The role of hyperbolic invariant objects: From Arnold diffusion to biological clocks

Gemma Huguet Casades
Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya

Advisors: Amadeu Delshams and Antoni Guillamon

Al Ferran
i a la meva família

Agraïments

Durant els anys que he estat treballant en aquesta tesi he tingut la sort de conèixer i treballar amb molta gent interessant que m'han ajudat a madurar tant des del punt de vista científic com humà i als quals m'agradaria expressar la meva més sincera gratitud.

En primer lloc, m'agradaria agrair als meus directors de tesi, l'Amadeu i el Toni, les seves ensenyances i totes les hores dedicades pacientment, sovint a molts quilòmetres de distància i en horaris estranys, així com l'enorme confiança que han dipositat en mi.

I és que els que m'han seguit de prop saben prou bé que aquesta tesi l'he feta a cavall entre Barcelona i Austin. D'això n'és, en part, responsable el Professor Rafael de la Llave, afectuosament el Rafa. És difícil expressar amb paraules el meu agraïment per la seva inestimable ajuda, tant a nivell acadèmic com personal. Gràcies per donar-me l'oportunitat de venir a Austin i poder col·laborar amb tant bons matemàtics. El Capítol 2 d'aquesta tesi és fruit d'aquestes nombroses estades i l'hem fet en col·laboració amb el Professor Yannick Sire. Yannick, merci pour l'amitié et l'appui constant.

I would like to thank also Professor David Terman to give me the opportunity to visit the Mathematical Bioscience Institute at The Ohio State University, for whose hospitality I am very grateful. Thank you very much for the discussions on computational neuroscience and also for making my stay in Columbus so pleasant and fruitful. Thanks also to Professor Marian Gidea for many discussions and several invitations to Chicago.

Als companys del grup de Sistemes Dinàmics i als companys del departament de Matemàtica Aplicada I, per crear un ambient tant agradable per treballar i aprendre. Gràcies a les noies: l'Ari, la Eli, l'Estrella i la Maria. Moltes gràcies a la Professora Tere M-Seara, a la Tere, per l'interès i l'ajut que sempre ha mostrat per la meva tesi i per les seves paraules d'ànims quan les he necessitades. Moltes gràcies també als companys de despatx, l'Alejandro, el Marcel i el Martí per tantes paraules d'ànim i tantes discussions de matemàtiques. Also thanks to my friends at the Department of Mathematics at UT Austin for making my stays in Austin so pleasant and enjoyable. Y en especial, a Renato, por su generosa acogida y sincera amistad.

A la meva família i, en especial, als meus pares, gràcies pel seu suport i ànims. I també als meus amics, als quals he hagut de robar estones del nostre temps.

Finalment m'agradaria agrair al Ferran per la seva tendresa i vitalitat, per empènyer-me sempre a lluitar pel que vull i no deixar-me desanimar mai. Moltíssimes gràcies per haver-me acompanyat en aquest viatge.

Oficialment, em cal agrair la beca FPU del Ministeri d'Educació i Ciència, així com el projecte MTM2006-00478. La tesi la vaig acabar a Austin amb una plaça de *Research Engineering Student Assistant* finançada per la NSF Grant DMS 0354567.

Contents

Introduction

The framework for this thesis are the hyperbolic invariant objects (whiskered tori, limit cycles, NHIM, ...), which are the essential objects to study different problems ranging from Arnold diffusion to biological clocks. We deal with three different related topics and the approach is both theoretical and numeric with special attention to applications, specially in neurobiology:

- Existence of Arnold diffusion for a priori unstable Hamiltonian systems.
- Fast numerical algorithms to compute invariant tori and the associated whiskers in Hamiltonian systems using a parameterization method.
- Computation of Isochrons and Phase Resetting Curves (PRC) in neurobiological systems using a parameterization method.

The problems presented here cover a wide spectrum in the field of dynamical systems and they are all relevant problems in the corresponding area. Indeed, the problem of Arnold diffusion has been studied in the last ten years by important groups of researchers and mathematicians all over the world (Cheng, Gidea, Kaloshin, de la Llave, Mather, Neishtadt, Treschev, ...) as well as by several researchers in the UB-UPC dynamical systems group (Delshams, Fontich, Martín, Seara, Simó,...) and a large variety of methods leading to important results have been developed [DLS06a, Tre04, GL06b, GL06a, CY04a, CY04b, FM01, Sim94, SV01]. The numerical computation of the hyperbolic invariant objects present in the system, as well as their possible connections, constitutes one of the most important areas of research in the field of dynamical systems because these organize the dynamics around them. It is worth mentioning in this sense the contributions of the group of dynamical systems in Barcelona (de la Llave, Gómez, Haro, Jorba, Masdemont, Mondelo, Ollé, Simó, ...), specially in the applications to celestial mechanics and atrodynamics [HL07, Sim99, GJSM01b, GJSM01a]. Finally, the applications of dynamical systems to biology constitute a very leading area in the current research, as well a new and very interesting field to explore. The computation of PRC and its relation with the dynamical system behind the model has been explored in the recent years, specially in neuroscience (Ermentrout, Kopell, Josic, Shea-Brown, Holmes, Govaerts, Izhikevich, Mohelis, ..., see [Erm96, EK91, Izh07, Jos00]).

Interestingly enough, the problems studied in this thesis have in common the fact that they are tackled with a large set of non trivial and important tools as well as a bunch of consolidated methods in dynamical systems: hyperbolic invariant objects and their connections (limit cycles, isochrons, whiskered tori, NHIM, homoclinic orbits, ...),

normal forms, averaging methods, KAM theorem, Lie symmetries, ... It is important to remark that these techniques constitute a fundamental and wide background to solve relevant problems in a future research.

Moreover, it is worth mentioning that this thesis presents also a wide-ranging methodology since it contains both theory and numerics, just as a willingness of an interdisciplinary approach, with clear contributions to biology and, more precisely, to neuroscience. Hence, while the first part of the thesis is basically theoretical and uses analytical tools, the other two parts are essentially numeric even if they also contain analytical studies. The last part of the thesis arises from the purpose of relating biological clocks with some known tools in dynamical systems.

Indeed, a deep knowledge in dynamical systems both in theory and numerics, as well as a good preparation in neuroscience constitute a good platform to tackle some problems in neuroscience with tools that have never been applied to this area, leading to a more general and deep insight.

Next, we will discuss in more detail the problems studied in this thesis with the main contributions to each one.

Arnold diffusion for a priori unstable Hamiltonian systems

Hamiltonian systems appear naturally as models of many systems with negligible friction. We could mention, for instance, the models in celestial mechanics or the models for motion of charged particles in magnetic fields. One of the problems that appears naturally in the applications and has attracted the attention for a long time is whether the effect of periodic small perturbations on these models accumulate over time and lead to large effect (instability) or whether these effects average out (stability).

It is often the case that Hamiltonian systems exhibit both stable and unstable regimes. For instance, there exist rigorous results that prove that for the mentioned systems most of the trajectories are stable for all the time, it is the so well known KAM Theorem (Kolmogorov-Arnol'd-Moser) or they are stable for very long times (Nekhoroshev).

The orbits that do not lie on KAM invariant objects and evolve over time scales where Nekhoroschev theory breaks up may possibly drift arbitrarily far. Indeed, it was conjectured in [Arn63b] that this is a phenomenon that happens in rather general systems. In the celebrated paper [Arn64], Arnold constructed an example of a nearly integrable Hamiltonian system for which he proved the existence of trajectories that avoided the KAM tori and that performed long excursions. The mechanism is based on the existence of chains of whiskered tori such that the unstable manifold of one intersects the stable manifold of the next one (*transition chains*).

The example proposed turns out to be rather artificial because the perturbation has been chosen carefully so that it does not affect the foliation of invariant tori present in the unperturbed system, but it causes the stable and unstable manifolds to intersect transversally. In general, the perturbation destroys the foliation of persisting primary KAM tori creating gaps of size bigger than the splitting of the stable and unstable manifolds. In this case one can not form transition chains as in the Arnold mechanism. This is known

in the literature as the *large gap problem*.

Overcoming the *large gap problem* and identifying new mechanisms for diffusion has become an important direction of study itself. This problem has attracted the attention of both mathematicians and physicists due to its practical importance and mathematical depth. Recently, there has been a lot of progress in the mathematical theory of these problems and a wide variety of methods have been suggested.

One special case of the problem is the *a priori unstable* case, in which the unperturbed system presents hyperbolicity (it cannot be written in global action-angle variables). This is described by a Hamiltonian of the form

$$H_\varepsilon(p, q, I, \varphi, t) = \frac{1}{2}I^2 + \frac{1}{2}p^2 + (\cos q - 1) + \varepsilon h(p, q, I, \varphi, t; \varepsilon), \tag{1}$$

where $(p, q, I, \varphi, t) \in (\mathbb{R} \times \mathbb{T})^2 \times \mathbb{T}$.

The *large gap problem* for this type of Hamiltonian systems was solved (under certain conditions) through geometrical methods in [DLS00, DLS06a, DLS06b] (*scattering map*), as well as in [Tre04] (separatrix map); topological methods in [GL06b, GL06a] and variational methods in [CY04a, CY04b].

Furthermore, numerical experiments suggest that in the resonance zones the diffusion is more intense, see [Chi79]. Hence, a deep study of these zones turns out to be very important to understand the diffusion mechanism. Moreover, it is well known that the resonances play an important role not only in the problem of diffusion but also in other problems of physics and chemistry

In [DLS06a], the authors proved the existence of diffusing orbits for (1.4), when h is assumed to be a trigonometric polynomial in the angular variables (φ, t). The geometrical mechanism proposed is based on the incorporation in the transition chain of new invariant objects created by the perturbation. Indeed, one first identifies the normally hyperbolic invariant manifold (NHIM), which is an invariant object with an inner dynamics and associated stable and unstable manifolds. When one restricts to the dynamics on it, one can observe that new invariant objects, like secondary tori and invariant manifolds of lower dimensional tori, are created in the gaps between primary KAM tori and they are topologically different from the tori present in the unperturbed system. The *scattering map* of a NHIM, introduced by the same authors, is the essential tool for the heteroclinic connections of invariant objects of different topology.

This assumption on h is clearly non generic. In the first part of the thesis we generalize the result in [DLS06a] for the case of differentiable enough perturbations, whose Fourier series in the angular variables are not assumed to have a finite number of harmonics.

In order to prove this result we focused on the study of the resonant zones and the invariant objects generated therein. The method used for this study consists of a combination of a resonant averaging procedure in the NHIM and a KAM Theorem. Indeed, we define different domains in the NHIM and we show that, near the resonances, one can approximate the system by a system which is pendulum like, whereas for the rest of the phase space the system is close to a rotor. Finally, it is shown that one can consider the real system in the NHIM approximated by either a pendulum or a rotor and, therefore, using a KAM Theorem, that some of the tori present in the pendulum or the rotor are also present in the real system.

The main difficulty arises from the fact that a generic perturbation creates an infinite number of resonances and this makes impossible to isolate the resonant domains. Our solution to this problem was to consider a truncation of the Fourier series and to keep a good control of the error estimates. We also performed a detailed analysis of the motion in each resonance with error terms.

This allows us to give a complete description of the *geography of the resonances* generated by a generic perturbation. It is remarkable that we observe that our perturbation generates a Cantor-like set of gaps of different sizes among the primary tori.

Finally, in section 1.5 we present an example for which we can check the concrete conditions that guarantee the existence of diffusing orbits.

As future work, we plan to extend these results to the multiresonant case, that is when the Hamiltonian (1.1) contains several rotors and penduli, and also to study the speed of diffusion in terms of the invariant objects used in the transition chain.

Fast algorithms to compute invariant KAM tori and the associated whiskers

This part of the thesis has been done in collaboration with Rafael de la Llave and Yannick Sire.

The methods for diffusion discussed in chapter 1 require a careful study of the hyperbolic invariant objects present in the system as well as their possible connections, because they organize the solutions around them. The existence, description of their properties and approximation using computational tools of numerical analysis, constitutes one of the most important research areas in dynamical systems, not only because of their theoretical interest but also because of their applicability.

Among the most relevant invariant objects for the description of this mechanism one can find the so called whiskered tori, that is, quasi-periodic solutions with associated stable and unstable manifolds. Usually they are studied using KAM methods that require to have the system written down in action-angle variables, see [Lla01b] for a comparison of different KAM methods. In [LGJV05] de la Llave et al. introduce KAM techniques, based on the parametrization method, that allow to compute KAM tori both primary and secondary in a unified way and skipping the canonical transformation formalism. In [FLS07] Fontich, de la Llave and Sire generalize this result for whiskered tori. The method proposed in both papers is based on a quasi-Newton method and it lends itself to a very efficient numerical implementation.

In chapter 2 we develop numerical algorithms that use small requirements of storage and operations for the computation of invariant tori and the associated whiskers in Hamiltonian systems (symplectic maps and Hamiltonian vector fields). The algorithms are based on the parameterization method (developed by Cabré, Fontich and de la Llave in [CFL03a, CFL03b, CFL05]) and follow closely the proof of the KAM theorem given in [LGJV05] and [FLS07]. They essentially consist of solving a functional equation satisfied by the invariant tori by means of a Newton method.

In sections 2.5 and 2.7, we show how to compute efficiently both maximal invariant tori and whiskered tori. It is important to remark that the numerical algorithms presented

allow to compute in a unified way primary and secondary invariant KAM tori. Moreover, in section 2.6 we present fast algorithms for the iteration of the cocycles and the computation of the invariant bundles, which constitutes a previous step for the computation of invariant whiskered tori and the associated whiskers.

In section 2.10 we include examples of the numerical implementation to compute invariant tori, with precise information about the time running and efficiency. More precisely, in section 2.10.1 we compute primary and secondary tori for the standard map and in section 2.10.2 we apply the algorithms to compute primary maximal and hyperbolic KAM tori of the Froeschlé map.

Concerning the whiskers, the fast algorithms presented are designed to compute invariant manifolds of any rank, but we have restricted ourselves to rank-1 manifolds for this memory. Actually, we think that considering this simplification makes easier the comprehension of the main idea behind. Again, as in the case of invariant tori, the key point is that taking advantage of the geometry of the problem we can devise algorithms which implement a Newton step without having to store and much less invert a large matrix.

We include a discussion of the algorithms for the *large matrix method*, because they have been used in the next chapter for the computation of the Isochrons (see section 3.4). Algorithms with similar features have been developed in [HL06b, HL06a] for the case of quasi-periodic systems.

As we already mentioned, the method allows to compute KAM tori in a very efficient way. Bostered by this success, we plan to study the breakdown of maximal tori for symplectic 4-dimensional maps, like the Froeschlé map. Many researchers have worked on this problem (see for instance [HS95, CFL04]), but there are still few results in this area.

Moreover, the method proposed has the advantage that it is not required to have the system written in action-angle variables. This is very convenient when we are working in situations such that the action-angle variables are singular. This is the case in the problem of Arnol'd diffusion discussed in chapter 1, in which it is necessary to get the invariant tori (primary and secondary) close to the resonances. As we mentioned in the previous section, near the resonances one can approximate the system by a system which is a pendulum like, and the action-angle variables close to the separatrix become singular. Avoiding this step using the methods in papers [LGJV05] and [FLS07] could allow us to reduce the regularity r of the Hamiltonian (1.1) assumed in Theorem 1.2.1 and get better results in terms of differentiability.

Finally, this method can be also applied to compute quasi-periodic breathers of coupled map lattices. This is a kind of systems that appear frequently in physics and biology models when identical particles interact with each other.

Computation of Isochrons and PRC in Neuroscience models

The behavior of coupled neural oscillators in neuroscience has been the subject of a great deal of recent interest and there is a wide literature on this topic (see [Izh07] for a survey).

Many oscillators can be described by the their phase variable. Moreover, under certain conditions, the phase of the oscillation can be also defined outside the hyperbolic limit cycle via asymptotic phase. Thus, the stable manifold of a point x_0 on a limit cycle, represents solutions having equal phases, and it is often referred as the isochron of x_0.

One easily measurable property of a neural oscillator is its phase resetting curve (PRC). The PRC is found by perturbing the oscillation with a brief stimulus at different times in its cycle and measuring the resulting phase-shift from the unperturbed system. It is a very useful tool to explain how the coupling between neurons can affect the phase and lead them to a synchronized or desynchronized activity.

In section 3.3 we study how are the phase resetting curves around a limit cycle γ of a planar vector field X related to the fact that X is the infinitesimal generator of a Lie symmetry (that is, there exists a vector field Y and a scalar function μ such that $[Y, X] = \mu Y$). We show how the time variables involved in the Lie symmetry provide a natural way (a kind of *normal form*) to express the vector field around γ, similar to action-angle variables for integrable systems. In addition, the knowledge of the orbits of Y give a trivial way to compute the phase resetting curve, not only on γ as it is usual in the literature, but also in a neighborhood of it, thus obtaining what we call *phase resetting surfaces*.

Apart from the theoretical aim of unveiling relationships among different concepts, in this chapter we also perform the effective computation of these symmetries. The numerical scheme is based on the theoretical ground of the parameterization method (developed by Cabré, Fontich and de la Llave, [CFL05]) to compute invariant manifolds (the orbits of Y) in a neighborhood of γ. The algorithms implemented have the same features as the ones described in the previous chapter. However, we leave for the a future work the implementation of the fast algorithms.

Limit cycles in biological (more specifically, neuroscience) models encompass numerical problems that are often neglected or avoided; we present a discussion about them and give general solutions whenever it is possible.

Finally, we use all theoretical and numerical results to compute both the phase resetting curves and surfaces and the isochronous sections of limit cycles for well-known biological models. In this part of the chapter, we also explore how the phase resetting curves evolve (in the parameter space) between different bifurcation values.

In the future, we plan to extend this results to $\mathbb{R}^n$. We also plan to study the effect on the synchronization when these neurons are stimulated with an external stimuli of short period. Another case of high frequency stimulation is when the surrounding neurons of the neuron which is the object of study are in a *bursting* state. This is like and intrinsic stimuli of high frequency.

Main Contributions and Future work

In this section we summarize the main achievements in this thesis for easy reference as well as some open problems for the future work.

Arnold diffusion for a priori unstable Hamiltonian systems

- We prove the existence of diffusing orbits for a priori unstable Hamiltonian systems of the form (1.1) with a generic perturbation h which is regular enough.
- We give explicitly the generic conditions required to prove the existence of diffusion and we check them for a particular example.
- In the future, we plan to compute estimates of the diffusion time using topological methods.

Fast algorithms to compute invariant KAM tori and the associated whiskers

- We developed fast numerical algorithms to compute invariant to tori and the whiskers. They include maximal and hyperbolic invariant tori. The algorithms can deal in a unified way with primary and secondary KAM tori.
- We implemented the fast algorithms for the invariant tori and we applied them to compute primary and secondary maximal tori of the standard map and primary maximal and whiskered tori of the Froeshclé map.
- In the future, we plan to develop algorithms to compute the whiskers of rank greater than 1 in a very general case and implement them.
- We plan also to study the breakdown of maximal invariant tori for the case of symplectic 4D-maps.

Computation of Isochrons and PRC in Neuroscience models

- We give a relation between the Phase Resetting Curves and the Isochrons with the Lie symmetries.
- We extend the Phase Resetting Curves to a neighborhood of the limit cycle, obtaining what we call the *Phase Resetting Surface*.
- We compute numerically the Phase Resetting Curves and Surfaces for different models using a parameterization method.
- In the future, we plan to implement the faster algorithms described in chapter 2 for the computation of Isochrons and extend the results to $\mathbb{R}^n$.
- We plan to exploit the differences between PRS and PRC to show up the effect of stimulations far from the limit cycle as opposed to near the limit cycle.

Chapter 1

Arnold diffusion for a priori unstable Hamiltonian systems

1.1 Introduction

The goal of this chapter is to present a generalization of the geometric mechanism for global instability (popularly known as Arnold diffusion) in a priori unstable Hamiltonian systems introduced in [DLS06a]. That paper developed an argument to prove the existence of global instability in a-priori unstable nearly integrable Hamiltonian systems (the unperturbed Hamiltonian presents hyperbolicity, so that it can not be expressed globally in action-angle variables) and applied it to a model that presented the so called *large gap problem*. However, in that case, the perturbation was assumed to be a trigonometric polynomial in the angular variables. In this memoir we perform an accurate process of truncation of the Fourier series of the perturbation and we present a deeper study of the resonant zone. Using this, we are able to extend and simplify some of the results in [DLS06a] and apply them to an a priori unstable Hamiltonian system with a generic perturbation.

The phenomenon of global instability in Hamiltonian systems has attracted the attention of both mathematicians and physicists in the last years due to its remarkable importance for the applications. It deals, essentially, with the question of what is the effect on the dynamics when an autonomous mechanical system is submitted to a small periodic perturbation. More precisely, whether these perturbations accumulate over time giving rise to a long term effect or whether these effects average out.

The instability problem was formulated first by Arnold in 1964. In his celebrated paper [Arn64], Arnold constructed an example for which he proved the existence of trajectories that avoided the obstacles of KAM tori and performed long excursions. The mechanism is based on the existence of transition chains of whiskered tori, that is, sequences of tori such that the unstable manifold (whisker) of one of these tori intersects transversally the stable manifold (whisker) of the next one. By an obstruction argument, there is an orbit that follows this transition chain, giving rise to an unstable orbit.

The example proposed in [Arn64] turns out to be rather artificial because the perturbation was chosen in such a way that it preserved exactly the complete foliation of invariant tori existing in the unperturbed system. However, a generic perturbation of size

ε creates gaps at most of size $\sqrt{\varepsilon}$ in the foliation of persisting primary KAM tori, whereas it moves the whiskers only by an amount ε. These gaps are centered around resonances, that is resonant tori that are destroyed by the perturbation. This is what is known in the literature as the *large gap problem* (see, for instance, [Moe96] for a discussion about the large gap problem and, indeed, of the problem of diffusion).

In the last ten years there has been a notable progress in the comprehension of the mechanisms that give rise to the phenomenon of instability and a variety of methods have been suggested. As an example of this, we will mention that the large gap problem has been solved simultaneously by a variety of techniques: different geometrical methods [DLS00, DLS06a, DLS06b] (scattering map) and [Tre04, PT07] (separatrix map); topological methods [GL06b, GL06a] and variational methods [CY04a, CY04b].

Of particular interest for the present memoir are [DLS00, DLS06a, DLS06b]. The strategy in the mentioned papers is based on the incorporation of new invariant objects, created by the resonances, like secondary KAM tori and the stable and unstable manifolds of lower dimensional tori in the transition chain, together with the primary KAM tori. The scattering map, introduced by the same authors (see [DLS08] for a geometric study) is the essential tool for the heteroclinic connections between invariant objects of different topology.

In this thesis we extend the geometric mechanism introduced in the mentioned papers to a class of model systems for which the perturbation does not need to have a finite number of harmonics in the angular variables. In particular, the Hamiltonian studied has the following form

$$H_\varepsilon(p,q,I,\varphi,t) = \pm\left(\frac{1}{2}p^2 + V(q)\right) + \frac{1}{2}I^2 + \varepsilon h(p,q,I,\varphi,t;\varepsilon), \tag{1.1}$$

where $p \in (-p_0, p_0) \subset \mathbb{R}$, $I \in (I_-, I_+) \subset \mathbb{R}$ and $(q,\varphi,t) \in \mathbb{T}^3$.

The main result of this part of the thesis is Theorem 1.2.1, stated in Section 1.2.2 with the concrete hypotheses for the Hamiltonian (1.1), from which we can deduce the following short version:

Theorem 1.1.1. *Consider the Hamiltonian (1.1) and assume that V and h are $\mathcal{C}^{r+2}$ generic, with $r > r_0$, large enough. Then there is $\varepsilon^* > 0$ such that for $0 < |\varepsilon| < \varepsilon^*$ and for any interval $[I_-^*, I_+^*] \in (I_-, I_+)$, there exists a trajectory $\widetilde{x}(t)$ of the system (1.1) such that for some $T > 0$*

$$I(\widetilde{x}(0)) \leq I_-^*; \qquad I(\widetilde{x}(T)) \geq I_+^*.$$

Remark 1.1.2. A value of r_0 which follows from our argument is $r_0 = 242$ (see Remark 1.2.2).

Our strategy for the proof follows the geometric mechanism proposed in [DLS06a]. Indeed, in order to organize the different invariant objects that we will use to construct a transition chain, we will first identify the normally hyperbolic invariant manifold (NHIM) present in the system. This NHIM will have associated stable and unstable invariant manifolds that, generically, intersect transversally. Therefore, we can associate to this object two types of dynamics: the *inner* and the *outer*. The outer dynamics takes into account the asymptotic motions to the NHIM and is described by the scattering map.

The inner dynamics is the one restricted to the NHIM and contains Cantor families of primary and secondary KAM tori. The combination of these two types of dynamics will allow us to construct a transition chain.

The results in [DLS06a] can be applied straightforward for the persistence of the NHIM and the transversality of the associated stable and unstable manifolds. The arguments presented in this memoir focus on the inner dynamics and the study of the invariant objects present in the NHIM.

As we already mentioned for the Arnold example, at resonances, which correspond to the places where the frequency $I = -l/k$ for $(k, l) \in \mathbb{Z}^2$ is rational and the associated Fourier coefficient $h_{k,l}$ of the perturbation h is nonzero, the foliation of KAM tori in the NHIM is destroyed and gaps between the Cantor family of invariant tori in the NHIM of size $\mathcal{O}(\varepsilon^{1/2}|h_{k_0,l_0}|^{1/2})$ are created, for (k_0, l_0) such that $l/k = l_0/k_0$ and $\gcd(k_0, l_0) = 1$ (see equation (1.86)). For a generic $\mathcal{C}^{r+2}$ perturbation h, when we restrict it to the NHIM and we write it in adequate coordinates we are left with a $\mathcal{C}^r$ perturbation (see the subsection "restriction to NHIM" in Section 1.2.3.3), so that $|h_{k,l}| \sim |(k, l)|^{-r}$, and therefore the above gaps are of size $\mathcal{O}(\varepsilon^{1/2}|(k_0, l_0)|^{-r/2})$. Moreover, other invariant objects, like secondary tori and lower dimensional tori, are created inside the gap. They correspond to invariant objects of different topology that were not present in the unperturbed system but are generated by the resonances.

In order to study their existence and give an approximate expression for them we will use a combination of m steps of averaging and a KAM Theorem. Notice that in our case, since the perturbation is generic, we will have an infinite number of resonances. Our approach for this study will be to consider an adequate truncation up to some order M, depending on ε of the Fourier series of the perturbation h in such a way that we deal only with a finite number of harmonics $|(k, l)| \leq M$ and therefore of resonances.

Another remarkable difference with respect to the results obtained in [DLS06a] is that in that case the size of the gaps created in the foliation of invariant tori was uniform, whereas in our case, since the size is $\mathcal{O}(\varepsilon^{1/2}|(k_0, l_0)|^{-r/2})$, we have a heterogeneous sea of gaps of different sizes. Among them, we will distinguish between *small gaps* and *big gaps*, which are strongly related to the mentioned large gap problem. Indeed, big gaps are generated by resonances $-l_0/k_0$ of order one, such that $|(k_0, l_0)| < M_{\mathrm{BG}}$, equivalently $|(k_0, l_0)|^{-r/2} \geq \varepsilon^{1/2}$, where $M_{\mathrm{BG}} \sim \varepsilon^{-1/r}$ (see Section 1.3.3.3 for precise results).

From a more technical point of view (see Section 1.3.2 for details), we would like to remark that the main difficulties arise from the fact that in order to perform a resonant averaging procedure, we need to isolate resonances corresponding to $|(k, l)| \leq M$, for M depending on ε. Consequently, the width L of the resonant domain can not be chosen independently of ε, as it was the case in [DLS06a]. Moreover, along the averaging procedure we need to keep track of the $\mathcal{C}^\ell$ norms of the averaged terms and the remainders, and these blow up as a negative power of L. Hence, we will see that a good choice for L around a resonance $I = -l/k$ will be $L = L_k \sim \varepsilon^{1/n}/|k|$ (see hypotheses of Theorem 1.3.11), where n is the required regularity to apply KAM Theorem after the averaging procedure. Notice that L is not uniform along the resonances but depends on the size of the Fourier coefficients $h_{k,l}$.

Finally, after m steps of averaging, we will show that the remainder tail, that is the Fourier coefficients $h_{k,l}$ such that $|(k, l)| > M$ can be neglected. This will be ensured by

a fast enough decreasing rate of the coefficients and therefore a large enough regularity r of the perturbation. Thus, the required regularity r will be determined according to the number m of steps of averaging performed, as well as the needed regularity n to apply KAM Theorem after the averaging procedure.

We are using a version of the KAM Theorem that requires to have the Hamiltonian system written in action-angle variables. Since near the resonances we approximate the system by one which is close to a pendulum, the action variable becomes singular close to the separatrix. This fact, together with the requirement to have the invariant objects close enough (at a distance smaller than ε) implies that the perturbation of the averaged Hamiltonian has to be extremely small in the resonant regions. The immediate consequence of this fact is that, in the case we are studying, one has to perform at least $m = 10$ steps of averaging (see Theorem 1.3.28). The needed regularity n to apply KAM Theorem after m averaging steps is $n = 2m + 6$ (see Proposition 1.3.24). Since the regularity r required to ensure that the remainder tail is smaller than the averaging remainder turns out to be $r > (n-2)m + 2$, see Remark 1.3.20, one has to impose $r > 242$.

We do not claim that this is an optimal result. Actually, another version of the KAM Theorem that allowed us to avoid the change into action-angle variables could improve the results in terms of the needed regularity. However, it is worth mentioning that we managed to decrease the required steps of averaging in the resonant domains with respect to the results in [DLS06a]. Since in the resonances the behavior of KAM tori is different depending on how close they are to the separatrix (tori are flatter as they are further from the separatrix), we consider different regions where we perform different scalings. This strategy, which was already introduced in [DLS06a], has been improved in this thesis introducing a new sequence of domains in Theorem 1.3.30. When applied to the case with a finite number of resonances as in [DLS06a], $m = 9$ steps of averaging and $r \geq 26$ are enough. This clearly improves the needed regularity r which was $r \geq 60$ in [DLS06a] because m was chosen $m = 26$.

Sections 1.3.3.3, 1.3.3.4 and 1.3.3.5 contain a quantitative description of the *geography* of the resonances and a detailed study of the invariant objects generated by the resonances. The effect of the resonances in a system constitutes a fundamental problem not only for diffusion but also for many other physical applications and it has been an important object of study in the physical literature, see for instance [Chi79, Ten82]. The study performed in this part of the thesis contributes to understand better the different types of resonances and the geometric objects that one can find therein and can be very helpful in many physical problems.

Moreover, we think that this study can be extended to a class of models that present multiple resonances, see [DLS07].

We would like to emphasize that in our case, and this is different from the results in [DLS06a], only the resonances of order 1, that is, the ones that appear at the first step of averaging, create big gaps; whereas in [DLS06a], both resonances of order 1 and 2 could generate big gaps. This is because we are dealing with a perturbation that generically will have all the harmonics different from zero. This means that the effect of the resonances associated to the biggest Fourier coefficients (low frequencies) will be detected at the first step of averaging. Since the size of the gap depends on both the order of the resonance and the size of the Fourier coefficient associated to that resonance, the ones that appear

at the second step of averaging already correspond to small Fourier coefficients and the size of their gap will be smaller than ε. The immediate consequence of this fact it that in the forthcoming Theorem 1.2.1, we can give conditions **H3** explicitly in terms of the original perturbation h.

This chapter is organized in the following way. In Section 1.2 we state Theorem 1.2.1, which establishes the existence of diffusing orbits for the model considered under precise conditions. Since the required hypotheses are checked to be generic, Theorem 1.1.1 follows straightforwardly. The proof of Theorem 1.2.1 is given in Section 1.2, except for two technical results, Theorem 1.3.1 and Proposition 1.4.1, which are postponed to the following sections.

Thus, in Section 1.3 we prove Theorem 1.3.1, which provides a quantitative existence of invariant objects for the inner dynamics in the NHIM following the steps indicated in Section 1.2. In Section 1.4 we use the scattering map to prove Proposition 1.4.1 about the existence of heteroclinic connections between the invariant objects obtained in the previous section.

We would like to remark that, in contrast to [DLS06a], and thanks to the new results about the scattering map obtained in [DLS08], we use the Hamiltonian function generating the deformation of the scattering map instead of the scattering map itself, in order to compute the images of the leaves of a certain foliation under the scattering map.

Finally, in Section 1.5 we have included for illustration a concrete example, for which we check the hypotheses of Theorem 1.2.1. In the Appendix, we have brought some technical results used in this part of the thesis.

1.2 Statement of results

Before stating the main result we need to introduce some notation.

1.2.1 Notation and preliminaries

Let r be a positive integer and $\mathcal{D} \subset \mathbb{R}^n$ a compact set. We will denote the set of $\mathcal{C}^r$ functions from $\mathcal{D}$ to $\mathbb{R}^m$ by $\mathcal{C}^r(\mathcal{D}, \mathbb{R}^m)$. By a $\mathcal{C}^r$ function f in $\mathcal{D}$ we mean a continuous function in $\mathcal{D}$ and all the derivatives continuous in $\bar{\mathcal{D}}$. When $m = 1$, we simply write $\mathcal{C}^r(\mathcal{D})$. Given $f \in \mathcal{C}^r(\mathcal{D}, \mathbb{R}^m)$, we consider the standard $\mathcal{C}^r$ norm,

$$|f|_{\mathcal{C}^r(\mathcal{D})} = \sum_{i=1}^{m} \sum_{\ell=0}^{r} \sum_{|\alpha|=\ell} \sup_{x \in \mathcal{D}} \frac{|D^\alpha f_i(x)|}{\alpha!}, \tag{1.2}$$

where f_i denotes the i-th component of the function f, for $i = 1, \ldots, m$. We omit the domain in the notation when it does not lead to confusion.

We use the standard multi-index notation: if $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ one sets

$$|\alpha| = \alpha_1 + \cdots + \alpha_n$$

$$\alpha! = \alpha_1! \alpha_2! \cdots \alpha_n!$$

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

In the case that the function f depends only on a few of the variables, we will denote it in the same way, that is $|f|_{\mathcal{C}^r} = |f|_{\mathcal{C}^r(\mathcal{D})}$, and consider it as a function of more variables defined in the appropriate domain.

Note that we denote $|f|_{\mathcal{C}^0} = \sup_{x\in\mathcal{D}} |f(x)|$, which is the standard supremum norm, so the $|\cdot|_{\mathcal{C}^r(\mathcal{D})}$ norm can be expressed, equivalently, as

$$|f|_{\mathcal{C}^r(\mathcal{D})} := \sum_{i=1}^{m} \sum_{\ell=0}^{r} \sum_{|\alpha|=\ell} \frac{|D^\alpha f_i|_{\mathcal{C}^0(\mathcal{D})}}{\alpha!}.$$

The space of $\mathcal{C}^r$ functions on $\mathcal{D}$ endowed with the $\mathcal{C}^r$ norm is a Banach algebra (see [Con90]), that is, it is a Banach space with the property that given any two functions f, g in $\mathcal{C}^r(\mathcal{D})$, they satisfy

$$|fg|_{\mathcal{C}^\ell} \le |f|_{\mathcal{C}^\ell} |g|_{\mathcal{C}^\ell} .$$

Since we will deal with $\mathcal{C}^r$ functions defined on a compact domain $\mathcal{D} \subset \mathbb{R}^n \times \mathbb{T}^n$, we can also considerate the following seminorm, that takes into account the different regularities and the estimates for the derivatives in each type of variable:

$$|f|_{\ell_1,\ell_2} := \sum_{m_1=0}^{\ell_1} \sum_{m_2=0}^{\ell_2} \sum_{\substack{\alpha_1,\alpha_2\in(\mathbb{N}\cup\{0\})^n \\ |\alpha_1|=m_1,|\alpha_2|=m_2}} \frac{1}{\alpha_1!\alpha_2!} \sup_{(I,\varphi)\in\mathbb{R}^n\times\mathbb{T}^n} \left| \frac{\partial^{m_1+m_2} f}{\partial I^{\alpha_1} \partial \varphi^{\alpha_2}} \right|, \tag{1.3}$$

for $0 \le \ell_1 + \ell_2 \le r$.

Note that $|f|_{\mathcal{C}^\ell} = \sum_{m=0}^{\ell} |f|_{m,\ell-m}$, for $0 \le \ell \le r$.

We will use the following notation, which is rather usual. We will write $\alpha \preceq \beta$ and also $\alpha = \mathcal{O}(\beta)$ if there exists a constant C such that $\alpha \le C\beta$. When we have $\alpha \preceq \beta$ and $\beta \preceq \alpha$ we will write $\alpha \sim \beta$. However, in some informal discussions we will abuse notation and we will say that a quantity is of order ε^p, where $\varepsilon > 0$ and $p \in \mathbb{R}$, if it is between $C_0\varepsilon^p$ and $C_1\varepsilon^p$, for some $0 < C_0 < C_1$ independent of ε.

We will say that a function $f = \mathcal{O}_{\mathcal{C}^r(\mathcal{D})}(\eta)$ when

$$|f|_{\mathcal{C}^r(\mathcal{D})} \preceq \eta.$$

In general, we will deal with scalar functions, that is $f \in \mathcal{C}^r(\mathcal{D}, \mathbb{R}) = \mathcal{C}^r(\mathcal{D})$.

1.2.2 Set up and main result

We consider a 2π-periodic in time perturbation of a pendulum and a rotor described by the non-autonomous Hamiltonian,

$$\begin{aligned} H_\varepsilon(p,q,I,\varphi,t) &= H_0(p,q,I) + \varepsilon h(p,q,I,\varphi,t;\varepsilon) \\ &= P_\pm(p,q) + \frac{1}{2}I^2 + \varepsilon h(p,q,I,\varphi,t;\varepsilon), \end{aligned} \tag{1.4}$$

where

$$P_\pm(p,q) = \pm\left(\frac{1}{2}p^2 + V(q)\right) \tag{1.5}$$

and $V(q)$ is a 2π-periodic function. We will refer to $P_\pm(p,q)$ as the *pendulum*.

The term $\frac{1}{2}I^2$ describes a rotor and the final term h is the perturbation term and depends periodically on time, so it can be expressed via its Fourier series in the variables (φ, t)

$$h(p,q,I,\varphi,t;\varepsilon) = \sum_{(k,l)\in\mathbb{Z}^2} h_{k,l}(p,q,I;\varepsilon)e^{\mathrm{i}(k\varphi+lt)}. \tag{1.6}$$

It will be convenient to consider the autonomous system by introducing the extra variables (A,s):

$$\begin{aligned}\widetilde{H}_\varepsilon(p,q,I,\varphi,A,s) &= A + H_0(p,q,I) + \varepsilon h(p,q,I,\varphi,s;\varepsilon)\\ &= A + P_\pm(p,q) + \frac{1}{2}I^2 + \varepsilon h(p,q,I,\varphi,s;\varepsilon)\end{aligned} \tag{1.7}$$

where the pairs $(p,q)\in\mathbb{R}\times\mathbb{T}$, $(I,\varphi)\in\mathbb{R}\times\mathbb{T}$ and $(A,s)\in\mathbb{R}\times\mathbb{T}$ are symplectically conjugate.

The extra variable s makes the system autonomous and A is symplectically conjugate to the variable s and does not play any dynamical role. So, we are only interested in studying the dynamics of variables (p,q,I,φ,s), given by the system of equations:

$$\begin{aligned}\dot{p} &= \mp V'(p) && -\varepsilon\tfrac{\partial h}{\partial q}(p,q,I,\varphi,s;\varepsilon)\\ \dot{q} &= \pm p && +\varepsilon\tfrac{\partial h}{\partial p}(p,q,I,\varphi,s;\varepsilon)\\ \dot{I} &= && -\varepsilon\tfrac{\partial h}{\partial \varphi}(p,q,I,\varphi,s;\varepsilon)\\ \dot{\varphi} &= I && +\varepsilon\tfrac{\partial h}{\partial I}(p,q,I,\varphi,s;\varepsilon)\\ \dot{s} &= 1\end{aligned} \tag{1.8}$$

The domain of definition we consider is a compact set of type

$$\mathcal{D} := \mathcal{S}\times[I_-,I_+]\times\mathbb{T}^2\times[-\varepsilon_0,\varepsilon_0],$$

where $\mathcal{S}\subset\mathbb{R}\times\mathbb{T}$ is a neighborhood of the separatrix of the pendulum.

Then, the main Theorem of this part of the thesis is:

Theorem 1.2.1. *Consider a Hamiltonian of the form (1.4) where V and h are $\mathcal{C}^{r+2}$ in $\mathcal{D}$, with $r > r_0$, sufficiently large. Assume also that,*

H1 *The potential $V:\mathbb{T}\to\mathbb{R}$ has a unique global maximum, say at $q=0$, which is non-degenerate (i.e. $V''(0)<0$). We denote by $(p_0(t),q_0(t))$ an orbit of the pendulum $P_\pm(p,q)$ in (1.4), homoclinic to $(0,0)$.*

H2 *Consider the* Poincaré function*, also called* Melnikov potential*, associated to h (and to the homoclinic orbit (p_0,q_0)):*

$$\begin{aligned}\mathcal{L}(I,\varphi,s) &= -\int_{-\infty}^{+\infty}(h(p_0(\sigma),q_0(\sigma),I,\varphi+I\sigma,s+\sigma;0)\\ &\quad -h(0,0,I,\varphi+I\sigma,s+\sigma;0))d\sigma\end{aligned} \tag{1.9}$$

H2' *Given real numbers $I_- < I_+$, assume that, for any value of $I \in (I_-, I_+)$, there exists an open set $\mathcal{J}_I \in \mathbb{T}^2$, with the property that when $(I,\varphi,s) \in H_-$, where*

$$H_- = \bigcup_{I\in(I_-,I_+)} \{I\}\times\mathcal{J}_I \subset (I_-,I_+)\times\mathbb{T}^2, \tag{1.10}$$

the map

$$\tau \in \mathbb{R} \mapsto \mathcal{L}(I,\varphi - I\tau, s-\tau)$$

has a non-degenerate critical point τ which is locally given by the implicit function Theorem in the form $\tau = \tau^*(I,\varphi,s)$, with τ^* a smooth function.

H2" *Introduce the reduced Poincaré function $\mathcal{L}^*$ defined by*

$$\mathcal{L}(I,\varphi - I\tau^*(I,\varphi,0), -\tau^*(I,\varphi,0)) := \mathcal{L}^*(I,\varphi), \tag{1.11}$$

on

$$H_-^* = \{(I,\widetilde{\theta}) : \widetilde{\theta} = \varphi - Is, (I,\varphi,s) \in H_-\} = \bigcup_{I\in(I_-,I_+)} \{I\}\times\mathcal{J}_I^*, \tag{1.12}$$

and assume that

$$\widetilde{\theta} \mapsto \frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}(I,\widetilde{\theta})$$

for $\widetilde{\theta} = \varphi - Is \in \mathcal{J}_I^*$ is non-constant and positive (respectively negative).

H3 *Fix $0 < \nu \le 1/16$, for any $0 < \varepsilon < 1$ and for any $(k_0,l_0) \in \mathbb{Z}^2$ with $\gcd(k_0,l_0) = 1$ and $|(k_0,l_0)| < M_{\rm BG}$, where $|(k_0,l_0)| = \max(|k_0|,|l_0|)$ and $M_{\rm BG} = \varepsilon^{-(1+\nu)/r}$, introduce the 2π-periodic function*

$$U^{k_0,l_0}(\theta) = \sum_{\substack{t\in\mathbb{Z}-\{0\},\\ |t||(k_0,l_0)|<M_{\rm BG}}} h_{tk_0,tl_0}(0,0,-l_0/k_0;0)e^{{\rm i}t\theta},$$

where $\theta = k_0\varphi + l_0 s$, for which we assume:

H3' *The function U^{k_0,l_0} has a non-degenerate global maximum.*

H3" *For $|(k_0,l_0)| \prec \varepsilon^{-1/r}$ (i.e. $|(k_0,l_0)| = o(\varepsilon^{-1/r})$, we assume that the $2\pi k_0$-periodic function f given by*

$$f(\theta) = \frac{k_0 U'^{k_0,l_0}(\theta)\frac{\partial\mathcal{L}^*}{\partial\widetilde{\theta}}\left(\frac{-l_0}{k_0},\frac{\theta}{k_0}\right) + 2U^{k_0,l_0}(\theta)\frac{\partial^2\mathcal{L}^*}{\partial\widetilde{\theta}^2}\left(\frac{-l_0}{k_0},\frac{\theta}{k_0}\right)}{2\frac{\partial^2\mathcal{L}^*}{\partial\widetilde{\theta}^2}\left(\frac{-l_0}{k_0},\frac{\theta}{k_0}\right)} \tag{1.13}$$

is non-constant.

H3"' *For* $|(k_0, l_0)| \sim \varepsilon^{-1/r}$ *(i.e.* $|(k_0, l_0)| = \mathcal{O}(\varepsilon^{-1/r})$ *and* $\varepsilon^{-1/r} = \mathcal{O}(|(k_0, l_0)|)$*), we assume the non-degeneracy condition stated explicitly in equation* (1.150)*.*

Then, there exists $\varepsilon^* > 0$ *such that for* $0 < |\varepsilon| < \varepsilon^*$ *and for any interval* $[I_-^*, I_+^*] \in (I_-, I_+)$*, there exists a trajectory* $\widetilde{x}(t)$ *of the system (1.4) such that for some* $T > 0$

$$I(\widetilde{x}(0)) \leq I_-^*; \qquad I(\widetilde{x}(T)) \geq I_+^*.$$

(respectively:

$$I(\widetilde{x}(0)) \geq I_+^*; \qquad I(\widetilde{x}(T)) \leq I_-^*).$$

Remark 1.2.2. r_0 depends on the number m of some averaging steps performed in the proof: $r_0 = 2(m+1)^2$ and $m \geq 10$ (see hypothesis of Theorem 1.3.1 in Section 1.3). If we take, for instance, $m = 10$ then $r_0 = 242$ is enough.

Remark 1.2.3. Notice that for every fixed ε we have one condition **H3** for every (k_0, l_0) such that $|(k_0, l_0)| < M_{\mathrm{BG}}$, that depends explicitly on (k_0, l_0). Hence, the number of non-degeneracy conditions **H3** is finite but grows with ε.

Remark 1.2.4. Notice that by the definition of $\tau^*(I, \varphi, s)$, the function

$$\mathcal{L}(I, \varphi - I\tau^*(I, \varphi, s), s - \tau^*(I, \varphi, s))$$

is a solution of the equation

$$I\partial_\varphi f(I, \varphi, s) + \partial_s f(I, \varphi, s) = 0,$$

hence we can define the reduced Poincaré function in the following way

$$\mathcal{L}^*(I, \varphi - Is) = \mathcal{L}(I, \varphi - I\tau^*(I, \varphi, s), s - \tau^*(I, \varphi, s)).$$

Remark 1.2.5. The main feature of Theorem 1.2.1, as already said in the introduction, is that h is not required to be a trigonometric polynomial in the variables (φ, s), which is a non-generic assumption, as it was the case in [DLS06a].

For every fixed ε, conditions **H1** and **H2** are open and dense, that is they hold for an open and dense set of Hamiltonians in the $\mathcal{C}^{r+2}$ topology.

For every fixed ε the number of non-degeneracy conditions **H3** is finite but grows with ε (the number of conditions depends on $(k_0, l_0) \in \mathbb{Z}^2$ such that $\gcd(k_0, l_0) = 1$ and $|(k_0, l_0)| \preceq \varepsilon^{-1/r}$). When ε tends to 0 we have a countable number of conditions, hence the set of Hamiltonians satisfying them is a residual set in the $\mathcal{C}^{r+2}$ topology, that is, a countable intersection of open and dense sets in the $\mathcal{C}^{r+2}$ topology.

Therefore the hypotheses of the theorem are $\mathcal{C}^{r+2}$ generic. So, the short version of Theorem 1.2.1 stated in Theorem 1.1.1 in the Introduction follows straightforwardly.

1.2.3 Proof of Theorem 1.2.1

The proof of this theorem follows the geometric mechanism stated in [DLS06a] and it is organized in four parts that we first sketch now:

1. The first part deals with the existence of a NHIM, which jointly with its associated stable and unstable manifolds, organizes all the dynamics, and is a consequence of hypothesis **H1**. By hypothesis **H2'**, its associated stable and unstable manifolds will intersect transversally, so we can associate to this object two types of dynamics: the inner and the outer.

2. The outer dynamics, which is the one that takes into account the asymptotic motions to the NHIM, is studied in the second part. We will see that we can associate a scattering map to the NHIM and give formulas for the Hamiltonian function which determines the deformation of this scattering map.

3. The third part of the proof consists of studying the inner dynamics, that is the one restricted to the NHIM. The goal is to show that, by hypotheses **H3'**, there exists a discrete foliation of invariant tori of different topology, which are close enough. Among these tori, some of them are primary, so they are just a continuation of the ones that existed for the integrable system ($\varepsilon = 0$), and some of them are secondary, these ones are contractible to a periodic orbit, so they correspond to motions with topologies that were not present in the unperturbed system but they are created by the resonances. The method of proof will be a combination of an averaging procedure and a quantitative version of KAM Theorem, which requires the Hamiltonian to be differentiable enough.

4. The last part of the proof consists of showing that the combination of both types of dynamics gives rise to a construction of a transition chain, that is, a sequence of whiskered tori in which the stable manifold of one torus intersects transversally the unstable manifold of the next one. To this end, one needs to show that the discrete foliation of whiskered tori which are invariant under the (inner) flow is not invariant under the scattering map or outer map. This is ensured by hypotheses **H2"**, **H3"** and **H3"'** in Theorem 1.2.1, which indeed provide the transversality of this discrete foliation to the scattering map. Finally we prove, using a standard obstruction property, that there is an orbit that follows this transition chain.

Next we give a proof of Theorem 1.2.1 organized in the four parts that we have mentioned. The first two parts follow readily from [DLS06a] and Theorems stated in [DLS06a] apply straightforwardly because hypotheses **H1** and **H2'** required for the proof of the mentioned results are the same as in our case. Moreover, for the second part we use the symplectic properties developed in [DLS08] to generalize the computation of the scattering map using its Hamiltonian function. So, for these parts, we only refer in Section 1.2.3.1 and 1.2.3.2, to the results in [DLS06a] and [DLS08] that we are using.

However, for the third part, the results obtained in [DLS06a] do not apply directly because in our case we are not assuming that the perturbation has a finite number of harmonics. Therefore, it has been necessary to develop a new methodology in order to prove that when we have a $\mathcal{C}^{r+2}$ perturbation h, with r large enough, and hypotheses **H3'** are fulfilled, for every ε we can truncate adequately its Fourier series and deal only with a finite number of harmonics and therefore a finite number of resonances to get a discrete foliation of tori close enough. Moreover, explicit approximate expressions for these tori

are obtained as the level sets of a certain function. The mentioned results are stated and proved rigorously in Section 1.3, giving rise to Theorem 1.3.1 and they constitute the essential result of this part of the thesis. In Section 1.2.3.3 we just refer to the results in Section 1.3 needed to prove part 3 of Theorem 1.2.1.

Once we have fixed in part 3, for every ε, the number of resonances, part 4 follows readily from the finite hypotheses **H2"**, **H3"** and **H3"'** as in [DLS06a]. The main difference is that, in contrast to [DLS06a] and thanks to the new results about the symplectic properties of the scattering map obtained in [DLS08], we can use the Hamiltonian function generating the deformation of the scattering map instead of the scattering map itself, in order to compute the images of the leaves of a certain foliation under the scattering map. The results with their proof are stated in Section 1.4. In Section 1.2.3.4 we just refer to the results in Section 1.4 needed to prove part 4 of Theorem 1.2.1.

1.2.3.1 First Part: Existence of a NHIM and its associated stable and unstable manifolds

The method of proof is based on the existence of an invariant object, a NHIM (see, for instance, [HPS77, Fen74, Fen77, Fen79, Lla00, Wig90] for the standard theory of NHIMs used in this thesis), which jointly with its associated stable and unstable manifolds, organizes all the dynamics.

We start discussing the geometric features of the unperturbed case which will survive under the perturbation. For the case $\varepsilon = 0$, the Hamiltonian (1.4) is integrable and consists of two uncoupled systems: a rotor and a pendulum. So, the cartesian product of invariant objects of each of these subsystems will give an invariant object of the full system. Then, by hypothesis **H1**, if we consider the product of the hyperbolic fixed point $(p, q) = (0, 0)$ of the pendulum $P_{\pm}(p, q)$ in (1.5) with all the other variables, we have that for the values I_-, I_+ given in Theorem 1.2.1, the set

$$\widetilde{\Lambda} = \{\widetilde{x} = (p, q, I, \varphi, s) \in (\mathbb{R} \times \mathbb{T})^2 \times \mathbb{T} : p = q = 0, I \in [I_-, I_+]\} \tag{1.14}$$

is a 3-dimensional invariant manifold and normally hyperbolic for the flow of the Hamiltonian system (1.8) for $\varepsilon = 0$. The associated stable and unstable invariant manifolds of $\widetilde{\Lambda}$ are the ones inherited from the separatrices of the pendulum (stable and unstable manifolds of the hyperbolic fixed point) and they agree:

$$W^s\widetilde{\Lambda} = W^u\widetilde{\Lambda} = \{(p_0(\tau), q_0(\tau), I, \varphi, s) : \tau \in \mathbb{R}, I \in [I_-, I_+], (\varphi, s) \in \mathbb{T}^2\} \tag{1.15}$$

where $(p_0(\tau), q_0(\tau))$ is the orbit of the pendulum $P_{\pm}$, provided by hypothesis **H1**, which is homoclinic to the hyperbolic fixed point $(0, 0)$.

The Hamiltonian system (1.8) for $\varepsilon = 0$ restricted to the manifold $\widetilde{\Lambda}$ is given simply by

$$\dot{I} = 0, \quad \dot{\varphi} = I, \quad \dot{s} = 1.$$

The dynamics on this manifold is very simple: all the solutions lie on an 2-dimensional invariant torus $I = \text{cte}$. Therefore, the normally hyperbolic invariant manifold is foliated by a one-parameter family of 2-dimensional invariant tori indexed by I, with associated frequency $(I, 1)$.

For $0 < |\varepsilon| \ll 1$, by the theory of NHIM (see the references above), the manifold $\widetilde{\Lambda}$ persists, giving rise to another manifold $\widetilde{\Lambda}_\varepsilon$ with associated local stable and unstable manifolds $W^{s,\mathrm{loc}}\widetilde{\Lambda}_\varepsilon$ and $W^{u,\mathrm{loc}}\widetilde{\Lambda}_\varepsilon$, which can be prolonged to $W^s\widetilde{\Lambda}_\varepsilon$ and $W^u\widetilde{\Lambda}_\varepsilon$, respectively. Both $\widetilde{\Lambda}_\varepsilon$ and its local stable and unstable manifolds, $W^{s,\mathrm{loc}}\widetilde{\Lambda}_\varepsilon$ and $W^{u,\mathrm{loc}}\widetilde{\Lambda}_\varepsilon$, are ε-close in the $\mathcal{C}^r$ sense to the unperturbed ones:

$$\widetilde{\Lambda}_\varepsilon = \widetilde{\Lambda} + \mathcal{O}_{\mathcal{C}^r}(\varepsilon); \quad W^{s,\mathrm{loc}}\widetilde{\Lambda}_\varepsilon = W^{s,\mathrm{loc}}\widetilde{\Lambda} + \mathcal{O}_{\mathcal{C}^r}(\varepsilon); \quad W^{u,\mathrm{loc}}\widetilde{\Lambda}_\varepsilon = W^{u,\mathrm{loc}}\widetilde{\Lambda} + \mathcal{O}_{\mathcal{C}^r}(\varepsilon). \tag{1.16}$$

The result of the persistence of the NHIM $\widetilde{\Lambda}_\varepsilon$ and its stable and unstable manifolds is formulated in Theorem 7.1 of [DLS06a], where the perturbation h in (1.4) was assumed to be a trigonometric polynomial. However, the only assumption required for the proof was the fact that the perturbation h and the potential V were $\mathcal{C}^{r+2}$, so Theorem 7.1 can be applied straightforwardly in our case.

1.2.3.2 Second Part: Outer Dynamics

The outer dynamics, which is the one that takes into account the asymptotic motion to the NHIM $\widetilde{\Lambda}_\varepsilon$, is described by the scattering map. It is possible to construct a scattering map associated to the NHIM $\widetilde{\Lambda}_\varepsilon$, as long as its stable and unstable manifolds intersect transversally.

In Proposition 9.2 in [DLS06a] it is proved that if hypothesis **H2'** in Theorem 1.2.1 is satisfied, then the stable and unstable manifolds $W^s\widetilde{\Lambda}_\varepsilon$ and $W^u\widetilde{\Lambda}_\varepsilon$ of the NHIM intersect transversally along a homoclinic manifold Γ_ε, which is also called a *homoclinic channel* (see [DLS08] for more details). So, we will be able to locally define the scattering map associated to Γ_ε and compute it in first order perturbation theory using the results in [DLS08]. Again, hypothesis **H2'** required for Proposition 9.2 in [DLS08] does not depend on the number of harmonics of the perturbation h, so the results stated also hold for the case we are considering in this memoir.

Therefore, the manifold $\widetilde{\Lambda}_\varepsilon$ defined in (1.16) from (1.14) has a scattering map associated to a homoclinic manifold Γ_ε, defined in the following way

$$\begin{array}{rccc} S_\varepsilon: & H_- \subset \widetilde{\Lambda}_\varepsilon & \to & \widetilde{\Lambda}_\varepsilon \\ & x_- & \mapsto & x_+ \end{array} \tag{1.17}$$

such that $x_+ = S(x_-) \Leftrightarrow \exists\, z \in \Gamma_\varepsilon$ such that

$$\mathrm{dist}(\Phi_{t,\varepsilon}(z), \Phi_{t,\varepsilon}(x_\pm)) \to 0 \qquad \text{for } t \to \pm\infty,$$

where $\Phi_{t,\varepsilon}$ is the flux of the Hamiltonian (1.4). Indeed,

$$|\Phi_{t,\varepsilon}(z) - \Phi_{t,\varepsilon}(x_\pm)| \leq \mathrm{cte}\, e^{-\mu|t|/2} \qquad \text{for } t \to \pm\infty,$$

where $\mu = \sqrt{-V''(0)} > 0$ is the characteristic exponent of the saddle point $(0,0)$ of the pendulum $P_\pm(p,q)$ in (1.5).

Heuristically, the scattering map maps points of the manifold $\widetilde{\Lambda}_\varepsilon$ to points of the manifold $\widetilde{\Lambda}_\varepsilon$, such that the motion of z synchronizes with that of x_- (and x_+) in the past (and in the future).

Moreover, in Proposition 9.2 in [DLS06a] it is given a perturbative formula for the difference of the actions I of the points $x_+ = S_\varepsilon(x_-)$ and x_-. Concretely, expressing the points $x_\pm$ in terms of the parametrization of $\widetilde{\Lambda}_\varepsilon$, given in Theorem 7.1 in [DLS06a] we have that

$$I(x_\pm) = I + \mathcal{O}_{\mathcal{C}^1}(\varepsilon),\ \varphi(x_\pm) = \varphi + \mathcal{O}_{\mathcal{C}^1}(\varepsilon),\ s(x_\pm) = s,$$

and

$$I(x_+) - I(x_-) = \varepsilon \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(I, \widetilde{\theta}) + \mathcal{O}_{\mathcal{C}^1}(\varepsilon^2), \tag{1.18}$$

for $\widetilde{\theta} = \varphi - Is$, where $\mathcal{L}^*$ is the reduced Poincaré function defined in hypothesis **H2"**.

The method used in [DLS06a], based on the fact that I is a slow variable, allowed only to compute the leading term of the action component of the scattering map, but not the φ component since it is not a slow variable.

Remark 1.2.6. Notice that there is a wrong sign in formula (9.9) in [DLS06a].

However, in a recent paper [DLS08] the authors showed that the scattering map is exact symplectic and introduced geometric methods that allow to compute perturbatively an expression for both fast and slow variables.

Thus, using the method proposed in Section 5 in [DLS08], we can give perturbative formulas for the Hamiltonian $\mathcal{S}_\varepsilon$ generating the deformation of the scattering map S_ε.

It follows straightforwardly from Theorem 31 in [DLS08] that the reduced Poincaré function $\mathcal{L}^*$ introduced in (1.11) corresponds to the Hamiltonian $-\mathcal{S}_0$, so that we obtain

$$\mathcal{S}_\varepsilon(I, \varphi, A, s) = -\mathcal{L}^*(I, \widetilde{\theta}) + \mathcal{O}(\varepsilon). \tag{1.19}$$

with $\widetilde{\theta} = \varphi - Is$.

Hence, the first order perturbative term of the scattering map is given by

$$S_\varepsilon(I, \varphi, A, s) = (I, \varphi, A, s) + \varepsilon \mathrm{J} \nabla \mathcal{S}_0(I, \varphi, A, s) + \mathcal{O}(\varepsilon^2), \tag{1.20}$$

where J is the canonical matrix of the symplectic form $\omega = \mathrm{d}I \wedge \mathrm{d}\varphi + \mathrm{d}A \wedge \mathrm{d}s$ and $\nabla = (\partial_I, \partial_\varphi, \partial_A, \partial_s)$. The extra variable A, conjugated to the angle s, was introduced to make apparent the symplectic character of the scattering map.

Notice that equation (1.18) is just the I component of equation (1.20).

We would like to remark that $S_\varepsilon = Id + \mathcal{O}(\varepsilon)$. In particular, one iteration of S_ε can only jump distances of order ε in the action direction I.

Remark 1.2.7. For the mechanism of diffusion we are interested in comparing the inner dynamics in $\widetilde{\Lambda}_\varepsilon$ with the outer dynamics provided by the scattering map S_ε. Although the computation up to first order of the scattering map for the I component is enough for our purposes, it is more natural to study the action of the scattering map in terms of the Hamiltonian $\mathcal{S}_\varepsilon$.

1.2.3.3 Third Part: Inner Dynamics

In this section we study the inner dynamics, that is, the dynamics of the flow of the Hamiltonian (1.4) restricted to the NHIM $\widetilde{\Lambda}_\varepsilon$. The main result is Theorem 1.3.1, which

states that there exists a discrete sequence of invariant tori $\mathcal{T}_i$ in the NHIM $\widetilde{\Lambda}_\varepsilon$, which are distributed along the actions in the interval (I_-, I_+) introduced in Theorem 1.2.1 and which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced in terms of the action variable, for some $0 < \eta \leq 1/32$. Moreover, Theorem 1.3.1 provides explicit approximate expressions for the invariant tori, which are of two types depending on the region of the phase of space where invariant tori lie: the big gaps region and the flat tori region.

The big gaps region is defined as

$$\mathcal{D}_{\mathrm{BG}} = \{(I, \varphi, s) \in (I_-, I_+) \times \mathbb{T}^2; |I + l/k| \leq L/|k|; |(k, l)| < M_{\mathrm{BG}}\} \tag{1.21}$$

where L and M_{BG} are defined in (1.55) and (1.87), respectively. For the purposes of this exposition it will be enough to use $L \approx \varepsilon^{1/n}$ and $M_{\mathrm{BG}} \approx \varepsilon^{-1/r}$, where n is the regularity of the Hamiltonian required for the application of KAM Theorem ($n = 26$ will be enough, see hypotheses of Theorem 1.3.1) and r ($r > n$) is the regularity of the Hamiltonian required for Theorem 1.2.1. The flat tori region is the complementary region.

In the flat tori region, there exists a Cantorian foliation of primary KAM tori, which are just a continuation of invariant tori $I = \mathrm{cte}$ present in $\widetilde{\Lambda}_0$ for the unperturbed Hamiltonian (1.4) for $\varepsilon = 0$.

The big gaps region is formed by gaps of size bigger or equal than ε in the Cantorian foliation of primary KAM tori. These gaps are bigger than the size ε of the heteroclinic jumps provided by the scattering map (1.20). This is what is known in the literature as the *large gap problem*. Inside these regions, apart from primary KAM tori which are bent, there appear other invariant objects, which were not present in the unperturbed case, like secondary KAM tori and lower-dimensional tori, which are not detected by a direct application of KAM Theorem, but require a more careful analysis based on an averaging procedure.

In order to prove Theorem 1.3.1 we will restrict the Hamiltonian (1.4) to the NHIM $\widetilde{\Lambda}_\varepsilon$ and perform an averaging procedure before applying a quantitative version of KAM Theorem. The fundamental difference with respect to [DLS06a] is that for every fixed ε it will be necessary to truncate adequately the perturbation in order to deal with a finite number of harmonics depending on ε. The truncated Hamiltonian possesses an heterogeneous sea of finite gaps of different sizes, depending on the size of the harmonics of the perturbation.

Restriction to the NHIM $\widetilde{\Lambda}_\varepsilon$.

Following the same arguments given in sections 8.1 and 8.2 in [DLS06a], we have that the flow restricted to $\widetilde{\Lambda}_\varepsilon$ is Hamiltonian. More precisely, by proposition 8.2 of [DLS06a], we can construct a $\mathcal{C}^r$ system of coordinates (J, φ, s) on $\widetilde{\Lambda}_\varepsilon$, where

$$J = \mathcal{J}(I, \varphi, s; \varepsilon) = I + \mathcal{O}_{\mathcal{C}^{r-1}}(\varepsilon), \tag{1.22}$$

such that the symplectic form on any $\Lambda_\varepsilon^s = \{(J', \varphi', s') \in \widetilde{\Lambda}_\varepsilon : s' = s\}$ has the standard expression $\omega_{|\Lambda_\varepsilon^s} = dJ \wedge d\varphi$. Since $\widetilde{\Lambda}_\varepsilon = \widetilde{\Lambda}$ for $\varepsilon = 0$ according to equation (1.16), by proposition 8.4 in [DLS06a], the restriction of the Hamiltonian H_ε in (1.4) to $\widetilde{\Lambda}_\varepsilon$ expressed in this action-angle coordinates (J, φ, s) has the form

$$k(J, \varphi, s; \varepsilon) = Z(J) + \varepsilon R(J, \varphi, s; \varepsilon) \tag{1.23}$$

with

$$Z(J) = J^2/2 \quad \text{and} \quad R(J, \varphi, s; 0) = h(0, 0, J, \varphi, s; 0), \tag{1.24}$$

where h is the perturbation in H_ε given in (1.6) and R is $\mathcal{O}_{\mathcal{C}^r}(1)$.

Remark 1.2.8. Notice that, by (1.24), $R_{k,l}(J; 0) = h_{k,l}(0, 0, J; 0)$, where $h_{k,l}$ and $R_{k,l}$ are the Fourier coefficients in the angle variables (φ, s) of the perturbation h and its restriction R to $\widetilde{\Lambda}_\varepsilon$, respectively.

Averaging procedure.

We start performing an averaging procedure to the restricted Hamiltonian (1.23), as it was done in [DLS06a], which follows the argument used in the proof of KAM Theorem in [Arn63a], but paying attention to resonant regions. In [DLS06a] the perturbation was assumed to be a trigonometric polynomial, so there was only a finite number of resonances. However, in Hamiltonian (1.4) the perturbation h has an infinite number of harmonics, in the same way as R in (1.23), which give rise to an infinite number of resonances, so the results in [DLS06a] do not apply directly.

The main result for the implementation of an averaging procedure for a generic perturbation will be Theorem 1.3.11 in Section 1.3.2. This theorem makes precise the hypotheses required to truncate the Fourier series of the perturbation R in (1.23) with respect to the angle variables and develop a global averaging procedure that casts the Hamiltonian (1.23) into a global normal form that has different expressions in the non-resonant and resonant regions. The main property of the normal form is that it is almost ready to apply on it a quantitative version of KAM Theorem.

The precise statement and rigorous proof of Theorem 1.3.11 are postponed to Section 1.3.2. In the following we only describe its main features and the results needed to apply KAM Theorem.

There are three parameters that play an important role in the averaging procedure of Theorem 1.3.11. One is the number of steps of averaging m to be performed, which imposes a restriction on the differentiability r of the perturbation: $r > 2(m+1)^2$. This number of averaging steps is chosen later in the application of KAM Theorem. The other two are M, which is the order of truncation of the Fourier series and L, which determines the size of the resonant regions. Both of them are chosen to depend on ε in the following way: $M \sim \varepsilon^{-\rho}$ and $L \sim \varepsilon^{\alpha}$ with $\rho, \alpha > 0$, which will be chosen conveniently during this averaging procedure.

For every fixed ε, we truncate the Fourier series of the perturbation R in (1.23) with respect to the angle variables (φ, s) up to order M in the following way

$$R = R^{[\leq M]} + R^{[>M]},$$

where

$$R^{[\leq M]}(J, \varphi, s; \varepsilon) = \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} R_{k,l}(J; \varepsilon) e^{\mathrm{i}(k\varphi + ls)}, \tag{1.25}$$

and

$$R^{[>M]}(J, \varphi, s; \varepsilon) = \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|>M}} R_{k,l}(J; \varepsilon) e^{\mathrm{i}(k\varphi + ls)}, \tag{1.26}$$

and we deal only with $R^{[\leq M]}$, which is the trigonometric polynomial of degree M, as a perturbation. The error introduced in Hamiltonian (1.23) coming from the neglected tail of the Fourier series will have to be estimated later on.

Since the truncated Hamiltonian $R^{[\leq M]}$ has a finite number of harmonics, an averaging procedure of m steps has to take into account a finite number of resonances, which are the set of rational numbers $J = -l/k$ with $|l| + |k| \leq mM$ (see definitions 1.3.6 and 1.3.4 for more precision).

This averaging procedure divides the phase space (J, φ, s) in two types of domains. On the one hand, the **non-resonant regions up to order** m $\mathcal{D}^m_{\rm nr}$, which are the set of points (J, φ, s) such that its action variable J is at a distance greater than $2L_k$ of any resonance $J = -l/k$, where $L_k = L/|k|$. On the other hand, the **resonant regions up to order** m $\mathcal{D}^m_{\rm r}$, which are the set of points (J, φ, s) such that its action variable J is at a distance smaller than L_k of any resonance $J = -l/k$ (see definitions 1.3.7 and 1.3.9 for more precision).

To avoid overlapping between all the resonant domains, the distance between a resonance $-l_0/k_0$ and any other $-l/k$ must be greater than $2(L_{k_0} + L_k)$. Since the resonances considered satisfy $|k| \leq mM$ we need to impose $4L < 1/mM$, which requires $\rho \leq \alpha$ in terms of exponents of ε and this corresponds to the left hand side inequality of hypothesis (1.56) in Theorem 1.3.11.

Along the averaging procedure, one needs to control the $\mathcal{C}^\ell$ norms of the averaged terms and the remainders, for $0 \leq \ell \leq n$ and $2m < n < r$, where n is the regularity which will be needed for the KAM Theorem and r is the regularity of the perturbation R in (1.23). It turns out that the estimates for the $\mathcal{C}^\ell$ norm blow up as a negative power of $L \sim \varepsilon^\alpha$. Since the averaged terms and the remainder contain a power of ε in front of them, bounds for them can be kept small provided that α is small enough, that is for $\alpha < 1/n$. This corresponds to the right hand side inequality of hypothesis (1.56) in Theorem 1.3.11 and also implies $\rho < 1/n$, which is formula (1.52) in the hypotheses of Theorem 1.3.11.

In all this averaging procedure, there was an initial error coming from the neglected tail of the truncation of order M of the perturbation R in Hamiltonian (1.23), whose $\mathcal{C}^\ell$ norm can be bounded by $\varepsilon/M^{r-\ell-2}$, where r is the regularity of the perturbation R. To keep it smaller than the $\mathcal{C}^\ell$ norm of the remainder after m steps of averaging, one has to impose a lower bound on ρ, which implies $r \geq (1/\rho - 2)m + 2$ in order to make compatible lower and upper bounds for ρ, and this is hypothesis (1.53) in Theorem 1.3.11.

These conditions on m, ρ, α and r are stated in the hypotheses of Theorem 1.3.11. In it, it is proved that one can develop a global averaging procedure that casts the Hamiltonian into a global normal form (1.57), that has different expressions in the non-resonant and resonant regions (these correspond to expressions (1.58) and (1.59) in theses of Theorem 1.3.11). In the non-resonant regions one can perform non resonant averaging transformations in such a way that the averaged Hamiltonian is very close to a rotor. On the other hand, near the resonances, the resonant averaging transformations cast the system to a one degree of freedom (d.o.f.) Hamiltonian, which is close to an integrable pendulum, provided that the perturbation satisfies some non-degeneracy conditions like **H3'**.

Summing up, we end up with a Hamiltonian that consists of an integrable part $\bar{Z}^m$ (the averaged Hamiltonian) plus a perturbation $\varepsilon^{m+1}\bar{R}^m$ which is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{m+1-\alpha(\ell+2m)})$, for $\ell = 0, \ldots, n - 2m$, where m is the number of steps of averaging performed. Recall that

the integrable Hamiltonian has different expressions in resonant regions and non-resonant regions.

The integrable part of the Hamiltonian gives us an approximate equation $\bar{Z}^m = \text{cte}$ for the invariant tori. The next step is to show which tori survive and what is the distance between them when we add the perturbation term.

Quantitative version of KAM Theorem.

The main tool for this section will be KAM Theorem 1.3.22, which is a result about the existence of invariant tori of a periodic perturbation of a Hamiltonian expressed in action-angle variables. It is a direct adaptation of Theorem 8.12 in [DLS06a]. We will use Theorem 1.3.22 to show that there exists a discrete foliation of invariant tori which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $0 < \eta \leq 1/32$, and give approximate explicit expressions for them.

Since the integrable Hamiltonian (1.57) after m steps of averaging has different expressions in resonant and non-resonant regions (up to order m) introduced along the averaging procedure, we perform this study separately. In the end, we will show that all these regions can be grouped in two according to the expressions for the invariant tori obtained in each one, which are the big gaps region (1.21) and its complementary the flat tori region, already mentioned at the beginning of this subsection. Notice that the big gaps region (1.21) is formed by the resonances $J = -l/k$ of order 1, such that $|(k,l)| \leq M_{\text{BG}}$, whereas flat tori region is composed by the non resonant regions up to order m and the resonant regions up to order m such that $J = -l/k$ and $|(k,l)| > M_{\text{BG}}$, where $M_{\text{BG}} = \varepsilon^{-(1+\nu)/r}$, for $0 < \nu \leq 1/16$.

Non-resonant regions are studied in Section 1.3.3.2. In Proposition 1.3.24, we apply Theorem 1.3.22 directly to Hamiltonian (1.57), which is already written in action-angle variables, and we conclude that for these regions there exist flat primary KAM tori given in (1.79) as the level sets of a flat function $F = I + \mathcal{O}(\varepsilon^{1+\eta})$, which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$, provided that $m \geq 2$ and $n \geq 2m + 6$.

Resonant regions are studied in Section 1.3.3.3. As we already said, for these regions Hamiltonian (1.57) is not written in action-angle variables but it is close to an integrable pendulum provided that hypotheses **H3'** are satisfied. The integrable pendulum has rotational and librational orbits as well as separatrices, which separate these two types of motion. Rotational orbits have the same topology as the primary tori in the integrable system and librational orbits are contractible to a periodic orbit, so they correspond to motions with topologies that were not present in the unperturbed system and they are called secondary tori. Librational orbits cover all the region inside the separatrix loop, giving rise to a gap between primary tori, and the size of this gap depends on the order of the corresponding resonance and the size of the Fourier coefficient associated to it.

When gaps are of size smaller than ε, which is the size of the heteroclinic jumps provided by the scattering map (1.17), they are called small gaps. In Section 1.3.3.4, we study the resonant regions with small gaps $\mathcal{D}_{\text{SG}}$ and in Proposition 1.3.26 we show that we can apply the same argument as in the case of non resonant regions to conclude that for these regions there exist flat primary KAM tori given in (1.88) as the level sets of a flat function $F = I + \mathcal{O}(\varepsilon^{1+\eta})$, which is the same as in the non-resonant case, and which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$, provided that $m \geq 2$ and $n \geq 2m + 6$.

Notice that tori in the non resonant regions and resonant regions with small gaps are given by the level sets of the same function $F = I + \mathcal{O}(\varepsilon^{1+\eta})$ and they are flat up to $\mathcal{O}(\varepsilon^{1+\eta})$, for some $\eta > 0$. Both regions form the flat tori region.

Resonant regions with big gaps $\mathcal{D}_{\rm BG}$ are studied in Section 1.3.3.5. They correspond to resonances $J = -l/k$ such that $|(k,l)| < M_{\rm BG}$, where $M_{\rm BG} = \varepsilon^{-(1+\nu)/r}$, for $0 < \nu \leq 1/16$. The size of the gap for these resonances is $C\varepsilon^{1/2}|(k,l)|^{-r/2}$, where C is a constant independent of ε and (k,l). Note that there is no uniform size of the gaps since it runs from order $\varepsilon^{1/2}$ for resonances with low $|(k,l)|$ to $\varepsilon^{1+\nu/2}$ for resonances with $|(k,l)| \sim M_{\rm BG}$.

Our criterium for the choice of the big gaps has been motivated by the size of the heteroclinic jumps provided by the scattering map (1.20): small gaps are of size smaller than ε, so they can always be traversed just connecting two primary tori by the scattering map, whereas this is not the case for big gaps. For these big gaps, we will show that we can find other invariant objects, like secondary tori, which fill the region inside the gaps and they get rather close to the frontier of the gaps among the primary KAM tori.

Remark 1.2.9. We would like to remark that our result about resonances that create big gaps is remarkably different of the one obtained in [DLS06a], where it was considered the case of a perturbation h with a finite number of harmonics. In that case there was a uniform size for the gaps created by the resonances of order 1 which was $C\varepsilon^{1/2}$. Moreover, for resonances of order 2 the uniform size of the associated gap was $C\varepsilon$. Hence, both resonances of order 1 and 2 were considered as big gaps.

In the case of resonances with big gaps, we will need to write the integrable Hamiltonian $\bar{Z}^m$ in action-angle variables before applying KAM Theorem 1.3.22. Since this change of coordinates becomes singular close to the separatrix of the pendulum, we will need to define different action-angle variables inside and outside the separatrix, and we will exclude a thin neighborhood of the separatrix.

Moreover, since the behavior of the tori outside is different depending on their distance to the separatrix (tori are flatter as they are further from the separatrix) we consider different regions in the outside part of the separatrix, where we perform different scalings. This strategy, which was already introduced in [DLS06a], has been improved introducing a new sequence of domains in Theorem 1.3.30, which reduce the differentiability requirements.

The main result for the implementation of the above strategy for resonances with big gaps is Theorem 1.3.30 jointly with Corollary 1.3.31 which make explicit the relationship between the minimum distance between the surviving tori and the number m of steps of averaging performed.

In Theorem 1.3.28 we use both Theorem 1.3.30 and Corollary 1.3.31 to show that many of the invariant tori (both primary and secondary) of the integrable averaged Hamiltonian persist under the perturbation forming a sequence of tori given in (1.95) as the level sets of a function F, close to the averaged Hamiltonian with a distance between consecutive tori of order $\varepsilon^{1+\eta}$, for some $\eta > 0$, in terms of the action variable, provided that $m \geq 10$ and $n \geq 2m + 6$.

Propositions 1.3.24 and 1.3.26 and Theorem 1.3.28 can be joined in a unique result about the existence of nearby invariant tori for the inner dynamics, which is Theorem 1.3.1. This Theorem also gives explicit approximate expressions for the invariant tori,

which are of two types depending on the region of the phase of space where invariant tori lie: the big gaps region and the flat tori region.

We refer to sections 1.3.2 and 1.3.3 for the referenced theorems where one can find the complete and rigorous proof.

1.2.3.4 Fourth Part: Construction of a transition chain and obstruction property

In order to finish the proof, it remains to check that the finite sequence of invariant tori provided by Theorem 1.3.1 form a transition chain along the NHIM $\widetilde{\Lambda}_\varepsilon$, traversing both big gaps and flat tori regions, and to show that there are orbits that follow it closely. These are the orbits claimed in Theorem 1.2.1.

The scattering map S_ε associated to the homoclinic channel Γ_ε, defined in (1.17), is the main tool to detect that there exist transverse heteroclinic connections between these tori, which are objects of different topology. Indeed, by Lemma 10.4 in [DLS06a], we know that two submanifolds, like the invariant tori $\mathcal{T}_i$, of a NHIM $\widetilde{\Lambda}_\varepsilon$, have a transverse heteroclinic intersection if they are transversal under the scattering map as submanifolds of $\widetilde{\Lambda}_\varepsilon$.

The main result of this section is Proposition 1.4.1 where it is proved the existence of transition chains, that is chains of invariant tori $\mathcal{T}_i$, both primary and secondary, such that their image under the scattering map S_ε in (1.20) intersects transversally $\mathcal{T}_{i+1}$ on $\widetilde{\Lambda}_\varepsilon$, that is

$$S_\varepsilon(\mathcal{T}_i) \pitchfork_{\widetilde{\Lambda}_\varepsilon} \mathcal{T}_{i+1}. \tag{1.27}$$

In Section 1.2.3.2 we have obtained an explicit expression (1.20) up to first order for the scattering map S_ε using the first order calculation of the Hamiltonian function $\mathcal{S}_\varepsilon$. In Section 1.2.3.3 we have shown that on the NHIM $\widetilde{\Lambda}_\varepsilon$ there exists a discrete foliation of KAM tori $\mathcal{T}_i$ (primary and secondary) which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$. Moreover, we have obtained explicit expressions for tori $\mathcal{T}_i$, both primary and secondary, and we have seen that these invariant objects are given approximately by the level sets of the averaged Hamiltonian.

In Lemma 1.4.2 in Section 1.4.1, we will give an expression for the action of the scattering map S_ε on a foliation given by the level sets of a certain function, using the expression for the Hamiltonian function $\mathcal{S}_\varepsilon$ generating the deformation of the scattering map, introduced in Section 1.2.3.2. Moreover, we will give conditions to assure transversality between the foliation in $\widetilde{\Lambda}_\varepsilon$ and its image under the scattering map S_ε.

As we have seen in the previous section, the different types of tori that appear in our problem have different quantitative properties and therefore the dominant terms in the expression of these invariant objects as the level sets of a certain function will be different whether they lie in a flat tori region or a big gaps region. Lemma 1.4.2 is applied in Lemma 1.4.5 for the case of the flat tori region, and in Lemma 1.4.7 for the case of the big gaps region. We will show that the sufficient conditions on the perturbation (1.4) for the transversality are hypotheses **H2"**, **H3"** and **H3"'** in Theorem 1.2.1.

Putting all these results together in Proposition 1.4.1, we have that, by hypothesis **H2"** and the non-degeneracy conditions **H3"** and **H3"'**, the scattering map S_ε maps pieces of these tori transversally in Λ_ε to other tori at a distance $\mathcal{O}(\varepsilon)$, that is $S_\varepsilon(\mathcal{T}_i) \pitchfork \mathcal{T}_{i+1}$, where

$\mathcal{T}_i$ and $\mathcal{T}_{i+1}$ are invariant tori at a distance smaller than ε. Therefore, by Lemma 10.4 in [DLS06a] we have that $W^u_{\mathcal{T}_i} \pitchfork W^s_{\mathcal{T}_{i+1}}$ and we have constructed a transition chain.

Finally, we use the well known result that given a transition chain $\{\mathcal{T}_i\}_{i=0}^N$, we can find an orbit visiting all the elements of the chain. In our case, as it was the case in [DLS06a] we have incorporated in the chain objects with different topologies, so applying Lemma 11.1 in [DLS06a] to the transition chain obtained, we get that there is $\varepsilon^* > 0$ such that for $0 < |\varepsilon| < \varepsilon^*$, and for any interval $[I^*_-, I^*_+] \in (I_-, I_+)$, $\widetilde{x}(t)$ satisfies that, for some $T > 0$

$$I(\widetilde{x}(0)) \leq I^*_-; \qquad I(\widetilde{x}(T)) \geq I^*_+$$

(respectively:

$$I(\widetilde{x}(0)) \geq I^*_+; \qquad I(\widetilde{x}(T)) \leq I^*_-)$$

as we wanted to prove.

1.3 Inner Dynamics

The main goal of this section is to prove Theorem 1.3.1 about the existence of a sequence of invariant tori $\mathcal{T}_i$ in the NHIM $\widetilde{\Lambda}_\varepsilon$, which are distributed along all the actions in the interval (I_-, I_+) and are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$. The method of proof will consist of the combination of two parts: averaging methods and KAM Theorem.

In Section 1.3.2 we will consider the restricted Hamiltonian (1.23) and perform, in Theorem 1.3.11, a global averaging procedure that casts the Hamiltonian into a global normal form, which has different expressions in the non-resonant and resonant regions. In the non-resonant regions, averaging transformations cast the system close to strongly integrable and, in general, in the non-resonant regions close to an integrable pendulum.

In Section 1.3.3 we will use KAM Theorem 1.3.22 to show that many of the invariant tori of the averaged Hamiltonian persist when we add the error terms of the normal form and they are close enough in terms of the action variables. For the flat tori region, which consists of non-resonant regions and resonant regions with small gaps, we can apply KAM Theorem 1.3.22 almost straightforwardly and this is done in Propositions 1.3.24 and 1.3.26, respectively. For the big gaps region, we will show in Theorem 1.3.28 that we can apply KAM Theorem after we have written the Hamiltonian in action-angle coordinates.

1.3.1 Main result

The main result about the existence of invariant tori in the NHIM $\widetilde{\Lambda}_\varepsilon$ is stated in the following Theorem:

Theorem 1.3.1. *Consider a Hamiltonian of the form* (1.4) *and assume that* $r > 2(m+1)^2$*, with* $m \geq 10$*, as well as hypothesis* **H3'** *on* (k_0, l_0)*. Choose* $\eta = \min((1-\alpha n)/2, \nu/2)$*, where* $\alpha < 1/n$ *and* $0 < \nu \leq 1/16$*. Then, for* ε *small enough, there exists a finite sequence of invariant tori* $\{\mathcal{T}_i\}_{i=0}^N$ *in* $\widetilde{\Lambda}_\varepsilon$ *which are distributed along all the actions in the interval* (I_-, I_+)*, such that*

1. *They are defined by equation* $F(I,\varphi,s;\varepsilon)\equiv E_i$, *where* F *is a* $\mathcal{C}^{4-\varrho}$ *function, for any* $\varrho>0$, *which has the form* (1.88) *and* (1.95) *depending on the region where the invariant tori lie: the flat tori region or a connected component of the big gaps region defined in* (1.82), *respectively. In the flat tori region, the invariant tori are primary whereas in the big gaps region invariant tori can be primary and secondary. In the big gaps region, for values of* $E_i>0$ *equation* (1.95) *provides two primary tori* $\mathcal{T}^{\pm}_{E_i}$, *whereas for* $E_i<0$ *it gives a secondary tori* $\mathcal{T}_{E_i}$.

2. *They can be also written as a graph of the variable* I *over the angle variables* (φ,s): $I=\lambda_E(\varphi,s;\varepsilon)$ *with* λ_E *given in* (1.89) *for the flat tori region. In the case of the big gaps region, the equations for them are given for two different invariant tori* $\mathcal{T}_i^{\pm}$ *(two different components in the case of secondary KAM tori) in the form* $I=\lambda_E^{\pm}(\varphi,s;\varepsilon)$, *with* $\lambda_E^{\pm}$ *given in* (1.96).

3. *These tori are* $\mathcal{O}(\varepsilon^{1+\eta})$*-closely spaced in terms of the action variable* I, *for some* $0<\eta\leq 1/32$. *In the connected component* (1.82) *of the big gaps region they are* $\mathcal{O}(\varepsilon^{3/2+1+\eta}|(k_0,l_0)|^{-r/2+1})$*-closely spaced in terms of energies, where* $-l_0/k_0$ *is the associated resonance.*

4. $\mathcal{T}_0$ *and* $\mathcal{T}_N$ *are* $\mathcal{O}_{\mathcal{C}^2}(\varepsilon^{1+\eta})$*-close to* I_- *and* I_+, *respectively.*

The proof of the Theorem 1.3.1 is a combination of an averaging procedure (Section 1.3.2) and a KAM Theorem (Section 1.3.3). In Section 1.3.4 we put the results obtained in the previous sections together to give a proof of Theorem 1.3.1.

1.3.2 Averaging procedure

In this section we proceed to obtain a suitable global normal form of the restricted Hamiltonian (1.23), according to the procedure described in Section 1.2.3.3. We use the standard formalism of Lie Series, so we are considering canonical transformations obtained as the time-one map of a Hamiltonian. A very pedagogical treatment of this method can be found in [LM88]. As we have already mentioned, we consider a truncation of the Fourier Series of the perturbation and we deal with trigonometric polynomials of a finite order. We first introduce a Banach space with a suitable norm, which allows an efficient study of the estimates for the different terms that appear in the averaging procedure.

1.3.2.1 Preliminaries. Functional Spaces

We consider the space of functions defined on $\mathcal{I}\times\mathbb{T}^2$, $\mathcal{I}\subset\mathbb{R}$ compact set, which consists of trigonometric polynomials of order M on $(\varphi,s)\in\mathbb{T}^2$, and $\mathcal{C}^r$ with respect to $J\in\mathcal{I}\subset\mathbb{R}$. We denote this space $\mathcal{T}_M(\mathcal{I}\times\mathbb{T}^2)$. A function $u\in\mathcal{T}_M(\mathcal{I}\times\mathbb{T}^2)$ is of the form

$$u(J,\varphi,s)=\sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} u_{k,l}(J)e^{\mathrm{i}(k\varphi+ls)}. \tag{1.28}$$

Remark 1.3.2. Note that the product of two elements $u \in \mathcal{T}_M(\mathcal{I}\times\mathbb{T}^2)$ and $v \in \mathcal{T}_N(\mathcal{I}\times\mathbb{T}^2)$ is another trigonometric polynomial in the variables $(\varphi, s) \in \mathbb{T}^2$ but of degree $M+N$, that is, $uv \in \mathcal{T}_{M+N}(\mathcal{I}\times\mathbb{T}^2)$.

Clearly, the space $\mathcal{T}_M(\mathcal{I}\times\mathbb{T}^2)$ is a closed subset of $\mathcal{C}^r(\mathcal{I}\times\mathbb{T}^2)$. Therefore, $\mathcal{T}_M(\mathcal{I}\times\mathbb{T}^2)$ is a Banach space with the $\mathcal{C}^r$ norm introduced in (1.2).

Moreover, since the functions u are trigonometric polynomials in (φ, s), we can consider the expression (1.28) and deal with the Fourier norm:

$$\|u\|^{[\leq M]}_{\mathcal{C}^\ell(\mathcal{I}\times\mathbb{T}^2)} := \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} |u_{k,l}|_{\mathcal{C}^n(\mathcal{I})}|(k,l)|^{m-n} \tag{1.29}$$

where $|u_{k,l}|_{\mathcal{C}^n(\mathcal{I})}$ is defined in (1.2) and $|(k,l)| = \max(|k|,|l|)$, and $|\cdot|$ denotes the standard modulo. When there is no possibility of confusion we will abreviate it as $\|\cdot\|_{\mathcal{C}^\ell}$.

On the other hand, to understand better the behavior of the function u with respect to the variable J when it gets closer to the resonances, we will use the Fourier norm with a weight $L \leq 1$:

$$\|u\|^{[\leq M]}_{\mathcal{C}^\ell(\mathcal{I}\times\mathbb{T}^2),L} := \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} |u_{k,l}|_{\mathcal{C}^n(\mathcal{I}),L}|(k,l)|^{m-n} \tag{1.30}$$

where $|(k,l)|$ is as before and

$$|u_{k,l}|_{\mathcal{C}^n(\mathcal{I}),L} := \sum_{i=0}^{n} L^i \frac{|D^i u_{k,l}|_{\mathcal{C}^0(\mathcal{I})}}{i!}.$$

As before, when there is no confusion we will abbreviate these norms as $|\cdot|_{\mathcal{C}^n,L}$ and $\|\cdot\|_{\mathcal{C}^\ell,L}$, respectively.

Note that when $L = 1$, we recover the Fourier norm (1.29).

The basic properties of these norms are collected in Appendix 1.6.2. In particular they are related by

$$L^\ell |u|_{\mathcal{C}^\ell} \leq \|u\|_{\mathcal{C}^\ell,L} \leq CM^2 |u|_{\mathcal{C}^\ell}, \tag{1.31}$$

where C is a constant that depends on ℓ and $0 < L \leq 1$.

For the seminorm $|\cdot|_{j,\ell-j}$ defined in (1.3) one has that for all $0 \leq j \leq \ell$,

$$L^j\, |u|_{j,\ell-j} \leq \|u\|_{\mathcal{C}^\ell,L}\,. \tag{1.32}$$

Note that in the case that the function $u \in \mathcal{T}_M(\mathcal{I}\times\mathbb{T}^2)$ does not depend on the action variable J, we have that

$$|u|_{\mathcal{C}^\ell} = |u|_{0,\ell},$$

therefore by equation (1.32),

$$|u|_{\mathcal{C}^\ell} \leq \|u\|_{\mathcal{C}^\ell,L}\,. \tag{1.33}$$

Moreover, given $u \in \mathcal{T}_M(\mathcal{I}\times\mathbb{T}^2)$ and $v \in \mathcal{T}_N(\mathcal{I}\times\mathbb{T}^2)$, we have that $uv \in \mathcal{T}_{M+N}(\mathcal{I}\times\mathbb{T}^2)$ and for $0 < L \leq 1$ and $0 \leq \ell \leq r$,

$$\|uv\|^{[\leq M+N]}_{\mathcal{C}^\ell,L} \leq \|u\|^{[\leq M]}_{\mathcal{C}^\ell,L}\, \|v\|^{[\leq N]}_{\mathcal{C}^\ell,L}\,. \tag{1.34}$$

We will say that a function f is $\mathcal{O}_{\mathcal{C}^r,L}(\eta)$ when $\|f\|_{\mathcal{C}^r,L} \preceq \eta$.

1.3.2.2 The homological equation

In this section, we will use the standard formalism of Lie series to perform a resonant averaging procedure. We first start discussing the infinitesimal equations for averaging, which will serve as a motivation for the phenomenon of resonances and therefore for the resonant averaging.

We begin with a Hamiltonian $K(J, A, \varphi, s) = K_0(J, A) + \varepsilon K_1(J, A, \varphi, s)$, where $(J, A, \varphi, s) \in \mathbb{R}^2 \times \mathbb{T}^2$ and $K_0(J, A) = A + J^2/2$. We start looking for a canonical transformation g, given by the time-one map flow of a Hamiltonian εG (generating function), that eliminates, when it is possible, the dependence on the angle variables (φ, s) up to order ε. Therefore,

$$\begin{aligned} K \circ g &= K + \{K, \varepsilon G\} + \frac{1}{2}\{\{K, \varepsilon G\}, \varepsilon G\} + \dots \\ &= K_0 + \varepsilon(K_1 + \{K_0, G\}) + \mathcal{O}(\varepsilon^2) \end{aligned}$$

where $\{,\}$ denotes the Poisson bracket in the canonical coordinates (J, A, φ, s):

$$\{f, g\} = \frac{\partial f}{\partial \varphi}\frac{\partial g}{\partial J} + \frac{\partial f}{\partial s}\frac{\partial g}{\partial A} - \frac{\partial f}{\partial J}\frac{\partial g}{\partial \varphi} - \frac{\partial f}{\partial A}\frac{\partial g}{\partial s}.$$

So, we need to solve for G the infinitesimal equation

$$K_1 + \{K_0, G\} = \bar{K},$$

where $\bar{K}$ is chosen to be as simple as possible. In Fourier coefficients this equation has the form

$$K_{k,l}(J) - \mathrm{i}(\omega(J) \cdot (k, l))G_{k,l}(J) = \bar{K}_{k,l}(J) \tag{1.35}$$

where $K_{k,l}(J)$, $G_{k,l}(J)$ and $\bar{K}_{k,l}(J)$ are the Fourier coefficients of K_1, G and $\bar{K}$, respectively, for $(k, l) \in \mathbb{Z}^2$, and $\omega(J) \in \mathbb{R}^2$ is of the form

$$\omega(J) = \left(\frac{\partial K_0}{\partial J}, \frac{\partial K_0}{\partial A}\right) = (J, 1).$$

This vector $\omega(J)$ is called *resonant* when $(J, 1) \cdot (k, l) = Jk + l = 0$, for $(k, l) \neq (0, 0)$; and the values $J = -l/k$, with $k \neq 0$, for which this equation vanishes and $K_{k,l}(-l/k) \neq 0$ are called *resonances*. Looking at equation (1.35) it is clear that these are the places where we can not choose $G_{k,l}(J)$ in order to have $\bar{K}_{k,l}(J) \equiv 0$. So, for these values of J and, in order to keep smoothness, the ones in a neighborhood around them, we will choose $\bar{K}_{k,l}(J)$ to be the Fourier term $K_{k,l}(-l/k)$. Note that we cannot have $\bar{K}_{0,0}(J) \equiv 0$ for any J either, so we will also keep the Fourier coefficient $K_{0,0}(J)$.

The precise result with the estimates for the functions is formulated in the following Lemma:

Lemma 1.3.3. *Given $M > 0$, let $K(J, \varphi, s)$ be a Hamiltonian, which is a $\mathcal{C}^{r+1}$ function with respect to J and a trigonometric polynomial in (φ, s) of degree M, so it can be expressed in the following way*

$$K(J, \varphi, s) = \sum_{(k,l)\in\mathcal{N}} K_{k,l}(J)e^{\mathrm{i}(k\varphi + ls)},$$

with $\mathcal{N} = \{(k,l) \in \mathbb{Z}^2, |k|+|l| \leq M\}$. *We refer to resonances as the elements of the finite set of rational numbers*

$$\mathcal{R} = \{-l/k \in \mathbb{Q} : (k,l) \in \mathcal{N}, k \neq 0, K_{k,l}(-l/k) \neq 0\}. \tag{1.36}$$

For any $(k,l) \in \mathcal{N}$, *we consider* $(\widetilde{k},\widetilde{l}) \in \mathbb{Z}^2$ *such that* $-l/k = -\widetilde{l}/\widetilde{k}$ *and* $\gcd(\widetilde{k},\widetilde{l}) = 1$ *and we define* $L_k = L_{\widetilde{k}} = L/|\widetilde{k}|$, *being* L *some constant small enough such that for all* $-l/k \in \mathcal{R}$, *the real intervals* $[-l/k - 2L_k, -l/k + 2L_k]$ *are all disjoint.*

Then, there exist a function $G = G^{[\leq M]}$ *of class* $\mathcal{C}^r$ *with respect to* J *and* $\bar{K} = \bar{K}^{[\leq M]}$ *of class* $\mathcal{C}^{r+1}$, *which are both trigonometric polynomials in* (φ, s), *such that they solve the homological equation*

$$K + \{K_0, G\} = \bar{K}, \tag{1.37}$$

and verify:

1. *If* $|J + l/k| \geq 2L_k$ *for any* $(k,l) \in \mathcal{N}$, *then*

$$\bar{K}(J,\varphi,s) = K_{0,0}(J). \tag{1.38}$$

2. *If* $|J + l_0/k_0| \leq L_{k_0}$ *for some* $(k_0,l_0) \in \mathcal{N}$, *then*

$$\begin{aligned} \bar{K}(J,\varphi,s) &= K_{0,0}(J) + \sum_{\substack{t\in\mathbb{Z}-\{0\} \\ |t|(|k_0|+|l_0|)\leq M}} K_{tk_0,tl_0}(-l_0/k_0) e^{\mathrm{i}t(k_0\varphi + l_0 s)} \\ &=: K_{0,0}(J) + U_{k_0,l_0}(k_0\varphi + l_0 s). \end{aligned} \tag{1.39}$$

3. *The function* $\bar{K}$ *verifies*

$$\left\|\bar{K}\right\|_{\mathcal{C}^\ell,L} \leq C_\ell \left\|K\right\|_{\mathcal{C}^\ell,L}, \tag{1.40}$$

for $\ell = 0, \ldots, r+1$, *where* C_ℓ *is a constant independent of* L.

4. *The function* G *verifies*

$$\|G\|_{\mathcal{C}^\ell,L} \leq \frac{C_\ell}{L} \|K\|_{\mathcal{C}^{\ell+1},L} \tag{1.41}$$

for $\ell = 0, \ldots, r$, *where* C_ℓ *is a constant independent of* L.

Proof. We want to solve for each $(k,l) \in \mathcal{N}$ the equation

$$K_{k,l}(J) - \mathrm{i}(Jk + l) G_{k,l}(J) = \bar{K}_{k,l}(J), \tag{1.42}$$

where the unknowns are the Fourier coefficients of the generating function G and the averaged Hamiltonian $\bar{K}$.

So, we first choose:

1. $\bar{K}_{0,0}(J) = K_{0,0}(J)$,
2. for $(0,l) \in \mathcal{N}$, $l \neq 0$, $\bar{K}_{0,l}(J) = 0$,

3. if $(k,l) \neq (0,0) \in \mathcal{N}$, $k \neq 0$, we choose $\bar{K}_{k,l}(J)$ as

$$\bar{K}_{k,l}(J) = K_{k,l}(-l/k)\psi\left(\frac{1}{L_k}(J + l/k)\right), \tag{1.43}$$

where $\psi(x)$ is a fixed $\mathcal{C}^\infty$ function such that: $\psi(x) = 1$, if $x \in [-1,1]$, and $\psi(x) = 0$, if $x \notin [-2,2]$. With this choice we have that $\bar{K}_{k,l}$ verifies:

(a) If $|J + l/k| \leq L_k$ then $\bar{K}_{k,l}(J) = K_{k,l}(-l/k)$,

(b) if $|J + l/k| \geq 2L_k$ then $\bar{K}_{k,l}(J) = 0$.

Once we have defined $\bar{K}$ as above, it is clear that it is a $\mathcal{C}^{r+1}$ function, and its Fourier coefficients satisfy:

$$\begin{aligned}
\left|\bar{K}_{k,l}\right|_{\mathcal{C}^n,L} &= \sum_{i=0}^{n} L^i \frac{\left|D^i \bar{K}_{k,l}\right|_{\mathcal{C}^0}}{i!} \\
&\leq \sum_{i=0}^{n} \frac{L^i}{i!} \frac{|K_{k,l}(-l/k)|\,|D^i\psi|_{\mathcal{C}^0}}{L_k^i} \\
&\leq |K_{k,l}|_{\mathcal{C}^0}\,|k|^n \sum_{i=0}^{n} \frac{|D^i\psi|_{\mathcal{C}^0}}{i!} \\
&= |K_{k,l}|_{\mathcal{C}^0}\,|k|^n\,|\psi|_{\mathcal{C}^n}
\end{aligned} \tag{1.44}$$

Using this inequality for the Fourier coefficients it is easy to see that $\bar{K}$ verifies the desired bound (1.40). More precisely,

$$\begin{aligned}
\left\|\bar{K}\right\|_{\mathcal{C}^\ell,L} &= \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} \left|\bar{K}_{k,l}\right|_{\mathcal{C}^n,L} |(k,l)|^{m-n} \\
&\leq \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} |\psi|_{\mathcal{C}^n}\,|K_{k,l}|_{\mathcal{C}^0,L}\,|k|^n |(k,l)|^{m-n} \\
&\leq |\psi|_{\mathcal{C}^\ell} \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} |K_{k,l}|_{\mathcal{C}^0,L}\,|(k,l)|^m \\
&\leq (\ell+1)\,|\psi|_{\mathcal{C}^\ell}\,\|K\|_{\mathcal{C}^\ell,L}
\end{aligned}$$

for $\ell = 0,\ldots,r+1$, so choosing $C_\ell = (\ell+1)\,|\psi|_{\mathcal{C}^\ell}$, which is independent of L, we get the desired bound.

Now, we choose G to verify equation (1.42) and we get:

1. $G_{0,0} = 0$ and $G_{k,l}(J) = 0$ if $(k,l) \notin \mathcal{N}$,

2. $G_{0,l}(J) = K_{0,l}(J)/\mathrm{i}l$,

3. if $(k,l) \in \mathcal{N}$, $k \neq 0$, we choose $G_{k,l}(J)$ as:

   (a) If $J \neq -l/k$ then $G_{k,l}(J) = \mathrm{i}\dfrac{\bar{K}_{k,l}(J) - K_{k,l}(J)}{Jk+l}$,

   (b) $G_{k,l}(-l/k) = \lim\limits_{J\to -l/k} \dfrac{K_{k,l}(J) - \bar{K}_{k,l}(J)}{\mathrm{i}(Jk+l)} = \dfrac{K'_{k,l}(-l/k)}{\mathrm{i}k}$.

Then $G(J,\varphi,s)$ is a trigonometric polynomial in (φ,s) of degree M, and of class $\mathcal{C}^r$ with respect to J. To bound the function G, we first need to bound its Fourier coefficients in terms of $|\cdot|_{\mathcal{C}^\ell,L}$ norm for $0 \leq \ell \leq r$. Given a fixed $(k_0,l_0) \in \mathcal{N}$, by the definition of $\bar{K}$ and G, we have:

1. $\forall J$, $|G_{0,l}|_{\mathcal{C}^n,L} \leq |K_{0,l}|_{\mathcal{C}^n,L}/|l|$, for $\ell = 0,\ldots,r$.

2. If $|J + l_0/k_0| \leq L_{k_0}$, then $|G_{k_0,l_0}|_{\mathcal{C}^n,L} \leq (n+1)\dfrac{|K_{k_0,l_0}|_{\mathcal{C}^{n+1},L}}{L|k_0|}$, for $n = 0,\ldots,r$.

   This estimate comes from

$$\begin{aligned}
|G_{k_0,l_0}|_{\mathcal{C}^n,L} &= \sum_{i=0}^{n} L^i \frac{|D^i G_{k_0,l_0}|_{\mathcal{C}^0}}{i!} \\
&\leq \sum_{i=0}^{n} \frac{L^i}{i!} \frac{|D^{i+1} K_{k_0,l_0}|_{\mathcal{C}^0}}{|k_0|} \\
&\leq \frac{(n+1)}{L|k_0|} \sum_{i=0}^{n} \frac{L^{i+1}}{(i+1)!} \left|D^{i+1} K_{k_0,l_0}\right|_{\mathcal{C}^0} \\
&\leq (n+1)\frac{|K_{k_0,l_0}|_{\mathcal{C}^{n+1},L}}{L|k_0|}.
\end{aligned}$$

3. If $|J+l_0/k_0| \geq 2L_{k_0}$ then $|G_{k_0,l_0}|_{\mathcal{C}^n} \leq \dfrac{n+1}{2L} \sum\limits_{i=0}^{\ell} |K_{k_0,l_0}|_{\mathcal{C}^i,L}\, |k_0|^{n-i}$, for $n = 0,\ldots,r+1$.

   This estimate is obtained using Leibniz rule for derivatives in the following way:

$$\begin{aligned}
|G_{k_0,l_0}|_{\mathcal{C}^n,L} &= \sum_{i=0}^{n} L^i \frac{|D^i G_{k_0,l_0}|_{\mathcal{C}^0}}{i!} \\
&\leq \sum_{i=0}^{n} \frac{L^i}{i!} \sum_{j=0}^{i} \binom{i}{j} \frac{|D^j K_{k_0,l_0}|_{\mathcal{C}^0}}{(2L_{k_0})^{i-j+1}|k_0|} \\
&\leq \sum_{i=0}^{n} \frac{1}{L} \sum_{j=0}^{i} L^j \frac{|D^j K_{k_0,l_0}|_{\mathcal{C}^0}}{j!} |k_0|^{i-j} \\
&\leq \frac{n+1}{L} \sum_{i=0}^{n} L^i \frac{|D^i K_{k_0,l_0}|_{\mathcal{C}^0}}{i!} |k_0|^{n-i} \\
&\leq \frac{n+1}{L} \sum_{i=0}^{n} |K_{k_0,l_0}|_{\mathcal{C}^i,L}\, |k_0|^{n-i}.
\end{aligned}$$

4. If $L_{k_0} \leq |J + l_0/k_0| \leq 2L_{k_0}$, then

$$|G_{k_0,l_0}|_{\mathcal{C}^n,L} \leq \frac{n+1}{L} \sum_{i=0}^{n} |K_{k_0,l_0}|_{\mathcal{C}^i,L} |k_0|^{n-i} + (n+1) |K_{k,l}|_{\mathcal{C}^0} |k|^n |\psi|_{\mathcal{C}^n},$$

for $n = 0, \ldots, r$ and C is a constant independent of L.

This estimate can be obtained in the same way as the previous one using the estimate obtained for $\bar{K}_{k,l}$ in (1.44), in the following way

$$\begin{aligned} |G_{k_0,l_0}|_{\mathcal{C}^n,L} &= \sum_{i=0}^{n} L^i \frac{|D^i G_{k_0,l_0}|_{\mathcal{C}^0}}{i!} \\ &\leq \sum_{i=0}^{n} \frac{L^i}{i!} \left| D^i \left(\mathrm{i} \frac{\bar{K}_{k_0,l_0}}{Jk_0 + l_0} \right) \right|_{\mathcal{C}^0} + \sum_{i=0}^{n} \frac{L^i}{i!} \left| D^i \left(\mathrm{i} \frac{K_{k_0,l_0}}{Jk_0 + l_0} \right) \right|_{\mathcal{C}^0} \\ &\leq \frac{(n+1)^2}{L} |K_{k,l}|_{\mathcal{C}^0} |k|^n |\psi|_{\mathcal{C}^n} + \frac{n+1}{L} \sum_{i=0}^{n} |K_{k_0,l_0}|_{\mathcal{C}^i,L} |k_0|^{n-i}. \end{aligned}$$

In order to finish the proof, we will use these estimates for the Fourier coefficients of G to bound the function G.

If $|J + l_0/k_0| \leq L_{k_0}$ for some $-l_0/k_0 \in \mathcal{R}$, then for any other $(k,l) \in \mathcal{N}$, $(k,l) \neq (tk_0, tl_0)$, $t \in \mathbb{Z}$, J satisfies that $|J + l/k| \geq 2L_k$. Therefore, we will distinguish three types of Fourier coefficients $G_{k,l}$ of G, which are the ones described in points 1,2 and 3 in this proof. Using their corresponding bounds we have

$$\begin{aligned} \|G\|_{\mathcal{C}^\ell,L} &= \sum_{m=0}^{\ell} \sum_{n=0}^{m} 2^\ell \left(\sum_{l=-M}^{M} |G_{0,l}|_{\mathcal{C}^n,L} |l|^{m-n} + \sum_{\substack{(k,l)\in\mathcal{N} \\ (k,l)\neq t(k_0,l_0)}} |G_{k,l}|_{\mathcal{C}^n,L} |(k,l)|^{m-n} \right. \\ &\quad \left. + \sum_{t\in\mathbb{Z}-\{0\} cr |t|(|k_0|+|l_0|)\leq M} |G_{tk_0,tl_0}|_{\mathcal{C}^n,L} |t(k_0,l_0)|^{m-n} \right) \\ &\leq \sum_{m=0}^{\ell} \sum_{n=0}^{m} 2^\ell \left(\sum_{l=-M}^{M} \frac{|K_{0,l}|_{\mathcal{C}^n,L}}{|l|} |l|^{m-n} \right. \\ &\quad + \sum_{\substack{(k,l)\in\mathcal{N} \\ (k,l)\neq t(k_0,l_0)}} \left(\frac{(n+1)}{L} \sum_{i=0}^{n} |K_{k,l}|_{\mathcal{C}^i,L} |k|^{n-i} \right) |(k,l)|^{m-n} \\ &\quad \left. + \sum_{\substack{t\in\mathbb{Z}-\{0\} \\ |t|(|k_0|+|l_0|)\leq M}} \frac{(n+1)}{L|k_0|} |K_{tk_0,tl_0}|_{\mathcal{C}^{n+1},L} |t(k_0,l_0)|^{m-n} \right) \end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^{\ell}\left(\sum_{l=-M}^{M} |K_{0,l}|_{\mathcal{C}^n,L}\,|l|^{m-n}\right.\\
&\quad + \sum_{\substack{(k,l)\in\mathcal{N}\\(k,l)\neq t(k_0,l_0)}} \frac{(n+1)}{L}\sum_{i=0}^{n} |K_{k,l}|_{\mathcal{C}^i,L}\,|(k,l)|^{m-i}\\
&\quad \left. + \sum_{\substack{t\in\mathbb{Z}-\{0\}\\|t|(|k_0|+|l_0|)\leq M}} \frac{(n+1)}{L}\,|K_{tk_0,tl_0}|_{\mathcal{C}^{n+1},L}\,|t(k_0,l_0)|^{m-n-1}\right)\\
&\leq \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^{\ell}\left(\sum_{l=-M}^{M} |K_{0,l}|_{\mathcal{C}^n,L}\,|l|^{m-n}\right.\\
&\quad + \frac{(n+1)(m+1)}{L} \sum_{\substack{(k,l)\in\mathcal{N}\\(k,l)\neq t(k_0,l_0)}} |K_{k,l}|_{\mathcal{C}^n,L}\,|(k,l)|^{m-n}\\
&\quad \left. + \frac{n+1}{L} \sum_{\substack{t\in\mathbb{Z}-\{0\}\\|t|(|k_0|+|l_0|)\leq M}} |K_{tk_0,tl_0}|_{\mathcal{C}^{n+1},L}\,|t(k_0,l_0)|^{m-n-1}\right)\\
&\leq \|K\|_{\mathcal{C}^\ell,L} + \frac{(\ell+1)}{L}\left((\ell+1)\,\|K\|_{\mathcal{C}^\ell,L} + \|K\|_{\mathcal{C}^{\ell+1},L}\right)\\
&\leq \frac{C_\ell}{L}\,\|K\|_{\mathcal{C}^{\ell+1},L}
\end{aligned}
$$

for $\ell = 0,\ldots,r$, where $C_\ell = 3(\ell+1)^2$ is a constant independent of L.

Analogously, if $L_{k_0} \leq |J + l_0/k_0| \leq 2L_{k_0}$ for some $-l_0/k_0 \in \mathcal{R}$, then for any other $(k,l) \in \mathcal{N}$, $(k,l) \neq (tk_0, tl_0)$, $t \in \mathbb{Z}$, J satisfies that $|J + l/k| \geq 2L_k$. In this case, we will distinguish three types of Fourier coefficients $G_{k,l}$ of G, which are the ones described in points 1,3 and 4 in this proof. Using the same argument as in the previous case, jointly with the bounds for the Fourier coefficients, we have that

$$
\begin{aligned}
\|G\|_{\mathcal{C}^\ell,L} &\leq \|K\|_{\mathcal{C}^\ell,L} + \frac{(\ell+1)^2}{L}\,\|K\|_{\mathcal{C}^\ell,L} + \frac{(\ell+1)^2}{L}\,|\psi|_{\mathcal{C}^\ell}\,\|K\|_{\mathcal{C}^\ell,L} + \frac{(\ell+1)^2}{L}\,\|K\|_{\mathcal{C}^\ell,L}\\
&\leq \frac{C_\ell}{L}\,\|K\|_{\mathcal{C}^\ell,L}
\end{aligned}
$$

for $\ell = 0,\ldots,r$, where $C_\ell = 4(\ell+1)^2$ is a constant independent of L.

And finally, if $|J + l/k| \geq 2L_k$ for any $(k,l) \in \mathcal{N}$, the Fourier coefficients $G_{k,l}$ of G are just the ones described in points 1 and 3. Arguing as before we have

$$
\|G\|_{\mathcal{C}^\ell,L} \leq \|K\|_{\mathcal{C}^\ell,L} + \frac{(\ell+1)^2}{L}\,\|K\|_{\mathcal{C}^\ell,L} \leq \frac{C_\ell}{L}\,\|K\|_{\mathcal{C}^\ell,L}\,,
$$

for $\ell = 0,\ldots,r$, where $C_\ell = 2(\ell+1)^2$ is a constant independent of L.

So putting all these estimates together we get the desired bound (1.41) for the whole domain. □

1.3.2.3 The main averaging result

In this section we apply repeatedly the procedure stated in the previous section to the truncated Fourier series of the perturbation $R^{[\leq M]}$ in (1.25), to get a suitable normal form.

We start the averaging procedure with the Hamiltonian

$$k_0(J,\varphi,s;\varepsilon) = Z^0(J,\varphi,s;\varepsilon) + \varepsilon R^0(J,\varphi,s;\varepsilon),$$

where $Z^0(J,\varphi,s;\varepsilon) = J^2/2$ and $R^0(J,\varphi,s;\varepsilon) = R^{[\leq M]}(J,\varphi,s;\varepsilon)$, which is a trigonometric polynomial of degree M in the angle variables (φ,s).

We will search for a canonical transformation g_0, given by the time-1 map flow of Hamiltonian εG_0 provided by Lemma 1.3.3 that eliminates, when it is possible, the dependence on the angle variables (φ,s) at order ε.

According to expression (1.36), we will refer to *resonances of order 1* as the elements of

$$\mathcal{R}_1 = \{-l/k \in \mathbb{Q} \cap (I_-,I_+), |k|+|l| \leq M, k \neq 0, R^0_{k,l}(-l/k;0) \neq 0\},$$

where $R^0_{k,l}$ are the Fourier coefficients of R^0. For each resonance $-l/k$ in $\mathcal{R}_1$ we will define a strip of size $2L/|k|$, for $L \sim \varepsilon^\alpha$ and $\alpha > 0$, centered on the resonance. We will call *resonant region of order 1* the union of these strips, where the averaging transformation g_0 can not eliminate the dependence on the angle variables, and *non resonant region up to order 1* the complementary region in $\widetilde{\Lambda}_\varepsilon$, where $k_0 \circ g_0$ reduces to contain only the harmonic $R^0_{0,0}(J;0)$ at order ε.

Hence, the Hamiltonian $k_1 = k_0 \circ g_0$ is now of the form

$$k_1(J,\varphi,s;\varepsilon) = Z^1(J,\varphi,s;\varepsilon) + \varepsilon^2 R^1(J,\varphi,s;\varepsilon), \tag{1.45}$$

where the normal form Z^1 is a $\mathcal{C}^r$ function, which has different expressions in the resonant and non resonant regions, and the remainder $\varepsilon^2 R^1$ is a $\mathcal{C}^{r-2}$ function.

Proceeding by induction, we obtain a sequence of Hamiltonians k_{q-1}, for $q \geq 1$, which are normalized up to order ε^{q-1}, that is, in adequate symplectic coordinates Hamiltonian k_{q-1} takes the form

$$k_{q-1}(J,\varphi,s;\varepsilon) = Z^{q-1}(J,\varphi,s;\varepsilon) + \varepsilon^q R^{q-1}(J,\varphi,s;\varepsilon), \tag{1.46}$$

where, as before, the normal form Z^{q-1} is a $\mathcal{C}^{r-2(q-2)}$ function, which has different expressions in the resonant and non resonant regions up to order $q-1$, and the remainder $\varepsilon^q R^{q-1}$ is a $\mathcal{C}^{r-2(q-1)}$ function.

The set of resonances of order q and its associated resonant and non resonant regions up to order q, are defined recursively in the following way:

1.3.2.4 Resonances. Resonant and non resonant regions

Definition 1.3.4. The *set of resonances of order* $q \geq 1$ is the set of rational numbers $r \in \mathcal{R}_q \backslash (\mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{q-1})$, where $\mathcal{R}_q$ is the set of rational numbers $r \in \mathbb{Q} \cap (I_-,I_+)$ which

admit a representation $r = -l/k$ for some integers k, l satisfying $|l| + |k| \leq qM$, such that $R^{q-1}_{k,l}(-l/k; 0) \neq 0$; in symbols,

$$\mathcal{R}_q = \mathcal{R}_q(M) = \left\{ -\frac{l}{k} \in \mathbb{Q} \cap (I_-, I_+) : |k| + |l| \leq qM, k \neq 0, R^{q-1}_{k,l}(-l/k; 0) \neq 0 \right\}, \quad (1.47)$$

where $R^{q-1}_{k,l}$ are the Fourier coefficients of the remainder R^{q-1} in (1.46).

Roughly speaking, we call resonances of order q the places in J where it is impossible to apply q-th order averaging to $k_{q-1} = Z^{q-1} + \varepsilon^q R^{q-1}$.

Remark 1.3.5. Notice that, by hypotheses **H3'** in Theorem 1.2.1, for all $-l_0/k_0 \in \mathbb{Q} \cap (I_-, I_+)$ such that $|(k_0, l_0)| < M_{\rm BG}$ there exists $t^* \in \mathbb{Z}^2$ such that $h_{t^*k_0, t^*l_0}(0, 0, -l_0/k_0; 0) \neq 0$ and therefore, by equation (1.24), $R_{t^*k_0, t^*l_0}(-l_0/k_0; 0) \neq 0$. Hence, by definition 1.3.4 for resonances of order 1, as long as $M_{\rm BG} \leq M$, all the rational numbers $-l/k$ with $|(k, l)| < M_{\rm BG}$ are resonant of order 1.

Definition 1.3.6. The set $\mathcal{R}_{[\leq q]}(M)$ of resonances up to order q is the union of sets of resonances of order i, for $i = 1, \ldots, q$; in symbols,

$$\mathcal{R}_{[\leq q]} = \mathcal{R}_{[\leq q]}(M) = \bigcup_{i=1,\ldots,q} \mathcal{R}_i(M) \subset \mathbb{Q}. \quad (1.48)$$

For this set of resonances we define different strips in $\widetilde{\Lambda}_\varepsilon$ of a width depending on a parameter L, which is $L \sim \varepsilon^\alpha$, with $\alpha > 0$. This divides the phase space in two types of regions:

Definition 1.3.7. The non-resonant region up to order q $\mathcal{D}^q_{\rm nr}$ is the set of points $(J, \varphi, s) \in \widetilde{\Lambda}_\varepsilon$ which are at a distance greater than $2L_k$ in terms of the J variable of any resonance $-l/k \in \mathcal{R}_{[\leq q]}$, where $L_k = L/|k|$; in symbols,

$$\mathcal{D}^q_{\rm nr} = \mathcal{D}^q_{\rm nr}(M, L) = \left\{ (J, \varphi, s) \in (I_-, I_+) \times \mathbb{T}^2 : \left| J + \frac{l}{k} \right| \geq 2L_k, \text{ for } -\frac{l}{k} \in \mathcal{R}_{[\leq q]} \right\}. \quad (1.49)$$

Definition 1.3.8. The resonant region of order q $\mathcal{D}_{r,q}$ is the set of points $(J, \varphi, s) \in \widetilde{\Lambda}_\varepsilon$ which are at a distance smaller than $L_k = L/|k|$ in terms of the J variable from any resonance $-l/k \in \mathcal{R}_q \backslash (\mathcal{R}_1 \cup \ldots \cup \mathcal{R}_{q-1})$; in symbols,

$$\begin{aligned} \mathcal{D}_{r,q} = \mathcal{D}_{r,q}(M, L) \quad = \quad & \left\{ (J, \varphi, s) \in (I_-, I_+) \times \mathbb{T}^2 : \left| J + \tfrac{l}{k} \right| \leq L_k, \right. \\ & \left. \text{for some } -\tfrac{l}{k} \in \mathcal{R}_q \backslash (\mathcal{R}_1 \cup \ldots \cup \mathcal{R}_{q-1}) \right\}. \end{aligned} \quad (1.50)$$

The union of resonant regions of order i, for $i = 1, \ldots, q$ gives us the resonant region up to order q, which can be defined in the following way:

Definition 1.3.9. The resonant region up to order q $\mathcal{D}^q_{\rm r}$ is the set of points $(J, \varphi, s) \in \widetilde{\Lambda}_\varepsilon$ which are at a distance smaller than $L_k = L/|k|$ in terms of the J variable from any resonance $-l/k \in \mathcal{R}_{[\leq q]}$; in symbols,

$$\mathcal{D}^q_{\rm r} = \mathcal{D}^q_{\rm r}(M, L) = \left\{ (J, \varphi, s) \in (I_-, I_+) \times \mathbb{T}^2 : \left| J + \frac{l}{k} \right| \leq L_k, \text{ for some } -\frac{l}{k} \in \mathcal{R}_{[\leq q]} \right\}. \quad (1.51)$$

The dependence of these domains on M and L: $\mathcal{D}^q_{\rm nr} = \mathcal{D}^q_{\rm nr}(M,L)$, $\mathcal{D}_{r,q} = \mathcal{D}_{r,q}(M,L)$ and $\mathcal{D}^q_{\rm r} = \mathcal{D}^q_{\rm r}(M,L)$, will be suppressed to simplify notation.

Remark 1.3.10. Note that, by remark 1.3.5, the big gaps region $\mathcal{D}_{BG}$ introduced in (1.21) is contained in the resonant region of order 1 $\mathcal{D}_{r,1}$.

The precise result to obtain a global normal form for the reduced Hamiltonian by applying repeatedly the averaging procedure, jointly with the estimates for the bounds of the normal form terms and the expression of the order of truncation M and the constant L as functions of ε are stated in the following Theorem 1.3.11:

Theorem 1.3.11. *Let n, m be any given integers satisfying $1 \leq 2m \leq n$. Given ρ a real number satisfying*

$$\rho < \frac{1}{n}, \tag{1.52}$$

and r an integer verifying

$$r \geq (1/\rho - 2)m + 2, \tag{1.53}$$

consider a $\mathcal{C}^r$ Hamiltonian of the form

$$k(J, \varphi, s; \varepsilon) = \frac{J^2}{2} + \varepsilon R(J, \varphi, s; \varepsilon), \tag{1.54}$$

satisfying $\varepsilon R(J, \varphi, s; \varepsilon) = \mathcal{O}_{\mathcal{C}^r}(\varepsilon)$.

Introduce $M \sim \varepsilon^{-\rho}$, for any $-l/k \in \mathcal{R}_{[\leq m]}(M)$, introduced in (1.48), consider $L_k = L/|k|$, where

$$L = C\varepsilon^{\alpha} \tag{1.55}$$

with

$$\rho \leq \alpha < 1/n \tag{1.56}$$

and C a constant independent of ε, such that for $-l/k \in \mathcal{R}_{[\leq m]}$, the real intervals $\mathcal{I}_{-l/k} \equiv [-l/k - 2L_k, l/k + 2L_k]$ are disjoint. Then, there exists a symplectic change of variables, depending on time, $(\mathcal{B}, \phi, s) \mapsto (J, \varphi, s)$, periodic in φ and s, and of class $\mathcal{C}^{r-2m}$, which is ε-close to the identity in the $\mathcal{C}^{n-2m-1}$ sense, such that transforms the Hamiltonian system associated to $k(J, \varphi, s; \varepsilon)$ into a Hamiltonian system of Hamiltonian

$$\bar{k}^m(\mathcal{B}, \phi, s; \varepsilon) = \bar{Z}^m(\mathcal{B}, \phi, s; \varepsilon) + \varepsilon^{m+1}\bar{R}^m(\mathcal{B}, \phi, s; \varepsilon), \tag{1.57}$$

where the function $\bar{Z}^m$ is of class $\mathcal{C}^{r-2m+2}$ and $\bar{R}^m$ is of class $\mathcal{C}^{r-2m}$ and they verify:

1. *If $\mathcal{B} \notin \bigcup_{-l/k \in \mathcal{R}_{[\leq m]}} \mathcal{I}_{-l/k}$, then*

$$\bar{Z}^m(\mathcal{B}, \phi, s; \varepsilon) = \frac{1}{2}\mathcal{B}^2 + \varepsilon \widetilde{Z}^m(\mathcal{B}; \varepsilon), \tag{1.58}$$

for any $(\mathcal{B}, \phi, s) \in \mathcal{D}^m_{\rm nr}$ ($\mathcal{D}^m_{\rm nr}$ was introduced in (1.49)).

2. *If $\mathcal{B} \in \mathcal{I}_{-l_0/k_0}$ for some $-l_0/k_0 \in \mathcal{R}_i \setminus (\mathcal{R}_1 \cup \ldots \mathcal{R}_{i-1})$, for some $1 \leq i \leq m$, then*

$$\bar{Z}^m(\mathcal{B}, \phi, s; \varepsilon) = \frac{1}{2}\mathcal{B}^2 + \varepsilon \widetilde{Z}^m(\mathcal{B}; \varepsilon) + \varepsilon^i U_m^{k_0,l_0}(k_0\phi + l_0 s; \varepsilon), \tag{1.59}$$

for any $(\mathcal{B}, \phi, s) \in \mathcal{D}_{\mathrm{r},i}$, ($\mathcal{D}_{\mathrm{r},i}$ was introduced in (1.50)).

In a particular case of a resonance $-l_0/k_0$ of order 1, $U_m^{k_0,l_0}(k_0\phi + l_0 s; 0)$ does not depend on m and is given by

$$U_m^{k_0,l_0}(\theta; 0) = \sum_{\substack{t \in \mathbb{Z} - \{0\} \\ |t|(|k_0|+|l_0|) \leq M}} R_{tk_0,tl_0}(-l_0/k_0; 0) e^{\mathrm{i}t\theta} \tag{1.60}$$

where $\theta = k_0\phi + l_0 s$ and $R_{k,l}(J; \varepsilon)$ are the Fourier coefficients of the perturbation $R(J, \varphi, s; 0)$ with respect to (φ, s).

3. *The function $\varepsilon\widetilde{Z}^m(\mathcal{B}; \varepsilon)$ in (1.58) and (1.59) is a polynomial of degree m in ε, whose term of order $q+1$ is of class $\mathcal{C}^{r-2q}$ and of size $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{q+1-\alpha(\ell+2q)})$, for $\ell = 0, \ldots, n-2q$ and $q = 0, \ldots, m-1$. The function $\varepsilon^i U_m^{k_0,l_0}(k_0\phi + l_0 s; \varepsilon)$ in (1.59) is a polynomial of degree m in ε and a trigonometric polynomial in $\theta = k_0\phi + l_0 s$, which is $\mathcal{O}_{\mathcal{C}^\ell,\theta}(\varepsilon^{i-2\alpha(i-1)}|(k_0, l_0)|^{-r+2(i-1)})$, for $\ell = 0, \ldots, n-2(i-1)$. Finally, the function $\varepsilon^{m+1}\bar{R}(\mathcal{B}, \phi, s; \varepsilon)$ in (1.57) is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{m+1-\alpha(\ell+2m)})$, for $\ell = 0, \ldots, n-2m$.*

Remark 1.3.12. Notice that, by equation (1.24) and remark 1.2.8, the function $U_m^{k_0,l_0}$ given in (1.60) for a resonance $-l_0/k_0 \in \mathcal{R}_1$ corresponds to the function U^{k_0,l_0} in hypothesis **H3** when $M = M_{\mathrm{BG}}$.

Remark 1.3.13. Note that the bound on the trigonometric polynomial $\varepsilon^i U_m^{k_0,l_0}(\theta; \varepsilon)$, where $\theta = k_0\phi + l_0 s$, is more precise because it incorporates the size of its Fourier coefficients. We use the notation $\mathcal{O}_{\mathcal{C}^\ell,\theta}$ to emphasize that we are bounding the derivatives with respect to the variable θ.

Remark 1.3.14. Notice that although the remainder term $\varepsilon^{m+1}\bar{R}^m$ is $\mathcal{C}^{r-2q}$, it is bounded in the supremum norm $|\cdot|_{\mathcal{C}^\ell}$ for ℓ only up to $n-2q$, for $n < r$, which is enough for the future application of KAM Theorem.

1.3.2.5 Proof of Theorem 1.3.11

The proof of this theorem will follow by the repeated application of the inductive Lemma 1.3.18 m times. Before stating it, we need two previous Lemmas that we will use to prove Lemma 1.3.18 and finally Theorem 1.3.11.

Lemma 1.3.15. *Let $G(J, \varphi, s)$ a Hamiltonian and assume that G is $\mathcal{C}^r$ trigonometric polynomial of order M defined in a compact domain $\mathcal{I} \times \mathbb{T}^2$, with $\mathcal{I} \subset \mathbb{R}$, such that $\sup\{x \in \mathcal{I} \times \mathbb{T}^2\} \leq D$. Consider the $\mathcal{C}^{r-1}$ change of variables defined on $\mathcal{I} \times \mathbb{T}^2$,*

$$(J, \varphi, s) = g_t(\mathcal{B}, \phi, s),$$

given by the time t flow of the Hamiltonian $\varepsilon^p G(J,\varphi,s)$, for some $p\in\mathbb{N}$. Assume that G is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{\eta_\ell})$, η_ℓ being some positive number. Then,

$$\max_{0\le t\le 1}|g_t|_{\mathcal{C}^\ell}\le D_\ell$$

for $\ell=0,\dots,r-1$, D_ℓ being some constant, which depends on the domain and ℓ, but not on ε.

Proof. By the fundamental Theorem of calculus we can write

$$g_t(x)=x+\int_0^t\frac{\partial g_\tau}{\partial\tau}(x)\mathrm{d}\tau=x+\int_0^t \mathrm{J}\nabla G\circ g_\tau(x)\mathrm{d}\tau,$$

where $x=(\mathcal{B},\phi,s)\in\mathcal{I}\times\mathbb{T}^2$ and J is the canonical matrix of the symplectic form $\omega=\mathrm{d}J\wedge\mathrm{d}\varphi+\mathrm{d}A\wedge\mathrm{d}s$. The extra variable A, conjugated to the angle s, was introduced to make apparent the symplectic character of the change of variables.

Using formula (1.175) in Appendix 1.6.3 we obtain

$$\begin{aligned}|g_t|_{\mathcal{C}^\ell} &\le |\mathrm{Id}|_{\mathcal{C}^\ell}+\int_0^1|\mathrm{J}\nabla G\circ g_\tau|_{\mathcal{C}^\ell}\,\mathrm{d}\tau\\ &\le |\mathrm{Id}|_{\mathcal{C}^\ell}+C_\ell\int_0^1\Big(|\mathrm{J}\nabla G|_{\mathcal{C}^1}\,|g_\tau|_{\mathcal{C}^\ell}+|\mathrm{J}\nabla G|_{\mathcal{C}^\ell}\,|g_\tau|^\ell_{\mathcal{C}^{\ell-1}}\Big)\,\mathrm{d}\tau\end{aligned}\tag{1.61}$$

for $\ell=2,\dots,r-1$, where C_ℓ is a constant depending on ℓ; and

$$|g_t|_{\mathcal{C}^1}\le|\mathrm{Id}|_{\mathcal{C}^1}+\int_0^1|\mathrm{J}\nabla G|_{\mathcal{C}^1}\,|g_\tau|_{\mathcal{C}^1}\,\mathrm{d}\tau.$$

Let us define $a_\ell=\max\limits_{0\le t\le 1}|g_t|_{\mathcal{C}^\ell}$. Then,

$$a_1\le D+\delta_1 a_1,$$

and

$$a_\ell\le D+\delta_1 a_\ell+C_\ell\delta_\ell a^\ell_{\ell-1},\qquad\text{for }\ell\ge 2,$$

with $\delta_\ell=|G|_{\mathcal{C}^{\ell+1}}$. Hence,

$$a_\ell\le\frac{D+\delta_\ell a^\ell_{\ell-1}}{1-\delta_1}\qquad\text{for }\ell\ge 2.$$

Since $\delta_1\sim\varepsilon^{\eta_2}\ll 1$ and $\delta_\ell\sim\varepsilon^{\eta_{\ell+1}}\ll 1$, it is easy to check by induction that $a_\ell\le D_\ell$, for $\ell\ge 1$, D_ℓ being some constant independent of ε. □

Since the averaging procedure is based on the method of Lie transforms, the transformed Hamiltonian will be expressed in terms of Poisson brackets. In the following Lemma 1.3.16 we give an estimate for the bound of the Poisson bracket of two functions, where the second one is a generating function, in terms of the bounds of each one.

Lemma 1.3.16. *Let ρ, α be two positive real numbers, such that $\rho \leq \alpha$ and $M \sim \varepsilon^{-\rho}$. Given $F^p(J,\varphi,s)$ and $G^q(J,\varphi,s)$, two trigonometric polynomials in (φ,s), assume that $F^p(J,\varphi,s)$ is a $\mathcal{C}^n$, $n>0$ function in J and a trigonometric polynomial of degree $M_p = (p+1)M$ and $G^q(J,\varphi,s)$ is a $\mathcal{C}^m$, $m>0$ function in J and a trigonometric polynomial of degree $M_q = (q+1)M$, that satisfy $\|\varepsilon^{p+1}F^p\|_{\mathcal{C}^\ell,L} \preceq \varepsilon^{p+1-\alpha(2p)}$ and $\|\varepsilon^{q+1}G^q\|_{\mathcal{C}^\ell,L} \preceq \varepsilon^{q+1-\alpha(2q+1)}$, for $\ell = 0,\dots,n$, with $\varepsilon>0$. Then $\{F^p,G^q\}$ is a $\mathcal{C}^r$ function in J, for $r=\min(n,m)-1$ and a trigonometric polynomial of degree $M_{\widetilde{p}} = (\widetilde{p}+1)M$ in (φ,s), where $\widetilde{p}=p+q+1$, and $\varepsilon^{\widetilde{p}+1}F^{\widetilde{p}} = \{\varepsilon^{p+1}F^p, \varepsilon^{q+1}G^q\}$ satisfies*

$$\left\|\varepsilon^{\widetilde{p}+1}F^{\widetilde{p}}\right\|_{\mathcal{C}^\ell,L} \preceq \varepsilon^{\widetilde{p}+1-\alpha(2\widetilde{p})},$$

for $\ell = 0,\dots,r$.

Proof. From

$$\{F^p,G^q\} = \frac{\partial F^p}{\partial\varphi}\frac{\partial G^q}{\partial J} - \frac{\partial F^p}{\partial J}\frac{\partial G^q}{\partial\varphi},$$

we have

$$\begin{aligned}\{F^p,G^q\} \;=\; & \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_p}} \mathrm{i}kF^p_{k,l}(J)e^{\mathrm{i}(k\varphi+ls)} \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_q}} \frac{\partial G^q_{k,l}(J)}{\partial J}e^{\mathrm{i}(k\varphi+ls)} \\ & - \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_p}} \frac{\partial F^p_{k,l}(J)}{\partial J}e^{\mathrm{i}(k\varphi+ls)} \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_q}} \mathrm{i}kG^p_{k,l}(J)e^{\mathrm{i}(k\varphi+ls)}\end{aligned}$$

It is clear from this expression that $\{F^p,G^q\}$ is a trigonometric polynomial of degree $M_p+M_q=(p+q+2)M$.

On the other hand, using equation (1.34), it follows that

$$\begin{aligned}&\left\|\{\varepsilon^{p+1}F^p,\varepsilon^{q+1}G^q\}\right\|_{\mathcal{C}^\ell,L} \leq \\ &\left\|\varepsilon^{p+1}\sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_p}} \mathrm{i}kF^p_{k,l}(J)e^{\mathrm{i}(k\varphi+ls)}\right\|_{\mathcal{C}^\ell,L} \left\|\varepsilon^{q+1}\sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_q}} \frac{\partial G^q_{k,l}(J)}{\partial J}e^{\mathrm{i}(k\varphi+ls)}\right\|_{\mathcal{C}^\ell,L} \\ &+ \left\|\varepsilon^{p+1}\sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_p}} \frac{\partial F^p_{k,l}(J)}{\partial J}e^{\mathrm{i}(k\varphi+ls)}\right\|_{\mathcal{C}^\ell,L} \left\|\varepsilon^{q+1}\sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_q}} \mathrm{i}kG^p_{k,l}(J)e^{\mathrm{i}(k\varphi+ls)}\right\|_{\mathcal{C}^\ell,L}\end{aligned}$$

$$\leq \left\| \varepsilon^{p+1} \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_p}} F^p_{k,l}(J) e^{\mathrm{i}(k\varphi+ls)} \right\|_{\mathcal{C}^{\ell+1},L} \frac{1}{L} \left\| \varepsilon^{q+1} \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_q}} G^q_{k,l}(J) e^{\mathrm{i}(k\varphi+ls)} \right\|_{\mathcal{C}^{\ell+1},L}$$

$$+\frac{1}{L} \left\| \varepsilon^{p+1} \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_p}} F^p_{k,l}(J) e^{\mathrm{i}(k\varphi+ls)} \right\|_{\mathcal{C}^{\ell+1},L} \left\| \varepsilon^{q+1} \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M_q}} G^q_{k,l}(J) e^{\mathrm{i}(k\varphi+ls)} \right\|_{\mathcal{C}^{\ell+1},L}$$

$$\leq \frac{2}{L} \left\|\varepsilon^{p+1}F^p\right\|_{\mathcal{C}^{\ell+1},L} \left\|\varepsilon^{q+1}G^q\right\|_{\mathcal{C}^{\ell+1},L}.$$

And using the hypotheses on $\varepsilon^{q+1}F^p$ and $\varepsilon^{p+1}G^p$ in this Lemma and the fact that $L = C\varepsilon^\alpha$, where C is a constant independent of ε, we have

$$\begin{aligned} \left\|\varepsilon^{p+1}F, \varepsilon^{q+1}G\right\}\right\|_{\mathcal{C}^\ell,L} &\preceq \varepsilon^{-\alpha}\varepsilon^{p+1-\alpha(2p)}\varepsilon^{q+1-\alpha(2q+1)} \\ &= \varepsilon^{p+q+2-\alpha(2(p+q+1))} \\ &= \varepsilon^{\widetilde{p}+1-\alpha(2\widetilde{p})}. \end{aligned}$$

□

Remark 1.3.17. This Lemma will be applied a certain number of times and expresses the fact that given two functions $\varepsilon^{p+1}F^p$ and $\varepsilon^{q+1}G^q$, which are trigonometric polynomials in (φ, s) of degree $M_p = (p+1)M$ and $M_q = (q+1)M$, respectively, with bounds $\|\varepsilon^{p+1}F^p\|_{\mathcal{C}^\ell,L} \preceq \varepsilon^{p+1-\alpha(2p)}$ and $\|\varepsilon^{q+1}G^q\|_{\mathcal{C}^\ell,L} \preceq \varepsilon^{q+1-\alpha(2q+1)}$, its Poisson bracket is a function $\varepsilon^{\widetilde{p}+1}F^{\widetilde{p}}$, with $\widetilde{p} = p+q+1$, that is, $\varepsilon^{\widetilde{p}+1}F^{\widetilde{p}} = \{\varepsilon^{p+1}F^p, \varepsilon^{q+1}G^q\}$ is a trigonometric polynomial in (φ, s) of degree $M_{\widetilde{p}} = (\widetilde{p}+1)M$ with a bound $\left\|\varepsilon^{\widetilde{p}+1}F^{\widetilde{p}}\right\|_{\mathcal{C}^\ell,L} \preceq \varepsilon^{\widetilde{p}+1-\alpha(2\widetilde{p})}$.

Notice that this process of "$\varepsilon^{q+1}G^q$ Poisson-bracketing" can be iterated: $\varepsilon^{\widehat{p}+1}F^{\widehat{p}} = \{\varepsilon^{\widetilde{p}+1}F^{\widetilde{p}}, \varepsilon^{q+1}G^q\}$, with $\widehat{p} = \widetilde{p}+q+1$, is a trigonometric polynomial in (φ, s) of degree $M_{\widehat{p}} = (\widehat{p}+1)M$ with a bound $\left\|\varepsilon^{\widehat{p}+1}F^{\widehat{p}}\right\|_{\mathcal{C}^\ell,L} \preceq \varepsilon^{\widehat{p}+1-\alpha(2\widehat{p})}$.

We state and prove now the iterative Lemma for averaging, which is Lemma 1.3.18. It basically tells us that given a Hamiltonian already in normal form up to some order ε^q, we can produce another Hamiltonian which is normalized up to order ε^{q+1}. The averaged Hamiltonian is given rather explicitly both in the resonant regions and in the non-resonant ones, which are redefined at every step according to the new resonances that will come up.

Lemma 1.3.18. *Let $r > n > 1$ and $0 \leq 2q < n$ be any given integers. Consider a Hamiltonian of the form*

$$k_q(J, \varphi, s; \varepsilon) = Z^q(J, \varphi, s; \varepsilon) + \varepsilon^{q+1}R^q(J, \varphi, s; \varepsilon),$$

satisfying the following hypotheses:

1. *$Z^0(J, \varphi, s; \varepsilon) = \frac{J^2}{2}$ and, for $q \geq 1$, $Z^q(J, \varphi, s; \varepsilon)$ is a $\mathcal{C}^{r-2q+2}$ function that verifies:*

   There exist finite sets $\mathcal{R}_i \subset \mathbb{Q}$, $i = 1, \ldots, q$, depending on $M \sim \varepsilon^{-\rho}$, where ρ is a positive number satisfying $\rho < \frac{1}{n}$, and a number $L = C\varepsilon^\alpha > 0$, with $\rho \leq \alpha < \frac{1}{n}$ and C a constant independent of ε, such that:

1a For $-l/k \in \mathcal{R}_{[\le q]} \equiv \bigcup_{i=1\dots q} \mathcal{R}_i$, the intervals $\mathcal{I}_{-l/k} \equiv [-l/k - 2L_k, -l/k + 2L_k]$, with $L_k = L/|k|$, are disjoint.

1b If $J \notin \bigcup_{-l/k \in \mathcal{R}_{[\le q]}} \mathcal{I}_{-l/k}$, then

$$Z^q(J, \varphi, s; \varepsilon) = \frac{J^2}{2} + \varepsilon \widetilde{Z}^q(J; \varepsilon),$$

for any $(J, \varphi, s) \in \mathcal{D}^q_{\rm nr}$ ($\mathcal{D}^q_{\rm nr}$ was introduced in (1.49)), where $\varepsilon \widetilde{Z}^q(J; \varepsilon)$ is a polynomial of degree q in ε whose term of order $p+1$ is $\mathcal{O}_{\mathcal{C}^\ell, L}(\varepsilon^{p+1-\alpha(2p)})$, for $\ell = 0, \dots, r-2p$ and $p = 0, \dots, q-1$.

1c If $J \in \mathcal{I}_{-l_0/k_0}$, for some $-l_0/k_0 \in \mathcal{R}_i \setminus \mathcal{R}_1 \cup \dots \cup \mathcal{R}_{i-1}$ for some $1 \le i \le q$, then

$$Z^q(J, \varphi, s; \varepsilon) = \frac{J^2}{2} + \varepsilon \widetilde{Z}^q(J; \varepsilon) + \varepsilon^i U_q^{k_0, l_0}(k_0 \varphi + l_0 s; \varepsilon)$$

for any $(J, \varphi, s) \in \mathcal{D}_{{\rm r},i}$ ($\mathcal{D}_{{\rm r},i}$ was introduced in (1.50)), where $\varepsilon \widetilde{Z}^q(J; \varepsilon)$ is a polynomial of degree q in ε and $U_q^{k_0,l_0}(\theta; \varepsilon)$ is a polynomial of degree $q-i$ in ε and a trigonometric polynomial in $\theta = k_0\varphi + l_0 s$. The term of order $p+1$ in ε of Z^q is $\mathcal{O}_{\mathcal{C}^\ell, L}(\varepsilon^{p+1-\alpha(2p)})$, for $\ell = 0, \dots, r-2p$ and $p = 0, \dots, q-1$.

2. $\varepsilon^{q+1} R^q(J, \varphi, s; \varepsilon)$ is a $\mathcal{C}^{r-2q}$ function and is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{q+1-\alpha(\ell+2q)})$, for $\ell = 0, \dots, n-2q$. The term of order $i+1$ of the Taylor expansion with respect to ε of $\varepsilon^{q+1} R^q(J, \varphi, s; \varepsilon)$ is a trigonometric polynomial in (φ, s) of degree $M_i = (i+1)M$ and is $\mathcal{O}_{\mathcal{C}^\ell, L}(\varepsilon^{i+1-\alpha(2i)})$, for $\ell = 0, \dots, r-q-i$ and for $i = q, \dots, r-q$.

Denote $K = R^q(J, \varphi, s; 0)$, which is the term of the perturbation of order exactly $q+1$ in ε. Following definition 1.3.4, introduce the set

$$\mathcal{R}_{q+1} = \{-l/k \in \mathbb{Q} \cap (I_-, I_+), |k| + |l| \le M_q, k \ne 0, R^q_{k,l}(-l/k; 0) \ne 0\}, \tag{1.62}$$

where $M_q = (q+1)M$ and $R^q_{k,l}$ are the Fourier coefficients of R^q.

Choose a new value of C, independent of ε, in $L = C\varepsilon^\alpha$, such that the intervals $\mathcal{I}_{-l/k} \equiv [-l/k - 2L_k, -l/k + 2L_k]$, with $L_k = L/|k|$, are disjoint for $-l/k \in \mathcal{R}_{[\le q+1]}$.

Let $G(J, \varphi, s) = G_q(J, \varphi, s)$ be the $\mathcal{C}^{r-2q-1}$ trigonometric polynomial of order M_q given by Lemma 1.3.3, verifying (1.37) with $K = R^q(J, \varphi, s; 0)$.

Then, the $\mathcal{C}^{r-2q-2}$ change of variables

$$(J, \varphi, s) = g_q(\mathcal{B}, \phi, s),$$

given by the time one flow of the Hamiltonian $\varepsilon^{q+1} G_q(\mathcal{B}, \phi, s)$, transforms the Hamiltonian $k_q(J, \varphi, s; \varepsilon)$ into a Hamiltonian $k_{q+1} = k_q \circ g_q$ of the form

$$k_{q+1}(\mathcal{B}, \phi, s; \varepsilon) = Z^{q+1}(\mathcal{B}, \phi, s; \varepsilon) + \varepsilon^{q+2} R^{q+1}(\mathcal{B}, \phi, s; \varepsilon),$$

with

$$Z^{q+1}(\mathcal{B}, \phi, s; \varepsilon) = Z^q(\mathcal{B}, \phi, s; \varepsilon) + \varepsilon^{q+1} \bar{R}^q(\mathcal{B}, \phi, s; 0)$$

where $\bar{R}^q(\mathcal{B}, \phi, s; 0) = \bar{K}(\mathcal{B}, \phi, s)$ given in Lemma 1.3.3, is a $\mathcal{C}^{r-2q}$ function, such that

i. If $\mathcal{B} \notin \bigcup_{-l/k \in \mathcal{R}_{[\leq q+1]}} \mathcal{I}_{-l/k}$, *then*

$$\bar{R}^q(\mathcal{B}, \phi, s; 0) = R^q_{0,0}(\mathcal{B}; 0),$$

for any $(\mathcal{B}, \phi, s) \in \mathcal{D}^{q+1}_{\rm nr}$ *and* $\varepsilon^{q+1}\bar{R}^q$ *is* $\mathcal{O}_{\mathcal{C}^\ell, L}(\varepsilon^{q+1-\alpha(2q)})$, *for* $\ell = 0, \ldots, r-2q$.

ii. If $\mathcal{B} \in \mathcal{I}_{-l_0/k_0}$, *for some* $-l_0/k_0 \in \mathcal{R}_i \setminus \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{i-1}$ *for some* $1 \leq i \leq q+1$, *then*

$$\bar{R}^q(\mathcal{B}, \phi, s) = R^q_{0,0}(\mathcal{B}; 0) + \sum_{\substack{t \in \mathbb{Z} - \{0\}, \\ |t|(|k_0|+|l_0|) \leq M_q}} R^q_{tk_0, tl_0}(-l_0/k_0; 0) e^{{\rm i}t\theta}, \tag{1.63}$$

for any $(\mathcal{B}, \phi, s) \in \mathcal{D}_{{\rm r},i}$, *where* $R^q_{k,l}(J; \varepsilon)$ *are the Fourier coefficients of the function* $R^q(J, \varphi, s; \varepsilon)$ *with respect to* (φ, s). *Moreover,* $\varepsilon^{q+1}\bar{R}^q$ *is* $\mathcal{O}_{\mathcal{C}^\ell, L}(\varepsilon^{q+1-\alpha(2q)})$, *for* $\ell = 0, \ldots, r-2q$.

Moreover, the Hamiltonian $Z^{q+1}(\mathcal{B}, \phi, s; \varepsilon)$ *verifies properties [1b] and [1c] up to order* $q+1$, *and* $R^{q+1}(\mathcal{B}, \phi, s; \varepsilon)$ *verifies property 2 replacing* q *by* $q+1$.

Remark 1.3.19. Note that all the terms of order $p+1$, for $p \geq 0$, in the Taylor expansion in ε that appear in Lemma 1.3.18 are $\mathcal{C}^{r-2p}$ functions in J and trigonometric polynomials in the variables (φ, s) and they are bounded independently of ε in the Fourier weighted norm $\|\cdot\|_{\mathcal{C}^\ell, L}$ defined in (1.30) for ℓ up to $r-2p$. However, the whole remainder term $\varepsilon^{q+2}R^{q+1}$ is not a trigonometric polynomial in the variables (φ, s), so we can not use the Fourier weighted norm. In this case we estimate their supremum norm $|\cdot|_{\mathcal{C}^\ell}$ defined in (1.2), but only for ℓ up to $n-2q$, as in Theorem 1.3.11 (see Remark 1.3.14).

Proof. We will apply Lemma 1.3.3 with $K = R^q(J, \varphi, s; 0)$, which is a $\mathcal{C}^{r-2q}$ function, as well as a trigonometric polynomial in (φ, s) of degree $M_q = (q+1)M$. Accordingly, by definition 1.3.4 resonances of order $q+1$ correspond to the set of rational numbers $r \in \mathcal{R}_{q+1} \setminus (\mathcal{R}_1 \cup \ldots \cup \mathcal{R}_q)$.

Taking $L = C\varepsilon^\alpha$, with α satisfying $\alpha < 1/n$ and $C = C_q$ chosen adequately, the real intervals $\mathcal{I}_{-l/k} = [-l/k - 2L_k, -l/k + 2L_k]$, with $L_k = L/|k|$, for $-l/k \in \mathcal{R}_{[\leq q+1]}$ are disjoint. More precisely, the distance d_{k,k_0} between any two resonances $-l_0/k_0, -l/k \in \mathcal{R}_{[\leq q+1]}$ is greater or equal than $1/(|k_0||k|)$. In order to avoid overlapping between all these intervals, the distance d_{k,k_0} must be greater than $2L_{k_0} + 2L_k$. Taking into account that we only consider resonances with denominators $|k|, |k_0| \leq (q+1)M$, the condition that ensures that these intervals are separated is $1/((q+1)M) > 4L$, which requires $\rho \leq \alpha$ in terms of exponents of ε. This is guaranteed by the hypothesis on α and ρ in this Lemma.

Hence, we can apply Lemma 1.3.3, obtaining a $\mathcal{C}^{r-2q-1}$ function $G_q(J, \varphi, s)$ and a $\mathcal{C}^{r-2q}$ function $\bar{K} = R^q(J, \varphi, s)$, which are also trigonometric polynomials in (φ, s) of degree M_q.

Under the canonical change of variables $(J, \varphi, s) = g_q(\mathcal{B}, \phi, s)$, where g_q is the time-one

flow of $\varepsilon^{q+1}G_q$, the extended autonomous Hamiltonian $A + k_q$ becomes

$$\begin{aligned} A + k_{q+1} &= (A + k_q) \circ g_q \\ &= (A + Z^q + \varepsilon^{q+1} R^q) \circ g_q \\ &= A + Z^q + \varepsilon^{q+1}(\{A + Z^0, G_q\} + R^q(\cdot, 0)) \\ &\quad +(Z^q - Z^0) \circ g_q - (Z^q - Z^0) \\ &\quad +(A + Z^0) \circ g_q - (A + Z^0) - \{A + Z^0, \varepsilon^{q+1} G_q\} \\ &\quad +\varepsilon^{q+1}(R^q \circ g_q - R^q) + \varepsilon^{q+1}(R^q - R^q(\cdot, 0)) \\ &:= A + Z^q + \varepsilon^{q+1} \bar{R}^q + \varepsilon^{q+2} R^{q+1}, \end{aligned}$$

where

$$\bar{R}^q = \{A + Z^0, G_q\} + R^q(\cdot, 0), \tag{1.64}$$

and

$$\begin{aligned} \varepsilon^{q+2} R^{q+1} &= (Z^q - Z^0) \circ g_q - (Z^q - Z^0) \\ &\quad +(A + Z^0) \circ g_q - (A + Z^0) - \{A + Z^0, \varepsilon^{q+1} G_q\} \\ &\quad +\varepsilon^{q+1}(R^q \circ g_q - R^q) + \varepsilon^{q+1}(R^q - R^q(\cdot, 0)). \end{aligned} \tag{1.65}$$

We first see that the the normal form term $\varepsilon^{q+1}\bar{R}^q$ is bounded in the $\|\cdot\|_{\mathcal{C}^\ell, L}$ norm by $\varepsilon^{q+1-\alpha(2q)}$, for $\ell = 0, \ldots, n - 2q$.

Indeed, using (1.38) and (1.39) from Lemma 1.3.3 we have:

i. If $\mathcal{B} \notin \bigcup_{-l/k \in \mathcal{R}_{[\leq q+1]}} \mathcal{I}_{-l/k}$, then

$$\bar{R}^q(\mathcal{B}, \phi, s) = R^q_{0,0}(\mathcal{B}; 0) \tag{1.66}$$

for any $(\mathcal{B}, \phi, s) \in \mathcal{D}^{q+1}_{\mathrm{nr}}$ and, by formula (1.40) and the second part of hypothesis 2 for $i = q$ of Lemma 1.3.18, we have

$$\left\|\varepsilon^{q+1}\bar{R}^q\right\|_{\mathcal{C}^\ell, L} \leq \left\|\varepsilon^{q+1} R^q(\cdot; 0)\right\|_{\mathcal{C}^\ell, L} \preceq \varepsilon^{q+1-\alpha(2q)}, \tag{1.67}$$

for $\ell = 0, \ldots, r - 2q$.

ii. If $\mathcal{B} \in \mathcal{I}_{-l_0/k_0}$, for some $-l_0/k_0 \in \mathcal{R}_i \setminus \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{i-1}$ for some $1 \leq i \leq q + 1$, then, by equation (1.39) in Lemma 1.3.3,

$$\bar{R}^q(\mathcal{B}, \phi, s) = R^q_{0,0}(\mathcal{B}; 0) + \sum_{\substack{t \in \mathbb{Z}^2 - \{0\}, \\ |t|(|k_0| + |l_0|) \leq M_q}} R^q_{tk_0, tl_0}(-l_0/k_0; 0) e^{\mathrm{i}t\theta} \tag{1.68}$$

for any $(\mathcal{B}, \phi, s) \in \mathcal{D}_{\mathrm{r},i}$, where $R^q_{k,l}(J; 0)$ are the Fourier coefficients of the function $R^q(J, \varphi, s; 0)$ with respect to (φ, s).

As before, by formula (1.40) from Lemma 1.3.3 and the second part of hypothesis 2 of this Lemma for $i = q$, we have

$$\left\|\varepsilon^{q+1}\bar{R}^q\right\|_{\mathcal{C}^\ell, L} \preceq \left\|\varepsilon^{q+1} R^q(\cdot; 0)\right\|_{\mathcal{C}^\ell, L} \preceq \varepsilon^{q+1-\alpha(2q)} \tag{1.69}$$

for $\ell = 0, \ldots, r - 2q$.

Note that, since $\alpha < 1/n$, for the bounds obtained in (1.67) and (1.69) the power of ε is a positive number greater than q.

To finish the proof, we only need to estimate the remainder term $\varepsilon^{q+2}R^{q+1}$ in (1.65) and its Taylor expansion coefficients with respect to ε.

We will first estimate the the remainder term $\varepsilon^{q+2}R^{q+1}$ in (1.65). Since it is not a trigonometric polynomial we will estimate it in terms of the supremum norm $|\cdot|_{\mathcal{C}^\ell}$. Using the integral bound for the Taylor remainder and definitions (1.64) and (1.65) of $\bar{R}^q$ and $\varepsilon^{q+2}R^{q+1}$, respectively, we have

$$\begin{aligned}
\left|\varepsilon^{q+2}R^{q+1}\right|_{\mathcal{C}^\ell} &\leq \int_0^1 \left|\{Z^q - Z^0, \varepsilon^{q+1}G_q\} \circ g_t\right|_{\mathcal{C}^\ell} \mathrm{d}t \\
&\quad + \int_0^1 \left|(1-t)(\{\{A + Z^0, \varepsilon^{q+1}G_q\}, \varepsilon^{q+1}G_q\} \circ g_t)\right|_{\mathcal{C}^\ell} \mathrm{d}t \\
&\quad + \int_0^1 \left|\{\varepsilon^{q+1}R^q, \varepsilon^{q+1}G_q\} \circ g_t\right|_{\mathcal{C}^\ell} \mathrm{d}t \\
&\quad + \left|\varepsilon^{q+1}(R^q - R^q(\cdot;0))\right|_{\mathcal{C}^\ell} \\
&= \int_0^1 \left|\{Z^q - Z^0, \varepsilon^{q+1}G_q\} \circ g_t\right|_{\mathcal{C}^\ell} \mathrm{d}t \\
&\quad + \int_0^1 \left|(1-t)\{\varepsilon^{q+1}(\bar{R}^q - R^q(\cdot;0)), \varepsilon^{q+1}G_q\} \circ g_t\right|_{\mathcal{C}^\ell} \mathrm{d}t \\
&\quad + \int_0^1 \left|\{\varepsilon^{q+1}R^q, \varepsilon^{q+1}G_q\} \circ g_t\right|_{\mathcal{C}^\ell} \mathrm{d}t \\
&\quad + \left|\varepsilon^{q+1}(R^q - R^q(\cdot;0))\right|_{\mathcal{C}^\ell},
\end{aligned}$$

for $\ell = 0, \ldots, n - 2(q+1)$.

Using Faa-di Bruno formulae (1.174) we obtain

$$\begin{aligned}
\left|\varepsilon^{q+2}R^{q+1}\right|_{\mathcal{C}^\ell} &\preceq \left|\{Z^q - Z^0, \varepsilon^{q+1}G_q\}\right|_{\mathcal{C}^\ell} \int_0^1 |g_t|_{\mathcal{C}^\ell}^\ell \mathrm{d}t \\
&\quad + \left|\{\varepsilon^{q+1}(\bar{R}^q - R^q(\cdot;0)), \varepsilon^{q+1}G_q\}\right|_{\mathcal{C}^\ell} \int_0^1 (1-t)\,|g_t|_{\mathcal{C}^\ell}^\ell \mathrm{d}t \\
&\quad + \left|\{\varepsilon^{q+1}R^q, \varepsilon^{q+1}G_q\}\right|_{\mathcal{C}^\ell} \int_0^1 |g_t|_{\mathcal{C}^\ell}^\ell \mathrm{d}t \\
&\quad + \left|\varepsilon^{q+1}(R^q - R^q(\cdot;0))\right|_{\mathcal{C}^\ell},
\end{aligned} \tag{1.70}$$

for $\ell = 0, \ldots, n - 2(q+1)$.

By formula (1.41) from Lemma 1.3.3 and the second part of hypothesis 2 for $i = q$ of this Lemma, and using that $L \sim \varepsilon^\alpha$, we get that

$$\left\|\varepsilon^{q+1}G_q\right\|_{\mathcal{C}^\ell,L} \leq \frac{C}{L}\left\|\varepsilon^{q+1}R^q(\cdot;0)\right\|_{\mathcal{C}^{\ell+1},L} \preceq \varepsilon^{q+1-\alpha(2q+1)},$$

for $\ell = 0, \ldots, r - 2q - 1$. Hence, using the equivalence relation between $\|\cdot\|_{\mathcal{C}^\ell,L}$ and $|\cdot|_{\mathcal{C}^\ell}$ norms in (1.31), $\varepsilon^{q+1}G_q$ satisfies

$$\left|\varepsilon^{q+1}G_q\right|_{\mathcal{C}^\ell} \preceq \varepsilon^{q+1-\alpha(\ell+2q+1)}, \tag{1.71}$$

for $\ell = 0, \dots, n-2q-1$. So, we can apply Lemma 1.3.15 with $G = \varepsilon^{q+1} G_q$ and $\eta_\ell = q+1-\alpha(\ell+2q+1) > 0$ in $\mathcal{D} = (I_-, I_+) \times \mathbb{T}^2$, and we have that $\max_{0\le t\le 1} |g_t|_{\mathcal{C}^\ell}$ is bounded for $\ell = 0, \dots, n-2(q+1)$, independently of ε by some constant D_ℓ, where D_ℓ depends on ℓ and the size of the domain $\mathcal{D}$.

In the expression (1.70), the terms $Z^q - Z^0$, G_q, $\bar{R}^q$ and $R^q(\cdot;0)$ are trigonometric polynomials in the variables (φ, s). In order to bound their corresponding Poisson brackets in the $|\cdot|_{\mathcal{C}^\ell}$ norm, we will first estimate their $\|\cdot\|_{\mathcal{C}^\ell,L}$ norm and apply Lemma 1.3.16. Finally, using the equivalence relation (1.31) between $|\cdot|_{\mathcal{C}^\ell}$ and $\|\cdot\|_{\mathcal{C}^\ell,L}$ norms, we will bound their corresponding Poisson bracket in the $|\cdot|_{\mathcal{C}^\ell}$ norm. On the other hand, the terms R^q and therefore $R^q - R^q(\cdot;0)$ are not trigonometric polynomials, so we can not use the $\|\cdot\|_{\mathcal{C}^\ell,L}$ norm. For this reason we will bound the $|\cdot|_{\mathcal{C}^\ell}$ norm for the Poisson brackets directly.

The terms $\varepsilon^{q+1} R^q(\cdot;0)$ and $\varepsilon^{q+1}\bar{R}^q$ in (1.70) are both bounded in the $\|\cdot\|_{\mathcal{C}^\ell,L}$ norm by $\varepsilon^{q+1-\alpha(2q)}$, for $\ell = 0, \dots, r-2q$, because of the second part of hypothesis 2 for $i = q$ and points (i) and (ii) already proved, respectively. Note that both terms are of type $\varepsilon^{q+1}F^q$, according to remark 1.3.17.

The term $Z^q - Z^0 = \varepsilon\bar{R}^0 + \varepsilon^2\bar{R}^1 + \dots$, is a polynomial with respect to ε, so it can be bounded by its main term $\varepsilon\bar{R}^0$. Hence, using the bound for the term of order $p = 1$ of Z^q given in hypotheses 1b and 1c, we have

$$\left\| Z^q - Z^0 \right\|_{\mathcal{C}^\ell,L} \preceq \left\| \varepsilon\bar{R}^0 \right\|_{\mathcal{C}^\ell,L} \preceq \varepsilon, \tag{1.72}$$

for $\ell = 0, \dots, r-2(q-1)$. Note that $\varepsilon\bar{R}^0$ is of type εF^0, according to Remark 1.3.17.

The estimate for the $|\cdot|_{\mathcal{C}^\ell}$ norm of the term $(R^q - R^q(\cdot;0))$ can be obtained from the bound for the Taylor remainder and the first part of hypothesis 2. More precisely,

$$\left|\varepsilon^{q+1}(R^q - R^q(\cdot;0))\right|_{\mathcal{C}^\ell} \le \varepsilon^{q+2} \left|R^q\right|_{\mathcal{C}^{\ell+1}} \preceq \varepsilon^{q+2-\alpha(\ell+1-2q)}, \tag{1.73}$$

for $\ell = 0, \dots, n-2q-1$.

Moreover using the bounds for $\varepsilon^{q+1}R^q$ and $\varepsilon^{q+1}G_q$ in the $|\cdot|_{\mathcal{C}^\ell}$ norm, and Leibniz rule for derivatives we have

$$\begin{aligned}
&\left|\{\varepsilon^{q+1}R^q, \varepsilon^{q+1}G_q\}\right|_{\mathcal{C}^\ell} \\
&\preceq \sum_{i=0}^{\ell} \binom{\ell}{i} \left(\left|\varepsilon^{q+1}\frac{\partial R^q}{\partial\varphi}\right|_{\mathcal{C}^i} \left|\varepsilon^{q+1}\frac{\partial G_q}{\partial J}\right|_{\mathcal{C}^{\ell-i}} + \left|\varepsilon^{q+1}\frac{\partial R^q}{\partial J}\right|_{\mathcal{C}^i} \left|\varepsilon^{q+1}\frac{\partial G_q}{\partial\varphi}\right|_{\mathcal{C}^{\ell-i}} \right) \\
&\preceq \sum_{i=0}^{\ell} \binom{\ell}{i} \left|\varepsilon^{q+1}R^q\right|_{\mathcal{C}^{i+1}} \left|\varepsilon^{q+1}G_q\right|_{\mathcal{C}^{\ell-i+1}}.
\end{aligned}$$

Hence, if $q = 0$, we have

$$\left|\{\varepsilon R^0, \varepsilon G_0\}\right|_{\mathcal{C}^\ell} \preceq \varepsilon\varepsilon^{1-\alpha(\ell+2)} \preceq \varepsilon^{2-\alpha(\ell+2)},$$

otherwise,

$$\begin{aligned}
\left|\{\varepsilon^{q+1}R^q, \varepsilon^{q+1}G_q\}\right|_{\mathcal{C}^\ell} &\preceq \sum_{i=0}^{\ell} \binom{\ell}{i} \varepsilon^{q+1-\alpha(i+1+2q)} \varepsilon^{q+1-\alpha(\ell-i+1+2q+1)} \\
&\preceq \varepsilon^{2(q+1)-\alpha(\ell+2(2q+1)+1)},
\end{aligned}$$

for $\ell = 0, \dots, n - 2(q-1)$.

Putting together in (1.70) the estimates in (1.71), (1.72) and (1.73), as well as the estimate for $\{\varepsilon^{q+1}R^q, \varepsilon^{q+1}G_q\}$ and $\varepsilon^{q+1}\bar{R}^q$ (these last two not relevant for $q \neq 0$), and using Lemma 1.3.16 and the equivalence relation (1.31) one gets the following bound for (1.65):

$$\left|\varepsilon^{q+2}R^{q+1}\right|_{\mathcal{C}^\ell} \preceq \varepsilon^{q+2-\alpha(\ell+2(q+1))},$$

for $\ell = 0, \dots, n - 2(q+1)$.

Finally, all the terms in the Taylor expansion of $\varepsilon^{q+2}R^{q+1}(\mathcal{B}, \phi, s, \varepsilon)$ with respect to ε, are obtained from a finite number of algebraic operations and a process of "$\varepsilon^{q+1}G_q$ Poisson bracketing", as stated in Remark 1.3.17, to the Taylor coefficients in ε of Z^q and of $\varepsilon^{q+1}R^q$, which are all of them of the form $\varepsilon^{p+1}F^p$. Applying Lemma 1.3.16 iteratively, we conclude that the Taylor expansion coefficient of order $i+1$ of $\varepsilon^{q+1}R^{q+2}(\mathcal{B}, \phi, s, \varepsilon)$ with respect to ε is of the type $\varepsilon^{i+1}F^i$ according to Remark 1.3.17, that is a trigonometric polynomial of order $M_i = (i+1)M$ in the angle variables, satisfying $\mathcal{O}_{\mathcal{C}^{\ell},L}(\varepsilon^{i+1-\alpha(2i)})$ for $\ell = 0, \dots, r-q-i$ and for $i = q, \dots, r-q$. Again, by condition $\alpha < 1/n$, the power of ε is a positive number greater than i. □

Proof of Theorem 1.3.11.

The proof is by induction in q. To begin induction process, we consider $R^{[\leq M]}$, which is the truncated Fourier series of the perturbation R up to some order $M_0 = M$ as in (1.25). The order of truncation M is $M \sim \varepsilon^{-\rho}$, with ρ satisfying hypothesis (1.52). We want to apply Lemma 1.3.18 for $q = 0$ to the Hamiltonian

$$k_0(J, \varphi, s; \varepsilon) = Z^0(J, \varphi, s; \varepsilon) + \varepsilon R^0(J, \varphi, s; \varepsilon),$$

where $Z^0(J, \varphi, s; \varepsilon) = J^2/2$ and $R^0(J, \varphi, s; \varepsilon) = R^{[\leq M]}(J, \varphi, s; \varepsilon)$.

We introduce the finite set

$$\mathcal{R}_1 = \{-l/k \in \mathbb{Q} \cap (I_-, I_+), |k| + |l| \leq M, k \neq 0, R^0_{k,l}(-l/k; 0)\},$$

where $R^0_{k,l}$ are the Fourier coefficients of R^0. According to definition 1.3.4 we will refer to resonances of order 1 the elements of the set $\mathcal{R}_1$.

By hypothesis (1.56), we can apply Lemma 1.3.18 for $q = 0$, which provides a symplectic change of variables $(\mathcal{B}, \phi, s) \mapsto (J, \varphi, s) = g_0(\mathcal{B}, \phi, s)$ of class $\mathcal{C}^{r-2}$ and we get a Hamiltonian of the form

$$k_1(J, \varphi, s; \varepsilon) = Z^1(J, \varphi, s; \varepsilon) + \varepsilon^2 R^1(J, \varphi, s; \varepsilon), \tag{1.74}$$

where Z^1 is a $\mathcal{C}^r$ function and $\varepsilon^2 R^1$ is a $\mathcal{C}^{r-2}$ function, verifying properties [1b],[1c] and [2] of Lemma 1.3.18 with $q = 1$.

Moreover, formula (1.60) follows easily from expression (1.63) in Lemma 1.3.18.

Proceeding by induction, we assume that we have applied Lemma 1.3.18 up to order q, for $0 < q < m$, so in adequate symplectic coordinates, the Hamiltonian k_q of this Theorem takes the form

$$k_q(J, \varphi, s; \varepsilon) = Z^q(J, \varphi, s; \varepsilon) + \varepsilon^{q+1} R^q(J, \varphi, s; \varepsilon),$$

and satisfies hypotheses 1 and 2 of Lemma 1.3.18. Again, by hypothesis (1.56), Lemma 1.3.18 can be applied again to the Hamiltonian k_q, providing a Hamiltonian

$$k_{q+1}(J,\varphi,s;\varepsilon) = Z^{q+1}(J,\varphi,s;\varepsilon) + \varepsilon^{q+2}R^q(J,\varphi,s;\varepsilon)$$

satisfying properties 1 and 2 of Lemma 1.3.18 replacing q by $q+1$ and a new constant C in $L = C\varepsilon^\alpha$ to accommodate new resonances.

Applying the inductive Lemma m times, we get a Hamiltonian k_m

$$k_m(J,\varphi,s;\varepsilon) = Z^m(J,\varphi,s;\varepsilon) + \varepsilon^{m+1}R^m(J,\varphi,s;\varepsilon),$$

that consists of an integrable Hamiltonian Z^m, which satisfies thesis 1 and 2 of Theorem 1.2.1 for $\bar{Z}^m = Z^m$ and a perturbation $\varepsilon^{m+1}R^m$ of order $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{m+1-\alpha(\ell+2m)})$, $0 \leq \ell \leq n-2m$.

Moreover Lemma 1.3.18 gives us estimates for the terms of the integrable part Z^m of the Hamiltonian k_m in the Fourier weighted norm $\|\cdot\|_{\mathcal{C}^\ell,L}$ defined in (1.30). More precisely, we know that Z^m is a polynomial of degree m in ε, whose term of order $q+1$ is $\mathcal{O}_{\mathcal{C}^\ell,L}(\varepsilon^{q+1-\alpha(2q)})$, for $\ell = 0,\dots,r-2q$ and $q = 0,\dots,m-1$.

So, for the complete proof of Theorem 1.3.11 it will be necessary to express these estimates in the supremum norm $|\cdot|_{\mathcal{C}^\ell}$ defined in (1.2). Note that for the case of non resonant regions the integrable part $\bar{Z}^m$, given in expression (1.58), consists of a term $\mathcal{B}^2/2$ plus the term $\varepsilon\widetilde{Z}^m(\mathcal{B};\varepsilon)$, which is a polynomial of degree m in ε whose coefficients depend only on the action variable $\mathcal{B}$. In the case of resonant regions of order i, $1 \leq i \leq m$, the integrable part Z^m given in expression (1.59) consists of two different types of terms. On the one hand there is the term $\mathcal{B}^2/2$ plus $\varepsilon\widetilde{Z}^m$, which is the same as in the non resonant case, and on the other hand there is the term $\varepsilon^i U_m^{k_0,l_0}(\theta;\varepsilon)$, which is a polynomial of degree m in ε and a trigonometric polynomial in the variable $\theta = k_0\phi + l_0 s$. We will estimate these terms separately.

We will start estimating the term $\widetilde{Z}^m$ in expressions (1.58) and (1.59) in the supremum norm. By the definition of the Fourier weighted norm in (1.30) and the estimate for the term of order $q+1$ of εZ^m in the Fourier weighted norm obtained in Lemma 1.3.18, jointly with the equivalence relation between $\|\cdot\|_{\mathcal{C}^\ell,L}$ and $|\cdot|_{\mathcal{C}^\ell}$ norms (1.31), one obtains that the term of order $q+1$ of $\widetilde{Z}^m$ in expressions (1.58) and (1.59) is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{q+1-\alpha(\ell+2q)})$, for $\ell = 0,\dots,n-2q$ and $q = 0,\dots,m-1$.

The function $U_m^{k_0,l_0}(\theta;\varepsilon)$ in expression (1.59) is a polynomial of degree $m-i$ in ε and a trigonometric polynomial in $\theta = k_0\phi + l_0 s$. So, $\varepsilon^i U_m^{k_0,l_0}(\theta;\varepsilon)$ can be bounded by its main term $\varepsilon^i U_m^{k_0,l_0}(\theta;0)$, which is a trigonometric polynomial in the variables (ϕ,s) and independent of the action variable $\mathcal{B}$. Hence, by definition of the Fourier weighted norm in (1.30) and the estimate for the term of order i of Z^m in the $\|\cdot\|_{\mathcal{C}^\ell,L}$ norm for $\ell = r-2(i-1)$ obtained in Lemma 1.3.18, we have

$$\left\|\varepsilon^i U_m^{k_0,l_0}(\cdot;0)\right\|_{\mathcal{C}^{r-2(i-1)},L} = \varepsilon^i \sum_{\substack{t\in\mathbb{Z}-\{0\}\\ |t|(|k_0|+|l_0|)\leq M_q}} |U_{tk_0,tl_0}|_{\mathcal{C}^0}|t(k_0,l_0)|^{r-2(i-1)} \preceq \varepsilon^{i-\alpha(2(i-1))},$$

where $U_{k,l}$ are the Fourier coefficients of the function $U_m^{k_0,l_0}(\theta;0)$, $M_q = (q+1)M$ and $|(k,l)| = \max(|k|,|l|)$. From this expression it is clear that

$$|U_{tk_0,tl_0}|_{\mathcal{C}^0} \leq C\varepsilon^{i-\alpha(2(i-1))}/|t(k_0,l_0)|^{r-2(i-1)},$$

for some constant C independent of ε. Hence, bounding derivatives with respect to the variable θ we have

$$\begin{aligned}
|U_m^{k_0,l_0}(\cdot;\varepsilon)|_{\mathcal{C}^\ell,\theta} &\leq C \sum_{\substack{t\in\mathbb{Z}-\{0\}\\ |t|(|k_0|+|l_0|)\leq M_q}} |U_{tk_0,tl_0}|_{\mathcal{C}^0}\,|t|^\ell \\
&\leq C \sum_{\substack{t\in\mathbb{Z}-\{0\}\\ |t|(|k_0|+|l_0|)\leq M_q}} \frac{\varepsilon^{i-\alpha(2(i-1))}}{|t(k_0,l_0)|^{r-2(i-1)}}|t|^\ell \\
&\leq C\frac{\varepsilon^{i-\alpha(2(i-1))}}{|(k_0,l_0)|^{r-2(i-1)}} \sum_{\substack{t\in\mathbb{Z}-\{0\}\\ |t|(|k_0|+|l_0|)\leq M_q}} \frac{1}{|t|^{r-2(i-1)-\ell}} \\
&\leq C\frac{\varepsilon^{i-\alpha(2(i-1))}}{|(k_0,l_0)|^{r-2(i-1)}},
\end{aligned}$$

for $\ell = 0,\dots,n-2(i-1)$, as claimed in point 3 of Theorem 1.3.11.

Finally, it remains to prove that the tail $\varepsilon R^{[>M]}$ of the Fourier series of the perturbation εR that we have truncated at order $M \sim \varepsilon^{-\rho}$ at the beginning of this proof is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{m+1-\alpha(\ell+2m)})$, for $0 \leq \ell \leq n-2m$. Since the perturbation R in Hamiltonian (1.54) of Theorem 1.3.11 is a $\mathcal{O}_{\mathcal{C}^r}(1)$ function, the Fourier coefficients $R_{k,l}(J,\varepsilon)$ of $R(J,\varphi,s,\varepsilon)$ decrease at a rate of order $1/|(k,l)|^r$, for $(k,l) \longrightarrow \infty$. So, by equation (1.167) in Proposition 1.6.2 we have the following bound for $\varepsilon R^{[>M]}$,

$$|\varepsilon R^{[>M]}|_{\mathcal{C}^\ell} \preceq \frac{\varepsilon}{M^{r-\ell-2}} \preceq \varepsilon^{1+\rho(r-\ell-2)}, \tag{1.75}$$

for $\ell = 0,\dots,n-2m$.

From Lemma 1.3.18, we know that the changes of coordinates g_q, for $q = 0,\dots,m-1$ are bounded as $\mathcal{C}^\ell$ functions by some constant D_ℓ, independent of ε. Then, using this fact and formula (1.75), by Faadi-Bruno formula (1.174) we have

$$|R^{[>M]} \circ g|_{\mathcal{C}^\ell} \preceq \varepsilon^{1+\rho(r-\ell-2)},$$

where $g = g_{m-1} \circ \cdots \circ g_0$.

To get $|\varepsilon R^{[>M]} \circ g|_{\mathcal{C}^\ell} \preceq \varepsilon^{m+1-\alpha(\ell+2m)}$, we need $\varepsilon^{1+\rho(r-\ell-2)} \leq \varepsilon^{m+1-\alpha(\ell+2m)}$, that is

$$\rho \geq \frac{m-\alpha(\ell+2m)}{(r-\ell-2)}, \tag{1.76}$$

for $\ell = 0,\dots,n-2m$. In order that bounds (1.56) and (1.76) were compatible, we need to choose $r \geq \left(\frac{1}{\rho}-2\right)m+2$, which is condition (1.53) in the hypotheses of this Theorem.

Finally the choice $\bar{Z} = Z^m$ and $\bar{R} = R^m + R^{[>M]} \circ g$, with $g = g_m \circ \cdots \circ g_0$, gives the desired averaged Hamiltonian (1.57) which satisfies theses 1,2 and 3. □

Remark 1.3.20. Choosing $\rho = 1/(n+\delta)$, with $0 < \delta < 1/m$, so that condition (1.56) is fulfilled for any α between ρ and $1/n$, we have that r must satisfy

$$r \geq (n-2+\delta)m+2,$$

where m is the number of steps of averaging performed. So, as long as the regularity r of the Hamiltonian satisfies

$$r > r_{\min} := (n-2)m + 2, \tag{1.77}$$

there exist ρ, α satisfying condition (1.56) and therefore (1.52) of Theorem 1.3.11 and henceforth, m steps of averaging can be performed to provide estimates of class $\mathcal{C}^{n-2m}$, contained in the theses of Theorem 1.3.11.

Remark 1.3.21. It is important to note that the averaging procedure is valid in the full domain $(I_-, I_+) \times \mathbb{T}^2 \subset \widetilde{\Lambda}_\varepsilon$. Indeed, we have performed an averaging procedure to the Hamiltonian $k(J, \varphi, s; \varepsilon)$ in all $(I_-, I_+) \times \mathbb{T}^2$, except at the subsets $\mathcal{D}_{\mathrm{t}}(L)$, where

$$\mathcal{D}_{\mathrm{t}}(L) = \{(J, \varphi, s) \in (I_-, I_+) \times \mathbb{T}^2; L_k \leq |J + l/k| \leq 2L_k, \text{ for } -l/k \in \mathcal{R}_{[\leq m]}\}.$$

To provide an averaging procedure in the full domain $(I_-, I_+) \times \mathbb{T}^2$, we apply again Theorem 1.3.11 with $\widetilde{L}_k = \widetilde{L}/|k|$, where $\widetilde{L} = L/2$. The region $\mathcal{D}_{\mathrm{t}}(L)$ is now contained in the non resonant region corresponding to $\widetilde{L}_k$, $\mathcal{D}^m_{\mathrm{nr}}(M, \widetilde{L})$ defined in definition 1.3.7. So the averaged Hamiltonian in $\mathcal{D}_{\mathrm{t}}$ is also given by Theorem 1.3.11, with slightly different constants.

1.3.3 KAM Theorem

Up to this point, once we choose m, by Theorem 1.3.11 we can perform m steps of averaging and we obtain a Hamiltonian (1.57) which consists of an integrable Hamiltonian $\bar{Z}^m$ plus a perturbation $\varepsilon^{m+1}\bar{R}^m$ which is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{m+1-\alpha(\ell+2m)})$, for $\ell = 0, \ldots, n-2m$. Notice that $n \geq 2m$ is required as well as a large r and that the integrable Hamiltonian has different expressions in resonant regions and non-resonant regions as specified in Theorem 1.3.11.

The integrable part of the Hamiltonian gives us an approximate equation $\bar{Z}^m = \text{cte}$ for the invariant tori in $\widetilde{\Lambda}_\varepsilon$. To finish the proof of Theorem 1.3.1 it remains to determine which tori survive and what is the distance between them when we add the perturbation term $\varepsilon^{m+1}\bar{R}^m$. By choosing an adequate m large enough the goal is to show that we can cover the whole region $(I_-, I_+) \times \mathbb{T}^2 \subset \widetilde{\Lambda}_\varepsilon$ with invariant tori which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$, and obtain an approximate expression for them.

To that end, we will use KAM Theorem 1.3.22 stated in Section 1.3.3.1, which is a result about the existence of invariant tori for a periodic perturbation of a Hamiltonian expressed in action-angle variables. It is a direct adaptation of Theorem 8.12 in [DLS06a].

Since the integrable Hamiltonian (1.57) has different expressions in the resonant and non-resonant regions, we perform this study separately.

Non-resonant regions are studied in Section 1.3.3.2. In Proposition 1.3.24, we apply Theorem 1.3.22 directly to Hamiltonian (1.57) for $m \geq 2$ and we conclude that for these regions there exist primary KAM tori which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$.

Resonant regions are studied in Section 1.3.3.3. As it has been described in Section 1.2.3.3, we will see that for these regions gaps of different sizes are created in the foliation of primary KAM tori. According to the size of the gaps, we will distinguish two types of resonant regions: the resonant regions with big gaps, where gaps are of size bigger or

equal than ε, which is the size of the heteroclinic jumps provided by the scattering map, and the resonant region with small gaps, where gaps are of size smaller than ε.

In the referred Section 1.3.3.3, we will see that the resonant regions with big gaps introduced in (1.21) correspond to the resonances $J = -l/k$ of order 1 such that $|(k,l)| < M_{\mathrm{BG}} = \varepsilon^{-(1+\nu)/r}$, for $0 < \nu \leq 1/16$, whereas resonant regions with small gaps correspond to the rest of the resonances.

The case of resonant regions with small gaps is studied in 1.3.3.4. It will not be different from the non-resonant case and it will be enough to apply KAM Theorem 1.3.22 to Hamiltonian (1.57) for $m \geq 2$ to obtain primary tori $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$. This is done in Proposition 1.3.26. Resonant regions with small gaps constitute, jointly with the non resonant regions, what we call the flat tori region introduced in Section 1.2.3.3.

The case of resonant regions with big gaps is significantly different and it will be done in Section 1.3.3.5. In this case the integrable Hamiltonian $\varepsilon \bar{Z}^m$ is like a pendulum, and we will need to write it first in action-angle variables before applying KAM Theorem 1.3.22 to Hamiltonian (1.57) for $m \geq 10$. We will see that in these regions we can find other invariant objects, like secondary tori, which fill the region inside the gaps and they get rather close to the frontier of the gaps among the primary KAM tori. The precise result, jointly with the approximate equations for the invariant tori is given in Proposition 1.3.28.

Finally, Theorem 1.3.1 follows directly from Propositions 1.3.24, 1.3.26 and Theorem 1.3.28.

1.3.3.1 The KAM Theorem

The following result is about the existence of invariant tori for a periodic perturbation of a Hamiltonian system expressed in action-angle variables and it is standard in KAM theory (see [Lla01b] for a tutorial on this theory). We skip its proof since it is simply an adaptation of Theorem 8.12 in [DLS06a], which relies on a quantitative KAM Theorem of Herman [Her83, Theorem 5.4, p. 198] for exact symplectic mappings of the annulus.

Theorem 1.3.22. *Let $K(I,\varphi,s;\varepsilon)$ be Hamiltonian of the form*

$$K(I,\varphi,s;\varepsilon) = K_0(I;\varepsilon) + K_1(I,\varphi,s;\varepsilon), \tag{1.78}$$

for $(I,\varphi,s) \in \mathcal{I} \times (\mathbb{R}/2\pi k\mathbb{Z}) \times \mathbb{T}$, for some $k \in \mathbb{N}$. Assume that

i. K is a $\mathcal{C}^{n_0+\beta}$ function of the variables (I,φ,s), with $n_0 \geq 5$ and $0 < \beta < 1$,

ii. For any $s \in \mathbb{T}$, $|K_1(\cdot,s;\varepsilon)|_{\mathcal{C}^{n_0+\beta}} \leq \delta$ and $\left|K_0''(\cdot;\varepsilon)\right|_{\mathcal{C}^0} \geq M > 0$, where $\delta = \delta(\varepsilon)$ and $M = M(\varepsilon)$ depend on ε.

Then, for ε sufficiently small and fixed, there exists a finite set of values $I_i \in \mathcal{I}$, such that the Hamiltonian $K(I,\varphi,s;\varepsilon)$ has invariant tori $\mathcal{T}_i$, such that:

a. The torus $\mathcal{T}_i$ can be written as a graph of the variable I over the angle variables (φ,s):

$$\mathcal{T}_i = \{(I,\varphi,s) \in \mathcal{I} \times \mathbb{T}^2 : I = I_i + \Psi_i(\varphi,s;\varepsilon)\},$$

where $\Psi_i(\varphi,s;\varepsilon)$ is a $\mathcal{C}^{n_0-2+\beta}$ function and $|\Psi_i(\cdot;\varepsilon)|_{\mathcal{C}^2} \leq \mathrm{cte}\, M^{-1}\delta^{1/2}$.

b. The motion on the torus is $\mathcal{C}^{n_0-4+\beta}$ conjugate to a rigid translation of frequency $(\omega(I_i), 1)$, where $\omega(I_i)$ is a Diophantine number of constant type and Markov constant $\kappa = K\delta^{1/2}$, *that is*

$$|\omega(I_i)k - l|^{-1} \leq C\kappa^{-1}|(k,l)| \quad \forall (k,l) \in \mathbb{Z}^2.$$

c. The union of neighborhoods of size $CM^{-1}\delta^{1/2}$ of these tori cover all the region $\mathcal{I} \times (\mathbb{R}/2\pi k\mathbb{Z}) \times \mathbb{T}$.

Remark 1.3.23. The version of KAM Theorem requires to have the system written in action-angle variables. We would like to mention that recently there have appeared some nice results on KAM theory without action-angle variables (see [LGJV05] and [FLS07]) for analytic maps, which could be adapted but some extra work is required.

1.3.3.2 Non-resonant region

In this section we apply directly Theorem 1.3.22 to the averaged Hamiltonian (1.57) in the non-resonant region up to order m $\mathcal{D}^m_{\rm nr}$ introduced in (1.49) where we use $L/2$ instead of L, according to Remark 1.3.21, so that

$$\mathcal{D}^m_{\rm nr} = \mathcal{D}^m_{\rm nr}(M, L/2) = \{(J, \varphi, s) \in (I_-, I_+) \times \mathbb{T}^2 : |J + l/k| \geq L_k, \text{for} - l/k \in \mathcal{R}_{[\leq m]}\},$$

where $L_k = L/|k|$, with $L \sim \varepsilon^\alpha$ and $\alpha < 1/n$, as required in Theorem 1.3.11.

Going back to the original variables (I, φ, s), using the changes given by Theorem 1.3.11 and equation (1.22), we obtain the following result about the existence of invariant tori of Hamiltonian (1.4):

Proposition 1.3.24. *Assume that $m \geq 2$, $n \geq 2m + 6$ and $r > (n-2)m + 2$. Choose any $0 < \eta \leq (1 - \alpha n)/2$, where $\alpha < 1/n$ as required in Theorem 1.3.11. Then, for ε small enough, in any connected component of the non resonant region up to order m $\mathcal{D}^m_{\rm nr}$, there exists a finite set of values E_i such that:*

i. For any E_i there exists a torus $\mathcal{T}_i$ invariant by the flow of the Hamiltonian (1.4) *contained in $\widetilde{\Lambda}_\varepsilon$, which is given in $\widetilde{\Lambda}_\varepsilon$, by the level set of a $\mathcal{C}^{n-2m-2-\varrho}$ function F, for any $\varrho > 0$, of the form*

$$F(I, \varphi, s; \varepsilon) = I + \mathcal{O}_{\mathcal{C}^2}(\varepsilon^{1+\eta}), \tag{1.79}$$

corresponding to $F(I, \varphi, s; \varepsilon) \equiv E_i$.

ii. The torus $\mathcal{T}_i$ contained in $\widetilde{\Lambda}_\varepsilon$ can be written as a graph of the variable I over the angle variables (φ, s):

$$\mathcal{T}_i = \{(I, \varphi, s) \in \mathcal{D}^m_{\rm nr}, I = \lambda_{E_i}(\varphi, s; \varepsilon)\},$$

with

$$\lambda_E(\varphi, s; \varepsilon) = E + U_E(\varphi, s; \varepsilon); \tag{1.80}$$

where $U_E(\varphi, s; \varepsilon)$ is a $\mathcal{C}^{n-2m-2-\varrho}$ function, for any $\varrho > 0$, and $|U_E|_{\mathcal{C}^2} \preceq \varepsilon^{1+\eta}$.

iii. These tori are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced in terms of the variable I.

Proof: By equation (1.57) and (1.58) in Theorem 1.3.11, in one connected component of the non-resonant region $\mathcal{D}^m_{\rm nr}$, the Hamiltonian (1.23) expressed in the averaged variables $(\mathcal{B},\phi,s)$ has the following expression

$$k_m(\mathcal{B},\phi,s;\varepsilon)=\frac{\mathcal{B}^2}{2}+\varepsilon\widetilde{Z}^m(\mathcal{B},\varepsilon)+\varepsilon^{m+1}\bar{R}^m(\mathcal{B},\phi,s;\varepsilon), \tag{1.81}$$

where $\varepsilon\widetilde{Z}^m(\mathcal{B};\varepsilon)$ is a polynomial of degree m in ε, whose coefficient in terms of ε of order $q+1$, for $q=0\dots m-1$, is a $\mathcal{C}^{r-2q}$ function and is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{q+1-\alpha(\ell+2q)})$, for $\ell=0,\dots,n-2q$. Moreover, $\varepsilon^{m+1}\bar{R}^m(\mathcal{B},\phi,s;\varepsilon)$ is a $\mathcal{C}^{r-2m}$ function, which is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{m+1-\alpha(\ell+2m)})$ for $\ell=0\dots n-2m$.

Our next step is to apply KAM Theorem 1.3.22 to the Hamiltonian (1.81), which is of the form (1.78), for $K_0=\mathcal{B}^2/2+\varepsilon\widetilde{Z}^m(\mathcal{B},\varepsilon)$ and $K_1=\varepsilon^{m+1}\bar{R}^m(\mathcal{B},\phi,s;\varepsilon)$ and 2π-periodic in φ and s. Assuming that $n\geq 2m+6$, it satisfies properties (i) and (ii) with $n_0=n-2m-1$, $\beta=1-\varrho$, for any $\varrho>0$, $\delta=\varepsilon^{m+1-\alpha n}$ and M independent of ε. Therefore we can apply KAM Theorem 1.3.22 and we conclude that the non-resonant region $\mathcal{D}^m_{\rm nr}$ contains KAM tori given by

$$\mathcal{B}=\mathcal{B}_i+\Psi_i(\phi,s;\varepsilon),$$

where Ψ_i is a $\mathcal{C}^{n-2m-2-\varrho}$ function, for any $\varrho>0$, and $|\Psi_i|_{\mathcal{C}^2}\preceq\varepsilon^{(m+1-\alpha n)/2}$. These tori are $\mathcal{O}(\varepsilon^{(m+1-\alpha n)/2})$-closely spaced in terms of the averaged variable $\mathcal{B}$.

For a fixed value of ε, we have that since $\alpha<1/n$ for $n\geq 2m+6$, then $\varepsilon^{(m+1-\alpha n)/2}\leq\varepsilon^{(m+2\eta)/2}$, for any $0<\eta\leq(1-\alpha n)/2$. Hence, using that $m\geq 2$, we get the claimed estimate $\varepsilon^{1+\eta}$.

After applying KAM Theorem to Hamiltonian (1.81), we can go back to the original variables (I,φ,s). Using that the changes $(J,\varphi,s)\mapsto(\mathcal{B},\phi,s)$ and $(I,\varphi,s)\mapsto(J,\varphi,s)$ are ε-close to the identity by Theorem 1.3.11 and equation (1.22), respectively, tori obtained in the region $\mathcal{D}^m_{\rm nr}$ are given by

$$I=I_i+U_i(\varphi,s;\varepsilon)$$

where the function U_i verifies the same properties as Ψ_i, and they are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced in terms of the variable I. We get the results claimed for $E_i=I_i$. □

1.3.3.3 Resonant region

In this section, we analyze Hamiltonian (1.23) in the resonant region up to order m $\mathcal{D}^m_{\rm r}$ defined in (1.51).

We will perform an accurate study in this resonant region $\mathcal{D}^m_{\rm r}$ and we will estimate the size of the gaps created in the foliation of primary KAM tori. We will see that this size depends on the order j of the resonance, for $0\leq j\leq m$, and on the size of the harmonic associated to the corresponding resonance. According to this, we will define two types of regions: the small gaps regions $\mathcal{D}_{\rm SG}$ where the size of the gap is smaller than ε and the big gaps regions $\mathcal{D}_{\rm BG}$ where the size of the gap is bigger or equal than ε.

We will work in one connected component of the resonant domain $\mathcal{D}^m_{\rm r}$ which, according to (1.51), is of the form

$$\{(J,\varphi,s)\in[-l_0/k_0-L_{k_0},-l_0/k_0+L_{k_0}]\times\mathbb{T}^2\}, \tag{1.82}$$

for some $-l_0/k_0 \in \mathcal{R}_j \setminus (\mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{j-1})$, for $1 \leq j \leq m$, where $L_{k_0} = L/|k_0|$, with $L \sim \varepsilon^\alpha$ and $\alpha < 1/n$, as required in Theorem 1.3.11.

By formulas (1.57) and (1.59) of Theorem 1.3.11, in component (1.82), Hamiltonian (1.23) expressed in the averaged variables $(\mathcal{B}, \phi, s)$, which are ε-close to the original variables (J, φ, s), can be written as

$$\begin{aligned} k_m(\mathcal{B}, \phi, s; \varepsilon) &= \frac{1}{2}\mathcal{B}^2 + \varepsilon \widetilde{Z}^m(\mathcal{B}; \varepsilon) + \varepsilon^j U_m^{k_0,l_0}(k_0\phi + l_0 s; \varepsilon) + \varepsilon^{m+1}\bar{R}^m(\mathcal{B}, \phi, s; \varepsilon), \\ &:= \bar{Z}^m(\mathcal{B}, \phi, s; \varepsilon) + \varepsilon^{m+1}\bar{R}^m(\mathcal{B}, \phi, s; \varepsilon), \end{aligned} \tag{1.83}$$

where $\widetilde{Z}^m(\mathcal{B}; \varepsilon)$ and $U_m^{k_0,l_0}(k_0\phi + l_0 s; \varepsilon)$ are polynomials of degree $m-1$ and $m-j$ in ε, respectively, and $U_m^{k_0,l_0}(k_0\phi + l_0 s; \varepsilon)$ is a trigonometric polynomial in $\theta = k_0\phi + l_0 s$.

For $q = 0, \ldots, m-1$, the coefficient of order $q+1$ in ε of $\varepsilon\widetilde{Z}^m$ is a $\mathcal{C}^{r-2q}$ function which is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{q+1-\alpha(\ell+2q)})$ for $\ell = 0, \ldots, n-2q$. The function $\varepsilon^j U_m^{k_0,l_0}(\theta; \varepsilon)$, for $\theta = k_0\phi + l_0 s$, satisfies

$$\left|\varepsilon^j U_m^{k_0,l_0}(\cdot; \varepsilon)\right|_{\mathcal{C}^\ell} \preceq \varepsilon^{j-2\alpha(j-1)}|(k_0, l_0)|^{-r+2(j-1)}, \tag{1.84}$$

for $\ell = 0, \ldots, n-2(j-1)$ and $|(k_0, l_0)| = \max(|k_0|, |l_0|)$.

Moreover, $\varepsilon^{m+1}\bar{R}^m$ is a $\mathcal{C}^{r-2m}$ function which is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{m+1-\alpha(\ell+2m)})$, for $\ell = 0 \ldots n-2m$.

From expression (1.83) it is clear that the integrable part $\bar{Z}^m$ is like a pendulum. The integrable pendulum has rotational and librational orbits as well as separatrices, which separate these two types of motion. It is straightforward to see that the size of the gap, created by the separatrix loop, associated to the resonance $-l_0/k_0 \in \mathcal{R}_j \setminus \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{j-1}$, in terms of the J variables, can be bounded from above by $\sqrt{2}\varepsilon^{j/2}\left|U_m^{k_0,l_0}(\cdot; \varepsilon)\right|_{\mathcal{C}^0}^{1/2}$.

From expression (1.84) with $\ell = 0$, we have that the size of the gap for a resonance $-l_0/k_0$ of order j is

$$\mathcal{O}(\varepsilon^{(j-2\alpha(j-1))/2}|(k_0, l_0)|^{(-r+2(j-1))/2}). \tag{1.85}$$

In the particular case of a resonance $-l_0/k_0$ of order 1 ($j = 1$), the size of the gap is

$$\mathcal{O}(\varepsilon^{1/2}|(k_0, l_0)|^{-r/2}). \tag{1.86}$$

For any $\nu > 0$, the resonances of order 1 such that $|(k_0, l_0)| \geq M_{\mathrm{BG}}$, where

$$M_{\mathrm{BG}} = \varepsilon^{-(1+\nu)/r}, \tag{1.87}$$

create gaps of size $\mathcal{O}(\varepsilon^{1+\nu/2})$, that is smaller than ε. Notice that M_{BG} is smaller than the order of truncation $M \sim \varepsilon^{-\rho}$, where ρ is chosen to be $\rho = 1/(n+\delta)$, with $0 < \delta < 1/m$ (see Remark 1.3.20) and $0 < \nu \leq 1/16$.

On the other hand, we know that resonances $-l_0/k_0$ of order greater than 1 satisfy $M_{\mathrm{BG}} \leq |(k_0, l_0)| \leq mM$ (see Remark 1.3.5). Hence, according to (1.85) the size of the gap created by a resonance $-l_0/k_0$ or order j, for $j = 2, \ldots, m$ is

$$\mathcal{O}(\varepsilon^{(j+1+\nu-(\alpha+(1+\nu)/r)2(j-1))/2})).$$

Using the condition $\alpha < 1/n$, with $r > n \geq 2m$ and $m \geq 2$, the size of the gap is

$$\mathcal{O}(\varepsilon^{(j+1-4\alpha(j-1))/2}).$$

For $j \geq 2$ the size of the gaps is smaller than $\varepsilon^{1+\eta}$, for $\eta = (1-4\alpha)/2$. Notice that $0 < \eta < 1$ thanks to the condition on α.

As we already said, we will distinguish between two types of resonant regions depending whether the size of the gaps created in the foliation of primary KAM tori are bigger or smaller than the size ε of the heteroclinic jumps provided by the scattering map (1.17).

- **Resonant regions with big gaps $\mathcal{D}_{\mathrm{BG}}$.** Gaps of size of order equal or greater than ε are created in the foliation of primary invariant tori. According to (1.86) they correspond to resonances $-l_0/k_0$ of order 1 with $\gcd(k_0, l_0) = 1$, satisfying $|(k_0, l_0)| < M_{\mathrm{BG}}$, where $M_{\mathrm{BG}} = \varepsilon^{-(1+\nu)/r}$, for $0 < \nu \leq 1/16$. See definition (1.21).

- **Resonant regions with small gaps $\mathcal{D}_{\mathrm{SG}}$.** Gaps between primary tori are smaller than ε. They correspond to the resonant regions of resonances $-l_0/k_0$ of order 1 such that $|(k_0, l_0)| \geq M_{\mathrm{BG}}$ and resonances of order greater or equal than 2.

Remark 1.3.25. We would like to emphasize that our result about resonances is remarkably different of the one obtained in [DLS06a], where it was considered the case of a perturbation h with a finite number of harmonics. In that case there was a uniform size for the gaps created by the resonances. For instance, the gaps created by the resonances of order 1 and 2 were $C\varepsilon^{1/2}$ and $C\varepsilon$, respectively. In our case we have a heterogeneous sea of gaps of different sizes. Moreover, in our case the resonances that create big gaps are just some of the resonances of order 1, whereas in [DLS06a], both resonances of order 1 and 2 created big gaps.

1.3.3.4 Resonant regions with small gaps

In this section, we will study the resonant regions with small gaps $\mathcal{D}_{\mathrm{SG}}$, which correspond to either resonances $-l_0/k_0$ of order 1 such that $|(k_0, l_0)| > M_{\mathrm{BG}}$, where M_{BG} is defined in (1.87), or resonances of order greater than 1.

We will work in one connected component, and we will see that we can apply almost directly Theorem 1.3.22 to Hamiltonian (1.83) in order to prove that they are covered by primary tori which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, for some $\eta > 0$. In order to get better estimates, we will first perform a simply change of coordinates, which is not symplectic but conformally symplectic.

Going back to the original variables (I, φ, s) using the changes given by Theorem 1.3.11 and equation (1.22), we obtain the following result about the existence of invariant primary KAM tori of Hamiltonian (1.4):

Proposition 1.3.26. *Assume that $m \geq 2$, $n \geq 2m+6$ and $r > (n-2)m+2$. Choose any $0 < \eta \leq \nu/2$, for $0 < \nu \leq 1/16$. Then, for ε small enough, in any connected component of $\mathcal{D}_{\mathrm{SG}}$, which is of the form (1.82), for some $-l_0/k_0 \in \mathcal{R}^{[\leq m]}$ and $|(k_0, l_0)| \geq M_{\mathrm{BG}}$, $L_{k_0} = L/|k_0|$ with $L \sim \varepsilon^{\alpha}$ and $\alpha < 1/n$, as required in Theorem 1.3.11, there exists a finite set of values E_i such that:*

i. For any E_i there exists a torus $\mathcal{T}_i$ invariant by the flow of the Hamiltonian (1.4) contained in $\widetilde{\Lambda}_\varepsilon$, which is given in $\widetilde{\Lambda}_\varepsilon$ by the equation $F(I, \varphi, s; \varepsilon) \equiv E_i$, where F is

a $\mathcal{C}^{n-2m-2-\varrho}$ function, for any $\varrho > 0$, of the form

$$F(I, \varphi, s; \varepsilon) = I + \mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{1+\eta-\ell(1+\nu)/r}), \tag{1.88}$$

for $\ell = 0, 1, 2$.

ii. The torus $\mathcal{T}_i$ can be written as a graph of the variable I over the angle variables (φ, s):

$$\mathcal{T}_i = \{(I, \varphi, s) \in \mathcal{D}_{\mathrm{SG}}; I = \lambda_{E_i}(\varphi, s; \varepsilon)\},$$

with

$$\lambda_E(\varphi, s; \varepsilon) = E + U_E(\varphi, s; \varepsilon) \tag{1.89}$$

where U_E is a $\mathcal{C}^{n-2m-2-\varrho}$ function, for any $\varrho > 0$, and $U_E = \mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{1+\eta-\ell(1+\nu)/r})$, for $\ell = 0, 1, 2$.

iii. These tori are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced in terms of the variable I.

Proof: By Theorem 1.3.11, in any connected component of $\mathcal{D}_{\mathrm{SG}}$, Hamiltonian (1.23) expressed in the averaged variables $(\mathcal{B}, \phi, s)$ has the expression (1.83).

We will first perform the change of variables depending on the time s and (k_0, l_0) given by

$$b = k_0(\mathcal{B} + l_0/k_0), \quad \theta = k_0\phi + l_0 s, \quad s = s. \tag{1.90}$$

The change (1.90) was already performed in [DLS06a], but only for resonances with big gaps. In this thesis we perform it in the case of resonances with small gaps because in our case (k_0, l_0) depend on ε and this will allow us to get better estimates for the functions defining the searched tori. Note that this change of coordinates is not symplectic but conformally symplectic, hence the new system of equations verified by (b, θ, s) is also Hamiltonian of Hamiltonian:

$$\bar{K}(b, \theta, s; \varepsilon) = \bar{K}^0(b; \varepsilon) + \varepsilon^j \bar{V}(\theta; \varepsilon) + \varepsilon^{m+1} \bar{K}^1(b, \theta, s; \varepsilon), \tag{1.91}$$

with

$$\begin{aligned}
\bar{K}^0(b, \varepsilon) &= b^2/2 + \varepsilon k_0^2 \widetilde{Z}^m(-l_0/k_0 + b/k_0; \varepsilon), \\
\bar{V}(\theta; \varepsilon) &= k_0^2 U_m^{k_0, l_0}(\theta; \varepsilon), \\
\bar{K}^1(b, \theta, s; \varepsilon) &= k_0^2 \bar{R}^m(-l_0/k_0 + b/k_0, \frac{\theta - l_0 s}{k_0}, s; \varepsilon).
\end{aligned}$$

Note that $\bar{K}^0$ is of class $\mathcal{C}^{r-2m+2}$ with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n - 2m + 2$, $\bar{V}$ is analytic because it is a polynomial in ε of degree $m - j$ and a trigonometric polynomial in θ, and $\bar{K}^1$ is of class $\mathcal{C}^{r-2m}$ with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n - 2m$.

Hamiltonian (1.91) is of the form (1.78), with $K_0 = \bar{K}^0(b, \varepsilon)$, which is a $\mathcal{C}^{r-2m+2}$ function and

$$K_1 = \varepsilon^j \left(\bar{V}(\theta; \varepsilon) + \varepsilon^{m+1-j} \bar{K}^1(\mathcal{B}, \phi, s; \varepsilon)\right),$$

which is a $\mathcal{C}^{r-2m}$ function and 2π-periodic in s.

Recall from Theorem 1.3.11 that $U_m^{k_0,l_0}(k_0\phi+l_0 s;\varepsilon)$ is a polynomial in ε of degree $m-j$ and a trigonometric polynomial in $\theta = k_0\phi + l_0 s$, which has the following bound

$$\left|\varepsilon^j U_m^{k_0,l_0}(\cdot;\varepsilon)\right|_{\mathcal{C}^\ell} \preceq \varepsilon^{j-2\alpha(j-1)}|(k_0,l_0)|^{-r+2(j-1)}, \tag{1.92}$$

for $\ell = 0,\dots,n-2m$. Moreover, $\varepsilon^{m+1}\bar{R}^m$ is a $\mathcal{C}^{r-2m}$ function with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n-2m$ given by

$$\left|\varepsilon^{m+1}\bar{R}^m(\cdot;\varepsilon)\right|_{\mathcal{C}^\ell} \preceq \varepsilon^{m+1-\alpha(\ell+2m)}. \tag{1.93}$$

Hence, from the estimates for the $\mathcal{C}^{n-2m}$ norm of functions $\varepsilon^j U_m^{k_0,l_0}$ in (1.92) and $\varepsilon^{m+1}\bar{R}^m$ in (1.93) with $\ell = n-2m$, one gets the following estimates for the derivatives of K_1 with respect to (b,θ),

$$|K_1(\cdot,\cdot,s;\varepsilon)|_{\mathcal{C}^{n-2m}} \preceq \varepsilon^{j-2\alpha(j-1)}|(k_0,l_0)|^{-r+2j} + \varepsilon^{m+1-\alpha n}|k_0|^{2-(n-2m)},$$

with $\alpha < 1/n$ and $n \geq 2m+6$. Assuming that $m \geq 2$, for any $1 \leq j \leq m$, the worse estimate comes from $j = 1$

$$|K_1(\cdot,\cdot,s;\varepsilon)|_{\mathcal{C}^{n-2m}} \preceq \varepsilon|(k_0,l_0)|^{-r+2}.$$

Hence, choosing any $m \geq 2$ and assuming that $n \geq 2m+6$, Hamiltonian (1.83) satisfies properties (i) and (ii) in Theorem 1.3.22 with $n_0 = n-2m-1$ and $\beta = 1-\varrho$, for any $\varrho > 0$, $\delta = \varepsilon|(k_0,l_0)|^{-r+2}$ and $|K_0''(\cdot;\varepsilon)| \geq M > 0$, for M independent of ε. Therefore, we can apply Theorem 1.3.22 and we conclude that regions $\mathcal{D}_{\rm BG}$ contain KAM tori given by

$$b = b_i + \beta_i(\theta, s;\varepsilon)$$

where β_i is a $\mathcal{C}^{n-2m-2-\varrho}$ function, for any $\varrho > 0$, and

$$|\beta_i|_{\mathcal{C}^2} \preceq \varepsilon^{1/2}|(k_0,l_0)|^{(-r+2)/2}.$$

These tori are $\mathcal{O}(\varepsilon^{1/2}|(k_0,l_0)|^{(-r+2)/2})$-closely spaced in terms of the variable b.

Going back to the variables $(\mathcal{B},\phi,s)$ using the transformation in (1.90), we get that the tori are given by

$$\mathcal{B} = \mathcal{B}_i + \Psi_i(\phi, s;\varepsilon)$$

where $\Psi_i(\phi,s;\varepsilon) = k_0^{-1}\beta_i(k_0\phi + l_0 s, s;\varepsilon)$ is a $\mathcal{C}^{n-2m-2-\varrho}$ function, for any $\varrho > 0$, and

$$|\Psi_i|_{\mathcal{C}^\ell} \preceq \varepsilon^{1/2}|(k_0,l_0)|^{(-r+2\ell)/2},$$

for $\ell = 0,1,2$. These tori are $\mathcal{O}(\varepsilon^{1/2}|(k_0,l_0)|^{-r/2})$-closely spaced in terms of the variable $\mathcal{B}$.

Taking into account that the resonances with small gaps, which are the ones that we are considering, are the resonances of order greater or equal than 1 satisfying $|(k_0,l_0)| \geq M_{\rm BG} = \varepsilon^{-(1+\nu)/r}$, for $0 < \nu \leq 1/16$, we have that

$$|\Psi_i|_{\mathcal{C}^\ell} \preceq \varepsilon^{1+\nu/2-\ell(1+\nu)/r},$$

for $\ell = 0, 1, 2$. Moreover, tori are $\mathcal{O}(\varepsilon^{1+\nu/2})$-closely spaced in terms of the variable $\mathcal{B}$.

As in the non-resonant regions we can go back to the original variables (I, φ, s). Using that the changes $(J, \varphi, s) \mapsto (\mathcal{B}, \phi, s)$ and $(I, \varphi, s) \mapsto (J, \varphi, s)$ are ε-close to the identity by Theorem 1.3.11 and equation (1.22), the tori obtained in the region $\mathcal{D}_{\mathrm{SG}}$ are given by

$$I = I_i + U_i(\varphi, s; \varepsilon),$$

where the function U_i verifies the same properties as Ψ_i, and they are $\mathcal{O}(\varepsilon^{1+\nu/2})$-closely spaced in terms of the averaged variable I. We get the claimed results taking $E_i = I_i$. □

Remark 1.3.27. Notice that invariant tori in the small gaps region $\mathcal{D}_{\mathrm{SG}}$ are given by a certain function F in (1.88) that, as in the case of non-resonant regions (see (1.79)), is of the form

$$F(I, \varphi, s; \varepsilon) = I + \mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta}), \tag{1.94}$$

for some $\eta > 0$. The estimates for the first and second order derivatives are different from the ones obtained in the non-resonant case. However, it will be enough for the future application of the scattering map to have them bounded by ε^{δ}, for some $\delta > 0$.

1.3.3.5 Resonant regions with big gaps

In this section, we will see that the resonant regions with big gaps $\mathcal{D}_{\mathrm{BG}}$ which correspond to resonances of order 1 such that $|(k_0, l_0)| < M_{\mathrm{BG}}$ are covered with invariant objects (either primary tori or secondary tori) which are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced in terms of the action variable, for some $\eta > 0$.

To that end, we will apply Theorem 1.3.22 to Hamiltonian (1.81) as we did in the previous cases. The main difference is that in this case the integrable Hamiltonian is not written down into action-angle variables, so we will need to perform a change of coordinates before applying KAM Theorem. Furthermore, we will perform two useful changes of coordinates, which are not symplectic but conformally symplectic.

Finally, going back to the original variables (I, φ, s) using the changes given by Theorem 1.3.11 and equation (1.22), we obtain the following result about the existence of invariant tori of Hamiltonian (1.4):

Theorem 1.3.28. *Let m, n, r integers satisfying $m \geq 10$, $n \geq 2m+6$ and $r > (n-2)m+2$. Let us consider the Hamiltonian* (1.23) *in any connected component of $\mathcal{D}_{\mathrm{BG}}$, which is of the form (1.82), for some $-l_0/k_0$ of order 1 such that $|(k_0, l_0)| < M_{\mathrm{BG}}$, with M_{BG} defined in (1.87), $L_{k_0} = L/|k_0|$ with $L \sim \varepsilon^{\alpha}$ and $\alpha < 1/n$, as required in Theorem 1.3.11.*

Assume that the function $U_m^{k_0,l_0}(k_0\phi + l_0 s; 0)$ in Hamiltonian (1.83) has a global maximum which is non degenerate (this assumption corresponds to the hypothesis **H3'** *on (k_0, l_0) in Theorem 1.2.1). Choose any $0 < \eta \leq \nu/2$, for $0 < \nu \leq 1/16$. Then, for ε small enough, there exists a finite set of values E_i in some range of energies $-\varepsilon|(k_0, l_0)|^{-r+2} \leq E \leq L^2$ such that:*

i. For any E_i there exist invariant objects by the flow of the Hamiltonian (1.4) *contained in $\widetilde{\Lambda}_\varepsilon$, which are given in $\widetilde{\Lambda}_\varepsilon$ by the equation $F(I, \varphi, s; \varepsilon) \equiv E_i$, where F is a*

$\mathcal{C}^{4-\varrho}$ function, for any $\varrho > 0$, of the form

$$\begin{aligned} F(I,\varphi,s;\varepsilon) &= \frac{(k_0 I + l_0 + \mathcal{O}_{\mathcal{C}^2}(|k_0|\varepsilon))^2}{2}(1+\mathcal{O}_{\mathcal{C}^2}(\varepsilon)) + \varepsilon k_0^2 U_m^{k_0,l_0}(k_0\varphi + l_0 s;\varepsilon) \\ &\quad + \mathcal{O}_{\mathcal{C}^2}(\varepsilon^{3/2+\eta}|(k_0,l_0)|^{-r+2}). \end{aligned} \tag{1.95}$$

For values of $E_i > 0$ the invariant objects given by equation $F \equiv E_i$ are two primary KAM tori $\mathcal{T}_{E_i}^{\pm}$, whereas for $E_i < 0$ is a secondary tori $\mathcal{T}_{E_i}$. In this case we denote $\mathcal{T}_{E_i}^{\pm}$ each of the components of

$$\mathcal{T}_{E_i} \cap \{(I,\varphi,s) \in \mathcal{D}_{\mathrm{BG}}; \rho \le k_0\varphi + l_0 s \le 2\pi - \rho\},$$

for some $0 < \rho < 2\pi$.

ii. There exists $\rho \ge 0$, such that the two primary KAM tori (components of the secondary tori) $\mathcal{T}_{E_i}^{\pm}$ contained in $\widetilde{\Lambda}_\varepsilon$ can be written as a graph of the variable I over the angle variables (φ, s):

$$\mathcal{T}_{E_i}^{\pm} = \{(I,\varphi,s) \in [-l_0/k_0 - L_{k_0}, -l_0/k_0 + L_{k_0}] \times [\rho, 2\pi - \rho] \times \mathbb{T};\ I = \lambda_{E_i}^{\pm}(\varphi,s;\varepsilon)\},$$

where

$$\lambda_E^{\pm}(\varphi,s;\varepsilon) = -l_0/k_0 + \frac{1}{k_0}\mathcal{Y}_{\pm}(\theta,E) + \mathcal{O}_{\mathcal{C}^2}(\varepsilon), \tag{1.96}$$

for $\rho \le \theta = k_0\varphi + l_0 s \le 2\pi - \rho$, where

$$\mathcal{Y}_{\pm}(x,E) = \pm(1+\varepsilon b)\ell(\theta,E) + \varepsilon\widetilde{\mathcal{Y}}_{\pm}(\ell(\theta,E)), \tag{1.97}$$

$\ell(\theta,E) = \sqrt{2(E - \varepsilon k_0^2 U_m^{k_0,l_0}(\theta;0))}$ and $\widetilde{\mathcal{Y}}_{\pm}$ satisfies (1.119).

iii. These invariant tori are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced in terms of the variable I and $\mathcal{O}(\varepsilon^{3/2+\eta}|(k_0,l_0)|^{-r/2+1})$ in terms of energies.

Remark 1.3.29. In Remark 1.3.12 we pointed out that the function $U_m^{k_0,l_0}(k_0\varphi + l_0 s;0)$ given explicitly in (1.60), corresponds to the function $U^{k_0,l_0}(\theta)$ for $\theta = k_0\varphi + l_0 s$ in hypothesis **H3** on (k_0,l_0) in Theorem 1.2.1 when $M = M_{\mathrm{BG}}$. Hence, it is clear that hypothesis **H3'** on (k_0,l_0) of Theorem 1.2.1, which assumes that the function $U^{k_0,l_0}(\theta)$ has a global maximum which is non-degenerate, implies the same assumption for $U_m^{k_0,l_0}(k_0\varphi + l_0 s;0)$.

1.3.3.6 Proof of Theorem 1.3.28

Beginning of the Proof.

By Theorem 1.3.11, in any connected component of the resonant domain $\mathcal{D}_{\mathrm{BG}}$, which is of the form (1.82), Hamiltonian (1.23) expressed in the averaged variables $(\mathcal{B},\phi,s)$ can be written as

$$\begin{aligned} k_m(\mathcal{B},\phi,s;\varepsilon) &= \frac{1}{2}\mathcal{B}^2 + \varepsilon\widetilde{Z}^m(\mathcal{B};\varepsilon) + \varepsilon U_m^{k_0,l_0}(k_0\phi + l_0 s;\varepsilon) + \varepsilon^{m+1}\bar{R}^m(\mathcal{B},\phi,s;\varepsilon) \\ &:= \bar{Z}^m(\mathcal{B},\phi,s;\varepsilon) + \varepsilon^{m+1}\bar{R}^m(\mathcal{B},\phi,s;\varepsilon), \end{aligned} \tag{1.98}$$

defined on the domain

$$\{(\mathcal{B},\phi,s)\in\mathbb{R}\times\mathbb{T}^2;\ |\mathcal{B}+l_0/k_0|\leq \bar{L}_{k_0}\},\tag{1.99}$$

where $|L_{k_0}-\bar{L}_{k_0}|\leq \text{cte}\,\varepsilon$.

In this domain, $\varepsilon\widetilde{Z}^m(\mathcal{B};\varepsilon)$ is a $\mathcal{C}^{r-2m+2}$ function in the variable $\mathcal{B}$ and it is a polynomial of degree $m-1$ in ε, whose coefficient of order $q+1$, for $q\geq 1$ is a $\mathcal{C}^{r-2q}$ function and $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^{q+1-\alpha(\ell+2q)})$, for $\ell=0,\dots,n-2q$. The coefficient of order 1 is a $\mathcal{C}^r$ function with a bounded $\mathcal{C}^\ell$ norm up to $\ell=r$, which is $\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon)$.

Moreover $U_m^{k_0,l_0}(k_0\phi+l_0 s;\varepsilon)$ is a polynomial of degree $m-1$ in ε and a trigonometric polynomial in $\theta=k_0\phi+l_0 s$, satisfying $\varepsilon U_m^{k_0,l_0}(\theta;\varepsilon)=\mathcal{O}_{\mathcal{C}^\ell}(\varepsilon|(k_0,l_0)|^{-r})$. Its main term $U_m^{k_0,l_0}(\theta;0)$ is given in expression (1.60) of Theorem 1.3.11.

Finally, $\varepsilon^{m+1}\bar{R}^m$ is a $\mathcal{C}^{r-2m}$ function in the variables $(\mathcal{B},\phi,s)$ with a bounded $\mathcal{C}^\ell$ norm up to $\ell=n-2m$, which is

$$\left|\varepsilon^{m+1}\bar{R}^m(\cdot;\varepsilon)\right|_{\mathcal{C}^\ell}\preceq\varepsilon^{m+1-\alpha(\ell+2m)}.\tag{1.100}$$

By the hypothesis in Theorem 1.3.28, the function $U_m^{k_0,l_0}(\theta,0)$ (the first order term in ε of the function $U_m^{k_0,l_0}(\theta,\varepsilon)$) has a global maximum which is non-degenerate and this implies that the integrable part $\bar{Z}^m$ of the Hamiltonian (1.98) is like an integrable pendulum.

As it has been done in Section 8.5.2 in [DLS06a], we perform two useful changes of coordinates which are not symplectic but conformally symplectic. The first one is the same as we performed in (1.90) for the resonances with small gaps. It depends on the time s and the resonance (k_0,l_0) and we recall it now:

$$b=k_0(\mathcal{B}+l_0/k_0),\quad \theta=k_0\phi+l_0 s,\quad s=s,\tag{1.101}$$

hence the system of equations verified by (b,θ,s) is also Hamiltonian of Hamiltonian:

$$\bar{K}(b,\theta,s;\varepsilon)=\bar{K}^0(b;\varepsilon)+\varepsilon\bar{V}(\theta;\varepsilon)+\varepsilon^{m+1}\bar{K}^1(b,\theta,s;\varepsilon),\tag{1.102}$$

with

$$\begin{aligned}\bar{K}^0(b,\varepsilon)&=b^2/2+\varepsilon k_0^2\widetilde{Z}^m(-l_0/k_0+b/k_0;\varepsilon),\\ \bar{V}(\theta;\varepsilon)&=k_0^2U_m^{k_0,l_0}(\theta;\varepsilon),\\ \bar{K}^1(b,\theta,s;\varepsilon)&=k_0^2\bar{R}^m(-l_0/k_0+b/k_0,\frac{\theta-l_0 s}{k_0},s;\varepsilon).\end{aligned}\tag{1.103}$$

Note that $\bar{K}^0$ is of class $\mathcal{C}^{r-2m+2}$ with a bounded $\mathcal{C}^\ell$ norm up to $\ell=n-2m+2$ and $\bar{V}$ is analytic because it is a trigonometric polynomial in θ and a polynomial of degree $m-1$ in ε. $\bar{K}^1$ is a function of class $\mathcal{C}^{r-2m}$ with a bounded $\mathcal{C}^\ell$ norm up to $\ell=n-2m$, which is $2\pi k_0$-periodic in θ and 2π-periodic in s.

The integrable part $\bar{K}^0(b;\varepsilon)+\varepsilon\bar{V}(\theta;\varepsilon)$ of the Hamiltonian (1.102) is a one degree of freedom Hamiltonian close to a pendulum-like Hamiltonian

$$\frac{b^2}{2}+\varepsilon\bar{V}(\theta;0).$$

By hypothesis **H3'** on (k_0, l_0) this pendulum-like Hamiltonian has a hyperbolic saddle at $(0, \theta_1)$ and by the implicit function Theorem the whole integrable Hamiltonian $\bar{K}^0(b;\varepsilon) + \varepsilon\bar{V}(\theta;\varepsilon)$ has also a saddle at $(b(\varepsilon), \theta_1(\varepsilon))$. The function $b(\varepsilon)$ is of class $\mathcal{C}^{r-2m+1}$ in ε and of the form $b(\varepsilon) = \mathcal{O}(|k_0|\varepsilon)$ whereas $\theta_1(\varepsilon)$ is analytic in ε and of the form $\theta_1(\varepsilon) = \theta_1 + \mathcal{O}(\varepsilon)$.

To make the analysis of this system easier we perform a second change of variables, which depends on ε and consists of the following translation

$$y = b - b(\varepsilon), \quad x = \theta - \theta_1(\varepsilon), \quad s = s, \tag{1.104}$$

in such a way that the integrable part of the Hamiltonian expressed in these new variables has a saddle point at $(0,0)$ and the energy of the saddle and the separatrices is 0. More precisely, we obtain the $\mathcal{C}^{r-2m}$ Hamiltonian with respect to (y, x, s) with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n - 2m$

$$K(y, x, s; \varepsilon) = h^0(y;\varepsilon) + \varepsilon U^{k_0,l_0}(x;\varepsilon) + \varepsilon^{m+1} S(y, x, s; \varepsilon) \tag{1.105}$$

which consists of an integrable part corresponding to the terms up to order ε^m, which is the following $\mathcal{C}^{r-2m+2}$ function with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n - 2m + 2$,

$$K_0(y, x; \varepsilon) = h^0(y;\varepsilon) + \varepsilon U^{k_0,l_0}(x;\varepsilon), \tag{1.106}$$

and a perturbation $\varepsilon^{m+1} S(y, x, s; \varepsilon)$, which is a $\mathcal{C}^{r-2m}$ function with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n - 2m$.

The function $h^0(y;\varepsilon)$ in the integrable part K_0 is a $\mathcal{C}^{r-2m+2}$ function in y with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n - 2m + 2$ of the form

$$h^0(y;\varepsilon) = \frac{y^2}{2}\widehat{h}(y;\varepsilon) = \frac{y^2}{2}(1 + \varepsilon k_0^2 \widetilde{h}(y;\varepsilon)), \tag{1.107}$$

for some $\mathcal{C}^{r-2m}$ function in (y, ε), $\widetilde{h}(y;\varepsilon)$, with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n - 2m$ in y. The function U^{k_0,l_0} in K_0 is given by

$$U^{k_0,l_0}(x;\varepsilon) = k_0^2(U_m^{k_0,l_0}(x + \theta_1(\varepsilon);\varepsilon) - U_m^{k_0,l_0}(\theta_1(\varepsilon);\varepsilon)), \tag{1.108}$$

and it satisfies

$$\left|\varepsilon U^{k_0,l_0}(\cdot;\varepsilon)\right|_{\mathcal{C}^\ell} \preceq \varepsilon |k_0|^2 |(k_0, l_0)|^{-r} \tag{1.109}$$

for $\ell = 0, \dots n$.

We also notice that the following conditions are satisfied,

$$h^0(0;\varepsilon) = \frac{\partial h^0}{\partial y}(0;\varepsilon) = 0, U^{k_0,l_0}(0;\varepsilon) = \frac{\partial U^{k_0,l_0}}{\partial x}(0;\varepsilon) = 0, \frac{\partial^2 U^{k_0,l_0}}{\partial x^2}(0;\varepsilon) < 0,$$

as well as that $x = 0$ is a global maximum of U^{k_0,l_0}.

The perturbation term $\varepsilon^{m+1} S(y, x, s; \varepsilon)$ is given by

$$S(y, x, s; \varepsilon) = k_0^2 \bar{R}^m \left(-\frac{l_0}{k_0} + \frac{y + b(\varepsilon)}{k_0}, \frac{x + \theta_1(\varepsilon) - l_0 s}{k_0}, s; \varepsilon \right)$$

and by equation (1.100) it can be bounded in the variables (y, x) by

$$\left|\varepsilon^{m+1} S(\cdot, s; \varepsilon)\right|_{\mathcal{C}^\ell} \preceq |k_0|^{2-\ell} \varepsilon^{m+1-\alpha(\ell+2m)} \tag{1.110}$$

for $\ell = 0, \ldots, n - 2m$.

Since we will want to apply some of the results in [DLS06a], it will be convenient for us to have K_0 written in another way adapted to the notation in [DLS06a]. Motivated by the size $\varepsilon|k_0|^2|(k_0, l_0)|^{-r}$ of $\varepsilon U^{k_0,l_0}$ estimated in formula (1.109), we introduce here the parameter $\gamma \in \mathbb{R}$, $2 > \gamma \geq 1$, depending on (k_0, l_0) and ε, such that

$$\varepsilon^\gamma = \varepsilon|k_0|^2|(k_0, l_0)|^{-r}, \tag{1.111}$$

in such a way that $\varepsilon U^{k_0,l_0}(\cdot; \varepsilon) = \mathcal{O}_{\mathcal{C}^\ell}(\varepsilon^\gamma)$, for $\ell = 0, \ldots, n$.

Notice that $\gamma = 1$ for small values of (k_0, l_0), that is $|(k_0, l_0)| \sim 1$, and in general $1 < \gamma < 2 + \nu$ for $|(k_0, l_0)| \sim \varepsilon^{-\varrho}$, for any $0 < \varrho < (1 + \nu)/r$, where $0 < \nu \leq 1/16$.

With this choice of γ, we will denote K_0 the one degree of freedom $\mathcal{C}^{r-2m+2}$ Hamiltonian in (y, x)

$$K_0(y, x; \varepsilon) = h^0(y; \varepsilon) + \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(x; \varepsilon), \tag{1.112}$$

where

$$\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(x; \varepsilon) = \varepsilon U^{k_0,l_0}(x; \varepsilon), \tag{1.113}$$

with $2 + \nu > \gamma \geq 1$ and $\widetilde{U}^{k_0,l_0}(\cdot; \varepsilon) = \mathcal{O}_{\mathcal{C}^\ell}(1)$, for $\ell = 0, \ldots, n$.

The energy level $K_0(y, x; \varepsilon) = 0$ consists of the saddle $(0, 0)$ and its separatrices.

The Hamiltonian $K(y, x, s; \varepsilon)$ introduced in (1.105) is $2\pi k_0$-periodic in x and 2π-periodic in s and is defined in the domain D_{k_0} given by

$$D_{k_0} = \{(y, x, s) \in \mathbb{R} \times \mathbb{R}/(2\pi k_0 \mathbb{Z}) \times \mathbb{T}, |y| \leq \bar{L}\}, \tag{1.114}$$

where $\bar{L} = k_0 \bar{L}_{k_0}$, whereas the integrable part $K_0(y, x; \varepsilon)$ in (1.112) is 2π-periodic in x and independent of s, therefore the region D_{k_0} can be seen as k_0 copies of the region

$$D = \{(y, x, s) \in \mathbb{R} \times \mathbb{T}^2, |y| \leq \bar{L}\}.$$

This effect is colloquially described as saying that the resonance $-l_0/k_0$ has k_0 eyes. As k_0 increases, these eyes form long necklaces.

The region D (and also D_{k_0}) is filled by the energy surfaces of the Hamiltonian K_0,

$$\mathcal{T}_E^0 = \{(y, x, s) \in [-\bar{L}, \bar{L}] \times \mathbb{T}^2 : K_0(y, x; \varepsilon) = E\}$$

which are invariant under the flow of the Hamiltonian K_0.

As we already said, the energy surface $\mathcal{T}_0^0$ corresponding to $E = 0$ consists of the saddle $(0, 0)$ and its separatrices forming a separatrix loop. Therefore, this separatrix loop $\mathcal{T}_0^0$ separates two types of topological invariant objects. The energy surfaces corresponding to the values $E > 0$ are primary tori and the ones corresponding to the the values $E < 0$ are called secondary tori, which are tori of different topology than the primary ones because they are contractible to points. Secondary tori cover all the region inside the separatrix

loop $\mathcal{T}_0^0$. In the next section we will discuss the persistence of primary and secondary tori when we add the perturbation term.

Primary and Secondary tori near the resonances with big gaps.

In this section, we will show that many of the invariant tori $\mathcal{T}_E^0$ of the Hamiltonian $K_0(y, x; \varepsilon)$ in (1.112), inside the region D_{k_0} given in (1.114), both primary and secondary, survive when we add the perturbation term $\varepsilon^{m+1}S(y, x, s; \varepsilon)$. Moreover, we will estimate the number of steps of averaging m required to get invariant tori with a distance of $\mathcal{O}(\varepsilon^{1+\eta})$ between them, for some $\eta > 0$, in terms of the original variables (I, φ, s).

To establish this we will write the Hamiltonian (1.112) into action-angle variables and apply KAM Theorem 1.3.22. Since the unperturbed Hamiltonian $K_0(y, x; \varepsilon)$ is a pendulum, we can not define global action-angle variables because the change of coordinates becomes singular on the separatrix. Therefore, we will define different action-angle variables inside and outside the separatrix and we will exclude a thin neighborhood around it.

We will find convenient to consider different regions in the domain D_{k_0} in terms of the values of the energy E, in which the behavior of the tori is different.

Recall that tori $\mathcal{T}_E^0$ in D_{k_0} are given approximately by the energy surfaces of Hamiltonian K_0, that is

$$K_0(y, x, s; \varepsilon) = E,$$

and we will see that excluding an small interval they can be seen as a graph of the action variable y over the angle variables (x, s).

Introducing $\delta = \varepsilon^\gamma$, we consider the foliation given by the level sets

$$h^0(y; \varepsilon) + \delta \widetilde{U}^{k_0,l_0}(x; \varepsilon) = E, \tag{1.115}$$

where $h^0(y; \varepsilon)$ is of the form (1.107) and on $x = 0$ there is a non-degenerate global maximum of $\widetilde{U}^{k_0,l_0}(x; \varepsilon)$, which verifies $-c \leq \widetilde{U}^{k_0,l_0}(\cdot; \varepsilon) \leq 0$ and $\widetilde{U}^{k_0,l_0}(\cdot; \varepsilon) \simeq -ax^2$ as $x \to 0$, with $a > 0$.

Since $h^0(y; \varepsilon) + \delta \widetilde{U}^{k_0,l_0}(x; \varepsilon) \simeq \frac{y^2}{2} + \delta \widetilde{U}^{k_0,l_0}(x; \varepsilon)$, the main term in the solution of (1.115) is

$$y = \pm \ell(x, E), \tag{1.116}$$

where

$$\ell(x, E) = \sqrt{2(E - \delta \widetilde{U}^{k_0,l_0}(x; \varepsilon))}. \tag{1.117}$$

Writing y in (1.115) as a function of (1.116), we can apply the implicit function Theorem to equation (1.115) and we get a solution $y = \mathcal{Y}_\pm(x, E)$ given by

$$\mathcal{Y}_\pm(x, E) = \pm(1 + \varepsilon b)\ell(x, E) + \varepsilon \widetilde{\mathcal{Y}}_\pm(\ell(x, E)), \tag{1.118}$$

where

i. b is independent of δ, and $\widetilde{\mathcal{Y}}_\pm(0) = \widetilde{\mathcal{Y}}'_\pm(0) = 0$.

ii. $\varepsilon\widetilde{\mathcal{Y}}_\pm$ is a $\mathcal{C}^{r-2m+2}$ function and

$$\begin{aligned}\left|\varepsilon\widetilde{\mathcal{Y}}_\pm \circ \ell\right|_{\mathcal{C}^s(\mathcal{I}_{E_0})} &\preceq \varepsilon, \quad s = 0, 1,\\ \left|\varepsilon\widetilde{\mathcal{Y}}_\pm \circ \ell\right|_{\mathcal{C}^s(\mathcal{I}_{E_0})} &\preceq \varepsilon E_0^{-s+1/2}, \quad 2 \leq s \leq n - 2m + 2,\end{aligned} \tag{1.119}$$

where $\mathcal{I}_{E_0} := \{(x, E), x \in \mathbb{T}, E \geq E_0 > 0\}$.

This result is stated explicitly in Lemma 8.34 in [DLS06a]. For more details and a rigorous proof we refer the reader to it.

From expression (1.117) it is clear that the size of the energy affects which are the dominant terms in $\ell(x, E)$. Thus, if $E \gg \delta = \varepsilon^\gamma$ the tori $\mathcal{T}_E^0$ are rather flat because the term $\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(x;\varepsilon)$ is very small compared with E, whereas if $E \leq \varepsilon^\gamma$, the term $\sqrt{E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(x;0)}$ and therefore the size of y oscillates between E and ε^γ and it has the effect of bending the tori up to the point that they are bunched near the critical point (see Figure 1.1).

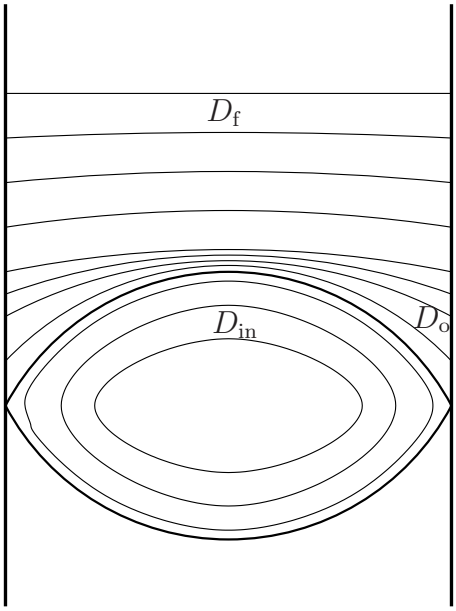

Figure 1.1: Schematic representation for the bending effect

Hence D_{k_0} will be divided in three regions in a similar way as in [DLS06a]: $D_{\rm f}$ is the region far from the separatrix, $D_{\rm o}$ close to the separatrix but outside the region bounded by the separatrix loop and $D_{\rm in}$ close to the separatrix but inside the separatrix loop, in the following way:

$$D_{\rm f} = \{(y, x, s) \in D_{k_0} : K_0(y, x; \varepsilon) = E, \varepsilon^\gamma \leq E \leq \bar{L}^2\} \tag{1.120}$$

$$D_{\rm o} = \{(y, x, s) \in D_{k_0} : K_0(y, x; \varepsilon) = F, \varepsilon^\beta \leq F \leq \varepsilon^\gamma\} \tag{1.121}$$

$$D_{\rm in} = \{(y, x, s) \in D_{k_0} : K_0(y, x; \varepsilon) = G, -\varepsilon^\gamma \leq G \leq -\varepsilon^\beta\} \tag{1.122}$$

where $1 \leq \gamma < 2+\nu$ as in (1.111) and β is arbitrary provided that $\beta > \gamma$ (see Figure 1.1).

Theorem 1.3.30 establishes the existence of primary tori in $D_{\rm f} \cup D_{\rm o}$ and secondary tori in $D_{\rm in}$ at a certain distance between them that depends on m and close to the level sets of the averaged Hamiltonian $K_0(y, x; \varepsilon)$.

Theorem 1.3.30. *Consider the $\mathcal{C}^{r-2m}$ reduced Hamiltonian $K(y, x, s; \varepsilon)$ given in (1.105) inside the region D_{k_0} defined in* (1.114)*. Consider $\beta > \gamma$, with γ as in (1.111) and assume that $r > (n-2)m+2$, $n \geq 2m+6$ and $m \geq 14(\beta-\gamma)+3\gamma/2$. Then, for $|\varepsilon|$ small enough, one has:*

1. ***Primary tori far from resonance**. There exists a set of values $E_1 < \cdots < E_{l_E}$ verifying $\varepsilon^\gamma \leq E_i \leq \bar{L}^2$, such that*

   (a) *The frequencies $\omega(E_i)$ are Diophantine numbers of constant type and Markov constant $K E_i^{-1/4} \varepsilon^{\frac{m+1-\alpha(6+2m)}{2}} |k_0|^{-2}$.*

   (b) *For any value E_i, there exist two primary invariant tori $\mathcal{T}^{\pm}_{E_i}$ of Hamiltonian (1.105) contained in $D_{\rm f}$.*

   (c) *The motion of the tori $\mathcal{T}^{\pm}_{E_i}$ is $\mathcal{C}^1$-conjugated to a rigid translation of frequencies $(\omega(E_i), 1)$.*

   (d) *This tori can be written as*

$$\mathcal{T}^{+}_{E_i} = \{(y, x, s) \in D_{\rm f}, K_{E_i}(y, x, s; \varepsilon) = E_i, y > 0\}$$

$$\mathcal{T}^{-}_{E_i} = \{(y, x, s) \in D_{\rm f}, K_{E_i}(y, x, s; \varepsilon) = E_i, y < 0\}$$

   where $K_{E_i}(y, x, s; \varepsilon)$ is a $\mathcal{C}^{4-\varrho}$ function, for any $\varrho > 0$, given by

$$K_{E_i}(y, x, s; \varepsilon) = K_0(y, x; \varepsilon) + \mathcal{O}_{\mathcal{C}^2}\left(\varepsilon^{\frac{m+1-\alpha(6+2m)}{2}} E_i^{1/4} |k_0|^{-2}\right) \tag{1.123}$$

   (e) *$D_{\rm f} \subset \bigcup_i B(\mathcal{T}^{\pm}_{E_i}, \varepsilon^{\frac{m+1-\alpha(6+2m)}{2}} E_i^{1/4} |k_0|^{-2})$, where*

$$B(\mathcal{T}^{\pm}_{E}, \delta) = \{(y, x, s) \in D_{k_0}, |K_0(y, x; \varepsilon) - E| \leq \delta\}$$

2. ***Primary tori close to resonance.** There exists a set of values $F_1 < \cdots < F_{l_F}$ verifying $\varepsilon^\beta \leq F_i \leq \varepsilon^\gamma$, such that*

   (a) *The frequencies $\omega(F_i)$ are Diophantine numbers of constant type and Markov constant $K \varepsilon^{\frac{m+1-\alpha(6+2m)-\gamma/2+6\gamma}{2}} F_i^{-3} |k_0|^{-2}$.*

   (b) *For any value F_i, there exist two primary invariant tori $\mathcal{T}^{\pm}_{F_i}$ of Hamiltonian (1.105) contained in $D_{\rm o}$.*

   (c) *The motion of the tori $\mathcal{T}^{\pm}_{F_i}$ is $\mathcal{C}^1$-conjugated to a rigid translation of frequencies $(\omega(F_i), 1)$.*

(d) This tori can be written as

$$\mathcal{T}^+_{F_i} = \{(y,x,s) \in D_{\rm o}, K_{F_i}(y,x,s;\varepsilon) = F_i, y > 0\}$$
$$\mathcal{T}^-_{F_i} = \{(y,x,s) \in D_{\rm o}, K_{F_i}(y,x,s;\varepsilon) = F_i, y < 0\}$$

where $K_{F_i}(y,x,s;\varepsilon)$ is a $\mathcal{C}^{4-\varrho}$ function, for any $\varrho > 0$, given by

$$K_{F_i}(y,x,s;\varepsilon) = K_0(y,x;\varepsilon) + \mathcal{O}_{\mathcal{C}^2}\left(\varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+14\gamma}{2}} F_i^{-7}|k_0|^{-2}\right) \quad (1.124)$$

(e) $D_{\rm o} \subset \bigcup_i B(\mathcal{T}^\pm_{F_i}, \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+10\gamma}{2}} F_i^{-5}|k_0|^{-2})$, where

$$B(\mathcal{T}^\pm_E, \delta) = \{(y,x,s) \in D_{k_0}, |K_0(y,x;\varepsilon) - E| \le \delta\}$$

3. ***Secondary tori close to resonance.*** *There exists a set of values $G_1 < \cdots < G_{l_G}$ verifying $-\varepsilon^\gamma \le G_i \le -\varepsilon^\beta$, such that*

   (a) The frequencies $\omega(G_i)$ are Diophantine numbers of constant type and Markov constant $K\varepsilon^{\frac{m+1-\alpha(6+2m)-\gamma/2+6\gamma}{2}} G_i^{-3}|k_0|^{-2}$.

   (b) For any value G_i, there exist a secundary invariant torus $\mathcal{T}^\pm_{G_i}$ of Hamiltonian (1.105) contained in $D_{\rm in}$, contractible to the set

$$\{(0,a,s), a \in \mathbb{R}, s \in \mathbb{R}/(2\pi k_0 \mathbb{Z})\} \subset D_{\rm in}$$

   (c) The motion on the torus $\mathcal{T}_{G_i}$ is $\mathcal{C}^1$-conjugated to a rigid translation of frequencies $(\omega(G_i), 1)$.

   (d) This torus can be written as

$$\mathcal{T}_{G_i} = \{(y,x,s) \in D_{\rm in}, K_{G_i}(y,x,s;\varepsilon) = G_i\}$$

   where $K_{G_i}(y,x,s;\varepsilon)$ is a $\mathcal{C}^{4-\varrho}$ function, for any $\varrho > 0$, given by

$$K_{G_i}(y,x,s;\varepsilon) = K_0(y,x;\varepsilon) + \mathcal{O}_{\mathcal{C}^2}\left(\varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+14\gamma}{2}} G_i^{-7}|k_0|^{-2}\right) \quad (1.125)$$

   (e) $D_{\rm in} \subset \bigcup_i B(\mathcal{T}^\pm_{G_i}, \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+10\gamma}{2}} G_i^{-5}|k_0|^{-2})$.

The following Corollary makes more explicit the assertions about the proximity of these tori as a function of m, and it also gives properties of the KAM tori when expressed as graphs of the action y in terms of the angle variables (x,s).

Corollary 1.3.31. *Consider the reduced Hamiltonian $K(y,x,s;\varepsilon)$ given in* (1.105) *inside the region D_{k_0} defined in* (1.114)*. Consider $\beta = \gamma/2 + 1 + \nu/2$, with $1 \le \gamma < 2 + \nu$ as in* (1.111) *and $0 < \nu \le 1/16$. Assume that $r > (n-2)m + 2$, $n \ge 2m + 6$ and $m \ge 10$. Then, the tori obtained in Theorem 1.3.30 verify:*

1. *For any value E_i, the primary tori $\mathcal{T}^\pm_{E_i}$ can be written as graphs of the action y over the angles (x,s):*

$$\mathcal{T}^\pm_{E_i} = \{(y,x,s) \in D_{\rm f}, y = f^\pm_{E_i}(x,s;\varepsilon)\}.$$

2. *For any value F_i, the primary tori $\mathcal{T}^{\pm}_{F_i}$ can be written as graphs of the action y over the angles (x,s):*
$$\mathcal{T}^{\pm}_{F_i} = \{(y,x,s) \in D_{\rm o}, y = f^{\pm}_{F_i}(x,s;\varepsilon)\}.$$

3. *There exists $\rho_0 > 0$ such that for any $0 < \rho_0 \leq \rho \leq \pi$, and for any value G_i, each of the components of*
$$\mathcal{T}_{G_i} \cap \{(y,x,s) : x \in I_\rho\}, \qquad I_\rho = \bigcup_{l=0}^{k_0-1} [2\pi l + \rho, 2\pi(l+1) - \rho],$$
that we will denote by $\mathcal{T}^{\pm,\rho}_{G_i}$, can be written as a graph of the action y over the angles (x,s):
$$\mathcal{T}^{\pm,\rho}_{G_i} = \{(y,x,s) \in D_{\rm i}, x \in I_\rho, y = f^{\pm}_{G_i}(x,s;\varepsilon)\}$$

4. *All these functions $f_v = f^{\pm}_v$ are at least of class $\mathcal{C}^2$ with respect to (x,s), and, denoting by D the derivatives with respect to x and s, for $v = E_i$, $i = 1,\dots,l_E$, $v = F_i$, $i = 1,\dots,l_F$, and $v = G_i$, $i = 1,\dots,l_G$, they verify:*

   (a) *There exists a function $\mathcal{Y}(x,E)$ given explicitly in* (1.118) *such that:*
   $$|f_v - \mathcal{Y}(x,v)|_{\mathcal{C}^1} \preceq |k_0|^{-2}\varepsilon^{1+\nu/2} \tag{1.126}$$

   (b) $|Df_v| \preceq \varepsilon^{\gamma/2}$, $|D^2 f_v| \preceq \varepsilon^{\gamma/2}$.

   (c) *For any two consecutive values v and $\overline{v}$ we have:*
   $$|v - \bar{v}| \preceq \varepsilon^{\beta},$$
   and
   $$|f_v - f_{\overline{v}}|_{\mathcal{C}^1} \preceq \frac{|v - \overline{v}|}{\varepsilon^{\gamma/2}} \preceq \varepsilon^{1+\nu/2}.$$

Proof of Theorem 1.3.30. The proof follows the strategy established in [DLS06a], with the same scaling in the domains $D_{\rm o}$ and $D_{\rm in}$. The main difference is that we will perform a sequence of scalings in the far domain $D_{\rm f}$, whereas in [DLS06a] there was no scaling in this region. This sequence of scalings in $D_{\rm f}$ will reduce the number of averaging steps m needed to get tori close enough in the region $D_{\rm f}$, and therefore the required differentiability r.

We will first give a detailed proof of part 1) of this Theorem. Notice that in $D_{\rm f}$ defined in (1.120), the energy E ranges from ε^γ to $\bar{L}^2 \sim \varepsilon^{2\alpha}$. Hence, we consider a value of E, let us say E_i, in the interval $[\varepsilon^\gamma, \varepsilon^{2\alpha}]$ and a small neighborhood around it of the form $[c_a E_i, c_b E_i] \subseteq [\varepsilon^\gamma, \varepsilon^{2\alpha}]$, where c_a, c_b are constants independent of ε and E_i, such that $c_a < 1$ and $c_b > 1$. Thus, we introduce the following domain contained in $D_{\rm f}$:
$$D_{E_i} = \{(y,x,s) \in D_{\rm f} : K_0(y,x;\varepsilon) = E,\ c_a E_i \leq E \leq c_b E_i\}. \tag{1.127}$$
By the equation of K_0 in (1.112) and the expression of h^0 in (1.107), the main term in y is given in (1.117). Therefore, in D_{E_i} the coordinate y is of size $\mathcal{O}(\sqrt{E_i})$ and it is natural to perform the scaling
$$y = \sqrt{E_i} Y, \tag{1.128}$$

which transforms the Hamiltonian system of Hamiltonian $K(y,x,s;\varepsilon)$ given in (1.105), which is $\mathcal{C}^{r-2m}$ with respect to the variables (y,x,s) with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n-2m$, into a Hamiltonian system of $\mathcal{C}^{r-2m}$ Hamiltonian with respect to (Y,x,s) with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n-2m$,

$$\mathcal{K}(Y,x,s;\sqrt{E_i},\varepsilon) = \sqrt{E_i}\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) + \frac{\varepsilon^{m+1}}{\sqrt{E_i}} S(\sqrt{E_i}Y,x,s;\varepsilon), \tag{1.129}$$

with

$$\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) = \frac{Y^2}{2}\widehat{h}(\sqrt{E_i}Y;\varepsilon) + \frac{\varepsilon^\gamma}{E_i}\widetilde{U}^{k_0,l_0}(x;\varepsilon), \tag{1.130}$$

where $\widehat{h}(y;\varepsilon) = 1+\mathcal{O}(\varepsilon)$ is given in (1.107) and, consequently, $\mathcal{K}_0$ is a $\mathcal{C}^{r-2m+2}$ function with respect to (Y,x) with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n-2m+2$, because $\widehat{h}(y;\varepsilon)$ is $\mathcal{C}^{r-2m+2}$ with respect to y with a bounded $\mathcal{C}^\ell$ norm up to $\ell = n-2m+2$.

The scaling (1.128) transforms the domain D_{E_i} in (1.127) into

$$\begin{aligned}\widetilde{D} &= \{(Y,x,s)\in\mathbb{R}\times\mathbb{R}/2\pi k_0\mathbb{Z}\times\mathbb{T} : \mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) = E/E_i,\ c_aE_i\le E\le c_bE_i\}\\ &= \{(Y,x,s)\in\mathbb{R}\times\mathbb{R}/2\pi k_0\mathbb{Z}\times\mathbb{T} : \mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) = e,\ c_a\le e\le c_b\}.\end{aligned} \tag{1.131}$$

Notice that since $\varepsilon\le E_i\le\varepsilon^{2\alpha}$ depends on ε the domain D_{E_i} depends on ε, whereas $\widetilde{D}$ is now independent of ε.

Next we will define the action-angle variables (A,ψ) associated to the Hamiltonian $\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon)$ in the domain $\widetilde{D}$. Note that the Hamiltonian $\mathcal{K}(Y,x,s;\sqrt{E_i},\varepsilon)$ is $2\pi k_0$-periodic in x and 2π-periodic in s, whereas $\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon)$ is 2π-periodic in x and independent of s. Therefore, the domain $\widetilde{D}$ is nothing else but k_0 copies of the domain $D^*\times\mathbb{T}$, where

$$D^* = \{(Y,x)\in\mathbb{R}\times\mathbb{T} : \mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) = e,\ c_a\le e\le c_b\}. \tag{1.132}$$

Notice that, by expression (1.130) for $\mathcal{K}_0$, the equation

$$\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) = e$$

has the same form as equation (1.115) with $\delta = \varepsilon^\gamma/E_i$ and it defines two functions $Y = \mathcal{Y}_\pm(x,e)$ on D^*, given in (1.118), which are of the form

$$\mathcal{Y}_\pm(x,e) = \pm\sqrt{2\left(e - \frac{\varepsilon^\gamma}{E_i}\widetilde{U}^{k_0,l_0}(x;\varepsilon)\right)}(1+\mathcal{O}_{\mathcal{C}^{n-2m+2}}(\varepsilon)).$$

Since, by construction of $\widetilde{U}^{k_0,l_0}(x;\varepsilon)$, on $x=0$ there is a global maximum such that $-c\le\widetilde{U}^{k_0,l_0}(x;\varepsilon)\le 0$, in the domain D^* we have

$$0\le c_a\le e\le e - \frac{\varepsilon^\gamma}{E_i}\widetilde{U}^{k_0,l_0}(x;\varepsilon)\le e + c\frac{\varepsilon^\gamma}{E_i}\le c_b + \text{cte}\,,$$

therefore $c_a\le\mathcal{Y}_\pm(x,e)\le c_b+\text{cte}$ and is $\mathcal{O}_{\mathcal{C}^{n-2m+2}}(1)$.

We consider in D^* the action-angle variables

$$\begin{array}{rcl} A &=& \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} \mathcal{Y}_\pm(x,e)\mathrm{d}x, \\ \psi &=& \dfrac{2\pi}{T(e)}\tau(x,e), \end{array} \qquad rcl \tag{1.133}$$

where $\tau(x,e)$ is the time along the orbit of the Hamiltonian $\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon)$ with energy e given by

$$\tau(x,e) = \int_0^x \frac{\partial \mathcal{Y}_\pm}{\partial e}(u,e)\mathrm{d}u. \tag{1.134}$$

We have chosen the origin of time at $x=0$ and with this choice $T(e)=\tau(2\pi,e)$ is the period of the periodic orbit.

From expression (1.133) it is obvious that the variable A is $\widetilde{c}_a \le A \le \widetilde{c}_b$ and is $\mathcal{O}_{\mathcal{C}^{n-2m+3}}(1)$, for some constants $\widetilde{c}_a$ and $\widetilde{c}_b$.

The action-angle variables (A,ψ) introduced in (1.133) have already been studied in Proposition 8.35 of [DLS06a] for the case when they become singular, that is when the domain D^* depends on ε. In our case, we can adapt the result in Proposition 8.35 of [DLS06a] when the domain D^* does not depend on ε. We obtain that we can express the integrable Hamiltonian $\sqrt{E_i}\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon)$ in (1.130) into action-angle variables (A,ψ) in the domain D^* and the change of coordinates is away from the singularity in this domain. More precisely, there exists a $\mathcal{C}^{r-2m+2}$ change of variables in D^*

$$\begin{array}{rcl} \mathcal{X} : D^{**} &\to& D^* \\ (A,\psi) &\mapsto& (Y,x) \end{array} \tag{1.135}$$

given in (1.133) with $D^{**} = \{(A,\psi) : \widetilde{c}_a \le A \le \widetilde{c}_b, \psi \in \mathbb{T}\} = [\widetilde{c}_a,\widetilde{c}_b]\times\mathbb{T}$ and $\widetilde{c}_a, \widetilde{c}_b$, are suitable constants independent of ε and E_i, such that:

i. $\mathcal{K}_0(\mathcal{X}(A,\psi);\sqrt{E_i},\varepsilon) = \mathcal{G}(A;\sqrt{E_i},\varepsilon)$.

ii. $|\mathcal{X}|_{\mathcal{C}^{n_0}(\mathcal{D}^{**})} \le K$, $\left|\mathcal{X}^{-1}\right|_{\mathcal{C}^{n_0}(D^*)} \le K$, $0 \le n_0 \le n-2m+2$.

iii. $|\mathcal{G}|_{\mathcal{C}^3(D^{**})} \le K$ and $\left|\mathcal{G}''\right|_{\mathcal{C}^0(D^{**})} \ge K$

where K is a constant independent of ε and E_i.

Now, we consider the Hamiltonian $\mathcal{K}$ in (1.129) expressed in action-angle variables,

$$\widetilde{\mathcal{K}}(A,\psi,s;\sqrt{E_i},\varepsilon) = \sqrt{E_i}\mathcal{G}(A;\sqrt{E_i},\varepsilon) + \frac{\varepsilon^{m+1}}{\sqrt{E_i}}\widetilde{S}(A,\psi,s;\sqrt{E_i},\varepsilon), \tag{1.136}$$

where $\widetilde{\mathcal{K}} = \mathcal{K}\circ\mathcal{X}$ and $\widetilde{S} = S\circ\mathcal{X}$.

The Hamiltonian (1.136) is of the form (1.78) with $K_0 = \sqrt{E_i}\mathcal{G}(A;\sqrt{E_i},\varepsilon)$ and $K_1 = \varepsilon^{m+1}E_i^{-1/2}\widetilde{S}(A,\psi,s;\sqrt{E_i},\varepsilon)$ and $2\pi k_0$-periodic in ψ.

The functions $\mathcal{G}$ and $\widetilde{S}$ are $\mathcal{C}^{r-2m+2}$ and $\mathcal{C}^{r-2m}$ with bounded $\mathcal{C}^\ell$ norms up to $\ell = n-2m+2$ and $\ell = n-2m$ in the variables (A,ψ), respectively. Since we have assumed

in the hypotheses of Theorem 1.3.30 that $r > n \geq 2m+6$, $\mathcal{G}$ and $\widetilde{S}$ have a bounded $\mathcal{C}^6$ norm in the variables (A, ψ). Therefore, using Faa-di Bruno formula (1.173) and the bound for the $\mathcal{C}^6$ norm in the variables (y, x) for $\varepsilon^{m+1}S$ in expression (1.110) jointly with the bounds for the change of coordinates $\mathcal{X}$ in item ii) we have that, for any $s \in \mathbb{T}$,

$$\left| \frac{\varepsilon^{m+1}}{\sqrt{E_i}} \widetilde{S}(\cdot, s; \sqrt{E_i}, \varepsilon) \right|_{\mathcal{C}^6(D^{**}_{k_0})} \preceq |k_0|^{-4} E_i^{-1/2} \varepsilon^{m+1-\alpha(6+2m)},$$

where $D^{**}_{k_0} = [\widetilde{c}_a, \widetilde{c}_b] \times \mathbb{R}/2\pi k_0 \mathbb{Z}$ and by item iii) in this proof we have that

$$\sqrt{E_i} \left| \mathcal{G}''(\cdot; \sqrt{E_i}, \varepsilon) \right|_{\mathcal{C}^0(D^{**})} \geq K \sqrt{E_i}.$$

Therefore, we can apply KAM Theorem 1.3.22 to Hamiltonian (1.136) with $n_0 = 5$, $\beta = 1 - \varrho$, for any $\varrho > 0$, $\delta(\varepsilon) = |k_0|^{-4} E_i^{-1/2} \varepsilon^{m+1-\alpha(6+2m)}$ and $M(\varepsilon) = K\sqrt{E_i}$ and we obtain:

1. There exist a set of values A_l, such that the Hamiltonian $\mathcal{K} \circ \mathcal{X}$ has invariant tori given by
$$\mathcal{T}_l = \{(A, \psi, s) \in D^{**}_{k_0} \times \mathbb{T} : A = A_l + \mathcal{A}_l(\psi, s; \sqrt{E_i}, \varepsilon)\}$$
where $\mathcal{A}_l$ are $\mathcal{C}^{4-\varrho}$ functions in the variables (ψ, s), for any $\varrho > 0$ and
$$\left| \mathcal{A}_l(\cdot; \sqrt{E_i}, \varepsilon) \right|_{\mathcal{C}^2(\mathbb{R}/2\pi k_0 \mathbb{Z} \times \mathbb{T})} \preceq |k_0|^{-2} E_i^{-3/4} \varepsilon^{(m+1-\alpha(6+2m))/2}.$$

2. The motion of these tori is $\mathcal{C}^{2-\varrho}$-conjugate to a rigid translation of frequencies $(\omega(A_l), 1)$, where $\omega(A_l)$ is a Diophantine number of constant type and Markov constant $K|k_0|^{-2} E_i^{-1/4} \varepsilon^{(m+1-\alpha(6+2m))/2}$.

3. The union of neighborhoods of size $|k_0|^{-2} E_i^{-3/4} \varepsilon^{(m+1-\alpha(6+2m))/2}$ of these tori cover all the region $D^{**}_{k_0} \times \mathbb{T}$.

In the variables $(Y, x, s) = (\mathcal{X}_i(A, \psi), s)$, the torus $\mathcal{T}_l$ satisfies $\mathcal{K}_0(Y, x; \sqrt{E_i}, \varepsilon) = \mathcal{G}(A_l + \mathcal{A}_l(\psi, s; \sqrt{E_i}, \varepsilon))$, so that, introducing $\mathcal{G}(A_l; \sqrt{E_i}, \varepsilon) = e_l$ and using the estimates in items (ii) and (iii) in this proof, one obtains that the tori are given by

$$\begin{aligned} \mathcal{K}_0(Y, x; \sqrt{E_i}, \varepsilon) &= \mathcal{G}(A_l; \sqrt{E_i}, \varepsilon) + |\mathcal{G}|_{\mathcal{C}^3} |\mathcal{A}_l|_{\mathcal{C}^2} \left| \mathcal{X}^{-1} \right|_{\mathcal{C}^2} \\ &= e_l + \mathcal{O}_{\mathcal{C}^2} \left(|k_0|^{-2} E_i^{-3/4} \varepsilon^{(m+1-\alpha(6+2m))/2} \right) \end{aligned} \tag{1.137}$$

Multiplying by E_i and performing the scaling $y = \sqrt{E_i} Y$ one obtains that the tori are given by

$$K_0(y, x; \varepsilon) = E_i e_l + \mathcal{O}_{\mathcal{C}^2} \left(|k_0|^{-2} E_i^{1/4} \varepsilon^{\frac{m+1-\alpha(6+2m)}{2}} \right).$$

Finally $E_l = E_i e_l$, the tori are given by

$$K_0(y, x; \varepsilon) = E_l + \mathcal{O}_{\mathcal{C}^2} \left(|k_0|^{-2} E_l^{1/4} \varepsilon^{\frac{m+1-\alpha(6+2m)}{2}} \right).$$

By compactness of $D_{\rm f}$, the covering $\{{\rm int}(D_{E_i})\}_{i=1}^{\infty}$ of $D_{\rm f}$ admits a finite subcovering $D_{\rm f} = \bigcup_{i=0}^{N} {\rm int}(D_{E_i})$, and we get the claimed results in part 1 of Theorem 1.3.30.

The proof of parts 2) and 3) of this Theorem follows as in [DLS08]. The only difference is that we introduce a sequence of domains as we did in this proof in the far region and we perform adequate scalings which allow us to get better estimates for the functions describing the searched tori. More precisely, consider the region $D_{\rm o}$ (the case for $D_{\rm in}$ is analogous) and introduce the domain

$$D_{F_i} = \{(y,x,s) \in D_{\rm o} : K_0(y,x;\varepsilon) = F,\ c_a F_i \leq F \leq c_b F_i\},$$

analogous to (1.127) in part 1). Since the energy $F_i \leq \varepsilon^{\gamma}$ in $D_{\rm o}$ (see (1.121)), from the expression for the main term of y given by $\ell(x,E)$ in (1.117), the coordinate y ranges from $\sqrt{F_i}$ to $\varepsilon^{\gamma/2}$. Hence we perform the scaling $y = \varepsilon^{\gamma/2} Y$ and we proceed as in Lemma 8.36 in [DLS06a]. We obtain that the original system is transformed into a Hamiltonian system of $\mathcal{C}^{r-2m}$ Hamiltonian with respect to (Y,x,s) of the form

$$\mathcal{K}(Y,x,s;\sqrt{E_i},\varepsilon) = \sqrt{E_i}\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) + \varepsilon^{m+1-\gamma/2} S(\sqrt{E_i}Y,x,s;\varepsilon),$$

with

$$\mathcal{K}_0(Y,x;\sqrt{E_i},\varepsilon) = \frac{Y^2}{2}\widehat{h}(\sqrt{E_i}Y;\varepsilon) + \widetilde{U}^{k_0,l_0}(x;\varepsilon)$$

where $\widehat{h}(y;\varepsilon) = 1 + \mathcal{O}(\varepsilon)$ is given in (1.107). The Hamiltonian is defined now on the domain

$$\begin{aligned}\widetilde{D} &= \{(Y,x,s) \in \mathbb{R}\times\mathbb{R}/2\pi k_0\mathbb{Z}\times\mathbb{T} : \mathcal{K}_0(Y,x;\varepsilon^{\gamma/2}) = F/F_i,\ c_a F_i \leq F \leq c_b F_i\}\\ &= \{(Y,x,s) \in \mathbb{R}\times\mathbb{R}/2\pi k_0\mathbb{Z}\times\mathbb{T} : \mathcal{K}_0(Y,x;\varepsilon^{\gamma/2}) = e,\ c_a F_i/\varepsilon^{\gamma} \leq e \leq c_b F_i/\varepsilon^{\gamma}\}\end{aligned}$$

Next, we define the action-angle variables in the domain $\widetilde{D}$ by formulas (1.133). The only change is that we need to take into account that instead of expression (8.77) in [DLS06a] we have

$$c_a\frac{F_i}{\varepsilon^{\gamma}} \leq e - \widetilde{U}^{k_0,l_0}(x;\varepsilon) \leq c_b\frac{F_i}{\varepsilon^{\gamma}} + c \leq {\rm cte}\,,$$

and the perturbation $\varepsilon^{m+1-\gamma/2} S(Y,x,s;\varepsilon^{\gamma/2})$ can be bounded in the $\mathcal{C}^6$ norm in the variables (Y,x) by $\varepsilon^{m+1-\alpha(6+2m)}|k_0|^{-4}$.

Therefore we can apply Proposition 8.38 in [DLS06a] and proceed as in the proof of parts 2) and 3) of Theorem 8.30 in [DLS06a] replacing ε^j by ε^{γ}, ε^{m+1} by $\varepsilon^{m+1-\alpha(6+2m)}|k_0|^{-4}$ and $\varepsilon^{\beta-\gamma}$ by $F_i\varepsilon^{-\gamma}$. Finally, by compactness of $D_{\rm o}$ we get the claimed results. We skip the proof of these two parts and we refer the reader to Section 8.5.4 in [DLS06a] for it.

Once we have obtained the approximate expression for the invariant tori, taking into account that $E_i \preceq \varepsilon^{2\alpha}$ and $F_i, G_i > \varepsilon^{\beta}$, the condition $m > 14(\beta-\gamma) + 3\gamma/2$ guarantees that invariant tori are $\mathcal{O}_{\mathcal{C}^2}(\varepsilon^{\gamma})$-close to the level sets of the unperturbed Hamiltonian $K_0(y,x;\varepsilon)$ given in (1.112).

We have proved that the invariant tori in the domains $D_{\rm f}$, $D_{\rm o}$ and $D_{\rm in}$, given by the implicit equations (1.123), (1.124) and (1.125), are of the form

$$K_0(y,x,s;\varepsilon) = E + \nu g(y,x,s,E;\varepsilon) \tag{1.138}$$

with $|g|_{\mathcal{C}^2} \leq$ cte and $E = E_i$ and $\nu = \varepsilon^{\frac{m+1-\alpha(6+2m)}{2}} E_i^{1/4} |k_0|^{-2}$, $E = F_i$ and $\nu = \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+14(\gamma)}{2}} F_i^{-7} |k_0|^{-2}$ and $E = G_i$ and $\nu = \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+14(\gamma)}{2}} G_i^{-7} |k_0|^{-2}$, respectively. □

Proof of Corollary 1.3.31. It is totally analogous to the proof of corollary 8.31 in [DLS06a] and it follows from Theorem 1.3.30 just applying the implicit function Theorem.

We apply KAM Theorem, with $m \geq 10$ and $\beta = \gamma/2 + 1 + \nu/2$, where $1 \leq \gamma < 2 + \nu$ and $0 < \nu \leq 1/16$, which satisfy the hypotheses of Theorem 1.3.30 and we obtain an approximate expression for the invariant KAM tori. In order to write them as graphs of the action y over the angles (x, s), we apply Lemma 8.39 in [DLS06a] to the implicit equation (1.138) in the different cases.

We notice that by the choice $m \geq 10$ and $\beta = \gamma/2 + 1 + \nu/2$, where $0 < \nu \leq 1/16$, and using that $|k_0| \geq 1$, $E_i \preceq \varepsilon^{2\alpha}$ and $F_i, G_i > \varepsilon^\beta$, the condition $|\nu| \leq \varepsilon^\beta$ is verified in the three cases. Thus, we obtain results in items 1), 2), 3) 4b) and

$$|f_v - \mathcal{Y}(x, v)|_{\mathcal{C}^1} \preceq |k_0|^{-2} \varepsilon^{\beta - \gamma/2} = |k_0|^{-2} \varepsilon^{1+\nu/2},$$

as claimed in 4a).

Finally, from results 1e), 2e) and 3e) in Theorem 1.3.30 and definitions of D_{f}, D_{o} and D_{in} given in (1.120), (1.121) and (1.122) we obtain

$$\begin{aligned}
|E_i - E_{i+1}| &\preceq \varepsilon^{\frac{m+1-\alpha(6+2m)}{2}} (E_i^{1/4} + E_{i+1}^{1/4}) |k_0|^{-2} \\
|F_i - F_{i+1}| &\preceq \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+10\gamma}{2}} (F_i^{-5} + F_{i+1}^{-5}) |k_0|^{-2} \\
|G_i - G_{i+1}| &\preceq \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2+10\gamma}{2}} (|G_i|^{-5} + |G_{i+1}|^{-5}) |k_0|^{-2}
\end{aligned}$$

and taking into account that $E_1 \sim F_{l_F} \sim \varepsilon^\gamma$ and $F_1 \sim G_{l_G} \sim \varepsilon^\beta$ we get

$$\begin{aligned}
|E_1 - F_{l_F}| &\preceq \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2-10(\beta-\gamma)}{2}} |k_0|^{-2} \\
|F_1 - G_{l_G}| &\preceq (\varepsilon^\beta + \varepsilon^{\frac{m+1-\alpha(6+2m)+\gamma/2-10(\beta-\gamma)}{2}}) |k_0|^{-2}.
\end{aligned}$$

Since $\beta = \gamma/2 + 1 + \nu/2$, $m \geq 10$ and $|k_0| \geq 1$, all these exponents are bigger than β as claimed in item 4c). The last estimate in item 4c) follows from the inequalities above and the following bounds given by Lemma 8.39 in [DLS06a]

$$\left| \frac{\partial f_E}{\partial E} \right| \preceq \varepsilon^{-\gamma/2}, \qquad \left| \frac{\partial D f_E}{\partial E} \right| \preceq \varepsilon^{-\gamma/2}.$$

□

Remark 1.3.32. Note that in the case of a finite number of resonances as in [DLS06a] and $\alpha = 0$, the continuous scaling performed decreases the number of steps of averaging m required to have the tori at a distance smaller than ε between them, and therefore it reduces the needed regularity. In [DLS06a], if we choose $\beta = j/2 + 1 + \eta$ for $j = 1, 2$ and η small enough, then $m = 9$ steps of averaging are enough and $r \geq 26$. This improves the needed regularity for r which was $r \geq 52$ because m was chosen $m = 26$.

Going back to the original variables.

Theorem 1.3.30 gives KAM tori, both primary and secondary, in the variables (y, x, s). From equations (1.123), (1.124) and (1.125) in Theorem 1.3.30, we know that these tori are given approximately by the level sets of the Hamiltonian $K_0(y, x; \varepsilon)$ in (1.112).

We can write them in the original variables (I, φ, s) using the change given by Theorem 1.3.11 and changes (1.22), (1.101) and (1.104). More precisely, we have that the relation with the original variables is given in first order by

$$y = k_0 I + l_0 + \mathcal{O}_{\mathcal{C}^2}(|k_0|\varepsilon), \quad x = k_0\varphi + l_0 s.$$

Using expression (1.106) and (1.107) these invariant objects are given by the level sets of a $\mathcal{C}^{4-\varrho}$ function F, for any $\varrho > 0$, which has the form

$$\begin{aligned} F(I, \varphi, s; \varepsilon) &= \frac{(k_0 I + l_0 + \mathcal{O}_{\mathcal{C}^2}(|k_0|\varepsilon))^2}{2}(1 + \mathcal{O}_{\mathcal{C}^2}(\varepsilon)) + \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta; \varepsilon) \\ &\quad + \mathcal{O}_{\mathcal{C}^2}(\varepsilon^{\gamma/2+1+\nu/2}), \end{aligned} \tag{1.139}$$

where $\theta = k_0\varphi + l_0 s$. By the definition of γ in (1.111) jointly with $\widetilde{U}^{k_0,l_0}$ and U^{k_0,l_0} in (1.113) and (1.108), respectively, we get the expression (1.95) given in Theorem 1.3.28.

Moreover, from items (1), (2) and (3), together with the estimates in item (4a) in Corollary 1.3.31 we have that KAM tori can be written as graphs in the variables (y, x, s) of functions of the form

$$y = f_E^{\pm}(x, s; \varepsilon) = \mathcal{Y}_{\pm}(x, E) + \mathcal{O}_{\mathcal{C}^1}(|k_0|^{-2}\varepsilon^{1+\nu/2}).$$

Using the mentioned changes, we obtain that the tori inside the region $\mathcal{D}_{\mathrm{BG}}$, are given in the original variables (I, φ, s) by

$$I = \lambda_E^{\pm}(\varphi, s; \varepsilon) = -l_0/k_0 + \frac{1}{k_0}\mathcal{Y}_{\pm}(\theta, E) + \mathcal{O}_{\mathcal{C}^1}(|k_0|^{-1}\varepsilon^{1+\nu/2})$$

with $\theta = k_0\varphi + l_0 s$, where $\mathcal{Y}_{\pm}$ is given (1.118).

Finally, from item (4c) in Corollary 1.3.31 and by the definition of γ given in (1.111), we obtain the claimed results in item (iii) of Theorem 1.3.28. □

1.3.4 Proof of Theorem 1.3.1

The proof of Theorem 1.3.1 follows directly from the results obtained in Propositions 1.3.24, 1.3.26 and Theorem 1.3.28.

Choosing $n = 2m + 6$ and assuming $m \geq 10$ and $r > 2(m + 1)^2$, the hypotheses on r, n and m in the mentioned Propositions and Theorem are satisfied. Moreover, the choice $\eta = \min((1 - \alpha n)/2, \nu/2)$, fits clearly with the assumptions on η in Propositions 1.3.24 and 1.3.26, and also with the one in Theorem 1.3.28 with the condition $\nu \leq 1/16$.

By Propositions 1.3.24 and 1.3.26, the tori obtained in the non resonant region and in the resonant region with small gaps are primary and they are given by the level sets

of the same function $F = I + \mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta})$, so they are flat up to $\mathcal{O}_{\mathcal{C}^2}(\varepsilon^{1+\eta-2(1+\nu)/r})$. Both regions form the flat tori region. The explicit approximate expressions for the invariant tori are given implicitly by the function (1.88) and as a graph of the action I over the variables (φ, s) by (1.89), both functions in Proposition 1.3.26.

By hypotheses **H3'**, Theorem 1.3.28 provides a sequence of invariant KAM tori (both primary and secondary) for the big gaps region. In a connected component of this region of the form (1.82), these tori are given by the level sets of a function F in (1.95) and as a graph of the action I over the angle variables (φ, s), in (1.96). Moreover, the distance between consecutive tori is $\mathcal{O}(\varepsilon^{1+\eta})$ in terms of the action variable and $\mathcal{O}(\varepsilon^{\gamma/2+1+\nu/2})$ in terms of the energy. By the definition of γ given in (1.111), we get the claimed result. □

1.4 Construction of a transition chain

In the previous section, we have proved that in the NHIM $\widetilde{\Lambda}_\varepsilon$ there exists a discrete foliation of invariant tori $\mathcal{T}_i$ (primary and secondary) with graphs at a distance $\mathcal{O}_{\mathcal{C}^1}(\varepsilon^{1+\eta})$, for some $\eta > 0$. We have also shown that these tori are close to being the level sets of the averaged Hamiltonian, and we have given its first order perturbative calculation for the flat tori region $\mathcal{D}_{\mathrm{F}}$ and the big gaps region $\mathcal{D}_{\mathrm{BG}}$.

The goal of this section is to prove Proposition 1.4.1, which states that, assuming that the non-degeneracy conditions **H2"**, **H3"** and **H3"'** in Theorem 1.2.1 hold, there exists transversality between the foliation of invariant tori in $\widetilde{\Lambda}_\varepsilon$ provided by Theorem 1.3.1 and its image under the scattering map S_ε given in (1.20) and it is possible to construct a transition chain.

Recall that, as we said in Section 1.2.3.4, by Lemma 10.4 in [DLS06a] two submanifolds, like the invariant tori $\mathcal{T}_i$ of the NHIM $\widetilde{\Lambda}_\varepsilon$, have a transverse heteroclinic intersection provided they are transversal under the scattering map as submanifolds of $\widetilde{\Lambda}_\varepsilon$. Hence, Proposition 1.4.1 provides a transition chain through applications of the scattering map.

Proposition 1.4.1. *Consider a Hamiltonian* (1.4) *satisfying the hypotheses of Theorem 1.2.1. Pick two KAM tori $\mathcal{T}_\pm$ such that $|I(x_\pm) - I_\pm| \leq \varepsilon^{1+\eta}$ for some $x_\pm \in \mathcal{T}_\pm$ and $\eta > 0$ (these tori exist thanks to Theorem 1.3.1). Then, there exists a transition chain $\{\mathcal{T}_i\}_{i=0}^{N(\varepsilon)}$, where $N(\varepsilon) = C/\varepsilon$, in such a way that*

1. *The transition chain is obtained through applications of the scattering map. That is,*
$$S_\varepsilon(\mathcal{T}_i) \pitchfork_{\widetilde{\Lambda}_\varepsilon} \mathcal{T}_{i+1}.$$
2. $\mathcal{T}_0 = \mathcal{T}_-$, $\mathcal{T}_{N(\varepsilon)} = \mathcal{T}_+$.

Proof. The proof of Proposition 1.4.1 is postponed to Section 1.4.2 and is based on the results in the following Section 1.4.1.

1.4.1 The scattering map and the transversality of heteroclinic intersections

The main result of this section is Lemma 1.4.2, stated below, which considers a foliation $\mathcal{F}_F$ whose leaves are the level sets of a certain function F and provides an expression for

the action of the scattering map S_ε on this foliation in terms of the Hamiltonian function $\mathcal{S}_\varepsilon$ given in (1.19), generating its deformation. Moreover, it gives criteria to establish transversality between the foliation $\mathcal{F}_F$ and its image under the scattering map S_ε.

Lemma 1.4.2. *Consider the foliation $\mathcal{F}_F$ whose leaves L_E^F are the level sets of a certain function F:*

$$L_E^F = \{(I,\varphi,s) \in (I_-,I_+)\times\mathbb{T}^2, F(I,\varphi,s;\varepsilon) = E\},\ E \in (E_1,E_2).$$

Let S_ε be the scattering map introduced in (1.17)*, and $\mathcal{S}_\varepsilon = \mathcal{S}_0 + \varepsilon\mathcal{S}_1 + \mathcal{O}(\varepsilon^2)$ its Hamiltonian function given in* (1.19)*. In particular, notice that $S_0 = -\mathcal{L}^*$, where $\mathcal{L}^*$ is the reduced Poincaré function introduced in* (1.11)*. Then the image sets of the leaves L_E^F of $\mathcal{F}_F$ under the scattering map S_ε satisfy the equation*

$$F\circ S_\varepsilon = F + \varepsilon\{F,\mathcal{S}_0\} + \frac{\varepsilon^2}{2}(\{\{F,\mathcal{S}_0\},\mathcal{S}_0\} + \{F,\mathcal{S}_1\}) + \mathcal{O}(\varepsilon^{2+\varrho}), \tag{1.140}$$

where $\{F,\mathcal{S}_i\} = \partial_\varphi F\partial_I\mathcal{S}_i - \partial_I F\partial_\varphi\mathcal{S}_i$ is the Poisson bracket of the functions F and $\mathcal{S}_i$ and $\varrho > 0$. Moreover, assuming that

$$\frac{|\{F, F\circ S_\varepsilon^{-1}\}|}{|\nabla F|^2} \geq C\varepsilon, \tag{1.141}$$

where C is a constant independent of ε and E, and $F\circ S_\varepsilon^{-1}$ is given by

$$F\circ S_\varepsilon^{-1} = F - \varepsilon\{F,\mathcal{S}_0\} + \frac{\varepsilon^2}{2}(\{\{F,\mathcal{S}_0\},\mathcal{S}_0\} - \{F,\mathcal{S}_1\}) + \mathcal{O}(\varepsilon^{2+\varrho}), \tag{1.142}$$

the angle between the surfaces $L_{E'}^F$ and $S_\varepsilon(L_E^F)$ at the intersection points is bounded from below by $C\varepsilon$. Therefore, foliations $\mathcal{F}_F$ and $\mathcal{F}_{F\circ S_\varepsilon^{-1}}$ intersect transversally.

Remark 1.4.3. For the case of a function F which is $\mathcal{O}_{\mathcal{C}^2}(1)$, the scattering map increases (decreases) the energy E by order ε, provided that the first order term $\{F,\mathcal{L}^*\}$ in (1.140) satisfies

$$\{F,\mathcal{L}^*\} \not\equiv 0.$$

Remark 1.4.4. Using expression (1.140) and $\mathcal{S}_0 = -\mathcal{L}^*$, the condition for the transversality of the foliations (1.141) reads out

$$\frac{|\{F,\{F,\mathcal{L}^*\}\} + \varepsilon/2(-\{F,\{\{F,\mathcal{L}^*\},\mathcal{L}^*\}\} + \{F,\{F,\mathcal{S}_1\}\}) + \mathcal{O}(\varepsilon^{1+\varrho})|}{|\nabla F|^2} \geq C. \tag{1.143}$$

Notice that if F is $\mathcal{O}_{\mathcal{C}^2}(1)$ the term ε^2 can be neglected and the condition reduces to

$$\frac{|\{F,\{F,\mathcal{L}^*\}\}|}{|\nabla F|^2} \geq C. \tag{1.144}$$

Proof: In Section 1.2.3.2 we have shown that there exists a Hamiltonian function $\mathcal{S}_\varepsilon$ generating the deformation of the scattering map S_ε and we have given its first order perturbative computation in equation (1.19). Hence, taking into account that $\mathcal{S}_\varepsilon = \mathcal{S}_0 + \varepsilon\mathcal{S}_1 + \mathcal{O}(\varepsilon^2)$, it is clear that (see [CH82] for instance) $F \circ S_\varepsilon$ is given by

$$F \circ S_\varepsilon = F + \varepsilon\{F, \mathcal{S}_0\} + \frac{\varepsilon^2}{2}(\{\{F, \mathcal{S}_0\}, \mathcal{S}_0\} + \{F, \mathcal{S}_1\}) + \mathcal{O}(\varepsilon^{2+\varrho}),$$

for $\varrho > 0$, with $\mathcal{S}_0 = -\mathcal{L}^*$.

In order to show the transversality between the foliations $\mathcal{F}_F$ and $\mathcal{F}_{F\circ S_\varepsilon^{-1}}$, we need to obtain lower bounds for the angle of intersection. More precisely, the angle α between the normal vectors to the tangent planes to the surfaces $S_\varepsilon(L_E^F)$ and $L_{E'}^F$ is given by

$$\sin(\alpha) = \frac{|\nabla(F \circ S_\varepsilon^{-1}) \times \nabla F|}{|\nabla(F \circ S_\varepsilon^{-1})||\nabla F|} = \frac{|\{F, F \circ S_\varepsilon^{-1}\}|}{|\nabla(F \circ S_\varepsilon^{-1})||\nabla F|},$$

where $\{F, S_\varepsilon^{-1}\}$ is given in expression (1.142). From this expression one can see that $\sin(\alpha)$ is $\mathcal{O}(\varepsilon)$ and condition (1.141) gives the required transversality. □

As we have argued in the previous section the tori in $\widetilde{\Lambda}_\varepsilon$ have different behavior depending whether they are close or far from the separatrix. Thus, the tori in the flat tori region and in the big gaps region far from the resonance are rather flat, whereas they are bent in the big gaps region close to a resonance. The fact that the tori are not flat has the consequence that the dominant effect of comparing a torus with a torus in the image of the scattering map will include some extra terms. For this reason, we will divide the study in three cases: on the one hand, the flat tori region and on the other hand the resonant region with big gaps, where we will distinguish between far and close to the resonance.

1.4.1.1 The flat tori region

In Lemma 1.4.5, we apply Lemma 1.4.2 to the flat tori region $\mathcal{D}_{\mathrm{F}}$. By Theorem 1.3.1, in one connected component of this region the invariant tori are given by the leaves L_E^F of a foliation $\mathcal{F}_F$, where F is of the form (1.88). Moreover they can be written as a graph of the action I over the angle variables (φ, s): $I = \lambda_E(\varphi, s; \varepsilon)$, where λ_E is given in (1.89).

Lemma 1.4.5. *Let us consider a foliation $\mathcal{F}_F$ contained in a connected component of the flat tori region $\mathcal{D}_{\mathrm{F}}$, where the function F is of the form* (1.88)*, so that the equation $F(I, \varphi, s; \varepsilon) = E$ defines a smooth surface given as a graph $\lambda_E(\varphi, s; \varepsilon)$, with λ_E as in* (1.89)*.*

Assume that hypothesis **H2"** *is fulfilled. More precisely, the reduced Poincaré function $\mathcal{L}^*$ defined in* (1.11) *verifies, for any value of $(I, \varphi, s) \in H_- \cap \mathcal{D}_{\mathrm{F}}$ that the function*

$$\widetilde{\theta} \mapsto \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(I, \widetilde{\theta})$$

for $\widetilde{\theta} = \varphi - Is$ is positive (resp. negative) and non-constant for $\widetilde{\theta}$ on some set $\mathcal{J}_E^*$ (see (1.12)*). Then the foliations $\mathcal{F}_F$ and $\mathcal{F}_{F\circ S^{-1}}$ intersect transversally.*

More precisely, any surface $S_\varepsilon(L_E^F)$ intersects at some point the surface $L_{E'}^F$ for any E', $E'-E=\mathcal{O}(\varepsilon)$. The angle between the surfaces $S_\varepsilon(L_E^F)$ and $L_{E'}^F$ at the intersection can be bounded from below by $C\varepsilon$, where C is a constant independent of ε and E.

Proof: We will apply Lemma 1.4.2 with $F(I,\varphi,s;\varepsilon)=I+\mathcal{O}_{\mathcal{C}^2}(\varepsilon^{1+\eta-2(1+\nu)/r})$ and $I=\lambda_E(I,\varphi,s;\varepsilon)=E+\mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta})$ for some $0<\eta\leq 1/32$. We will see that provided hypothesis **H2"** is fulfilled, condition (1.141) of Lemma 1.4.2 is satisfied.

We first apply the scattering map to the implicit surface

$$L_E^F=\{(I,\varphi,s)\in\mathcal{D}_{\mathrm{F}},F(I,\varphi,s)=E\}.$$

Using the expression (1.140), we can compute $S_\varepsilon(L_E^F)$, which is given by

$$F\circ S_\varepsilon=E-\varepsilon\{F,\mathcal{L}^*\}+\mathcal{O}(\varepsilon^2). \tag{1.145}$$

where,

$$\begin{aligned}\{F,\mathcal{L}^*\} &= \frac{\partial F}{\partial\varphi}\frac{\partial\mathcal{L}^*}{\partial I}-\frac{\partial\mathcal{L}^*}{\partial\varphi}\frac{\partial F}{\partial I}\\ &= -(1+\mathcal{O}_{\mathcal{C}^1}(\varepsilon^{1+\eta-2(1+\nu)/r}))\frac{\partial\mathcal{L}^*}{\partial\widetilde{\theta}}+\mathcal{O}_{\mathcal{C}^1}(\varepsilon^{1+\eta-2(1+\nu)/r})\\ &= -\frac{\partial\mathcal{L}^*}{\partial\widetilde{\theta}}+\mathcal{O}_{\mathcal{C}^1}(\varepsilon^{1+\eta-2(1+\nu)/r})\end{aligned}$$

with $\widetilde{\theta}=\varphi-Is$. Evaluating on $I=E+\mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta})$, equation (1.145) reads out

$$F(I,\varphi,s;\varepsilon)=E+\varepsilon\frac{\partial\mathcal{L}^*}{\partial\widetilde{\theta}}(E,\varphi-Es)+\mathcal{O}(\varepsilon^{1+\varrho}).$$

for $\varrho>0$.

By hypothesis **H2"** in Theorem 1.2.1 the scattering map increases (resp. decreases) the energy by order ε. Hence, the surface $S_\varepsilon(L_E^F)$ intersects surfaces $L_{E'}^F$, for $E'>E$ and $E'-E=\mathcal{O}(\varepsilon)$.

Moreover, in order to see that they intersect transversally we need to check that condition (1.141) is satisfied. Notice that in this case, by Remark 1.4.4, condition (1.144) implies (1.141). Thus, we first compute

$$\begin{aligned}\{F,\{F,\mathcal{L}^*\}\} &= \left(\frac{\partial F}{\partial I}\right)^2\frac{\partial^2\mathcal{L}^*}{\partial\varphi^2}+\mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta-2(1+\nu)/r})\\ &= (1+\mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta-2(1+\nu)/r}))\frac{\partial^2\mathcal{L}^*}{\partial\widetilde{\theta}^2}+\mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta-2(1+\nu)/r})\\ &= \frac{\partial^2\mathcal{L}^*}{\partial\widetilde{\theta}^2}+\mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta-2(1+\nu)/r}).\end{aligned}$$

Since, by assumption, the function $\frac{\partial\mathcal{L}^*}{\partial\widetilde{\theta}}(E,\widetilde{\theta})$ is non-constant for $\widetilde{\theta}$ in $\mathcal{J}_E^*$, there exists an interval $\bar{\mathcal{J}}_E\subset\mathcal{J}_E^*$ where

$$\left|\frac{\partial^2\mathcal{L}^*}{\partial\widetilde{\theta}^2}\right|\geq C>0,$$

and using

$$|\nabla F| = 1 + \mathcal{O}_{\mathcal{C}^1}(\varepsilon^{1+\eta-2(1+\nu)/r}),$$

we have that condition (1.141) is satisfied and the angle between the surfaces $S_\varepsilon(L_E^F)$ and $L_{E'}^F$ at the intersection can be bounded from below by $C\varepsilon$, where C is a constant independent of ε and E. □

Remark 1.4.6. By Theorem 1.3.1, two consecutive tori are, at most, at distance of $\mathcal{O}(\varepsilon^{1+\eta})$, for some $\eta > 0$, in terms of the I variable. Moreover, these tori are $\mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta})$ close to the level sets of the action I.

Hence, we conclude that the image under the scattering map of a torus $\mathcal{T}_i$ in the flat tori region, given by $I = I_i + \mathcal{O}(\varepsilon^{1+\eta})$ intersects transversally another torus of this region given by $I = I_{i+1} + \mathcal{O}(\varepsilon^{1+\eta})$ with $|I_{i+1} - I_i| = \mathcal{O}(\varepsilon)$:

$$S_\varepsilon(\mathcal{T}_i) \pitchfork \mathcal{T}_{i+1}.$$

1.4.1.2 Big gaps region

In Lemma 1.4.7 we apply Lemma 1.4.2 in one connected component of the big gaps region $\mathcal{D}_{\mathrm{BG}}$. By Theorem 1.3.1, the invariant tori are given by the leaves L_E^F of a foliation $\mathcal{F}_F$ for a certain function F of the form (1.95). Moreover, they can be written as a graph of the action I over the angle variables (φ, s): $I = \lambda_E^\pm(\varphi, s; \varepsilon)$, with $\lambda_E^\pm$ as in (1.96). Recall that in this foliation, the leaves with $E > 0$ are primary KAM tori whereas the leaves with $E < 0$ are secondary.

The dominant terms in F and in the expressions $\lambda_E^\pm$ of these tori depend on the resonance $-l_0/k_0$ and the distance to the separatrix, which is measured in terms of E. Thus, on the one hand tori are bent when they approach the separatrix, that is when $E \to 0$, and on the other hand tori are flatter when the size $\varepsilon|(k_0, l_0)|^{-1/r}$ of the gap decreases, which is controlled by k_0 and therefore by γ (see (1.111) for a definition of γ).

In the following Lemma 1.4.7 we consider the different cases and we prove that conditions **H2"**, **H3"** and **H3"'** ensure the existence of a transversal intersection between the foliation $\mathcal{F}_F$ and its image under the scattering map.

Lemma 1.4.7. *Let us consider a connected component of the big gaps region $\mathcal{D}_{\mathrm{BG}}$ defined in* (1.82)*. Recall from formula* (1.95) *that, in this component, the function F is of the form*

$$F(I, \varphi, s; \varepsilon) = \frac{(k_0 I + l_0)^2}{2}(1 + \mathcal{O}_{\mathcal{C}^2,I}(\varepsilon)) + \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta; \varepsilon) + \mathcal{O}_{\mathcal{C}^2}(\varepsilon^{\gamma/2+1+\eta}), \tag{1.146}$$

where $\theta = k_0\varphi + l_0 s$, and for some $0 \le \rho < \pi$ and some range of energies $-\varepsilon^\gamma \le E \le L^2$, the equation $F(I, \varphi, s; \varepsilon) = E$ defines two smooth surfaces $L_E^{F\pm}$ given as graphs $I = \lambda_E^\pm(\varphi, s; \varepsilon)$, with $\lambda_E^\pm$ given in (1.96)*, of the form*

$$\lambda_E^\pm(\varphi, s; \varepsilon) = -\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_\pm(x, E) + \mathcal{O}_{\mathcal{C}^1}(\varepsilon^{1+\eta}), \tag{1.147}$$

where

$$\mathcal{Y}_\pm(x, E) = \pm(1 + \varepsilon b)\ell(\theta, E) + \varepsilon\widetilde{\mathcal{Y}}_\pm(\ell(\theta, E)), \tag{1.148}$$

for $\rho \leq \theta = k_0\varphi + l_0 s \leq 2\pi - \rho$ *and* $\ell(\theta, E) = \sqrt{2(E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))}$ *with* $\widetilde{U}^{k_0,l_0}(\theta;\varepsilon)$ *defined in* (1.113) *and* $\widetilde{\mathcal{Y}}_\pm$ *satisfying* (1.119).

Assume that hypothesis **H2"** *is fulfilled, more precisely, that the reduced Poincaré function* $\mathcal{L}^*$ *verifies, for any value of* $(I, \varphi, s) \in H_- \cap \mathcal{D}_{\mathrm{BG}}$, *that the function*

$$\widetilde{\theta} \mapsto \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(I, \widetilde{\theta})$$

for $\widetilde{\theta} = \varphi - Is$ *is positive (resp. negative) and non-constant for* $\widetilde{\theta} \in \mathcal{J}_I^*$.

For $|(k_0, l_0)| \prec \varepsilon^{-1/r}$ *assume hypothesis* **H3"** *on* (k_0, l_0) *in Theorem 1.2.1, which is that the function*

$$\theta \to \frac{k_0 \widetilde{U}'^{k_0,l_0}(\theta;0)\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}\left(\frac{-l_0}{k_0}, \frac{\theta}{k_0}\right) + 2\widetilde{U}^{k_0,l_0}(\theta;0)\frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}\left(\frac{-l_0}{k_0}, \frac{\theta}{k_0}\right)}{2\frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}\left(\frac{-l_0}{k_0}, \frac{\theta}{k_0}\right)} \tag{1.149}$$

is non-constant.

For $|(k_0, l_0)| \sim \varepsilon^{-1/r}$ *we assume the following hypothesis, which is condition* **H3"'** *on* (k_0, l_0) *in Theorem 1.2.1:*

There exists a constant C, *independent of* E *and* ε, *and an interval* $\mathcal{J} \subset \mathcal{J}^*_{-l_0/k_0}$ *such that given any* E, ε *in this region and* $\theta \in \mathcal{J}$,

$$\begin{aligned}
&\left| \frac{1}{2(E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} \left(2E \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right.\right. \\
&-\varepsilon^\gamma \left[k_0 \widetilde{U}'^{k_0,l_0}(\theta;0) \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) + 2\widetilde{U}^{k_0,l_0}(\theta;0) \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right] \\
&\left.\left. \pm \varepsilon k_0 \sqrt{2(E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right) \right| \geq C.
\end{aligned} \tag{1.150}$$

Then, the foliations $\mathcal{F}_F$ *and* $\mathcal{F}_{F \circ S^{-1}}$ *intersect transversally.*

More precisely, any surface $S_\varepsilon(L_E^{F,-})$ *intersects at some point the surface* $L_{E'}^{F,-}$ *for any* $E' < E$ *and* $|E' - E| \leq Ck_0\varepsilon \max(|E|^{1/2}, \varepsilon^{\gamma/2})$. *Analogously, any surface* $S_\varepsilon(L_E^{F,+})$ *intersects at some point the surface* $L_{E'}^{F,+}$ *for any* $E' > E$ *and* $|E' - E| \leq Ck_0\varepsilon \max(|E|^{1/2}, \varepsilon^{\gamma/2})$. *In some cases, it is possible that a certain surface* $S_\varepsilon(L_E^{F,-})$ *intersects the surface* $L_{E'}^{F,-}$ *with* $E' > E$ *and* $|E' - E| \leq Ck_0\varepsilon \max(|E|^{1/2}, \varepsilon^{\gamma/2})$

The angle between the surfaces $S_\varepsilon(L_E^{F,\pm})$ *and* $L_{E'}^{F,\pm}$ *at the intersection is bounded from below by* $C\varepsilon$, *where* C *is a constant independent of* ε *and* E.

Remark 1.4.8. Lemma 10.16 in [DLS06a] gives a computable sufficient condition that guarantees that hypothesis **H3"'** on (k_0, l_0) is verified. Indeed, let

$$\begin{aligned}
a(\theta) &= \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}), \\
b(\theta) &= -\frac{1}{2}\left(k_0 \widetilde{U}'^{k_0,l_0}(\theta;0) \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) + 2\widetilde{U}^{k_0,l_0}(\theta;0) \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right), \\
c(\theta) &= \pm \frac{\sqrt{2}}{2} \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}),
\end{aligned}$$

if there exist θ_1, θ_2 and θ_3 in some interval $\mathcal{J}$ verifying

$$\left| \begin{array}{ccc} \widetilde{a}(\theta_1) & \widetilde{a}(\theta_2) & \widetilde{a}(\theta_3) \\ \widetilde{b}(\theta_1) & \widetilde{b}(\theta_2) & \widetilde{b}(\theta_3) \\ \widetilde{c}(\theta_1) & \widetilde{c}(\theta_2) & \widetilde{c}(\theta_3) \end{array} \right| \neq 0, \tag{1.151}$$

where

$$\begin{aligned} \widetilde{a}(\theta) &= a(\theta)^2 \\ \widetilde{b}(\theta) &= 2a(\theta)b(\theta) - c(\theta)^2 \\ \widetilde{c}(\theta) &= b(\theta)^2 - c(\theta)^2 \widetilde{U}^{k_0,l_0}(\theta;0), \end{aligned} \tag{1.152}$$

then there exists a constant C and three intervals $\theta_i \in \mathcal{J}_i \subset \mathcal{J}$, $i = 1, 2, 3$ such that for any $\theta \in \mathcal{J}_i$

$$\left| \frac{a(\theta)E + b(\theta)\varepsilon^\gamma + c(\theta)\varepsilon k_0 \sqrt{E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0)}}{E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0)} \right| \geq C,$$

which is hypothesis **H3"'** on (k_0, l_0).

Proof: We will apply Lemma 1.4.2 to the foliation $\mathcal{F}_F$ given in (1.146).

We first apply the scattering map to the implicit surface

$$L_E^F = \{(I, \varphi, s) \in \mathcal{D}_{\mathrm{BG}}, F(I, \varphi, s) = E\}.$$

Using the expression (1.140), we can compute $S_\varepsilon(L_E^F)$, which is given by

$$(F \circ S_\varepsilon)(I, \varphi, s; \varepsilon) = E - \varepsilon\{F, \mathcal{L}^*\} + \frac{\varepsilon^2}{2}(\{\{F, \mathcal{L}^*\}, \mathcal{L}^*\} + \{F, \mathcal{S}_1\}) + \mathcal{O}(\varepsilon^{2+\rho}), \tag{1.153}$$

where

$$\begin{aligned} \{F, \mathcal{L}^*\} &= -\frac{\partial F}{\partial I}\frac{\partial \mathcal{L}^*}{\partial \varphi} + \frac{\partial F}{\partial \varphi}\frac{\partial \mathcal{L}^*}{\partial I} \\ &= -k_0(k_0 I + l_0)(1 + \mathcal{O}_{\mathcal{C}^2}(\varepsilon))\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(I, \widetilde{\theta}) + \varepsilon^\gamma k_0 \widetilde{U}'^{k_0,l_0}(\theta, \varepsilon)\frac{\partial \mathcal{L}^*}{\partial I}(I, \widetilde{\theta}) \\ &\quad + \mathcal{O}_{\mathcal{C}^1}(\varepsilon^{\gamma/2+1+\eta}) \end{aligned}$$

where $\widetilde{\theta} = \varphi - Is$ and $\theta = k_0\varphi + l_0 s$. Evaluating on $I = \lambda_E^\pm(\varphi, s; \varepsilon)$ introduced in (1.96) we have that

$$\begin{aligned} \{F, \quad & \mathcal{L}^*\} \\ & = -k_0 \mathcal{Y}_\pm(\theta, E)(1 + \mathcal{O}_{\mathcal{C}^2}(\varepsilon))\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}\left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_\pm(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_\pm(\theta, E)\right)s\right) \\ & + \varepsilon^\gamma k_0 \widetilde{U}'^{k_0,l_0}(\theta, \varepsilon)\frac{\partial \mathcal{L}^*}{\partial I}\left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_\pm(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_\pm(\theta, E)\right)s\right) \\ & + \mathcal{O}_{\mathcal{C}^1}(\varepsilon^{1+\eta}) \\ & = \mp k_0 \ell(\theta, E)\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}\left(-\frac{l_0}{k_0} + \frac{1}{k_0}\ell(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \frac{1}{k_0}\ell(\theta, E)\right)s\right) \\ & + \mathcal{O}_{\mathcal{C}^1}(k_0(\varepsilon\ell(\theta;\varepsilon) + \varepsilon^{1+\eta} + \varepsilon^\gamma)). \end{aligned} \tag{1.154}$$

Note that, in this case, the size of the main term of $\{F, \mathcal{L}^*\}$ depends on $\ell(\theta, E)$, which ranges from $\varepsilon^{\gamma/2}$ up to L, depending on the energy E. Roughly speaking, when the size of the energy is big, the term $\varepsilon\{F, \mathcal{L}^*\}$ in (1.153) is bigger than ε^2, so the terms of this order can be neglected. However, when the energy is small, the size of the term $\varepsilon\{F, \mathcal{L}^*\}$ can be of order ε^2 and in this case we need to take into account the terms of order ε^2 that appear in (1.153).

Hence, we choose μ such that $0 \leq \mu < \gamma$ and we distinguish two cases: the case when tori are close to the resonance, which corresponds to small values of the energy ($-\varepsilon^\gamma \leq E \leq \varepsilon^\mu$) and the case when they are reasonably far from a resonance, which corresponds to greater values of the energy ($\varepsilon^\mu \leq E \leq L^2$).

Far from the resonance: $\varepsilon^\mu \leq E \leq L^2$**.**

The case far from a resonance is analogous to the flat tori region, studied in the previous section, because in this case

$$\begin{aligned} \ell(\theta, E) &= \sqrt{2(E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;\varepsilon))} = \sqrt{2E}\sqrt{1 - \frac{\varepsilon^\gamma}{E}\widetilde{U}^{k_0,l_0}(\theta;\varepsilon)} \\ &= \sqrt{2E}(1 + \mathcal{O}(\varepsilon^{\gamma-\mu})), \end{aligned}$$

consequently, since $\sqrt{2E} \geq \sqrt{2}\varepsilon^{\mu/2}$ and $\varepsilon^{\mu/2} > \varepsilon^{\gamma/2} > k_0\varepsilon$, the function $\{F, \mathcal{L}^*\}$ becomes

$$\{F, \mathcal{L}^*\} = \mp k_0\sqrt{2E}\left(\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}\left(-\frac{l_0}{k_0} + \frac{\sqrt{2E}}{k_0}, \varphi - (-\frac{l_0}{k_0} + \frac{\sqrt{2E}}{k_0})s\right) + \mathcal{O}(k_0\varepsilon^{\gamma-\mu})\right),$$

and therefore in (1.153) for $S_\varepsilon(L_E^{F,\pm})$, the dominant terms become

$$F = E \pm \varepsilon k_0\sqrt{2E}\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}\left(-\frac{l_0}{k_0} + \frac{\sqrt{2E}}{k_0}, \varphi - (-\frac{l_0}{k_0} + \frac{\sqrt{2E}}{k_0})s\right) + \mathcal{O}(k_0 E^{1/2}\varepsilon^{1+\gamma-\mu}).$$

Then, if the function $\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(I, \widetilde{\theta})$ is not constant and positive as a function of $\widetilde{\theta}$, and using that $\nabla F = \pm k_0\sqrt{2E} + \mathcal{O}(\varepsilon^\gamma)$, Lemma 1.4.2 gives us the desired result.

Close to the resonance: $-\varepsilon^\gamma \leq E \leq \varepsilon^\mu$**.**

The case close to a resonance is more technical because the size of the energy is now comparable to the term $\varepsilon^\gamma \widetilde{U}^{k_0,l_0}$ and therefore $|\ell(\cdot, E)| = \max(\varepsilon^{\gamma/2}, |E|^{1/2})$.

As we already mentioned, in this case we need to take into account the term of order ε^2 in the expression (1.153) because in some cases it will be comparable to $\varepsilon\{F, \mathcal{L}^*\}$. Fortunately, among all the terms in $\varepsilon^2/2(\{\{F, \mathcal{L}^*\}, \mathcal{L}^*\} + \{F, \mathcal{S}_1\})$ we will see that there is a dominant one. To that end we notice first that all the terms that appear in the derivatives up to second order for F with respect to (I, φ, s) are $\mathcal{O}(I, \varepsilon^\gamma)$, except

$$\frac{\partial^2 F}{\partial I^2} = k_0^2(1 + \mathcal{O}_{\mathcal{C}^2}(\varepsilon)). \tag{1.155}$$

Hence, in the expression $\{\{F, \mathcal{L}^*\}, \mathcal{L}^*\} + \{F, \mathcal{S}_1\}$ all the terms are of order ε^ϱ, for some $\varrho > 0$, except

$$\frac{\partial^2 F}{\partial I^2}\left(\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(I, \widetilde{\theta})\right)^2.$$

Therefore, using this feature and (1.154), the expression (1.153) for $S_\varepsilon(L_E^{F,\pm})$ is given by

$$\begin{aligned} F(I,\varphi,s;\varepsilon) &= E \pm \varepsilon k_0\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))}\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \\ &\quad +\frac{\varepsilon^2}{2}k_0^2\left(\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right)^2+\mathcal{O}_{\mathcal{C}^1}(\varepsilon^{2+\varrho}) \\ &= E+\varepsilon\mathcal{M}_\pm(\theta;\varepsilon)+\mathcal{O}_{\mathcal{C}^1}(\varepsilon^{2+\varrho}), \end{aligned} \tag{1.156}$$

where

$$\mathcal{M}_\pm(\theta;\varepsilon)=k_0\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\left(\pm\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))}+\varepsilon k_0\frac{1}{2}\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right). \tag{1.157}$$

In the above expression (1.157) there appear two quantities that can be comparable or not depending on k_0. On the one hand, there is $\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))}$, which is related to the size of the gap and the other one there is $\varepsilon k_0\frac{1}{2}\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0})$, which is related to the size of the heteroclinic jumps provided by the scattering map S_ε. Hence we distinguish three situations depending on k_0:

i. If $\varepsilon^{\gamma/2}\prec k_0\varepsilon$, that is $|(k_0,l_0)|\succ\varepsilon^{-1/r}$ (see definition for γ in (1.111)) we have that the expression (1.156) reduces to

$$F(I,\varphi,s;\varepsilon)=E+\frac{\varepsilon^2}{2}k_0^2\left(\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}\left(-\frac{l_0}{k_0},\frac{\theta}{k_0}\right)\right)^2+\mathcal{O}_{\mathcal{C}^1}(k_0^2\varepsilon^{2+\varrho}),$$

for any $\varrho>0$. So, tori are essentially flat and this is equivalent to the flat tori case. Hence, condition **H2"** assures that the foliations intersect transversally.

ii. If $k_0\varepsilon\prec\varepsilon^{\gamma/2}$, that is $|(k_0,l_0)|\prec\varepsilon^{-1/r}$ (see definition for γ in (1.111)), we have that the expression (1.156) for $S_\varepsilon(L_E^{F,\pm})$ reduces to

$$F(I,\varphi,s;\varepsilon)=E\pm\varepsilon k_0\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))}\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}\left(-\frac{l_0}{k_0},\frac{\theta}{k_0}\right)+\mathcal{O}_{\mathcal{C}^1}(k_0^2\varepsilon^{2+\varrho}), \tag{1.158}$$

for any $\varrho>0$.

This is the case when the size of the gaps in the foliation of primary tori is bigger than the size of the heteroclinic jumps provided by the scattering map. Hence, if we consider the surface $L_E^{F,-}$, by hypothesis **H2"** we have that

$$-\varepsilon k_0\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))}\frac{\partial \mathcal{L}^*}{\partial\widetilde{\theta}}\left(-\frac{l_0}{k_0},\frac{\theta}{k_0}\right)$$

is a negative function, and therefore by equation (1.158) $S(L_E^{F,-})$ intersects surfaces $L_{E'}^{F,-}$ with $E'<E$ and $|E'-E|\preceq k_0\varepsilon\max(\varepsilon^{\gamma/2},|E|^{1/2})$. An analogous result is obtained for $L_E^{F,+}$ with $E'>E$.

iii. If $\varepsilon^{\gamma/2} \sim k_0\varepsilon$, which is the case when $|(k_0,l_0)| \sim \varepsilon^{-1/r}$ we have that the terms $\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))}$ and $\frac{1}{2}\varepsilon k_0 \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0})$ in the expression (1.157) are comparable. This case is the hardest to study because the size of the gap has the same order than the heteroclinic jumps. This causes that there are different geometries for $S_\varepsilon(L_E^{F,\pm})$ that could happen depending on the numerical values of the leading coefficients.

If we focus in the case of $S_\varepsilon(L_E^{F,-})$, by hypothesis **H2"**, the main term $\mathcal{M}_-$ in F given in (1.157) can have different signs depending on the size of $\ell(\theta;\varepsilon)$. According to that, we distinguish the following cases:

(a) The first case is when

$$\left| -\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} + \varepsilon k_0 \frac{1}{2}\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \right|_{\mathcal{C}^1} \leq \varepsilon^{1+\varrho}.$$

This case corresponds to points in $L_E^{F,-}$ that are $\mathcal{O}_{\mathcal{C}^2}(\varepsilon^{2+\varrho})$-close to homoclinic jumps $S_\varepsilon(L_E^{F,-}) \pitchfork L_E^{F,-}$. They are not good for diffusion.

(b) The second case is when

$$\left| -\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} + \varepsilon k_0 \frac{1}{2}\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \right|_{\mathcal{C}^1} > \varepsilon^{1+\varrho}.$$

This case corresponds to points in heteroclinic jumps $S_\varepsilon(L_E^{F,-}) \pitchfork L_{E'}^{F,-}$ and we can distinguish two situations that can take place.

On the one hand, if

$$-\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} > \varepsilon k_0 \frac{1}{2}\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}},$$

which is the case when the heteroclinic jumps are smaller than the gap, $S_\varepsilon(L_E^{F,-})$ intersects surfaces $L_{E'}^{F,-}$ with $E' < E$ and $|E'-E| \preceq k_0\varepsilon \max(\varepsilon^{\gamma/2}, |E|^{1/2})$. Thus, for small values of energy $E > 0$, the scattering map will connect a surface with energy $E > 0$ with a surface $E' < 0$, which corresponds to a heteroclinic connection of a primary tori with a secondary one.

On the other hand, when

$$-\sqrt{2(E-\varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} < \varepsilon k_0 \frac{1}{2}\frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}},$$

which is the case when the heteroclinic jumps are bigger than the gaps created between primary tori, we obtain that $S_\varepsilon(L_E^{F,-})$ will intersect the surfaces $L_{E'}^{F,-}$ with $E' > E$ and $|E'-E| \preceq k_0\varepsilon \max(\varepsilon^{\gamma/2}, |E|^{1/2})$. In this case the scattering map will connect two tori with positive energy, that is, two primary tori, and cross the gap with just one application of the scattering map.

Once we have a heteroclinic connection that crosses the separatrix loop, we can consider $S_\varepsilon(L_E^{F,+})$, which corresponds to the upper branch of the level set $F(I,\varphi,s;\varepsilon) = E$, $E > 0$. In this case, by hypothesis **H2"**, in expression (1.156) the main term $\mathcal{M}_+$ in F given in (1.157) is always positive, so $S_\varepsilon(L_E^{F,+})$ will intersect surfaces $L_{E'}^{F,+}$ with $E' > E$ and $|E' - E| \preceq k_0\varepsilon \max(\varepsilon^{\gamma/2}, |E|^{1/2})$.

Now, we want to check that the intersections for the cases (ii) and (iii) take place transversally by means of condition (1.141). For the case described in item (ii) in this proof, condition (1.144) implies condition (1.141). So, we first compute $\{F,\{F,\mathcal{L}^*\}\}$, where $\{F,\mathcal{L}^*\}$ is given by equation (1.154). Using again that the derivatives up to second order for F with respect to (I,φ,s) are $\mathcal{O}(I,\varepsilon^\gamma)$, except the one in (1.155), we get

$$\begin{aligned}\{F,\{F,\mathcal{L}^*\}\} &= \left(\frac{\partial F}{\partial I}\right)^2 \frac{\partial^2 \mathcal{L}^*}{\partial \varphi^2} - \frac{\partial F}{\partial \varphi}\frac{\partial^2 F}{\partial I^2}\frac{\partial \mathcal{L}^*}{\partial \varphi} + \mathcal{O}_{\mathcal{C}^0}(\varepsilon^{1+\eta}) \\ &= (k_0 I + l_0)^2 k_0^2 \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(\widetilde{\theta}, E) - \varepsilon^\gamma \widetilde{U}'^{k_0,l_0}(\theta,\varepsilon) k_0 k_0^2 \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(\widetilde{\theta}, E) \\ &\quad + \mathcal{O}_{\mathcal{C}^0}(k_0^2 \varepsilon \ell(\theta;\varepsilon) + \varepsilon^{1+\eta}).\end{aligned}$$

Using that

$$|\nabla F| = (k_0 I + l_0)k_0 + \mathcal{O}(k_0\varepsilon),$$

and evaluating on $I = \lambda_E^\pm(\varphi, s;\varepsilon)$, we have that the condition (1.141) is satisfied provided that

$$\begin{aligned}&\left| \frac{\pm 1}{2(E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} \left(2E \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right.\right. \\ &\left.\left. -\varepsilon^\gamma \left[k_0 \widetilde{U}'^{k_0,l_0}(\theta;0) \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) + 2\widetilde{U}^{k_0,l_0}(\theta;0) \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right] \right) \right| \geq C\end{aligned}$$

By Lemma 10.10 in [DLS06a], hypothesis **H3"** on (k_0, l_0) implies the previous condition and therefore the angle between the surfaces $S_\varepsilon(L_E^F)$ and $L_{E'}^F$ at the intersection can be bounded from below by $C\varepsilon$, for some suitable constant independent of ε.

For the particular case $|(k_0, l_0)| \sim \varepsilon^{-1/r}$ described in item (iii), we will check condition (1.143). Using the expression (1.157) for the dominant term $\mathcal{M}_\pm$ in F and proceeding as in the previous case, we have that the dominant term in the numerator of (1.141) is given by

$$\begin{aligned}&\left(\frac{\partial F}{\partial I}\right)^2 \frac{\partial^2 \mathcal{L}^*}{\partial \varphi^2} - \frac{\partial F}{\partial \varphi}\frac{\partial^2 F}{\partial I^2}\frac{\partial \mathcal{L}^*}{\partial \varphi} + \varepsilon \frac{\partial F}{\partial I}\frac{\partial \mathcal{L}^*}{\partial \varphi}\frac{\partial^2 \mathcal{L}^*}{\partial \varphi^2} \\ = \; &(k_0 I + l_0)^2 k_0^2 \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(\widetilde{\theta}, E) - \varepsilon^\gamma \widetilde{U}'^{k_0,l_0}(\theta,\varepsilon) k_0 k_0^2 \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(\widetilde{\theta}, E) \\ &+ \varepsilon (k_0 I + l_0) k_0^2 \frac{\partial \mathcal{L}^*}{\partial \varphi}(\widetilde{\theta}, E) \frac{\partial^2 \mathcal{L}^*}{\partial \varphi^2}(\widetilde{\theta}, E)\end{aligned}$$

Using that

$$|\nabla F| = (k_0 I + l_0)k_0 + \mathcal{O}(k_0\varepsilon),$$

and evaluating on $I = \lambda_E^{\pm}(\varphi, s; \varepsilon)$, we have that the condition (1.141) is satisfied provided that

$$\left| \frac{\pm 1}{2(E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} \left(2E \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right.\right.$$
$$-\varepsilon^\gamma \left[k_0 \widetilde{U}'^{k_0,l_0}(\theta;0) \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) + 2\widetilde{U}^{k_0,l_0}(\theta;0) \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right]$$
$$\left.\left. \pm \varepsilon k_0 \sqrt{2(E - \varepsilon^\gamma \widetilde{U}^{k_0,l_0}(\theta;0))} \frac{\partial \mathcal{L}^*}{\partial \widetilde{\theta}}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \frac{\partial^2 \mathcal{L}^*}{\partial \widetilde{\theta}^2}(-\frac{l_0}{k_0}, \frac{\theta}{k_0}) \right) \right| \geq C,$$

for some constant C. By hypothesis **H3"'** on (k_0, l_0) in Theorem 1.2.1 we know that the previous condition is satisfied for $\theta \in \mathcal{J} \subset \mathcal{J}^*_{-l_0/k_0}$. Consequently, the angle of intersection can be bounded again from below by $C\varepsilon$, for some suitable constant independent of ε. □

Remark 1.4.9. By Theorem 1.3.1 we know that the tori in a connected component of the big gaps region are given by the expression $I = \lambda_E^{\pm}(\varphi, s; \varepsilon)$, for $E = E_i$ and $-\varepsilon^\gamma \leq E_i \leq L^2$, with $\lambda_E^{\pm}$ given in (1.96). Moreover, they satisfy

$$|E_i - E_{i+1}| \leq \varepsilon^{\gamma/2+1+\eta} \leq \max(|E_i|^{1/2}, \varepsilon^{\gamma/2})$$

and they are $\mathcal{O}(\varepsilon^{1+\eta})$-closely spaced, in terms of the I variable.

Hence, we conclude that the image under the scattering map of a torus $\mathcal{T}_i$, $S_\varepsilon(\mathcal{T}_i)$ in the big gaps region, given by $I = \lambda_{E_i}^{\pm}(\varphi, s; \varepsilon)$, intersects transversally another torus $\mathcal{T}_{i+1}$ of this region given by $\bar{I} = \lambda_{E_{i+1}}^{\pm}(I, \varphi, s; \varepsilon)$, with $|E_{i+1} - E_i| = \mathcal{O}(\varepsilon^{\gamma/2+1+\eta})$ (equivalently $|I - \bar{I}| \leq \varepsilon^{1+\eta}$):

$$S_\varepsilon(\mathcal{T}_i) \pitchfork \mathcal{T}_{i+1}.$$

1.4.2 Proof of Proposition 1.4.1

The proof is just a combination of the results obtained in Section 1.4.1.

We start with a torus $\mathcal{T}_0$, which is $\mathcal{O}(\varepsilon^{1+\eta})$-close to the submanifold $I = I_-$. Assume that this torus belongs to the flat tori region with averaged energy E_0. Then, we apply Lemma 1.4.5 and Remark 1.4.6 and we get that $S_\varepsilon(\mathcal{T}_0)$ intersects transversally all primary tori with averaged energy in the mentioned interval $(E_0 - C\varepsilon, E_0 + C\varepsilon)$. We pick a primary KAM torus $\mathcal{T}_1$ provided by Theorem 1.3.1 with energy E_1 in the interval and we repeat the argument until we reach a big gaps region. Assuming that we have applied it K times, we have that the torus $\mathcal{T}_0$ has heteroclinic connections with all the tori whose energy lies in the interval $(E_0 - KC\varepsilon, E_0 + KC\varepsilon)$, or equivalently, in the interval $(I_- - K^*C\varepsilon, I_- + K^*C\varepsilon)$ in terms of action variables.

When the domain $(I_- - K^*C\varepsilon, I_- + K^*C\varepsilon) \times \mathbb{T}^2$ for which the torus $\mathcal{T}_0$ has a heteroclinic connection overlaps with a big gaps region $[-l_0/k_0 - L_{k_0}, -l_0/k_0 + L_{k_0}] \times \mathbb{T}^2$ we use Lemma 1.4.7 and Remark 1.4.9 to show that we can cross the gap created by the resonance $-l_0/k_0$ just connecting either a primary KAM torus with a secondary one and again with a primary one or two primary KAM tori. Hence, we can construct a piece of chain that starts in $\mathcal{T}_0$ and reaches all the way to $\mathcal{T}_i$, where $\mathcal{T}_i$ is a primary KAM torus whose equation is $I = -l_0/k_0 + L_{k_0} + \mathcal{O}(\varepsilon)$ and is contained again in the flat tori region.

Therefore, we can keep constructing a transition chain just repeating the procedure stated before for the primary KAM torus $\mathcal{T}_i$ until we reach $\mathcal{T}_{N(\varepsilon)}$.

The case when $\mathcal{T}_0$ belongs to a big gaps region is analogous. □

1.5 Example

Consider the Hamiltonian

$$H_\varepsilon(p,q,I,\varphi,t) = \pm\left(\frac{p^2}{2}+\cos q-1\right)+\frac{I^2}{2}+\varepsilon\cos q\; g(\varphi,t), \tag{1.159}$$

which is a generalization of the famous example introduce by V.I. Arnol'd in [Arn64]. The function g is chosen as a periodic function with an *infinite number of harmonics* in the angles (φ,t), say

$$g(\varphi,t)=\sum_{(k,l)\in\mathbb{N}^2} a_{k,l}\cos(k\varphi+lt), \tag{1.160}$$

with $a_{k,l}=\rho^k r^l$ and $0<\rho,r<1$ real numbers to be chosen small enough.

The Hamiltonian of one degree of freedom $P_\pm(p,q)=\pm\,(p^2/2+\cos q-1)$ is the standard pendulum when we choose the $+$ sign, and its separatrix for positive p is given by

$$q_0(t)=4\arctan e^{\pm t},\quad p_0(t)=2/\cosh t.$$

An important feature of the Hamiltonian (1.159) is that the 3-dimensional NHIM

$$\tilde\Lambda=\{(0,0,I,\varphi,s):(I,\varphi,s)\in\mathbb{R}\times\mathbb{T}^2\}$$

is preserved without any deformation for any ε: $p=q=0\Rightarrow\dot p=\dot q=0$. However, in contrast with the example in [Arn64], the perturbation does not vanish on $\tilde\Lambda$. Indeed, the Hamiltonian (1.159) restricted to $\tilde\Lambda$ takes the form $I^2/2+\varepsilon g(\varphi,t)$. Hence, the 2-dimensional whiskered tori

$$\mathcal{T}^0_I=\{(0,0,I,\varphi,s):(\varphi,s)\in\mathbb{T}^2\}$$

are not preserved for $\varepsilon\neq 0$, and resonances (1.48) take place at $I=-l/k$ for each $(k,l)\in\mathbb{N}^2$, $\gcd(k,l)=1$. Therefore, we have a dense set of gaps of size $\mathcal{O}(\varepsilon^{1/2}\sqrt{a_{k,l}})$ and, among them the ones such that $\sqrt{a_{k,l}}\succcurlyeq\varepsilon^{1/2}$ give rise to resonances with big gaps and the example (1.159) presents the large gap problem for $I<0$.

Hence, for any finite range of I, $[I_-,I_+]\subset\mathbb{R}^-$ we will prove the existence of diffusing orbits.

The Melnikov potential (1.9) of the Hamiltonian (1.159) is given by

$$\mathcal{L}(I,\varphi,s)=\sum_{(k,l)\in\mathbb{N}^2} A_{k,l}(I)\cos(k\varphi+ls),$$

with

$$A_{k,l}(I)=2\pi\frac{(kI+l)}{\sinh\frac{\pi}{2}(kI+l)}a_{k,l}\quad\text{for } I\neq -l/k,\quad A_{k,l}(I)(-l/k)=4a_{k,l}. \tag{1.161}$$

Next, we will see that for $0 < \rho < r \ll 1$ we can find open sets of $(I, \varphi, s) \in [I_-, I_+] \times \mathbb{T}^2$, such that the function $\tau \in \mathbb{R} \mapsto \mathcal{L}(I, \varphi - I\tau, s - \tau)$ has non-degenerate critical points at $\tau = \tau^*(I, \varphi, s)$ which verify the hypothesis **H2'**.

Recall that hypothesis **H2'** deals with the existence of transverse intersections of the stable and unstable manifolds of $\widetilde{\Lambda}_\varepsilon$. Hence, the non-degenerate critical points of the function $\tau \mapsto \mathcal{L}(I, \varphi - I\tau, s - \tau)$ give rise to transverse intersections.

In order to check hypothesis **H2'**, we will use the results in the example given in Chapter 13 of [DLS06a] by means of the following argument. Assuming that ρ, r are small enough, the function $g(\varphi, s)$ is well approximated by its truncated first order trigonometric polynomial $g^{[\leq 1]}(\varphi, s) = 1 + \rho \cos \varphi + r \cos s$. More precisely,

$$\begin{aligned} g(\varphi, s) &= 1 + \rho \cos \varphi + r \cos s + \mathcal{O}_2(\rho, r) \\ &:= g^{[\leq 1]}(\varphi, s) + g^{[>1]}(\varphi, s). \end{aligned}$$

Hence, as long as $0 < \rho, r \ll 1$, if hypothesis **H2'** is verified for the trigonometric polynomial $g^{[\leq 1]}(\varphi, s)$, it will be also verified for the perturbation $g(\varphi, s)$.

Notice that the Fourier coefficients $A_{k,l}(I)$ are nothing else but the Fourier coefficients $a_{k,l}$ multiplied by a certain function depending on I that decreases exponentially as $|I|$ goes to infinity. Hence, arguing as we did for the perturbation g, we approximate the function $\mathcal{L}(I, \varphi, s)$ by its first order trigonometric polynomial $\mathcal{L}^{[\leq 1]}(I, \varphi, s) = A_{0,0} + A_{1,0}(I) \cos \varphi + A_{0,1} \cos s$, that is

$$\begin{aligned} \mathcal{L}(I, \varphi, s) &= A_{0,0} + A_{1,0}(I) \cos \varphi + A_{0,1} \cos s + \mathcal{O}_2(\rho, r) \\ &:= \mathcal{L}^{[\leq 1]}(I, \varphi, s) + \mathcal{L}^{[>1]}(I, \varphi, s). \end{aligned} \tag{1.162}$$

Recall that we are looking for non-degenerate critical points of

$$\mathcal{L}(\tau) := \mathcal{L}(I, \varphi - I\tau, s - \tau) = \sum_{(k,l) \in \mathbb{N}^2} A_{k,l}(I) \cos(k\varphi + ls - \tau(Ik + l)), \tag{1.163}$$

with $A_{k,l}(I)$ as in (1.161).

Using that the Melnikov function $\mathcal{L}$ is well approximated by $\mathcal{L}^{[\leq 1]}$, fixed (I, φ, s), we only need to study the evolution of $\mathcal{L}^{[\leq 1]}$ along the straight lines

$$R : \tau \in \mathbb{R} \mapsto (\varphi - I\tau, s - \tau) \in \mathbb{T}^2 \tag{1.164}$$

on the torus.

This study has already been performed in the example in Chapter 13 in [DLS06a], where the reader can find more details. We just mention that since $\rho, r > 0$, for any fixed I, the function $(\varphi, s) \mapsto \mathcal{L}^{[\leq 1]}(I, \varphi, s)$ possesses exactly four non-degenerate critical points: a maximum at $(0, 0)$, a minimum at (π, π) and two saddles at $(0, \pi)$ and $(\pi, 0)$ (see Figure 1.2). Around the two extremum points, its level curves are closed (and indeed convex) curves which fill out a basin ending at the level curve of one of the saddle points.

Therefore, any straight line (1.164) that enters into some extremum basin is tangent to one of the convex closed level curves, giving rise to a non-degenerate extremum of $\tau \in \mathbb{R} \mapsto \mathcal{L}^{[\leq 1]}(I, \varphi - Is, s - \tau)$. So, degenerate extrema of $\tau \in \mathbb{R} \mapsto \mathcal{L}^{[\leq 1]}(I, \varphi - Is, s - \tau)$

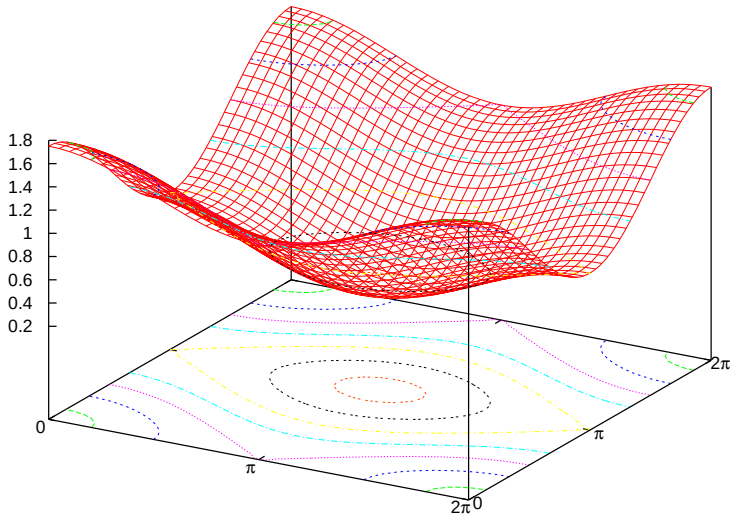

Figure 1.2: Graph and level curves of the Melnikov potential $\mathcal{L}^{[\leq 1]}(I,\varphi,s)$ with $\rho = 1/16$, $r = 1/8$ and $I = -l/k$

can only exist for straight lines that never enter inside such extremum basins. It is clear that this never happens for irrational values of I because it implies a dense straight line (and infinite non-degenerate extrema for $\tau \in \mathbb{R} \mapsto \mathcal{L}^{[\leq 1]}(I, \varphi - Is, s - \tau)$). On the other hand, straight lines with rational slopes enter inside both extremum basins at least twice, except for the slopes $0, \pm 1, \pm\infty$, which correspond to the rational values $I = 0, -1/2, -1$. In the case of slopes $0, \pm\infty$, straight lines enter at least once in one of the extremum basins. In the case of slopes ± 1, it might happen that the straight line (1.164) passes through both saddle points and never enters inside an extremum basin. However, we have assumed $\rho < r$ so that the values of Melnikov potential are different at the saddle points and therefore the function $\tau \in \mathbb{R} \mapsto \mathcal{L}^{[\leq 1]}(I, \varphi - Is, s - \tau)$ has non-degenerate extrema for any $I \in [I_-, I_+] \subset \mathbb{R}^-$.

When we take into account $\mathcal{L}^{[>1]}$ in the Melnikov potential $\mathcal{L}$ in (1.162), it is clear that in the compact subset $[I_-, I_+] \times \mathbb{T}^2$, as long as $0 < \rho, r \ll 1$, the function $\tau \in \mathbb{R} \mapsto \mathcal{L}(I, \varphi - I\tau, s - \tau)$ has non-degenerate extrema, and for every I we can find a smooth function $\tau = \tau^*(I, \varphi, s)$ defined in an open set of $(\varphi, s) \in \mathbb{T}^2$.

Moreover, since $\mathcal{L}$ is periodic with respect to (φ, s) and non-constant with non-degenerate extrema along any straight line, $\partial_\varphi \mathcal{L}^*$, where $\mathcal{L}^*$ is given in (1.11) is also periodic and non-constant and indeed changes sign. Therefore, for every I, there exists a nonempty set $\mathcal{J}_I$ where $\partial_\varphi \mathcal{L}^* > 0$ (and a nonempty set $\mathcal{J}_I^-$ where $\partial_\varphi \mathcal{L}^* < 0$), so hypothesis **H2"** is fulfilled. Indeed the set of points where $\partial_\varphi \mathcal{L}^*$ vanishes is a discrete set.

Conditions **H3'**, **H3"** and **H3"'** can also be checked in the example (1.160) at the resonances $I = -l_0/k_0$.

If we consider $I = -l_0/k_0$ for any $(k_0, l_0) \in \mathbb{N}^2$, $k_0 \neq 0$ and $\gcd(k_0, l_0) = 1$, the function

U^{k_0,l_0} in hypothesis **H3** on (k_0, l_0) has the following expression

$$U^{k_0,l_0}(\theta) = \sum_{t=1}^{M} a_{tk_0,tl_0} \cos(t\theta) = a_{k_0,l_0} \cos(\theta) + \mathcal{O}_2(\rho^{k_0}, r^{l_0}), \tag{1.165}$$

where $\theta = k_0\varphi + l_0 s$.

Therefore, $\theta_1 = 0$ and $\theta_2 = \pi$ are the unique critical points for the function $U^{k_0,l_0}(\theta)$. Hence hypothesis **H3'** on (k_0, l_0) is clearly verified.

Next, for $I = -l_0/k_0$ we want to check hypothesis **H3"** on (k_0, l_0). This condition requires to show that the function f in (3.22) is not constant. To that end, we will consider two values of θ and we will show that their images for this function are different. For instance, notice that the function f in (3.22) takes the same values as U^{k_0,l_0} evaluated on its critical points θ_1 and θ_2 as long as $\frac{\partial^2 \mathcal{L}^*}{\partial \varphi^2}(I, \theta_i/k_0) \neq 0$, for $i = 1, 2$. Hence, hypothesis **H3"** on (k_0, l_0) is clearly satisfied if the function U^{k_0,l_0} has two extrema θ_i taking different values which satisfy $\frac{\partial^2 \mathcal{L}^*}{\partial \varphi^2}(I, \theta_i/k_0) \neq 0$, which is the case.

Similarly, we can check hypothesis **H3"** on (k_0, l_0). In this case we need to show that the determinant (1.151) given in Remark 1.4.9 does not vanish. It is clearly non-zero if we choose, for the two first columns, the two critical points θ_1 and θ_2 discussed above, and for the third column $\theta_3 \neq 0, \pi$, such that $\frac{\partial^2 \mathcal{L}^*}{\partial \varphi^2}(-l_0/k_0, \theta_3/k_0) = 0$, but otherwise $U'^{k_0,l_0}(\theta_3) \neq 0$ and $\frac{\partial \mathcal{L}^*}{\partial \varphi}(-l_0/k_0, \theta_3/k_0) \neq 0$.

Hence, we apply Theorem (1.2.1) and we conclude that

Proposition 1.5.1. *Given the Hamiltonian* (1.159) *with* g *as in* (1.160), $0 < \rho < r \ll 1$ *and* $[I_-, I_+] \subset \mathbb{R}^-$, *for* $|\varepsilon| \leq \varepsilon^*(\rho, r)$ *there exist orbits following the mechanism described in the previous sections and such that* $I(0) \leq I_-$, $I(T) \geq I_+$, *for any* $T > 0$.

1.6 Appendix

1.6.1 Appendix: Double Fouries Series

Proposition 1.6.1. *Let* f *be a* $\mathcal{C}^r$ *function with respect to* $(J, \varphi, s, \varepsilon)$, $r \geq 1$ *and* 2π*-periodic with respect to* (φ, s). *Then its Fourier coefficients* $f_{k,l}(J, \varepsilon)$, $(k, l) \in \mathbb{Z}^2$, *satisfy, for* $\ell = 0, \ldots, r$

$$|f_{k,l}|_{\mathcal{C}^\ell} \leq C \frac{|f|_{\mathcal{C}^r}}{|(k,l)|^{r-\ell}}, \tag{1.166}$$

where C *is a constant that depends only on* r *and* ℓ *and* $|(k, l)| = \max(|k|, |l|)$.

Proof. From the expression for the Fourier coefficients of a function f

$$f_{k,l}(J; \varepsilon) = \frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} f(J, \varphi, s; \varepsilon) e^{\mathrm{i}(k\varphi + ls)} \mathrm{d}\varphi \mathrm{d}s,$$

taking into account that f is $\mathcal{C}^r$ in the variables (φ, s), we can integrate $r = n + m$ times by parts (n times with respect to φ and m times with respect to s) and express the Fourier coefficient $f_{k,l}(J, \varepsilon)$, with $(k, l) \neq (0, 0)$ in the form

$$f_{k,l}(J; \varepsilon) = (-1)^r \frac{1}{(2\pi)^2} \frac{1}{(\mathrm{i}k)^n (\mathrm{i}l)^m} \int_{\mathbb{T}^2} \frac{\partial^r f(J, \varphi, s; \varepsilon)}{\partial \varphi^n \partial s^m} e^{\mathrm{i}(k\varphi + ls)} \mathrm{d}\varphi \mathrm{d}s,$$

so that,

$$|f_{k,l}|_{\mathcal{C}^0} \leq \frac{1}{|k|^n|l|^m}\left|\frac{\partial^r f}{\partial\varphi^n\partial s^m}\right|_{\mathcal{C}^0} \leq \frac{n!m!|f|_{0,r}}{|k|^n|l|^m},$$

for any $0 \leq n, m \leq r$ such that $n+m=r$, where $|\cdot|_{\mathcal{C}^\ell}$ is the standard $\mathcal{C}^\ell$ norm defined in (1.2) and $|\cdot|_{\ell_1,\ell_2}$ is the seminorm defined in (1.3). Therefore,

$$|f_{k,l}|_{\mathcal{C}^0} \leq \frac{r!|f|_{0,r}}{|(k,l)|^r} \leq \frac{r!|f|_{\mathcal{C}^r}}{|(k,l)|^r}.$$

where $|(k,l)| = \max(|k|,|l|)$.

Now, taking into account that $D^\ell f_{k,l}(J;\varepsilon)$ is the Fourier coefficient of the function $\frac{\partial^\ell f(J,\varphi,s;\varepsilon)}{\partial J^\ell}$, which is a $\mathcal{C}^{r-\ell}$ function, and using the same argument as before we have that

$$|D^\ell f_{k,l}|_{\mathcal{C}^0} \leq \frac{\ell!(r-\ell)!|f|_{\ell,r-\ell}}{|(k,l)|^{r-\ell}} \leq \frac{\ell!(r-\ell)!|f|_{\mathcal{C}^r}}{|(k,l)|^{r-\ell}}.$$

From the definition of $|\cdot|_{\mathcal{C}^\ell}$ norm in (1.2) we have the estimate

$$|f_{k,l}|_{\mathcal{C}^\ell} = \sum_{i=0}^{\ell}\frac{|D^i f_{k,l}|_{\mathcal{C}^0}}{i!} \leq \sum_{i=0}^{\ell}\frac{(r-i)!|f|_{\mathcal{C}^r}}{|(k,l)|^{r-i}} \leq C\frac{|f|_{\mathcal{C}^r}}{|(k,l)|^{r-\ell}},$$

where C is a constant that only depends on ℓ and r, $C = r! + (r-1)! + \ldots + (r-\ell)!$, as we wanted to see. □

We consider the truncation of its Fourier series at order M in the following way:

$$f(J,\varphi,s;\varepsilon) = f^{[\leq M]}(J,\varphi,s;\varepsilon) + f^{[>M]}(J,\varphi,s;\varepsilon),$$

where

$$f^{[\leq M]}(J,\varphi,s;\varepsilon) = \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\leq M}} f_{k,l}(J;\varepsilon)e^{\mathrm{i}(k\varphi+ls)},$$

and

$$f^{[>M]}(J,\varphi,s;\varepsilon) = \sum_{\substack{(k,l)\in\mathbb{Z}^2\\ |k|+|l|> M}} f_{k,l}(J;\varepsilon)e^{\mathrm{i}(k\varphi+ls)}.$$

Proposition 1.6.2. *Let f be of class $\mathcal{C}^r$ with respect to $(J,\varphi,s,\varepsilon)$, $r \geq 1$ and 2π-periodic with respect to (φ,s). The M-th order remainder $f^{[>M]}$ of the Fourier series of f is bounded in the standard $\mathcal{C}^\ell$ norm, for $\ell = 0,\ldots,r-3$ by*

$$\left|f^{[>M]}\right|_{\mathcal{C}^\ell} \leq C\frac{|f|_{\mathcal{C}^r}}{M^{r-(\ell+2)}}, \tag{1.167}$$

where C is a constant that depends only on r and ℓ.

Proof. The proof is very simple and follows from the estimate (1.166) for the Fourier coefficients of a $\mathcal{C}^r$ function obtained in the previous propositon. More precisely,

$$\begin{aligned}
\left|f^{[>M]}\right|_{\mathcal{C}^\ell} &\leq \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|>M}} |f_{k,l}|_{\mathcal{C}^\ell} \\
&\leq C \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|>M}} \frac{|f|_{\mathcal{C}^r}}{|(k,l)|^{r-\ell}} \\
&\leq C \sum_{t=M+1}^{\infty} 4t \frac{|f|_{\mathcal{C}^r}}{t^{r-\ell}} \\
&\leq 4C\,|f|_{\mathcal{C}^r} \int_M^\infty t^{\ell-r+1}\mathrm{d}t \\
&= 4\frac{C}{r-\ell-2}\,|f|_{\mathcal{C}^r}\, M^{\ell-r+2},
\end{aligned}$$

where C is a constant that depends only on r and ℓ. □

1.6.2 Appendix: weighted norms

We consider functions $u \in \tau_M(\mathcal{I}\times\mathbb{T}^2)$, where $\mathcal{I}\subset\mathbb{R}$, introduced in (1.28), and we can consider the different types of norms introduced in this memoir: the standard $\mathcal{C}^r$ norm introduced in (1.2), the Fourier norm introduced in (1.29) and the Fourier norm with a weight introduced in (1.30).

The equivalence relations between all these norms are given in the following Lemmas:

Lemma 1.6.3. *The norms $|\cdot|_{\mathcal{C}^\ell}$ and $\|\cdot\|_{\mathcal{C}^\ell}$ defined in (1.2) and (1.29), respectively, are equivalent and satisfy the following equivalence relation for $u\in\tau_M(\mathcal{I}\times\mathbb{T}^2)$ and $0<L\leq 1$,*

$$L^\ell |u|_{\mathcal{C}^\ell} \leq \|u\|_{\mathcal{C}^\ell,L} \leq CM^2|u|_{\mathcal{C}^\ell}$$

where C is a constant depending on ℓ.

Proof. The first inequality is obvious using that $L\leq 1$. For the second one, using again that $L\leq 1$ we have

$$|u_{k,l}|_{\mathcal{C}^n,L} = \sum_{i=0}^n L^i \frac{|D^i u_{k,l}|_{\mathcal{C}^0}}{i!} \leq \sum_{i=0}^n \frac{|D^i u_{k,l}|_{\mathcal{C}^0}}{i!} = |u_{k,l}|_{\mathcal{C}^n},$$

for $0\leq n\leq \ell$. Therefore, the result follows directly from the estimate (1.166) for the $\mathcal{C}^\ell$

norm of the Fourier coefficients of a $\mathcal{C}^r$ function u, for $\ell = 0, \dots, r$. More precisely,

$$\begin{aligned}
\|u\|_{\mathcal{C}^\ell,L} &= \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\le M}} |u_{k,l}|_{\mathcal{C}^n,L}\,|(k,l)|^{m-n} \\
&\le \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\le M}} |u_{k,l}|_{\mathcal{C}^n}\,|(k,l)|^{m-n} \\
&\le \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\le M}} \widetilde{C}\frac{|u|_{\mathcal{C}^\ell}}{|(k,l)|^{\ell-n}}|(k,l)|^{m-n} \\
&\le \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\le M}} \widetilde{C}\,|u|_{\mathcal{C}^\ell} \\
&\le CM^2\,|u|_{\mathcal{C}^\ell}
\end{aligned}$$

as we wanted to prove. □

Lemma 1.6.4. *For the seminorm $|\cdot|_{j,\ell-j}$ defined in (1.2), one has that for all $0 \le j \le \ell$,*

$$L^j\,|u|_{j,\ell-j} \le \|u\|_{\mathcal{C}^\ell,L} \tag{1.168}$$

Proof. Again, It follows directly from the fact that $L < 1$ and therefore,

$$L^j|u_{k,l}|_{\mathcal{C}^n} \le \sum_{i=0}^{n} L^i\frac{|D^i u_{k,l}|_{\mathcal{C}^0}}{i!} = |u_{k,l}|_{\mathcal{C}^n,L}.$$

for $0 \le n \le j$. □

Lemma 1.6.5. *For $0 < L \le 1$, and $0 \le \ell \le r$ we have that for any $u \in \tau_M(\mathcal{I}\times\mathbb{T}^2)$ and $v \in \tau_N(\mathcal{I}\times\mathbb{T}^2)$*

$$\|uv\|_{\mathcal{C}^\ell,L} \le \|u\|_{\mathcal{C}^\ell,L}\,\|v\|_{\mathcal{C}^\ell,L}\,. \tag{1.169}$$

Proof. Let us define

$$\|u\|_{n,m} = \sum_{\substack{(k,l)\in\mathbb{Z}^2,\\ |k|+|l|\le M}} |u_{k,l}|_{\mathcal{C}^n,L}|(k,l)|^{m-n},$$

then,

$$\|u\|_{\mathcal{C}^\ell,L} = \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell\|u\|_{n,m}. \tag{1.170}$$

The $\alpha - th$ Fourier coefficient of uv, where $\alpha \in \mathbb{Z}^2$, is

$$(uv)_\alpha = \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\le N\\ |\alpha-\beta|\le M}} u_{\alpha-\beta}v_\beta.$$

Using the Leibniz rule for derivatives we have

$$\begin{aligned}
|(uv)_\alpha|_{\mathcal{C}^n,L} &= \sum_{i=0}^{n} \frac{1}{i!} L^i |D^i (uv)_\alpha|_{\mathcal{C}^0} \\
&\leq \sum_{i=0}^{n} \frac{1}{i!} \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\leq N \\ |\alpha-\beta|\leq M}} L^i |D^i u_{\alpha-\beta} v_\beta|_{\mathcal{C}^0} \\
&\leq \sum_{i=0}^{n} \frac{1}{i!} \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\leq N \\ |\alpha-\beta|\leq M}} \sum_{j=0}^{i} \left(\begin{array}{c} i \\ j \end{array} \right) L^{i-j} |D^{i-j} u_{\alpha-\beta}|_{\mathcal{C}^0} L^j |D^j v_\beta|_{\mathcal{C}^0} \\
&= \sum_{i=0}^{n} \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\leq N \\ |\alpha-\beta|\leq M}} \sum_{j=0}^{i} L^{i-j} \frac{|D^{i-j} u_{\alpha-\beta}|_{\mathcal{C}^0}}{(i-j)!} L^j \frac{|D^j v_\beta|_{\mathcal{C}^0}}{j!} \\
&= \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\leq N \\ |\alpha-\beta|\leq M}} \sum_{i=0}^{n} \sum_{j=0}^{i} L^{i-j} \frac{|D^{i-j} u_{\alpha-\beta}|_{\mathcal{C}^0}}{(i-j)!} L^j \frac{|D^j v_\beta|_{\mathcal{C}^0}}{j!} \\
&\leq \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\leq N \\ |\alpha-\beta|\leq M}} |u_{\alpha-\beta}|_{\mathcal{C}^n,L} \, |v_\beta|_{\mathcal{C}^n,L} \, .
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
|\alpha|^{m-n} &\leq (|\alpha-\beta|+|\beta|)^{m-n} = \sum_{i=0}^{m-n} \left(\begin{array}{c} m-n \\ i \end{array} \right) |\alpha-\beta|^i |\beta|^{m-n-i} \\
&\leq \max\left(|\alpha|^{m-n}, \sum_{i=0}^{m-n} \left(\begin{array}{c} m-n \\ i \end{array} \right) |\alpha-\beta|^{m-n} |\beta|^{m-n} \right) \\
&= \max(|\alpha|^{m-n}, 2^{m-n} |\alpha-\beta|^{m-n} |\beta|^{m-n}).
\end{aligned}$$

Hence, using these two inequalities, we have that

$$\begin{aligned}
\|uv\|_{n,m} &= \sum_{\substack{\alpha\in\mathbb{Z}^2,\\ |\alpha|\le M+N}} |(uv)_\alpha|_{\mathcal{C}^n,L}|\alpha|^{m-n} \\
&\le \sum_{\substack{\alpha\in\mathbb{Z}^2,\\ |\alpha|\le M+N}} \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\le N\\ |\alpha-\beta|\le M}} |u_{\alpha-\beta}|_{\mathcal{C}^n,L}\,|v_\beta|_{\mathcal{C}^n,L}\,|\alpha|^{m-n} \\
&\le \sum_{\substack{\alpha\in\mathbb{Z}^2,\\ |\alpha|\le M+N}} |u_0|_{\mathcal{C}^n,L}\,|v_\alpha|_{\mathcal{C}^n,L}\,|\alpha|^{m-n} + |u_\alpha|_{\mathcal{C}^n,L}\,|\alpha|^{m-n}\,|v_0|_{\mathcal{C}^n,L} \\
&\quad + \sum_{\substack{\beta\in\mathbb{Z}^2,|\beta|\le N\\ |\alpha-\beta|\le M}} |u_{\alpha-\beta}|_{\mathcal{C}^n,L}\,|v_\beta|_{\mathcal{C}^n,L}\,2^{m-n}|\alpha-\beta|^{m-n}|\beta|^{m-n} \\
&\le 2^{m-n}\sum_{\substack{\alpha\in\mathbb{Z}^2,\\ |\alpha|\le M}} |u_\alpha|_{\mathcal{C}^n,L}\,|\alpha|^{m-n} \sum_{\substack{\beta\in\mathbb{Z}^2\\ |\beta|\le M}} |v_\beta|_{\mathcal{C}^n,L}\,|\beta|^{m-n} \\
&= 2^{m-n}\|u\|_{n,m}\|v\|_{n,m}.
\end{aligned}$$

Going back to the definition of $\|uv\|_{\mathcal{C}^\ell,L}$ in (1.170), we have

$$\begin{aligned}
\|uv\|_{\mathcal{C}^\ell,L} &= \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell\|uv\|_{n,m} \\
&\le \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell 2^{m-n}\|u\|_{n,m}\|v\|_{n,m} \\
&\le \sum_{m=0}^{\ell}\sum_{n=0}^{m} 2^\ell 2^\ell\|u\|_{n,m}\|v\|_{n,m} \\
&\le \|u\|_{\mathcal{C}^\ell,L}\,\|v\|_{\mathcal{C}^\ell,L}\,,
\end{aligned}$$

as claimed. □

1.6.3 Appendix: Faa-di Bruno formula

Let g be a $\mathcal{C}^s(U,V)$ function, with $U\subset\mathbb{R}$ and $g(U)\subset W\subset\mathbb{R}$ and f be a $\mathcal{C}^r(W,\mathbb{R})$ function with $r,s>0$. Then $f\circ g$ is a $\mathcal{C}^t(U,\mathbb{R})$ function, where $t=\min(r,s)$. By a repeated application of the chain rule, one gets

$$D^\ell(f\circ g)(x) = \sum_{k=1}^{\ell}\sum_{j_1+\cdots+j_k=\ell} c_{k,j_1,\cdots,j_k} D^k f(g(x))D^{j_1}g(x)\cdots D^{j_k}g(x), \tag{1.171}$$

for $\ell=1,\ldots,t$, where $c_{k,j_1,\cdots,j_k}$ are combinatorial coefficients. The formula (1.171) is called **Faa-di Bruno formula** (see [LO99]).

From equation (1.171), it is easy to see that there exists a constant C_t depending on t such that

$$|f\circ g|_{\mathcal{C}^t} \le C_t\,|f|_{\mathcal{C}^t}\,|g|^t_{\mathcal{C}^t}\,. \tag{1.172}$$

Since we are interested in multi-valued functions, we introduce now a generalized bound. Thus, let us consider a function g in $\mathcal{C}^s(U,V)$, with $U \subset \mathbb{R}^n$ and $g(U) \subset W \subset \mathbb{R}^m$ and a function f in $\mathcal{C}^r(W,\mathbb{R})$ with $r,s>0$. As before, $f \circ g$ is a $\mathcal{C}^t(U,\mathbb{R})$ function, where $t=\min(r,s)$. Similarly, we can get an expression for the derivatives of $f \circ g$, such that for $\ell = 1,\dots,t$,

$$|f\circ g|_{\mathcal{C}^\ell} \leq C_\ell \sum_{k=1}^{\ell} \sum_{j_1+\cdots+j_k=\ell} |f|_{\mathcal{C}^k} |g|_{\mathcal{C}^{j_1}} \cdots |g|_{\mathcal{C}^{j_k}} , \tag{1.173}$$

for $\ell = 0,\dots,t$, where C_ℓ is a constant depending on ℓ. As before, we can consider the following less precise but more compact bound,

$$|f\circ g|_{\mathcal{C}^\ell} \leq C_\ell \, |f|_{\mathcal{C}^\ell} \, |g|^\ell_{\mathcal{C}^\ell} , \tag{1.174}$$

for $\ell = 1,\dots,t$, where C_ℓ is a constant depending on ℓ.

For some other results related to this, we refer the reader to [LO99].

In some cases, it will be more convenient to use another estimate for the $|\cdot|_{\mathcal{C}^\ell}$ norm instead of the one obtained in (1.174). In formula (1.173) we can separate the term corresponding to $k=1$ in the following way

$$|f\circ g|_{\mathcal{C}^\ell} \leq C_\ell \left(|f|_{\mathcal{C}^1} \, |g|_{\mathcal{C}^\ell} + \sum_{k=2}^{\ell} \sum_{j_1+\cdots+j_k=\ell} |f|_{\mathcal{C}^k} \, |g|_{\mathcal{C}^{j_1}} \cdots |g|_{\mathcal{C}^{j_k}} \right),$$

for $\ell = 1,\dots,t$ and we can bound it in the $|\cdot|_{\mathcal{C}^\ell}$ norm

$$|f\circ g|_{\mathcal{C}^\ell} \leq C_\ell (|f|_{\mathcal{C}^1} \, |g|_{\mathcal{C}^\ell} + |f|_{\mathcal{C}^\ell} \, |g|^\ell_{\mathcal{C}^{\ell-1}}), \tag{1.175}$$

for $\ell = 1,\dots,t$, where C_ℓ is a constant depending on ℓ.

Chapter 2

Fast numerical algorithms to compute invariant tori and the associated whiskers in Hamiltonian systems

2.1 Introduction

The goal of this chapter is to describe efficient algorithms to compute invariant manifolds in Hamiltonian systems. The invariant manifolds we are considering are invariant tori and the associated whiskers, which include the usual stable and unstable invariant manifolds.

For us, an invariant torus will mean a torus with a quasi-periodic dynamics on it and a dimension equal to the number of independent frequencies, which we will assume *Diophantine*. Invariant tori have been an important object of study because they provide landmarks that organize the long term behavior of the system. There are several variants of these tori; in this thesis we will consider *maximal tori* and *whiskered tori*.

Tori of the maximal dimension are quasi-periodic solutions of n frequencies in systems with n-degrees of freedom. Since they are barriers in the phase space for $n > 2$, their importance lies on the fact that they provide long term stability. In contrast, whiskered tori are tori with $n - \ell$ independent frequencies in systems with n-degrees of freedom. They have associated smooth invariant manifolds and, among them, the most well known are probably, due to their importance, the stable and unstable manifolds. Stable and unstable manifolds are tangent to the spaces invariant under the linearization spanned by ℓ-independent directions which contract exponentially fast in the future and in the past, respectively. Moreover, the persistence of these manifolds under a perturbation has been studied classically using the theory of normally hyperbolic invariant manifolds (NHIM) (see [Fen72, Fen74, Fen77, HPS77]. Nevertheless, the (un)stable invariant manifolds are not the only ones attached to the invariant tori. One can study, for instance the *slow* invariant manifolds associated to the less contractive (expansive) directions or several combination of spaces satisfying non resonance conditions.

The whiskered tori and their invariant manifolds organize the long term behavior and provide routes which lead to long scale instability. Indeed, the well known paper [Arn64]

showed that, in some particular example, one can use the heteroclinic intersections among these manifolds to produce large scale motions. In [Arn63a] this is conjectured as a generic mechanism. The transitions between different kinds of whiskered invariant tori have been the basis of many of the theoretical models of Arnol'd diffusion [DLS06a, DH08] (see also the first chapter of this memoir).

One of the applied areas where these objects have been studied for a long time is in astrodynamics. Thus, the monographs [Sim99, GJSM01b, GJSM01a] show that, when one computes these invariant objects in realistic models of the Solar System, one can obtain orbits of practical importance for the design of space missions.

The numerical method we use, closely related to the theoretical proof, is based on the parameterization method [CFL03a, CFL03b], which consists of deriving a functional equation satisfied by the parameterization of the invariant manifold and then implement a Newton method for it. The parameterization method is well suited for the numerical implementation because it uses functions of the same number of variables as the dimension of the objects that we want to compute.

The main goal of this chapter to design very efficient numerical algorithms to perform a Newton step. What we mean by efficient is that, if the functional equation is discretized using N Fourier coefficients, one Newton step requires only storage of $O(N)$ and takes only $O(N \ln N)$ operations. Note that a straightforward implementation of the Newton method (which is called the *large matrix method*) requires to store an $N \times N$ matrix and solve the linear equation, which requires $O(N^3)$ operations. We include a comparison with the standard Newton method in Section 2.4.1.

For the case of quasi-periodic systems, algorithms with the same features were discussed in [HL06b, HL06a, HL07] and, for some Lagrangian systems (some of which do not admit a Hamiltonian interpretation) in [CL08].

One key ingredient in this calculation—and indeed in all subsequent calculations—is the phenomenon of "automatic reducibility." This phenomenon, which also lies at the basis of the rigorous proofs [LGJV05, Lla01b, FLS07], shows that the preservation of the symplectic structure leads to the fact that the Newton equations can be reduced—by explicit changes of variables—to upper triangular difference equations with diagonal constant coefficients. These equations can be solved very efficiently in Fourier coefficients.

The changes of variables are algebraic expressions involving derivatives of the parameterization. We note that derivatives can be computed fast in a Fourier representation whereas the algebraic expressions can be computed fast in the real space representation. Therefore, the algorithm to produce a Newton step consists of a small number of steps, each of which is diagonal either in real space or in Fourier space. Of course, the FFT algorithm allows us to switch from real space to Fourier space in $O(N \log N)$ computations.

We also note that the algorithms mirror very closely the proofs of the KAM theorem. In [LGJV05] and [FLS07] we can find proofs that the algorithms considered here converge to a true solution of the problem provided that the initial approximation solves the invariance equation with good enough accuracy and satisfies some appropriate non-degeneracy conditions. In numerical analysis this is typically known as *a posteriori estimates*.

It is important to remark that the algorithms that we will present can compute in unified way both primary and secondary tori. We recall here that *secondary tori* are invariant tori which are contractible to a torus of lower dimension, whereas this is not

the case for primary tori. We note that the tori which appear in integrable systems in action-angle variables are always primary. In quasi-integrable systems, the tori which appear through Lindstedt series or other perturbed expansions starting from those of the integrable system are always primary. Secondary tori, however, are generated by resonances. They are the very prominent "islands" which have been studied in resonances. As an example of the importance of secondary tori we will mention that in the recent paper [DLS06a] they constituted the essential object to overcome the "large gap problem" and prove the existence of diffusion. In the previous chapter one can find a detailed and deep analysis of these objects.

In this thesis, we will discuss algorithms basically for diffeomorphisms. For the case of vector fields, we note that taking time-1 maps, we can reduce the problem of vector fields to the problem of diffeomorphisms. Nevertheless in some practical applications it will be convenient to present a direct treatment of case of vector fields. For this reason, in some cases, we include algorithms that are specially designed for flows.

This chapter is organized as follows. In Section 2.2 we summarize the notions of mechanics we will use. In Section 2.3 we formulate the invariance equations for the objects of interest (invariant tori, invariant bundles and invariant manifolds) and we will present some generalities of the numerical algorithms. In Section 2.5 we specify the fast algorithm to compute maximal tori –both primary and secondary– and we compare it with a straightforward Newton method (Section 2.4).

In Section 2.6 we present fast algorithms for the iteration of cocycles over rotations and for the calculation of their invariant bundles. The main idea is to use a renormalizable algorithm which allows to pass from a cocycle to a longer cocycle. The calculation of invariant bundles for cocycles is a preliminary step for the algorithms for the calculation of whiskered invariant tori. Indeed, these algorithms require the computation of the projections over the linear subspaces of the linear cocycle. In Section 2.7.1 we present an alternative procedure to compute the projections based on a Newton method. Algorithms for whiskered tori are discussed in Section 2.7.

In Section 2.8 we discuss fast algorithms to compute rank-1 (un)stable manifolds of whiskered tori. Again, the key point is that taking advantage of the geometry of the problem we can devise algorithms which implement a Newton step without having to store—and much less invert—a large matrix. We first discuss the order by order method, which serves as a comparison with the more efficient methods based on the reducibility. We present algorithms that compute at the same time the torus and the whiskers and algorithms that given a torus and the linear space compute the invariant manifold tangent to it.

Of course, Newton methods require a good initial guess. Typically, one uses a continuation method starting from an integrable case, where the solutions are trivial and can be computed analytically. However, in the case of secondary KAM tori, which do not exist in the integrable case, one requires other types of methods. In Section 2.9 we include a discussion of the different possibilities.

Finally, in Section 2.10 we include examples of the numerical implementation we have carried out. In Section 2.10.1 we computed maximal invariant tori, both primary and secondary, of the standard maps and in Section 2.10.2 we computed maximal and hyperbolic invariant tori of the Froeschlé map. We also provide details of storage and running

times.

2.2 Setup and conventions

We will be working on systems defined on an Euclidean phase space endowed with a symplectic form. Hence, the phase space we will consider is $(U \subset \mathbb{R}^d) \times \mathbb{T}^d$. Nevertheless, we will not assume that the coordinates in the phase space are action-angle variables.

There are several reasons why we do not want to assume that the system is given in action-angle variables. As we will see, this assumption does not simplify our work much. We also note that there are several systems (even quasi-integrable ones) which are very smooth in Cartesian coordinates but less smooth in action-angle variables (e.g., neighborhoods of elliptic fixed points, hydrogen atoms in crossed electric and magnetic fields, several problems in celestial mechanics, ...).

Therefore, we will assume that the symplectic form Ω is given by an antisymmetric matrix J

$$\Omega(u, v) = \langle u, Jv \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard Euclidean inner product.

An important particular case is when

$$J^2 = -\operatorname{Id}\,, \tag{2.1}$$

but most of our calculations do not need this assumption. One important case, where the identity (2.1) is not satisfied, is in surfaces of section chosen arbitrarily in the energy surface.

We will also assume that the symplectic form is exact, that is $\Omega = d\alpha$ for some 1-form α.

We will consider that the maps F are not only symplectic, that is $F^*\Omega = \Omega$, where $F^*\Omega$ is the pull-back of Ω under f, but also exact symplectic, that is

$$F^*\alpha = \alpha + dP,$$

where P is a 0-form, called the *primitive function*.

Similarly, we will assume that the flows we consider are globally Hamiltonian, that is the vector field X is given by

$$X = J\nabla H$$

where H is a globally defined function.

We will also assume that the frequencies ω that we consider are Diophantine, as it is standard in the KAM theory. We recall here that the notion of Diophantine is different for flows and for diffeomorphisms. Hence, we define

$$\begin{aligned} \mathcal{D}^{\text{aff}}(\nu, \tau) &= \left\{ \omega \in \mathbb{R}^d \;\middle|\; |\omega \cdot k|^{-1} \leq \nu |k|^\tau \;\forall\; k \in \mathbb{Z}^d - \{0\} \right\} \\ \mathcal{D}(\nu, \tau) &= \left\{ \omega \in \mathbb{R}^d \;\middle|\; |\omega \cdot k - n|^{-1} \leq \nu |k|^\tau \;\forall\; k \in \mathbb{Z}^d - \{0\},\; n \in \mathbb{Z} \right\}, \end{aligned} \tag{2.2}$$

which correspond to the sets of Diophantine frequencies for flows and maps, respectively.

It is well known that for non-Diophantine frequencies substantially complicated behavior can appear [Her92, FKW01]. Observing convincingly these Liouvillian behaviors seems a very ambitious challenge for numerical exploration.

2.3 Equations for invariance

In this section we will discuss the functional equations for the objects of interest, that is, invariant tori and the associated whiskers.

2.3.1 Equations for invariant tori

We recall that quasi-periodic functions of continuous and discrete time are functions that can be written in Fourier series, respectively, as

$$\begin{aligned} x(t) &= \sum_{k\in\mathbb{Z}^\ell} \hat{x}_k e^{2\pi i k\cdot\omega t} \\ x_n &= \sum_{k\in\mathbb{Z}^\ell} \hat{x}_k e^{2\pi i k\cdot\omega n} \ , \end{aligned} \tag{2.3}$$

where $\omega \in \mathbb{R}^\ell$ and the components of ω are independent over the integers.

Note that we allow some components of x to be angles. In that respect, it suffices to take some of the components of (2.3) modulo 1.

It is natural to describe the quasi-periodic functions using the "hull" function $K : \mathbb{T}^\ell \to \mathbb{R}^d \times \mathbb{T}^d$ defined as

$$K(\theta) = \sum_{k\in\mathbb{Z}^\ell} \hat{x}_k e^{2\pi i k\theta},$$

so that we can write

$$\begin{aligned} x(t) &= K(\omega t) \\ x_n &= K(n\omega). \end{aligned}$$

The geometric interpretation of the hull function is that it gives an embedding of $\mathbb{T}^\ell$ into the phase space. In our applications, the torus will be invariant, so that the embedding will actually be an immersion.

It is clear that quasi-periodic functions will be orbits for a vector field X or a map F if and only if the hull function K satisfies:

$$\begin{aligned} \omega\frac{\partial}{\partial\theta}K - X\circ K &= 0 \\ F\circ K - K\circ T_\omega &= 0, \end{aligned} \tag{2.4}$$

where T_ω denotes a rigid rotation, given by $T_\omega(\theta) = \theta + \omega$.

A modification of the invariance equations (2.4) which we will study is

$$\begin{aligned} \omega\frac{\partial}{\partial\theta}K - X\circ K - (J\circ K_0)(DX\circ K_0)\lambda &= 0 \\ F\circ K - K\circ T_\omega - (J(K_0)DK_0)\circ T_\omega\lambda &= 0, \end{aligned} \tag{2.5}$$

where the unknowns are now $K : \mathbb{T}^\ell \to \mathbb{R}^d \times \mathbb{T}^d$ (as before) and $\lambda \in \mathbb{R}^\ell$. Here, K_0 is an approximate solution of the equation (2.4).

It has been shown in [FLS07] that, for exact symplectic maps, if (K, λ) satisfy the equation (2.5) with K_0 close to K, then $\lambda = 0$ and K is a solution of the invariance equation (2.4). Of course, for approximate solutions of the invariance equation (2.4), there is no reason why λ should vanish. The vanishing of λ depends on global considerations that are discussed in Section 2.3.1.1.

The advantage of equation (2.5) is that it makes easier to implement a Newton method in the cases that, for the approximate solutions, certain cancelations do not apply. This is particularly important for the case of secondary tori that we will discuss in Section 2.3.1.2.

The equations (2.4) and (2.5) will be the centerpiece of our treatment. We will discretize them using Fourier series and study numerical methods to solve the discretized equations.

It is important to remark that there are "a posteriori" rigorous results for equations (2.4). That is, there are theorems that ensure that given a function which satisfies (2.4) rather approximately and which, at the same time, satisfies some non-degeneracy conditions, then there is a true solution nearby. These results, stated in [FLS07], (whose proof follows closely the algorithms we discuss) give us some extra quantities to monitor so that we can be confident that the numerical solutions computed are not spurious effects induced by the truncation.

Remark 2.3.1. Notice that the dimension of the torus ℓ is smaller than the dimension of the phase space $2d$ for whiskered tori. In the case of maximal tori $\ell = d$. Hence, the method suggested has the numerical advantage that that it looks for a function K, which is a function of ℓ variables, to compute objects of dimension ℓ. This is important because the cost of handling functions grows exponentially fast with the number of variables. Indeed, to discretize a function of ℓ variables into $\mathbb{R}^n$ in a grid of side h one needs to store $(1/h)^\ell \cdot n$ values.

Remark 2.3.2. Recall that taking time-1 maps one can reduce the problem of vector fields to the problem of diffeomorphisms. Indeed, since autonomous Hamiltonians systems preserve energy, we can take a surface of section and deal with the return map. This reduces by 1 the number of variables needed to compute invariant tori.

Remark 2.3.3. The equations (2.4) do not have unique solutions. Observe that if K is a solution, for any $\sigma \in \mathbb{R}^\ell$, $K \circ T_\sigma$ is also a solution. In [LGJV05] and [FLS07] it is shown that, in many circumstances, this is the only non uniqueness in a neighborhood. Hence, it is easy to get rid of it by imposing some normalization.

2.3.1.1 Some global considerations

Since $K : \mathbb{T}^\ell \to \mathbb{R}^d \times \mathbb{T}^d$ and both the domain and the range have some topology, there will be some topological considerations in the way that the torus $\mathbb{T}^\ell$ gets embedded in the phase space. Indeed, the angle variables of $\mathbb{T}^\ell$ can get wrapped around the angle variables in different ways.

A concise way of characterizing the topology of the embedding is to consider the lift

$$\widehat{K} : \mathbb{R}^\ell \to \mathbb{R}^d \times \mathbb{R}^d,$$

in such a way that K is obtained from $\widehat{K}$ by identifying variables in the domain and in the range that differ by an integer.

It is therefore clear that, $\forall\ e \in \mathbb{Z}^\ell$

$$\begin{aligned} \widehat{K}^p(\theta + e) &= \widehat{K}^p(\theta) \\ \widehat{K}^q(\theta + e) &= \widehat{K}^q(\theta) + I(e) \end{aligned} \tag{2.6}$$

where $\widehat{K}^p$, $\widehat{K}^q$ denote the different components. It is easy to see that $I(e)$ is a linear function of e

$$I(e) = \left(\sum_{j=1}^{\ell} I_{ij} e_j \right)_{i=1,\ldots,d}$$

with $I_{ij} \in \mathbb{Z}$.

We note that if a function $\widehat{K}^q$ satisfies

$$\widehat{K}^q(\theta + e) = \widehat{K}^q(\theta) + Ie\ ,$$

the function

$$\widetilde{K}^q(\theta) \equiv \widehat{K}^q(\theta) - I\theta \tag{2.7}$$

is periodic. The numerical methods will always be based on studying the periodic functions $\widetilde{K}^q$, but we will not emphasize this unless it can lead to confusion.

Of course, the integer valued matrix I remains constant if we modify the embedding slightly. Hence, it remains constant under continuous deformation. For example, in the integrable case with $\ell = d$, invariant tori satisfy $\widehat{K}^q(\theta) = \theta$, so that we have $I = \mathrm{Id}$ and, therefore, for all the invariant tori which can be continued to tori of the integrable case we also have $I = \mathrm{Id}$.

2.3.1.2 Secondary tori

Nevertheless, one can produce other ℓ-dimensional tori where the range of I is less then ℓ. It is easy to see that if $\mathrm{rank}(I) < \ell$ we can contract $K(\mathbb{T}^\ell)$ to a diffeomorphic copy of $\mathbb{T}^{\mathrm{rank}(I)}$. Even in the case of maximal dimensional tori $\ell = d$, one can have tori with contractible directions. The most famous example are the "islands" generated in twist maps around resonances. These tori are known as *secondary tori* and they do not exist in the integrable case. Indeed, they are created by the perturbation and therefore they cannot be obtained by continuation of the objects present in the integrable system, as it is standard in perturbative KAM theory.

Perturbative proofs of existence of secondary tori are done in [LW04] and in [DLS06a]. The properties of these tori are studied in great detail in the first part of this thesis. They were shown to have an important role in Arnol'd diffusion [DLdlS03, DLS06a, GL06b] to overcome the large gap problem described in the first part of this thesis.

In [HL00] there are heuristic arguments and numerical simulations arguing that in systems of coupled oscillators, the tori with contractible directions are much more abundant than the invariant tori which can be continued to the integrable limit.

In view of these reasons, we will pay special attention to the computation of these secondary tori in the numerical examples in Section 2.10.

One of the novelties of the method described is that it can deal in a unified way with both primary and secondary KAM tori. We want to remark some features of the method presented here which are crucial for the computation of secondary tori:

- The method does not require neither the system to be close to the integrable nor to be written in action-angle variables. Recall that secondary tori do not exist in the integrable case and they are generated by the resonances, and they can appear very close to the separatrices, where the action variable becomes singular (see the chapter 1 in this thesis).

- The modification of the invariance equation (2.4) by a term containing λ. This allows to adjust some global averages required to solve the Newton equations (see Section 2.5)

- The periodicity of the function K can be adjusted by the matrix I introduced in (2.6). Hence, the rank of the matrix I has to be chosen according to the number of contractible directions.

2.3.2 Equations for the invariant whiskers

Invariant tori usually have associated invariant bundles and whiskers. We are interested in computing the invariant manifolds which contain the torus and are tangent to the invariant bundles of the linearization around the torus. These include the stable and unstable manifolds, but also other invariant manifolds associated to other invariant bundles of the linearization, like the slow manifolds, associated to the less contracting directions.

Even if we can develop algorithms for a general case, which can deal even with the resonant cases, we postpone it for a future work and we will focus on the algorithms for the manifolds of rank-1, which will be discussed in Section 2.8. We think that considering this particular case we can state in a more clear and simpler way the main idea behind the algorithms. Moreover, they are very important in practice. The implementation of these algorithms is now being pursued.

Following the parameterization method, we will look for a function $W : \mathbb{T}^\ell \times (U \subset \mathbb{R}^{d-\ell}) \to \mathbb{R}^{2d}$ and a scalar λ satisfying the following invariance equation:

$$\begin{aligned} &F(W(\theta, s)) = W(\theta + \omega, \lambda s) \\ &\left(\omega \frac{\partial}{\partial \theta} + \lambda s \frac{\partial}{\partial s}\right) W(\theta, s) = X(W(\theta, s)), \end{aligned} \tag{2.8}$$

for the case of maps and flows, respectively.

Notice that equation (2.8) implies that in the variables (θ, s) the motion on the torus consist of a rigid rotation of frequency ω whereas the motion on the whiskers consists of a contraction by a constant λ (e^λ in the case of flows). We assume then that $|\lambda| < 1$ ($\lambda < 0$ for flows). The expanding case $|\lambda| > 1$ ($\lambda > 0$ for flows) is analogous.

2.3.3 Fourier-Taylor discretization

2.3.3.1 Fourier series discretization

Since the invariant tori are parameterized by a function which is periodic in the angle variable θ, it natural to discretize the functions K using Fourier modes and retaining a finite number of them,

$$K(\theta) = \sum_{k=0}^{N} c_k e^{2\pi k\theta}. \tag{2.9}$$

Since we will deal with real functions, then $c_k = c_{-k}^*$ and one can just consider the cosine and sine Fourier series,

$$K(\theta) = a_0 + \sum_{k=1}^{N} a_k \cos(2\pi k\theta) + b_k \sin(2\pi k\theta). \tag{2.10}$$

Therefore, in order to store K, we can either keep the values of the function in a grid of $2N$ points or $N+1$ Fourier modes of the Fourier series.

The main practical shortcoming of Fourier series discretization is that they are not adaptive and that for discontinuous functions, they converge very slowly and not uniformly. These shortcomings are not very serious for our applications. The fact that the tori are invariant under smooth maps makes them very homogeneous, so that adaptivity does not help.

The fact that the Fourier series converge slowly for functions with discontinuities is a slight problem. It is known rigorously that, when KAM tori have a certain regularity, they are analytic [LGJV05, FLS07]. Nevertheless, for certain values of the parameters the equation (2.4) admits solutions—corresponding to Aubry-Mather sets—which are discontinuous (the theory is much more developed for twist maps). Therefore, as we increase the parameters, the problem switches from having analytic solutions to having discontinuous solutions (this is the so-called breakdown of analyticity [Aub83, ALD83, Gre79, McK82, CFL04, OP08]).

For values of parameters which are close to the transition of analyticity breakdown, the Fourier discretization tends to behave in rather surprising ways and leads to spurious solutions. In Section 2.10.3 we will discuss how to detect these spurious solutions.

An important advantage of the Fourier discretization is that the coboundary equations, which play a very important role in KAM theory and in our treatment, are straightforward to solve in Fourier coefficients. That is, given η periodic we want to find φ also periodic solving

$$\begin{aligned} \omega \cdot \frac{\partial}{\partial \theta} \varphi &= \eta \\ \varphi - \varphi \circ T_\omega &= \eta. \end{aligned} \tag{2.11}$$

As it is well known equations (2.11) have a solution provided that $\hat{\eta}_0 \equiv \int_{\mathbb{T}^\ell} \eta = 0$. The

Fourier coefficients $\hat{\varphi}_k$ of the solution φ are given then by

$$\begin{aligned} \hat{\varphi}_k &= \frac{\hat{\eta}_k}{2\pi i \omega \cdot k} \\ \hat{\varphi}_k &= \frac{\hat{\eta}_k}{1 - e^{2\pi i k \cdot \omega}}, \end{aligned} \tag{2.12}$$

where $\hat{\eta}_k$ are the Fourier coefficients of the function η. Notice that the solution φ is unique up to the addition of a constant (the average $\hat{\varphi}_0$ of φ is arbitrary). The equation (2.11) and its solution (2.12) are very standard in KAM theory (see the exposition in [Lla01b]). Very detailed estimates can be found in [Rüs75], when ω is Diophantine, which is our case.

On the other hand, we note that evaluation of $F \circ K$ is very fast if we discretize on a grid (we just need to evaluate the function F for each of the points on the grid). Hence, our iterative step will consist in the application of several operations all of which are fast either in the Fourier mode representation or in the grid representation.

Of course, using the Fast Fourier Transform we can transform from a grid representation to the Fourier coefficients in $O(N \log N)$ operations. Indeed, as a practical matter, the FFT is a very studied algorithm and there are extremely efficient implementations that take into account not only operation counts but also several other characteristics (memory access, cache, etc.) of modern computers.

2.3.3.2 Algebraic operations and elementary transcendental functions with Fourier series

Algebraic operations (sum, product, division) and elementary transcendental functions (sin, cos, exp, log, power, ...) of Fourier series can be computed either by manipulation of the Fourier coefficients or by using FFT.

For example, the product $h = f \cdot g$ of two Fourier series can be computed either by the Cauchy formula

$$h_k = \sum_{i=0}^{k} f_{k-i} g_i, \tag{2.13}$$

or by applying the inverse FFT to the coefficients of f and g, computing the product function on each point of the grid in real space and then applying the FFT. The first method clearly takes $O(N^2)$ operations while the second only $O(N \ln N)$.

A modification of the FFT algorithm which leads to some improvement is to consider a series of Fourier terms as a series of length $2N$, compute the inverse FFT on $2N$ points, perform the product and then take the FTT back. Note that, at this point, except for round-off errors, this algorithm is exact for trigonometric polynomials of degree $2N$. The final step is to truncate again to a polynomial of degree N.

The analysis of algorithms of multiplication from the point of view of theoretical computer science have been undertaken in [Knu97], but to our knowledge, there are few studies of the effects of truncation roundoff. An empirical study for the case of one variable was undertaken in [CL08].

In the case of several variables, the issues of numerical stability remain, but we also note that, from the point of efficiency, the way that the multiple loops involved in the evaluation of (2.13) are organized becomes crucial. These considerations depend on details of the computer architecture and are poorly understood. Some empirical studies can be found in [Har08].

2.3.3.3 Fourier-Taylor series

For the computation of whiskers of invariant tori we will use Fourier-Taylor expansions of the form

$$W(\theta, s) = \sum_{i=0}^{\infty} W_n(\theta) s^n, \tag{2.14}$$

where W_n are periodic functions in θ that we will approximate using Fourier series (2.10).

In order to manipulate these type of series we will use the so called *automatic differentiation algorithms* (see [Knu97]). For the basic algebraic operations and the elementary transcendental functions (exp, sin, cos, log, power, etc.), they provide an expression for the Taylor coefficients of the result in terms of the coefficients of each of the terms.

2.4 Numerical algorithms to solve the invariance equation for invariant tori

In this section, we will design a Newton method to solve equation (2.4) and discuss different algorithms to deal with the linearized equations.

Given an approximate solution K of the invariance equation (2.4), that is

$$\begin{aligned} \omega \frac{\partial}{\partial \theta} K - X \circ K &= E \\ F \circ K - K \circ T_\omega &= E \end{aligned} \tag{2.15}$$

with E small, the Newton method tries to compute Δ in such a way that setting $K \leftarrow K + \Delta$ and expanding the LHS of (2.15) in Δ up to order $|\Delta|^2$, it cancels the error E.

Performing a straightforward calculation, we obtain that the Newton procedure consists in finding Δ satisfying

$$\begin{aligned} \omega \cdot \frac{\partial}{\partial \theta} \Delta - (DX \circ K)\Delta &= -E \\ (DF \circ K)\Delta - \Delta \circ T_\omega &= -E. \end{aligned} \tag{2.16}$$

For the modified invariance equations (2.5), given an approximate solution K such that

$$\begin{aligned} \omega \cdot \frac{\partial}{\partial \theta} K - X \circ K - (J(K_0))(DX \circ K_0)\lambda &= E \\ F \circ K - K \circ T_\omega - (J(K_0) DK_0) \circ T_\omega \lambda &= E \end{aligned} \tag{2.17}$$

the Newton method consists in looking for (Δ, δ) in such a way that $K + \Delta$ and $\lambda + \delta$ eliminate the first order error. The linearized equations in this case are

$$\begin{aligned} \omega \cdot \frac{\partial}{\partial \theta}\Delta - DX \circ K\Delta - (J(K_0))(DX \circ K_0)\delta = -E \\ DF \circ K\Delta - \Delta \circ T_\omega - (J(K_0)DK_0) \circ T_\omega \delta = -E, \end{aligned} \tag{2.18}$$

where one can take $K_0 = K$.

The role of the parameter δ is now clear. It allows us to adjust some global averages that we need to be able to solve the equations (2.18).

As it is well known, the Newton method converges quadratically. Hence, the new error $\widetilde{E}$ for $K + \Delta$ will be such that $\|\widetilde{E}\| \leq C\|E\|^2$.

The main problem of the Newton method is that it needs a good initial guess. We will discuss several possibilities in Section 2.9. Of course, any reasonable algorithm can be used as an input to the Newton method. Indeed, our problems have enough structure that one can use Lindstedt series, variational methods, approximation by periodic orbits, frequency methods besides the customary continuation methods.

The linearized equations for the Newton method can be treated in different ways.

2.4.1 The large matrix method

The most straightforward method to implement the Newton method is

Algorithm 2.4.1. *Discretize the equations* (2.4) *using truncated Fourier series up to order N and apply the Newton method to the discretization.*

A slight variation is

Algorithm 2.4.2. *Discretize the equation* (2.4) *on a grid of $2N$ points and compute E. Discretize* (2.16) *using truncated Fourier series up to order N, solve the equation using a linear solver and apply the solution.*

The difference between algorithms 2.4.1 and 2.4.2 is that the first requires that the approximate derivative we are inverting is the derivative of the discretization.

We note that this extra symmetry is implementable using symbolic manipulation methods. Either of these algorithms requires a storage of a full $N \times N$ matrix. The solution of N linear equations requires $O(N^3)$ operations. There are several variations which are worth noting. It is sometimes convenient to use

$$K \leftarrow K + h\Delta$$

with $0 < h < 1$. This, of course, converges more slowly for very small h.

As we mentioned before in Remark 2.3.3, the solutions of the equations are not unique. One can cope with this by imposing normalizations in some cases. A general purpose solution is to use the singular value decomposition [GVL96]. The pseudo-inverse method leads to Δ's which reduce the residual as much as possible, which is all that is needed by the Newton method. We also note that, in contrast to Gaussian elimination which

is numerically unstable (the numerical instability can be partially mitigated by pivoting) the SVD computation, is numerically stable. In terms of speed, the SVD method is only a factor ≈ 4 slower than Gaussian elimination. For the cases that we will consider in this memoir, we think that the SVD is vastly superior to Gaussian elimination.

2.4.2 The Newton method and uniqueness

As we have mentioned in Remark 2.3.3, the solutions of (2.4) are not unique. Therefore, the implementations of the Newton method have to be careful not to lead to non-invertible matrices.

As we mentioned in Section 2.3.1.1, we will be looking for $\widetilde{K}(\theta) = \widehat{K}(\theta) - I\theta$ introduced in (2.7). Note that for $K_\sigma = K \circ T_\sigma$ we have

$$\widetilde{K}_\sigma = \widetilde{K} \circ T_\sigma - I\sigma$$

Hence, if $\{\nu_i\}_{i=1}^L$ is a basis for Range(I) we can impose the condition

$$\int \widetilde{K} \cdot \nu_i = 0 \qquad \text{for} \quad i = 1, \ldots, L \tag{2.19}$$

and we only need to deal with periodic functions which satisfy (2.19).

In the case that the dimension of the range of I is ℓ—the dimension of the torus—this leads to a unique solution (in the non-degenerate cases, according to the results of [LGJV05]) and we can expect that the matrices appearing in the discretization of the Newton method are invertible.

Two important examples of this situation are primary Lagrangian tori and some whiskered tori. In the case of secondary tori, as we will see, it is advantageous to use the extra variable λ to make progress in the Newton method.

2.5 Fast Newton methods for Lagrangian tori

In this section we follow [LGJV05] and develop a very effective Newton method. We present an algorithm so that the Newton step does not need to store any $N \times N$ matrix and only requires $O(N \log N)$ operations. We first discuss it for maps.

The key observation is that the Newton equations (2.16) and (2.18) are closely related to the dynamics and that, therefore, we can use geometric identities to find a linear change of variables that reduces the Newton equations to upper diagonal difference equations with constant coefficients. This idea is stated in the following Proposition:

Proposition 2.5.1. *Given an approximation K of the invariance equation as in* (2.15),

denote

$$\begin{aligned} \alpha(\theta) &= \frac{\partial}{\partial\theta}K \\ N(\theta) &= \left([\alpha(\theta)]^T\alpha(\theta)\right)^{-1} \\ \beta(\theta) &= \alpha(\theta)N(\theta) \\ \gamma(\theta) &= (J\circ K(\theta))\beta(\theta) \\ M(\theta) &= [\alpha(\theta) \mid \gamma(\theta)] \end{aligned} \tag{2.20}$$

where by $[\cdot \mid \cdot]$ *we denote the* $2d\times 2d$ *matrix obtained by juxtaposing the two* $2d\times d$ *matrices that are in the arguments.*

Then, we have

$$(DF\circ K(\theta))M(\theta) = M(\theta+\omega)\begin{pmatrix}\mathrm{Id} & A(\theta)\\ 0 & \mathrm{Id}\end{pmatrix} + \widehat{E}(\theta) \tag{2.21}$$

where

$$A(\theta) = \beta(\theta+\omega)^T[(DF\circ K(\theta))\gamma(\theta) - \gamma(\theta+\omega)], \tag{2.22}$$

and $\|\widehat{E}\| \le \|\nabla E\|$ *in the case of* (2.16) *or* $\|\widehat{E}\| \le \|\nabla E\| + |\lambda|$ *in the case of* (2.18)*.*

Remark 2.5.2. In case that condition $J^2 = -I$ is satisfied, we have that

$$\beta(\theta+\omega)^T\gamma(\theta+\omega) = 0,$$

and $A(\theta)$ has a simpler expression given by

$$A(\theta) = \beta(\theta+\omega)^T(DF\circ K(\theta))\gamma(\theta).$$

We omit the definition of the norms used in the bounds for $\widehat{E}$. For these precisions, we refer to the paper [LGJV05], where the convergence of the algorithm is established. For our purposes, we just note that, when we use (2.21) in the Newton step, we only incurr in higher order error.

It is interesting to pay attention to the geometric interpretation of the identity (2.21). Note that, taking derivatives with respect to θ in (2.15) we obtain that

$$(DF\circ K)\partial_\theta K - \partial_\theta K\circ T_\omega = \partial_\theta E,$$

which means that the vectors $\partial_\theta K$ are invariant under $DF\circ K$ (up to a certain error). Moreover, $(J\circ K)\partial_\theta KN$ are the symplectic conjugate vectors of $\partial_\theta K$, so that the preservation of the symplectic form clearly implies (2.21). The geometric interpretation of the matrix $A(\theta)$ is a shear near the approximately invariant torus. See Figure 2.1.

Next, we will see that the result stated in Proposition 2.5.1 allows to design a very efficient algorithm for the Newton step.

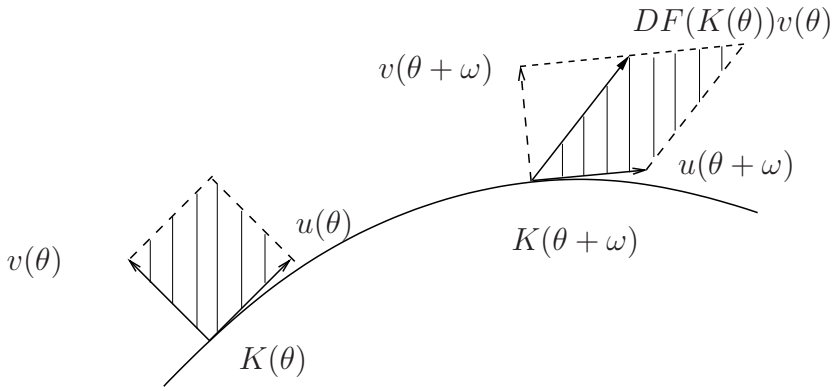

Figure 2.1: Geometric representation of the automatic reducibility

Notice first that if we change the unknowns $\Delta = MW$ in (2.16) and (2.18) and we use (2.21) we obtain

$$\begin{aligned} M(\theta+\omega)\begin{pmatrix}\mathrm{Id} & A(\theta)\\ 0 & \mathrm{Id}\end{pmatrix} W(\theta) - M(\theta+\omega)W(\theta+\omega) \\ - J\circ K_0(\theta+\omega)DK_0(\theta+\omega)\delta = -E(\theta)\end{aligned} \tag{2.23}$$

Of course, the term involving δ has to be omitted when considering (2.16).

Note that, multiplying (2.23) by $M(\theta+\omega)^{-1}$ we are left with the system of equations

$$\begin{aligned} &W_1(\theta) + A(\theta)W_2(\theta) - B_1(\theta)\delta - W_1(\theta+\omega) = -\widetilde{E}_1(\theta)\\ &W_2(\theta) - W_2(\theta+\omega) - B_2(\theta)\delta = -\widetilde{E}_2(\theta)\end{aligned} \tag{2.24}$$

where

$$\begin{aligned} \widetilde{E}(\theta) &= M(\theta+\omega)^{-1}E(\theta)\\ B(\theta) &= M(\theta+\omega)^{-1}J\circ K_0(\theta+\omega)DK_0(\theta+\omega)\end{aligned}$$

and the subindexes $i = 1, 2$ indicate coordinates. Of course, in the case of (2.16) we just omit the δ.

Notice that when K is close to K_0, we expect that B_2 is close to the d-dimensional Identity matrix and B_1 is small.

The next step is to solve equations (2.24) for W (and δ). Notice that equations (2.24) are a pair of equations of the form considered in (2.11) and they can be solved very efficiently in Fourier space.

More precisely, the second equation of (2.24) is uncoupled and allows us to determine W_2 (up to a constant) and δ. Indeed, we choose δ so that the term $B_2(\theta)\delta - \widetilde{E}_2$ has zero average. This allows to solve the equation for W_2 according to (2.12). We denote $W_2(\theta) = \widetilde{W_2}(\theta) + \overline{W}_2$ where $\widetilde{W_2}(\theta)$ has average zero and $\overline{W}_2 \in \mathbb{R}$.

Once we have $\widetilde{W}_2$, we can substitute W_2 in the first equation. We get $\overline{W}_2$ imposing that the average of

$$B_1(\theta)\delta - A(\theta)\widetilde{W}_2(\theta) - A(\theta)\overline{W}_2(\theta) - \widetilde{E}_1(\theta)$$

is zero and then we can find W_1 up to a constant according to (2.12).

Hence, we are led to the following algorithm:

Algorithm 2.5.3. *Consider given F, ω, K_0 and an approximate solution K (resp. K, λ). Perform the following calculations*

1. (1.1) *Compute $F \circ K$*

   (1.2) *Compute $K \circ T_\omega$*

2. *Set $E = F \circ K - K \circ T_\omega$ (resp. set $E = F \circ K - K \circ T_\omega - J \circ K_0 DK_0 \lambda$)*

3. *Following* (2.20)

   (3.1) Compute $\alpha(\theta) = \partial_\theta K$

   (3.2) Compute $N(\theta) = \left([\alpha(\theta)]^T \alpha(\theta)\right)^{-1}$

   (3.3) Compute $\beta(\theta) = \alpha(\theta) N(\theta)$

   (3.4) Compute $\gamma(\theta) = (J \circ K(\theta))\beta(\theta)$

   (3.5) Compute $M(\theta) = [\alpha(\theta) \mid \gamma(\theta)]$

   (3.6) Compute $M(\theta + \omega)$

   (3.7) Compute $M(\theta + \omega)^{-1}$

   (3.8) Compute $\widetilde{E}(\theta) = M(\theta + \omega)^{-1} E(\theta)$

   (3.9) Compute

   $$A(\theta) = [\beta(\theta)]^T DF(K(\theta))\gamma(\theta)$$

   as indicated in (2.22)

4. (4.1) *Solve for W_2 satisfying*

   $$W_2 - W_2 \circ T_\omega = -\widetilde{E}_2 - \int \widetilde{E}_2$$

   (resp.)

   (4.1′) *Solve for δ satisfying*

   $$\int \widetilde{E}_2 - \left[\int B_2\right]\delta = 0$$

   (4.2′) *Solve for W_2 satisfying*

   $$W_2 - W_2 \circ T_\omega = -\widetilde{E}_2 + B_2\delta$$

   Set W_2 such that the average is 0.

5. (5.1) *Compute* $A(\theta)W_2(\theta)$

   (5.2) *Solve for* $\overline{W}_2$ *satisfying*

$$0 = \int \widetilde{E}_1(\theta) + \int A(\theta)W_2(\theta) + \left[\int A(\theta)\right]\overline{W}_2$$

   (5.3) *Find* W_1 *solving*

$$W_1 - W_1 \circ T_\omega = -\widetilde{E}_1 - A(W_2 + \overline{W}_2)$$

   Normalize it so that $\int W_1 = 0$
   (resp.)

   (5.1′) *Compute* $A(\theta)W_2(\theta)$

   (5.2′) *Solve for* $\overline{W}_2$ *satisfying*

$$0 = \int \widetilde{E}_1(\theta) - \int B_1(\theta)\delta + \int A(\theta)W_2(\theta) + \left[\int A(\theta)\right]\overline{W}_2$$

   (5.3′) *Find* W_1 *solving*

$$W_1 - W_1 \circ T_\omega = -\widetilde{E}_1 - A(W_2 + \overline{W}_2) + B_1\delta$$

   Normalize it so that $\int W_1 = 0$

6. *The improved* K *is* $K(\theta) + M(\theta)W(\theta)$
   (resp. the improved λ *is* $\lambda + \delta$*).*

Notice that steps (1.2), (3.1), (3.6), (4.1) (resp. (4.2′)), (5.3) (resp. (5.3′)) are diagonal in Fourier series, the other steps are diagonal in the real space representation.

2.5.1 The Newton method for Lagrangian tori in Hamiltonian flows

As we mentioned in Remark 2.3.2 it is possible to reduce the treatment of differential equations to that of maps in numerically efficient ways. Nevertheless, it is interesting to present a direct treatment of the differential equation case of (2.4) or (2.5).

The main idea of the algorithm for flows is contained in the following Proposition:

Proposition 2.5.4. *Using the same notation as in Proposition 2.5.1 we have*

$$\omega\frac{\partial}{\partial\theta}M(\theta) - DX \circ K(\theta)M(\theta) = M(\theta)\begin{pmatrix}0 & S(\theta)\\ 0 & 0\end{pmatrix} + \widehat{E}(\theta) \tag{2.25}$$

where

$$S(\theta) = \beta^T(\theta)[-DX(K(\theta))J + JDX(K(\theta))]\beta(\theta) \tag{2.26}$$

and, as before, $\|\widehat{E}\| \le \|\nabla E\|$ *in the case of* (2.16) *or* $\|\widehat{E}\| \le \|\nabla E\| + |\lambda|$ *in the case of* (2.18)*.*

As before we refer the reader to the papers [LGJV05] for the definition of the norms and the proof of the convergence of the algorithm.

Again it is not difficult to see how to obtain the result stated in Proposition 2.5.4. Thus, considering the approximate invariance equation (2.15) in the case of flows and taking derivatives with respect to θ we obtain

$$(\omega\partial_\theta)\partial_\theta K - (DX \circ K)\partial_\theta K = \partial_\theta E. \tag{2.27}$$

Then, if we consider the change of variables M defined in Proposition 2.5.1 it is clear that the first n-columns of

$$\widetilde{M}(\theta) = \omega\frac{\partial}{\partial\theta}M(\theta) - DX \circ K(\theta)M(\theta)$$

are equal to zero (up to a certain error). Finally by equation (2.27) and the Hamiltonian character of the vector field it follows (2.25).

As in the case of symplectic maps we use equation (2.25) to transform the linearized Newton equation so that it can be solved in a very efficient way. Hence, if we change the unknowns $\Delta = MW$ and we use (2.25), equations (2.16) and (2.18) for flows reduce to

$$\begin{aligned} &M(\theta)\begin{pmatrix} 0 & S(\theta) \\ 0 & 0 \end{pmatrix} W(\theta) + M(\theta)\omega\frac{\partial}{\partial\theta}W(\theta) \\ &\quad - JDX \circ K_0(\theta)\delta = -E(\theta) \end{aligned} \tag{2.28}$$

and by multiplying by $M(\theta)^{-1}$ on both sides we are left with the system of equations

$$\begin{aligned} &\omega\partial_\theta W_1(\theta) + S(\theta)W_2(\theta) - B_1(\theta)\delta = -\widetilde{E}_1(\theta) \\ &\omega\partial_\theta W_2(\theta) - B_2(\theta)\delta = -\widetilde{E}_2(\theta) \end{aligned} \tag{2.29}$$

where

$$\begin{aligned} \widetilde{E}(\theta) &= M(\theta)^{-1}E(\theta) \\ B(\theta) &= M(\theta)^{-1}J(DX \circ K_0(\theta)) \end{aligned}$$

Notice that in the case of (2.16) we just omit the δ.

The equations (2.29) reduce to solving an equation of the form (2.11). Hence, we determine first δ by imposing that the RHS of the second equation has average zero. Then, the second equation determines W_2 up to a constant which is fixed by the first equation by imposing that the average on the RHS is zero. Finally, we obtain $W_1(\theta)$ up to constant.

The algorithm for flows is the following

Algorithm 2.5.5. *Consider given $X = J\nabla H$, ω, K_0 and an approximate solution K (resp. K, λ). Perform the following calculations*

1. (1.1) *Compute $\omega\partial_\theta K$*

(1.2) *Compute* $X \circ K$

2. *Set* $E = \omega\partial_\theta K - X \circ K$ *(resp. set* $E = \omega\partial_\theta K - X \circ K - (J \circ K_0)(DX \circ K_0)\lambda$*)*

3. *Following* (2.20)

(3.1) Compute $\alpha(\theta) = \partial_\theta K$

(3.2) Compute $\beta(\theta) = J\alpha(\theta)$

(3.3) Compute $N(\theta) = ([\beta(\theta)]^t\beta(\theta))^{-1}$

(3.4) Compute $\gamma(\theta) = \beta(\theta)N(\theta)$

(3.5) Compute $M(\theta) = [\alpha(\theta) \mid \gamma(\theta)]$

(3.6) Compute $M(\theta)^{-1}$

(3.7) Compute $\widetilde{E}(\theta) = M(\theta)^{-1}E(\theta)$

(3.8) Compute

$$S(\theta) = \beta^T(\theta)[-DX(K(\theta))J + JDX(K(\theta))]\beta(\theta)$$

as indicated in (2.26)

4. (4.1) *Solve for* W_2 *satisfying*

$$\omega\partial_\theta W_2 = -\widetilde{E}_2 - \int \widetilde{E}_2$$

(resp.)

(4.1′) *Solve for* δ *satisfying*

$$\int \widetilde{E}_2 - \left[\int B_2\right]\delta = 0$$

(4.2′) *Solve for* W_2 *satisfying*

$$\omega\partial_\theta W_2 = -\widetilde{E}_2 + B_2\delta$$

Set W_2 *such that the average is* 0.

5. (5.1) *Compute* $S(\theta)W_2(\theta)$

(5.2) *Solve for* $\overline{W}_2$ *satisfying*

$$0 = \int \widetilde{E}_1(\theta) + \int S(\theta)W_2(\theta) + \left[\int S(\theta)\right]\overline{W}_2$$

(5.3) *Find* W_1 *solving*

$$\omega\partial_\theta W_1 = -\widetilde{E}_1 - S(W_2 + \overline{W}_2)$$

Normalize it so that $\int W_1 = 0$
(resp.)

(5.1′) *Compute* $S(\theta)W_2(\theta)$

(5.2′) *Solve for* $\overline{W}_2$ *satisfying*

$$0 = \int \widetilde{E}_1(\theta) + \int B_1(\theta)\delta - \int S(\theta)W_2(\theta) - \left[\int S(\theta)\right]\overline{W}_2$$

(5.3′) *Find* W_1 *solving*

$$\omega\partial_\theta W_1 = -\widetilde{E}_1 - S(W_2 + \overline{W}_2) + B_1\delta$$

Normalize it so that $\int W_1 = 0$

6. *The improved* K *is* $K(\theta) + M(\theta)W(\theta)$
(resp. the improved λ *is* $\lambda + \delta$*).*

Notice that steps (1.1), (3.1), (4.1) (resp. (4.2′)), (5.3) (resp. (5.3′)) are diagonal in Fourier series, the other steps are diagonal in the real space representation.

2.6 Fast iteration of cocyles over rotations. Computation of hyperbolic bundles

It is clear from the previous sections that the linearized Newton equations of the invariance equations are very closely tied to the long term behavior of the equations of variation describing the propagation of infinitesimal disturbances around an invariant object. This connection will become more apparent in our discussion of whiskered tori in Section 2.7. Indeed, the relation between structural stability and exponential rates of growth has been one of the basic ingredients of the theory of Anosov systems [Ano69].

In this section we will study algorithms related to the iteration of cocycles over rotations. These will be ingredients of further algorithms, some of which are independent of the goals of this chapter. Hence, we have striven to make this section independent of the rest.

2.6.1 Some standard definitions

Given a function $M : \mathbb{T}^\ell \to GL(n, \mathbb{R})$ and a vector $\omega \in \mathbb{R}^\ell$, we consider the *cocycle* over the rotation T_ω associated to the matrix M. This is the function $\mathcal{M} : \mathbb{Z} \times \mathbb{T}^\ell \to GL(n, \mathbb{R})$ defined by

$$\mathcal{M}(n, \theta) = \begin{cases} M(\theta + (n-1)\omega) \cdots M(\theta) & n \geq 1 \\ \mathrm{Id} & n = 0 \\ M^{-1}(\theta + (n+1)\omega) \cdots M^{-1}(\theta) & n \leq 1 \end{cases} \tag{2.30}$$

Equivalently, a cocycle is defined by the relation

$$\begin{aligned} \mathcal{M}(0, \theta) &= \mathrm{Id} \\ \mathcal{M}(1, \theta) &= M(\theta) \\ \mathcal{M}(n + m, \theta) &= \mathcal{M}(n, T_\omega^m(\theta)) \cdot \mathcal{M}(m, \theta) \end{aligned} \tag{2.31}$$

We will say that M is the generator of $\mathcal{M}$. Note that if $M(\mathbb{T}^d) \subset G$ where $G \subset GL(n, \mathbb{R})$ is a group, then the cocycle $\mathcal{M}$ also has range in the group G.

The main example of a cocycle for this chapter is

$$M(\theta) = DF \circ K(\theta),$$

for K a parameterization of an invariant circle satisfying (2.4). Other examples appear in discrete Schrödinger equations [Pui02]. In the mentioned examples, we have that the cocycles lie in the symplectic group and in the unitary group, respectively.

Similarly, given a matrix valued function $M(\theta)$, a *continuous time cocycle* $\mathcal{M}(\theta, t)$ is defined to be the only solution to:

$$\begin{aligned} \frac{d}{dt}\mathcal{M}(\theta, t) &= M(\theta + \omega t)\mathcal{M}(\theta, t) \\ \mathcal{M}(\theta, 0) &= \mathrm{Id} \end{aligned} \tag{2.32}$$

From the uniqueness of differential equations we have

$$\begin{aligned} \mathcal{M}(\theta, t + s) &= \mathcal{M}(\theta + \omega t, s)\mathcal{M}(\theta, t) \\ \mathcal{M}(\theta, 0) &= \mathrm{Id} \end{aligned} \tag{2.33}$$

Note that (2.32) and (2.33) are continuous time analogues of (2.30) and (2.31), respectively. Moreover, if $M(\mathbb{T}^d) \subset \mathfrak{G}$, where $\mathfrak{G}$ is the Lie algebra of the Lie group G, then $\mathcal{M}(\theta, t) \in G$.

The main example for us of a continuous time cocycle will be

$$M(\theta) = DX \circ K(\theta),$$

where K is a solution of the invariance equation (2.4) and X is a Hamiltonian vector field. In this case, $\mathcal{M}(\theta, t)$ is symplectic.

2.6.2 Hyperbolicity of cocycles

One of the most crucial properties of cocycles is hyperbolicity (or spectral dichotomies) [MS89, SS74, SS76a, SS76b, Sac78]

Definition 2.6.1. Given $0 < \lambda < \mu$ we say that a cocycle $\mathcal{M}_n(\theta)$ (resp. $\mathcal{M}(\theta, t)$) has a λ, μ dichotomy when we can find a constant $c > 0$ and a splitting depending on θ,

$$\mathbb{R}^d = \mathcal{E}^s_\theta \oplus \mathcal{E}^u_\theta$$

which is characterized by:

$$\begin{aligned} v \in \mathcal{E}^s_\theta &\Leftrightarrow |\mathcal{M}_n(\theta)v| \leq c\lambda^n |v| \ , \qquad \forall n \geq 0 \\ v \in \mathcal{E}^u_\theta &\Leftrightarrow |\mathcal{M}_n(\theta)v| \leq c\mu^n |v| \ , \qquad \forall n \leq 0 \end{aligned} \tag{2.34}$$

or, in the continuous time case

$$\begin{aligned} v \in \mathcal{E}^s_\theta &\Leftrightarrow |\mathcal{M}(\theta, t)v| \leq c\lambda^t |v| \ , \qquad \forall t \geq 0 \\ v \in \mathcal{E}^u_\theta &\Leftrightarrow |\mathcal{M}(\theta, t)v| \leq c\mu^t |v| \ , \qquad \forall t \leq 0. \end{aligned} \tag{2.35}$$

The notation $\mathcal{E}^s$ and $\mathcal{E}^u$ is meant to suggest that very important cases is when the splitting is between stable and unstable bundles. This is the case when $\lambda < 1 < \mu$, and we say, then, that the cocycle is hyperbolic. Nevertheless this is not a necessary condition.

A system can have several dichotomies, but for the purposes of this thesis the Definition 2.6.1 will be enough.

It is well known that the mappings $\theta \to \mathcal{E}^{s,u}_\theta$ are C^r if $M \in C^r$ for $r \in \mathbb{N} \cup \{\infty, \omega\}$ [HL06b]. This result uses heavily that we are considering cocycles over a rotation.

For simplicity, in the following definitions we will present mainly the case of discrete time and we will develop the small changes needed for the continuous case in remarks.

One fundamental problem for subsequent applications is the computation of the invariant splittings (and, of course, to ascertain their existence). The computation of the invariant bundles is closely related to the computation of iterations of the cocycle.

Given a typical vector $v \in \mathcal{E}^u_\theta$, we expect that, for $n \gg 1$, $\mathcal{M}_n(\theta)v$ will be a vector in $\mathcal{E}^u_{T^n_\omega(\theta)}$. This suggests a method to compute $\mathcal{E}^u_\theta$ by iteration, which may be practical if $\mathcal{E}^u_\theta$ bundle is 1-dimensional.

Notice that this is very similar in spirit to the calculation of the leading eigenvalue of a matrix by the power method. In our case, however, we will see that there are issues related to the θ dependence.

2.6.2.1 Equivalence of cocycles, reducibility

Definition 2.6.2. We say that $\widetilde{M}(\theta)$ is equivalent to $M(\theta)$ if there exists $Q : \mathbb{T}^\ell \to GL(n, \mathbb{R})$ in such a way that

$$\widetilde{M}(\theta) = Q(\theta + \omega)^{-1} M(\theta) Q(\theta). \tag{2.36}$$

It is easy to check that $\widetilde{M}$ being equivalent to M is an equivalence relation.

If $\widetilde{M}$ is equivalent to a constant cocycle, we say that $\widetilde{M}$ is "reducible."

The important point is that, when (2.36) holds, we have

$$\widetilde{\mathcal{M}}_n(\theta) = Q(\theta + n\omega)^{-1} \mathcal{M}_n(\theta) Q(\theta). \tag{2.37}$$

In particular, if M is a constant matrix A we have

$$\widetilde{\mathcal{M}}_n(\theta) = Q^{-1}(\theta + n\omega) A^n Q(\theta),$$

so that iterations of reducible cocycles are easy to compute.

2.6.3 Algorithms for fast iteration of cocycles

In its simplest form, the algorithm for fast iteration of cocyles is:

Algorithm 2.6.3. *Given $M(\theta)$, compute*

$$\widetilde{M}(\theta) = M(\theta + \omega) M(\theta). \tag{2.38}$$

Set $\widetilde{M} \to M$, $2\omega \to \omega$ and iterate the procedure.

The important property is that

$$\widetilde{\mathcal{M}}_n(\theta) = \mathcal{M}_{2n}(\theta) \ .$$

Hence, if $M^{(k)}(\theta)$ is the result of applying k times the renormalization procedure described in Algorithm 2.6.3, we have

$$\mathcal{M}_n^{(k)}(\theta) = \mathcal{M}_{2^k n}(\theta) \ .$$

If we discretize the cocycle $M(\theta)$ taking N points (or N Fourier modes) and denote by $C(N)$ the number of operations required to perform a step of Algorithm 2.6.3, then we can compute 2^k iterates at a cost of operations $kC(N)$.

Notice that the important point is that the cost to compute 2^k iterations is proportional to k. Of course, the proportionality constant depends on N. The form of this dependence on N depends on the details of how do we compute the shift and the product.

There are several alternatives to perform the transformation (2.38). The main difficulty arises from the fact that, if we have points on a equally spaced grid, then $\theta + \omega$ will not be in the same grid. We have at least three alternatives:

i. Keep the discretization by points in a grid and compute $M(\theta + \omega)$ by interpolating with nearby points.

ii. Keep the discretization by points in a grid but note that the shift is diagonal in Fourier space. Of course, the multiplication of the matrix is diagonal in real space.

iii. Keep the discretization in Fourier space but use the Cauchy formula for the product.

The cost factor of each of these alternatives is, respectively,

$$\begin{aligned} C_1(N) &= O(N) \\ C_2(N) &= O(N \log N) \\ C_3(N) &= O(N^2). \end{aligned} \tag{2.39}$$

Besides differing in their cost, the above algorithms may depend on their stability and roundoff properties.

Clearly, the main idea of the method is to precompute some blocks of the iteration, store them and use these blocks in iteration. One can clearly choose different strategies to group different blocks. It seems quite possible that different methods can lead to different numerical stability. At the moment, we lack a definitive theory of stability which allows us to choose the blocks.

Next, we will present an alternative consisting of using the QR decomposition for the iterates. As it has been argued by many people, it seems that the QR algorithm to compute the iterates is rather stable [Ose68, DVV02].

Algorithm 2.6.4. *Given $M(\theta)$ and a QR decomposition of $M(\theta)$,*

$$M(\theta) = Q(\theta)R(\theta),$$

perform the following operations:

- *Compute* $S(\theta) = R(\theta+\omega)Q(\theta)$
- *Compute pointwise a* QR *decomposition of* S*:* $S(\theta) = \bar{Q}(\theta)\bar{R}(\theta)$.
- *Compute* $\widetilde{Q}(\theta) = Q(\theta+\omega)\bar{Q}(\theta)$

  $$\widetilde{R}(\theta) = \bar{R}(\theta)R(\theta+\omega)$$

  $$\widetilde{M}(\theta) = \widetilde{Q}(\theta)\widetilde{R}(\theta)$$

- *Set* $M \leftarrow \widetilde{M}$

  $$R \leftarrow \widetilde{R}$$

  $$Q \leftarrow \widetilde{Q}$$

  $$2\omega \rightarrow \omega$$

and iterate the procedure.

Since the QR decomposition can be computed fast, indeed it takes $O(N \log N)$, then the cost of the implementation depends on the issues discussed in (2.39). One can also do the same procedure based on SVD, which is somewhat slower than the QR.

2.6.4 The "straddle the saddle" phenomenon and preconditioning

The iteration of cocycle has several pitfalls compared with the iteration of matrices. The main complication from the numerical point of view is that the (un)stable bundle does depend on the base point.

In this section we describe a geometric phenomenon that causes some instability in the iteration of cocycles. This instability –which is genuine– becomes catastrophic when we apply some of the fast iteration methods described in Section 2.6.3.

Since

$$\mathcal{M}_n(\theta) = \mathcal{M}_{n-1}(\theta+\omega)M(\theta),$$

we see that we can think of computing $\mathcal{M}_n(\theta)$ as applying $\mathcal{M}_{n-1}(\theta+\omega)$ to the column vectors of $M(\theta)$.

A column of M, which we will denote by $v(\theta)$, can be thought geometrically as an embedding from $\mathbb{T}^\ell$ to $\mathbb{R}^n$. If the stable manifold of $\mathcal{M}_{n-1}(\theta)$ has codimension ℓ or smaller, there could be points $\theta^* \in \mathbb{T}^\ell$ such that $v(\theta^*) \in \mathcal{E}^s_{\theta^*}$ but for points such that $0 < d(\theta, \theta^*) < \varepsilon$, then $v(\theta) \notin \mathcal{E}^s_\theta$.

Clearly,

$$\mathcal{M}_{n-1}(\theta^*+\omega)v(\theta^*)$$

will decrease exponentially. Nevertheless, for all θ in a neighborhood of θ^*

$$\mathcal{M}_{n-1}(\theta+\omega)v(\theta)$$

will grow exponentially. Of course, the direction in which the growth takes place, will depend on the projection of $\mathcal{M}_{n-1}(\theta+\omega)v(\theta)$ on $\mathcal{E}^u_{\theta+\omega}$.

For example, in the case that $n = 2$, $\ell = 1$ and the stable and unstable directions are one dimensional, the unstable components will have different signs and the vectors $\mathcal{M}_{n-1}(\theta+\omega)v(\theta)$ will align in opposite directions. An illustration of this phenomenon happens in Figure 2.2.

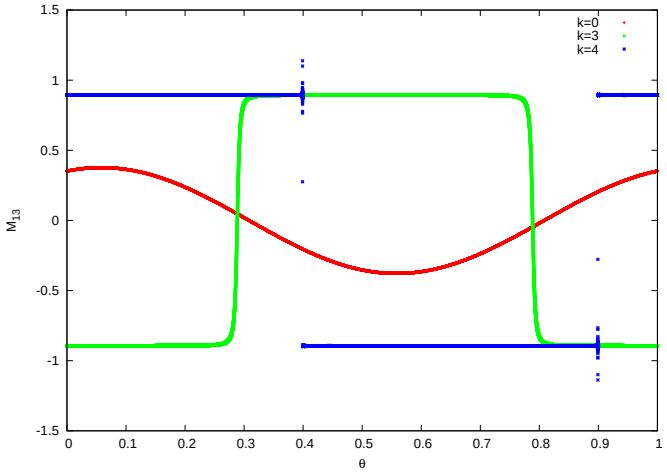

Figure 2.2: The straddle the saddle phenomenon. We plot one of the components of the cocycle $\mathcal{M}_{2^k}(\theta)$ for the values $k = 0, 3, 4$. The case $k = 0$ was scaled by a factor 200.

The transversal intersection of the range of $v(\theta)$ with $\mathcal{E}^s_\theta$ is indeed a true phenomenon, and it is a true instability.

Unfortunately, if $v(\theta)$ is very discontinuous, the discretization by Fourier series or the interpolation by splines will be extremely inaccurate so that the Algorithm 2.6.3 will be extremely inaccurate.

This phenomenon is easy to detect when it happens because the derivatives grow exponentially fast in some localized spots.

2.6.4.1 Eliminating straddle the saddle

Fortunately, once the phenomenon is detected, it can be eliminated. The main idea is that one can find an equivalent cocycle which does not have the problem (or presents it in a smaller extent).

In more geometric terms we observe that, even if the stable and unstable bundles are geometrically natural objects, the decomposition of a matrix into columns is coordinate dependent. Hence, if we choose some coordinate system which is reasonably close to the stable and unstable manifolds and we denote by Q the change of coordinates, then the cocycle

$$\widetilde{\mathcal{M}}(\theta) = Q(\theta+\omega)^{-1} M(\theta) Q(\theta),$$

is close to a constant.

This may seem somewhat circular, but the circularity can be broken using continuation methods. Given a cocycle which is close to constant, fast iteration methods work and they allow us to compute the splitting. Then, if we have computed Q for some M, we can use it to precondition the computation of neighboring M.

The algorithms for the computation of bundles will be discussed next. As we will see some of them are based on iteration.

2.6.5 Final computation of rank-1 bundles and the dynamics associated to them

The algorithms described in the previous section provide a fast way to iterate the cocycle. We will see that this iteration method, which is a kind of power method, gives the dominant eigenvalue $m(\theta)$ and the corresponding eigenvector $v(\theta)$.

The methods based on iteration rely strongly on the fact that the cocycle has one dominating eigenvalue which is simple. This is the case in the numerical examples we considered in Section 2.10.

Consider that we have performed k iterations of the cocycle (of course we perform scalings at each step) and we have computed $\mathcal{M}_n(\theta)$, with $n = 2^k$. Then, one can easily read the dominant rank-1 bundle from the QR decomposition of the cocycle $\mathcal{M}_n(\theta)$, just taking the column of Q associated to the larger value in the diagonal of the upper triangular matrix R. Indeed, one obtains a vector $v(\theta + 2^k\omega)$ (and therefore $v(\theta)$ by performing a shift of angle $-2^k\omega$) of modulus 1 spanning the unstable manifold. Since,

$$M(\theta)v(\theta) = m(\theta)v(\theta + \omega),$$

then

$$m(\theta) = ([M(\theta)v(\theta)]^T[M(\theta)v(\theta)])^{1/2}.$$

As it is standard in the power method, we perform scalings at each step. Indeed we divide all the entries in the matrix $M(\theta)$ by the maximum value among the entries of the matrix.

Hence, for the simplest case that there is one dominant eigenvalue, the method produces a section v (spanning the unstable subbundle) and a real function $m(\theta)$, which represents the dynamics on the rank 1 unstable subbundle, such that

$$M(\theta)v(\theta) = m(\theta)v(\theta + \omega).$$

Following [HL06a], under certain non-resonant conditions which are satisfied in the case of the stable and unstable subspaces, one can reduce the 1-dimensional cocycle associated to m and ω to a constant. Hence, we look for a positive function $p(\theta)$ and a constant λ such that

$$m(\theta)p(\theta) = \lambda p(\theta + \omega). \tag{2.40}$$

If $m(\theta) > 0$, we take logarithms on both sides of the equation (2.40). Then,

$$\log m(\theta) + \log p(\theta) = \log \lambda + \log p(\theta + \omega),$$

and taking $\log \lambda$ the average of $\log m(\theta)$, then the problem reduces to solve for $\log p(\theta)$ an equation of the form (2.11). The case $m(\theta) < 0$ is analogous. Of course, $p(\theta)$ and λ can be obtained just exponentiating.

2.7 Fast Newton methods for whiskered Lagrangian tori

In this section we follow [FLS07] and we develop an efficient Newton method to solve the invariance equations (2.4) and (2.5) for the case of whiskered tori, that is invariant tori with associated stable and unstable manifolds. We focus on the case of maps (the case for vector fields is similar).

As in the case of maximal KAM tori, we will assume that the motion on the torus is a rigid rotation with a Diophantine frequency $\omega \in \mathbb{Z}^\ell$. Then, in the case of whiskered tori we have that the direction tangent to the torus and, by the symplectic structure, its symplectic conjugate direction, span the 2ℓ-dimensional center subspace and the other $2d-2\ell$ directions are hyperbolic, that is, they contract exponentially either in the future or in the past.

Hence, we have that for every $\theta \in \mathbb{T}^\ell$ the tangent space $T_{K(\theta)}\mathcal{M}$, where $\mathcal{M} = \mathbb{R}^d \times \mathbb{T}^d$, has an invariant analytic splitting, that is

$$T_{K(\theta)}\mathcal{M} = \mathcal{E}^c_{K(\theta)} \oplus \mathcal{E}^s_{K(\theta)} \oplus \mathcal{E}^u_{K(\theta)} \tag{2.41}$$

such that, there exist $0 < \mu_1, \mu_2 < 1$, $\mu_3 > 1$ satisfying $\mu_1\mu_3 < 1$, $\mu_2\mu_3 < 1$ and $C > 0$ such that for all $n \geq 1$ and $\theta \in \mathbb{T}^\ell$

$$\begin{aligned} v \in \mathcal{E}^s_{K(\theta)} &\iff |\mathcal{M}(n,\theta)v| \leq C\mu_1^n|v| \qquad \forall n \geq 1 \\ v \in \mathcal{E}^u_{K(\theta)} &\iff |\mathcal{M}(n,\theta)v| \leq C\mu_2^n|v| \qquad \forall n \leq 1 \\ v \in \mathcal{E}^c_{K(\theta)} &\iff |\mathcal{M}(n,\theta)v| \leq C\mu_3^n|v| \qquad \forall n \in \mathbb{Z} \end{aligned} \tag{2.42}$$

where $\mathcal{M}$ is the cocycle associated to the matrix $M(\theta) = DF(K(\theta))$ and the frequency ω (see Definition 2.30).

We associate to this splitting the projections $\Pi^c_{K(\theta)}$, $\Pi^s_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ over the invariant spaces $\mathcal{E}^c_{K(\theta)}$, $\mathcal{E}^s_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$, respectively, which are analytic with respect to θ.

In Section 2.6 we have given a method to compute the rank-1 bundles by iterating the cocycle. Of course, once we have computed the vector $v(\theta)$ spanning the rank-1 (un)stable bundle it is very easy to obtain the projections. In Section 2.7.1 we discuss an alternative to compute the projections by means of a Newton method. In that case we do not need to assume that the bundle is 1-dimensional.

Recall that we want to design a Newton method to solve the invariance equation (2.4) and (2.5). Hence, we are left with solving the linearized equations, which are (2.16) and (2.18). The main difference with respect to maximal tori is that we first will project them on the invariant subspaces $\mathcal{E}^c, \mathcal{E}^u$ and $\mathcal{E}^s$, and then solve an equation for each subspace.

Thus, let us denote

$$\begin{aligned} \Delta^{s,c,u}(\theta) &= \Pi^{s,c,u}_{K(\theta)}\Delta(\theta) \\ E^{s,c,u}(\theta) &= \Pi^{s,c,u}_{K(\theta)}E(\theta), \end{aligned} \tag{2.43}$$

such that $\Delta(\theta) = \Delta^s(\theta)+\Delta^c(\theta)+\Delta^u(\theta)$. Then, by the invariant properties of the splitting, the linearized equations for the Newton method (2.16) and (2.18) split in

$$\begin{aligned} DF(K(\theta))\Delta^c(\theta) - \Delta^c \circ T_\omega(\theta) &= -E^c(\theta) \\ DF(K(\theta))\Delta^s(\theta) - \Delta^s \circ T_\omega(\theta) &= -E^s(\theta) \\ DF(K(\theta))\Delta^u(\theta) - \Delta^u \circ T_\omega(\theta) &= -E^u(\theta) \end{aligned} \tag{2.44}$$

and

$$\begin{aligned} DF(K(\theta))\Delta^c(\theta) - \Delta^c \circ T_\omega(\theta) + \Pi^c_{K(\theta)}(J \circ K_0(\theta))DK_0(\theta)\delta &= -E^c(\theta) \\ DF(K(\theta))\Delta^s(\theta) - \Delta^s \circ T_\omega(\theta) + \Pi^s_{K(\theta)}(J \circ K_0(\theta))DK_0(\theta)\delta &= -E^s(\theta) \\ DF(K(\theta))\Delta^u(\theta) - \Delta^u \circ T_\omega(\theta) + \Pi^u_{K(\theta)}(J \circ K_0(\theta))DK_0(\theta)\delta &= -E^u(\theta), \end{aligned} \tag{2.45}$$

respectively

We solve for Δ^c and δ the equation on the center subspace using the algorithm described in Section 2.5. Notice that once δ is obtained the equations (2.45) for the hyperbolic spaces reduce to the equations (2.44). More precisely,

$$DF(K(\theta))\Delta^{s,u}(\theta) - \Delta^{s,u} \circ T_\omega(\theta) = -\widetilde{E}^{s,u}(\theta) \tag{2.46}$$

where

$$\widetilde{E}^{s,u} = E^{s,u}(\theta) + \Pi^{s,u}_{K(\theta)}(J \circ K_0(\theta))DK_0(\theta)\delta.$$

Equations (2.44) and (2.45) for the stable and unstable spaces can be solved iteratively using the contraction properties of the cocycles on the hyperbolic spaces given in (2.42). Indeed, a solution for equations (2.46) is given explicitly by

$$\Delta^s(\theta) = \widetilde{E}^s \circ T_{-\omega}(\theta) + \sum_{k=1}^{\infty} (DF \circ K \circ T_{-\omega}(\theta) \times \cdots \times DF \circ K \circ T_{-k\omega}(\theta))(\widetilde{E}^s \circ T_{-(k+1)\omega}(\theta)) \tag{2.47}$$

for the stable equation, and

$$\Delta^u(\theta) = -\sum_{k=0}^{\infty} (DF^{-1} \circ K(\theta) \times \cdots \times DF^{-1} \circ K \circ T_{k\omega}(\theta))(\widetilde{E}^u \circ T_{k\omega}(\theta)) \tag{2.48}$$

for the unstable direction. Of course, the contraction of the cocycles guarantees the uniform convergence of these series.

In Section 2.7.2 we discuss fast algorithms to compute sums of the form (2.47) and (2.48). They are based on the idea exposed in Section 2.6 for fast iteration of the cocycles.

Hence, the algorithm for whiskered tori will be a combination of different algorithms that we summarize here:

Algorithm 2.7.1. *Consider given F, ω, K_0 and an approximate solution K (resp. K, λ), perform the following operations:*

- *Compute the projections associated to the cocycle $M(\theta) = DF \circ K(\theta)$ and ω using the algorithms described either in Section 2.6.5 or 2.7.1.*
- *Project the linearized equation to the center manifold and use the algorithm 2.5.3 to obtain Δ^s and δ.*
- *Project the linearized equation to the hyperbolic space and use the algorithms described in Section 2.7.2 to obtain $\Delta^{s,u}$.*
- *Set $K + \Delta^s + \Delta^u + \Delta^c \to K$ and $\lambda + \delta \to \lambda$ and iterate.*

2.7.1 A Newton method to compute the projections

In this section we will discuss a Newton method to compute the projections $\Pi^c_{K(\theta)}$, $\Pi^s_{K(\theta)}$ and $\Pi^u_{K(\theta)}$ associated to the linear spaces $\mathcal{E}^c_{K(\theta)}$, $\mathcal{E}^s_{K(\theta)}$ and $\mathcal{E}^u_{K(\theta)}$, respectively, where K is an (approximate) invariant tori.

We will design a Newton method to compute $\Pi^s_{K(\theta)}$ and $\Pi^{cu}_{K(\theta)} = \Pi^c_{K(\theta)} + \Pi^u_{K(\theta)}$. Similar arguments allow to design a Newton method to compute $\Pi^u_{K(\theta)}$ and $\Pi^{cs}_{K(\theta)} = \Pi^c_{K(\theta)} + \Pi^s_{K(\theta)}$. Then, of course, Π^c is given by

$$\Pi^c_{K(\theta)} = \Pi^{cs}_{K(\theta)} \Pi^{cu}_{K(\theta)} = \Pi^{cu}_{K(\theta)} \Pi^{cs}_{K(\theta)} \ .$$

Let us discuss first a Newton method to compute $\Pi^s_{K(\theta)}$ and $\Pi^{cu}_{K(\theta)}$. In order to simplify notation, from now on, we will omit $K(\theta)$ in the notation for the projections.

We will look for the maps $\Pi^s : \mathbb{T}^\ell \to GL(2d, \mathbb{R})$ and $\Pi^{cu} : \mathbb{T}^\ell \to GL(2d, \mathbb{R})$ satisfying the following equations:

$$\Pi^{cu}(\theta + \omega) M(\theta) \Pi^s(\theta) = 0 \tag{2.49}$$
$$\Pi^s(\theta + \omega) M(\theta) \Pi^{cu}(\theta) = 0 \tag{2.50}$$
$$\Pi^s(\theta) + \Pi^{cu}(\theta) = \mathrm{Id} \tag{2.51}$$
$$[\Pi^s(\theta)]^2 = \Pi^s(\theta) \tag{2.52}$$
$$[\Pi^{cu}(\theta)]^2 = \Pi^{cu}(\theta) \tag{2.53}$$
$$\Pi^s(\theta) \Pi^{cu}(\theta) = 0 \tag{2.54}$$
$$\Pi^{cu}(\theta) \Pi^s(\theta) = 0 \tag{2.55}$$

where $M(\theta) = DF(K(\theta))$.

Notice that the system of equations (2.49)–(2.55) is redundant. Indeed, it is easy to see that equations (2.53), (2.54) and (2.55) follow from equations (2.51) and (2.52). Therefore, the system of equations that needs to be solved is reduced to equations (2.49)–(2.52).

We are going to design a Newton method to solve equations (2.49)–(2.50) and use equations (2.51),(2.52) as constraints. Assume that we are given an approximate solution (Π^s, Π^{cu}) of the equations (2.49)–(2.50), which is exactly a projection. That is,

$$\Pi^{cu}(\theta + \omega) M(\theta) \Pi^s(\theta) = E^{cu}(\theta) \tag{2.56}$$
$$\Pi^s(\theta + \omega) M(\theta) \Pi^{cu}(\theta) = E^s(\theta) \tag{2.57}$$
$$\Pi^s(\theta) + \Pi^{cu}(\theta) = \mathrm{Id} \tag{2.58}$$
$$[\Pi^s(\theta)]^2 = \Pi^s(\theta) \tag{2.59}$$

Notice that the error for equation (2.56) has components only on the center and unstable "approximated" subspaces and we denote it by E^{cu}. The same happens with the equation (2.57) but on the "approximated" stable subspace. We assume that E^{cu} and E^s are both small.

We will look for Δ^s and Δ^{cu} in such a way that setting $\Pi^s \leftarrow \Pi^s + \Delta^s$ and $\Pi^{cu} \leftarrow \Pi^{cu} + \Delta^{cu}$, the new projections solve equations (2.49) and (2.50) up to order $|E|^2$ where $|E| = |E^s| + |E^{cu}|$.

Hence, the Newton method consists of looking for Δ^s and Δ^{cu} solving the following equations

$$\begin{aligned} &\Delta^{cu}(\theta+\omega)M(\theta)\Pi^s(\theta) + \Pi^{cu}(\theta+\omega)M(\theta)\Delta^s(\theta) = -E^{cu}(\theta) \\ &\Delta^s(\theta+\omega)M(\theta)\Pi^{cu}(\theta) + \Pi^s(\theta+\omega)M(\theta)\Delta^{cu}(\theta) = -E^s(\theta) \end{aligned} \tag{2.60}$$

with the constraints

$$\begin{aligned} &\Delta^s(\theta) + \Delta^{cu}(\theta) = 0 \\ &\Pi^s(\theta)\Delta^s(\theta) + \Delta^s(\theta)\Pi^s(\theta) = \Delta^s(\theta)\ . \end{aligned} \tag{2.61}$$

Notice that by (2.61) we only need to compute Δ^s. Of course, $\Delta^{cu} = -\Delta^s$. Next, we will try to work out equations (2.60) and (2.61) so that we can give an explicit expression for Δ^s.

Denote

$$\begin{aligned} \Delta^s_s &= \Pi^s\Delta^s \\ \Delta^s_{cu} &= \Pi^{cu}\Delta^s \end{aligned}$$

then

$$\Delta^s = \Delta^s_s + \Delta^s_{cu} \tag{2.62}$$

and (2.61) reads

$$\Delta^s_s(\theta) + \Delta^s(\theta)\Pi^s(\theta) = \Delta^s_s(\theta) + \Delta^s_{cu}(\theta), \tag{2.63}$$

from where

$$\Delta^s(\theta)\Pi^s(\theta) = \Delta^s_{cu}(\theta)\ . \tag{2.64}$$

Notice that by (2.58), (2.64) and (2.62) we have that

$$\Delta^s(\theta)\Pi^{cu}(\theta) = \Delta^s(\theta) - \Delta^s(\theta)\Pi^s(\theta) = \Delta^s(\theta) - \Delta^s_{cu}(\theta) = \Delta^s_s(\theta). \tag{2.65}$$

Now, using (2.61), equations (2.56)–(2.57) transform to

$$-\Delta^s(\theta+\omega)M(\theta)\Pi^s(\theta) + \Pi^{cu}(\theta+\omega)M(\theta)\Delta^s(\theta) = -E^{cu}(\theta)$$

and

$$\Delta^s(\theta+\omega)M(\theta)\Pi^{cu}(\theta) - \Pi^s(\theta+\omega)M(\theta)\Delta^s(\theta) = -E^s(\theta)$$

Let us introduce some notation. Denote

$$\begin{aligned} N_s(\theta) &= \Pi^s(\theta+\omega)M(\theta)\Pi^s(\theta) \\ N_{cu}(\theta) &= \Pi^{cu}(\theta+\omega)M(\theta)\Pi^{cu}(\theta), \end{aligned}$$

then, using that $\Pi^{cu}(\theta+\omega)M(\theta)\Pi^{s}(\theta)$ and $\Pi^{s}(\theta+\omega)M(\theta)\Pi^{cu}(\theta)$ are small by (2.56) and (2.57) and that $[\Pi^{s}(\theta)]^2=\Pi^{s}(\theta)$ and $[\Pi^{u}(\theta)]^2=\Pi^{u}(\theta)$ by (2.58) and (2.59) the equations we need to solve reduce to

$$-\Delta_s^s(\theta+\omega)\Pi^s(\theta+\omega)N_s(\theta)+N_{cu}(\theta)\Pi^{cu}(\theta)\Delta^s(\theta)=-E^{cu}(\theta)$$

and

$$\Delta^s(\theta+\omega)\Pi^{cu}(\theta+\omega)N_{cu}(\theta)-N_s(\theta)\Pi^s(\theta)\Delta^{cu}(\theta)=-E^s(\theta)$$

Finally, by (2.64) and (2.65) we obtain that

$$-\Delta_{cu}^s(\theta+\omega)N_s(\theta)+N_{cu}(\theta)\Delta_{cu}^s(\theta)=-E^{cu}(\theta) \tag{2.66}$$

$$\Delta_s^s(\theta+\omega)N_{cu}(\theta)-N_s(\theta)\Delta_s^s(\theta)=-E^s(\theta) \tag{2.67}$$

Equations (2.66)-(2.67) are of the form (2.70) for $A(\theta)=N_{cu}(\theta)$, $B(\theta)=N_s(\theta)$ and $\eta(\theta)=-E^{cu}(\theta)$ in the case of equation (2.66) and $A(\theta)=N_s(\theta)$, $B(\theta)=N_{cu}(\theta)$ and $\eta(\theta)=+E^s(\theta)$ in the case of equation 2.67. Notice that $\|N_s(\theta)\|<1$ and $\|N_{cu}^{-1}\|<1$. Hence, they can be solved iteratively using the fast iterative algorithms described in section 2.7.2.

The explicit expression for Δ_{cu}^s and Δ_s^s is given by

$$\begin{aligned}\Delta_{cu}^s(\theta)=-\Big[&N_{cu}^{-1}(\theta)E^{cu}(\theta)+\sum_{n=1}^{\infty}N_{cu}^{-1}(\theta)\times\cdots\times\\&N_{cu}^{-1}(\theta+n\omega)E^{cu}(\theta+n\omega)N_s(\theta+(n-1)\omega)\times\cdots\times N_s(\theta)\Big]\end{aligned} \tag{2.68}$$

and

$$\begin{aligned}\Delta_s^s(\theta)=E^s(\theta-\omega)N_{cu}^{-1}(\theta-\omega)+&\sum_{n=1}^{\infty}N_s(\theta-\omega)\times\cdots\times\\&N_s(\theta-(n+1)\omega)N_{cu}^{-1}(\theta-(n+1)\omega)\times\cdots\times N_{cu}^{-1}(\theta-\omega).\end{aligned} \tag{2.69}$$

Remark 2.7.2. Notice that by the way $N_{cu}(\theta)$ is defined, it is a matrix with no full rank. Therefore, we denote $N_{cu}^{-1}(\theta)$ to refer to the "pseudoinverse" matrix.

Finally, let us check that $\Delta^s=\Delta_{cu}^s+\Delta_s^s$ also satisfies the constraints for this problem. In order to check that constraint (2.64) is satisfied we first notice that

$$N_s(\theta)\Pi^s(\theta)=N_s(\theta)$$

and

$$N_{cu}^{-1}(\theta-\omega)\Pi^s(\theta)=0\ .$$

Moreover, from (2.56) and using (2.59) one can see that

$$E^{cu}(\theta)\Pi^s(\theta)=\Pi^{cu}(\theta+\omega)M(\theta)[\Pi^s(\theta)]^2=E^{cu}(\theta)\ .$$

Then, from expressions (2.68) and (2.69) it is clear that the constraint (2.64) is satisfied.

Now, using the first equation in (2.61) we get $\Delta^u(\theta) = -(\Delta^s_s(\theta) + \Delta^s_{cu}(\theta))$ and the improved projections are

$$\widetilde{\Pi}^s(\theta) = \Pi^s(\theta) + \Delta^s_s(\theta) + \Delta^s_{cu}(\theta)$$
$$\widetilde{\Pi}^{cu}(\theta) = \Pi^{cu}(\theta) + \Delta^u(\theta).$$

The new error for equations (2.49) and (2.50) is now $\|\widetilde{E}\| \leq C\|E\|^2$ where $\|E\| = \|E^{cu}\| + \|E^s\|$. Of course equation (2.51) is clearly satisfied but (2.52) is satisfied up to an error which is quadratic in $\|E\|$.

However it is easy to get an exact solution for (2.52) and the correction is quadratic in Δ^s (and therefore in Δ^{cu}). Thus, we just take the SVD decomposition of $\widetilde{\Pi}^s$ and we set the values in the singular value decomposition either to 1 or 0.

In this way we obtain new projections Π^s_{new} and $\Pi^u_{\text{new}} = \text{Id} - \Pi^s_{\text{new}}$ satisfying

$$\|\Pi^s_{\text{new}} - \widetilde{\Pi}^s\| < [\Delta^s]^2$$
$$\|\Pi^u_{\text{new}} - \widetilde{\Pi}^u\| < [\Delta^u]^2,$$

so that the error for equations (2.49) and (2.50) is still quadratic in $\|E\|$ and, moreover, they satisfy equations (2.52) and, of course, (2.51) exactly.

Hence, setting $\Pi^s \leftarrow \Pi^s_{\text{new}}$ and $\Pi^{cu} \leftarrow \Pi^{cu}_{\text{new}}$ we can repeat the procedure described in this section and perform another Newton step.

So, the algorithm for the Newton method to compute the projections is:

Algorithm 2.7.3. *Consider given F, k, ω and an approximate solution (Π^s, Π^{cu}) of equations* (2.49)-(2.50)*. Perform the following calculations:*

1. *Compute $M(\theta) = DF \circ K(\theta)$*
2. (2.1) *Compute $E^{cu}(\theta) = \Pi^{cu}(\theta + \omega)M(\theta)\Pi^s(\theta)$*

   (2.2) *Compute $E^s(\theta) = \Pi^s(\theta + \omega)M(\theta)\Pi^{cu}(\theta)$*
3. (3.1) *Compute $N_s(\theta) = \Pi^s(\theta + \omega)M(\theta)\Pi^s(\theta)$*

   (3.2) *Compute $N_{cu}(\theta) = \Pi^{cu}(\theta + \omega)M(\theta)\Pi^{cu}(\theta)$*
4. (4.1) *Solve for Δ^s_s satisfying*

$$N_s(\theta)\Delta^s_s(\theta) - \Delta^s_s(\theta + \omega)N_{cu}(\theta) = E^s(\theta)$$

   (4.2) *Solve for Δ^s_{cu} satisfying*

$$N_{cu}(\theta)\Delta^s_{cu}(\theta) - \Delta^s_{cu}(\theta + \omega)N_s(\theta) = -E^{cu}(\theta)$$

5. (5.1) *Compute $\widetilde{\Pi}^s(\theta) = \Pi^s(\theta) + \Delta^s_s(\theta) + \Delta^s_{cu}(\theta)$.*

   (5.2) *Compute the SVD decomposition of $\widetilde{\Pi}^s(\theta)$: $\widetilde{\Pi}^s(\theta) = U(\theta)\Sigma(\theta)V^T(\theta)$.*

(5.3) *Set the values in* $\Sigma(\theta)$ *equal to the closer integer (which will be either 0 or 1).*

(5.4) *Recompute* $\bar{\Pi}^s(\theta) = U(\theta)\Sigma(\theta)V^T(\theta)$.

6. *Set* $\bar{\Pi}^s \to \Pi^s$

$\mathrm{Id} - \bar{\Pi} s \to \Pi^{cu}$

and iterate the procedure.

2.7.2 Fast algorithms to solve the cohomology equation

In this section we present fast algorithms to solve for $\Delta(\theta)$ the cohomology equation

$$A(\theta)\Delta(\theta) - \Delta(\theta+\omega)B(\theta) = \eta(\theta) \tag{2.70}$$

for given $A(\theta)$, $B(\theta)$ and $\eta(\theta)$ satisfying either $\|A\| < 1$, $\|B^{-1}\| < 1$ or $\|A^{-1}\| < 1$, $\|B\| < 1$.

This type of equation appears in the computation of the projections using a Newton method (see equations (2.66)-(2.67)) as well as in the computation of whiskered tori (this is the resulting equation of the projection of the linearized equation of the Newton method onto the hyperbolic subspaces, see (2.44) and (2.46)).

The algorithms we present use the contraction of the cocycles and they consist of iterations. Interestingly enough, for the 1-dimensional case we present an amazing fast algorithm which does not use the contraction properties but Fourier transforms and solves it exactly. The main shortcoming is that it involves small divisors, whereas it is not the case for the iterative methods.

2.7.2.1 Fast algorithm for the 1-D cohomology equation

In this section we present an efficient algorithm for the 1-dimensional equation (2.70). It is an adaptation of Herman's "tricks" in [Her83].

Consider the following equation,

$$\frac{A(\theta)}{B(\theta)}\Delta(\theta) - \Delta(\theta+\omega) = \frac{\eta(\theta)}{B(\theta)} \tag{2.71}$$

which is obtained from (2.70) multiplying by $B^{-1}(\theta)$.

We will solve (2.71) in two steps:

1. Find $C(\theta)$ and $\lambda \in \mathbb{R}$ such that

$$\frac{A(\theta)}{B(\theta)} = \lambda \frac{C(\theta)}{C(\theta+\omega)} \tag{2.72}$$

2. Applying (2.72) in (2.71) and multiplying by $C(\theta+\omega)^{-1}$ we obtain

$$\lambda C(\theta)\Delta(\theta) - C(\theta+\omega)\Delta(\theta+\omega) = \tilde{\eta}(\theta) \tag{2.73}$$

where $\tilde{\eta}(\theta) = C(\theta+\omega)B^{-1}(\theta)\eta(\theta)$.

If we change the unknowns in (2.73) by $W(\theta) = C(\theta)\Delta(\theta)$, we are left with the equation

$$\lambda W(\theta) - W(\theta + \omega) = \tilde{\eta}(\theta). \tag{2.74}$$

Of course, if $|\lambda| \neq 1$, equation (2.74) can be solved in Fourier space. That is,

$$\widehat{W}_k = \frac{\widehat{\tilde{\eta}}_k}{\lambda - e^{2\pi i k\omega}} \ ,$$

and the solution is unique. Notice that equation (2.74) involves small divisors, which is not the case for the iterative methods discussed in the following section.

Finally, once we have $W(\theta)$ we get

$$\Delta(\theta) = C^{-1}(\theta)W(\theta)$$

Now, step 1 can be achieved by taking logarithms to (2.72). Assume that $A(\theta)/B(\theta) > 0$, otherwise we change the sign. Then,

$$\log A(\theta) - \log B(\theta) = \log \lambda + \log C(\theta) - \log C(\theta + \omega)$$

from where taking $\log \lambda$ the average of $\log A(\theta) - \log(B)$, the problem reduces to solve for $\log C(\theta)$ an equation of the form (2.11). Then $C(\theta)$ and λ can be obtained exponentiating. Notice that $\log C(\theta)$ is determined up to a constant. We will fix it by assuming average 0.

Hence, we have the following algorithm:

Algorithm 2.7.4. *Given $A(\theta)$, $B(\theta)$ and $\eta(\theta)$. Perform the following instructions:*

1. (1.1) *Compute* $L_A(\theta) = \log(A(\theta)) - \log(B(\theta))$

   (1.2) *Compute* $L_\lambda = \int L_A$

2. *Solve for* L_C *satisfying*

$$L_C(\theta) - L_C \circ T_\omega(\theta) = L_A(\theta) - L_\lambda$$

   Set L_C *such that the average is* 0.

3. (3.1) *Compute* $C(\theta) = \exp(L_C(\theta))$

   (3.2) *Compute* $\lambda = \exp(L_\lambda)$

4. *Compute* $\tilde{\eta}(\theta) = C(\theta + \omega)B^{-1}(\theta)\eta(\theta)$

5. *Solve for* W *satisfying*

$$\lambda W(\theta) - W(\theta + \omega) = \tilde{\eta}(\theta)$$

6. *The solution of* (2.70) *is* $\Delta(\theta) = C^{-1}(\theta)W(\theta)$

2.7.2.2 Fast iterative algorithms for the cohomology equation

In this section we will present a fast algorithm to solve equation (2.70) using iterative methods. The main idea is the same as the one described in section 2.6 for the fast iteration of the cocycles.

We will consider the case $\|A^{-1}\| < 1$ and $\|B\| < 1$. Then, multiplying (2.70) by $A^{-1}(\theta)$ on the RHS, we obtain

$$\Delta(\theta) = A^{-1}(\theta)\Delta(\theta+\omega)B(\theta) + A^{-1}(\theta)\eta(\theta). \tag{2.75}$$

Next, we compute $\Delta(\theta+\omega)$ by shifting (2.75) and substituting in (2.75), so that we get

$$\begin{aligned}\Delta(\theta) = {} & A^{-1}(\theta)\eta(\theta) \\ & + A^{-1}(\theta)A^{-1}(\theta+\omega)\eta(\theta+\omega)B(\theta) \\ & + A^{-1}(\theta)A^{-1}(\theta+\omega)\Delta(\theta+2\omega)B(\theta+\omega)B(\theta).\end{aligned}$$

Notice that if we define

$$\bar{\eta}(\theta) = A^{-1}(\theta)\eta(\theta)$$

and

$$\begin{aligned}A_1^{-1}(\theta) &= A^{-1}(\theta)A^{-1}(\theta+\omega) \\ B_1(\theta) &= B(\theta+\omega)B(\theta) \\ \eta_1(\theta) &= \bar{\eta}(\theta) + A^{-1}(\theta)\bar{\eta}(\theta)B(\theta)\end{aligned}$$

we have that

$$\Delta(\theta) = \eta_1(\theta) + A_1^{-1}(\theta)\Delta(\theta+2\omega)B_1(\theta)$$

which has the same structure as (2.75) and we can repeat the same scheme. This leads to an iterative procedure to compute $A(\theta)$ that is going to converge superexponentially in the number of iterations. Thus, define

$$\begin{aligned}A_{n+1}^{-1}(\theta) &= A_n^{-1}(\theta)A_n^{-1}(\theta+2^n\omega) \\ B_{n+1}(\theta) &= B_n(\theta+2^n\omega)B_1(\theta) \\ \eta_{n+1}(\theta) &= \eta_n(\theta) + A_n^{-1}(\theta)\eta_n(\theta)B_n(\theta)\end{aligned}$$

for $n \geq 0$, with

$$\begin{aligned}A_0^{-1}(\theta) &= A^{-1}(\theta) \\ B_0(\theta) &= B(\theta) \\ \eta_0(\theta) &= \bar{\eta}(\theta).\end{aligned}$$

Then,

$$\Delta(\theta) = \eta_{n+1}(\theta) + A_{n+1}^{-1}(\theta)\Delta(\theta+2^{n+1}\omega)B_{n+1}(\theta) \;\forall\; n \geq 0$$

and

$$\Delta(\theta) = \lim_{n\to\infty} \eta_{n+1}(\theta).$$

Of course, the convergence of the algorithm is guaranteed by the contraction of A^{-1} and B. And the cost of computing 2^n iterations of the algorithm is proportional to n.

The iterative algorithm is the following:

Algorithm 2.7.5. *Given $A(\theta)$, $B(\theta)$ and $\eta(\theta)$ perform the following operations:*

1. *Compute* $\Delta(\theta) = A^{-1}(\theta)\eta(\theta)$
2. *Compute*
   - *(2.1)* $\tilde{\Delta}(\theta) = A^{-1}(\theta)\Delta(\theta+\omega)B(\theta) + \Delta(\theta)$
   - *(2.2)* $\tilde{A}^{-1}(\theta) = A^{-1}(\theta)A^{-1}(\theta+\omega)$
   - *(2.3)* $\tilde{B}(\theta) = B(\theta+\omega)B(\theta)$
3. *Set* $\tilde{\Delta} \to \Delta$
   $\tilde{A}^{-1} \to A$
   $\tilde{B} \to B$
   $2\omega \to \omega$
4. *Iterate points* $2-3$

The case when $\|A\| < 1$ and $\|B^{-1}\| < 1$ can be done similarly. In this case, we multiply (2.70) by $B^{-1}(\theta)$ on the LHS so that we obtain

$$\Delta(\theta+\omega) = A(\theta)\Delta(\theta)B^{-1}(\theta) - \eta(\theta)B^{-1}(\theta).$$

Before applying the iterative scheme we shift by $-\omega$. In this way

$$\Delta(\theta) = A(\theta')\Delta(\theta')B^{-1}(\theta') - \eta(\theta')B^{-1}(\theta')$$

where $\theta' = T_{-\omega}\theta$.

Define

$$\bar{\eta}(\theta') = \eta(\theta')B^{-1}(\theta')$$

and

$$\begin{aligned} A_{n+1}(\theta') &= A_n(\theta')A_1(\theta' - 2^n\omega) \\ B_{n+1}(\theta') &= B^{-1}(\theta' - 2^n\omega)B^{-1}(\theta') \\ \eta_{n+1}(\theta') &= \eta_n(\theta') + A_n(\theta')\eta_n(\theta')B_n(\theta') \end{aligned}$$

for $n \geq 0$ with

$$\begin{aligned} A_0(\theta') &= A(\theta') \\ B_0(\theta') &= B^{-1}(\theta') \\ \eta_0(\theta') &= \bar{\eta}(\theta'), \end{aligned}$$

then

$$\Delta(\theta) = A_{n+1}(\theta')\Delta(\theta' - 2^{n+1}\omega)B_{n+1}^{-1}(\theta') - \eta_{n+1}(\theta')$$

and

$$\Delta(\theta) = \lim_{n\to\infty} -\eta_n(\theta').$$

Again the convergence is superexponential in n.

The iterative algorithm in this case is

Algorithm 2.7.6. *Given $A(\theta)$, $B(\theta)$ and $\eta(\theta)$, perform the following operations:*

1. *Compute* $\Delta(\theta) = -\eta(\theta)B^{-1}(\theta)$
2. *Compute*
   - *(2.1)* $\tilde{\Delta}(\theta) = A(\theta)\Delta(\theta - \omega)B^{-1}(\theta) + \Delta(\theta)$
   - *(2.2)* $\tilde{A}(\theta) = A(\theta)A(\theta - 2\omega)$
   - *(2.3)* $\tilde{B}^{-1}(\theta) = B^{-1}(\theta - 2\omega)B^{-1}(\theta)$
3. *Set*

$$\tilde{\Delta} \to \Delta$$
$$\tilde{A} \to A$$
$$\tilde{B} \to B$$
$$2\omega \to \omega$$

4. *Iterate parts 2–3*

This algorithm gives us $\Delta(\theta + \omega)$. Shifting it by $-\omega$ we get $\Delta(\theta)$.

2.8 Computation of rank-1 whiskers of an invariant torus

In this section we present algorithms to compute the whiskers associated to an invariant torus, that is the invariant manifolds that contain the torus and are tangent to the invariant bundles.

For the sake of simplicity and in order to state in a clear way the main idea behind the methods we will discuss the case when the invariant whiskers are one-dimensional, that is $d - \ell = 1$ and the bundle is trivial. However, this idea can be extended to compute invariant manifolds of any rank. We plan to come back to this issue in the future.

Recall that we will look for the whiskers by finding a parameterization for them, so we will seek a function $W : \mathbb{T}^\ell \times (U \subset \mathbb{R}^{d-\ell}) \to \mathbb{R}^{2d}$ and a scalar λ satisfying equation (2.8).

Notice that we are assuming that the dynamics on the manifold can be reduced to constant coefficients.

We will consider three different methods to solve equation (2.8). We will first discuss the order by order method, which has the main disadvantage that one needs to store and invert a full matrix so it has several shortcomings. The other two methods considered, are based on the philosophy of quasi-Newton methods and by using the phenomenon of " automatic reducibility" we are able to design a very efficient Newton method. The first method allows to compute simultaneously the invariant tori and the whiskers, whereas the second one is just a simplified version of the first one, when one already knows the invariant tori and the tangent bundle.

We considered only the case of maps because the same ideas work for the case of vector fields.

Similar algorithms for quasi-periodic maps were developed and implemented in [HL06a, HL07]

2.8.1 The order by order method

In this section we follow [CFL05]. We refer the reader to [CFL05] for the proof of the convergence of the series presented.

We will find a solution (W, λ) for the invariance equation (2.8), by discretizing it in Fourier-Taylor series. Hence, we will look for W as a power series

$$W(\theta, s) = \sum_{n=0}^{\infty} W_n(\theta) s^n, \tag{2.76}$$

and match similar coefficients in s^n on both sides of equation (2.8).

For $n = 0$, one obtains

$$F(W_0(\theta)) = W_0(\theta + \omega), \tag{2.77}$$

which is equation (2.4) for the invariant torus. Therefore, we have $W_0(\theta) = K(\theta)$, where K is a parametrization of the invariant torus.

For $n = 1$, we obtain

$$DF \circ K(\theta) W_1(\theta) = W_1(\theta + \omega)\lambda, \tag{2.78}$$

which tells us that $W_1(\theta)$ is an eigenfunction with eigenvalue λ of the operator $\mathcal{M}_1(\theta)$ defined in (2.30)

Equation (2.78) for W_1 states that the bundle spanned by W_1 is invariant for the linearization of F, and the dynamics on it is reduced to a contraction/expansion by a constant λ. Therefore, one can compute W_1 and λ using the algorithms given in Section 2.6.5.

Remark 2.8.1. Notice that if $W_1(\theta)$ is a solution of equation (2.78), then $bW_1(\theta)$, for any $b \in \mathbb{R}$, is also a solution. Even though all the choices of $W_1(\theta)$ are mathematically equivalent, the choice affects the numerical properties of the algorithm.

For $n \geq 2$, we obtain

$$DF \circ K(\theta) W_n(\theta) = W_n(\theta + \omega)\lambda^n + R_n(W_0, \ldots, W_{n-1}), \tag{2.79}$$

where R_n is an explicit polynomial in $W_0, \ldots, W_{n-1}$ whose coefficients are derivatives of F evaluated at $W_0 = K$.

Equation (2.79) can be solved provided that λ satisfies a certain non-resonance condition, which is that λ^n is not in the spectrum of the operator $\mathcal{M}_1(\theta)$. This condition is clearly satisfied in the case of (un)stable bundles which are 1-dimensional.

Equation (2.79), for $n \geq 2$ can be solved using the large matrix method explained in Section 2.4.1. Hence, we will discretize the equation (2.79) using Fourier series and reduce the problem to solve a linear system in Fourier space, where the unknowns are the Fourier coefficients of the matrix W_n.

Notice that once $W_0(\theta)$ and $W_1(\theta)$ are fixed, the solution $W_n(\theta)$ for $n \geq 2$ of equation (2.79) is uniquely determined.

Finally, if the non resonance condition is satisfied, we know from [CFL05] that the series constructed here converges to a true analytic solution of the problem.

2.8.2 A Newton method to compute simultaneously the invariant torus and the whiskers

Instead of solving equation (2.8) step by step we can construct a Newton method for it. Hence, we start with an initial approximation (W, λ) (resp. (W, λ, μ) satisfying

$$\begin{aligned} &F(W(\theta, s)) - W(\theta + \omega, \lambda s) = E(\theta, s) \\ &F(W(\theta, s)) - W(\theta + \omega, \lambda s) - (JDK_0 \circ T_\omega)(\theta)\mu = E(\theta, s) \end{aligned} \tag{2.80}$$

and we look for an improved solution

$$\begin{aligned} W &\leftarrow W + \Delta \\ \lambda &\leftarrow \lambda + \delta \\ \mu &\leftarrow \mu + \delta_\mu \end{aligned}$$

by solving the linearized equation

$$\begin{aligned} &[DF \circ W(\theta, s)]\Delta(\theta, s) - \Delta(\theta + \omega, \lambda s) - \delta s \partial_s W(\theta + \omega, \lambda s) \\ &\quad - JDK_0(\theta + \omega)\delta_\mu = -E(\theta, s). \end{aligned} \tag{2.81}$$

Remark 2.8.2. As in the previous cases the role of the parameter μ is to allow us to adjust some averages to solve the equations for the case $n = 0$.

We will try to solve equation (2.81) by discretizing it in Fourier-Taylor series. Notice that equation (2.81) is not diagonal when discretized in Fourier-Taylor series because of the term $DF \circ W$. However, there is a way to make it diagonal if we notice the following trick:

Consider the expression (2.80) and apply the ∂_θ and ∂_s operators so that we obtain

$$\begin{aligned} DF \circ W(\theta, s)\partial_\theta W(\theta, s) &= \partial_\theta W(\theta + \omega, \lambda s) + O(E) \\ DF \circ W(\theta, s)\partial_s W(\theta, s) &= \lambda \partial_s W(\theta + \omega, \lambda s) + O(E) \end{aligned}$$

From where we read that, up to a certain error, the vector $\partial_\theta W(\theta, s)$ remains invariant under $DF \circ W(\theta, s)$, whereas the vector $\partial_s W(\theta, s)$ is multiplied by a factor λ.

The vectors $(J \circ K_0(\theta))\partial_\theta W(\theta, s)N(\theta, s)$ and $(J \circ K_0(\theta))\partial_s W(\theta, s)\tilde{N}(\theta, s)$, where N and $\tilde{N}$ are some normalization matrices, are the symplectic conjugate vectors of $\partial_\theta W(\theta, s)$ and $\partial_s W(\theta, s)$, respectively. The preservation of the symplectic structure implies that

$$\begin{aligned} &(DF \circ W(\theta, s))(J \circ K_0(\theta))\partial_\theta W(\theta, s)N(\theta, s) = \\ &\qquad J \circ K_0(\theta)\partial_\theta W(\theta, s)N(\theta, s) + A(\theta, s)\partial_\theta W(\theta, s) + O(E) \\ &(DF \circ W(\theta, s))(J \circ K_0(\theta))\partial_s W(\theta, s)\tilde{N}(\theta, s) = \\ &\qquad \frac{1}{\lambda} J \circ K_0(\theta)\partial_s W(\theta, s)\tilde{N}(\theta, s) + B(\theta, s)\partial_s W(\theta, s) + O(E) \end{aligned}$$

where $A(\theta, s)$ and $B(\theta, s)$ represent the twist.

We summarize these properties in the following proposition:

Proposition 2.8.3. *Denote*

$$\begin{aligned}
\alpha(\theta, s) &= \partial_\theta W(\theta, s) \\
\beta(\theta, s) &= \partial_s W(\theta, s) \\
N(\theta, s) &= (\alpha(\theta, s)^t \alpha(\theta, s))^{-1} \\
\tilde{N}(\theta, s) &= (\beta(\theta, s)^t \beta(\theta, s))^{-1} \\
\gamma(\theta, s) &= (J \circ K_0(\theta))\alpha(\theta, s) N(\theta, s) \\
\eta(\theta, s) &= (J \circ K_0(\theta))\beta(\theta, s) \tilde{N}(\theta, s)
\end{aligned}$$

and form the following matrix

$$M(\theta, s) = [\alpha(\theta, s) \mid \gamma(\theta, s) \mid \beta(\theta, s) \mid \eta(\theta, s)] \tag{2.82}$$

where we denote by $[\cdot \mid \cdot \mid \cdot \mid \cdot]$ *the* $2d \times 2d$ *matrix obtained by juxtaposing the two* $2d \times \ell$ *matrices and the two* $2d \times (d - \ell)$ *matrices that are in the arguments.*

Then

$$DF \circ W(\theta, s) M(\theta, s) = M(\theta + \omega, \lambda s) \left(\begin{array}{cc|cc} \mathrm{Id} & A(\theta, s) & & \\ 0 & \mathrm{Id} & \multicolumn{2}{c}{\bigcirc} \\ \hline & & \lambda & B(\theta, s) \\ \multicolumn{2}{c|}{\bigcirc} & 0 & 1/\lambda \end{array}\right) + O(E)$$

where $A(\theta, s)$ *and* $B(\theta, s)$ *have an explicit expression that we omit here and* E *is the error in* (2.80).

Now if we change the unknowns $\Delta = MV$ in (2.81) and multiply by $M^{-1}(\theta + \omega, \lambda s)$ on the LHS, by Proposition 2.8.3 we are left with the following system of equations

$$\begin{aligned}
&\left(\begin{array}{cc|cc} \mathrm{Id} & A(\theta, s) & & \\ 0 & \mathrm{Id} & \multicolumn{2}{c}{\bigcirc} \\ \hline & & \lambda & B(\theta, \sigma) \\ \multicolumn{2}{c|}{\bigcirc} & 0 & 1/\lambda \end{array}\right) V(\theta, s) - V(\theta + \omega, \lambda s) - \tilde{C}(\theta, s)\delta_u \\
&\qquad = -\widetilde{E}(\theta, s) + s\delta H(\theta, s)
\end{aligned} \tag{2.83}$$

where

$$\begin{aligned}
\tilde{C}(\theta, s) &= M^{-1}(\theta + \omega, \lambda s) J D K_0(\theta) \\
\widetilde{E}(\theta, s) &= M^{-1}(\theta + \omega, \lambda s) E(\theta, s) \\
H(\theta, s) &= M^{-1}(\theta + \omega, \lambda s) \partial_s W(\theta + \omega, \lambda s)
\end{aligned}$$

Now, we expand the terms in (2.83) as a power series up to some order L and match coefficients of the same order on both sides of the equation. We use subindexes to denote coordinates and superindexes to denote the order. Hence, for order s^0 we have

$$V_1^0(\theta) - V_1^0(\theta+\omega) + A^0(\theta)V_2^0(\theta) - C_1^0(\theta)\delta_\mu = -\widetilde{E}_1^0(\theta) \tag{2.84}$$
$$V_2^0(\theta) - V_2^0(\theta+\omega) - C_2^0(\theta)\delta_\mu = -\widetilde{E}_2^0(\theta) \tag{2.85}$$
$$\lambda V_3^0(\theta) - V_3^0(\theta+\omega) + B^0(\theta)V_4^0(\theta) - C_3^0(\theta)\delta_\mu = -\widetilde{E}_3^0(\theta) \tag{2.86}$$
$$\frac{1}{\lambda}V_4^0(\theta) - V_4^0(\theta+\omega) - C_4^0(\theta)\delta_\mu = -\widetilde{E}_4^0(\theta) \tag{2.87}$$

Notice that (2.84) and (2.85) can be solved using the algorithm 2.5.3 described in section 2.5. Hence, we determine V_1^0, V_2^0 and δ_μ. Once we know δ_μ we can solve uniquely for V_3^0 and V_4^0 equations (2.86) and (2.87). These equations do not have any small divisors nor obstructions.

For order s^1 we have

$$\begin{aligned} V_1^1(\theta) - \lambda V_1^1(\theta+\omega) + A^0(\theta)V_2^1(\theta) + A^1(\theta)V_2^0(\theta) \\ = -\widetilde{E}_1^1(\theta) + \delta H_1^0(\theta) + \delta_\mu C_1^1(\theta) \end{aligned} \tag{2.88}$$
$$V_2^1(\theta) - \lambda V_2^1(\theta+\omega) = -\widetilde{E}_2^1(\theta) + \delta H_2^0(\theta) + \delta_\mu C_2^1(\theta) \tag{2.89}$$
$$\begin{aligned} \lambda V_3^1(\theta) - \lambda V_3^1(\theta+\omega) + B^0(\theta)V_4^1(\theta) + B^1(\theta)V_4^0(\theta) \\ = -\widetilde{E}_3^1(\theta) + \delta H_3^0(\theta) + \delta_\mu C_3^1(\theta) \end{aligned} \tag{2.90}$$
$$\frac{1}{\lambda}V_4^1(\theta) - \lambda V_4^1(\theta+\omega) = -\widetilde{E}_4^1(\theta) + \delta H_4^0(\theta) + \delta_\mu C_4^1(\theta) \tag{2.91}$$

Notice that once we choose δ, equations (2.88) and (2.89) are uniquely solvable for V_1^1 and V_2^1. Recall that δ_μ is known because it has been computed in the case of order 0 equations.

The same happens with equation (2.91). However, equation (2.90) has an average that depends on $V_4^1(\theta)$. In order to overcome this problem we perform the following trick. We denote by F and G the solution of

$$\frac{1}{\lambda}F(\theta) - \lambda F(\theta+\omega) = H_4^0(\theta)$$
$$\frac{1}{\lambda}G(\theta) - \lambda G(\theta+\omega) = D_4^1(\theta)$$

where

$$D_4^1(\theta) = -\widetilde{E}_4^1(\theta) + \delta_\mu C_4^1(\theta)$$

then

$$V_4^1(\theta) = \delta F(\theta) + G(\theta)$$

Taking averages of the equation (2.90) we have that

$$\overline{D_4^1} + \delta\overline{H_3^0} - \overline{B^0F}\delta - \overline{B^0G} - \overline{B^1V_4^0} = 0,$$

so we can solve for δ provided that $\overline{H_3^0} - \overline{B^0F} \neq 0$.

Now the other orders do not have any problem. For s^n, with $n \geq 2$, we have

$$\begin{aligned}
&V_1^n(\theta) - \lambda^n V_1^n(\theta+\omega) + \sum_{k=0}^{n} A^{n-k}(\theta) V_2^k(\theta) = -\widetilde{E}_1^n(\theta) + \delta H_1^{n-1}(\theta) + \delta_\mu C_1^n(\theta) \\
&V_2^n(\theta) - \lambda^n V_2^n(\theta+\omega) = -\widetilde{E}_2^n(\theta) + \delta H_2^{n-1}(\theta) + \delta_\mu C_2^n(\theta) \\
&\lambda V_3^n(\theta) - \lambda^n V_3^n(\theta+\omega) + \sum_{k=0}^{n} B^{n-k}(\theta) V_4^k(\theta) = -\widetilde{E}_3^n(\theta) + \delta H_3^{n-1}(\theta) + \delta_\mu C_3^n(\theta) \\
&\frac{1}{\lambda} V_4^n(\theta) - \lambda^n V_4^n(\theta+\omega) = -\widetilde{E}_4^n(\theta) + \delta H_4^{n-1}(\theta) + \delta_\mu C_4^n(\theta)
\end{aligned} \tag{2.92}$$

and equations (2.92) can be solved uniquely for V_1^n, V_2^n, V_3^n and V_4^n, for $n = 2 \dots L$, where L is the fixed degree for the Taylor expansion. Hence, we have obtained $\delta, \delta_\mu \in \mathbb{R}$ and

$$V(\theta, s) = \sum_{n=0}^{L} V^n(\theta) s^n,$$

so that the improved solution is

$$\begin{aligned}
W &\leftarrow W + MV \\
\lambda &\leftarrow \lambda + \delta \\
\mu &\leftarrow \mu + \delta_\mu.
\end{aligned}$$

The algorithm for the simultaneous computation of the whiskers and the invariant torus is

Algorithm 2.8.4. *Consider given F, ω, k_0 and a fixed order L. Given an approximate solution (W, λ, μ), perform the following calculations*

1. *Compute $E(\theta, s) = F \circ W(\theta, s) - W(\theta + \omega, \lambda s) - ((J \circ K) DK_0(\theta + \omega)\mu$*
2. *Compute*
   - *(2.1)* $\alpha(\theta, s) = \partial_\theta W(\theta, s)$
   - *(2.2)* $\beta(\theta, s) = \partial_s W(\theta, s)$
   - *(2.3)* $N(\theta, s) = [\alpha(\theta, s)^t \alpha(\theta, s)]^{-1}$
   - *(2.4)* $\tilde{N}(\theta, s) = [\beta(\theta, s)^t \beta(\theta, s)]^{-1}$
   - *(2.5)* $\gamma(\theta, s) = (J \circ K)\alpha(\theta, s) N(\theta, s)$
   - *(2.6)* $\eta(\theta, s) = (J \circ K)\beta(\theta, s)\tilde{N}(\theta, s)$
   - *(2.7)* $M(\theta, s) = [\alpha(\theta, s) \mid \gamma(\theta, s) \mid \beta(\theta, s) \mid \eta(\theta, s)]$
   - *(2.8)* $[M(\theta, s)]^{-1}$
3. *Compute*

(3.1) $\tilde{C}(\theta, s) = M^{-1}(\theta + \omega, \lambda s) J D K_0(\theta + \omega)$

(3.2) $\widetilde{E}(\theta, s) = M^{-1}(\theta + \omega, \lambda s) E(\theta, s)$

(3.3) $H(\theta, s) = M^{-1}(\theta + \omega, \lambda s) \alpha(\theta + \omega, \lambda s)$

4. Compute

(4.1) $A(\theta, s) = [\alpha(\theta, s) N(\theta, s)]^T [DF \circ W(\theta, s)] \gamma(\theta, s)$

(4.2) $B(\theta, s) = [\beta(\theta, s) \tilde{N}(\theta, s)]^T [DF \circ W(\theta, s)] \eta(\theta, s)$

5.(5.1) Solve for δ_μ *satisfying*

$$\int \widetilde{E}_2^0 - \left[\int C_2^0 \right] \delta_\mu = 0$$

(5.2) Solve for V_2^0 *satisfying*

$$V_2^0 - V_2^0 \circ T_\omega = -\widetilde{E}_2^0 + C_2^0 \delta_\mu$$

Set V_2^0 *such that the average is* 0.

6.(6.1) Compute $A^0(\theta) V_2^0(\theta)$

(6.2) Solve for $\bar{V}_2^0$ *satisfying*

$$\int \widetilde{E}_1^0 - \int C_1^0(\theta) \delta_\mu + \int A^0 V_2^0 + \left[\int A^0 \right] \bar{V}_2^0 = 0$$

(6.3) Set $V_2^0 = V_2^0 + \bar{V}_2^0$

(6.4) Solve for $\bar{V}_1^0$ *satisfying*

$$V_1^0 - V_1^0 \circ T_\omega = -\widetilde{E}_1^0 - A^0 V_2^0 + C_1^0 \delta_\mu$$

(6.5) Normalize so that $\int V_1^0 = 0$

7. Solve for V_4^0 *satisfying*

$$\frac{1}{\lambda} V_4^0 - V_4^0 \circ T_\omega = -\widetilde{E}_4^0 + C_4^0 \delta_\mu$$

8. Solve for V_4^0 *satisfying*

$$\lambda V_3^0 - V_3^0 \circ T_\omega = -\widetilde{E}_3^0 + C_3^0 \delta_\mu - B^0 V_4^0$$

9.(9.1) Solve for F *satisfying*

$$\frac{1}{\lambda} F - \lambda F \circ T_\omega = H_4^0$$

(9.2) Solve for G *satisfying*

$$\frac{1}{\lambda} G - \lambda G \circ T_\omega = -\widetilde{E}_4^1 + \delta_\mu C_4^1$$

(9.3) Solve for δ satisfying

$$\left(-\overline{\widetilde{E}_4^1} + \delta_\mu \overline{C_4^1} - \overline{\overline{B^0 G} - \overline{B^1 V_4^0}}\right) + \delta(\overline{H_3^0} - \overline{\overline{B^0 F}}) = 0$$

(9.4) Set $V_4^1 = \delta F + G$

10. *(10.1) Solve for* V_3^1 *satisfying*

$$\lambda V_3^1 - \lambda V_3^1 \circ T_\omega = -\widetilde{E}_3^1 + \delta H_3^0 + \delta_\mu C_3^1 - B^0 V_4^1 + B^1 V_4^0$$

(10.2) Normalize so that $\int V_3^1 = 0$

(10.3) Solve for V_2^1 *satisfying*

$$V_2^1 - \lambda V_2^1 \circ T_\omega = -\widetilde{E}_2^1 + \delta H_2^0 + \delta_\mu C_2^1$$

(10.4) Solve for V_1^1 *satisfying*

$$V_1^1 - \lambda V_1^1 \circ T_\omega = -\widetilde{E}_1^1 + \delta H_1^0 + \delta_\mu C_1^1 - A^0 V_2^1 + A^1 V_2^0$$

11. *For* $n = 2 \ldots L$ *do*

(11.1) Solve for V_2^n *satisfying*

$$V_2^n - \lambda^n V_2^n \circ T_\omega = -\widetilde{E}_2^n(\theta) + \delta H_2^{n-1} + \delta_\mu C_2^n$$

(11.2) Compute

$$\tilde{A}^n = \sum_{k=0}^{n} A^{n-k} V_2^k$$

(11.3) Solve for V_1^n *satisfying*

$$V_1^n - \lambda^n V_1^n \circ T_\omega = -\widetilde{E}_1^n + \delta H_1^{n-1} + \delta_\mu C_1^n - \tilde{A}^n$$

(11.4) Solve for V_4^n *satisfying*

$$\frac{1}{\lambda} V_4^n - \lambda^n V_4^n \circ T_\omega = -\widetilde{E}_4^n + \delta H_4^{n-1} + \delta_\mu C_4^n$$

(11.5) Compute

$$\tilde{B}^n = \sum_{k=0}^{n} B^{n-k} V_4^k$$

(11.6) Solve for V_3^n *satisfying*

$$\lambda V_3^n - \lambda^n V_3^n \circ T_\omega = -\widetilde{E}_3^n + \delta H_3^{n-1} + \delta_\mu C_3^n - \tilde{B}^n$$

12. *Compute*

$$V(\theta) = \sum_{n=0}^{L} V^n(\theta) S^n$$

13. *Set* $W \leftarrow W + MV$

$\lambda \leftarrow \lambda + \delta$

$\mu \leftarrow \mu + \delta_\mu$

2.8.3 A Newton method to compute the whiskers

Assuming that we have computed exactly an invariant tori $K(\theta)$ with the associated stable bundle $V^s(\theta)$ (resp. unstable bundle $V^u(\theta)$) and the rate of contraction (resp. expansion) λ, we can use a Newton method to compute the whiskers.

We consider the invariance equation (2.8), and we assume that we have an initial approximation W for the whiskers that expressed as a power series

$$W(\theta, s) = \sum_{n=0}^{\infty} W^n(\theta) s^n$$

is such that

$$W^0(\theta) = K(\theta) \quad \text{and} \quad W^1(\theta) = V^s(\theta)$$

(the case unstable is analogous).

Then, it is clear that the error E for the initial approximation W is such that

$$E(\theta, s) = \sum_{n \geq 2} E^n(\theta) s^n,$$

because the approximation is exact for the terms of order 0 and 1.

For a given function $G : \mathbb{T}^\ell \times \mathbb{R} \to \mathbb{R}^{2d}$ we denote

$$G(\theta, s) = G^{[<L]}(\theta, s) + G^{[\geq L]}(\theta, s)$$

where

$$G^{[<L]}(\theta, s) = \sum_{n=0}^{L-1} G^n(\theta) s^n$$

and

$$G^{[\geq L]}(\theta, s) = \sum_{n=L}^{\infty} G^n(\theta) s^n$$

Using this notation, the linearized equation for the Newton method is,

$$[DF \circ W(\theta, s)] \Delta^{[\geq 2]}(\theta, s) - \Delta^{[\geq 2]}(\theta + \omega, \lambda s) = -E^{[\geq 2]}(\theta, s).$$

Similarly, as we did in the previous section, we can perform the change of coordinates defined in Proposition 2.8.3 and, following the same notation, reduce the problem to solve an equation which is diagonal in Fourier-Taylor series, that is

$$M(\theta, s) V^{[\geq 2]}(\theta, s) - V^{[\geq 2]}(\theta + \omega, \lambda s) = -\widetilde{E}^{[\geq 2]}(\theta, s).$$

Notice that in this case, we do not have to solve the system of equations for order 0 and 1 and we can go straight to order $n \geq 2$. We use subindexes to denote coordinates

and superindexes to denote the order. Hence, for order $n \geq 2$ we need to solve the system of equations

$$\begin{aligned} &V_1^n(\theta) - \lambda^n V_1^n(\theta+\omega) + \sum_{k=2}^{n} A^{n-k}(\theta) V_2^k(\theta) = -\widetilde{E}_1^n(\theta) \\ &V_2^n(\theta) - \lambda^n V_2^n(\theta+\omega) = -\widetilde{E}_2^n(\theta) \\ &\lambda V_3^n(\theta) - \lambda^n V_3^n(\theta+\omega) + \sum_{k=2}^{n} B^{n-k}(\theta) V_4^k(\theta) = -\widetilde{E}_3^n \\ &\frac{1}{\lambda} V_4^n(\theta) - \lambda^n V_4^n(\theta+\omega) = -\widetilde{E}_4^n \end{aligned} \tag{2.93}$$

Notice that now the solution of (2.93) for $n = 2, 3$ provides an exact solution of the invariance equation up to order 4. That is, if we set

$$V^{[<4]}(\theta, s) = V^2(\theta, s) + V^3(\theta, s)$$

where V^2 and V^3 are obtained by solving equations (2.93), then the improved solution $\bar{W}$ given by

$$\bar{W}(\theta, s) = W(\theta, s) + M(\theta, s) V^{[<4]}(\theta, s),$$

satisfies that it approximates the solution of the invariance equations with an error $\bar{E}$ such that

$$\bar{E}(\theta, s) = \bar{E}^{[\geq 4]}(\theta, s).$$

This process can be iterated and at each step we solve the invariance equation exactly up to an order which is the double of the one we had for the initial approximation. Thus, if we assume that the initial guess W is such that the error in (2.80) satisfies that

$$E = E^{[\geq L]},$$

then the modified linearized equation for the Newton method is such that

$$M(\theta, s) V^{[\geq L]}(\theta, s) - V^{[\geq L]}(\theta + \omega, \lambda s) = -\widetilde{E}^{[\geq L]}(\theta, s)$$

and if we solve the system of equations (2.93) for $n = L \ldots (2L-1)$ then the improved $\bar{W}$ is

$$\bar{W}(\theta, s) = W(\theta, s) + M(\theta, s) V^{[<2L]}(\theta, s)$$

and is such that the new error $\bar{E}$ satisfies $\bar{E}(\theta, s) = \bar{E}^{[\geq 2L]}(\theta, s)$.

The algorithm in this case is

Algorithm 2.8.5. *Given F, ω as well as K, V^s, λ and an approximate solution W such that*

$$F \circ W(\theta, s) - W(\theta + \omega, \lambda s) = E^{[\geq L]}$$

with $L \geq 2$ and $W(\theta, 0) = K(\theta)$ and $\partial_s W(\theta, s) = V^s(\theta)$, perform the following calculations:

1. Compute $E^{[\geq L]} = F \circ W(\theta, s) - W(\theta + \omega, \lambda s)$

2. Compute

(2.1) $\alpha(\theta, s) = \partial_\theta W(\theta, s)$

(2.2) $\beta(\theta, s) = \partial_s W(\theta, s)$

(2.3) $N(\theta, s) = [\alpha(\theta, s)^t \alpha(\theta, s)]^{-1}$

(2.4) $\tilde{N}(\theta, s) = [\beta(\theta, s)^t \beta(\theta, s)]^{-1}$

(2.5) $\gamma(\theta, s) = (J \circ K)\alpha(\theta, s)N(\theta, s)$

(2.6) $\eta(\theta, s) = (J \circ K)\beta(\theta, s)\tilde{N}(\theta, s)$

(2.7) $M(\theta, s) = [\alpha(\theta, s) \mid \gamma(\theta, s) \mid \beta(\theta, s) \mid \eta(\theta, s)]$

(2.8) $[M(\theta, s)]^{-1}$

3. Compute

$$\widetilde{E}^{[\geq L]}(\theta, s) = M^{-1}(\theta + \omega, \lambda s)E^{[\geq L]}(\theta, s)$$

4. Compute

(4.1) $A(\theta, s) = [\alpha(\theta, s)N(\theta, s)]^T [DF \circ W(\theta, s)]\gamma(\theta, s)$

(4.2) $B(\theta, s) = [\beta(\theta, s)\tilde{N}(\theta, s)]^T [DF \circ W(\theta, s)]\eta(\theta, s)$

5. For $n = L \ldots 2L - 1$ *do*

(5.1) Solve for V_2^n *satisfying*

$$V_2^n - \lambda^n V_2^n \circ T_\omega = -\widetilde{E}_2^n(\theta)$$

(5.2) Compute

$$\tilde{A}^n = \sum_{k=L}^{n} A^{n-k} V_2^k$$

(5.3) Solve for V_1^n *satisfying*

$$V_1^n - \lambda^n V_1^n \circ T_\omega = -\widetilde{E}_1^n - \tilde{A}^n$$

(5.4) Solve for V_4^n *satisfying*

$$\frac{1}{\lambda} V_4^n - \lambda^n V_4^n \circ T_\omega = -\widetilde{E}_4^n$$

(5.5) Compute

$$\tilde{B}^n = \sum_{k=L}^{n} B^{n-k} V_4^k$$

(5.6) Solve for V_3^n satisfying

$$\lambda V_3^n - \lambda^n V_3^n \circ T_\omega = -\widetilde{E}_3^n - \tilde{B}^n$$

6. Compute

$$V(\theta) = \sum_{n=L}^{2L-1} V^n(\theta) s^n$$

7. Set $W \leftarrow W + MV$

2.9 Guesses

All the methods we have discussed are based on considering an initial approximation for the invariant object we are looking for and then using a Newton method to improve this approximation. Therefore, it remains to discuss how to obtain a good initial guess for the Newton method. We will discuss the case of maps. The case for vector fields is analogous.

A very standard method, specially when one studies families of maps indexed by a parameter, is to use a continuation method. Typically, one starts with a value of the parameter, let us say ε_0, for which the map F_{ε_0} can be studied analytically or a solution is known and then use this solution as an initial guess for the Newton method to compute the solution for the map F_ε for $|\varepsilon - \varepsilon_0|$ small enough. An improvement of this method could be to extrapolate an initial guess for F_ε from the previous computed solutions for different values of ε.

A very classical example is when one studies maps which are a perturbation of an integrable one, for which the dynamics is very simple and well known. Nevertheless, we recall that the methods presented in this thesis do not require the map to be close to the integrable case. Therefore, one needs to develop other methods to compute an initial approximation.

An alternative to get an initial approximation for the computation of primary invariant tori is using the Percival variational principle ([Per79]). It is easy to see that primary invariant tori correspond to minimizers of a certain functional $\mathcal{P}$. In this cases, one can find minimizers of this functional using minimization algorithms such as conjugate gradient methods. The problem of this method is that it does not provide smooth solutions and that the precision for the solution will be $C\varepsilon^{1/2}$, where ε is the error for the computation of the functional (roundoff and truncation). Nevertheless, one can use the minimizer obtained as an initial guess for the Newton method.

Another shortcoming of this method is that there is not a Percival variational principle for secondary tori. Moreover, secondary tori are invariant objects that one can not continue to the integrable case, because they are created by the perturbation.

An alternative to get an initial guess to compute secondary tori is to use averaging methods (see the first part of this thesis for an accurate discussion about this technique). Another possibility consists of computing the rotation number for several points and look for a point p that rotates with a frequency which is close to the Diophantine frequency of the invariant tori we are looking for. Of course, taking some iterates $\{F^n(p)\}_{n=0}^{\ell}$ of this point we can obtain an initial guess for the Newton method.

However, we already mentioned in the previous sections that in one step of the Newton method we need to perform some operations in Fourier space. Hence, we need to apply the FFT algorithm and this requires to have the values on an equidistant grid, but this is not the case when one considers iterates of a point p by a map F. In order to get rid of this problem, one can use cubic interpolation and calculate the values of the interpolating function at the points $\{\theta_i\}_{i=0}^{\ell}$, taking into account the periodicity of θ. This is the method we used in the numerical implementation we carried out. See Section 2.10.1.3 for a more detailed description. An algorithm that can perform Fourier Transforms on grids of non equally spaced points is the USFFT and we plan to study it in the future.

2.10 Numerical Examples

We have carried out the numerical implementation of the algorithms described in the previous section. In this section we will discuss some implementation details as well as some efficiency properties and we will show some examples of the results obtained.

The algorithms have been implemented in C language and have been run under the Linux environment. For the computation of the FFT we used the `fftw3` library (see http://www.fftw.org/) and we also used some of the functions available in the LAPACK and BLAS routines (see http://www.netlib.org/lapack/).

2.10.1 Computation of primary and secondary KAM tori for the standard map

The first example that we consider is the very well known standard map introduced by Chirikov [Chi79]. It is a 2D-symplectic map from the cylinder $\mathbb{R} \times \mathbb{T}$ to itself and it is given by

$$\begin{aligned} \bar{p} &= p - \varepsilon V'(q) \\ \bar{q} &= q + \bar{p} \qquad (\text{mod } 1) \end{aligned} \tag{2.94}$$

where ε is a positive parameter and V is a 1-periodic smooth function called the potential given by

$$V(q) = \frac{1}{(2\pi)^2} \cos(2\pi q). \tag{2.95}$$

We refer to (p, q) as the action-angle variables.

2.10.1.1 Computation of primary invariant tori

We started from the integrable case $\varepsilon = 0$ where we have a 1-parameter family of 1-dimensional invariant tori indexed by the frequency of rotation. The celebrated KAM Theorem (see [Lla01a] for a survey) ensures that those invariant tori present in the unperturbed system with a Diophantine frequency will persist under the perturbation for ε up to a certain critical value at which they break down. Regarding this question, there is a very well known method due to Greene ([Gre79]), which consists of looking to the stability of nearby periodic orbits.

In this example, we first considered the invariant curve with a frequency of rotation equal to the golden mean, that is $\omega_g = (\sqrt{5}-1)/2$, which is conjectured to be the most robust curve for the standard map. The Greene's method estimates that for this curve the critical value is close to 0.971635406.

We followed a continuation method, starting from the integrable case $\varepsilon = 0$ and then increasing the parameter ε by steps of size 0.01 up to where our Newton method fails to converge, which turned out to be $\varepsilon = 0.96$. For the computations we have used $N = 2^{11}$ Fourier coefficients and each step of the continuation method takes 0.0305 seconds in average in a Intel(R) Core(TM)2, 2.15 GHz. The errors in the functional equations are smaller than 10^{-10}. We show some of the obtained curves in Figure 2.3.

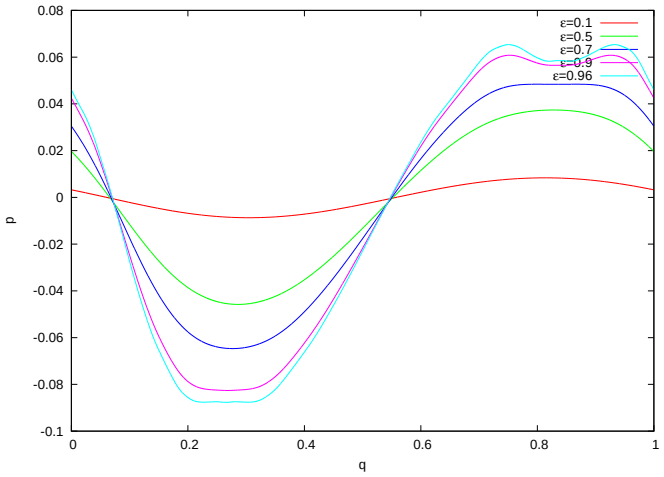

Figure 2.3: The invariant curve associated to the golden mean frequency ω_g for the standard map for some values of ε. Notice that they are shifted to have 0 offset.

We computed also invariant tori corresponding to frequencies of the form

$$\omega_\alpha = \frac{1}{\alpha + \omega_g}, \qquad \alpha \in \mathbb{N}. \tag{2.96}$$

More precisely, we applied our method to the cases $\alpha = 5$ and $\alpha = 50$ and we managed to continue the invariant curve up to $\varepsilon = 0.73$ and $\varepsilon = 0.068$, respectively, with similar time running estimates as in the previous case. In Figure 2.4 we display some of these curves for different values of ε.

2.10.1.2 Standard-like maps

In this study we also considered the case of a system of the form (2.94) but with a potential V which has an infinite number of harmonics. Hence, we chose V satisfying

$$V'(q) = -\frac{1}{1 - 0.9\sin(2\pi q)}, \tag{2.97}$$

and we studied the invariant tori for this system associated to the golden mean frequency.

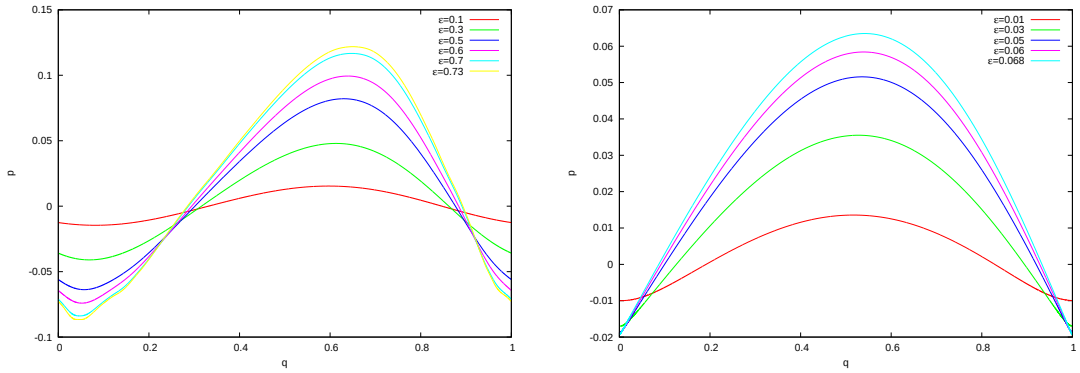

Figure 2.4: The invariant curves associated to the frequencies ω_5 (left) and ω_{50} (right) for the standard map for some values of ε. Notice that they are shifted to have 0 offset.

We computed the invariant curves corresponding to the golden mean frequency using a continuation method with smaller steps (in this case the step size was 0.001) for different values of ε (the program stopped at $\varepsilon = 23 \cdot 10^{-3}$). We used $N = 2^{11}$ Fourier modes and it took 0.0406 seconds to perform one step of the continuation method. In Figure 2.5 we show some of the curves computed for different values of ε.

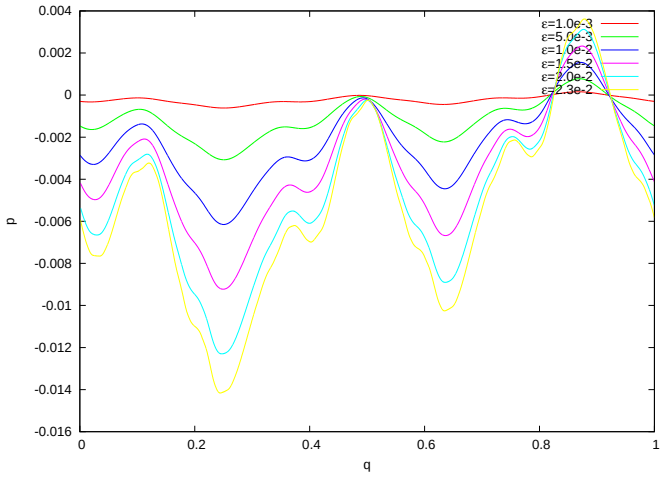

Figure 2.5: The invariant curves associated to the frequency ω_g for the standard-like map with a potential satisfying (2.97) for some values of ε. Notice that they are shifted to have 0 offset.

2.10.1.3 Computation of secondary invariant tori

As we already mentioned, our method can compute also secondary KAM tori. Recall that, in this case, we cannot use a continuation method starting from the integrable case $\varepsilon = 0$ because these tori do not exist in the unperturbed case, but they are created by

the perturbation. Hence, in order to get an initial approximation we will use some of the techniques described in Section 2.9.

Another thing that we need to take into account are the periodicities of the functions that appear in the Newton procedure. Thus, the matrix I introduced in (2.6) has rank 0.

Let us recall, first, that for $\varepsilon \neq 0$ the standard map has two fixed points corresponding to $(0,0)$ and $(0,1/2)$ and, with a simple linear stability analysis, it is easy to see that if ε is small they are hyperbolic and elliptic fixed points, respectively. Moreover, in the neighborhood of the elliptic fix point delimited by the stable and unstable manifolds of the hyperbolic fix point, there are born a family of invariant curves of different topology of the ones that existed in the unperturbed case, indeed, they are contractible to a point. They are the secondary KAM tori.

In order to get an initial approximation for these tori, we will first compute the rotation number for several points on the axis $p = 0$ between the hyperbolic and the elliptic fixed points. Recall that the rotation number provides information about how much turns in average every iterate of the standard map. We use a method by C. Simó that computes an upper and a lower bound for the rotation number. For the sake of completeness we include it here. It is summarized as follows:

- Compute n iterates of a point x^0 on the axis $p = 0$. For each iterate x^i, $i = 0, \ldots, n$, let us compute the number of turns n_i and the angle θ_i (modulus 2π) that form the points x^0, x^* and x^i, where x^* is the elliptic fixed point.

- Sort the angles θ_i in increasing order and keep the information by arranging the indexes i of the angles (we used the "quicksort" algorithm).

- Take two consecutive indexes i and j in the arrangement and perform the following computation,

  - If $i < j \Rightarrow \rho > \frac{n_j - n_i}{j - i}$.
  - If $i > j \Rightarrow \rho < \frac{n_i - n_j}{i - j}$.

Hence, we can obtain $\rho_{\min}$ and $\rho_{\max}$ such that $\rho_{\min} \leq \rho \leq \rho_{\max}$. Then we can approximate the rotation number ρ, either by $\rho_{\min}$ or $\rho_{\max}$ with an error of order $1/n^2$, where n is the number of iterates considered.

In Figure 2.6 we have plotted the computed rotation numbers with 2000 iterates for $\varepsilon = 0.1$ and $\varepsilon = 0.5$. In the cases where $\rho_{\min}$ is bigger than $\rho_{\max}$, it means that the invariant curve has been destroyed.

We fixed a Diophantine frequency in the interval of allowed frequencies and using a bisection method we have found a point x^0 on the axis $p = 0$ which rotates at a frequency which is close to the Diophantine frequency. Then, we computed several iterates of this point and interpolating using splines, we have obtained an approximation of the parameterization of secondary invariant tori evaluated on an equidistant grid. We have used the cubic interpolation routines `spline` and `seval` from [FMM77] taking into account the periodicity of θ.

In particular, we computed the invariant tori for the frequencies $3/40\omega_g \approx 0.04635026$ and $0.18\omega_g \approx 0.111246$ starting at $\varepsilon = 0.1$ and $\varepsilon = 0.5$, respectively. Using a continuation

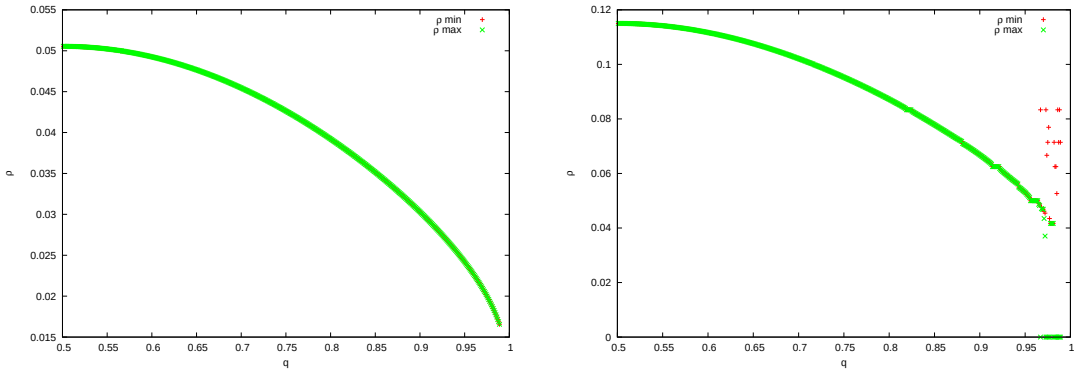

Figure 2.6: The rotation number for different values on the $p = 0$ axis under the standard map (2.94) for $\varepsilon = 0.1$ (left) and $\varepsilon = 0.5$ (right). Notice that the hyperbolic and elliptic fixed points correspond to $(0, 1)$ and $(1/2, 0)$, respectively.

method (with a step size of 0.001) we computed the invariant tori associated to the corresponding frequency up to $\varepsilon = 0.401$ and $\varepsilon = 0.853$, respectively. We used $N = 2^9$ Fourier modes and each step of the continuation method takes 0.01469 seconds in average in an Intel(R) Core(TM)2, 2.15 GHz. In Figure 2.7 we show the computed secondary tori.

Again, as in the case of primary tori the invariant tori are computed with errors in the functional equations (2.4) smaller than 10^{-10}.

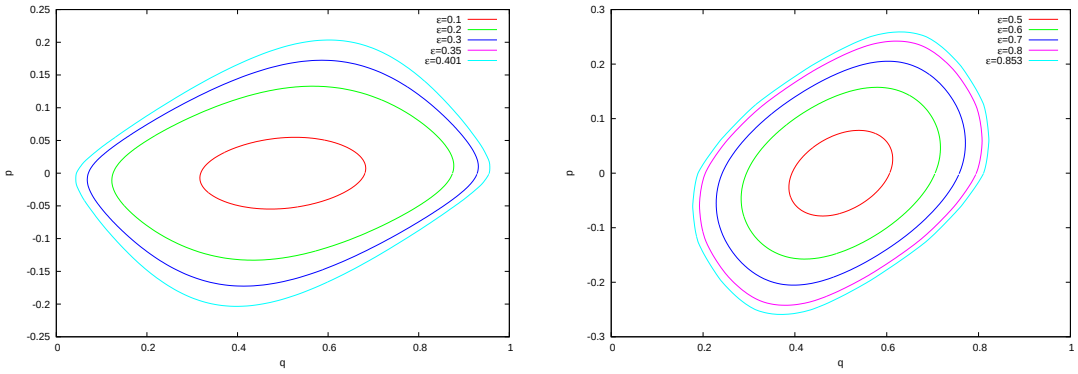

Figure 2.7: The secondary tori associated to the frequencies $3/40\omega_g$ (left) and $0.18\omega_g$ (right) of the standard map for some values of the parameter ε.

2.10.2 4D symplectic maps: The Froeschlé map

In this section we will discuss the computation of maximal and hyperbolic invariant tori for the Froeschlé map. The Froeschlé map is 4D symplectic map defined on $\mathbb{T}^2 \times \mathbb{R}^2$, consisting of two coupled standard maps. It was introduced by Froeshlé in [Fro72] and it

is given by

$$\begin{aligned}
\bar{p_1} &= p_1 - \varepsilon(\frac{\lambda_1}{2\pi}\sin(2\pi q_1) + \frac{\lambda_{12}}{2\pi}\sin(2\pi(q_1+q_2)) \\
\bar{p_2} &= p_2 - \varepsilon(\frac{\lambda_2}{2\pi}\sin(2\pi q_2) + \frac{\lambda_{12}}{2\pi}\sin(2\pi(q_1+q_2)) \\
\bar{q_1} &= q_1 + \bar{p_1} \qquad (\text{mod } 1) \\
\bar{q_2} &= q_2 + \bar{p_2} \qquad (\text{mod } 1)
\end{aligned} \tag{2.98}$$

where λ_{12} is a coupling parameter. The potential for this case is given by

$$V(q_1,q_2) = -\left(\frac{\lambda_1}{(2\pi)^2}\cos(2\pi q_1) + \frac{\lambda_2}{(2\pi)^2}\cos(2\pi q_2) + \frac{\lambda_{12}}{(2\pi)^2}\cos(2\pi(q_1+q_2)\right)$$

Notice that when $\lambda_{12} = 0$, the problem reduces to two uncoupled 2-dimensional standard maps.

For the Froeschlé map, we can consider maximal invariant tori, which are 2-dimensional invariant tori or hyperbolic invariant tori, that is 1-dimensional invariant tori with associated stable and unstable manifolds.

2.10.2.1 Computation of maximal tori

We will follow the behavior of the invariant tori associated to a certain frequency $\omega = (\omega_1, \omega_2)$ as ε increases. Of course, the choice of the frequency vector strongly influences the dynamics. In this study, we will just restrict to one particular case in order to show the way the method works.

In order to get a 2-dimensional frequency vector which is Diophantine, one needs to use some results in number theory. In this example, we considered one of the rotation vectors studied in [CFL04], which is given by

$$\omega_u = \left(\frac{1}{s}, s-1\right) = (0.754877\ldots, 0.324717\ldots) \tag{2.99}$$

where $s = 1.8392\ldots$ is the real root of the polynomial of degree 3,

$$t^3 - t^2 - t - 1 = 0.$$

In the mentioned paper the authors studied tori with other frequencies. We refer also to [HS95] for a study of the breakdown of 2-dimensional tori of the Froeschlé map.

We studied the case when $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 0.001$, starting from $\varepsilon = 0$. Recall that for $\varepsilon = 0$, the system consists of two uncoupled integrable standard maps. Hence, the invariant tori are given by the cross product of the invariant tori of each of the subsystems, which are trivial ($p_i = \text{ct}$, $i = 1, 2$). Using a continuation method we computed the invariant tori associated to the frequency (2.99), for different values of ε up to $\varepsilon = 0.446$. In [CFL04], the breakdown for this torus was estimated for $\varepsilon = 0.55$.

Recall that now we are computing and storing a function K of two variables. For the computations we used $N \times N = 2^8 \times 2^8$ Fourier modes and one step of the continuation method takes on average 4.5 seconds in an Intel(R) Core(TM)2, 2.15 GHz.

We only computed primary KAM tori, hence the matrix I in (2.6) is the Identity. For the case of secondary tori, there are different possibilities: we can take $I = diag(1, 0)$ or $I = diag(0, 1)$, which correspond to invariant tori that in the case of $\lambda_{12} = 0$ (the standard maps are uncoupled) and $\lambda_1 \neq 0$ or $\lambda_2 \neq 0$ consist of the cartesian product of a primary tori of one of the standard maps and a secondary tori of the other standard map. Another possibility is to consider $I = diag(0, 0)$, which in the case of $\lambda_{12} = 0$ (the standard maps are uncoupled) and $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$, the invariant tori consist of the cartesian product of two secondary tori, each one corresponding to a secondary invariant torus of each of the standard maps. Now, of course it is clear how to obtain an initial guess to compute these tori: we start with $\lambda_{12} = 0$ and we obtain the secondary invariant torus for each of the subsystems using the method described in Section 2.10.1.3. Then, we use a continuation method increasing λ_{12}.

In Figure 2.8, the angular components (q_1, q_2) of the invariant tori obtained are drawn and one can see the metamorphoses with respect to the parameter ε.

While the breakdown of invariant curves is rather well understood for the case of 2-D maps, there are very few results concerning higher dimensions. In the future, we plan to polish the programs to study the breakdown of 2-D invariant tori of the Froeschlé map for a wide range of frequencies.

2.10.2.2 Computation of whiskered tori

Recall that whiskered invariant tori for the Froeschlé map are 1-D invariant tori with associated rank-1 stable and unstable manifolds. Starting with $\lambda_2 = 0$, $\lambda_{12} = 0$ and $\lambda_1 = 1$ and ε small (in the computations we started with $\varepsilon = 0.01$), so that the two standard maps are uncoupled, the whiskered tori are given by the cross product of the hyperbolic fixed point of one of the coupled standard maps and a primary invariant torus of the second one. Moreover, the stable and unstable manifolds of the invariant torus are inherited from the ones of the hyperbolic fixed point. Therefore, the invariant torus has constant tangent bundles, that is they are independent of θ.

In the next step, we set $\lambda_2 = 1$ and $\lambda_{12} = 0.1$ and performed a continuation method increasing ε by an step size of 0.01. For the continuation method, the (un)stable bundle computed previously is used as an approximation for the (un)stable bundle of the increased parameter to perform a change of coordinates to the cocycle in order to avoid the *straddle the saddle phenomenon* discussed in Section 2.6.4. The computations performed used $N = 2^{12}$ Fourier modes and reached up to $\varepsilon = 0.87$. Since these tori have rank-1 bundles, we used the algorithms described in Section 2.6.5 and 2.7.2.1. In Figure 2.9 we display the figures obtained. They correspond to the Froeschlé map obtained by performing the change $(p, q) = (-p, -q)$ to the map (2.98).

2.10.3 Some final remarks on the analyticity breakdown

Since we work with truncated Fourier series it may happen that the Newton method converges to a solution of the truncated problem which does not correspond to a truncation of the real solution of the problem. These *spurious* solutions are more abundant and more difficult to detect when the parameters of the problem approach the breakdown.

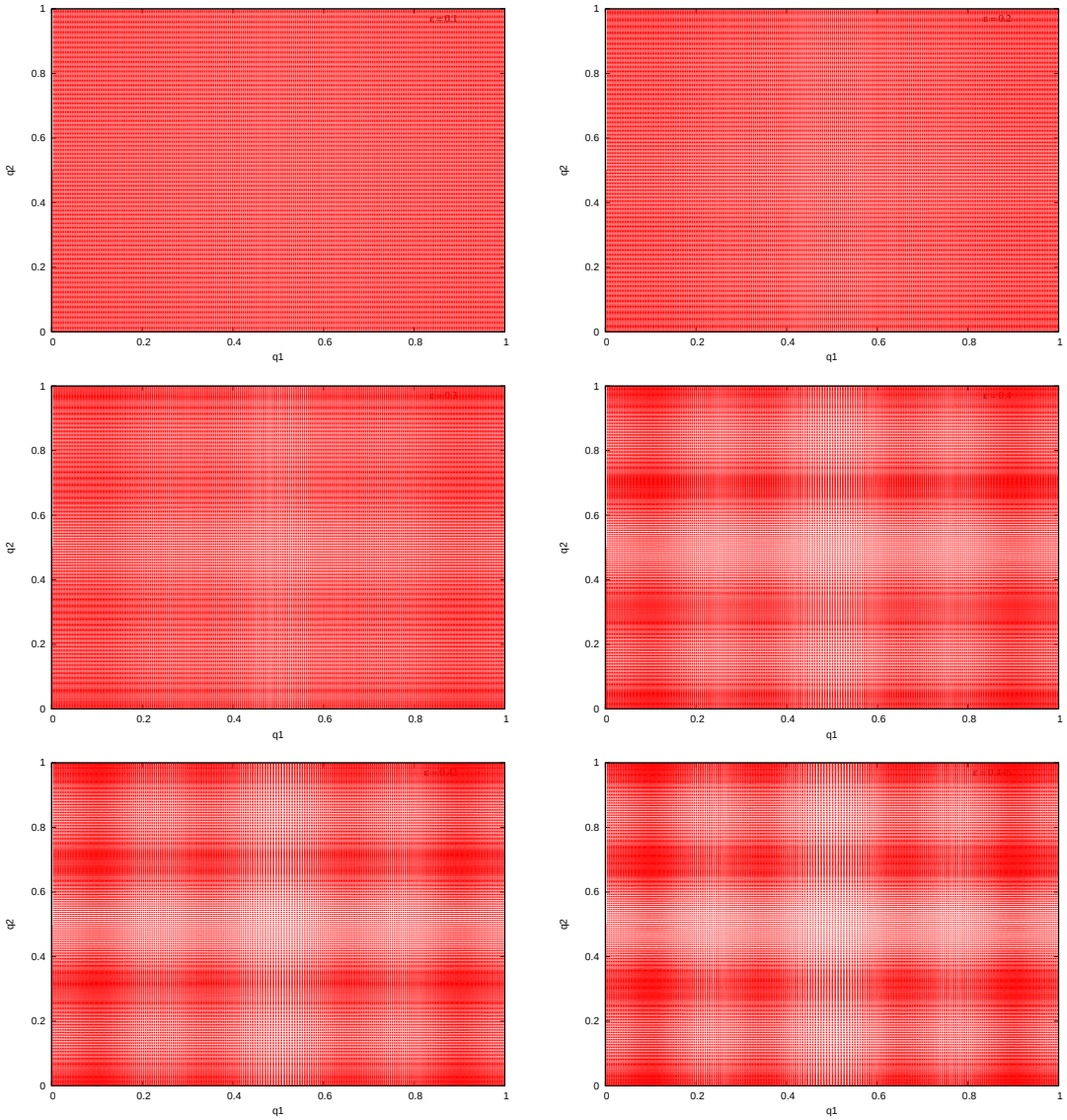

Figure 2.8: The angular variables (q_1, q_2) for the primary maximal tori of the Froeschlé map associated to the frequency ω_u for different values of the parameter ε

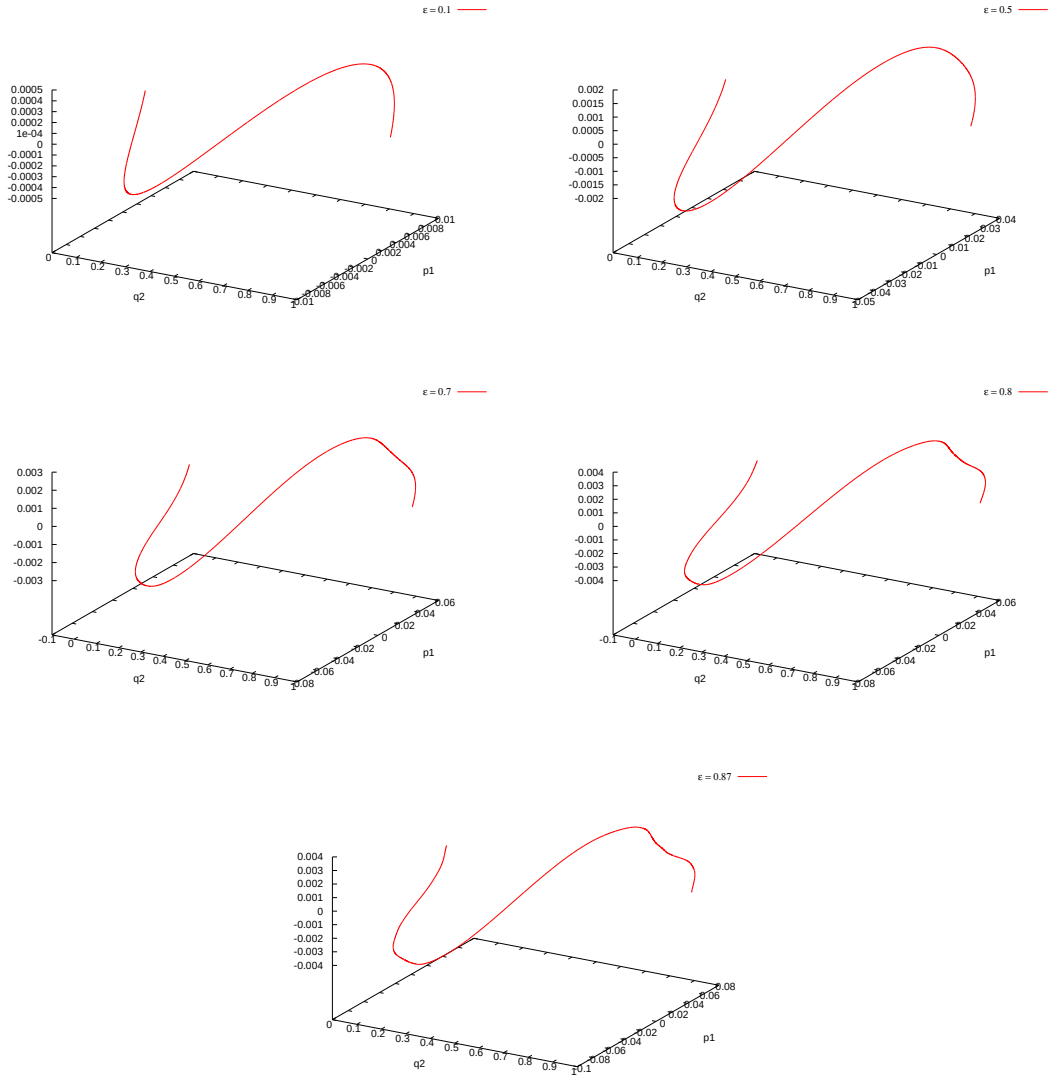

Figure 2.9: The $q_1q_2p_1$ projection of the primary hyperbolic invariant tori of the Froeschlé map associated to the golden mean frequency for different values of the parameter ε.

In the future, we plan to study the breakdown of invariant tori using the methods described in the previous sections. So, it is important to introduce a control to avoid the spurious solutions. One possibility is to use adaptative steps for the continuation method with a certain criterion. More precisely, consider that we have computed an invariant tori for a certain value of the parameter and we want to use it as an initial guess for the next step. First, we shift it by $1/2N$, so that we can obtain the initial guess on a grid of $2N$ points. Then, we apply the Newton method. In case that the method converges, we keep the step size and we move forward. Otherwise, we step back and reduce the step size.

Another quantity that it is important to characterize when we get close to the breakdown are the Sobolev norms, since they become larger. We plan to come to back to this issue in the future research.

Chapter 3

A computational and geometric approach to phase resetting curves and surfaces

3.1 Introduction

The behavior of coupled oscillators in biology and, more intensively, in neuroscience has been the subject of a great deal of recent interest and there is a wide literature on this topic (see [Izh07] for a survey), mainly because many oscillators can be described by their phase variable. Moreover, under generic conditions, the phase of the oscillation can be also defined outside the hyperbolic limit cycle via asymptotic phase. Thus, the stable manifold of a point x_0 on a limit cycle is the union of points having equal phases, and it is often referred to as the isochron of x_0.

To study synchronization, a useful measurable property of a neural oscillator is its phase resetting curve (PRC). The PRC is found by perturbing the oscillation with a brief stimulus at different times on its cycle and measuring the resulting phase-shift from the unperturbed system. It is a very useful tool to explain how the coupling between neurons can affect the phase and lead them to a synchronized or non-synchronized activity.

Phase resetting curves (PRCs) constitute a powerful resource in time-control problems in biological processes. For instance, in the study of circadian rhythms, phase resetting curves are indicators for the experimentalists to know the peaks of the phase advancement and for the practitioners to administrate drugs (see for instance [CKWJ03], [DN94] or [AP07] for different contexts); that is, to know the optimal phase advancement.

Different methods are known to compute the PRCs, see [Izh07, Ch. 10] for a survey. One of the most effective is the so-called *adjoint method*, see [EK91] (also [BHM04] for a review). Recently, Govaerts and Sautois (see [GS06]) have developed a new algorithm to solve the adjoint method problem accompanied with the implementation of continuation methods to study PRCs along families of vector fields with a persistent limit cycle.

Typically, solutions to the models of interest tend asymptotically to a limit cycle. However, one may be interested in computing the phase advancement in the transient state, when the dynamics has not relaxed back to the limit cycle. This occurs when the period of stimulation is too short and is favored by factors like a slow attraction to the limit

cycle, a large stimulus amplitude, other external stimuli, random fluctuations, bursting-like stimuli, ... Thus, the study of the phase advancement under a certain stimulus in a *neighborhood of the limit cycle*, not only *on the limit cycle*, is also interesting.

Since the method that we develop in this memoir gives a "natural" parameterization of an entire neighborhood of a limit cycle, and the way we obtain the phase resetting curves is independent on whether a point is on the limit cycle or not, we can extend the computation of the phase resetting curves to a neighborhood and obtain what we call *phase resetting surfaces* (PRS from now on); that is, we can evaluate the phase advancement even when the stimulus is performed out of the limit cycle. The restriction of our method to the limit cycle gives the numerical scheme also used in [GS06].

In the examples, we also include a discussion on the relationship between the excitability types and the types of the corresponding PRCs. This issue was introduced by Ermentrout in [Erm96]; models with strictly positive or mainly positive PRC are called usually "Type 1 PRC" or "Class 1", whereas models whose PRC changes sign and present a negative regime (delay in the phase) are known as "Type 2 PRC" or "Class 2". The PRC type have effects on the synchronization of an oscillator with a periodic pulse train. For instance, for Type 1 models, that is with a PRC mostly positive, they easily synchronize with fast inputs but they cannot synchronize with slower inputs. This is because they can advance the phase to catch up with faster inputs but they cannot delay the phase. This is not the case for Type 2 models because they can advance or delay the phase. We study, in parametric families, the evolution from one type to another. It can be observed how the negative parts of the PRC for a "Type 2 PRC" oscillator shrink as some bifurcation parameter evolves until it almost vanishes near to a "Type 1 excitability" value (that is, close to a frequency zero limit cycle bifurcation), corresponding also to a "Type I PRC" value, see Example 3.9.2 and (3.40) ($I_{app} = 10$). We are more concerned with the observation and biological consequences of these changes between PRCs inside a family of vector fields, rather than carrying out a systematic computation of PRCs with respect to some parameter. As mentioned above, this job has been done in [GS06].

In the examples we also show (see Figs. 3.2 to 3.6) that, depending on the geometry of the isochrons, the shape of the phase resetting surface may be different from the shape of the phase resetting curve in such a way that systems with "Type I PRC" can present regions with negative PRS, thus allowing, away from the limit cycle, a delay of the phase which is not supposed for "Type I PRC" oscillators.

The mathematical formalism that we use to obtain information about isochrons, PRCs and PRSs is based on Lie symmetries, see [Sab05] or [FGG07] for new applications to limit cycles. Indeed, given a vector field X with a limit cycle γ, it is proved the equivalence between the existence of a Lie symmetry $[Y, X] = \mu Y$ and the foliation of a neighborhood of γ by isochrons (that is, with asymptotic phase well defined). Parallel, to this theory, Cabré-Fontich-de la Llave have developed a method to parameterize invariant manifolds around an invariant object, see for instance [CFL05] for the part related to limit cycles. This so-called *parameterization method* is much easier to implement than the computation of Lie symmetries. In this memoir we relate the two approaches; in fact, we prove that the coordinate curves of the parameterization are exactly the orbits of the unknown vector field Y.

The chapter is organized as follows: in Section 3.2 we give the necessary background

(on isochronous sections, Lie symmetries, parameterization method and phase resetting curves) to tackle the rest of the chapter. In Section 3.3, we relate the Lie symmetries with the parameterization method. Sections 3.4 and 3.6 are devoted to develop the effective method to compute the isochrons and the phase resetting curves (and surfaces), using the parameterization method. For the purpose of comparison, in Section 3.5, we explain the adjoint method and prove that our result is also a solution of the adjoint equation. In Section 3.7, we take care of the details of the numerical implementation of the method, an aspect which is often overlooked although it is not trivial for slow-fast systems. We devote Section 3.8 to envision the application of our method to higher dimensions. We end the chapter with some examples, in Section 3.9, and a final discussion (Section 3.10).

3.2 Background and statement of the problem

In this section we go through the background about the main tools that will be related later on. In general, these tools are defined for vector fields in $\mathbb{R}^d$, although for the purposes of this chapter we will restrict to $d = 2$ from Section 3.3 on.

3.2.1 Isochronous sections of a limit cycle

Let us consider an autonomous system of ODEs

$$\dot{x} = X(x),\ x \in U \subseteq \mathbb{R}^d, d \geq 2, \tag{3.1}$$

having a periodic orbit γ of period T, parameterized by $\theta = t/T$ as

$$\begin{array}{rccc} \gamma: & \mathbb{T} = \mathbb{R}/\mathbb{Z} & \to \mathbb{R}^d \\ & \theta & \to \gamma(\theta) \end{array} \tag{3.2}$$

in order to have period 1, that is $\gamma(\theta) = \gamma(\theta + 1)$.

For the numerical purposes of this memoir we will assume that X is an analytic vector field and so, all the functions and manifolds that we will associate to it. Nevertheless, the theoretical background that we are quoting in this section is still valid for lower regularity vector fields.

Definition 3.2.1. We say that a point $q \in \Omega \subset \mathbb{R}^d$, where Ω is an open domain containing the limit cycle γ, is in *asymptotic phase* with a point $p \in \gamma$ if

$$\lim_{t \to +\infty} |\Phi_t(q) - \Phi_t(p)| = 0, \text{ or}$$

$$\lim_{t \to -\infty} |\Phi_t(q) - \Phi_t(p)| = 0,$$

where Φ_t is the flow associated to the vector field X.

The set of points having the same asymptotic phase is called *isochron*.

Definition 3.2.2. We will say that a limit cycle γ is isochronous if there exists an open neighborhood Ω containing γ such that every point in Ω is in phase with a point on γ.

Remark 3.2.3. Notice that the isochrons are mapped to isochrons by the flow Φ_t of the vector field X. Hence, they are Φ_T-invariant, that is $\Phi_T(q)$ belongs to the isochron of q.

This extends the notion of phase of oscillation to a neighborhood in the basin of attraction of the limit cycle. Hence, in a neighborhood Ω of the limit cycle γ there exists a unique scalar function

$$\begin{array}{rccc} \vartheta: & \Omega \subset \mathbb{R}^d & \to \mathbb{T} = [0,1) \\ & x & \mapsto \vartheta(x) \end{array} \tag{3.3}$$

such that

$$\lim_{t \to +\infty(-\infty)} |\Phi_t(x) - \gamma(t + T\vartheta(x))| = 0.$$

The value $\vartheta(x)$ is the asymptotic phase of x and the isochrons are the level sets of $\vartheta(x)$, since it is constant on each isochron.

3.2.2 Isochrons, stable manifolds and Lie symmetries

From a seminal paper by Winfree ([Win75]) and the theoretical answers given by Guckenheimer in a subsequent paper ([Guc75]), it is known that phase sets or isochrons and stable manifolds of hyperbolic limit cycles have a common link: if the limit cycle is stable, then the isochrons are the leaves of the stable manifold, that is $W^s(\gamma(\theta))$, for $\theta \in \mathbb{T}$.

Remark 3.2.4. Notice that the case of a hyperbolic unstable limit cycle is equivalent to the stable case just reversing the time. However, when one works in dissipative systems it only makes sense to talk about attractors because the other invariant objects cannot be seen when one integrates forward the system. Hence, from now on, we will only mention the stable case.

Not much is said about the computation of these isochrons probably perhaps they can often be reduced to the computation of stable manifolds, which have been thoroughly studied.

New papers recovering the problem of the existence of isochrons for a generic non-hyperbolic limit cycle in the plane have appeared recently ([CL04, Sab05]). From Chicone and Liu's work, [CL04], we know that a limit cycle γ of a $\mathcal{C}^2$ planar vector field is isochronous if and only if it is hyperbolic or it is a non-hyperbolic limit cycle satisfying $\pi''(p) \neq 0$ and $\tau'(p) = 0$, where τ is the time of the first return to a Poincaré section Σ at $p \in \gamma$ and π is the corresponding Poincaré map. Moreover, in [Sab05], Sabatini proves that a limit cycle γ of a $\mathcal{C}^2$ planar vector field X is isochronous if and only if the vector field X is an infinitesimal generator of another C^2 planar vector field Y transversal to X; that is, if

$$[Y, X] = \mu\, Y, \tag{3.4}$$

for some C^2 function $\mu : \mathbb{R}^2 \longrightarrow \mathbb{R}$, where $[,]$ stands for the Lie bracket of the two vector fields. Moreover, it is stated that, in this case, the orbits of Y crossing the limit cycle γ are its isochrons because, by the Lie symmetry, the flow of X sends orbits of Y to orbits of Y.

Finally, Freire et al., see [FGG07], give a closed formula for the characteristic exponent of a limit cycle γ of a planar $\mathcal{C}^1$ vector field X in terms of μ in (3.4). More precisely, they prove that the characteristic exponent of γ is given by

$$\lambda = \int_0^T \mu(\gamma(t/T))dt. \tag{3.5}$$

Although the result of Sabatini is a nice geometrical characterization of isochronous limit cycles, the difficulty arises when trying to find μ and Y. In this memoir, we link this result with the parameterization method developed in [CFL05] to compute the two-dimensional stable manifold containing γ, and we implement it numerically to obtain local expansions of μ and Y.

3.2.3 The parameterization method

In this subsection we introduce the parameterization method described in [CFL05]. For limit cycles in planar systems, it consists of looking for a parameterization of the two-dimensional stable manifold in terms of the phase variable θ on the limit cycle and another variable σ which moves along the isochron/leave of the stable manifold and corresponds to the time of the orbits of the vector field Y in (3.4). Hence, we will be looking for a map

$$\begin{aligned} K : \mathbb{T} \times U \subset \ \mathbb{T} \times \mathbb{R} \ &\to \mathbb{R}^2 \\ (\theta, \sigma) \ &\mapsto K(\theta, \sigma), \end{aligned} \tag{3.6}$$

where U is an open interval containing 0, and a scalar λ such that they satisfy the equation

$$\left(\frac{1}{T} \frac{\partial}{\partial \theta} + \frac{\lambda \sigma}{T} \frac{\partial}{\partial \sigma} \right) K(\theta, \sigma) = X(K(\theta, \sigma)), \tag{3.7}$$

where T is the period of the limit cycle.

On the computational side, in [CFL05], the authors provide a method to solve the invariance equation (3.7) that leads immediately to practical numerical algorithms. We have implemented it in several biological models (specially from neuroscience) in which the control of the phase advancements becomes crucial. In Section 3.6 we review the method and the algorithms to solve equation (3.7) and in Section 3.7 we discuss the numerical implementation. Similar implementations of this method have been performed in [HL06a] for the computation of stable and unstable manifolds of invariant tori in quasi-periodic maps.

3.2.4 Phase response curves and surfaces

As mentioned in the Introduction, phase resetting curves are a key tool to study phase advancement in oscillators. Here we introduce the basic background.

Let us consider an oscillator of the form (3.1) with a stable limit cycle γ of period T (let us say, for instance, a periodically spiking neuron) which is stimulated at a phase $\theta = t_s/T$ with an arbitrary perturbation.

The effect of the perturbation is to produce a phase shift that can be an advance or a delay depending on the time of the stimulus t_s relative to the phase of the oscillation θ, leading to a change of the period. The representation of this phase shift is usually called *Phase Response Curve* or *Phase Resetting Curve* (PRC). They are typically defined as

$$\Delta\vartheta = (T - T_{new})/T \tag{3.8}$$

where T_{new} is the period for the perturbed limit cycle.

In this memoir we will focus on the particular case of infinitesimally small perturbations in duration and amplitude. In this case, the perturbation consists of a pulse that instantaneously displaces the trajectory away from the limit cycle in a certain direction by a certain amplitude. Mathematically, we consider

$$\dot{x} = X(x) + \varepsilon\delta(t - t_s) \tag{3.9}$$

where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_d) \in \mathbb{R}^d$ and $\delta(t)$ is the Dirac delta function.

When $|\varepsilon| \ll 1$, it is common in the theory of weakly coupled neural oscillators (Ermentrout and Kopell, 1990) to construct the so called *infinitesimal Phase Resetting Curve* (iPRC). Using the scalar function ϑ given in (3.3) that associates to every point in a neighborhood of the limit cycle a phase in $[0, 1)$, it is easy to see that the iPRC for an instantaneous perturbation as in (3.9) is mathematically equivalent to

$$\Delta\vartheta(x) = \varepsilon \cdot \nabla\vartheta(x) = \left(\frac{\partial\vartheta}{\partial x_1}(x), \ldots, \frac{\partial\vartheta}{\partial x_d}(x)\right).$$

for $x \in \gamma$, see [Izh07, Ch. 10] for the details.

Note that the pulse in (3.9) can be in any direction in $\mathbb{R}^d$. Usually, one studies the PRCs for the directions given by a vector basis of $\mathbb{R}^d$. For instance, for the planar case ($d = 2$), we will consider the PRCs corresponding to $\varepsilon = (1, 0)$ and $\varepsilon = (0, 1)$ and we will refer to them as PRC_1 and PRC_2, respectively. For models in neuroscience, one is usually interested only on the PRC for perturbations in the direction of the voltage, that is, $\partial\vartheta(x)/\partial V$, for $x \in \gamma$.

Although in the literature the phase shift is only computed on the limit cycle, that is $x \in \gamma$, the isochrons allow to naturally extend it in a neighborhood of the limit cycle and introduce a new concept that we call *Phase Resetting Surface* (also PRS from now on). In general, PRS are not considered in the literature because the methods to obtain the PRCs are not easily extendable.

The phase resetting surface tabulates the change in the phase produced by a perturbation as a function of the phase θ and the distance σ to the limit cycle computed on the isochron at which it is received. Notice that the PRC is just the section $\sigma = 0$ of the PRS. Hence, PRSs are a generalization of the PRCs for $\sigma \neq 0$. This tool can be very useful if we want to stimulate the oscillator repeatedly, without needing to wait for the oscillator to relax back to the limit cycle attractor. This required time to relax back is specially inconvenient when the attraction to the limit cycle is too slow or the amplitude of the stimulus is too large.

The classical method for computing PRCs was given by Ermentrout and Kopell in [EK91] and is commonly known as the *Adjoint method*. In Section 3.4 we describe a new

alternative method that allows to compute not only the PRC but also the PRS, using the Lie symmetries formalism and the numerical scheme provided by the parameterization method.

3.3 Lie symmetries and normal forms around limit cycles

In this section we establish a relation between the existence of a Lie symmetry and a 2-dimensional invariant manifold parameterized by the phase θ and the variable σ for the limit cycle of a planar vector field. The main result is given by Theorem 3.3.1.

Theorem 3.3.1. *Let γ be a hyperbolic T-periodic orbit of a planar analytic vector field X parameterized by θ according to* (3.2)*. Then, there exists a transversal vector field Y and a scalar function μ, both analytic, such that in a neighborhood Ω of the periodic orbit γ*

$$[Y, X] = \mu Y,$$

if and only if there exists a manifold $\mathcal{M}$ which is invariant under the flow of X and can be parameterized by an analytic map $K : \mathbb{T} \times U \subset \mathbb{R} \to \mathbb{R}^2$, satisfying

$$\left(\frac{1}{T}\partial_\theta + \left(\int_0^\sigma \mu(K(\theta,\tau))d\tau\right)\partial_\sigma\right) K(\theta,\sigma) = X \circ K(\theta,\sigma). \tag{3.10}$$

Moreover, $Y \circ K = \partial_\sigma K$, or equivalently

$$K(\theta,\sigma) = \psi_\sigma(\gamma(\theta)), \tag{3.11}$$

where ψ_σ is the flow of the vector field Y.

Proof: Let us look at the first implication. Let us consider $K(\theta,\sigma) = \psi_\sigma(\gamma(\theta))$, where ψ_σ is the flow associated to the vector field Y and $\gamma(\theta)$ is the parameterization of the periodic orbit of the vector field X.

Then, notice that

$$\partial_\sigma(X \circ K(\theta,\sigma)) = DX \circ K(\theta,\sigma)\partial_\sigma K(\theta,\sigma) = (DX \circ K(\theta,\sigma))(Y \circ K(\theta,\sigma)).$$

Using the Lie symmetry $DXY - DYX = \mu Y$, we have

$$\frac{\partial}{\partial\sigma} X \circ K(\theta,\sigma) = (DY \circ K(\theta,\sigma))(X \circ K(\theta,\sigma)) + \mu(K(\theta,\sigma))(Y \circ K(\theta,\sigma)) \tag{3.12}$$

Hence, $X \circ K(\theta,\sigma)$ is a solution of the linear equation (3.12) with initial condition

$$X \circ K(\theta,0) = \gamma(\theta). \tag{3.13}$$

Let Ψ_σ be the fundamental solution of the homogeneous equation

$$\frac{\partial}{\partial\sigma} X \circ K(\theta,\sigma) = (DY \circ K(\theta,\sigma))(X \circ K(\theta,\sigma)), \tag{3.14}$$

then, the variation of parameters formula tells us that the solution of (3.12) with initial condition (3.13) is given by

$$X \circ K(\theta, s) = \Psi_\sigma X(\gamma(\theta)) + \Psi_\sigma \int_0^\sigma \Psi_s^{-1} \mu(K(\theta, s))(Y \circ K(\theta, s)) ds$$

Notice that $\Psi_s^{-1} Y(K(\theta, s))$ is independent of s, that is

$$\begin{aligned} \partial_s(\Psi_s^{-1} Y(K(\theta, s))) &= -\Psi_s^{-1}(DY \circ K(\theta, s))\Psi_s \Psi_s^{-1}(Y \circ K(\theta, s)) \\ &\quad +\Psi_s^{-1}(DY \circ K(\theta, s))(Y \circ K(\theta, s)) \\ &= -\Psi_s^{-1}(DY \circ K(\theta, s))(Y \circ K(\theta, s)) \\ &\quad +\Psi_s^{-1}(DY \circ K(\theta, s))(Y \circ K(\theta, s)) \\ &= 0 \end{aligned}$$

then we can take $\Psi_s^{-1} Y(K(\theta, s)) = \Psi_\sigma^{-1} Y(K(\theta, \sigma))$ and we are led with the following expression for $X \circ K(\theta, \sigma)$,

$$X \circ K(\theta, \sigma) = \Psi_\sigma X(\gamma(\theta)) + \Psi_\sigma \Psi_\sigma^{-1} Y(K(\theta, \sigma)) \int_0^\sigma \mu(K(\theta, s)) ds.$$

Finally, using that the parameterization K is given by the orbits of the vector field Y on the limit cycle γ, see equation (3.11), we have

$$\partial_\theta K(\theta, \sigma) = \partial_\theta \psi_\sigma(\gamma(\theta)) = D\psi_\sigma(\gamma(\theta)) T X(\gamma(\theta)) = T\Psi_\sigma X(\gamma(\theta)),$$

and the expression for $X \circ K(\theta, \sigma)$ reads out

$$X \circ K(\theta, \sigma) = \left(\frac{1}{T} \partial_\theta + \left(\int_0^\sigma \mu(K(\theta, \tau)) d\tau \right) \partial_\sigma \right) K(\theta, \sigma),$$

as we wanted to see.

The implication the other way follows in the following way. Let K being a parameterization of the stable manifold $\mathcal{M}$ satisfying equation (3.10). Consider Y the vector field whose orbits for the points on the limit cycle $\gamma(\theta)$ are given by $\{K(\theta, \sigma) | \sigma \in \mathbb{R}\}$. Let σ be the integration time along the orbits of the vector field Y. Therefore,

$$Y \circ K(\theta, \sigma) = \partial_\sigma K(\theta, \sigma). \tag{3.15}$$

The fact that the curves $\{K(\theta_0, \sigma) | \sigma \in \mathbb{R}\}$ are transversal to the orbits of X implies also that Y is transversal to X.

We next prove that X is a normalizer of the vector field Y. From equation (3.10), taking derivatives with respect to σ, we get

$$\left(\frac{1}{T} \partial_\theta + \left(\int_0^\sigma \mu(K(\theta, \tau)) d\tau \right) \partial_\sigma \right) \partial_\sigma K + (\mu \circ K) \partial_\sigma K = (DX \circ K) \partial_\sigma K,$$

and using (3.15), we get

$$\left(\frac{1}{T} \partial_\theta + \left(\int_0^\sigma \mu(K(\theta, \tau)) d\tau \right) \partial_\sigma \right) (Y \circ K) + (\mu \circ K)(Y \circ K) = (DX \circ K)(Y \circ K).$$

By the chain rule,

$$(DY \circ K)\left(\frac{1}{T}\partial_\theta + \left(\int_0^\sigma \mu(K(\theta,\tau))d\tau\right)\partial_\sigma\right)K + (\mu \circ K)(Y \circ K) \;=\; (DX \circ K)(Y \circ K),$$

and again, by the invariance equation (3.10), we obtain

$$\begin{aligned}(DX \circ K)(Y \circ K) - (DY \circ K)(X \circ K) &= (\mu \circ K)(Y \circ K)\\ [Y, X] &= \mu Y,\end{aligned}$$

as we wanted to prove. □

3.3.1 Simplifying the invariance equation (3.10)

We would like to remark that there is a certain freedom for the choice of Y and μ. Thus, given a vector field Y and a scalar function μ such that $[Y, X] = \mu Y$, then for a non-vanishing smooth scalar function f it turns out that

$$[f\,Y, X] = \left(\mu - \frac{\nabla f^T X}{f}\right)(f\,Y)\,.$$

Using this freedom, it will be convenient to choose μ to be constant. From (3.5) we know that the characteristic exponent for the periodic orbit is given by

$$\lambda = \int_0^T \mu dt = \mu T.$$

Then, it is natural to choose $\mu = \lambda/T$, where λ is the characteristic exponent of the periodic orbit γ.

Hence, the invariance equation for the parameterization of the invariant manifold $\mathcal{M}$ is given by (3.7), that we recall here

$$\left(\frac{1}{T}\partial_\theta + \frac{\lambda\sigma}{T}\partial_\sigma\right)K(\theta,\sigma) = X \circ K(\theta,\sigma).$$

From equation (3.7) it is clear that $\mathcal{M} := \mathit{Range}(K)$ is an invariant manifold for the flow of X. Moreover, the motion generated by the vector field X on $\mathcal{M}$ expressed in the variables (θ, σ) parameterizing $\mathcal{M}$, is given by

$$\begin{aligned}\dot\theta &= 1/T,\\ \dot\sigma &= \lambda\sigma/T.\end{aligned} \tag{3.16}$$

That is, the variable θ rotates at a constant speed $1/T$ and the variable σ moves exponentially. Hence,

$$\Phi_t(K(\theta_0,\sigma_0)) = K(\theta_0 + t/T, \sigma_0 e^{\lambda t/T}),$$

where Φ_t is the flow of the vector field X.

Therefore the orbit of a point $K(\theta_0, \sigma)$, for any $\sigma \in \Omega$, approaches exponentially fast to the orbit of the point $K(\theta_0, 0)$, which corresponds to the point $\gamma(\theta_0)$ on the limit cycle. Hence,

$$\{K(\theta_0, \sigma) | \sigma \in \Omega\} \subset W^s_{\gamma(\theta_0)},$$

that is, the point $K(\theta_0, \sigma)$ is contained in the isochron of $\gamma(\theta_0)$. Moreover $W^s_\gamma = \cup_{\theta \in [0,1)} W^s_{\gamma(\theta)}$.

Since, in this particular case the invariant stable manifold for the periodic orbit γ is 2-dimensional in $\mathbb{R}^2$, the parameterization in terms of the phase variable θ that gives the position on the limit cycle and σ, which is a variable that moves along a transversal direction and corresponds to the integration time along the orbits of the vector field Y, is also a parameterization of the phase space $\mathbb{R}^2$ in a neighborhood of the limit cycle. The expression of the vector field X in the variables (θ, σ) can be considered as the normal form for a planar vector field around a limit cycle, reminiscent of the action-angle variables for conservative systems.

3.4 Computation of Phase Resetting Curves and Surfaces

The parameterization K and the vector field Y jointly with the characteristic exponent λ allow us to compute the *isochrons* and the *Phase Resetting Curves and Surfaces* (PRS).

3.4.1 Computing the Isochrons

We already mentioned that the orbit of the points given by $K(\theta_0, \sigma)$, for any $\sigma \in U$ approach exponentially fast the orbit of the point $K(\theta_0, 0) = \gamma(\theta_0)$.

Therefore a parameterization of the isochron of the point $\gamma(\theta_0)$ is given by the analytic map

$$\begin{aligned} K(\theta_0, \cdot) : U \subset \mathbb{R} &\longrightarrow \mathbb{R}^2 \\ \sigma &\longmapsto K(\theta_0, \sigma). \end{aligned}$$

3.4.2 Computing the PRS

We already mentioned in section 3.2.4 that from the mathematical point of view, the change of phase due to a pulse stimulation at a point $p = K(\theta, \sigma)$ in a neighborhood Ω of the limit cycle γ is given by

$$\nabla \vartheta(p) = \left(\frac{\partial \vartheta}{\partial x}(p), \frac{\partial \vartheta}{\partial y}(p) \right).$$

In order to compute $\nabla \vartheta(p)$ we consider the following argument: on the one hand, the isochrons are given by the level sets of the function $\vartheta : \mathbb{R}^2 \to \mathbb{R}$, introduced in (3.3), which associates a phase to each point in a neighborhood of the limit cycle. On the other hand,

they are the orbits of a vector field Y satisfying (3.4). Hence, it is clear that $\nabla\vartheta(p)$ has the same direction as $Y^{\perp}(p)$, which corresponds to the vector orthogonal to Y on p given by

$$Y(p) = Y(K(\theta,\sigma)) = \partial_\sigma K(\theta,\sigma).$$

We only need to add some normalization. Notice that for a trajectory $\phi_t(p)$, $p \in \Omega$ where ϕ_t is the flow of the vector field X, we have

$$\frac{d\vartheta}{dt}(\phi_t(p)) = 1/T,$$

therefore

$$\frac{d\vartheta}{dt}(\phi_t(p)) = \nabla\vartheta(\phi_t(p)) \cdot \frac{d}{dt}\phi_t(p) = \nabla\vartheta(\phi_t(p)) \cdot X(\phi_t(p)) = 1/T.$$

Using this normalization we have that for any $p \in \Omega$, the PRC is given by

$$\nabla\vartheta(p) = \frac{Y^{\perp}(p)}{T < Y^{\perp}(p), X(p) >}, \tag{3.17}$$

where $<,>$ denotes the dot product.

The PRC is just the PRS restricted to the points on the limit cycle, that is $\sigma = 0$, then for $p = K(\theta,0) \in \gamma$

$$\nabla\vartheta(K(\theta,0)) = \frac{Y^{\perp}(K(\theta,0))}{T < Y^{\perp}(K(\theta,0)), X(K(\theta,0)) >},$$

where $K(\theta,0) = K_0(\theta) = \gamma(\theta)$ and $Y(K(\theta,0))$ is given by

$$Y(K(\theta,0)) = \partial_\sigma K(\theta,0) = K_1(\theta).$$

Therefore,

$$\nabla\vartheta(\gamma(\theta)) = \frac{K_1^{\perp}(\theta)}{< K_1^{\perp}(\theta), X(\gamma(\theta)) >}. \tag{3.18}$$

3.5 The relation with the Adjoint method

As we already mentioned in the introduction, the reference method in neuroscience which is commonly used to compute Phase Resetting Curves, is the *Adjoint Method* (see [EK91], Hoppenstead). It essentially computes the gradient of the asymptotic phase at the points $p \in \gamma$, that is $\nabla\vartheta(p)$, by looking for a T-periodic solution of the equation

$$\frac{d\nabla\vartheta(\gamma(t/T))}{dt} = -DX^T(\gamma(t/T))\nabla\vartheta(\gamma(t/T)), \tag{3.19}$$

where $DX^T(\gamma(t/T))$ is the transpose of the real matrix $DX(\gamma(t/T))$, with the condition

$$\nabla\vartheta(\gamma(t/T)) \cdot X(\gamma(t/T)) = \frac{1}{T},$$

which in particular must hold for $t = 0$.

This procedure has been automated in the program *XPPAUT*, see [Erm02].

However, the Adjoint problem can be extended to a neighborhood of the limit cycle. The idea of this generalization is summarized in the following Proposition:

Proposition 3.5.1. *Let γ be a hyperbolic T-periodic orbit of a planar analytic vector field X parameterized by θ according to* (3.2)*. Assume that there exists a transversal vector field Y satisfying* (3.4) *in a neighborhood Ω. Then, given a trajectory $\phi_t(p)$, $p \in \Omega$ we have that*

$$\nabla\vartheta(\phi_t(p)) = \frac{Y^{\perp}(\phi_t(p))}{T < Y^{\perp}(\phi_t(p)), X(\phi_t(p)) >} \tag{3.20}$$

solves the Adjoint Problem

$$\frac{d\nabla\vartheta(\phi_t(p))}{dt} = -DX^T(\phi_t(p))\nabla\vartheta(\phi_t(p)), \tag{3.21}$$

with the condition

$$\nabla\vartheta(\phi_t(p)) \cdot X(\phi_t(p)) = \frac{1}{T} \tag{3.22}$$

Proof: Notice first that, by construction, condition (3.22) is clearly satisfied.

Let us prove then that (3.20) is a solution of (3.21). In order to check this statement, we first introduce the matrix J given by

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{3.23}$$

such that $Y^{\perp} = JY$. Notice that for a 2×2 real matrix A we have

$$(JA) - (JA)^T = \mathrm{tr}(A)J. \tag{3.24}$$

Now, we consider the derivative of $\vartheta(\phi_t(p))$ with respect to the time. In order to simplify notation we set $x = \phi_t(p)$, $g(x) :=< Y^{\perp}(x), X(x) >$ and $\tau(x) = \mathrm{tr}(DX)(x)$. Using that $\dfrac{d}{dt}Y(\phi_t(p)) = DY(\phi_t(p))X(\phi_t(p))$, we have

$$\frac{d}{dt}\nabla\vartheta(x) = \frac{J\,DY(x)\,X(x)}{T\,g(x)} - \frac{Y^{\perp}(x)\,(< JDY(x)X(x), X(x) > + < JY, DX(x)\,X(x) >)}{T\,g(x)^2}$$

Using now that the Lie symmetry gives $DX\,Y - DY\,X = \mu Y$, that expression (3.20) reads out as $\nabla\vartheta(x) = (JY(x))/(T\,g(x))$ and dot product properties (namely, $< JY(x), DX(x)\,X(x) > = < DX(x)^T\,JY(x), X(x) >$), we obtain

$$\begin{aligned} \frac{d}{dt}\nabla\vartheta(x) &= \frac{J\,DX(x)\,Y(x) - \mu(x)\,JY(x)}{T\,g(x)} \\ &- \frac{\nabla\vartheta(x)\left(< JDX(x)Y(x) - \mu(x)\,JY(x) + DX(x)^T\,JY(x), X(x) >\right)}{g(x)} \end{aligned}$$

Applying equation (3.24) and $(J\,DX(x))^T = -DX(x)^T\,J$, we are led to

$$\frac{d}{dt}\nabla\vartheta(x) = \frac{(-DX(x)^T + \tau(x) - \mu(x))JY(x)}{T\,g(x)} - \frac{\nabla\vartheta(x)\,(< (\tau(x) - \mu(x))JY(x), X(x) >)}{g(x)}$$

In fact, again since $\nabla\vartheta(x) = (JY(x))/(T\,g(x))$, it can be written as:

$$\frac{d}{dt}\nabla\vartheta(x) = (-DX(x)^T + \tau(x) - \mu(x))\nabla\vartheta(x) - T\nabla\vartheta(x)\,(<(\tau(x)-\mu(x))\nabla\vartheta(x), X(x)>)$$

Finally, using the already proved condition (3.22) we have

$$\frac{d}{dt}\nabla\vartheta(x) = (-DX(x)^T + \tau(x) - \mu(x))\nabla\vartheta(x) - \nabla\vartheta(x)\,(\tau(x)-\mu(x)) = -DX(x)^T\nabla\vartheta(x),$$

as we wanted to prove. □

Remark 3.5.2. It is clear that the classical Adjoint method considers $p \in \gamma$, then $\phi_t(p) = \gamma(t/T)$ with $\gamma(0) = p$.

Remark 3.5.3. We will see in the section devoted to the numerical implementation, that we will obtain a local approximation of the PRS semi-analytically by computing the parameterization K and using formula (3.20). In order to obtain a PRS in a bigger domain, we will globalize the local approximation just integrating the adjoint problem equation (3.21) backwards.

3.6 Solving the invariance equation

In [CFL05], the authors provide a method to solve the invariance equation (3.7) and they prove its convergence. In this section, we review the basic steps of the method and we refer the reader to [CFL05] for more details and the proof of the theorems.

In order to solve the invariance equation (3.7), we will discretize it in Fourier-Taylor series. Hence, we will first look for a K as a power series

$$K(\theta,\sigma) = \sum_{n=0}^{\infty} K_n(\theta)\sigma^n, \tag{3.25}$$

where the components of K_n are periodic functions of period 1, and then match the coefficients in σ^n on both sides of equation (3.7).

For $n = 0$, one obtains

$$\frac{1}{T}\frac{d}{d\theta}K_0(\theta) = X(K_0(\theta)). \tag{3.26}$$

which admits the solution $K_0(\theta) = \gamma(\theta)$, where γ is a parameterization of the limit cycle given in (3.2).

Remark 3.6.1. Notice that if $K_0(\theta)$ is a solution, then $K_0(\theta+\omega)$ is also a solution for any $\omega \in [0,1)$. Therefore, there is some ambiguity in parameterizing the phase of an oscillation, that can be avoided fixing the initial point corresponding to the zero phase. It can be fixed anywhere on the limit cycle. In the context of tonic spiking in neuroscience, for instance, it is common to fix $\theta = 0$ at the peak of the spike.

For $n = 1$, we obtain

$$\frac{1}{T}\frac{d}{d\theta}K_1(\theta) + \frac{\lambda}{T}K_1(\theta) = DX \circ K_0(\theta)K_1(\theta), \tag{3.27}$$

which tells us that $K_1(\theta)$ is an eigenfunction with eigenvalue $-\lambda$ of the operator $\mathcal{L}$ defined by

$$\mathcal{L} := \frac{d}{d\theta} - TDX \circ K_0(\theta).$$

Using Proposition 5.2 in [CFL05], we know that $K_1(\theta)$ is a solution of the equation (3.27) with eigenvalue $-\lambda$ if and only if $K_1(0)$ is an eigenvector of the monodromy matrix Φ_1 with eigenvalue e^{λ}. The monodromy matrix can be computed by solving the first variational equation

$$\frac{d}{d\theta}\Phi_\theta = DX \circ K_0 \Phi_\theta,$$

with $\Phi_0 = \mathrm{Id}$ and taking the value Φ_1.

Recall that for planar vector fields, the other eigenvector is given by the vector field $X(K_0(0)) = X(K_0(1))$ with associated eigenvalue 1.

Finally, it is easy to see that $K_1(\theta) = e^{\lambda\theta/T}\Phi_\theta K_1(0)$ is a solution of equation (3.27).

Remark 3.6.2. For the numerical computations when the eigenvalue e^{λ} is very small, we will use that $\lambda = \int_0^T \mathrm{div}(X(\gamma(t/T)))dt$.

Remark 3.6.3. Notice that if $K_1(\theta)$ is a solution of equation (3.27), then $bK_1(\theta)$, for any $b \in \mathbb{R}$, is also a solution. Even though all the choices of $K_1(\theta)$ are mathematically equivalent, the choice affects the numerical properties of the algorithm. See Remark 3.7.2 for a more detailed discussion.

For $n \geq 2$, we have

$$\frac{1}{T}\frac{d}{d\theta}K_n + \frac{n\lambda}{T}K_n = (DX \circ K_0)K_n + R_n \tag{3.28}$$

where R_n is an explicit polynomial in $K_0, \ldots, K_{n-1}$ whose coefficients are derivatives of X evaluated at K_0. These coefficients will be computed using the methods of automatic differentiation (see for instance [Gri00] and [JZ05]).

By Proposition 5.2 in [CFL05] the equation (3.28) for $n \geq 2$ can be solved provided that $e^{n\lambda}$ is not an eigenvalue of the monodromy matrix Φ_1 associated to γ. Notice that this assumption is satisfied for planar vector fields, provided that the limit cycle is hyperbolic, that is $\lambda \neq 0$.

Once $K_0(\theta)$ and $K_1(\theta)$ are fixed (see Remarks 3.6.1 and 3.6.3), the solution $K_n(\theta)$ for $n \geq 2$ of equation (3.28) is uniquely determined. Taking into account that K_n are periodic solutions in θ, we will discretize the equation (3.28) using Fourier series and reduce the problem to solve a linear system in the Fourier space, see Section 3.7 for more details.

Finally, by Theorem 5.4 in [CFL05] we know that, provided that λ satisfies the mentioned conditions, the series constructed here converges to a true analytic solution of the problem.

3.7 Numerical implementation of the method

In the previous sections we have described the method, but there are many numerical details which are important and nontrivial. They do not depend on the method but they are inherent to the problem. In this section we provide some details about the implementation we have carried out.

3.7.1 Fourier-Taylor discretization

In order to solve equation (3.7), we will discretize the invariance equation using **Fourier-Taylor series** and study numerical methods to solve the discretized equations.

As we already mentioned in Section 3.6, we first seek K as a power series

$$K(\theta, \sigma) = \sum_{n=0}^{\infty} K_n(\theta)\sigma^n,$$

where $K_n(\theta)$ are 1-periodic functions in θ. Thus, using Fourier formalism, the $K_n(\theta)$ can be written as

$$K_n(\theta) = \sum_{k\in\mathbb{Z}} c_k^n e^{2\pi i k\theta}.$$

Since we deal only with real functions, we only need to store half of the coefficients or, equivalently, store the cosine and sine Fourier series:

$$K_n(\theta) = a_0^n + \sum_{k>0} a_k^n \cos(2\pi k\theta) + b_k^n \sin(2\pi k\theta),$$

where $a_0 = c_0^{re}$, $a_k = 2c_k^{re}$ and $b_k = -2c_k^{im}$ for $k > 0$.

In the numerical implementation we need to truncate these expansions. In order to decide up to which order N we compute the Fourier series we require that the residuals are of size of order $10^{-15} - 10^{-20}$. That is, we truncate the Fourier series up to some order N in such a way that the norm of the last 10 per cent of Fourier coefficients is smaller that the considered order, in symbols

$$|K_n^{tail}| = \sum_{k=\lfloor 0.9N/2 \rfloor}^{N/2} |a_k^n| + |b_k^n| < 1.0e - 15. \tag{3.29}$$

Remark 3.7.1. One of our goals is to apply this method to classical systems in neuroscience. The main practical shortcoming in these cases is that the Fourier series are not adaptable to the usual presence of spikes (slow-fast systems), where the Fourier coefficients decrease very slowly and not uniformly. Although these systems can be analytic, from this numerical point of view they behave as if they were not. In these cases, other methods of discretization which are more adaptive like splines or wavelets could give some improvements.

Since the vector field is assumed to be analytic, so it can be written using algebraic operations and elementary transcendental functions, we can use automatic differentiation algorithms (see [Gri00], [JZ05]) to obtain Taylor expansions of the operators in (3.7).

In the whole process when we have a discretization of a periodic function in the real space, we use the Fast Fourier Transform (FFT) to compute the Fourier series. In this work we have used the `fftw3` library (see http://www.fftw.org/).

3.7.2 Discretization of the invariance equation and accuracy of the solution

We will solve the invariance equation (3.7) by solving equations (3.26), (3.27) and (3.28) for $n \geq 2$. Observe that equations (3.26) and (3.27) are special because they involve four unknowns $(K_0(\theta), T, K_1(\theta), \lambda)$. These two equations will be solved simultaneously using additional information. Instead, equation (3.28) can be treated the same way for any $n \geq 2$.

In order to solve equations (3.26) and (3.27) we will need to integrate the system of ODEs. The integration method used is a Taylor method (we have used the routines provided by Jorba and Zou, see [JZ05] and http://www.maia.ub.es/ angel/soft.html). We used adaptive step size and degree and a tolerance (absolute and relative) of 1.0e-16.

Recall that for $n = 0$ we need to look for a periodic solution. In order to compute it, we consider a Poincaré section and reduce the problem to find a zero of the Poincaré map that can be achieved using a Newton method. Note that for the Newton method we will need to integrate the variational equations together with the vector field. The solution for the variational equations will be used to solve equation (3.27) according to the method explained in Section 3.6.

Once we obtain the limit cycle $K_0(\theta)$ and $K_1(\theta)$ we store them for equidistant values of θ, that is $\theta_i = i/N$, for $i = 0, \ldots, N-1$.

For $n \geq 2$ the most straightforward method is to discretize (3.28) using a basis of N Fourier coefficients and, then, apply a linear solver. However, once we have obtained K_0 and K_1, we can perform a change of coordinates given by $(x, y) = g(\theta, \sigma) = K_0(\theta) + \sigma\, K_1(\theta)$. If we apply the method again to the system obtained after this change, then it turns out that the equation (3.28) becomes diagonal in Fourier series. Once we obtain the solution as a Fourier series we can go back to real space using the Fast Fourier Transform. Again, as in the previous cases we store K_n for equidistant values of θ.

An alternative method consists of applying a quasi-Newton method to the invariance equation.

By now, the results shown in Section 3.9 have been obtained using the straightforward method.

To check the accuracy of the solutions K_n obtained, we substitute them in the corresponding equation ((3.26) if $n = 0$, (3.27) if $n = 1$ and (3.28) if $n \geq 2$) for discrete values of $\theta_i = i/N$, for $i = 0, \ldots, N-1$. For each value of θ, this substitution provides an evaluation of the error $E_n(\theta)$. Finally, we evaluate the discrete ℓ_1 norm of $\{E_n(\theta_i)\}_{i=0}^{N-1}$ to get the accuracy; that is

$$\|E_n\| = \frac{1}{N} \sum_{i=0}^{N-1} |E_n(\theta_i)|. \tag{3.30}$$

Notice that the computation of $E_n(\theta_i)$ involves again a FFT.

3.7.3 Local approximation of the isochrons

Once we have solved the invariance equation up to order L we have obtained a local approximation of the stable invariant manifold. It remains to determine the domain of

convergence and the order of the error of the local approximation. Both concepts are very strongly related.

Given an approximate solution

$$K^{[\leq L]}(\theta,\sigma)=\sum_{n=0}^{L}K_n(\theta)\sigma^n, \tag{3.31}$$

where

$$K_n(\theta)=a_0^n+\sum_{k=1}^{N/2}a_k^n\cos(2\pi k\theta)+b_k^n\sin(2\pi k\theta),$$

The radius of convergence $r=1/l$ for the Taylor series is given by

$$l=\lim_{n\to\infty}\frac{\|K_{n+1}\|}{\|K_n\|},$$

where $\|\cdot\|$ denotes the ℓ_1 norm defined in (3.30). Thus, a direct strategy to compute l could be imposing that $|l_n-l_{n-1}|<\varepsilon$, for some prescribed $\varepsilon>0$, where $l_n=\|K_{n+1}\|/\|K_n\|$. Then, $r=1/l_n$ could be a numerical approximation of the theoretical radius of convergence. However, for numerical reasons, the radius can shrink in practical implementation. Consequently, we compute the convergence region in an alternative way.

For each θ we compute a value $\sigma_0(\theta)$ such that the approximate solution $K^{[\leq L]}(\theta,\sigma)$ given in (3.31) solves the invariance equation (3.7) up to a certain error E, that we established between $10^{-10}-10^{-12}$. That is, we fix θ and we compute the values of $\sigma\in\mathbb{R}$ such that

$$\left|\frac{1}{T}\sum_{n=0}^{L}K_n'(\theta)\sigma^n+\frac{\lambda}{T}\sum_{n=0}^{L}nK_n(\theta)\sigma^n-X\left(\sum_{n=0}^{L}K_n(\theta)\sigma^n\right)\right|<E, \tag{3.32}$$

where

$$K_n'(\theta)=2\pi\sum_{k=1}^{N/2}kb_k^n\cos(2\pi k\theta)-ka_k^n\sin(2\pi k\theta).$$

Remark 3.7.2. Recall that if $K(\theta,\sigma)$ is a solution of the invariance equation (3.7) so is $K(\theta+\omega,b\sigma)$, for any $\omega\in[0,1)$ and $b\in\mathbb{R}$. As we already mentioned in Remark 3.6.1 the choice of ω is related to the zero phase for the limit cycle. So, following the usual criterion in neuroscience, we will fix the zero phase for the oscillator at the spike. The choice of b is related to the domain of convergence. Hence, if we choose a large b the domain where we can evaluate the series will be small. Although mathematically we can choose any value of b, for the numerical stability it will be convenient to choose a value of b such that the coefficients K_n can be kept at order 1, so that one can avoid the round-off errors. Notice that if we consider bK_1 then new K_n is b^nK_n.

However, in some cases, the K_n do not converge uniformly and in these cases one can not find a global b. The immediate consequence of this fact is that for some values of θ, the K_n become smaller than the machine precision and one can not trust them. For these values, increasing the order L of the Taylor polynomial has no effect on increasing the domain where the local approximation is reliable.

3.7.4 Globalizing the manifold

In theory, the method presented here gives a parameterization of the whole manifold. However, we have seen that, numerically, given an error bound ($10^{-10} - 10^{-12}$), we can compute the isochron only up to a value $\sigma = \sigma_0(\theta)$ for each θ.

A standard way to extend local approximations obtained semi-analytically is to globalize them using the dynamics given by the vector field (see [Sim90]).

Typically, given a point $\gamma(\theta_0)$ on the limit cycle, one could take n points parameterized by (θ_0, σ) with $\sigma \in (\sigma_0(\theta_0)e^{\lambda}, \sigma_0(\theta_0))$, on the corresponding isochron and then perform iterates of the inverse time-T map Φ_{-T} for these points, where here Φ_t denotes the flow of X. However, in many cases (included models in neuroscience in which we are specially interested) this method has the disadvantage that we get too many points close to $\gamma(\theta_0)$ and just a few far from it. Moreover, some of them may escape very fast far from the limit cycle.

This last shortcoming can be avoided using that isochrons, even if they are not invariant, they are preserved by the flow, that is isochrons are carried into isochrons. Hence, we can consider inverse time-$T\Delta\theta$ maps $\Phi_{-T\Delta\theta}$ as well as, taking $\Delta\theta = 1/n$, $n \in \mathbb{N}$, n local invariant manifolds corresponding to $\gamma(\theta_0 + j\Delta\theta)$, for $j = 0, \ldots, n-1$. Then, to globalize the isochron corresponding to $\gamma(\theta_0)$ we obtain points $\{p_0, \ldots, p_m\}$ on it from points on the local approximation of other isochrons parameterized by $(\theta_0 + k_m\Delta\theta, \sigma_m)$, such that

$$p_m = \Phi_{-k_m T\Delta\theta}(K(\theta_0 + k_m\Delta\theta, \sigma_m)) \tag{3.33}$$

with $k_m \in \mathbb{N}$.

The method to decide which σ_m and k_m we choose to compute each p_m in order to get points on the globalized isochron less sparse, is based on a method given in [Sim90] (see also [KO98] for another alternative). For the sake of completeness, we explain the details of the method adapted to our purpose.

We want to extend the local isochron for a phase θ_0. We are going to approximate it by a sequence of points $\{p_0, \ldots, p_m\}$ on the isochron for which we will assume that they are at a distance smaller than some tolerance Δs, that is

$$\|p_m - p_{m-1}\| < \Delta s,$$

and the angle between three consecutive points is bigger than a certain tolerance $\Delta\alpha$,

$$(p_{m-1} - p_{m-2}) \cdot (p_m - p_{m-1}) \geq \cos(\Delta\alpha)\|p_{m-1} - p_{m-2}\|\|p_m - p_{m-1}\|.$$

Assume that we have computed up to p_m satisfying the previous conditions and we have a current value of σ_m and $\Delta\sigma_m$ such that $\sigma_m = \sigma_{m-1} + \Delta\sigma_m$ and a certain iterate k such that

$$\Phi_{-kT\Delta\theta}(K(\theta_0 + k\Delta\theta, \sigma_m)) = p_m.$$

We want to predict the new σ_{m+1} and therefore $\Delta\sigma_{m+1}$, such that

$$\Phi_{-kT\Delta\theta}(K(\theta_0 + k\Delta\theta, \sigma_m + \Delta\sigma_{m+1})) = p_{m+1},$$

satisfying that it is at a distance smaller than Δs from p_m.

Then, we consider

$$\Delta\sigma_{m+1} = \min\left(\frac{\Delta s}{\Delta s_m}, \frac{\Delta\alpha}{\Delta\alpha_m}, 1.5\right)\Delta\sigma_m 0.8$$

where Δs_m is the distance between p_{m-1} and p_m and $\Delta\alpha_m$ the angle between $\bar{v} = p_{m-1} - p_{m-2}$ and $\bar{w} = p_m - p_{m-1}$. The factor 0.8 can be seen as a security factor.

If $\sigma_{m+1} = \sigma_m + \Delta\sigma_{m+1}$ falls into the allowed range for σ, that is $\sigma_{m+1} < \sigma_0(\theta_k)$, where $\theta_k = \theta_0 + k\Delta\theta$, which means that we are in the range where the local approximation of the isochron for θ_k is good, we integrate for $\Phi_{-kT\Delta\theta}$ and we obtain p_{m+1}. In this case we define $k_{m+1} = k$ according to (3.33).

Otherwise, we keep dividing both σ_{m+1} and $\Delta\sigma_{m+1}$ by $e^{\lambda\Delta\theta}$ l times until $\sigma_{m+1} \leq \sigma_0(\theta_{k+l})$. We say then that $k_{m+1} = k + l$ and we compute p_{m+1} from (3.33). Typically, $l = 1$ but it can be greater. We replace k by $k + 1$.

If despite our choice of $\Delta\sigma_m$ the point p_{m+1} obtained fails to satisfy one of the conditions, we can either consider a smaller $\Delta\sigma_{m+1}$ (taking into account that the $\Delta\sigma$'s cannot be smaller than a certain value $\Delta_{\min}$) or keep the computed point and use an interpolation method for this part.

Remark 3.7.3. We can globalize the PRS in parallel with the isochrons: we approximate them locally according to (3.20) and we globalize them integrating the system (3.21) backwards together with the vector field.

3.7.5 Software

The algorithms have been implemented in C language and have been run under the Linux environment. They have been applied to compute isochronous sections and PRCs of limit cycles for planar vector fields which appear in models of neuroscience and neurobiology.

The program performs the following steps: (1) Computation of the limit cycle and its period, the monodromy matrix and the characteristic exponent. (2) Computation of the Fourier-Taylor expansions of the isochrons (3) Computation of the domain of convergence and the local approximation for the isochrons and the PRS (4) Globalization of the isochrons and the PRS. The figures are obtained using `gnuplot` and `Matlab`.

3.8 Isochronous sections, (un)stable manifolds and foliations in $\mathbb{R}^n$

Although computing isochronous sections of limit cycles in $\mathbb{R}^n$, for $n \geq 3$, is beyond the scope of this thesis, we would like to highlight the main differences with the planar case. The effective computation in higher dimensions is a goal for future work.

The theoretical extension to higher dimensions can be derived in a straightforward manner, though the practical implementation encompasses a plethora of new challenges. Here, we give the theoretical ideas and concepts, together with comments concerning practical issues.

Definition 3.8.1. Let us consider a smooth system of differential equations

$$\dot{x} = X(x), \qquad x \in \Omega, \quad \Omega \in \mathbb{R}^n, \quad n \geq 2,$$

with a hyperbolic limit cycle $\gamma : \mathbb{R} \mapsto \mathbb{R}^n$. An *isochronous section* of γ is a hypersurface Σ of $\mathbb{R}^n$ (dimension $n-1$) such that

$$x \in \Sigma \Leftrightarrow \varphi_T(x) \in \Sigma$$

where $\varphi_t(x)$ is a solution of X such that $\varphi_0(x) = x$, and T is the period of the limit cycle γ.

To extend the theoretical results, it is convenient to refer to *integrable systems* (see [Olv93]):

Definition 3.8.2. Let $Y_1, \ldots, Y_r$ be vector fields on a smooth manifold M. An *integral submanifold* of $\{Y_1, \ldots, Y_r\}$ is a submanifold $N \subset M$ whose tangent space $TN|_x$ is spanned by the vectors $\{Y_1|_x, \ldots, Y_r|_x\}$ for each $x \in N$. The system of vector fields $\{Y_1, \ldots, Y_r\}$ is *integrable* if through every point $x_0 \in M$ is contained in an integral submanifold.

Thus, having a neighborhood $\Omega \subset \mathbb{R}^n$ of γ filled by isochrons is equivalent to having an integrable system of vector fields $\{Y_1, \ldots, Y_r\}$, with $r = n-1$ defined in this neighborhood. One way to obtain these vector fields Y_j is to impose (see [FGG07]) that $[Y_j, X] = \mu_j Y_j$, for $j = 1, \ldots, n-1$. However, as Frobenius theorem shows, this requirement is not sufficient: the $n-1$ vector fields have to be in involution.

We recall both the definition of involution and Frobenius theorem (see also [Olv93]):

Definition 3.8.3. A system of vector fields $\{Y_1, \ldots, Y_r\}$ on M is *in involution* if there exist smooth real-valued functions $h_{ij}^k(x)$, $x \in M$, $i,j,k = 1, \cdots, r$, such that for each $i, j = 1, \cdots, r$,

$$[Y_i, Y_j] = \sum_{k=1}^{r} h_{ij}^k \cdot Y_k$$

Theorem 3.8.4. (Frobenius) *Let $Y_1, \ldots, Y_r$ be smooth vector fields on M. Then, the system $\{Y_1, \ldots, Y_r\}$ is integrable if and only if it is in involution.*

Summing up, adapting Frobenius theorem to our problem, we have:

Theorem 3.8.5. *Let us consider a smooth system of differential equations*

$$\dot{x} = X(x), \qquad x \in \Omega, \quad \Omega \in \mathbb{R}^n, \quad n \geq 2,$$

with a hyperbolic limit cycle $\gamma : \mathbb{R} \mapsto \mathbb{R}^n$. Suppose that there exist $n-1$ non-trivial vector fields $Y_1, \ldots, Y_{n-1}$ in involution such that

$$[Y_j, X] = \mu_j Y_j \quad j = 1, \ldots, n-1$$

for scalar functions $\mu_j : \mathbb{R} \mapsto \mathbb{R}^n$, $j = 1, \ldots, n-1$. Then, the isochronous sections are the maximal integral submanifolds of the integrable system of vector fields $\{Y_1, \ldots, Y_{n-1}\}$.

The integral submanifolds themselves are referred to as *leaves* of the *foliation* of Ω.

When $n > 2$, it may happen that a limit cycle has both stable and unstable manifolds. Let $n_s = \dim W_s(\gamma)$ and $n_u = \dim W_u(\gamma)$. Using the parameterization method or other computational techniques, we can obtain both $W_s(\gamma)$ and $W_u(\gamma)$ numerically. In other words, we would be computing the leaves of the partial foliations given by the systems $\{Y_1, \ldots, Y_{n_s}\}$ (resp., $\{Y_{n_s+1}, \ldots, Y_{n_s+n_u}\}$), where $\int_\gamma \mu_j < 0$ $(resp., > 0)$ if $j = 1, \ldots, n_s$ (resp., $j = n_s + 1, \ldots, n_s + n_u$). However, this does not give the whole foliation of the neighborhood.

A more suitable formulation to characterize the isochronous leaves can be obtained by using differential geometry notation. For the sake of simplicity, we only give a sketch of it and we restrict ourselves to $\mathbb{R}^2$ though it can be extended to $\mathbb{R}^n$.

Consider a vector field $X := P(x,y)\frac{\partial}{\partial x} + Q(x,y)\frac{\partial}{\partial y}$ and its associated 1-form $\omega_X := -Q(x,y)\,dx + P(x,y)\,dy$. We recall that $i_X\,\alpha$ denotes the contraction (interior product) of a k-form α with respect to X which gives a $(k-1)$-form. In particular, any vector field and its associated 1-form are related by $i_X\,dx \wedge dy = \omega_X$.

It is known from classical tensor calculus (see formula 1.62 in [Olv93]) that

$$i_{[X,Y]}\,\omega = X(i_Y\,\omega) - i_Y(X(\omega)), \tag{3.34}$$

for any k-form ω. If we use that $[Y, X] = \mu\,Y$ and choose $\omega = \omega_Y$, the left-hand side of (3.34) becomes

$$i_{[Y,X]}(\omega_Y) = i_{\mu\,Y}(\omega_Y) = \mu\,i_Y(\omega_Y).$$

Using now that $i_Y(\omega_Y) = 0$, we can write (3.34) as $i_Y(X(\omega_Y)) = 0$ or, equivalently,

$$X(\omega_Y) = \lambda(x,y)\,\omega_Y, \tag{3.35}$$

for some function λ. Thus, the problem of finding isochrons is equivalent to finding a 1-form ω_Y and a function λ satisfying (3.35). The leaves of ω_Y will then be the isochronous sections of the limit cycle.

3.9 Examples

In this section, we apply our method to representative examples, ranging from the most simple instances of Hopf and SNIC (saddle-node on an invariant curve) bifurcations and the classical van der Pol oscillator to more sophisticated neuronal models. Apart from obtaining isochrons, PRCs and PRSs, through these examples we want to illustrate different facts: (a) what are the clues to explain the transition from "Type 1" PRCs to "Type 2" PRCs; (b) the numerical problems that arise when dealing with slow-fast systems; and, (c) up to which degree PRSs show disagreement with PRCs in the same phase and how this can affect high frequency stimulation. We end the chapter with a discussion on these facts in Section 3.10.

We start with a direct application to the simplest vector fields that exhibit either a Hopf or a SNIC bifurcation, for which we can compute their limit cycle and the corresponding normalizing vector field analytically and we can also get an analytic expression for the PRC.

Example 3.9.1. We consider a simple example of a supercritical Hopf bifurcation

$$\begin{cases} \dot{x} &= \beta x - y - x(x^2+y^2), \\ \dot{y} &= x + \beta y - y(x^2+y^2), \end{cases} \tag{3.36}$$

which writes, in polar coordinates, as

$$\begin{cases} \dot{r} &= r(\beta - r^2), \\ \dot{\theta} &= 1. \end{cases}$$

For $\beta = 0$, there is a supercritical Hopf bifurcation giving rise, for $\beta > 0$, to a stable limit cycle γ of radius $\sqrt{\beta}$ and period 1. We parameterize γ by θ in the following way:

$$\gamma(\theta) = (\sqrt{\beta}\cos(\theta), \sqrt{\beta}\sin(\theta)).$$

It is not difficult to see that the vector field $Y(x,y) = (x,y)$ and the function $\mu(x,y) = -2(x^2+y^2)$ satisfy the condition (3.4).

Hence, taking into account that $Y^\perp = (-y,x)$ and $< Y^\perp, X >= x^2+y^2$, by equation (3.18) the phase shift for a point $p = (x,y) \in \Omega$ is given by

$$\nabla\vartheta(p) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\right).$$

Then, using the parameterization of the limit cycle, the PRC is just

$$\nabla\vartheta(\gamma(\theta)) = \frac{1}{\beta}(-\sqrt{\beta}\sin(\theta), \sqrt{\beta}\cos(\theta)).$$

That is, $PRC_1(\theta) = -\sin(\theta)/\sqrt{\beta}$, and $PRC_2(\theta) = \cos(\theta)/\sqrt{\beta}$.

Example 3.9.2. The easiest way to obtain a saddle node on an invariant cycle bifurcation bifurcation is through

$$\begin{cases} \dot{r} &= r(\beta - r^2), \\ \dot{\phi} &= m - \sin(\phi), \end{cases} \tag{3.37}$$

which, in cartesian coordinates, writes as

$$\begin{cases} \dot{x} &= \beta x - my - x(x^2+y^2) + \dfrac{y^2}{\sqrt{x^2+y^2}}, \\ \dot{y} &= m\,x + \beta y - y(x^2+y^2) - \dfrac{xy}{\sqrt{x^2+y^2}}. \end{cases}$$

We assume that $\beta > 0$ and $m > 1$. Therefore, there exists a unique and stable circular limit cycle γ of radius $\sqrt{\beta}$ that we parameterize by a phase θ satisfying $\dot{\theta} = 1/T$, $\theta \in [0,1)$ in the following way

$$\gamma(\theta) = (\sqrt{\beta}\cos(\Omega(\theta)), \sqrt{\beta}\sin(\Omega(\theta))),$$

where Ω is the phase transformation between θ and ϕ, given by the solution of the Cauchy problem

$$\frac{1}{T}\frac{d\Omega}{d\theta} = m - \sin(\Omega(\theta)); \quad \Omega(0) = 0.$$

The explicit solution can be obtained analytically,

$$\Omega(\theta) = 2\arctan\left(\frac{m\sin(\frac{T}{2}\sqrt{m^2-1}\,\theta)}{\sqrt{m^2-1}\cos(\frac{T}{2}\sqrt{m^2-1}\,\theta)+\sin(\frac{T}{2}\sqrt{m^2-1}\,\theta)}\right). \tag{3.38}$$

Again, as in Example 3.9.1, the vector field $Y(x,y)=(x,y)$ and the function $\mu(x,y)=-2(x^2+y^2)$ satisfy condition (3.4).

Hence, taking into account that $Y^\perp=(-y,x)$ and

$$< Y^\perp, X >= m(x^2+y^2) - y\sqrt{x^2+y^2},$$

by equation (3.18) PRCs are given by

$$\nabla\vartheta(\gamma(\theta)) = \frac{1}{\beta(m-\sin(\Omega(\theta)))}(-\sqrt{\beta}\sin(\Omega(\theta)), \sqrt{\beta}\cos(\Omega(\theta))).$$

That is,

$$PRC_1(\theta) = -\frac{\sin(\Omega(\theta))}{\sqrt{\beta}\,(m-\sin(\Omega(\theta))},$$

see Fig. 3.1, and $PRC_2(\theta) = \dfrac{\cos(\Omega(\theta))}{\sqrt{\beta}\,(m-\sin(\Omega(\theta))}$.

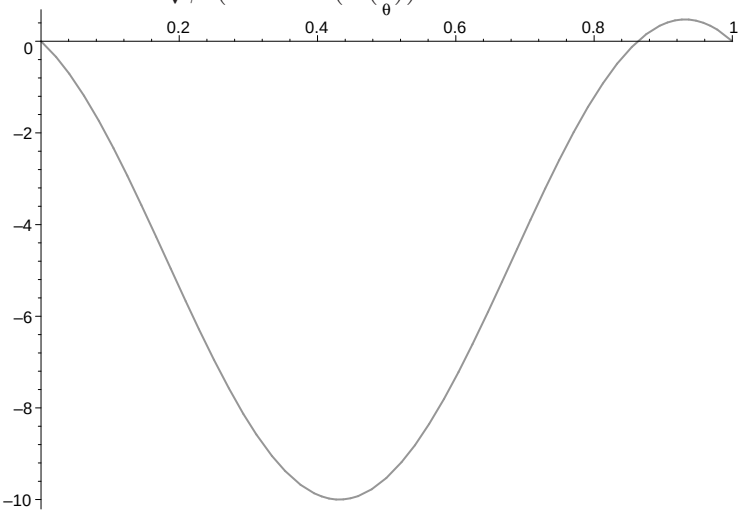

Figure 3.1: PRC_1 for system (3.37) with $m=1.1$, $\beta=1$. Phase has been scaled to $[0,1]$. Observe the slightly positive bump when $\theta\in(0.9,1)$ approx.

3.9.1 Numerical examples

We present here an application of our models to a set of examples we found relevant either to illustrate the properties or because they are representative of classical models:

1. The Van der Pol oscillator:

$$\begin{cases} \dot{x} &= -y + x - x^3, \\ \dot{y} &= x. \end{cases} \tag{3.39}$$

2. A reduced Hodgkin-Huxley-like system, with sodium and potassium currents, and only one gating variable:

$$\begin{cases} \dot{V} &= -\frac{1}{C_m}(g_{Na}m_\infty(V)(V-V_{Na}) + g_K n(V-V_K) + g_L(V-V_L) - I_{app}), \\ \dot{n} &= n_\infty(V) - n, \end{cases} \tag{3.40}$$

where V represents the membrane potential, n is a gating variable, the open-state probability functions are

$$m_\infty(V) = \frac{1}{1+\exp(-(V-V_{max,m})/k_m)}, \qquad n_\infty(V) = \frac{1}{1+\exp(-(V-V_{max,n})/k_n)},$$

and the parameters are $C_m = 1.$, $g_{Na} = 20.$, $V_{Na} = 60.$, $g_K = 10.$, $V_K = -90.$, $g_L = 8.$, $v_L = -80.$, $V_{max,m} = -20.$, $k_m = 15.$, $V_{max,n} = -25.$, $k_n = 5.$

3. The Selkov model (see [Sel68]), initially a model for self-oscillations in glycolysis, which has also been extensively used in models for circadian rhythms (see for instance [DN94] and [AP07]). It is given by

$$\begin{cases} \dot{x} &= 1 - x\,y, \\ \dot{y} &= a\,y\,(x - (1+b)/(1+b\,y)), \end{cases} \tag{3.41}$$

where the parameters are $a, b \in \mathbb{R}$.

4. The Morris-Lecar model (see [ML81]), initially conceived as a model for a barnacle giant muscle fiber, but well-studied in the neuroscience literature (after [RE98]) as a paradigm for the different bifurcations that give rise to limit cycles. The model is given by:

$$\begin{cases} \dot{V} &= \frac{1}{C}\left(I - g_L(V-V_L) - g_K\omega(V-V_K) - g_{Ca}m_\infty(V)(V-V_{Ca})\right), \\ \dot{w} &= \phi\dfrac{w_\infty(V)-w}{\tau_w(V)}, \end{cases} \tag{3.42}$$

where

$$\begin{aligned} m_\infty(V) &= \frac{1}{2}(1+\tanh((V-V_1)/V_2)), \\ w_\infty(V) &= \frac{1}{2}(1+\tanh((V-V_3)/V_4)), \text{ and} \\ \tau_w(V) &= (\cosh((V-V_3)/(2V_4)))^{-1}, \end{aligned}$$

and the parameters are $V_L = -60$, $V_K = -84$, $V_{Ca} = 120$, $V_1 = -1.2$, $V_2 = 18$, $V_3 = 2$, $V_4 = 30$, $g_L = 2$, $g_K = 8.0$, $g_{Ca} = 4.4$, $C = 20$ and $\phi = 0.04$.

All these examples share common characteristics with slight differences that will be remarked at the end of this section. Let us start, then, with the common features.

Model	Figure	$T \approx$	$\lambda \approx$	$L =$	$N =$	$E_{tail} \in$	$E_{loc} =$
1	3.2	6.663	-7.059	15	$2^8 = 256$	$(10^{-20}, 10^{-15})$	10^{-12}
2, $I_{app} = 10$	3.3	7.074	-27.66	5	$2^{11} = 2048$	$(10^{-17}, 10^{-12})$	10^{-8}
2, $I_{app} = 165$	3.4	1.630	-3.384	10	$2^9 = 512$	$(10^{-19}, 10^{-14})$	10^{-10}
3	3.5	6.344	-4.909	15	$2^9 = 512$	$(10^{-16}, 10^{-14})$	10^{-11}
4	3.6	99.27	-9.122	5	$2^{10} = 1024$	$(10^{-20}, 10^{-13})$	10^{-8}

Table 3.1: Parameter values for the different models: T=period of the orbit γ; λ=characteristic exponent associated to γ; L=order of the Taylor expansion; N=number of Fourier modes; $|K_n^{tail}|$=residuals for the K_n; E_{loc}=maximun error when computing local approximation of isochrons.

3.9.1.1 Common features: methods, parameter values and figure labelling

In all the cases, we are interested in studying the dynamics close to a hyperbolic limit cycle γ of period T that surrounds an unstable critical point p^*. The zero phase point on γ is the point which has a maximum value of the first component (x or V depending on the example). As in previous sections, we call λ the characteristic exponent of γ (so, the characteristic multiplier is e^{λ}). The computation of the periodic orbits has been performed using a Newton method with a tolerance of $1.0e-15$. In the neighborhood of γ, we have performed a Taylor expansion as in (3.31) up to order L and we have considered N Fourier modes for the K_n. With them we obtain residuals for the K_n as defined in (3.29), which are of order $|K_n|_{tail}$. The local approximation that we get for the isochrons defined in (3.32) is computed with an error smaller than E_{loc}, while the globalization of the manifold has been performed following (3.33) and using a Taylor method with a tolerance of order $1.0e-16$. In the globalization (see definitions after (3.33)), we require a distance of order $\Delta s = 1.0e-2$ between two consecutive points on the isochron and we fix $\Delta_{min} = 1.0e-8$ and $\Delta\alpha = 0.3$.

Values for each example of all the parameters defined in the last paragraph are given in Table 3.1. All the results will be given with 4 significant digits although all the computations have been performed with double precision.

For each model, we present a figure (Figs. 3.2 to 3.6) with different panels. In order to compact notation, we label each panel with a different symbol: (`Kn`), (`Iso`), (`PRC`), (`PRS`) and (`PR`θ).

In panels (`K`), the computed K_n, for some values of n, are shown. The fact that the orbits do not approach γ uniformly in θ has the effect that for certain values of θ, as n goes to infinity, the value of K_n goes to zero faster than for other values, see the slow-fast item in Section 3.10 for a discussion on this question.

In panels (`Iso`) we plot the isochrons corresponding to the phases j/N_ϕ, for $j = 0, \ldots, N_\phi - 1$, with $N_\phi = 16$, typically. We show the local approximation (green) computed semi-analytically using the parameterization method and the globalized isochron (red) using the dynamics given by the vector field. We restrict the computation to a rectangular domain $\mathcal{R}$ containing the limit cycle.

In panels (PRC) we plot the PRC_1 (green) and the PRC_2 (blue) for an infinitesimally small perturbation in the directions $(1,0)$ and $(0,1)$, respectively, with a certain amplitude specified in each figure caption, jointly with the x (or V) component (red line) of the oscillator (scaled for a better reference). Notice that the PRC_1, corresponding to the horizontal pulses, is just the section with $\sigma = 0$ of the PRSs that are given in panels (PRC).

In panels (PRS), we plot the PRS_1 in the variables θ and σ, but only for positive values of σ, avoiding negative values for the sake of clearness. The positive values of σ correspond to the points (x, y) in the external part of the limit cycle (depending on the orientation of the limit cycle, the sign of σ defined by the parameterization method out of the limit cycle can be also negative; in these figures we change $\sigma \mapsto -\sigma$ for the sake of homogeneity). Indeed, in these panels we plot the phase shift for the points (x, y) in the outer neighborhood of the limit cycle displayed in panels (Iso). Since we restrict to the rectangular domain $\mathcal{R}$ in the variables (x, y) (because they are the "real" phase space variables), when plotting the PRS in the variables (θ, σ), we come across with a non-regular domain. On the top of that, sometimes the discretization of local isochrons (Δs) used to globalize other isochrons undergoes the limit Δ_{min} and cannot longer extend the isochron. This is why some isochrons do not reach the border of the rectangular domain $\mathcal{R}$.

Like the isochrons, the PRS_1s are computed locally using semi-analytical methods and extended by integrating the system (3.21), see Section 3.7.4. Thus, the mesh is not completely regular, so we have used cubic interpolation with splines in order to show it in an regular grid on the plane (θ, σ).

Although the PRSs contain the maximum information about phase advancement, sometimes they are not easy to visualize. Accordingly, we have decided, in some cases, to show sections of the PRSs with fixed phases (that is, $\theta = \theta^*$ and parameterized by σ). Using this view, we can easily illustrate the differences in phase advancement between different points on the same isochron. These panels are labelled as (PRθ).

Before entering into the discussion, we go through the non-common features of each example.

Remark 3.9.3. For the Hodgkin-Huxley-like model (3.40), we have studied the system in two regimes: $I_{app} = 10$ (close to a SNIC bifurcation which occurs at $I_{app} \approx 4.513$) and $I_{app} = 165$ (closer to a Hopf bifurcation which occurs at $I_{app} \approx 213.8$). Bifurcation values are obtained through *XPPAUT*, see [Erm02].

For the case $I_{app} = 10$, which is the case close to SNIC, the system presents a slow-fast dynamics that will accentuate some of the problems that we already mentioned in the previous example. In this case, we computed the Taylor expansion up to order $L = 5$, because, as we can appreciate in Fig. 3.3, for some values of θ, as n increases the K_n tend to zero much faster. Moreover, since the functions K_n present very sharp spikes we need to consider up to $2^{10} = 1024$ or $2^{11} = 2048$ Fourier modes to get good approximations, that is with residuals smaller than a certain error. This implies solving linear systems with large matrices that are not very stable.

Notice that this limit cycle is "strongly" hyperbolic and the backwards integration can be somehow very unstable.

The system presents a fixed point at $(-26.83, 0.4093)$, which is computed using a

Newton method with a tolerance of $1.0e-13$.

Once can appreciate in Fig. 3.3 that the isochrons computed semi-analytically (green ones, hardly noticeable) are shorter than in the Hopf case (Fig. Fig. 3.4). In this case, in order to have a long enough local approximation for the isochrons we reduced the accuracy of the computation down to $1.0e-08$.

For the case $I_{app} = 165$, which is close to the Hopf bifurcation, in Fig. 3.4 one can observe that the slow-fast phenomenon that we mentioned above is softened.

The system presents a hyperbolic fixed point for at $(-21.00, 0.6899)$, which has been computed numerically with an error smaller than $1.0e-13$.

Remark 3.9.4. The Selkov model (3.41) presents an Andronov-Hopf bifurcation for $a = (1+b)/b$. We have studied here the case $a = 3$, $b = 1$. The unstable fixed point, that can be easily computed analytically, is located at $(1, 1)$.

In this example, we have decided to allow $\Delta_{max} = 1.0e-12$ because far from the limit cycle points escape.

Remark 3.9.5. For the Morris-Lecar model (3.42), we have considered the case $I_{app} = 91$, which presents similar issues as the 2D Hodgin-Huxley model close to a SNIC considered above.

The system has a fixed point at $(-26.26, 0.1320)$, which is computed using a Newton method with a tolerance of $1.0e-13$.

It is to be noticed that equation (3.42) presents a subcritical Hopf bifurcation at $I_{app} \approx 93.86$; the unstable limit cycle goes "back" in the parameter space up to $I_{app} \approx 88.29$, where it coalesces with a stable limit cycle in a bifurcation of a semistable limit cycle that disappears for lower I_{app}'s. The stable limit cycle, which comes from another bifurcation for some $I_{app} \gg 93.86$, is the one that we study. It can be checked that the period of this stable orbit is notably above of that of the unstable orbit, born at the subcritical Hopf bifurcation; thus, one may expect a more slow-fast dynamics, more similar to a limit cycle close to a SNIC bifurcation than to a Hopf bifurcation.

Some of the shortcomings of the Hodgkin-Huxley close to a SNIC are reproduced also here: it is necessary to solve linear systems with large matrices that are not very stable; the limit cycle is "strongly" hyperbolic and the backwards integration can be somehow very unstable; to have a long enough local approximations for the isochrons we need to reduce the accuracy of the computation.

Another specific observation is that the isochrons spiral around the unstable limit cycle in the interior of the stable one. This is not surprising since the two limit cycles have different periods and so, they cannot share the system of isochrons. Next short example illustrates analytically this fact.

Example 3.9.6. Consider the $\mathcal{C}^1$ system in polar coordinates

$$X := \begin{cases} \dot{r} &= r\, a(r), \\ \dot{\theta} &= b(r). \end{cases}$$

Suppose that $a(r_j) = 0$, $a'(r_j) \neq 0$ and $b(r_j) \neq 0$, for $j = 1, 2$, with $r_1 \neq r_2$, both positive.

It is straightforward to see that $\{r = r_1\}$ and $\{r = r_2\}$ are hyperbolic limit cycles of X with alternate stability. From [FGG07], we can deduce that, for each limit cycle, the

vector field

$$Y_j := \begin{cases} \dot{r} &= r, \\ \dot{\theta} &= (b(r) - b(r_j))/(a(r)) \end{cases}$$

satisfies $[Y_j, X] = \mu\, Y_j$, with $\mu(r) = r\, a'(r)$.

Let us take now Y_1. It is clear that, in general, $\dot{\theta}$ is not defined on $r = r_2$ since $a(r_2) = 0$. However, if $b(r_2) = b(r_1)$ (both limit cycles have the same period!), then $\dot{\theta}$ may be extended on $r = r_2$ and thus, $\{r = r_2\}$ may be contained in the domain Ω where the isochrons of $\{r = r_1\}$ are defined. This is the case, for instance, when the system is rigid ($\dot{\theta}$ =constant), for which the isochrons are straight lines from the origin.

In the case that $b(r_2) \neq b(r_1)$ (different periods like the numerical example illustrated in Fig. 3.6), $\dot{\theta}$ in Y_1 is not bounded close to $\{r = r_2\}$ and, then, the isochrons of $\{r = r_1\}$ spiral around.

3.10 Discussion

An integrated way to study the dynamics around a limit cycle.

We want to emphasize on the completeness of the method presented in this memoir. We provide a way to reconcile different concepts, from Lie symmetries to phase resetting curves and surfaces, through the implementation of the parameterization method. Although computing PRSs has been presented as our last goal, along the way we have related all the different concepts involved in the parameterization of a neighborhood of a periodic orbit in $\mathbb{R}^2$: choice of "canonical co-ordinates" inspired by the Lie symmetry, computation of isochrons, PRCs and, finally, PRSs. We have established, as well, a link between different parts of the scientific literature that are not usually connected: theoretical and numerical methods for invariant objects, qualitative theory of (mostly planar) ordinary differential equations and theoretical (neuro)biology.

We have also taken care of the numerical aspects involved in the method, which are not trivial and show up relationships among geometry, dynamics and numerical schemes. Concerning to the practical part of the effective computation, the above examples have shed light upon different biological and numerical issues that we would like to remark next.

From "Type 1" to "Type 2" PRCs.

As we mentioned in the Introduction, from [Erm96], PRCs are classified between models with strictly positive or mainly positive PRC ("Type 1" or "Class 1"), and models whose PRC changes sign ("Type 2" or "Class 2"). The "rule of thumb" proposed by Ermentrout is that Type 1 PRC correspond to models in which oscillations appear via saddle-node on invariant circle bifurcations, whereas Type 2 PRC correspond to supercritical Andronov-Hopf bifurcation.

Our examples confirm this rough classification and we have used them to give an idea how the transition between the two PRC types takes place. Another study, using only PRCs and continuation methods, has recently appeared in [GS06].

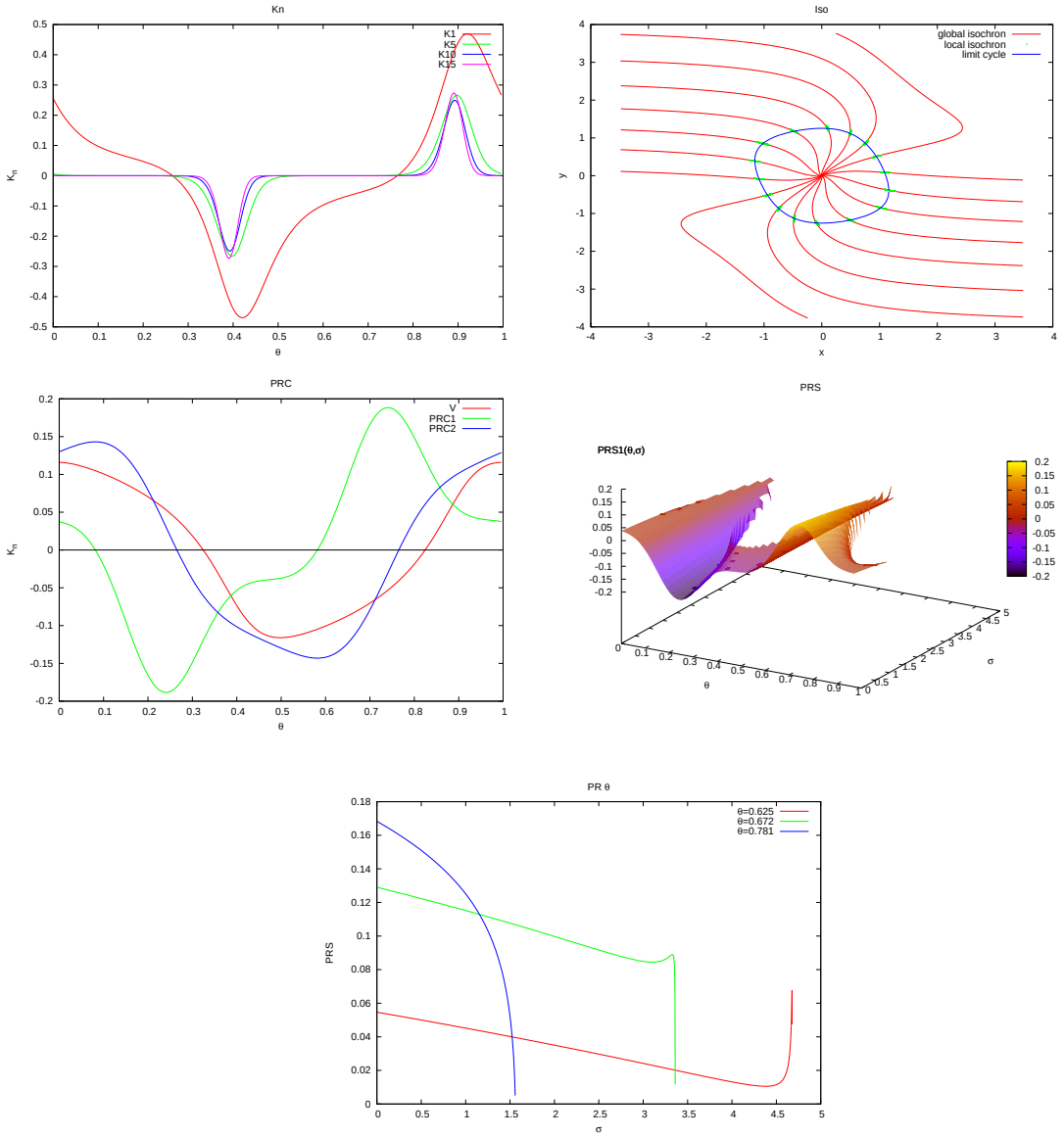

Figure 3.2: The van der Pol oscillator. In the panel PRC, we show PRC_1 and PRC_2 with amplitude 1 and x scaled by a factor 0.1. See Section 3.9.1.1 for a general explanation about the contents of each panel. In panel ($\mathtt{PR}\theta$), notice the diversity of phase advancements that can be obtained in the same isochron (three isochrons are shown: $\theta = 0.625, 0.672, 0.781$).

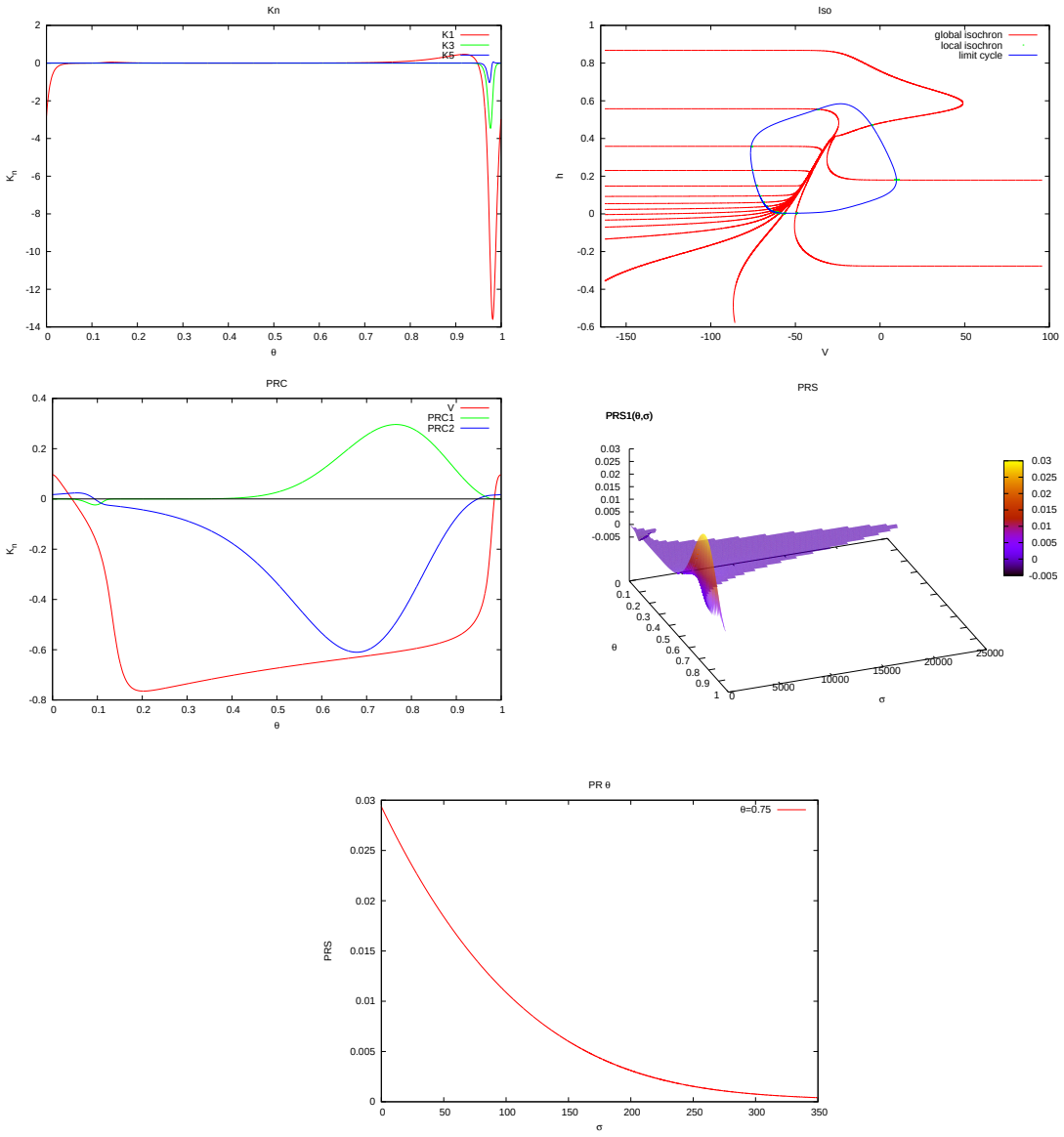

Figure 3.3: Two dimensional Hodgkin-Huxley close to a SNIC bifurcation. In the panel PRC, we show PRC_1 and PRC_2 with amplitude 10 and 0.1 respectively, and V scaled by a factor 0.01. See Section 3.9.1.1 for a general explanation about the contents of each panel. In panel (`Iso`), notice the heterogeneity in the distances between isochrons with equidistant phases, thus reflecting the slow-fast nature of the system.

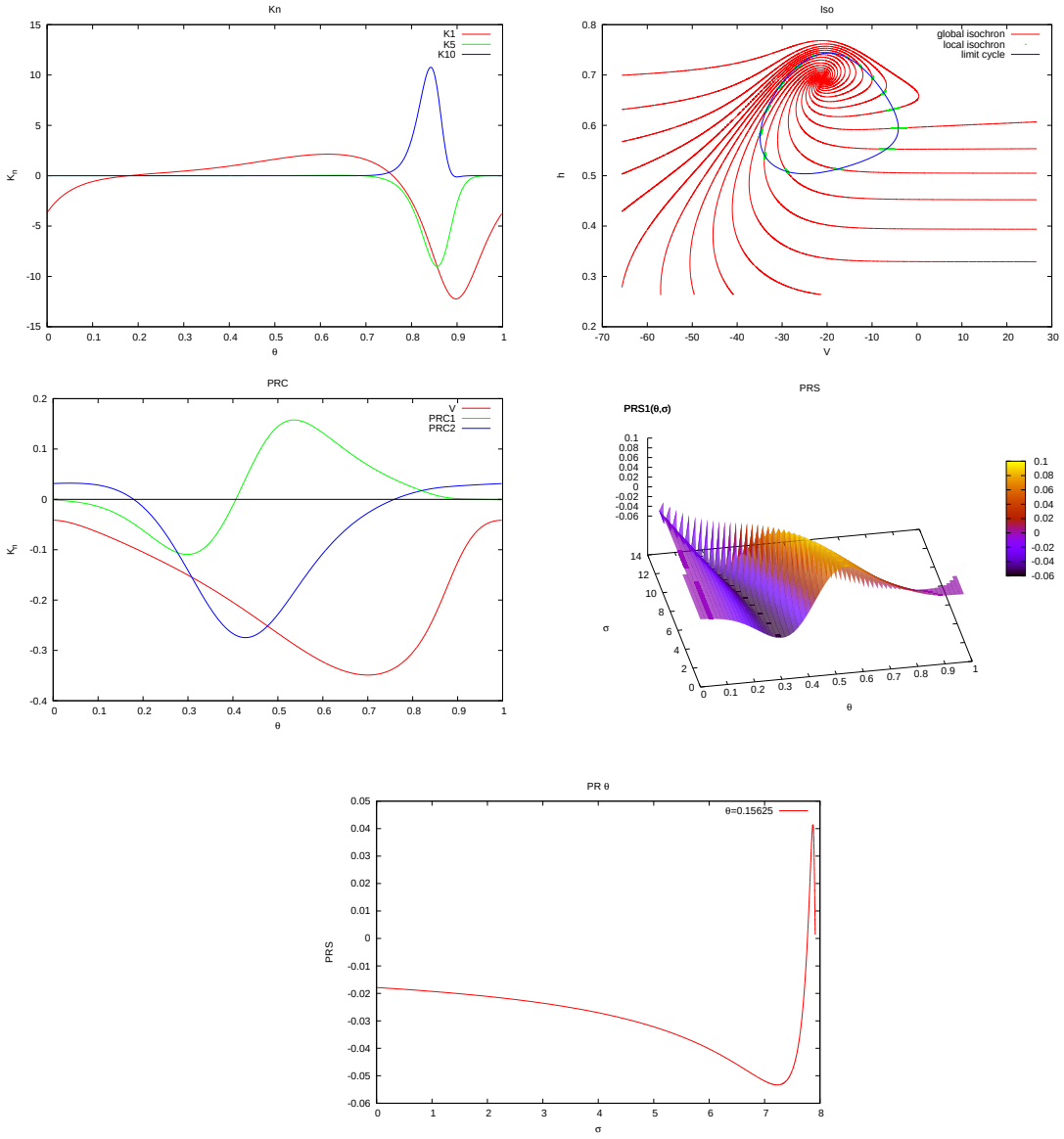

Figure 3.4: Two dimensional Hodgkin-Huxley close to a Hopf bifurcation. In the panel PRC, we show PRC_1 and PRC_2 with amplitude 2 and 0.02 respectively, and V scaled by a factor 0.01. See Section 3.9.1.1 for a general explanation about the contents of each panel.

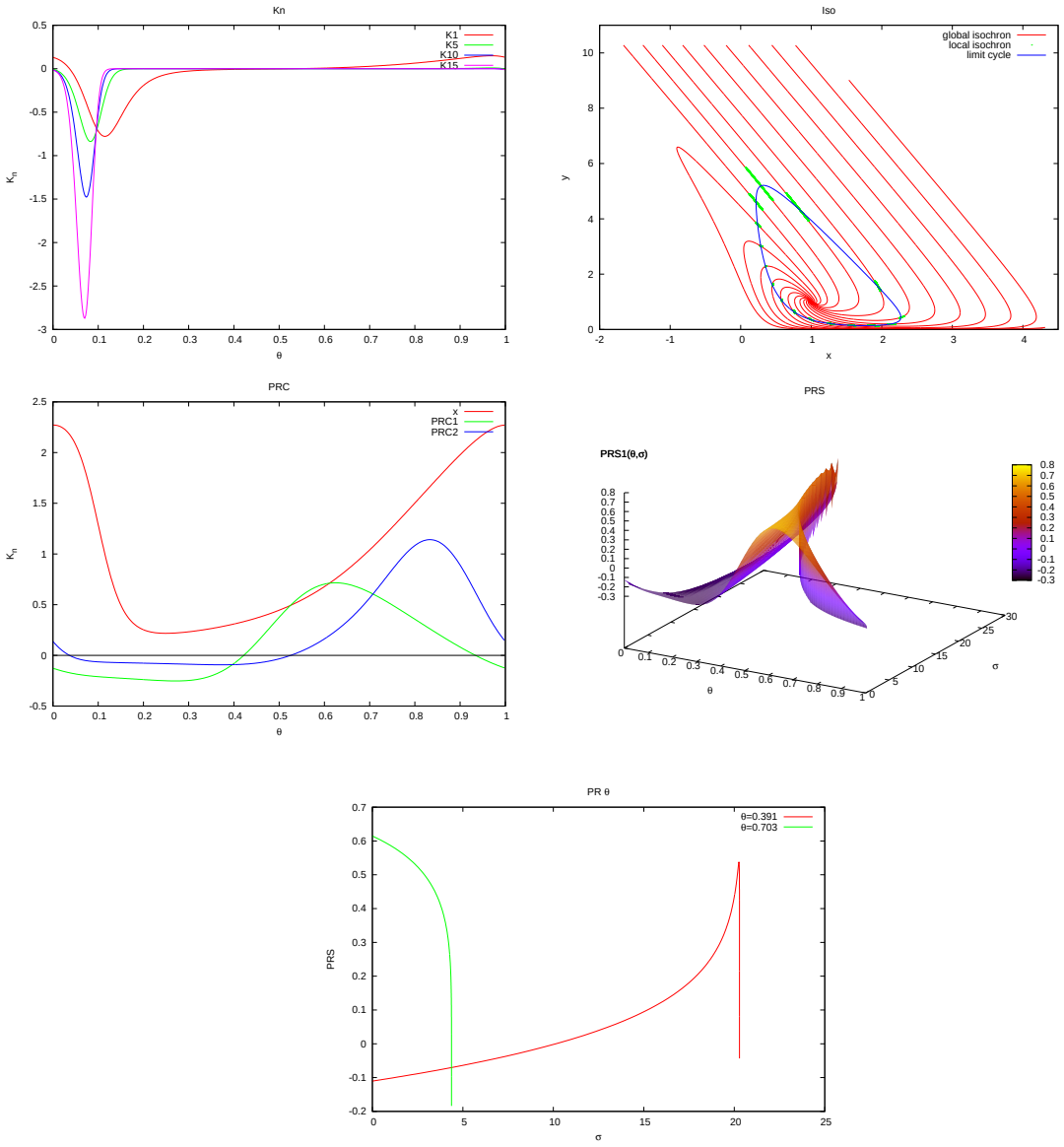

Figure 3.5: Selkov model with $a = 3$ and $b = 1$, not far from a Hopf bifurcation. See Section 3.9.1.1 for a general explanation about the contents of each panel.

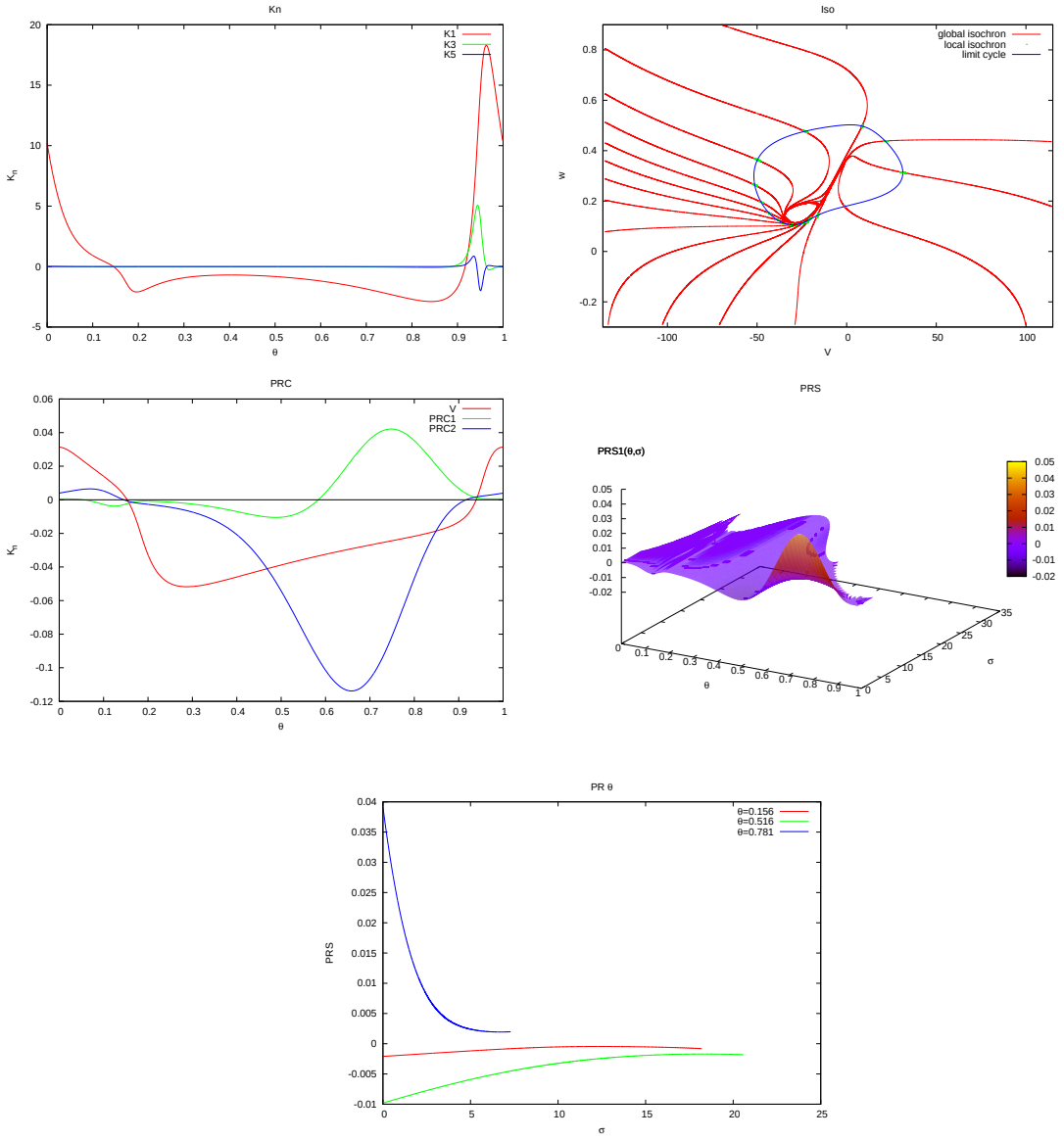

Figure 3.6: Morris-Lecar model. In the panel PRC, we show PRC_1 and PRC_2 with amplitude 1 and 0.01 respectively, and V scaled by a factor 0.001. See Section 3.9.1.1 for a general explanation about the contents of each panel. In panel (`Iso`), notice how isochrons spiral around the unstable limit cycle, see also Remark 3.9.5.

Analytically, we have seen for instance (see Example 3.9.1) that $PRC_1(\theta) = -1/\sqrt{\beta}\sin(\theta)$ close to a Hopf bifurcation and (see Example 3.9.2), $PRC_1(\theta) = -\sin(\Omega(\theta))/(\sqrt{\beta}(m - \sin(\Omega(\theta)))$, where Ω is given in (3.38), for a system that presents a SNIC bifurcation at $m = 1$ (see also Fig. 3.1).

These examples clearly show that the fact that systems with oscillations coming from a saddle-node bifurcation are of "Type 1", that is, the PRC is mainly positive (or mainly negative), is produced by the slow-fast dynamics; moreover, the time it spends on a negative (resp., positive) regime is very short compared to the time it spends out of it. This fact can also be observed when comparing (3.40) with $I_{app} = 10$ (see Fig. (3.3)) with (3.40) with $I_{app} = 165$ (see Fig. (3.4)). For instance, in Fig. (3.3), panel (`PRC`), we can appreciate a small negative portion of the PRC_1 (close to $\theta = 0.1$), whereas in the same panel of Fig. (3.4), the negative part has a bigger area. As the value I_{app} increases from $I_{app} = 10$ on, the negative portion of the PRC is enlarging up to the bifurcation point $I_{app} \approx 213.8$, where the PRC has practically zero integral.

Observe also that systems with a marked "Type II" tendency (mainly, (3.39), (3.40) with $I_{app} = 165$ or (3.41)) reach the extreme values of the PRCs when the isochrons have maximal curvature nearby the limit cycle (the correspondence is not exact because it also depends on the stimulus direction). Compare panels (`PRC`) with panels (`Iso`) in Figs. 3.2, 3.4 and 3.5. On the other hand, the chosen value for the Morris-Lecar system (see Fig. 3.6) shows an intermediate behavior, perhaps closer to "Type II" than to "Type I" (recall from (3.9.5) that it takes place close to a double limit cycle bifurcation).

Numerical drawbacks for slow-fast systems.

From a joint analysis of panels (`Kn`) and (`Iso`) we can deduce the effect of slow versus fast dynamics. In Fig. 3.2, for instance, one can appreciate two features close to values $\theta \approx 0.4$ or $\theta \approx 0.9$: (a) the $K_n \neq 0$ for every n; (b) the isochrons on these zones are more separated (though the phases are equidistant). In other words, the slow dynamics makes the convergence of the Fourier expansions of our method to slow down as well. The numerical problem is that the value of K_n goes to zero faster in some zones, and it is then impossible to choose a uniform b (see Remark 3.7.2) such that the K_n can be kept to have order 1 for all the values of θ. Thus, for these values of θ, increasing the degree L of the Taylor polynomial does not have any effect in the growth of the domain where the isochron can be computed semi-analytically. This situation turns out to be a serious issue for the cases when the systems present an accentuated slow-fast dynamics.

This effect is even more dramatic in system (3.40) close to the SNIC bifurcation, see Fig. 3.3. In this case, (observe panel (`Kn`)) the zone close to $\theta = 0$, where the K_n take values orders of magnitude above the rest of the cycle; that is, $\dfrac{\max\limits_{\theta} |K_n(\theta)|}{\min\limits_{\theta} |K_n(\theta)|} \gg 1$. This coincides again with the zone where the isochrons are more noticeable separated, panel (`Iso`) of Fig. 3.3.

On the other hand, the factor $\dfrac{\max\limits_{\theta} |K_n(\theta)|}{\min\limits_{\theta} |K_n(\theta)|}$ is more attenuated in (3.40) close to the

Hopf bifurcation, see Fig. 3.4.

Role of PRSs under high frequency stimulations.

As explained in Section 3.1, the phase advancement computed on the limit cycle (PRC) can differ from that computed out of the limit cycle (PRS). This difference will be important under different circumstances like a short period of stimulation, a slow attraction to the limit cycle, a large stimulus amplitude, environmental random fluctuations, bursting-like stimuli, etc. We fix our attention, now, to panels (PRS) and (PRθ) in all the figures. Our purpose is to highlight the differences in the phase advancement for points in a neighborhood of γ which share the same phase θ.

As with the maximum and minimum values of the PRCs, we pay attention to the zones where the curvature of the isochrons is extreme close to the limit cycle. This phenomenon can be timidly observed in Fig. 3.3, panel (PRθ), where the section $\theta = 0.75$ of the PRS is shown to be decreasing. More exaggerated variations can be obtained for $\theta = 0.15625$ in Fig. 3.4, panel (PRθ), or in Fig. 3.2. In the latter case, we show a *zoo* of possible sections ($\theta \in \{0.672, 0.625, 0.781\}$).

Changes of the phase advancement with respect to σ given a fixed θ, as the above examples show, combined with high frequency stimuli, rule out the possibility of controlling the whole phase advancement of an experiment using only the PRC. Thus, the PRSs become extremely useful. These differences, as our examples, show are more noticeable close to Type II oscillators (Hopf) because of the stronger curvature of the isochrons.

Our results agree with the fact that perturbations applied to Type II oscillators produce significant normal displacements from the limit cycle rather those applied to Type I. This fact has been also studied in [OC02] by Oprisan and Canavier, in the sense that the difference in angular velocity at displaced points compared to the angular velocity on the limit cycle is then more important. As pointed out in [OC02], this might affect the study of biological circuits comprising Type II neural oscillators, which appear frequently in identified central pattern-generating circuits.

Isochrons in higher dimensions.

Although in this memoir we only apply the method to compute isochrons and PRCs to planar differential systems, it can be applied to higher dimensions provided that the limit cycle is hyperbolic and stable. In higher dimensions, we would like to emphasize the interesting question about the existence of an isochronous foliation when a limit cycle is not stable (i.e., some of the characteristic multipliers are bigger than one). In Section 3.8 we have shown that for a given vector field $X \in \mathbb{R}^n$, $n \geq 2$, with a limit cycle γ, if there exist $n-1$ non-trivial vector fields $Y_1, \ldots, Y_{n-1}$ in involution satisfying the Lie symmetry equation (3.4), then γ is isochronous and the foliation can be defined. An effective/efficient method to compute the isochrons for a limit cycle with both stable and unstable manifolds is also a challenge to which the Lie symmetries approach can help. We have seen in Section 3.8 that the problem of finding isochrons is equivalent to solving equation (3.35); that is, finding a 1-form ω_Y and a function λ satisfying $X(\omega_Y) = \lambda(\mathrm{x})\,\omega_Y$. The leaves of ω_Y will then be the isochronous sections of the limit cycle. Developing numerical methods to solve equation (3.35) would then give the isochronous sections,

independently of the dimensions of the stable and unstable manifolds associated to the (hyperbolic) limit cycle. We want to emphasize, however, that this has more theoretical than practical interest since in models one usually encounters (hyperbolic) stable limit cycles. In this case, solving (3.35) would be equivalent to finding the stable manifold; the parameterization method could be useful as well, as in the bidimensional case.

Bibliography

[ALD83] S. Aubry and P. Y. Le Daeron. The discrete Frenkel-Kontorova model and its extensions. I. Exact results for the ground-states. *Phys. D*, 8(3):381–422, 1983.

[Ano69] D. V. Anosov. *Geodesic flows on closed Riemann manifolds with negative curvature.* American Mathematical Society, Providence, R.I., 1969.

[AP07] J.M. Gimble A.A. Ptitsyn, S. Zvonic. Digital signal processing reveals circadian baseline oscillation in majority of mammalian genes. *PLoS Comput Biol*, 3(6), 2007.

[Arn63a] V. I. Arnol'd. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations. *Russian Math. Surveys*, 18(5):9–36, 1963.

[Arn63b] V. I. Arnol′d. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russ. Math. Surveys*, 18:85–192, 1963.

[Arn64] V.I. Arnold. Instability of dynamical systems with several degrees of freedom. *Sov. Math. Doklady*, 5:581–585, 1964.

[Aub83] S. Aubry. The twist map, the extended Frenkel-Kontorova model and the devil's staircase. *Phys. D*, 7(3):240–258, 1983.

[BHM04] Eric Brown, Philip Holmes, and Jeff Moehlis. On the phase reduction and response dynamics of neural oscillator populations. *Neural Comp.*, 16:673–715, 2004.

[CFL03a] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. I. Manifolds associated to non-resonant subspaces. *Indiana Univ. Math. J.*, 52(2):283–328, 2003.

[CFL03b] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. II. Regularity with respect to parameters. *Indiana Univ. Math. J.*, 52(2):329–360, 2003.

[CFL04] A. Celletti, C. Falcolini, and U. Locatelli. On the break-down threshold of invariant tori in four dimensional maps. *Regul. Chaotic Dyn.*, 9(3):227–253, 2004.

[CFL05] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds. III. Overview and applications. *J. Differential Equations*, 218(2):444–515, 2005.

[CH82] Shui Nee Chow and Jack K. Hale. *Methods of bifurcation theory*, volume 251 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science]*. Springer-Verlag, New York, 1982.

[Chi79] B.V. Chirikov. A universal instability of many-dimensional oscillator systems. *Phys. Rep.*, 52(5):264–379, 1979.

[CKWJ03] C. Cajochen, K. Kräuchi, and A. Wirz-Justice. Role of melatonin in the regulation of human circadian rhythms and sleep. *J. Neuroendocrinology*, 15:432–437, 2003.

[CL04] Carmen Chicone and Weishi Liu. Asymptotic phase revisited. *J. Differential Equations*, 204(1):227–246, 2004.

[CL08] R. Calleja and R. de la Llave. Fast numerical computation of quasi-periodic equilibrium states in 1-d statistical mechanics, including twist maps. 2008. Manuscript.

[Con90] John B. Conway. *A course in functional analysis*, volume 96 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[CY04a] Chong-Qing Cheng and Jun Yan. Existence of diffusion orbits in a priori unstable Hamiltonian systems. *J. Differential Geom.*, 67(3):457–517, 2004.

[CY04b] Chong-Quin Cheng and Jun Yan. Arnold diffusion in Hamiltonian systems: 1 a priori unstable case. Preprint 04-265, `http://www.ma.utexas.edu/mp_arc/`, 2004.

[DH08] Amadeu Delshams and Gemma Huguet. The large gap problem in arnold diffusion for non polynomial perturbations of an a-priori unstable hamiltonian system. Manuscript, 2008.

[DLdlS03] Amadeu Delshams, Rafael Llave de la, and Tere M. Seara. A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: announcement of results. *Electron. Res. Announc. Amer. Math. Soc.*, 9:125–134 (electronic), 2003.

[DLS00] A. Delshams, R. de la Llave, and T.M. Seara. A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic geodesic flows of $\mathbb{T}^2$. *Comm. Math. Phys.*, 209(2):353–392, 2000.

[DLS06a] A. Delshams, R. de la Llave, and T. M. Seara. A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. *Mem. Amer. Math. Soc.*, 179(844):viii+141, 2006.

[DLS06b] Amadeu Delshams, Rafael de la Llave, and Tere M. Seara. Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows. *Adv. Math.*, 202(1):64–188, 2006.

[DLS07] Amadeu Delshams, Rafael de la Llave, and Tere M. Seara. Instability of high dimensional hamiltonian systems: Multiple resonances do not impede diffusion. Manuscript, 2007.

[DLS08] Amadeu Delshams, Rafael de la Llave, and Tere M. Seara. Geometric properties of the scattering map of a normally hyperbolic invariant manifold. *Adv. Math.*, 217(3):1096–1153, 2008.

[DN94] Antoni Diez-Noguera. A functional model of the circadian system based on the degree of intercommunication in a complex system. *Am J Physiol Regul Integr Comp Physiol*, 267:R1118–R1135, 1994.

[DVV02] Luca Dieci and Erik S. Van Vleck. Lyapunov spectral intervals: theory and computation. *SIAM J. Numer. Anal.*, 40(2):516–542 (electronic), 2002.

[EK91] G. B. Ermentrout and N. Kopell. Multiple pulse interactions and averaging in systems of coupled neural oscillators. *J. Math. Biol.*, 29(3):195–217, 1991.

[Erm96] G. B. Ermentrout. Type i membranes, phase resetting curves, and synchrony. *Neural Comp.*, 8:979–1001, 1996.

[Erm02] Bard Ermentrout. *Simulating, analyzing, and animating dynamical systems*, volume 14 of *Software, Environments, and Tools*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. A guide to XPPAUT for researchers and students.

[Fen72] Neil Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.*, 21:193–226, 1971/1972.

[Fen77] N. Fenichel. Asymptotic stability with rate conditions. II. *Indiana Univ. Math. J.*, 26(1):81–93, 1977.

[Fen79] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations*, 31(1):53–98, 1979.

[Fen74] N. Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[FGG07] Emilio Freire, Armengol Gasull, and Antoni Guillamon. Limit cycles and Lie symmetries. *Bull. Sci. Math.*, 131(6):501–517, 2007.

[FKW01] Bassam Fayad, Anatole Katok, and Alistar Windsor. Mixed spectrum reparameterizations of linear flows on $\mathbb{T}^2$. *Mosc. Math. J.*, 1(4):521–537, 644, 2001. Dedicated to the memory of I. G. Petrovskii on the occasion of his 100th anniversary.

[FLS07] E. Fontich, R. de la Llave, and Y. Sire. Construction of invariant whiskered tori by a parametrization method. part i: Maps and flows in finite dimensions. Preprint, 2007.

[FM01] Ernest Fontich and Pau Martín. Arnold diffusion in perturbations of analytic integrable Hamiltonian systems. *Discrete Contin. Dynam. Systems*, 7(1):61–84, 2001.

[FMM77] George E. Forsythe, Michael A. Malcolm, and Cleve B. Moler. *Computer methods for mathematical computations*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1977. Prentice-Hall Series in Automatic Computation.

[Fro72] C. Froeschlé. Numerical study of a four-dimensional mapping. *Astronom. and Astrophys.*, 16:172–189, 1972.

[GJSM01a] G. Gómez, À. Jorba, C. Simó, and J. Masdemont. *Dynamics and mission design near libration points. Vol. III*, volume 4 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Advanced methods for collinear points.

[GJSM01b] G. Gómez, À. Jorba, C. Simó, and J. Masdemont. *Dynamics and mission design near libration points. Vol. IV*, volume 5 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Advanced methods for triangular points.

[GL06a] Marian Gidea and Rafael de la Llave. Arnold diffusion with optimal time in the large gap problem. *Preprint*, 2006.

[GL06b] Marian Gidea and Rafael de la Llave. Topological methods in the instability problem of Hamiltonian systems. *Discrete Contin. Dyn. Syst.*, 14(2):295–328, 2006.

[Gre79] J. M. Greene. A method for determining a stochastic transition. *Jour. Math. Phys.*, 20:1183–1201, 1979.

[Gri00] Andreas Griewank. *Evaluating derivatives*, volume 19 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Principles and techniques of algorithmic differentiation.

[GS06] W. Govaerts and B. Sautois. Computation of the phase response curve: a direct numerical approach. *Neural Comput.*, 18(4):817–847, 2006.

[Guc75] J. Guckenheimer. Isochrons and phaseless sets. *J. Math. Biol.*, 1(3):259–273, 1974/75.

[GVL96] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

[Har08] Alex Haro. Automatic differentiation tools in computational dynamical systems. Manuscript, 2008.

[Her83] Michael-R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, volume 103 of *Astérisque*. Société Mathématique de France, Paris, 1983. With an appendix by Albert Fathi, With an English summary.

[Her92] M.-R. Herman. On the dynamics of Lagrangian tori invariant by symplectic diffeomorphisms. In *Progress in Variational Methods in Hamiltonian Systems and Elliptic Equations (L'Aquila, 1990)*, pages 92–112. Longman Sci. Tech., Harlow, 1992.

[HL00] A. Haro and R. de la Llave. New mechanisms for lack of equipartion of energy. *Phys. Rev. Lett.*, 89(7):1859–1862, 2000.

[HL06a] À. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: numerical algorithms. *Discrete Contin. Dyn. Syst. Ser. B*, 6(6):1261–1300 (electronic), 2006.

[HL06b] A. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: rigorous results. *J. Differential Equations*, 228(2):530–579, 2006.

[HL07] A. Haro and R. de la Llave. A parameterization method for the computation of whiskers in quasi periodic maps: numerical implementation and examples. *SIAM Jour. Appl. Dyn. Syst.*, 6(1):142–207, 2007.

[HPS77] M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds*, volume 583 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1977.

[HS95] A. Haro and C. Simó. A numerical study of the breakdown of invariant tori in 4D symplectic maps. In *XIV CEDYA/IV Congress of Applied Mathematics (Spanish)(Vic, 1995)*, page 9 pp. (electronic). Univ. Barcelona, Barcelona, 1995.

[Izh07] Eugene M. Izhikevich. *Dynamical systems in neuroscience: the geometry of excitability and bursting*. Computational Neuroscience. MIT Press, Cambridge, MA, 2007.

[Jos00] Krešimir Josić. Synchronization of chaotic systems and invariant manifolds. *Nonlinearity*, 13(4):1321–1336, 2000.

[JZ05] Àngel Jorba and Maorong Zou. A software package for the numerical integration of ODEs by means of high-order Taylor methods. *Experiment. Math.*, 14(1):99–117, 2005.

[Knu97] Donald E. Knuth. *The art of computer programming. Vol. 2: Seminumerical algorithms*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont, third revised edition, 1997.

[KO98] Bernd Krauskopf and Hinke Osinga. Growing 1D and quasi-2D unstable manifolds of maps. *J. Comput. Phys.*, 146(1):404–419, 1998.

[LGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[Lla00] R. de la Llave. Persistence of hyperbolic manifolds, 2000.

[Lla01a] Rafael de la Llave. Remarks on Sobolev regularity in Anosov systems. *Ergodic Theory Dynam. Systems*, 21(4):1139–1180, 2001.

[Lla01b] Rafael de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, pages 175–292. Amer. Math. Soc., Providence, RI, 2001.

[LM88] P. Lochak and C. Meunier. *Multiphase averaging for classical systems*, volume 72 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1988. With applications to adiabatic theorems, Translated from the French by H. S. Dumas.

[LO99] R. de la Llave and R. Obaya. Regularity of the composition operator in spaces of Hölder functions. *Discrete Contin. Dynam. Systems*, 5(1):157–184, 1999.

[LW04] R. de la Llave and C. E. Wayne. Whiskered and low dimensional tori in nearly integrable Hamiltonian systems. *Math. Phys. Electron. J.*, 10:Paper 5, 45 pp. (electronic), 2004.

[McK82] R. S. McKay. *Renormalisation in Area Preserving Maps*. PhD thesis, Princeton University, 1982.

[ML81] C. Morris and H. Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.*, 35:193–213, 1981.

[Moe96] Richard Moeckel. Transition tori in the five-body problem. *J. Differential Equations*, 129(2):290–314, 1996.

[MS89] Kenneth R. Meyer and George R. Sell. Mel′nikov transforms, Bernoulli bundles, and almost periodic perturbations. *Trans. Amer. Math. Soc.*, 314(1):63–105, 1989.

[OC02] S. Oprisan and C. Canavier. The influence of limit cycle topology on the phase resetting curve. *Neural Comp.*, 14:1027–1057, 2002.

[Olv93] Peter J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.

[OP08] A. Olvera and N. P. Petrov. Regularity properties of critical invariant circles of twist maps and their universality. *SIAM J. Appl. Dyn. Syst.*, 2008.

[Ose68] V. I. Oseledec. A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. *Trudy Moskov. Mat. Obšč.*, 19:179–210, 1968.

[Per79] I. C. Percival. A variational principle for invariant tori of fixed frequency. *J. Phys. A*, 12(3):L57–L60, 1979.

[PT07] G. N. Piftankin and D. V. Treshchëv. Separatrix mapping in Hamiltonian systems. *Uspekhi Mat. Nauk*, 62(2(374)):3–108, 2007.

[Pui02] J. Puig. Reducibility of linear differential equations with quasi-periodic coefficients: a survey. `http://www.ma.utexas.edu/mp_arc`, 02–246, 2002.

[RE98] J. Rinzel and G.B. Ermentrout. Analysis of neural excitability and oscillations. In C. Koch and I. Segev, editors, *Methods in Neural Modeling*, pages 135–169, Cambridge, MA, 1998. MIT Press.

[Rüs75] H. Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical Systems, Theory and Applications (Battelle Rencontres, Seattle, Wash., 1974)*, pages 598–624. Lecture Notes in Phys., Vol. 38, Berlin, 1975. Springer.

[Sab05] M. Sabatini. Isochronous sections via normalizers. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 15(9):3031–3037, 2005.

[Sac78] R.J. Sacker. Existence of dichotomies and invariant splittings for linear differential systems. IV. *J. Differential Equations*, 27(1):106–137, 1978.

[Sel68] E. Selkov. On the mechanism of single-frequency self-oscillations in glycolysis. a simple kinetic model. *Eur J Biochem*, 4:79–86, 1968.

[Sim90] Carles Simó. On the analytical and numerical approximation of invariant manifolds. *Les Méthodes Modernes de la Mecánique Céleste, D. Benest and C. Froeschlé (eds.). Editions Frontières*, pages 285–329, 1990.

[Sim94] Carles Simó. Averaging under fast quasiperiodic forcing. In *Hamiltonian mechanics (Toruń, 1993)*, volume 331 of *NATO Adv. Sci. Inst. Ser. B Phys.*, pages 13–34. Plenum, New York, 1994.

[Sim99] Carles Simó, editor. *Hamiltonian systems with three or more degrees of freedom*, Dordrecht, 1999. Kluwer Academic Publishers Group.

[SS74] R.J. Sacker and G.R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. I. *J. Differential Equations*, 15:429–458, 1974.

[SS76a] R.J. Sacker and G.R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. II. *J. Differential Equations*, 22(2):478–496, 1976.

[SS76b] R.J. Sacker and G.R. Sell. Existence of dichotomies and invariant splittings for linear differential systems. III. *J. Differential Equations*, 22(2):497–522, 1976.

[SV01] Carles Simó and Clàudia Valls. A formal approximation of the splitting of separatrices in the classical Arnold's example of diffusion with two equal parameters. *Nonlinearity*, 14(6):1707–1760, 2001.

[Ten82] Jeffrey Tennyson. Resonance transport in near-integrable systems with many degrees of freedom. *Phys. D*, 5(1):123–135, 1982.

[Tre04] D. Treschev. Evolution of slow variables in a priori unstable hamiltonian systems. *Nonlinearity*, 17(5):1803–1841, 2004.

[Wig90] S. Wiggins. *Global bifurcations and chaos: Analytical methods*, volume 73 of *Appl. Math. Sci.* Springer, New York, 1990.

[Win75] A. T. Winfree. Patterns of phase compromise in biological cycles. *J. Math. Biol.*, 1(1):73–95, 1974/75.