,r_0,s_0}, \bar{K}\}$ for $a = 13\,000$ km, $e_* = 0.2$ and $i_* \in [0, 90°]$ on a mesh of step $0.5°$.

To conclude, Figure 5.6 shows the relation between the computed values of $\bar{K}$ and $|f_\epsilon|_{a,s_0,r_0}$ for $a = 13\,000$ km, $e_* = 0.2$ and $i_* \in [0, 90°]$. As expected, the cut-off decreases exponentially with the norm of the perturbing function.

5.5 Effect of higher order geopotential terms

The results of Section 5.4 were obtained by considering as basic model for MEO the one based on the J_2 geopotential terms. It is well known that for the lowermost altitude at MEO ($a = 10000$ km) the secular dynamics is shaped by higher order terms (e.g., J_2^2) as well as higher harmonics in the Earth's geopotential. In the present section, we examine a model in which the J_2^2 terms obtained by second order averaging of the J_2 Hamiltonian term with respect to the particle's mean anomaly, as well as the first order averaging with respect to the J_3 and J_4 terms, are considered. The Hamiltonian is now as in Eq. (5.2.5), but with

$$\mathcal{H}_E^{(av)} = \mathcal{H}_{kep}^{(av)} + \mathcal{H}_{J_2}^{(av)} + \mathcal{H}_{J_3}^{(av)} + \mathcal{H}_{J_4}^{(av)} + \mathcal{H}_{J_2^2}^{(av)} \ , \tag{5.5.1}$$

where

$$\mathcal{H}_{J_3}^{(av)} = J_3 \frac{3\mu_E R_E^3 e \sin i}{2a^4\eta^5} \left(1 - \frac{5}{4}\sin^2 i\right)\sin(\omega)$$

$$\mathcal{H}_{J_4}^{(av)} = J_4 \frac{3\mu_E R_E^4}{8a^5\eta^7}\left(-1 - 1\frac{3e^2}{2} + \left(5 + \frac{15e^2}{2}\right)\sin^2 i\right.$$

$$\left. - \frac{35}{8}\left(1 + \frac{3e^2}{2}\right)\sin^4 i - \frac{15e^2\sin^2 i}{4}\left(1 - \frac{3\sin^2 i}{2}\right)\cos 2\omega\right)$$

while

$$
\begin{aligned}
\mathcal{H}_{J_2^2}^{(av)} &= -J_2^2 \frac{3\mu_E R_E^4}{8a^5\eta^7}\Bigg[\frac{5}{2} + \eta - \frac{1}{2}\eta^2 - \left(5 + 3\eta - \frac{1}{2}\eta^2\right)\sin^2 i \\
&+ \left(\frac{35}{16} + \frac{9}{4}\eta + \frac{5}{16}\eta^2\right)\sin^4 i - \left(\left(\frac{15}{4}(1+\eta) - \frac{23}{4}\eta^2 - \frac{7}{4}\eta^3\right)\sin^2 i \right. \\
&+ \left. \left(\frac{35}{8}(1-\eta) + \frac{55}{8}\eta^2 + \frac{15}{8}\eta^3\right)\sin^4 i\right)\frac{\cos(2\omega)}{1+\eta}\Bigg]
\end{aligned}
$$

with $\eta = \sqrt{1-e^2}$.

The Hamiltonian (5.5.1) can be obtained directly by eliminating the small body's mean anomaly through a Lie canonical transformation performed in two stages, as indicated in [141] (see [139] for details): in the first stage, called the *elimination of the parallax*, the Hamiltonian is transformed into a function of the form

$$
\begin{aligned}
\mathcal{H}^{(el)} &= \mathcal{H}_{kep} + \frac{1}{r^2}\Bigg(h_{J_2}^{(el)}(a,e,i,\omega,\Omega) \\
&+ h_{J_3}^{(el)}(a,e,i,\omega,\Omega) + h_{J_4}^{(el)}(a,e,i,\omega,\Omega) + h_{J_2^2}^{(el)}(a,e,i,\omega,\Omega)\Bigg) \ ,
\end{aligned}
$$

thus reducing the dependence of the Hamiltonian on the small body's mean anomaly M in only the factor $1/r^2$. In the second stage, we then eliminate this dependence with the usual procedure of Delaunay normalization ([145]). It should be stressed that this procedure yields equivalent results for the terms $\mathcal{H}_{J_2}^{(av)}$, $\mathcal{H}_{J_3}^{(av)}$ and $\mathcal{H}_{J_4}^{(av)}$ as the simple *scissor averaging* of Eq. (5.2.7), but it allows to formally introduce terms of higher order as $\mathcal{H}_{J_2^2}^{(av)}(a,e,i,\omega,\Omega)$. Also, an important difference is in the physical interpretation, since the Lie transformation, which is a near to identity transformation mapping the original canonical variables to new ones, still contains short-periodic terms. In the jargon of astrodynamics, this is called a transformation from *osculating* to *mean elements*. As already pointed out in Section 5.2.2, this means that, formally, all the results on Nekhoroshev stability in this and in previous sections refer to the stability of the mean elements, while the osculating elements perform short-period bounded oscillations around the secularly evolving values of the mean elements.

Returning to the latter question, Figure 5.7 allows to compare the results on Nekhoroshev stability using the Hamiltonian (5.2.5) with $\mathcal{H}_E^{(av)}$ computed as in (5.5.1), with those of the simple J_2-only model obtained as in (5.2.7).

Figure 5.7 provides information on both the domain of applicability of the Nekhoroshev theorem as well as the corresponding stability times: we consider the case of orbits with $e = 0.3$ and two different values of the semi-major axis, namely $a = 11000$ km (top panel in Fig. 5.7) and $a = 17000$ km (middle panel). The abscissa of each of the marked points indicates a value of the inclination for which applying Pöschel's theorem in the form of the algorithm of subsection 5.3.2 yields a positive result, i.e., that the Nekhoroshev stability criterion holds. The ordinate then indicates the corresponding time of Nekhoroshev stability. From these figures stem the following remarks:
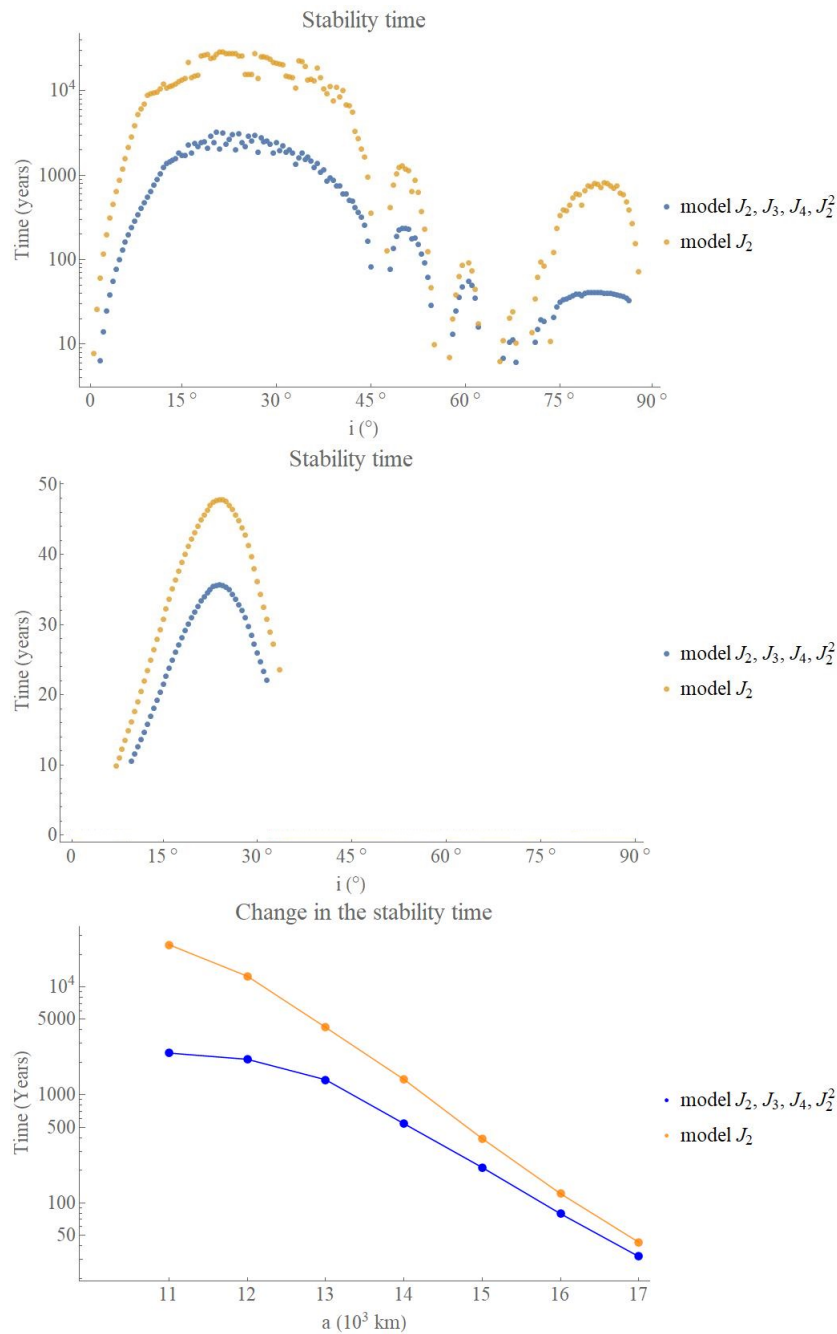
Fig. 5.7 Comparison of the domains and times of Nekhoroshev stability for different values of the inclination between the J_2 model (yellow dots) and the model including J_2^2, J_3, J_4 (blue dots), for a fixed eccentricity $e = 0.3$ and semimajor axis equal to $a = 11000$ km (top panel), or $a = 17000$ km (middle panel). A colored point indicates that Pöschel's criterion for Nekhoroshev stability is satisfied at the corresponding value of the inclination, shown in the abscissa. The ordinate shows the corresponding value of the Nekhoroshev stability time (in years). Bottom panel: comparison of the Nekhoroshev stability times as a function of the semimajor axis a in the J_2 model and in the extended model for $e = 0.3$ and $i = 20^o$.

i) the two models (named simply J_2 and 'extended') yield practically identical *domains* of stability. This is to be expected, since the domains of stability are mostly determined by the values of the (Diophantine) frequencies of the integrable part of the Hamiltonian h_0. In the extended model, the frequencies differ from those of the J_2-model by the addition of the terms $O(J_2^2)$, $O(J_3)$ and $O(J_4)$. All these terms are about 10^{-3} the size of the leading (J_2) terms, thus they only affect the frequencies at the third digit. This implies, in turn, that all Diophantine constants, cut-off in Fourier space etc., entering into the application of Pöschel's theorem remain practically invariant in computations with the extended model.

ii) On the other hand, the computed *times* of Nekhoroshev stability change, by about one order of magnitude at the lowest limit of the MEO zone ($a = 11000$ km), and marginally as we approch to the limit of the overall loss of the Nekhoroshev stability in Pöschel's sense $a > 17000$ km. The main reason for this difference lies in the integrability of the J_2 averaged model, which implies that only lunisolar perturbations affect the size of the term h_1 in the Hamiltonian of the simple J_2 model. In the extended model, instead, all three $O(J_2^2)$, $O(J_3)$ and $O(J_4)$ terms contribute to H_1 due to their containing $\cos(2\omega)$ and $\sin(\omega)$ terms depending on the canonical angles. It is noteworthy that the relative importance of these terms decreases as a power of the semimajor axis (see Eq.(5.5.1)), while the lunisolar perturbations increase with a, e.g. as a^2 for the quadrupolar terms. In particular, we find that the $\cos(2g)$ term due to J_2^2 and J_4 is dominant over the $\cos(2\omega)$ term generated by the lunisolar perturbation when $a = 11000$ km, but the former is only about 1.5% the size of the latter when $a = 20000$ km. As a result, the two different models converge as regards the times of stability as a increases, a tendency shown clearly in the bottom panel of Fig.5.7.

At any rate, it is important to note that in both models the computed times of stability correspond to 10^7 orbital revolutions in the lowermost limit of the MEO zone, reducing to about 10^5 orbital revolutions in the highermost limit where Nekhoroshev stability holds. These times are thus quite consistent with applications of the Nekhoroshev theorem in the practical context of the long term stability of satellite orbits, as they are larger by orders of magnitude compared to the satellites' operational lifetime.

5.6 Conclusions

The work presented in this chapter has a two-fold aim: on one side, we provide a specific algorithm by which we are able to formally specify the domains in the space of orbital elements (a, e, i) for which Nekhoroshev stability holds in the sense that all the necessary conditions for the applicability of Pöschel's theorem for non-resonant orbits are satisfied. On the other side, in those domains where Nekhoroshev stability holds we compute the associated Nekhoroshev times, and demonstrate that these times are long enough to be of use in practical Earth satellites applications. Our main results can be summarized as follows.

1. We examine in detail a secular model based on a 'scissor' averaged Hamiltonian over the fast angles, including the term J_2 as well as lunisolar perturbations. For this model,

we propose: i) a detailed 'book-keeping' algorithm allowing to write the Hamiltonian in a form suitable for the application of Pöschel's theorem, and ii) a 'preliminary normalization', which leads to a model devoid of the effects of trigonometric terms generated by the shifting of the secular equilibrium from the Earth's equator to the Laplace plane. Albeit technical, this step, explained in detail in subsection 5.2.4, turns to be crucial in suitably engineering the Hamiltonian so that relevant estimates on the real size of the secular perturbations can be obtained.

2. We then propose, in subsection 5.3.2, a particular algorithm by which the theorem of Pöschel can be transformed into a binary ("yes" or "no") criterion for the holding of Nekhoroshev stability in a small domain around any preselected value of the elements (a, e, i) within the MEO zone. Implementing this algorithm leads to the results of Figure 5.1: as intuitively expected, we find that the domains of Nekhoroshev stability shrink as the altitude (semimajor axis a) increases. This is due to the growing size of lunisolar perturbations as a increases. The most robust domain is found in the intervals $0 \leq e \leq 0.3$ and $10^o \leq i \leq 30^o$; the latter interval roughly corresponds to a domain well protected from major inclination-only dependent lunisolar resonances.

3. Using the same algorithm we can compute the times of Nekhoroshev stability, which span from 10^5 to 10^7 satellite orbital periods. These times are sufficiently high for applications related to the operational lifetime and end-of-life deployment of satellites, as well as to the long term orbital evolution of populations of space debris.

4. Finally, we examine a more extended model including the J_3 and J_4 harmonics of the Earth's potential as well as J_2^2 terms obtained by second order averaging of the Hamiltonian in closed form. While the complexity of the new model renders a full investigation of this extended model beyond our present scope, we provide some key comparisons with our basic (only J_2) model: i) the domains of stability remain practically the same in the two models, while ii) the times of stability are affected by about one order of magnitude at the lowermost limit of the MEO zone, a difference tending nevertheless to vanish as a increases. Section 5.5 discusses in detail the origin of these differences.

As a final remark, we should emphasize again that our study was limited only to the case in which the frequencies of motion satisfy suitable non-resonant conditions. Another limitation is that we disregarded the slow precession of the lunar nodes with respect to the ecliptic plane, by simply considering a constant inclination of the Moon, equal to the one of the ecliptic. Notwithstanding the arguments presented in section 5.2 as regards the precision of this model (presently motivated mostly by our computational limits), we still emphasize that we consider the results presented in this chapter as a first step that paves the way to several future directions of research. Among possible future extentions, we indicate: i) exploring the application of the resonant Nekhoroshev's theorem, which becomes relevant for particular values of the inclination associated with lunisolar resonances, and ii) removing the assumption of a fixed ellipse for the Moon's orbit. These possibilities leave open that Nekhoroshev stability might actually hold in domains larger than the ones found here, extending to altitudes $a > 18000$ km were many satellites reside (e.g., GPS and geosynchronous satellites).

5.7 Analytical expressions of $\mathcal{H}_b^{(av)}$ and $\mathcal{H}^{(sec)}$ in Section 5.2

5.7.1 Expansion of $\mathcal{H}_b^{(av)}$

In this additional section we provide an expression of $\mathcal{H}_b^{(av)}$ for a third body (index b, referring to the Moon or Sun) as a function of its orbital parameters $(a_b, e_b, i_b, \omega_b, \Omega_b)$ and the debris' parameters $(a, e, i, \omega, \Omega)$. Up to second order in the eccentricity we have:

$$
\begin{aligned}
\mathcal{H}_b^{(av)} = \frac{a^2}{16 a_b^3 (1 - e_b^2)^{3/2}} \Bigg(& -\frac{2 + 3e^2}{8}(1 + 3\cos(2i))(1 + 3\cos(2i_0)) \\
& - \frac{15}{4} e^2 (1 + 3\cos 2i_0) \sin i^2 \cos 2\omega - \frac{3}{2}(2 + 3e^2) \sin i^2 \sin i_0^{\;2} \cos 2(\Omega - \Omega_{b_0}) \\
& - 15 e^2 \cos(i/2)^4 \sin i_0^{\;2} \cos 2(\omega + \Omega - \Omega_{b_0}) - \frac{3}{2}(2 + 3e^2)\sin(2i)\sin(2i_0)\cos(\Omega - \Omega_{b_0}) \\
& + 30 e^2 \cos(i/2)^3 \sin(i/2) \sin(2i_0) \cos(2\omega + \Omega - \Omega_{b_0}) \\
& + \frac{15}{2} e^2 (-1 + \cos i)\sin i \sin i_0 \cos(2\omega - \Omega + \Omega_{b_0}) \\
& - 15 e^2 \sin(i/2)^4 \sin(i_0)^2 \cos 2(\omega - \Omega + \Omega_{b_0}) \Bigg) .
\end{aligned}
$$

$$(5.7.1)$$

5.7.2 List of the nonzero terms in $\mathcal{H}^{(sec)}$ for $j = 1, 2$

Assuming, as in Section 5.2, that both the lunar and solar orbits lie on a fixed ecliptic plane inclined with respect to the Earth's equatorial plane by an angle i_0, the frequencies ω_1 and ω_2 appearing in (5.2.10) are given by:

$$
\omega_1 = \omega_1^{(J_2)} + \omega_1^{(M)} + \omega_1^{(\odot)} , \qquad\qquad \omega_2 = \omega_2^{(J_2)} + \omega_2^{(M)} + \omega_2^{(\odot)} ,
$$

where

$$
\begin{aligned}
\omega_1^{(J_2)} &= -\frac{3}{4} R_E^2 J_2 \mu_E^4 \frac{(-1 + 5\cos i_*^{\;2} - 2\cos i_*)}{(\mu_E a)^{7/2}(1 - e_*^2)^2} \\
\omega_2^{(J_2)} &= \frac{3}{2} \frac{R_E^2 J_2 \mu_E^4}{(\mu_E a)^{7/2}(1 - e_*^2)^2} \cos i_* \\
\omega_1^{(M/\odot)} &= -\frac{3}{64} a^{3/2} \mu_{M/\odot} \frac{[3 + 2e_*^2 - 2(2 + 3e_*^2)\cos i_* + 5\cos 2i_*](1 + 3\cos 2i_0)}{\sqrt{1 - e_*^2}\sqrt{\mu_E}(a_{M/\odot}(1 - e_{M/\odot}))^3} \\
\omega_2^{(M/\odot)} &= \frac{3}{32} a^{3/2} \mu_{M/\odot} \frac{(2 + 3e_*^2)\cos i_* \, (1 + 3\cos 2i_0)}{\sqrt{1 - e_*^2}\sqrt{\mu_E}\left(a_{M/\odot}(1 - e_{M/\odot})\right)^3}.
\end{aligned}
$$

$$(5.7.2)$$

The coefficients $a_{\mathbf{l}}$ and $b_{\mathbf{l},\mathbf{k}}$ in (5.2.15) are given by:

$$
\begin{aligned}
a_{(2,0)} =& \frac{3}{4} \frac{J_2 R_E^2}{a^4 (1 - e_*^2)^{5/2}} (1 + 10 \cos i_* - 15 \cos^2 i_*) \\
& - \frac{3}{128} \frac{a}{\mu_E (1 - e_*^2)} \left(\frac{\mu_M}{R_M^3} + \frac{\mu_\odot}{R_\odot^3} \right) (1 + 3 \cos 2i_0)(21 + 4e_*^2 - 40 \cos i_* + 15 \cos 2i_*) \\
a_{(1,1)} =& \frac{3}{2} \frac{J_2 R_E^2}{a^4 (1 - e_*^2)^{5/2}} (5 \cos i_* - 1) \\
& - \frac{3}{32} \frac{a}{\mu_E (1 - e_*^2)} \left(\frac{\mu_M}{R_M^3} + \frac{\mu_\odot}{R_\odot^3} \right) (1 + 3 \cos 2i_0)(2 + 3e_*^2 - 10 \cos i_*) \\
a_{(0,2)} =& -\frac{3}{4} \frac{J_2 R_E^2}{a^4 (1 - e_*^2)^{5/2}} \\
& - \frac{3}{64} \frac{a}{\mu_E (1 - e_*^2)} \left(\frac{\mu_M}{R_M^3} + \frac{\mu_\odot}{R_\odot^3} \right) (1 + 3 \cos 2i_0) \left(2 + 3e_*^2 \right)
\end{aligned}
$$

$$
\begin{aligned}
b_{(0,0),(\pm 2,0)} =& -\frac{15}{32} (a^2 e_*^2 \sin^2 i_0 \cos^4 (i_*/2)) \left(\frac{\mu_M}{r_M^3} + \frac{\mu_\odot}{r_\odot^3} \right) \\
b_{(0,0),(\pm 2,\pm 1)} =& \frac{15}{16} \left[a^2 e_*^2 \sin (2i_0) \cos^3 (i_*/2) \sin (i_*/2) \right] \left(\frac{\mu_M}{r_M^3} + \frac{\mu_\odot}{r_\odot^3} \right) \\
b_{(0,0),(\pm 2,\mp 2)} =& -\frac{15}{128} \left[a^2 e_*^2 (i + 3 \cos (2i_0)) \sin^2 i_* \right] \left(\frac{\mu_M}{r_M^3} + \frac{\mu_\odot}{r_\odot^3} \right) \\
b_{(0,0),(\pm 2,\mp 3)} =& -\frac{15}{16} \left[a^2 e_*^2 \sin (2i_0) \sin^3 (i_*/2) \cos (i_*/2) \right] \left(\frac{\mu_M}{r_M^3} + \frac{\mu_\odot}{r_\odot^3} \right) \\
b_{(0,0),(\pm 2,\mp 4)} =& -\frac{15}{32} \left[a^2 e_*^2 \sin^2 i_0 \sin^4 (i_*/2) \right] \left(\frac{\mu_M}{r_M^3} + \frac{\mu_\odot}{r_\odot^3} \right) \\
b_{(0,0),(0,\pm 1)} =& -\frac{3}{64} \left[a^2 (2 + 3e_*^2) \sin (2i_0) \cos (2i_*) \right] \left(\frac{\mu_M}{r_M^3} + \frac{\mu_\odot}{r_\odot^3} \right) \\
b_{(0,0),(0,\pm 2)} =& -\frac{3}{64} \left[a^2 (2 + 3e_*^2) \sin^2 i_0 \sin^2 i_* \right] \left(\frac{\mu_M}{r_M^3} + \frac{\mu_\odot}{r_\odot^3} \right) .
\end{aligned}
$$

References

[1] Irene De Blasi and Susanna Terracini. Refraction periodic trajectories in central mass galaxies. *Nonlinear Anal.*, 218:Paper No. 112766, 40, 2022.

[2] Irene De Blasi and Susanna Terracini. On some refraction billiards. *To appear in Discrete and Continuous Dynamical Systems*, 2021.

[3] Vivina Barutello, Irene De Blasi, and Susanna Terracini. Symbolic dynamics and analytical non-integrability for a galactic billiard, In preparation, 2022.

[4] Irene De Blasi, Alessandra Celletti, and Christos Efthymiopoulos. Semi-analytical estimates for the orbital stability of earth's satellites. *Journal of Nonlinear Science*, 31(6):1–37, 2021.

[5] Alessandra Celletti, Irene De Blasi, and Christos Efthymiopoulos. Nekhoroshev estimates for satellites' orbital stability. *Preprint*, September 2022.

[6] Irene De Blasi, Alessandra Celletti, and Christos Efthymiopoulos. Satellites' orbital stability through normal forms. *Proceedings of the International Astronomical Union*, 15(S364):146–151, 2021.

[7] Airi Takeuchi and Lei Zhao. Conformal transformations and integrable mechanical billiards. *arXiv preprint arXiv:2110.03376*, 2021.

[8] Giovanni Gallavotti and Ian Jauslin. A theorem on ellipses, an integrable system and a theorem of boltzmann. *arXiv preprint arXiv:2008.01955*, 2020.

[9] Nick Delis, Christos Efthymiopoulos, and Constantinos Kalapotharakos. Effective power-law dependence of Lyapunov exponents on the central mass in galaxies. *Monthly Notices of the Royal Astronomical Society*, 448(3):2448–2468, 2015.

[10] Constantinos Kalapotharakos. The rate of secular evolution in elliptical galaxies with central masses. *Monthly Notices of the Royal Astronomical Society*, 389(4):1709–1721, 2008.

[11] Victor G. Baryakhtar, Vladimir Yanovsky, Sergei Naydenov, and A.V. Kurilo. Chaos in composite billiards. *Journal of Experimental and Theoretical Physics*, 103(2):292–302, 2006.

[12] Dentcho Genov, Shuang Zhang, and Xiang Zhang. Mimicking Celestial Mechanics in metamaterials. *Nature Physics*, 5(9):687–692, 2009.

[13] Wen Xiao, Sicen Tao, and Huanyang Chen. Mimicking the gravitational effect with gradient index lenses in geometrical optics. *Photonics Research*, 9(7):1197–1203, 2021.

[14] Irene Cavallari, Giovanni F Gronchi, and Giulio Baú. On the sun-shadow dynamics. *Physica D: Nonlinear Phenomena*, 432:133136, 2022.

[15] Sean Gasiorek. *On the dynamics of inverse magnetic billiards.* PhD thesis, University of California at Santa Cruz, 2019.

[16] George D. Birkhoff. *Dynamical systems*, volume 9. American Mathematical Soc., 1927.

[17] Serge Tabachnikov. *Geometry and billiards*, volume 30 of *Student Mathematical Library.* American Mathematical Society, Providence, RI; Mathematics Advanced Study Semesters, University Park, PA, 2005.

[18] Misha Bialy, Corentin Fierobe, Alexey Glutsyuk, Mark Levi, Alexander Plakhov, and Serge Tabachnikov. Open problems on billiards and geometric optics. *Arnold Mathematical Journal*, pages 1–12, 2022.

[19] Vadim Kaloshin and Alfonso Sorrentino. On the integrability of Birkhoff billiards. *Philos. Trans. Roy. Soc. A*, 376(2131):20170419, 16, 2018.

[20] Vadim Kaloshin and Alfonso Sorrentino. On the local Birkhoff conjecture for convex billiards. *Ann. of Math. (2)*, 188(1):315–380, 2018.

[21] Guan Huang, Vadim Kaloshin, and Alfonso Sorrentino. Nearly circular domains which are integrable close to the boundary are ellipses. *Geom. Funct. Anal.*, 28(2):334–392, 2018.

[22] Jürgen Möser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen, II*, pages 1–20, 1962.

[23] Serge Aubry and Pierre-Yves Le Daeron. The discrete Frenkel-Kontorova model and its extensions: I. Exact results for the ground-states. *Physica D: Nonlinear Phenomena*, 8(3):381–422, 1983.

[24] John Mather. Existence of quasi-periodic orbits for twist homeomorphisms of the annulus, topol-ogy21 (1982), 457-467. *MR 84g*, 58084.

[25] Christophe Golé. *Symplectic twist maps: global variational techniques*, volume 18. World Scientific, 2001.

[26] Tullio Levi-Civita. Sur la résolution qualitative du problème restreint des trois corps. *Acta Math.*, 30(1):305–327, 1906.

[27] Vladimir F. Lazutkin. The existence of caustics for a billiard problem in a convex domain. *Mathematics of the USSR-Izvestiya*, 7(1):185, 1973.

[28] Raphael Douady. Applications du théoreme des tores invariants. These de 3ème cycle. *Université de Paris VII*, 1982.

[29] Robert L. Devaney. *An introduction to chaotic dynamical systems.* Studies in Nonlinearity. Westview Press, Boulder, CO, 2003. Reprint of the second (1989) edition.

[30] Boris Hasselblatt and Anatole Katok. *A first course in dynamics.* Cambridge University Press, New York, 2003. With a panorama of recent developments.

[31] Nicola Soave. Symbolic dynamics: from the n-centre to the (n+ 1)-body problem, a preliminary study. *Nonlinear Differential Equations and Applications NoDEA*, 21(3):371–413, 2014.

[32] Vivina Barutello, Gian Marco Canneori, and Susanna Terracini. Symbolic dynamics for the anisotropic N-centre problem at negative energies. *Arch. Ration. Mech. Anal.*, 242(3):1749–1834, 2021.

[33] Carlo Miranda. Un'osservazione su un teorema di Brouwer. *Boll. UMI*, 3:5–7, 1940.

[34] Valery V. Kozlov. Integrability and nonintegrability in Hamiltonian mechanics. *Uspekhi Mat. Nauk*, 38(1(229)):3–67, 240, 1983.

[35] Alan B. Jenkin, John P. McVey, and Marlon E. Sorge. Orbital lifetime and collision risk reduction for inclined geosynchronous disposal orbits. *Acta Astronautica*, 161:153–165, August 2019.

[36] Seong-Hyeon Park, Hae-Dong Kim, and Gisu Park. Orbit, orbital lifetime, and reentry survivability estimation for orbiting objects. *Advances in Space Research*, 62(11):3012–3032, December 2018.

[37] Ha-Yeon Choi, Hae-Dong Kim, and Jae-Dong Seong. Analysis of Orbital Lifetime Prediction Parameters in Preparation for Post-Mission Disposal. *Journal of Astronomy and Space Sciences*, 32(4):367–377, December 2015.

[38] ESA's annual space environment report. Technical report, European Space Agency, Space Debris Office, 04 2022.

[39] Alessandra Celletti and Cătălin Galeş. On the dynamics of space debris: 1: 1 and 2: 1 resonances. *Journal of Nonlinear Science*, 24(6):1231–1262, 2014.

[40] Alessandra Celletti and Cătălin Galeş. Dynamical investigation of minor resonances for space debris. *Celestial Mechanics and Dynamical Astronomy*, 123(2):203–222, 2015.

[41] Alessandra Celletti and Cătălin Galeş. Dynamics of resonances and equilibria of low Earth objects. *SIAM Journal on Applied Dynamical Systems*, 17(1):203–235, 2018.

[42] Alessandra Celletti, Christos Efthymiopoulos, Fabien Gachet, Cătălin Galeş, and Giuseppe Pucacco. Dynamical models and the onset of chaos in space debris. *International Journal of Non-Linear Mechanics*, 90:147–163, 2017.

[43] Ioannis Gkolias and Camilla Colombo. Towards a sustainable exploitation of the geosynchronous orbital region. *Celestial Mechanics and Dynamical Astronomy*, 131(4):19, April 2019.

[44] Holger Krag, Stijn Lemmens, Tim Flohrer, and Heiner Klinkard. Analysing Global Achievements in Orbital Lifetime Reduction at the End of LEO Missions. In *6th European Conference on Space Debris*, volume 723 of *ESA Special Publication*, page 76, August 2013.

[45] William M. Kaula. Theory of satellite geodesy, Blaisdell publ. *Co., Waltham, Mass*, 1966.

[46] Jérôme Daquin, Aaron J. Rosengren, Elisa Maria Alessi, Florent Deleflie, Giovanni B. Valsecchi, and Alessandro Rossi. The dynamical structure of the meo region: long-term stability, chaos, and transport. *Celestial Mechanics and Dynamical Astronomy*, 124(4):335–366, 2016.

[47] Ioannis Gkolias, Jérôme Daquin, Fabien Gachet, and Aaron J. Rosengren. From order to chaos in Earth satellite orbits. *The Astronomical Journal*, 152(5):119, 2016.

[48] Jeffrey M. Aristoff, Joshua T. Horwood, and Kyle T. Alfriend. On a set of J_2 equinoctial orbital elements and their use for uncertainty propagation. *Celestial Mechanics and Dynamical Astronomy*, 133(3):1–19, 2021.

[49] Tao Nie and Pini Gurfil. Long-term evolution of orbital inclination due to third-body inclination. *Celestial Mechanics and Dynamical Astronomy*, 133(1):1–33, 2021.

[50] Fabien Gachet, Alessandra Celletti, Giuseppe Pucacco, and Christos Efthymiopoulos. Geostationary secular dynamics revisited: application to high area-to-mass ratio objects. *Celestial Mechanics and Dynamical Astronomy*, 128(2-3):149–181, 2017.

[51] Christos Efthymiopoulos. Canonical perturbation theory; stability and diffusion in Hamiltonian systems: applications in dynamical astronomy. In *Workshop Series of the Asociacion Argentina de Astronomia*, volume 3, pages 3–146, 2011.

[52] Nikolai N. Nekhoroshev. Stable lower estimates for smooth mappings and for gradients of smooth function. *Matematicheskii Sbornik*, 132(3):432–478, 1973.

[53] Nikolai N. Nekhoroshev. An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems. *Uspekhi Matematicheskikh Nauk*, 32(6):5–66, 1977.

[54] Jürgen Pöschel. Nekhoroshev estimates for quasi-convex Hamiltonian systems. *Mathematische Zeitschrift*, 213(1):187–216, 04 1993.

[55] Luigi Chierchia, Maria Elisabetta Faraggiana, and Massimiliano Guzzo. On steepness of 3-jet non-degenerate functions. *Annali di Matematica Pura ed Applicata (1923-)*, pages 1–15, 2018.

[56] Gabriella Schirinzi and Massimiliano Guzzo. On the formulation of new explicit conditions for steepness from a former result of N.N. Nekhoroshev. *Journal of Mathematical Physics*, 54(7):072702, 2013.

[57] Santiago Barbieri. On the algebraic properties of exponentially stable integrable Hamiltonian systems. *arXiv preprint arXiv:2011.09731*, 2020.

[58] Alessandra Celletti and Laura Ferrara. An application of the Nekhoroshev theorem to the restricted three-body problem. *Celestial Mechanics and Dynamical Astronomy*, 64(3):261–272, 1996.

[59] Antonio Giorgilli and Charalambos Skokos. On the stability of the Trojan asteroids. *Astronomy and Astrophysics*, 317:254–261, 1997.

[60] Daniel Steichen and Antonio Giorgilli. Long time stability for the main problem of artificial satellites. *Celestial Mechanics and Dynamical Astronomy*, 69(3):317–330, 1997.

[61] Mikhail L. Lidov. The evolution of orbits of artificial satellites of planets under the action of gravitational perturbations of external bodies. *planss*, 9:719–759, 1962.

[62] S. Hughes. Earth satellite orbits with resonant lunisolar perturbations I. Resonances dependent only on inclination. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 372(1749):243–264, 1980.

[63] Harish N.S. Krishnamoorthy, Zubin Jacob, Evgenii Narimanov, Ilona Kretzschmar, and Vinod M. Menon. Topological transitions in metamaterials. *Science*, 336(6078):205–209, 2012.

[64] Herbert Seifert. Periodische Bewegungen mechanischer Systeme. *Math. Z.*, 51:197–216, 1948.

[65] Morris W. Hirsch, Stephen Smale, and Robert L. Devaney. *Differential equations, dynamical systems, and an introduction to chaos.* Elsevier/Academic Press, Amsterdam, third edition, 2013.

[66] Misha Bialy, Andrey E. Mironov, and Serge Tabachnikov. Wire billiards, the first steps. *Adv. Math.*, 368:107154, 27, 2020.

[67] Paul Glendinning. Geometry of refractions and reflections through a biperiodic medium. *SIAM J. Appl. Math.*, 76(4):1219–1238, 2016.

[68] Manfredo P. Do Carmo. *Differential geometry of curves and surfaces: revised and updated second edition.* Courier Dover Publications, 2016.

[69] Nicola Soave and Susanna Terracini. Symbolic dynamics for the N-centre problem at negative energies. *Discrete Contin. Dyn. Syst.*, 32(9):3245–3301, 2012.

[70] Alessandra Celletti. *Stability and chaos in Celestial Mechanics.* Springer-Praxis, 2010.

[71] Vivina Barutello, Alberto Boscaggin, and Walter Dambrosio. On the minimality of Keplerian arcs with fixed negative energy. *Qual. Theory Dyn. Syst.*, 19(1):Paper No. 42, 21, 2020.

[72] Richard Montgomery. Minimizers for the Kepler problem. *Qual. Theory Dyn. Syst.*, 19(1):Paper No. 31, 12, 2020.

[73] Walter D. Neumann. Generalizations of the Poincaré Birkhoff fixed point theorem. *Bulletin of the Australian Mathematical Society*, 17(3):375–389, 1977.

[74] Artur Avila, Jacopo De Simoi, and Vadim Kaloshin. An integrable deformation of an ellipse of small eccentricity is an ellipse. *Annals of Mathematics*, pages 527–558, 2016.

[75] Richard H. Battin. *An introduction to the mathematics and methods of astrodynamics.* AIAA Education Series. American Institute of Aeronautics and Astronautics (AIAA), Reston, VA, revised edition, 1999. With a foreword by J. S. Przemieniecki.

[76] Sergey V. Bolotin and Robert S. Mackay. Periodic and chaotic trajectories of the second species for the n-centre problem. *Celestial Mechanics and Dynamical Astronomy*, 77(1):49–75, 2000.

[77] Moser Jürgen. Recent developments in the theory of Hamiltonian systems. *SIAM review*, 28(4):459–485, 1986.

[78] Gianmarco Capitanio. On geodesic envelopes and caustics. *arXiv preprint math/0411430*, 2004.

[79] Helmut Rüssmann. *Kleine Nenner. 1. Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes.* Vandenhoeck & Ruprecht, 1970.

[80] Michael R. Herman and Albert Fathi. *Sur les courbes invariantes par les difféo-morphismes de l'anneau*, volume 1. Société mathématique de France, 1983.

[81] Valeriĭ V. Kozlov and Dmitriĭ V. Treshchëv. *Billiards*, volume 89 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991. A genetic introduction to the dynamics of systems with impacts, Translated from the Russian by J. R. Schulenberger.

[82] Wang Sang Koon, Martin W. Lo, Jerrold E. Marsden, and Shane D. Ross. Heteroclinic connections between periodic orbits and resonance transitions in celestial mechanics. *Chaos*, 10(2):427–469, 2000.

[83] Zongcheng Li. Chaos induced by heteroclinic cycles connecting repellers and saddles in locally compact metric spaces. *Nonlinear Anal.*, 71(3-4):1379–1388, 2009.

[84] Xiaoying Wu. Heteroclinic cycles imply chaos and are structurally stable. *Discrete Dyn. Nat. Soc.*, pages Art. ID 6647132, 7, 2021.

[85] Lei Zhao. Projective dynamics and an integrable Boltzmann billiard model. *Communications in Contemporary Mathematics*, page 2150085, Sep 2021.

[86] Airi Takeuchi and Lei Zhao. Projective integrable mechanical billiards. *arXiv preprint arXiv:2203.12938*, 2022.

[87] Sergey V. Bolotin. Nonintegrability of the problem of n centers for $n > 2$. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, 3:65–68, 1984.

[88] Sergey V. Bolotin and Piero Negrini. Regularization and topological entropy for the spatial n-center problem. *Ergodic Theory Dynam. Systems*, 21(2):383–399, 2001.

[89] Sergey V. Bolotin and Piero Negrini. Chaotic behavior in the 3-center problem. *J. Differential Equations*, 190(2):539–558, 2003.

[90] Sergey V. Bolotin and Valery V. Kozlov. Topological approach to the generalized n-centre problem. *Uspekhi Mat. Nauk*, 72(3(435)):65–96, 2017.

[91] A. Knauf and I.A. Taĭmanov. On the integrability of the n-centre problem. *Math. Ann.*, 331(3):631–649, 2005.

[92] M. Klein and A. Knauf. Chaotic motion in Coulombic potentials. In *Mathematical physics, X (Leipzig, 1991)*, pages 308–312. Springer, Berlin, 1992.

[93] A. Knauf. The n-centre problem of celestial mechanics for large energies. *J. Eur. Math. Soc. (JEMS)*, 4(1):1–114, 2002.

[94] A. Knauf and I.A. Taĭmanov. Integrability of the n-center problem at high energies. *Dokl. Akad. Nauk*, 397(1):20–22, 2004.

[95] Alessandra Celletti, Cătălin Galeş, and Christoph Lhotka. Resonances in the Earth's space environment. *Communications in Nonlinear Science and Numerical Simulations*, 84:105185, May 2020.

[96] Giulia Schettino, Elisa Maria Alessi, Alessandro Rossi, and Giovanni B. Valsecchi. A frequency portrait of Low Earth Orbits. *Celestial Mechanics and Dynamical Astronomy*, 131(8):35, August 2019.

[97] Desmond G. King-Hele and Doreen M. C. Walker. Predicting the orbital lifetimes of Earth satellites. *Acta Astronautica*, 18:123–131, January 1988.

[98] Lin Liu and Xin Wang. On the orbital lifetime of high-altitude satellites. *Chinese Astronomy and Astrophysics*, 24(3):284–288, January 2000.

[99] Barbara E. Shute and Janice Chiville. The lunar-solar effect on the orbital lifetimes of artificial satellites with highly eccentric orbits. *Planetary and Space Science*, 14(4):361–369, April 1966.

[100] H. Robert Westerman. On satellite orbit lifetimes. *Astron. J.*, 68:385, August 1963.

[101] Alessandra Celletti and Cătălin Galeş. On the dynamics of space debris: 1:1 and 2:1 resonances. *J. Nonlinear Science*, 24(6):1231–1262, 2014.

[102] Giorgio E.O. Giacaglia. Lunar perturbations of artificial satellites of the Earth. *Celest. Mech.*, 9:239–267, 1974.

[103] William M. Kaula. Development of the lunar and solar disturbing functions for a close satellite. *Astron. J.*, 67:300–303, 1962.

[104] Mark T. Lane. On analytic modeling of lunar perturbations of artificial satellites of the Earth. *Celest. Mech. Dynam. Astr.*, 46(4):287–305, 1989.

[105] Slawomir Breiter. Lunisolar resonances revisited. *Celest. Mech. Dyn. Astr.*, 81:81–91, 2001.

[106] Slawomir Breiter. On the coupling of lunisolar resonances for Earth satellite orbits. *Celestial Mechanics and Dynamical Astronomy*, 80(1):1–20, 2001.

[107] Alessandra Celletti, Cătălin Galeş, and Giuseppe Pucacco. Bifurcation of lunisolar secular resonances for space debris orbits. *SIAM J. Appl. Dyn. Syst.*, 15:1352–1383, 2016.

[108] Alessandra Celletti, Cătălin Galeş, Giuseppe Pucacco, and Aaron Rosengren. Analytical development of the lunisolar disturbing function and the critical inclination secular resonance. *Celest. Mech. Dyn. Astron.*, 127(3):259–283, 2017.

[109] Chia-Chun Chao and R. Anne Gick. Long-term evolution of navigation satellite orbits: Gps/glonass/galileo. *Advances in Space Research*, 34:1221–1226, 2004.

[110] G. E. Cook. Luni-solar perturbations of the orbit of an Earth satellite. *Geophysical Journal International*, 6(3):271–291, 1962.

[111] Todd A. Ely and Kathleen C. Howell. Dynamics of artificial satellite orbits with tesseral resonances including the effects of luni-solar perturbations. *Dynamics and Stability of Systems*, 12(4):243–269, 1997.

[112] S. Hughes. Earth satellite orbits with resonant lunisolar perturbations. I. Resonances dependent only on inclination. *Proc. R. Soc. Lond. A*, 372:243–264, 1980.

[113] Anne Lemaître, Nicolas Delsate, and Stéphane Valk. A web of secondary resonances for large A/m geostationary debris. *Celest. Mech. Dyn. Astr.*, 104:383–402, 2009.

[114] Elisa M. Alessi, Florent Deleflie, Aaron J. Rosengren, Alessandro Rossi, Giovanni B. Valsecchi, Jérôme Daquin, and Klaus Merz. A numerical investigation on the eccentricity growth of GNSS disposal orbits. *Celest. Mech. Dyn. Astr.*, 125(1):71–90, 2016.

[115] Daniel Casanova, Alexis Petit, and Anne Lemaître. Long-term evolution of space debris under the J_2 effect, the solar radiation pressure and the solar and lunar perturbations. *Celest. Mech. Dyn. Astr.*, 123:223–238, 2015.

[116] Aaron J. Rosengren and Daniel J. Scheeres. Long-term dynamics of high area-to-mass ratio objects in high-Earth orbit. *Adv. Space Res.*, 52:1545–1560, 2013.

[117] Aaron J. Rosengren, Elisa M. Alessi, Alessandro Rossi, and Giovanni B. Valsecchi. Chaos in navigation satellite orbits caused by the perturbed motion of the moon. *Mon. Not. R. Astron. Soc.*, 449:3522–3526, 2015.

[118] Aaron J. Rosengren, Jérôme Daquin, Kleomenis Tsiganis, Elisa M. Alessi, Florent Deleflie, Alessandro Rossi, and Giovanni B. Valsecchi. Galileo disposal strategy: stability, chaos and predictability. *Monthly Notices of the Royal Astronomical Society*, 464(4):4063–4076, 2016.

[119] Alessandro Rossi. Resonant dynamics of medium Earth orbits: space debris issues. *Celestial Mechanics and Dynamical Astronomy*, 100(4):267–286, 2008.

[120] Despoina K. Skoulidou, Aaron J. Rosengren, Kleomenis Tsiganis, and George Voyatzis. Medium Earth orbit dynamical survey and its use in passive debris removal. *Advances in Space Research*, 63(11):3646–3674, 2019.

[121] Stéphane Valk and Anne Lemaître. Analytical and semi-analytical investigations of geosynchronous space debris with high area-to-mass ratios. *Advances in Space Research*, 41:1077–1090, 2008.

[122] Dirk Brouwer. Solution of the problem of artificial satellite theory without drag. *Astron. J.*, 64:378–397, 1959.

[123] Carl D. Murray and Stanley F. Dermott. *Solar system dynamics*. Cambridge university press, 1999.

[124] Aaron J. Rosengren, Despoina K. Skoulidou, Kleomenis Tsiganis, and George Voyatzis. Dynamical cartography of earth satellite orbits, submitted to adv. *Space Res*, 2018.

[125] Alessandra Celletti, Cătălin Galeş, Giuseppe Pucacco, and Aaron J. Rosengren. Analytical development of the lunisolar disturbing function and the critical inclination secular resonance. *Celestial Mechanics and Dynamical Astronomy*, 127(3):259–283, 2017.

[126] Aaron J. Rosengren, Daniel J. Scheeres, and Jay W. McMahon. The classical laplace plane as a stable disposal orbit for geostationary satellites. *Advances in Space Research*, 53(8):1219–1228, 2014.

[127] Christos Efthymiopoulos, Antonio Giorgilli, and George Contopoulos. Nonconvergence of formal integrals: Ii. improved estimates for the optimal order of truncation. *Journal of Physics A: Mathematical and General*, 37(45):10831, 2004.

[128] Francesco Fassò and Giancarlo Benettin. Composition of Lie transforms with rigorous estimates and applications to hamiltonian perturbation theory. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 40(3):307–329, 1989.

[129] Antonio Giorgilli. Notes on exponential stability of hamiltonian systems. 2002.

[130] Yoshihide Kozai. Secular perturbations of asteroids with high inclination and eccentricity. *Astron. J.*, 67:591–598, November 1962.

[131] Zoran Knezevic and Rade Pavlovic. Application of the Nekhoroshev theorem to the real dynamical system. *Novi Sad J. Math*, 38(3):181–188, 2008.

[132] Dario Bambusi and Alessandra Fusè. Nekhoroshev theorem for perturbations of the central motion. *Regular and Chaotic Dynamics*, 22(1):18–26, 2017.

[133] Gabriella Schirinzi and Massimiliano Guzzo. Numerical verification of the steepness of three and four degrees of freedom Hamiltonian systems. *Regular and Chaotic Dynamics*, 20(1):1–18, 2015.

[134] Vladimir I. Arnold. Instability of dynamical systems with several degrees of freedom. *Sov. Math. Doklady*, 5:581–585, 1964.

[135] Giancarlo Benettin and Giovanni Gallavotti. Stability of motions near resonances in quasi-integrable Hamiltonian systems. *Journal of statistical physics*, 44(3-4):293–338, 1986.

[136] Alessandra Celletti and Laura Ferrara. An application of Nekhoroshev theorem to the restricted three-body problem. *Celest. Mech. Dyn. Astr.*, 64:261–272, 1996.

[137] Alessandra Celletti and Antonio Giorgilli. On the stability of the Lagrangian points in the spatial restricted problem of three bodies. *Cel. Mech. Dyn. Astr.*, 50:31–58, 1991.

[138] Antonio Giorgilli and Charalampos Skokos. On the stability of the Trojan asteroids. *Astron. Astrophys.*, 317:254–261, 1997.

[139] Martin Lara. *Hamiltonian Perturbation Solutions for Spacecraft Orbit Prediction.* De Gruyter Studies in Mathematical Physics, 2021.

[140] Tao Nie and Pini Gurfil. Long-term evolution of orbital inclination due to third-body inclination. *Celest. Mech. Dyn. Astr.*, 133(1), 2021.

[141] André Deprit. Delaunay Normalisations. *Celestial Mechanics*, 26(1):9–21, January 1982.

[142] Martin Lara, Juan F. San-Juan, and Luis M. López-Ochoa. Delaunay variables approach to the elimination of the perigee in artificial satellite theory. *Celest. Mech. Dyn. Astron.*, 120(1):39–56, 2014.

[143] Martin Lara, Rosario López, Ivàn Pérez, and Juan F. San-Juan. Exploring the long-term dynamics of perturbed keplerian motion in high degree potential fields. *Communications in Nonlinear Science and Numerical Simulation*, 82:105053, 2020.

[144] Jesus Palacián. Normal forms for perturbed Keplerian systems. *Journal of Differential Equations*, 180(2):471–519, 2002.

[145] Jesus Palacián and Patricia Yanguas. Simplification of Perturbed Hamiltonians Through Lie Transformations. In J. Delgado, E. A. Lacomba, E. Pérez-Chavela, and J. Llibre, editors, *Hamiltonian Systems and Celestial Mechanics (HAMSYS-98)*, pages 284–302, October 2000.