A GEOMETRIC MECHANISM FOR DIFFUSION IN HAMILTONIAN SYSTEMS OVERCOMING THE LARGE GAP PROBLEM: HEURISTICS AND RIGOROUS VERIFICATION ON A MODEL

AMADEU DELSHAMS, RAFAEL DE LA LLAVE, AND TERE M. SEARA

ABSTRACT. We introduce a geometric mechanism for diffusion in a priori unstable nearly integrable dynamical systems. It is based on the observation that resonances, besides destroying the primary KAM tori, create secondary tori and tori of lower dimension. We argue that these objects created by resonances can be incorporated in transition chains taking the place of the destroyed primary KAM tori.

We establish rigorously the existence of this mechanism in a simple model that has been studied before.

The main technique is to develop a toolkit to study, in a unified way, tori of different topologies and their invariant manifolds, their intersections as well as shadowing properties of these bi-asymptotic orbits. This toolkit is based on extending and unifying standard techniques. A new tool used here is the scattering map of normally hyperbolic invariant manifolds.

An attractive feature of the mechanism is that the transition chains are shorter in the places where the heuristic intuition and numerical experimentation suggests that the diffusion is strongest.

CONTENTS

Date: April 14, 2003.

1. Introduction

The phenomenon of diffusion in phase space for Hamiltonian systems is important for applications and has attracted a great deal of attention both in the mathematical and in the physical literature.

The rather loose (and somewhat inadequate) name *diffusion* captures the intuition that there are trajectories that wander widely and explore large regions of phase space.

Even if diffusion is, presumably, a phenomenon that happens in rather general systems – as conjectured in [Arn63b], particular attention has been given to studying it in mechanical systems close to integrable. On the one hand, one hopes that the good understanding we have of integrable systems can be transferred to quasi-integrable systems. On the other hand, since integrable systems do not exhibit any diffusion, quasi-integrable systems—which also present KAM tori and other obstacles to diffusion—are a good test case for the generality of the phenomenon.

In the mathematical literature, there are different precise definitions of diffusion trying to capture the idea of "large" excursions. However, the dominant geometric paradigm for diffusion near integrable systems has been till recently the mechanism proposed by V. I. Arnol'd and documented, for example, in [Arn64, AA67].

This mechanism, henceforth referred as *classical Arnol'd diffusion*, is based on the existence of chains of whiskered tori—remnants of those of the integrable system—and such that the unstable manifold of one intersects the stable manifold of the next one. The classical Arnol'd diffusion has not been shown to be a generic phenomenon. On the other hand, several of its ingredients have been verified for some systems [CG94, CG98, Moe96] and some special families exhibiting it have been constructed [Dou88, DLC83, Gal99, FM01].

In this paper, however, we want to argue that the mechanism described in [Arn64] is not the only mechanism for diffusion in quasi-integrable systems and that there are other geometric mechanisms that can be studied with mathematical rigor. Of course, there are other possible approaches besides geometric perturbation theory, notably, variational methods which have produced significant progress in recent years, but we will not deal with them.

The first goal of this paper is to formulate heuristically two variants of the classical Arnol'd mechanism in which the diffusion takes place through orbits that follow transition chains of whiskered tori as in [Arn64]. Nevertheless, the transistion chains we consider contain objects which are not the KAM tori.

In the first of the mechanisms we propose, the whiskered tori involved in the transition chain are not only those tori that can be studied as perturbations of tori in the original system (called primary tori). In our mechanism, we incorporate tori (called secondary tori) which cannot be continuously deformed into tori invariant under the original system. They are topologically different from the tori present in the unperturbed system (see Definition 2). The generation of secondary tori due to resonances has been discussed in [LW89].

In the second mechanism we present here, some of the whiskered tori which enter in the transition chain are of lower dimensions, e.g, periodic orbits. Indeed, they are tori that survive resonances in which the primary tori are destroyed by a mechanism similar to that considered in [Poi99, Cap V, §81], [Tre89].

As we will see, these new objects dovetail very precisely in the gaps between KAM tori (see Section 8 and Fig. 1). Hence, they can be used as elements of transition chains that overcome the large gap problem.

The second goal (which takes up the bulk of this paper) is to verify rigorously the existence of these mechanisms in rather concrete systems. The verification will be rather explicit, and given a concrete systems there are finite calculations which establish it. In particular, we will study a model that has been studied already in [HM82].

We note that the model we discuss presents the *large gap problem* (namely that the size of the gaps between the KAM tori is larger than the size of first order change in the (un)stable manifolds) that has played an important role in Arnol'd diffusion (an excellent discussion of the problem and, indeed of the problem of diffusion can be found in [Moe96])

Our main rigorous result Theorem 7 establishes that the mechanisms that we present overcome the large gap problem in the model considered. Theorem 7 presents several rather explicit sufficient conditions that ensure that a system which verifies them has orbits that traverse a large gap region.

We believe that the mechanisms proposed here have the advantage that they fit better some of the intuition gathered from numerical and real experiments than the mechanism of [Arn64].

Note that the numerical, experimental and geometric intuition (see e.g. the classical [Chi79] or [CSUZ89, Mei92, JVMU99, Las93, LHRK02] among many others) is that resonances generate diffusion.

On the other hand, the mechanism of [Arn64] has difficulty dealing with resonances, which destroy the primary tori. Indeed, one of the main problems to establish rigorously the existence of the mechanism of [Arn64] is the *large gap problem*, which refers to the fact that the resonances create gaps in the families of primary whiskered tori whose size is bigger than the angle between the whiskers of the primary whiskered tori. In other words, the mechanism of [Arn64] for diffusion is difficult to verify—and can be presumably false—precisely in the places where experimental evidence suggests that diffusion should be most intense.

For the mechanisms that we propose here, we observe that resonances, even if they destroy the primary tori, they create secondary tori and tori of lower dimension that bridge them, so that the transition chains can continue. Indeed, in agreement with the physical intuition, the secondary tori created by resonances, lead to a larger increase in action in their elements of the transition chain. (See [Hal97, Hal99] for a discussion of the role of double resonances in diffusion).

Of course, we are far from believing that the mechanisms we present here are the only mechanisms for diffusion. Indeed, some other geometric mechanisms have been rigorously established for other systems. For example, [DLS00, BT99, DLS01] study geometrically a system that has been studied in [Mat95] by using variational methods. A variational approach to Arnol'd diffusion can be found in [BB02] and announcements of other variational methods are in [Xia98, Mat02]. The papers [Lla02, Moe02, EMR01, Tre02] study other geometric mechanisms. There are heuristic descriptions and numerical explorations of other geometric mechanisms in [LT83, Ten82, CLSV85].

In Section 2, we will describe the proposed mechanism in an informal and conjectural way. Hence, the reader interested only in results that are rigorously proved can skip this section altogether.

In Section 3, we present a class of model systems in which we will verify rigorously the mechanisms described in Section 2. In Section 4, we will state Theorem 7 that verifies the mechanism in the concrete model introduced in Section 3.

The subsequent sections are devoted to the proof of Theorem 7.

After introducing in Section 5 some notation, in Section 6 we describe the geometry of the problem, which has several elements in common with that of [DLS00]. Even if the present paper is logically independent of [DLS00] (we only use a few of the technical results), we have included in Section 12.2 a discussion of the similarities and differences between the phenomena considered and the methods used in this paper and in [DLS00].

In subsequent subsections we will prove Theorem 7, and we will develop tools to establish that, indeed, our system satisfies the proposed mechanism.

We note that the methods we present are rather constructive and are capable of establishing (or at least reduce to a finite calculation) the existence of the mechanisms in a concrete system given by explicit formulas. Indeed, in Section 13, we present some systems where it is possible to compute the formulas needed in closed form so that we can ensure that these systems have trajectories that cross the resonance gaps.

We call attention to the fact that the proof we present is quite modular and consists on well defined steps that depend only on the results of the previous ones. If one is willing to make assumptions that yield the results of one step, or finds a technique that verifies the steps independently, then the rest of the proof can be used.

We also note that most of the tools we need for the verification of the mechanism proposed here are variants of standard techniques in dynamical systems and specially in diffusion. Most of the results we use can be obtained through readily available techniques. Hence, the main novelty of this paper lies in the overall strategy and the dovetailing of different geometric structures (this is why we have decided to include the heuristic description of the mechanism. Undoubtedly, some experts will have little trouble filling the details themselves).

We hope that these improvements (sharpening the analytic tools so that they become reusable as parts of longer arguments) could be useful for the study of other mechanisms for diffusion. We also expect that the ingredients we have studied—mainly normally hyperbolic invariant manifolds, secondary tori—can be rearranged in other ways to provide other mechanisms for diffusion or that they can be integrated with other approaches, notably variational methods. On the other hand, we note that convexity of the problem does not play any role in our methods in contrast with variational method.

One fundamental tool that we adopted from [DLS00] is the use of scattering maps to normally hyperbolic invariant manifolds. The scattering map method allows us to discuss in a geometric way heteroclinic phenomena. In particular, we can discuss comfortably the intersection of invariant manifolds of tori of different topologies (even of different dimensions). This discussion does not seem straightforward in the usual Melnikov theory which often requires to put both objects experiencing the homoclinic phenomena in the same system of coordinates.

We have chosen to verify the results for one parameter families and formulate our results in such a way that these results apply for all $|\varepsilon| < \varepsilon^*$. Of course, the ε^* depends on the family and can tend to zero as the family considered approaches a particularly degenerate one which does not satisfy the assumptions of the theorem. We have also chosen to formulate our results by obtaining computable sufficient conditions on the family for the results to apply.

For one parameter families, it is customary to classify them— following the notation of [CG94]— in *a-priori stable* or *a-priori unstable* systems depending on whether the unperturbed system is strongly integrable (can be put in action-angle variables) or not.

Of course, this is not the only possible point of view and it is interesting to compare this formulation with other formulations of the problem of diffusion. For example, the method in [Arn64] considers families depending on two parameters (the second one exponentially small with respect to the first).

It is also natural and customary to make assertions for generic or for "cusp residual sets"(see[Mat02]). Nevertheless, in this paper we have adopted the concrete verification point of view and relegated a tentative discussion of genericity to remarks.

It is important to realize that, when one considers two parameter families—a fortiori if one consider generic results—the distinction between a-priori stable systems and a-priori unstable systems does not make sense. One can use the first parameter to move the system away from a-priori stable and then use the second parameter— or genericity—applying the methods of a-priori unstable systems. Hence, one can hope that the methods of one parameter families near a-priori unstable systems apply to obtain two parameters or generic results.

The study of one parameter families near a-priori stable systems seems to require considerations of another nature. There are indeed normal forms near resonances similar to the systems we consider here. Nevertheless there appear exponentially small phenomena in the parameter and dealing with them without introducing other exponentially small parameters or generic perturbations requires delicate analytical tools.

2. Heuristic discussion of the mechanism

In this section, we will describe heuristically what are the ideas that lead to the mechanism of diffusion proposed in this paper.

The discussion in this section will be completely non-rigorous. Nevertheless, we hope that it can serve as stimulus for future work.

Of course, the reader interested only in rigorous results can skip to Section 3 and the following after browsing through the definitions introduced in Section 2.1.

2.1. **Integrable systems, resonances, secondary tori.** We start by collecting some standard definitions and setting the notation we will use.

For our purposes, we describe a *strongly integrable system* as a system which has phase space $\mathbb{R}^d \times \mathbb{T}^d$, where $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, and on which the motion is given by:

$$\Phi_t(I, \varphi) = (I, \varphi + \omega(I)t) \tag{1}$$

in the case of flows or by:

$$F(I, \varphi) = (I, \varphi + \omega(I)) \tag{2}$$

in the case of maps.

According to this rather restrictive definition, the mathematical pendulum, a system in $\mathbb{R}^1 \times \mathbb{T}^1$ described by the Hamiltonian

$$P(p, q) = p^2/2 + \cos q - 1,$$

is not a strongly integrable system even if it has a conserved quantity—the Hamiltonian P—and all the motions are quite orderly. Some of the motions consist on librations, some others are rotations and the rotations and librations are separated by orbits that start and end in the critical point $p = 0, q = 0$. Note that these orbits (usually called *separatrices*) separate two different topological types of orbits. Hence there is no hope of writing global action-angle coordinates that straddle them.

Hence, sometimes, one introduces the less stringent definition of *integrable system* in which the representations (1) or (2) hold in open dense sets. These open sets are delimited by special submanifolds called *separatrices*. These separatrices are, at the same time, stable and unstable manifolds of lower dimensional tori invariant for the flow.

Following [CG94], it is customary to refer to systems as (1), (2) as *a priori stable*. Otherwise, systems which present separatrices and hyperbolicity, but which are strongly integrable piecewise in the complement of the separatrices, are called *a priori unstable*.

Remark 1. Of course, there are systems that present separatrices which are not the stable or unstable manifolds of other hyperbolic sets (e.g., take a pendulum with potential $V(q) = (\cos q - 1)^4$ so that the critical points are not hyperbolic points). Such situations seem not to have received much attention in the literature. (Nevertheless, see [BF02].) □

A *quasi-integrable system* is a system which is close to an integrable system in the topology of a space of sufficiently smooth functions. Sometimes, it will be convenient to consider families indexed by a parameter ε so that the system for $\varepsilon = 0$ is integrable. For $|\varepsilon|$ small enough, the system will be quasi-integrable.

We recall that, under appropriate differentiability and non-degeneracy conditions, if the unperturbed system is strongly integrable, the KAM theorem shows that most of the phase space of a quasi-integrable system is covered by invariant tori with a Diophantine frequency $\omega(I)$.

The KAM theorem does not deal with regions on which

$$\omega(I) \cdot k = \mathrm{O}(\varepsilon^{1/2}), \quad k \in \mathbb{Z}^d \setminus \{0\}. \tag{3}$$

The power 1/2 is optimal as can be seen in examples such as $\frac{I^2}{2} + \varepsilon \sin\varphi$, and it was established in [Sva80, Neĭ81, Pös82]. The $\mathrm{O}(\varepsilon^{1/2})$ in the size that needs to be excluded is non-uniform in k and goes to zero when $|k| \to \infty$.

We will call *resonances* the values of I for which $\omega(I) \cdot k = 0$, for some $k \in \mathbb{Z}^d \setminus \{0\}$. We will call *resonant regions* the regions (3) around the resonance that need to be excluded in the proof of KAM theorem.

It is known empirically that very often, in these resonant regions there are dynamical objects which are not present in the integrable system. Typically, the resonant regions contain poorly understood areas called *the chaotic sea* which includes homoclinic intersections and, more important for us, secondary tori. Indeed, secondary tori are the most visible objects in numerical simulations. In two dimensions, when the visualization is very clear, these secondary tori are the tori in the *islands* in the chaotic sea. (See Figure 1.)

More precisely:

Definition 2. *We say that a $(d-k)$-dimensional torus invariant under the flow is a secondary torus when it is retractable to a set $\{I\} \times \{(\varphi_1, \dots, \varphi_l)\} \times \mathbb{T}^{d-l-k}$.*

The fact that resonances generate secondary tori can be established rigorously under suitable non-degeneracy assumptions on the integrable system and the perturbation. We will develop one such proof in Section 8.5.

We also point out that besides the d-dimensional tori, KAM theory can establish the persistence of $d-k$-dimensional tori, $k < d$, when they have k stable and unstable directions. (See [Gra74, Zeh76]. Optimal measure estimates are available from [Val00].) These tori, called whiskered tori, can be primary or secondary.

It is also known that resonances generate lower dimensional whiskered tori. (See, for instance, [Poi99, Cap V, §81], [Tre89, Nie00, DG01] for primary tori and [LW89] for secondary tori.)

In our discussion, there will not be much difference between whiskered tori and fully dimensional tori. We will use the theory of normally hyperbolic invariant manifolds to find invariant submanifolds. The maximal KAM tori in these invariant submanifolds will correspond to whiskered tori for the full system. (This observation was already done in [Moe96, DLS00] and it is exploited in [Sor02]). On the other hand, distinction between primary and secondary tori will be very important for us.

2.2. **Heuristic description of the mechanism.** The systems we consider will be perturbations of integrable ones. To fix ideas, we will consider an unperturbed system which is integrable but not strongly integrable (i.e. it is a priori unstable), and we will perturb it.

Our unperturbed system will admit action-angle variables in open sets divided by *separatrices* i.e. invariant manifolds that end in lower dimensional tori (of codimension one). We will refer to these tori present in the unperturbed system as *primary whiskered tori*.

An important example of these systems, which has been considered many times in the literature and which we will revisit here, is the system consisting of a pendulum and an oscillator uncoupled. Note that this system exhibits a behavior which is very similar to the behavior we will find in a neighborhood of a resonance in a typical system.

Notice that when we consider all the whiskered tori of the integrable system, we obtain a normally hyperbolic invariant manifold (which we call the resonant manifold). In the integrable case, we note that the stable and unstable manifolds of the resonant manifold agree.

Note also that if we consider the motion of the integrable system restricted to the resonant manifold, it is strongly integrable. The whiskered tori of the full system are KAM tori inside the resonant manifold.

The motion restricted to the resonant manifold presents resonances, which in turn correspond to the double resonances of the original system.

As it is well known to specialists in Arnol'd diffusion, adding a perturbation of size ε preserves the resonant manifold and its stable and unstable invariant manifolds. For typical perturbations the foliation by tori in the perturbed manifold persists except for gaps or order $O(\varepsilon^{1/2})$. Moreover, the stable and unstable manifolds move by an amount of order $O(\varepsilon)$. It is rather straightforward to show that when there are tori at a distance $O(\varepsilon)$ one can easily construct heteroclinic intersections between the tori and, hence, construct transition chains along the resonant manifold.

However, near the gaps of order $O(\varepsilon^{1/2})$, these leading order considerations do not allow to conclude existence of transition chains. This is what is called the *large gap problem*.

The main idea of the method proposed here is to study carefully the objects generated by the double resonance. As it is more or less folklore, the resonances destroy the primary whiskered tori present in the original system but create objects of other types. In particular, they create secondary whiskered tori and whiskered tori of lower dimension.

The key point of the proposed mechanism is that the secondary whiskered tori and the stable and unstable manifolds of the lower dimensional tori come very close to the region covered by the primary tori. The secondary tori *bridge* the resonant region. Hence, by incorporating them in the transition chains we can overcome the large gap problem.

The heuristic description of the mechanism proposed here to overcome the large gap problem is:

1) Outside the regions of double resonance, KAM theorem applies, and we obtain the existence of codimension one whiskered tori which are extremely close to each other. Hence one can make transition chains following the classical Arnol'd mechanism, that is, chains formed by heteroclinic connections between the stable and unstable manifolds of primary whiskered tori.
2) Inside the regions of double resonance, and very close to their boundary, we can find KAM secondary whiskered tori.

2') Inside the regions of double resonance, and very close to their boundary, we can find also stable and unstable manifolds of whiskered tori of lower dimension than those of the integrable one (indeed, periodic orbits in our model).

3) The secondary whiskered tori, the lower dimensional whiskered tori and the primary whiskered tori lie on a normally hyperbolic invariant manifold which is a continuation of the resonant manifold for the integrable system.
4) This normally hyperbolic invariant manifold has stable and unstable invariant manifolds which, under some explicit non-degeneracy conditions on the perturbation, intersect transversally.
5) Under some explicit non-degeneracy conditions on the perturbation it is possible to produce transition chains (sequences of invariant whiskered tori in which the unstable manifold of one intersects transversely the stable manifold of the next) involving primary whiskered tori, secondary whiskered tori and lower dimensional whiskered tori.
6) Once we have the transition chains—even if they involve secondary tori or lower dimensional tori—it is possible to use variants of the usual *"obstruction property"* mechanism to show that there are orbits that follow the transition chains.

Perhaps the main conceptual novelty lies in point 2) and 2') above where we identify geometric structures in the double resonances which allow to bridge them.

Note that our mechanisms depend on establishing the existence of transverse intersections of stable and unstable manifolds for objects of different topological type. In such a case, one cannot easily use the methods of standard Melnikov theory which rely on using a similar coordinate representation for both of them. A technical tool which could be of independent interest is the scattering map, defined on the normally hyperbolic invariant manifold thanks to the fact that its stable and unstable invariant manifolds intersect transversally, and which allows us to discuss the heteroclinic intersection of invariant objects of different type.

Remark 3. We will see that the resonant regions, even if they have a size $O(\varepsilon^{1/2})$, are connected by only one element in the transition chain, whereas outside of the resonant regions, one step of the transition chain only makes a much smaller step in space.

If one makes the conjecture (see [Chi79], amplified in [CV89]) that the time spent in a transition chain is roughly proportional to the number of tori in the transition chain—there could be a factor $|\log \varepsilon|$—one obtains that the resonant region will be crossed over faster than other regions in the transition chain. Hence, the diffusion will be a slow drift punctuated by much larger jumps when the system approaches a resonance. This fits well with the descriptions in the empirical literature, in particular, the literature on Levy flights [SZF95, Zas02] or the experimental observations that confirm that diffusion is much more apparent near resonances [Las93, LHRK02], . Of course, the current state of the art is far from a proof of this conjecture, In particular, there do not seem to exist techniques to produce a rigorous an statistical theory. Nevertheless we find encouraging that our methods of proof move in the direction suggested by intuition and experimental observations and show some semblance of agreement with them. □

Remark 4. In many physical applications, one has to consider systems which consist of many identical systems with a local coupling. In [HL00] there is an empirical and heuristic study of the abundance of secondary tori in these coupled oscillators. The conclusion of numerical experiments and heuristic arguments of [HL00] is that secondary tori are much more abundant than KAM primary tori in systems of coupled oscillators. □

From the mathematical point of view, in this paper we will content ourselves with studying just one model chosen because the verification of the mechanism is the simplest possible among those for which the verification is non-trivial and this model was introduced in the paper [HM82], which ignored the large gap problem.

Most of the methods that we employ in our verification have been standard tools of the geometric approach to standard Arnol'd diffusion. Nevertheless, we call attention to the fact that we also make good use of the perturbation theory of normally hyperbolic invariant manifolds to provide an skeleton along which diffusion takes place. (This is a common feature with [DLS00, DLS01], which consider yet another mechanism of diffusion and with [Moe96].)

Indeed, the main object that organizes the mechanism here is a normally hyperbolic invariant manifold $\tilde{\Lambda}_\varepsilon$ close to the resonances. The stable and unstable manifolds of this manifold will intersect transversally. This will allow us to define two dynamics on the manifold $\tilde{\Lambda}_\varepsilon$. One is the dynamics restricted to the manifold, which we called the *inner map* in [DLS00]. Following again [DLS00], we will also define a *scattering map* in Section 9. Given an orbit $\tilde{z}(t)$ that performs a homoclinic excursion to $\tilde{\Lambda}_\varepsilon$, we can find two orbits in $\tilde{\Lambda}\varepsilon$ that approach $\tilde{z}(t)$ in the future and in the past. The scattering map associates the orbit asymptotic in the past to the orbit asymptotic in the future. It will be shown that, when one considers all the orbits $\tilde{z}(t)$ chosen in a homoclinic connection, we obtain a scattering map which is smooth.

One can obtain diffusing pseudo-orbits by applying alternatively the inner map and the scattering map. This corresponds to orbits that perform a homoclinic excursion when the perturbation in favorable but that, otherwise stay "parked" near $\tilde{\Lambda}_\varepsilon$. If the motion generated by the inner map contains invariant objects such as tori, it will be possible to use obstruction properties (see Section 11) to construct orbits that diffuse.

We note that there is a large literature on obstruction properties. We will verify that some methods in the literature apply even if the tori have different topology (of course, there are other methods which do not seem easy to adapt to our problem). We will follow the exposition in [DLS00], which uses mainly the Lambda Lemma result in [FM00], which makes it clear that neither the topology of the embedding of the torus or its dimension does not make any difference.

Of course, it is very likely that there are other methods to verify the mechanism presented here and indeed other mechanisms for diffusion. Notably, we expect that many of the results established here could be established for concrete systems using variational or more topological methods or proved for generic systems by other methods.

3. A simple model

In the rest of the paper we will verify rigorously the mechanism presented above in a specific model.

The model has been chosen in such a way that the verification of the proposed mechanism is as simple as possible among the nontrivial ones.

The model we consider has been studied in several other places, but to the best of our knowledge, the mechanism described in this paper has not been considered. In particular, this model was considered in the paper [HM82], but the large gap problem was not addressed in that paper.

We also note that the model we consider is a standard model of the behavior near resonances.

Even if the analysis in this paper will require several unduly restrictive assumptions, we hope that it could lay the foundations for further progress.

We will consider a mechanical system described by the non-autonomous Hamiltonian, periodic in time

$$\begin{aligned} H_\varepsilon(p,q,I,\varphi,t) &= H_0(p,q,I) + \varepsilon h(p,q,I,\varphi,t;\varepsilon) \\ &= P_\pm(p,q) + \frac{1}{2}I^2 + \varepsilon h(p,q,I,\varphi,t;\varepsilon) \end{aligned} \tag{4}$$

where we denote by

$$P_\pm(p, q) = \pm(\frac{1}{2}p^2 + V(q)) \tag{5}$$

and where $V(q)$ is a 2π-periodic function.

We will refer to $P_\pm(p, q)$ as the *pendulum.* (It is a physical pendulum when we take P_+, and $V(q) = \cos q - 1$, which is a good example to keep in mind.)

The term $\frac{1}{2}I^2$ of (4) describes a rotator whose frequency changes with the energy of the oscillation (equivalently with the action).

The final term h in (4) describes a small coupling between the rotator and the pendulum that depends periodically on time. We will assume—besides differentiability properties—that V reaches a maximum at one point—we will assume without loss of generality that it is 0—and that the maximum is non-degenerate (i.e. $V''(0) < 0$). We will assume for the sake of simplicity of exposition that 0 is the only local maximum. This later assumption can be eliminated at the only expense of complicating the notation.

Remark 5. The choice of sign in $P_\pm$ will not make any difference in our arguments which are based on hyperbolicity and KAM theory. Note however that the Hamiltonian with P_+ is convex for large I, p but with P_- is neither convex nor positive definite.

The assumption of positive definiteness seems to be very important for variational methods [Xia98, Mat02]. On the other hand, we note that the physical intuition [Chi79, TLL80] is that systems with subsystems of different signatures tend to be more unstable. □

It will also be convenient to consider the system (4) as described by an autonomous Hamiltonian flow on $(\mathbb{R} \times \mathbb{T})^3$—which we will call the symplectic extended phase space—endowed with the standard symplectic structure. The autonomous Hamiltonian will be:

$$\begin{aligned} \tilde{H}_\varepsilon(p, q, I, \varphi, A, s) &= A + H_\varepsilon(p, q, I, \varphi, s) \\ &= A + H_0(p, q, I) + \varepsilon h(p, q, I, \varphi, s; \varepsilon) \\ &= A + P_\pm(p, q) + \frac{1}{2}I^2 + \varepsilon h(p, q, I, \varphi, s; \varepsilon). \end{aligned} \tag{6}$$

We will follow the standard convention of assuming that the pairs $(p, q) \in \mathbb{R}\times\mathbb{T}$, $(I, \varphi) \in \mathbb{R}\times\mathbb{T}$, and $(A, s) \in \mathbb{R}\times\mathbb{T}$, will be symplectically conjugate variables.

Note that h does not involve A so that the equations of motion of the pair (A, s) are just

$$\dot{A} = -\partial_s \tilde{H}_\varepsilon(p, q, I, \varphi, s) = -\varepsilon \frac{\partial h}{\partial s}(p, q, I, \varphi, s; \varepsilon), \qquad \dot{s} = 1.$$

The introduction of the extra variables (A, s) is a standard device to formulate periodic in time perturbations as an autonomous system. The extra variable s makes the system autonomous and the variable A is symplectically conjugate to s to be able to treat the resulting system as a Hamiltonian one. So, even if the system described by (6) is, strictly speaking, a three degrees of freedom system, we refer to it as a two and a half degrees of freedom system.

Moreover, the variable A does not play any dynamical role. Note that A does not appear in any of the differential equations for any of the coordinates, including itself. Then, one can study the dynamics of the variables (p, q, I, φ, s), and then $\dot{A} = -\partial_s \tilde{H}_\varepsilon(p, q, I, \varphi, s)$ is just a quadrature. Hence, we will consider the equations:

$$
\begin{aligned}
\dot{p} &= \mp V'(q) && -\varepsilon \frac{\partial h}{\partial q}(p, q, I, \varphi, s; \varepsilon) \\
\dot{q} &= \pm p && +\varepsilon \frac{\partial h}{\partial p}(p, q, I, \varphi, s; \varepsilon) \\
\dot{I} &= && -\varepsilon \frac{\partial h}{\partial \varphi}(p, q, I, \varphi, s; \varepsilon) \\
\dot{\varphi} &= I && +\varepsilon \frac{\partial h}{\partial I}(p, q, I, \varphi, s; \varepsilon) \\
\dot{s} &= 1. &&
\end{aligned}
\tag{7}
$$

We will denote by $\mathcal{H}_\varepsilon$ the vector field (7) generated by the Hamiltonian of (6). We will denote by $\Phi_{t,\varepsilon}(\tilde{x})$ the flow generated by the vector field $\mathcal{H}_\varepsilon$ in the extended phase space $(\mathbb{R} \times \mathbb{T})^2 \times \mathbb{T}$.

Remark 6. Notice that introducing the extra variable A makes a certain difference in the geometric nature of the objects. For example, if we find a KAM torus in the non-autonomous system, it will become a family of tori—indexed by the variable A—in the autonomous system. Hence, even if the KAM tori in the non-autonomous system are, for typical perturbations, a Cantor set of tori, in the autonomous version, they always include one parameter families. □

The Hamiltonian $H_0(p, q, I)$ in (4) will be referred to as the unperturbed Hamiltonian. It is customary to describe H_0 as integrable. As we pointed out in Section 2.1, the Hamiltonian system of Hamiltonian H_0 is a priori unstable. Indeed, the system described by $P_\pm$ presents different topological types of oscillations which are separated by some special orbits (called separatrices) ending with zero velocity on the maximum of V.

4. Statement of rigorous results

In this section, we will formulate the main Theorem of this paper.

Actually, we prove somewhat more general results which we formulate later when we have introduced more notation. For example besides orbits which transverse the gaps, it follows from our proof the existence of a symbolic dynamics. We also use slightly weaker hypotheses that take advantage of several homoclinic intersections.

We will consider a neighborhood $\mathcal{S} \subset \mathbb{R} \times \mathbb{T}$ of the separatrix of the pendulum we are studying, and consider the compact set

$$\mathcal{D} := \mathcal{S} \times [I_-, I_+] \times \{\varphi \in \mathbb{T}\} \times \{s \in \mathbb{T}\} \times [-\varepsilon_0, \varepsilon_0] \tag{8}$$

to be the domain of definition of our problem. Hence, the C^r norms of functions will refer to sup norms defined on this set. Of course, in the case that the functions depend only on a few of the variables (e.g. V, which only depends on q), we can also consider them as function of more variables and define the norm in the appropriate domain.

The main rigorous result of this paper is:

Theorem 7. *Consider a system of the form* (6)

Assume:

- **H1**. *The terms V and h in* (6) *are uniformly C^r for $r \geq r_0$, sufficiently large.*
- **H2**. *The potential $V : \mathbb{T} \to \mathbb{R}$ has a unique global maximum at $q = 0$ which is non-degenerate (i.e., $V''(0) < 0$). We denote by $(q_0(t), p_0(t))$ an orbit of the pendulum $P_\pm(p, q)$ in* (5)*, homoclinic to* $(0, 0)$.
- **H3**. *h is a trigonometric polynomial in φ and s:*

$$h(p, q, I, \varphi, s; \varepsilon) = \sum_{k,l \in \mathcal{N}} \hat{h}_{k,l}(p, q, I; \varepsilon) e^{i(k\varphi + ls)}, \tag{9}$$

where $\mathcal{N} \subset \mathbb{Z}^2$ is a finite set.

- **H4**. *Consider the Poincaré function, also called Melnikov potential, associated to h (and to the homoclinic orbit (q_0, p_0) mentioned in* **H2***):*

$$\begin{aligned} \mathcal{L}(I, \varphi, s) = -\int_{-\infty}^{+\infty} \Big(& h\left(p_0(\sigma), q_0(\sigma), I, \varphi + I\sigma, s + \sigma; 0\right) \\ & - h(0, 0, I, \varphi + I\sigma, s + \sigma; 0) \Big) \, d\sigma. \end{aligned} \tag{10}$$

Assume that, for any value of $I \in (I_-, I_+)$ there exists an open set $\mathcal{J}_I \subset \mathbb{T}^2$, with the property that when $(I, \varphi, s) \in H_-$, *where*

$$H_- = \bigcup_{I\in(I_-,I_+)} \{I\} \times \mathcal{J}_I \subset (I_-, I_+) \times \mathbb{T}^2, \tag{11}$$

the map

$$\tau \in \mathbb{R} \mapsto \mathcal{L}(I, \varphi - I\tau, s - \tau) \equiv \Gamma(\tau; I, \varphi, s) \tag{12}$$

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.

Assume moreover that for every $(I, \varphi, s) \in H_-$, the function

$$\frac{\partial \mathcal{L}}{\partial \varphi}(I, \varphi - I\tau^*, s - \tau^*) \tag{13}$$

is non-constant and negative (respectively positive).

H5. *The perturbation terms $h(p, q, I, \varphi, s; 0)$, $\frac{\partial h}{\partial \varepsilon}(p, q, I, \varphi, s; 0)$ satisfy some non-degeneracy conditions(that will be stated explicitly in Section* 8.5.2 *and Section* 10.2 *as* **H5', H5"** *and* **H5"'** *respectively).*

Then, there is $\varepsilon^* > 0$ such that for $0 < |\varepsilon| < \varepsilon_*$ and for any interval $[I_-^*, I_+^*] \subset (I_-, I_+)$, there exists a trajectory $\tilde{x}(t)$ of the system (6) *such that for some $T > 0$*

$$I(\tilde{x}(0)) \le I_-^*; \quad I(\tilde{x}(T)) \ge I_+^* \tag{14}$$

(respectively:

$$I(\tilde{x}(0)) \ge I_+^*; \quad I(\tilde{x}(T)) \le I_-^*). \tag{15}$$

We will consider $I_- < I_+$ as fixed and somewhat large. In particular, $[I_-, I_+]$ can contain all the resonances $I = -l/k$, for $(k, l) \in \mathcal{N}$. Then, the trajectories that we construct cross over the resonant regions.

Hence, we overcome the *large gap problem* by showing the existence of orbits which traverse regions in which primary KAM tori are not present and indeed, there are no transverse heteroclinic intersections between them accessible to direct perturbation theory.

Remark 8. Note that given the fact that we can find an open set H_- as in (11), it will often happen that we can find two different open sets $H_-^\pm$ where the sign in (13) take place.

In a generic situation, the non-degeneracy condition (11) will only fail in a codimension 1 manifold. Moreover, generically, this critical manifold will not be a level set of the action.

Hence, for generic situations we expect that we will be able to find overlapping intervals $I^{\pm}$ such that $I^{\pm} \times \mathbb{T}^2$ intersects H_- and such that in $(I^{\pm} \times \mathbb{T}^2) \cap H_-$ we obtain the different alternatives in (13). Again, we emphasize that verifying that this generic situation happens in a concrete system, is reduced to a finite calculation.

As we will see in Theorem 93 this will allow us to construct very complicated orbits. □

Remark 9. One can obtain more general results in the case that the function

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

has a zero in $\mathcal{J}_I$, because in this case we can obtain subsets where it is positive and negative. This allows the orbit $\tilde{x}(t)$ to perform more arbitrary excursions. This happens, for example, if we change hypothesis **H4** to assuming that

$$(I_-, I_+) \times \mathbb{T}^2 \subset H_-. \tag{16}$$

In this case, we obtain that $\mathcal{L}(I, \varphi - I\tau^*, s - \tau^*)$ is a periodic function of φ and s and the fact that its derivative has a zero is guaranteed. □

In this paper, we will not address rigorously the issue of how abundant is the mechanism presented here. Nevertheless, in Section 12.3, we will present some heuristic remarks on how the abundance of the hypotheses.

Remark 10. Several of the following arguments are significantly easier if we assume (16). The reader may want to check this case first to gain intuition. For simplicity of heuristic intuition, the pictures in Figures 1, 2 we present to illustrate the results are done under hypothesis (16). Note that if there is a point (I_0, φ_0, s_0) for which the function $\Gamma(\tau; I_0, \varphi_0, s_0)$ has a non-degenerate critical point τ^*, by the implicit function theorem the same will be true for all the (I, φ, s) in an open neighborhood.

One, of course, expects that this open set has a boundary where the critical point becomes degenerate. So, the hypothesis (16) holds in open sets of models and it is verified in several cases of interest, which we will consider in detail in Section 13.

Even if (16) is not generic for one critical point (it is, however, open), as we will see in Section 12.3 when we consider several homoclinic and several critical points it is generic. □

Remark 11. In general, the pendulum $P_\pm$ will have two homoclinic orbits. To each of these homoclinic orbits we can associate a Poincaré function via (10).

For generic perturbations, each of the different Poincaré functions will have several non-degenerate critical points.

It can therefore happen that one can find several intervals $I^k \equiv [I_-^{*,k}, I_+^{*,k}]$ on which one can verify the hypotheses of Theorem 7.

We will see in Theorem 93 that if there is a closed interval $I^* = [I_-^*, I_+^*]$ contained in the union of the interior of the intervals I_k, then, we also obtain the conclusions of Theorem 7 for the interval I^*.

As we will see in more detail Section 12.3 this makes it much easier to verify the existence of the mechanism in concrete models. The reason is that if we consider just one smooth curve τ^* it often happens that this manifold has codimension one boundary in which the critical point becomes degenerate.

If the boundaries for different τ^*'s do not coincide, the diffusion can proceed. Moreover, we will see that if the projections in I of the region where the positive and negative alternatives of (13) hold, it is possible to find orbits whose I performs largely arbitrary excursions. $\square$

Remark 12. For $\varepsilon = 0$, the set $p = 0, q = 0$ is a normally hyperbolic invariant manifold whose stable and unstable manifolds coincide. For $0 < |\varepsilon| < 1$, this manifold and its stable and unstable manifolds are preserved due to the theory of normally hyperbolic invariant manifolds.

Hypothesis **H4** means that the first order in ε perturbation theory predicts the transversal intersection of the stable and unstable manifolds.

We have taken advantage of the symplectic geometry to express the sufficient condition of intersection in terms of only one function, called here Poincaré function or Melnikov potential. This is very natural geometrically since the stable and unstable manifolds of whiskered tori are Lagrangian. See [DG00, DG01].

As it turns out, the same Melnikov potential will appear in the conditions **H5**, a connection which is not apparent if one carries out the perturbation theory without using the symplectic structure. $\square$

Remark 13. We also note that in [CG94] it is shown that in the model (6), under the hypothesis that the perturbation contains a finite number of harmonics—which we also include at the moment—one can obtain diffusion in the I direction which is of a size independent of ε. The reason is that, outside the resonances $I = -l/k$, for

$(k,l) \in \mathcal{N}$, $k \neq 0$, which are a finite set, they show that the primary KAM tori are very close. (This is called in [CG94] the *gap bridging* mechanism.) Note that what is shown in [CG94] is not that there are orbits that transverse the resonant gaps, but rather that, in the regions that contain no resonances there are no large gaps between primary KAM tori. The argument of [CG94] is completely different from the mechanism considered here. In our case, the gaps indeed are large, but the mechanism introduced here overcomes them by using secondary tori and lower dimensional tori. □

Remark 14. In the paper [Gal99] one can find a treatment of diffusion along the action variables A. Since the KAM tori contain one parameter families in the variable A (see Remark 6), there are no gaps among them and it is possible to construct transition chains along them.

The phenomena discussed in this paper are very different from the phenomena established in [Gal99] since the diffusion we establish is present in the non-autonomous version of the problem which ignores the variable A. The non-autonomous system does present gaps among the KAM tori and, indeed, our transition chains have to include objects other than KAM tori. □

Remark 15. A value of r_0 which follows from our argument is $r_0 = 60$. Of course, this is not optimal even for the argument presented, and better regularities can presumably be obtained using different methods in part or in all of our argument. Clearly, topological, variational or generic results will use significantly lower regularity.

Note that the final result of the argument presented here, when we apply the obstruction criterion is a sequence of well aligned boxes. This is stable under C^0 small perturbations of the time one map. Hence, even if we use high smoothness to construct objects that provide us with an scaffolding for the diffusion, the final result is robust even under very rough perturbations which can destroy the scaffolding. □

Remark 16. We think that Hypothesis **H3** can presumably be eliminated in Theorem 7 at the price of complicating some of the estimates presented here.

We will present some heuristic considerations in Section 12.4 but we will postpone such considerations. □

Remark 17. It seems not difficult to consider situations in which (I, φ) are higher dimensional or in which s ranges in a torus or in which the oscillator is substituted by a rotator (an anisochronous system in the notation of [Gal94]). Considering (p, q) of higher dimension seems to be more difficult. □

Remark 18. Even if this paper is logically independent from [DLS00], the analysis here is inspired by that of [DLS00] and we will look for objects quite analogous to those considered in that paper. Indeed, we will keep many of the notations introduced there since this will make it easier to use some of the results of that paper in the argument here. Nevertheless, there will be quite significant differences between the way that they are organized. Roughly speaking, the fixed point of the pendulum in (6) will play the role that the periodic geodesic played in [DLS00]. One of the separatrices of the pendulum will play the role of the connecting geodesic in [DLS00].

In Section 12.2 we have undertaken a more systematic comparison between the objects and the methods of [DLS00] and this paper. □

In the subsequent sections we will give a detailed proof of Theorem 7. In fact we will establish this Theorem by checking that system (6) satisfies the mechanisms explained in Section 2.2.

5. Notation and definitions, resonances

We will say that a function satisfies $F = \mathrm{O}_{C^r}(\eta)$ when

$$||F||_{C^r} \leq \text{cte.}\, \eta.$$

Given a function depending on the ε variable—and others—in a sufficiently smooth fashion, we will introduce the following notation to denote the partial Taylor expansion in ε

$$F(x;\varepsilon) = F_0(x) + \varepsilon F_1(x) + \varepsilon^2 F_2(x) + \cdots + \varepsilon^n F_n(x) + \mathrm{O}(\varepsilon^{n+1}). \tag{17}$$

Note that if F is C^r in all its variables, including ε, the remainder will be $\mathrm{O}_{C^{r-n-1}}(\varepsilon^{n+1})$ considered as a function of all the variables. This counting of derivatives is somewhat wasteful, but avoiding it will require to keep track of separate regularities.

Given a trigonometric polynomial in the angle variables (φ, s)

$$F(p, q, I, \varphi, s) = \sum_{(k,l)\in\mathcal{N}} \hat{F}_{k,l}(p, q, I) e^{i(k\varphi + ls)}, \tag{18}$$

we denote by $\mathcal{N} = \mathcal{N}(F) \subset \mathbb{Z}^2$, the support of the Fourier transform.

It is obvious that:

$$
\begin{aligned}
&\mathcal{N}(h+g) \subset \mathcal{N}(h) \cup \mathcal{N}(g)\\
&\mathcal{N}(hg) \subset \mathcal{N}(h) + \mathcal{N}(g)\\
&\mathcal{N}(\nabla h) \subset \mathcal{N}(h)\\
&\mathcal{N}\left(\int h\, dI\right) \subset \mathcal{N}(h)
\end{aligned}
\tag{19}
$$

where by the addition of subsets in $\mathbb{Z}^2$ we denote the set consisting of sums of an element of the first set and an element of the second.

For a function $F(p, q, I, \varphi, s; \varepsilon)$ of the form (18), which moreover depends on a parameter ε, the set

$$
\mathcal{R}_1 = \mathcal{R}_1(F) := \{I = -l/k : (k, l) \in \mathcal{N}(F_1), k \neq 0\}
\tag{20}
$$

will be called the set of *primary resonances.* As it will be developed in more detail later (and is well known to experts) this is the region to avoid in first order averaging. (See Section 8.3 for a description of the averaging method and a clarification of why the set $\mathcal{R}_1$ plays a role.)

The region

$$
\begin{aligned}
\mathcal{R}_2 &= \mathcal{R}_2(F)\\
&:= \{I = -l/k : (k, l) \in (\mathcal{N}(F_1) + \mathcal{N}(F_1)) \cup \mathcal{N}(F_2), k \neq 0\}
\end{aligned}
\tag{21}
$$

will be called the set of *secondary resonances.* (Roughly speaking, $\mathcal{R}_2$ is the set where it is impossible to apply second order averaging.) These definitions can be generalized to the sets $\mathcal{R}_k$, of the resonances of order k which are the places where the k^{th} order averaging cannot be applied (see Section 8.3). The main resonant set for our model (6) will be

$$
\mathcal{R}_1(h) \cup \mathcal{R}_2(h).
\tag{22}
$$

As it turns out, in our problem, resonances of order higher than the second will play a very small role. When $h(p, q, I, \varphi, s; \varepsilon)$ is a trigonometric polynomial in (φ, s), $\mathcal{R}_1(h) \cup \mathcal{R}_2(h)$ is a finite set.

The main step in our proof will be to jump over regions close to the main resonance set (22). As we will show later in detail in Section 8.3, outside of the main resonant region, the KAM tori are close enough so there are no large gaps between them and the mechanism in [Arn64] applies.

6. Geometric features of the unperturbed problem

In this section we will discuss the geometric features of the unperturbed system which will survive under the perturbation. They will serve as landmarks to organize the motion of the perturbed system.

The main feature of Hamiltonian (6) for $\varepsilon = 0$ is that it consists of two uncoupled systems (a rotator and a pendulum) so that the cartesian product of invariant objects of each of the subsystems will give an invariant object of the full system.

To simplify the notation and without loss of generality, we have chosen coordinates in **H2** in such a way that the maximum of V happens at $q = 0$. Then, for the pendulum described by the Hamiltonian $P_\pm(p,q)$, the point $p = 0$, $q = 0$ is a fixed point and is hyperbolic as an orbit in $\mathbb{R} \times \mathbb{T}$, the phase space of the pendulum.

The stable and unstable manifolds of this point coincide. We will denote by

$$\gamma(t) := (p_0(t), q_0(t)) \tag{23}$$

one of the orbits of $P_\pm$ which is homoclinic to the hyperbolic point $(0,0)$.

Remark 19. As it is well known in mechanics for the case that $P_\pm(p,q) = \pm(p^2/2 + \cos q - 1)$ is the physical pendulum, the parameterizations of the separatrices have explicit formulas given by

$$q_0(t) = 4 \arctan e^{\pm t}, \quad p_0(t) = 2/\cosh t. \tag{24}$$

and

$$q_0(t) = 4 \arctan e^{\mp t}, \quad p_0(t) = -2/\cosh t. \tag{25}$$

In Section 13, we will present many other explicit calculations for the case of the pendulum.

□

When we consider the full system (7) for $\varepsilon = 0$, we see that, for any value of $I \in \mathbb{R}$, the product of the hyperbolic point of the pendulum with all the other variables will be a two dimensional invariant torus in the extended phase space

$$\mathcal{T}_I = \{\tilde{x} = (p, q, I, \varphi, s) : p = q = 0,\ (\varphi, s) \in \mathbb{T}^2\},$$

and one component of the stable and unstable manifolds of this torus, $W^{\mathrm{s}}_{\mathcal{T}_I}$, $W^{\mathrm{u}}_{\mathcal{T}_I}$, coincides along the three-dimensional manifold given by:

$$\{(p_0(\tau), q_0(\tau), I, \varphi, s) : \tau \in \mathbb{R},\ (\varphi, s) \in \mathbb{T}^2\},$$

where $(p_0(\tau), q_0(\tau))$ is given in (23).

Hence, for the values I_-, I_+, given in Theorem 7, we introduce the set

$$\begin{aligned} \tilde{\Lambda} &= \{\tilde{x} \in (\mathbb{R}\times\mathbb{T})^2 \times \mathbb{T} : p = q = 0, I \in [I_-, I_+]\} \\ &= \bigcup_{I\in[I_-,I_+]} \mathcal{T}_I \end{aligned} \tag{26}$$

The set $\tilde{\Lambda}$ is a 3-dimensional manifold with boundary. It is locally invariant and normally hyperbolic for the flow (7) for $\varepsilon = 0$. The manifold $\tilde{\Lambda}$ is topologically $[I_-, I_+] \times \mathbb{T}^2$. It has a global system of coordinates given by (I, φ, s). Hence, we will not need to distinguish very explicitly whether we are talking about the three numbers (I, φ, s) or the point in the manifold $\tilde{\Lambda}$. Nevertheless, we will emphasize as much as possible the geometric meaning of the calculations that need to be carried out in coordinates.

The unperturbed flow on $\tilde{\Lambda}$ is given by the quasi-periodic flow:

$$\Phi_{t,0}(0, 0, I_0, \varphi_0, s_0) = (0, 0, I_0, \varphi_0 + I_0 t, s_0 + t), \tag{27}$$

and we will be denote by

$$\tilde{\lambda}_t(I_0, \varphi_0, s_0) = (0, 0, I_0, \varphi_0 + I_0 t, s_0 + t).$$

the orbits in the torus $\mathcal{T}_{I_0}$.

One component of the stable and unstable invariant manifolds for $\tilde{\Lambda}$, $W^{\rm s}_{\tilde{\Lambda}}$, $W^{\rm u}_{\tilde{\Lambda}}$ coincides in a manifold $\tilde{\gamma}$ of orbits homoclinic to $\tilde{\Lambda}$

$$\tilde{\gamma} \subset (W^{\rm s}_{\tilde{\Lambda}} \setminus \tilde{\Lambda}) \cap (W^{\rm u}_{\tilde{\Lambda}} \setminus \tilde{\Lambda}).$$

and a parameterization of $\tilde{\gamma}$ is given by:

$$\tilde{\gamma} := \{(p_0(\tau), q_0(\tau), I, \varphi, s) \ : \ I \in [I_-, I_+],\ \tau \in \mathbb{R},\ (\varphi, s) \in \mathbb{T}^2\}, \tag{28}$$

where (p_0, q_0) is, as in (23), the chosen homoclinic orbit of the pendulum.

Hence, the meaning of the coordinate τ is the time of the flow along the unperturbed separatrix. We denote by

$$\begin{aligned} \tilde{\gamma}_t(\tau, I_0, \varphi_0, s_0) &= (p_0(\tau + t), q_0(\tau + t), I_0, \varphi_0 + I_0 t, s_0 + t) \\ &= \Phi_{t,0}(p_0(\tau), q_0(\tau), I_0, \varphi_0, s_0), \end{aligned}$$

the unperturbed flow on $\tilde{\gamma}$, our chosen component of the homoclinic manifold.

We note that, for any $\tau \in \mathbb{R}$

$$\operatorname{dist}(\tilde{\lambda}_t(I_0, \varphi_0, s_0), \tilde{\gamma}_t(\tau, I_0, \varphi_0, s_0)) \to 0 \quad \text{for} \quad t \to \pm\infty. \tag{29}$$

Let us note that the unperturbed system is a product, and that $(0, 0)$ is a critical point of the pendulum with characteristic exponents $\mu_\pm = \pm\mu$, where $\mu := (-V''(0))^{1/2}$. Moreover, the exponents of contraction in the tangent direction of $\tilde{\Lambda}$ are 0 (see (27)). Then, the stable and unstable manifolds of $\tilde{\Lambda}$ are characterized as the set of orbits whose distance to the orbits in $\tilde{\Lambda}$ is less than $C \exp(-\mu\,|t|)$ respectively as $t \to \pm\infty$, and we have

$$\operatorname{dist}(\tilde{\lambda}_t(I_0, \varphi_0, s_0), \tilde{\gamma}_t(\tau, I_0, \varphi_0, s_0)) \le C(\tau) \exp(-\mu\,|t|) \quad \text{for} \quad t \to \pm\infty.$$

7. Persistence of the normally hyperbolic invariant manifold and its stable and unstable manifolds

Since the manifold $\tilde{\Lambda}$ is normally hyperbolic, and is locally invariant for the flow (7) for $\varepsilon = 0$, by the theory of normally hyperbolic manifolds we have that the manifold persists under small perturbations. Moreover, if the system depends smoothly on parameters, the manifolds—they may be non unique—may be chosen to depend smoothly on parameters.

A formulation of the results of [Fen72, Fen74, Fen77] in the way that we will use them (very similar to the statement of Theorem 4.2 of [DLS00]) is:

Theorem 20. *Consider a Hamiltonian as in* (6). *Assume that H_ε is uniformly C^r, $r \ge 2$ in all its variables, including ε, in a neighborhood of $\tilde{\Lambda}$ and $\tilde{\gamma}$.*

Then, there exists $\varepsilon^* > 0$ such that for $|\varepsilon| < \varepsilon^*$, there is a C^{r-1} function

$$\tilde{\mathcal{F}} : \tilde{\Lambda} \times (-\varepsilon^*, \varepsilon^*) \longrightarrow (\mathbb{R} \times \mathbb{T})^2 \times \mathbb{T}$$

such that

$$\tilde{\Lambda}_\varepsilon := \tilde{\mathcal{F}}(\tilde{\Lambda} \times \{\varepsilon\}) \tag{30}$$

is locally invariant for the flow (7) *generated by the vector field $\mathcal{H}_\varepsilon$.*

In particular, $\tilde{\Lambda}_\varepsilon$ is ε-close to $\tilde{\Lambda}_0 = \tilde{\Lambda}$ in the C^{r-2} sense.

Moreover, $\tilde{\Lambda}_\varepsilon$ and the vector field $\mathcal{H}_\varepsilon$ can be extended so that it is a normally hyperbolic invariant manifold for the flow (7) *generated by $\mathcal{H}_\varepsilon$. In particular, it is possible to define local stable and unstable manifolds. That is, we can find a C^{r-1} function $\tilde{\mathcal{F}}^{\mathrm{s}}$ such that the (local) stable manifold of $\tilde{\Lambda}_\varepsilon$ takes the form*

$$W^{\mathrm{s,loc}}_{\tilde{\Lambda}_\varepsilon} = \tilde{\mathcal{F}}^{\mathrm{s}} \left(\tilde{\Lambda} \times (0, +\infty) \times \{\varepsilon\} \right). \tag{31}$$

If $\tilde{x} = \tilde{\mathcal{F}}(I, \varphi, s; \varepsilon) \in \tilde{\Lambda}_\varepsilon$, then $W^{\mathrm{s,loc}}_{\tilde{x}} = \tilde{\mathcal{F}}^{\mathrm{s}}(\{I\} \times \{\varphi\} \times \{s\} \times (0, \infty) \times \{\varepsilon\})$.

In particular, $W^{\mathrm{s,loc}}_{\tilde{\Lambda}_\varepsilon}$ is ε-close to $W^{\mathrm{s,loc}}_{\tilde{\Lambda}}$ in the C^{r-2} sense.

Analogous results hold for the (local) unstable manifold $W^{\mathrm{u,loc}}_{\tilde{\Lambda}_\varepsilon}$.

The proof of Theorem 20 is quite standard in the theory of normally hyperbolic invariant manifolds. It is a straightforward application of the theorems in [Fen74, Fen77, HPS77] (see also Appendix A of [DLS00]). A modern proof is in [Lla00].

A useful observation is that there is an easy way to obtain smooth dependence on parameters from the standard results on persistence. It suffices to consider the system obtained by taking the product of the original system and the identity in the direction of parameters. The fact that the invariant manifolds for the extended system are regular, gives the fact that the manifolds of the original system depend regularly on parameters.

Note that the coordinates along the unperturbed manifold $\tilde{\Lambda}$ under the unperturbed evolution just rotate or remain invariant. Since $||D\Phi_{t,0}|_{\tilde{\Lambda}}|| \leq C\,|t|$ (see (27)), we have that for every $\delta > 0$,

$$||\Phi_{t,0}|_{\tilde{\Lambda}}|| \leq C_\delta e^{\delta|t|}.$$

This shows that the tangential exponents along the manifold can be taken as small as we want.

On the other hand, because of assumption **H2**, the point $(p, q) = (0, 0)$ is a hyperbolic point for the pendulum. Since the system (7) for $\varepsilon = 0$ is a direct product of rotators—along $\tilde{\Lambda}$—and a pendulum, the hyperbolic directions of the pendulum become the stable/unstable bundles of the manifold $\tilde{\Lambda}$.

The fact that the tangential exponents are arbitrarily small allows us to conclude that the manifold is as regular as the flow when measured on the C^{r-1} classes, $r \in \mathbb{N}$. (If there was a non-trivial expansion exponent, then the regularity claimed for the manifold would be the infimum of $r - 1$ and a limiting regularity determined by the rates of expansion along the manifold and along the stable/unstable bundles.)

Remark 21. In the general theory of the persistence of overflowing locally invariant manifolds, the manifold obtained need not be unique since, in principle, it could depend on some of the choices made in the proof. Nevertheless, when the manifold $\tilde{\Lambda}_\varepsilon$ is invariant and not just locally invariant, it is unique. We will show later that the manifold $\tilde{\Lambda}_\varepsilon$ is invariant since the boundaries are KAM tori. Hence, for our case, the manifolds will turn out to be unique. Nevertheless, we will not take advantage of this fact since it does not simplify appreciably the proofs and it can fail in more complicated models that can be analyzed by the methods presented here. □

Remark 22. Note that, as we discuss after Theorem 20, the parameterization $\tilde{\mathcal{F}}$ is not uniquely defined even after the locally invariant manifold has been fixed. Indeed, if we compose $\tilde{\mathcal{F}}$ with any diffeomorphism of $\tilde{\Lambda}$ on the right we obtain a different parameterization.

We can make the parameterization unique by imposing that $\tilde{\mathcal{F}}$ does not change the (I, φ, s) coordinates:

$$\Pi_I \tilde{\mathcal{F}}(I, \varphi, s; \varepsilon) = I,\ \Pi_\varphi \tilde{\mathcal{F}}(I, \varphi, s; \varepsilon) = \varphi,\ \Pi_s \tilde{\mathcal{F}}(I, \varphi, s; \varepsilon) = s. \tag{32}$$

This can be expressed as saying that we write $\tilde{\Lambda}_\varepsilon$ as the graph of a function from $\tilde{\Lambda}$ into the (p, q) plane.

This normalization will be important for us when we compute the map $\tilde{\mathcal{F}}$ in coordinates. □

Remark 23. We recall that the precise definition of the stable manifolds in normal hyperbolicity theory is that they consist on the points that, under the evolution of the extended system, converge to the invariant manifold with an exponential rate close to that of the original system. It follows from the exponential convergence that $W^{\mathrm{s}}_{\tilde{\Lambda}_\varepsilon} = \cup_{\tilde{x} \in \tilde{\Lambda}_\varepsilon} W^{\mathrm{s}}_{\tilde{x}}$ where

$$W^{\mathrm{s}}_{\tilde{x}} = \{\tilde{y} \ : \ \operatorname{dist}(\Phi_{t,\varepsilon}(\tilde{x}), \Phi_{t,\varepsilon}(\tilde{y})) \leq C e^{-\tilde{\mu} t}, t > 0\},$$

where $\tilde{\mu} = \mu + \mathrm{O}(\varepsilon)$. Moreover $W^{\mathrm{s}}_{\tilde{x}} \cap W^{\mathrm{s}}_{\tilde{y}} = \emptyset$ when $\tilde{x} \neq \tilde{y}$, $\tilde{x}, \tilde{y} \in \tilde{\Lambda}_\varepsilon$. The stable manifolds of the points are as smooth as the flow (when smoothness is measured in the class of C^r spaces $r \in \mathbb{N} \cup \{\infty, \omega\}$). Nevertheless, the map $\tilde{x} \mapsto W^{\mathrm{s}}_{\tilde{x}}$ could be less regular than the flow, and therefore, the total manifold $W^{\mathrm{s}}_{\tilde{\Lambda}_\varepsilon} = \cup_{\tilde{x} \in \tilde{\Lambda}_\varepsilon} W^{\mathrm{s}}_{\tilde{x}}$ could, in general, be less smooth than the flow. In our case, however, since the tangent exponents are arbitrarily small, if the flow is C^{r-1}, $r - 1 \in \mathbb{N}$, then the resulting $W^{\mathrm{s}}_{\tilde{\Lambda}_\varepsilon}$ will be C^{r-1} for small ε.

Analogous results hold for the unstable manifold.

Note that the definition of the (un)stable manifolds for locally invariant manifolds can only be made in an extended system constructed in the proof for which the dynamics is defined for all times. In general, these (un)stable manifolds depend on the extended flow used.

One consequence of that is that $\tilde{\Lambda}_\varepsilon$, $W^{\mathrm{s}}_{\tilde{\Lambda}_\varepsilon}$ may fail to be C^∞ even if the flow is analytic. Nevertheless, when the manifolds are invariant, the stable and unstable manifolds are uniquely defined. In our case, as stated in Remark 21, the KAM tori will produce invariant boundaries for $\tilde{\Lambda}_\varepsilon$, hence stable and unstable manifolds will be uniquely defined. The proofs we present, however, do not take advantage of this fact. □

7.1. **Explicit calculations of the perturbed invariant manifold.** In the rest of this section, we will compute rather explicitly the expansions in ε of the map $\tilde{\mathcal{F}}$ for the Hamiltonian (6).

We note that the equation that $\tilde{\mathcal{F}}$ has to satisfy so that its range is invariant by the vector field $\mathcal{H}_\varepsilon$ given in (7) is

$$\mathcal{H}_\varepsilon \circ \tilde{\mathcal{F}} = D\tilde{\mathcal{F}}\mathcal{R} \tag{33}$$

where $\mathcal{R}$ is a vector field in $\tilde{\Lambda}_\varepsilon$. Note that (33) expresses that the vector field $\mathcal{H}_\varepsilon$ at a point in the range is tangent to the range of $\tilde{\mathcal{F}}$.

Using equation (33) one can find an expansion of the function $\tilde{\mathcal{F}}$.

Proposition 24. *The family of mappings $\tilde{\mathcal{F}}$ specified in Theorem 20 with the normalization* (32) *admits an expansion*

$$\tilde{\mathcal{F}} = \tilde{\mathcal{F}}_0 + \varepsilon\tilde{\mathcal{F}}_1 + \cdots + \varepsilon^m\tilde{\mathcal{F}}_m + \mathrm{O}_{C^{r-m-2}}(\varepsilon^{m+1}),$$

where $\tilde{\mathcal{F}}_0(I,\varphi,s) = (0,0,I,\varphi,s)$.

In the case that the flow satisfies assumptions **H1**–**H3***, then the functions $\tilde{\mathcal{F}}_1,\dots,\tilde{\mathcal{F}}_m$ are trigonometric polynomials in φ, s, and $\tilde{\mathcal{F}}_i$ are of class $\mathcal{C}^{r-1-i}$. Moreover, in such a case, $\mathcal{N}(\tilde{\mathcal{F}}_1) \subset \mathcal{N}(h_1)$, $\mathcal{N}(\tilde{\mathcal{F}}_2) \subset \mathcal{N}(h_1) + \mathcal{N}(h_1) \cup \mathcal{N}(h_2)$.*

Proof. Once we impose normalization (32) to $\tilde{\mathcal{F}}$, we can compute $\tilde{\mathcal{F}}_i$ by matching powers of ε in the equation (33). We know by Theorem 20 that the expansion exists.

Equating terms in the expansion on ε of the equation for invariance (33) we obtain, up to order two:

$$\begin{aligned}
\mathcal{H}_0 \circ \tilde{\mathcal{F}}_0 &= D\tilde{\mathcal{F}}_0\mathcal{R}_0 \\
(D\mathcal{H}_0 \circ \tilde{\mathcal{F}}_0)\tilde{\mathcal{F}}_1 + \mathcal{H}_1 \circ \tilde{\mathcal{F}}_0 &= D\tilde{\mathcal{F}}_0\mathcal{R}_1 + D\tilde{\mathcal{F}}_1\mathcal{R}_0 \\
(D\mathcal{H}_0 \circ \tilde{\mathcal{F}}_0)\tilde{\mathcal{F}}_2 + \frac{1}{2}(D^2\mathcal{H}_0 \circ \tilde{\mathcal{F}}_0)\tilde{\mathcal{F}}_1^{\otimes 2} &+ (D\mathcal{H}_1 \circ \tilde{\mathcal{F}}_0)\tilde{\mathcal{F}}_1 + \mathcal{H}_2 \circ \tilde{\mathcal{F}}_0 \\
&= D\tilde{\mathcal{F}}_0\mathcal{R}_2 + D\tilde{\mathcal{F}}_1\mathcal{R}_1 + \mathcal{D}\tilde{\mathcal{F}}_2\mathcal{R}_0.
\end{aligned} \tag{34}$$

In general, the equation obtained after matching the coefficients of ε^n on both sides of (33), is of the form:

$$(D\mathcal{H}_0 \circ \tilde{\mathcal{F}}_0)\tilde{\mathcal{F}}_n - D\tilde{\mathcal{F}}_n\mathcal{R}_0 - D\tilde{\mathcal{F}}_0\mathcal{R}_n = -\mathcal{H}_n \circ \tilde{\mathcal{F}}_0 + \mathcal{S}_n \tag{35}$$

where $\mathcal{S}_n$ is polynomial in $\mathcal{H}_0,\dots,\mathcal{H}_{n-1}$, their derivatives, $\tilde{\mathcal{F}}_0,\dots,\tilde{\mathcal{F}}_{n-1}$, and $\mathcal{R}_0,\dots,\mathcal{R}_{n-1}$.

Clearly, the first equation in (34) has as solution $\tilde{\mathcal{F}}_0 = (0,0,I,\varphi,s)$, and $\mathcal{R}_0 = (0,I,1)$. Therefore, if one can develop a method to solve equations for $\tilde{\mathcal{F}}_n$, $\mathcal{R}_n$ of the form of the linear Hamiltonian System

(35) equal to a preassigned right hand side, we can keep on solving the hierarchy of equations (35) to any order.

In our case, since the unperturbed motion due to $\mathcal{R}_0$ is quasi-periodic, the equations of the form (35) can be solved quite explicitly using Fourier coefficients. (There are more general theories [Lla00] that allow to solve equations of the form (35) even if the motion on the base is not quasi-periodic.)

The theory of the equations that we need is summarized in next Lemma 25. Clearly, applying it recursively, we obtain a proof of Proposition 24. □

Lemma 25. *Let $\tilde{\mathcal{F}}_0$, $\mathcal{H}_0$, $\mathcal{R}_0$ be as in Proposition 24. Given a C^s function, $s \geq 1$, $\eta : \tilde{\Lambda} \to \mathbb{R}^5$, we can find unique C^1 functions ξ, $\rho : \tilde{\Lambda} \to \mathbb{R}^5$ such that*

$$(D\mathcal{H}_0 \circ \tilde{\mathcal{F}}_0)\xi - D\xi\mathcal{R}_0 - D\tilde{\mathcal{F}}_0\rho = \eta \tag{36}$$

and

$$\Pi_I\xi = 0,\ \Pi_\varphi\xi = 0,\ \Pi_s\xi = 0. \tag{37}$$

Furthermore, ξ, ρ are of class C^s and we have, for a constant C that depends only on $\mathcal{H}_0, \mathcal{R}_0$,

$$||\xi||_{C^s} \leq C||\eta||_{C^s}$$
$$||\rho||_{C^s} \leq C||\eta||_{C^s}.$$

Moreover, if η is a trigonometric polynomial, so are ξ and ρ, and we have $\mathcal{N}(\xi) \subset \mathcal{N}(\eta)$, $\mathcal{N}(\rho) \subset \mathcal{N}(\eta)$.

Proof. Note that the unperturbed flow and its differential at $\tilde{\Lambda}$ preserve the I, φ, s directions. The plane p, q is invariant.

Moreover, $D\tilde{\mathcal{F}}_0$ in the coordinates we are using is a 5×3 matrix. It consists of a 3×3 identity along the I, φ, s directions and 0 along the other directions (p, q).

The two observations above, immediately give us that we can satisfy the normalization (37) in a unique way by setting:

$$\rho = \Pi_{I,\varphi,s}\eta \tag{38}$$

from which the regularity claims about ρ follow.

Using again the invariance of the (p, q) plane it suffices to study the equation

$$(D\mathcal{H}_0 \circ \tilde{\mathcal{F}}_0)|_{p,q}\Pi_{p,q}\xi - D\Pi_{p,q}\xi\mathcal{R}_0 = \Pi_{p,q}\eta \tag{39}$$

This equation (39) can be further reduced by noticing that $D\mathcal{H}_0$ preserves the two eigendirections of the equilibrium point of the pendulum. Hence, if we denote by Π^{s}, Π^{u}, the projections along the stable and the unstable components and by μ the eigenvalue, we obtain that (39) is equivalent to:

$$\begin{aligned} \mu\Pi^{\mathrm{u}}\xi - D\Pi^{\mathrm{u}}\xi\mathcal{R}_0 &= \Pi^{\mathrm{u}}\eta \\ -\mu\Pi^{\mathrm{s}}\xi - D\Pi^{\mathrm{s}}\xi\mathcal{R}_0 &= \Pi^{\mathrm{s}}\eta \end{aligned} \tag{40}$$

The equations (40) can be studied easily. For example, they can be studied noting that the operator on the linear Hamiltonian system of (40) is diagonal on Fourier series. Hence, if

$$\Pi^{\mathrm{s}}\eta(I,\varphi,s) = \sum_{(k,l)\in\mathcal{N}(\eta)} \hat{\eta}_{k,l}(I)e^{i(k\varphi+ls)},$$

the solution of (40) is:

$$\Pi^{\mathrm{s}}\xi(I,\varphi,s) = \sum_{(k,l)\in\mathcal{N}(\eta)} \hat{\eta}_{k,l}(I)[-\mu + i(kI+l)]^{-1}e^{i(k\varphi+ls)} \tag{41}$$

and analogously (with $+\mu$ in place of $-\mu$) for $\Pi^{\mathrm{u}}\xi$.

Note that, since $\pm\mu \neq 0$, there are no small denominators in the finite sum (41). Then, one can bound the norm of ξ by the norm of η. This finishes the proof of Lemma 25. □

Remark 26. Even if for us, the solution 41 is useful because it gives the Fourier coefficients that will later be the base of the study of resonances, we note that, in geometric perturbation theory, it is common to represent the solution as integral formulas.

For example, in our case, we have:

$$\Pi^{\mathrm{s}}\xi(I,\varphi,s) = \int_0^\infty \eta(I,\varphi - It, s-t)e^{-\mu t}\,dt$$

Out of the preceding formula, it is easy to read off the regularity properties even in the case that η is not a polynomial. □

8. The dynamics in $\tilde{\Lambda}_\varepsilon$

In this section, we will study the dynamics restricted to the manifold $\tilde{\Lambda}_\varepsilon$.

The upshot of the discussion will be that, under hypothesis of differentiability and **H5**, the invariant manifold $\tilde{\Lambda}_\varepsilon$ contains primary tori (see Propositions 47, 50, Theorem 56, Corollary 57), secondary tori (see Theorem 56, Corollary 57) and stable and unstable manifolds of periodic orbits (see Proposition 66). They are arranged in such a way that the gaps among them are a power of ε which can be made as large as desired by assuming enough differentiability. For subsequent developments, any power greater than 1 will be enough. Hence, we have developed all the section trying to obtain only gaps of order $\varepsilon^{3/2}$.

The results are summarized in Figure 1, which depicts a surface of section of the manifold $\tilde{\Lambda}_\varepsilon$.

The method of proof will be a sequence of different steps. Basically, they will be a combination of averaging methods and KAM theorems.

In Section 8.1 we study the geometry of the motion restricted to the manifold $\tilde{\Lambda}_\varepsilon$. In particular, we will see that the flow restricted to $\tilde{\Lambda}_\varepsilon$ is Hamiltonian. Moreover, for convenience, we will introduce a system of coordinates (J, φ, s), in which the symplectic form has the standard expression.

In Section 8.2 we compute very explicitly the reduced Hamiltonian $k(J, \varphi, s; \varepsilon) = k_0(J, \varphi, s) + \varepsilon k_1(J, \varphi, s) + \cdots + \varepsilon^m k_m(J, \varphi, s) + \mathrm{O}(\varepsilon^{m+1})$, which is the restriction of the Hamiltonian $\tilde{H}_\varepsilon$ to $\tilde{\Lambda}_\varepsilon$ expressed in these appropriate action-angle coordinates. In particular, we will show that if the perturbation $h = h_1 + \varepsilon h_2 + \cdots + \varepsilon^{m-1} h_m + \mathrm{O}(\varepsilon^m)$ in (6) is a trigonometric polynomial with respect to (φ, s), so are $k_0, k_1, \ldots, k_m$, and we will give rather explicit formulas to compute the k_i's in terms of the h_i's.

In Section 8.3 we develop a global averaging procedure that casts the reduced Hamiltonian $k(J, \varphi, s; \varepsilon)$ into a convenient global normal form $\bar{k}(\mathcal{B}, \alpha, s; \varepsilon)$ which is given by different formulas in the resonant regions and in the non-resonant region.

The non-resonant region is studied in Section 8.4. Since the normal form $\bar{k}(\mathcal{B}, \alpha, s; \varepsilon)$ is very close to a strongly integrable Hamiltonian, a quantitative version of the KAM theorem, which we will develop, will show that the non-resonant region contains KAM tori which leave very small gaps between them.

The resonant region is analyzed in Section 8.5. In this region the normal form $\bar{k}(\mathcal{B}, \alpha, s; \varepsilon)$ is very close to a pendulum. Note that the pendulum has rotational and librational motions covering open sets as well as separatrices. The rotations in the pendulum have the same topology as the primary tori in the integrable system. The librations are contractible to a periodic orbit. Hence the librations correspond to motions with topologies that are not present in the integrable system.

The heart of the matter is Section 8.5.3 which shows that many of these rotational and librational orbits of the pendulum persist when we consider the error terms of the normal form. The rotational orbits will become primary KAM tori and the librational orbits will become secondary tori. For our purposes, it will be important to show that both the secondary tori and the primary tori can be found up to distances which are very close to the separatrices.

An important technical tool for the proofs of persistence of tori close to separatrices appears in Section 8.5.4 where we develop a system of action-angle coordinates in a neighborhood of the separatrices. This change of coordinates is singular (the $\mathcal{C}^r$ norms blow up as a negative power of the distance to the separatrix). Since the remainder contains a power of ε, we will see that a KAM theorem can be applied provided that the distance to the separatrix is bigger than a constant power of ε. This power is arbitrarily high if the original system is differentiable enough.

In the language which is customary in the heuristic study of diffusion, what we have done is to develop upper bounds for the size of the chaotic sea. Related arguments for analytic mappings appear in [Neĭ84]. In the analytic case, the distance can be chosen to be exponentially small.

Furthermore, in Section 8.5.5, we study what happens to the separatrices of the pendulum when we include the error terms of the normal form. They become stable and unstable manifolds or lower dimensional tori (indeed, periodic orbits in our setting). See Figure 1 for a depiction of these invariant objects.

The distances between these invariant objects can be bounded thanks to the non-degeneracy assumption **H5** by a power of the perturbation. This power can be made arbitrarily large by assuming more differentiability. We will take it to be just $O(\varepsilon^{3/2})$ since this is enough for subsequent applications.

Remark 27. Even if hypothesis **H5** is formulated in **H5'**, **H5"** and **H5"'** (see Section 8.5.3 and Section 9) in terms of the reduced Hamiltonian $k(J, \varphi, s; \varepsilon)$, we will have explicit formulas (see (46), (47)) that give the restricted Hamiltonian in terms of H_ε, the original one.

We have striven to make conditions **H5'**, **H5"** and **H5"'** quite explicit so that they can be verified rather easily for a concrete system. Nevertheless, we point out that the fact that they hold for a generic set—indeed a set of finite codimension—is rather easy. □

8.1. **A system of coordinates for $\tilde{\Lambda}_\varepsilon$.** In this subsection, we will describe the construction of a system of coordinates in $\tilde{\Lambda}_\varepsilon$ that will help us to analyze the motion of the flow restricted to $\tilde{\Lambda}_\varepsilon$. In this system of coordinates, the symplectic form on $\tilde{\Lambda}_\varepsilon$ has the standard form and the flow is a Hamiltonian flow described by a time periodic one degree of freedom Hamiltonian.

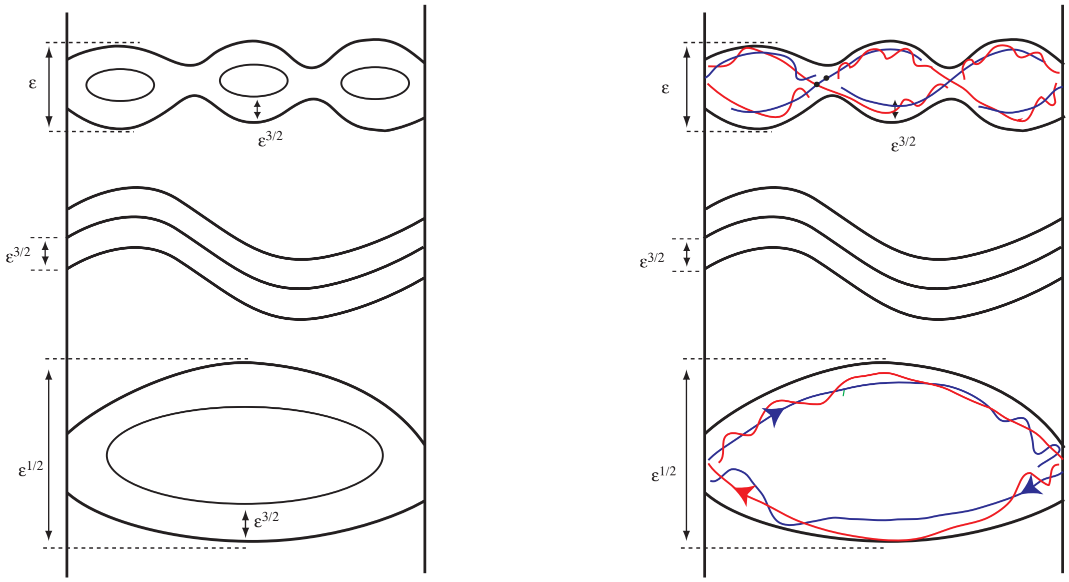

FIGURE 1. Surface of section of $\tilde{\Lambda}_\varepsilon$ illustrating the main invariant objects. The primary KAM tori and the secondary tori on the left. The primary tori and the stable and unstable manifolds of periodic orbits on the right.

We note that several of the notions that we use (e.g. resonances, which are related to the support of the Fourier transform) are dependent on the system of coordinates we use. Of course, we will check at the end that the geometric constructions are independent of the system of coordinates, but several of the analytic arguments using Fourier series, etc., depend on taking a system of coordinates.

Theorem 20 does not use much the structure of the flow (the only thing that we use of the structure of (6) is that $||D\Phi_{t,0}|_{\tilde{\Lambda}}|| \leq \text{cte.}\, t$). In the following paragraphs, we will develop some consequences of the facts that the perturbed and unperturbed flow are the autonomization of a Hamiltonian periodic flow.

The fact that the coordinate s evolves always according to $\dot{s} = 1$, implies that we can decompose the manifold $\tilde{\Lambda}_\varepsilon$ as

$$\tilde{\Lambda}_\varepsilon = \cup_{s \in \mathbb{T}^1} \Lambda^s_\varepsilon \times \{s\}$$

according to the values that the coordinate s takes.

Proposition 28. *If $r \geq 3$, in the manifold $\tilde{\Lambda}_\varepsilon$ there is a $\mathcal{C}^{r-2}$ system of coordinates (J, φ, s), where $J = \mathcal{J}(I, \varphi, s; \varepsilon)$, which is characterized uniquely by the following conditions:*

i) *The angle coordinates* φ *and* s *are the same as those of the unperturbed system.*
ii) *The symplectic form in* Λ_ε^s *is given by* $\omega|_{\Lambda_\varepsilon^s} = dJ \wedge d\varphi$.
iii) $\mathcal{J}(0, \varphi, s; \varepsilon) = 0$.

Moreover:

iv) *The function* $\mathcal{J}(I, \varphi, s; \varepsilon)$ *is* C^{r-2} *and is* ε^2*-close, in the* C^{r-3} *sense, to the function* I.
v) *There are expansions:*

$$\mathcal{J}(I, \varphi, s; \varepsilon) = I + \varepsilon^2 \mathcal{J}_2(I, \varphi, s) + \cdots + \varepsilon^m \mathcal{J}_m(I, \varphi, s) + \mathrm{O}_{\mathcal{C}^{r-m-3}}(\varepsilon^{m+1}),$$

where $\mathcal{J}_i$*, of class* $\mathcal{C}^{r-i-1}$*, are trigonometric polynomials in* (φ, s)*. Moreover* $\mathcal{N}(\mathcal{J}_2) \subset \mathcal{N}(h_1) + \mathcal{N}(h_1)$.

Proof.

It is clear that $\Phi_{t,\varepsilon}(\Lambda_\varepsilon^s \times \{s\}) = \Lambda_\varepsilon^{s+t} \times \{s+t\}$.

Since $\Lambda_0^s = \Lambda := \{(0, 0, I, \varphi) : (I, \varphi) \in [I_-, I_+] \times \mathbb{T}\}$ for any $s \in \mathbb{T}$, we have that $\omega|_{\Lambda_0^s}$ is a canonical symplectic form: $\omega|_{\Lambda_0^s} = dI \wedge d\varphi$.

Since Λ_ε^s is ε-close to Λ_0^s in the C^1 sense, it follows that $\omega_\varepsilon^s \equiv \omega|_{\Lambda_\varepsilon^s}$ is non-degenerate and, since it is closed, it is a symplectic form.

Since the restriction of vector fields, forms and the exterior differential to a manifold behave in a functorial way, we have that the restriction of the Hamiltonian $\tilde{H}_\varepsilon$ to the manifold $\Lambda_\varepsilon^s \times \{s\}$ generates the vector field $\mathcal{H}_\varepsilon|_{\Lambda_\varepsilon^s \times \{s\}}$, where $\mathcal{H}_\varepsilon$ is the vector field of (7), restricted to the manifold using the symplectic form given by $\tilde{H}_\varepsilon$.

The map $\tilde{\mathcal{F}} = (\mathcal{F}, \mathrm{Id}_\mathbb{T})$ given by Theorem 20 allows us to define a system of coordinates in $\tilde{\Lambda}_\varepsilon$ by transporting the coordinates defined in $\tilde{\Lambda}$. Note that $\tilde{\mathcal{F}}$ satisfies normalization (32), so the coordinates (I, φ, s) remain unaltered. We will use this to introduce symplectically conjugate coordinates as follows:

We can use $\mathcal{F}$ to pull back to Λ the form ω_ε^s. We note that $\mathcal{F}^*\omega_\varepsilon^s$ is close to $dI \wedge d\varphi$. Indeed, since the manifold Λ_ε^s is two-dimensional, we know that

$$\mathcal{F}^*\omega_\varepsilon^s = a(I, \varphi, s; \varepsilon) dI \wedge d\varphi, \tag{42}$$

where

$$a(I, \varphi, s; \varepsilon) = 1 + \{\Pi_p \tilde{\mathcal{F}}, \Pi_q \tilde{\mathcal{F}}\} = 1 + \varepsilon^2 \{\Pi_p \tilde{\mathcal{F}}_1, \Pi_q \tilde{\mathcal{F}}_1\} + \cdots$$

is of class $\mathcal{C}^{r-2}$. We claim in ii) that there is $J = \mathcal{J}(I, \varphi, s; \varepsilon)$, such that

$$\mathcal{F}^*\omega_\varepsilon^s = dJ \wedge d\varphi. \tag{43}$$

Indeed, the only function $\mathcal{J}$ satisfying this requirement is, up to an additive function of φ,

$$\mathcal{J}(I,\varphi,s;\varepsilon) = \int_0^I a(L,\varphi,s;\varepsilon)\,dL, \tag{44}$$

which is C^{r-2}. This additive function is identically zero to satisfy the normalization iii).

$\mathcal{J}$ is C^1-close to I on compact sets, so we have that (J,φ,s) will be a good system of coordinates on $\tilde\Lambda$.

We can now push forward this function $\mathcal{J}$ and use on $\tilde\Lambda_\varepsilon$ the coordinate system (J,φ,s).

By Proposition 24, we know that $\tilde{\mathcal{F}}_1$ is a trigonometric polynomial, $\mathcal{N}(\tilde{\mathcal{F}}_1) = \mathcal{N}(h_1)$, and that $\tilde{\mathcal{F}}_2$ is a trigonometric polynomial and $\mathcal{N}(\tilde{\mathcal{F}}_2) \subset N(h_1) + \mathcal{N}(h_1) \cup \mathcal{N}(h_2)$. Since $\tilde{\mathcal{F}}_i$ are trigonometric polynomials, we obtain that $J = I + \varepsilon^2 \mathcal{J}_2(I,\varphi,s) + \cdots + \varepsilon^m \mathcal{J}_m(I,\varphi,s;\varepsilon)$, where $\mathcal{J}_i$ are trigonometric polynomials in (φ,s), $\mathcal{N}(\mathcal{J}_2) \subset \mathcal{N}(\tilde{\mathcal{F}}_1) + \mathcal{N}(\tilde{\mathcal{F}}_1) \subset \mathcal{N}(h_1) + \mathcal{N}(h_1)$. $\square$

Remark 29. Notice that the explicit formula (44) requires that I, φ are one-dimensional.

Nevertheless, we point out that a similar result—with a higher loss of derivatives—can be established in higher dimensions using a global version of the Darboux Theorem, subject to the constraint that the change of coordinates does not change the variable φ.

Since the one-dimensional case is enough for the purposes of this paper, we will not purse the matter here. $\square$

8.2. **Calculation of the reduced Hamiltonian.** The goal of this section is to examine the reduced vector field $\mathcal{H}_\varepsilon|_{\Lambda_\varepsilon^s \times \{s\}}$ when written in the variables (J,φ,s) given by Proposition 28.

Proposition 30. *Let $k(J,\varphi,s;\varepsilon)$ be the Hamiltonian of the vector field $\mathcal{H}_\varepsilon$ restricted to $\tilde\Lambda_\varepsilon$, and expressed in the variables given by Proposition 28.*

If h is C^r in all the variables and $r \geq 3$, then k is C^{r-2} in all the variables. Therefore, for $r-2 > m \geq 0$, we can expand

$$k(J,\varphi,s;\varepsilon) = k_0(J) + \varepsilon k_1(J,\varphi,s) + \cdots + \varepsilon^m k_m(J,\varphi,s) + O_{C^{r-m-3}}(\varepsilon^{m+1}) \tag{45}$$

where $k_0(J) = \frac{J^2}{2}$, and k_i are of class C^{r-2-i}.

In case that

$$h(p,q,I,\varphi,s;\varepsilon) = h_1(p,q,I,\varphi,s) + \cdots + \varepsilon^{m-1} h_m(p,q,I,\varphi,s) + O_{C^{r-m}}(\varepsilon^m),$$

and $h_1,\dots,h_m$ are trigonometric polynomials in (φ,s), then

i) $k_1, \dots, k_m$ *are trigonometric polynomials in* (φ, s).

ii)
$$(k_1, k_2) = \mathcal{G}(h_1|_{p=0,q=0}, h_2|_{p=0,q=0}, \nabla_{p,q} h_1|_{p=0,q=0}, \nabla^2_{p,q} h_1|_{p=0,q=0})$$
where $\mathcal{G}$ *is a polynomial function which can be written explicitly.*

iii)
$$\mathcal{G}(0,0,0,0) = 0, \quad D_{h_1}\mathcal{G}(0,0,0,0) = D_{h_2}\mathcal{G}(0,0,0,0) = Id,$$
and all the other partial derivatives of $\mathcal{G}$ *evaluated at* $(0,0,0,0)$ *are zero.*

iv) $\mathcal{N}(k_1) \subset \mathcal{N}(h_1)$, *and* $\mathcal{N}(k_2) \subset \mathcal{N}(h_1) + \mathcal{N}(h_1) \cup \mathcal{N}(h_2)$.

Among the results of Proposition 30, the most important for us is that the restricted Hamiltonian is still a trigonometric polynomial up to order ε^m. The subsequent analysis will take some advantage of the fact that if we truncate to a fixed order in ε, we only need to deal with a Hamiltonian with a finite number of harmonics.

The analysis we will present later will have as hypothesis that k_1, k_2 avoid some manifolds of positive codimension in the space of trigonometric polynomials. The conclusion ii) of Proposition 30 shows that this hypothesis is implied by h_1, h_2 avoiding a manifold of positive codimension in the set of trigonometric polynomials. In particular, we conclude that the results of subsequent analysis are verified for a generic set of h's.

Proof.

It is almost obvious using that, by Proposition 24, one can write rather explicitly a formal power series in ε for the function $\tilde{\mathcal{F}}$, and that all the terms in the expansion are trigonometric polynomials.

We start by noting that since the Hamiltonian restricted to the invariant manifold generates the restricted flow when we consider the restricted form, we have that the function
$$K : [I_-, I_+] \times \mathbb{T} \times \mathbb{T} \times [-\varepsilon^*, \varepsilon^*] \to \mathbb{R}$$
defined by: $K(I, \varphi, s; \varepsilon) = H_\varepsilon \circ \tilde{\mathcal{F}}(I, \varphi, s; \varepsilon)$ defines the Hamiltonian as a function of the variables $(I, \varphi, s; \varepsilon)$.

Then, the desired Hamiltonian in the coordinate system (J, φ, s) will be:
$$k(J, \varphi, s; \varepsilon) = H_\varepsilon \circ \tilde{\mathcal{F}}(\mathcal{I}(J, \varphi, s; \varepsilon), \varphi, s; \varepsilon), \tag{46}$$
where we have denoted by $\mathcal{I}(J, \varphi, s; \varepsilon)$ the inverse of the function $I \mapsto \mathcal{J}(I, \varphi, s; \varepsilon)$ given in (44), which is a $\mathcal{C}^{r-2}$ function. Since the function $\mathcal{J}$ has an expansion in ε given in Proposition 28 the function $\mathcal{I}$ has an expansion with similar properties.

Therefore, if we expand in ε the expression for $k(J, \varphi, s; \varepsilon)$ given in (46), we obtain the results claimed in Proposition 30. Moreover,

$$k_1(J, \varphi, s) = DH_0 \circ \tilde{\mathcal{F}}_0(J, \varphi, s) \tilde{\mathcal{F}}_1(J, \varphi, s) + h_1 \circ \tilde{\mathcal{F}}_0(J, \varphi, s), \tag{47}$$

and one can obtain explicit formulas for k_2, etc.

□

8.3. **Isolating the resonances (resonant averaging).** The main result of this section will be Theorem 35. This Theorem makes precise the notion that, when the perturbation is a trigonometric polynomial, we can find a finite set of resonances where the averaging breaks down. When we are at a distance L—which we will take as a fixed number—from the resonances we can perform averaging transformations and reduce the system very close to strongly integrable. Near the resonances, the averaging transformations reduce the system very close to an integrable pendulum.

Remark 31. Even if we hope that the treatment of this section is quite self-contained, we mention that very pedagogical treatments of averaging theory can be found in [AKN88], [LM88]. Arguments very similar to those used here appear in the proof of KAM theorem using the strategy of [Arn63a]. See [Sva80] and specially [DG96]. The main difference between our study and that of KAM theory is that we will pay attention to the description of the phenomena that happen in the resonant regions. In KAM theory, the resonant regions are not analyzed and the only thing done with them is to estimate their size.

□

Remark 32. We will perform the averaging procedure to a rather high order (i.e. ε^{26}) and to exclude from the resonances a size L which is independent of the resonances and of the parameter ε.

Choosing the order in ε to be as high is done to simplify the exposition later. As we will see the only resonances that produce effects that affect subsequent arguments are those of order 1 and 2 (we will show that resonances of order j produce gaps among KAM tori of size $\varepsilon^{j/2}$ and there is a standard method to transverse those gaps of order smaller than ε). Hence, any order of averaging greater or equal than 2 would have been enough. Nevertheless, working harder at the averaging step will make other arguments in Section 8.5.2 simpler. We decided that this future simplification was worth that the very slight complication in the present chapter. □

Remark 33. Choosing the width of the resonant regions to be a fixed number L is rather wasteful. It is well known that one can choose the width to be a suitable power of ε multiplied by the size of the Fourier coefficients, which decrease with $|k|$.

For the purposes of this paper, choosing that width to be a constant is enough and simplifies the exposition. A more precise choice would be needed to eliminate the hypothesis **H3** from Theorem 7 that requires the perturbation to be a trigonometric polynomial. A heuristic discussion of these choices is included in Section 12.4. $\square$

8.3.1. *The infinitesimal equations for averaging.* Now we turn to the implementation of the above strategy. We first start discussing the first order equations, which will serve as motivation for the phenomena of resonances.

We recall that if we take g to be the canonical transformation obtained by the time-one flow of a periodic in time Hamiltonian εG then, given a function K, we have that $K \circ g = K + \varepsilon\{K, G\} + \mathrm{O}(\varepsilon^2)$ where $\{\ ,\ \}$ denotes the Poisson bracket in the canonical coordinates (J, φ, A, s). If $K = K_0 + \varepsilon K_1$, we see that $K \circ g = K_0 + \varepsilon(K_1 + \{K_0, G\}) + \mathrm{O}(\varepsilon^2)$. Hence it is natural to consider equations of the form:

$$K + \{K_0, G\} = \bar{K}, \tag{48}$$

where $K(J, \varphi, s)$ is a given Hamiltonian, $K_0 = A + \frac{J^2}{2}$, and the unknowns are G, the generator of the infinitesimal transformation, and $\bar{K}$, the averaged Hamiltonian, which will be chosen to be as simple as possible. Notice that, in this case, $\{K_0, G\} = J\frac{\partial G}{\partial \varphi} + \frac{\partial G}{\partial s}$ and then, if we write (48) in Fourier coefficients, we have,

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

where $K_{k,l}(J)$, $G_{k,l}(J)$, $\bar{K}_{k,l}(J)$, are the Fourier coefficients of K, G, $\bar{K}$, for $(k, l) \in \mathcal{N}$, the support of the Fourier transform of K.

When

$$Jk + l = 0 \tag{49}$$

we can not choose $G_{k,l}(J)$ in order to have $\bar{K}_{k,l}(J) = 0$. When $k = 0$, $l = 0$, equation (49) happens for all values of J, when $k = 0$, $l \neq 0$, happens for no value of J. For all the cases $k \neq 0$, (49) happens when $J = -l/k$.

We will refer to resonances as the places $J = -l/k$, $(k, l) \in \mathcal{N}$, $k \neq 0$.

Notice that when J is not resonant, we can reduce the system to contain only $\bar{K}_{0,0}(J)$, which is an integrable system.

To preserve smoothness, we will have also to keep a good part of these terms for neighboring values of the action. The tapering off of the elimination involves somewhat arbitrary choices.

In the following Lemma 34 we introduce a specific choice and provide estimates for it.

Lemma 34. *Let*

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

be a Hamiltonian, with $\mathcal{N} = \mathcal{N}(K) \subset \mathbb{Z}^2$ a finite set. Assume that K is of class $\mathcal{C}^{n+1}$ with respect to J.

Choose L to be a number such that the real intervals $[-l/k - 2L, -l/k + 2L]$ for all $(k,l) \in \mathcal{N}$, $k \neq 0$, are disjoint.

Then, there exist G of class $\mathcal{C}^n$ with respect to J, and $\bar{K}$ of class $\mathcal{C}^{n+1}$, such that they solve the homological equation (48) and that verify:

(1) *If $|J + l/k| \geq 2L$ for any $(k,l) \in \mathcal{N}$, $k \neq 0$, then*
$$\bar{K}(J,\varphi,s) = K_{0,0}(J).$$
(2) *If $|J + l_0/k_0| \leq L$ for some $(k_0,l_0) \in \mathcal{N}$, $k_0 \neq 0$, then*
$$\begin{aligned}\bar{K}(J,\varphi,s) &= K_{0,0}(J) + \sum_{t=-N}^{N} K_{tk_0,tl_0}(-l_0/k_0)e^{it(k_0\varphi+l_0s)}\\ &=: K_{0,0}(J) + U_{k_0,l_0}(k_0\varphi + l_0 s),\end{aligned}$$
where $0 < N < \infty$ is such that $(tk_0, tl_0) \in \mathcal{N}$, for $|t| \leq N$.

(3) *The function $\bar{K}$ verifies: $\left|\bar{K}\right|_{\mathcal{C}^{n+1}} \leq (1 + \frac{C}{L^{n+1}})\,|K|_{\mathcal{C}^{n+1}}$, where C is a constant independent of L.*

(4) *The function G verifies $|G|_{\mathcal{C}^n} \leq \dfrac{C}{L^n}\,|K|_{\mathcal{C}^{n+1}}$.*

(5) *$\mathcal{N}(G)$ and $\mathcal{N}(\bar{K})$ are finite sets. Moreover $\mathcal{N}(G)$, $\mathcal{N}(\bar{K}) \subset \mathcal{N}(K)$.*

Proof.

If we write (48) in Fourier coefficients, we have, for $(k,j) \in \mathcal{N}$:

$$K_{k,l}(J) + i(Jk+l)G_{k,l}(J) = \bar{K}_{k,l}(J), \tag{50}$$

Then, in order to solve equation (50), we choose:

(1) If $(0,0) \in \mathcal{N}$, $\bar{K}_{0,0}(J) = K_{0,0}(J)$.

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

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

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

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

(a) If $|J + l/k| \le L$ then $\bar{K}_{k,l}(J) = K_{k,l}(-l/k)$,
(b) if $|J + l/k| \ge 2L$ then $\bar{K}_{k,l}(J) = 0$.

Once we have defined $\bar{K}$ as above, it is clear that it is a $\mathcal{C}^{n+1}$ function, and that it verifies the desired bounds, where the constant C only depends on the cut-off function ψ, and on n.

Now, we choose G to verify equation (50):

(1) $G_{0,0}(J) = 0$, and $G_{k,l}(J) = 0$ if $(k, l) \notin \mathcal{N}$.
(2) $G_{0,l}(J) = -\dfrac{K_{0,l}(J)}{il}$,
(3) If $(k, l) \in \mathcal{N}$, $k \neq 0$, we choose $G_{k,l}(J)$ as:
   (a) If $J \neq -\dfrac{l}{k}$ then $G_{k,l}(J) = \dfrac{\bar{K}_{k,l}(J) - K_{k,l}(J)}{i(Jk + l)}$,
   (b) $G_{k,l}(-l/k) = \lim\limits_{J\to -l/k} \dfrac{\bar{K}_{k,l}(J) - K_{k,l}(J)}{i(Jk + l)} = \dfrac{-K'_{k,l}(-l/k)}{ik}$.

To bound the function G, we note that, given a fixed $(k_0, l_0) \in \mathcal{N}$, by the definition of $\bar{K}$ and G, we have:

(1) $\forall J$, then $|G_{0,l}(J)|_{\mathcal{C}^n} \le C\dfrac{|K_{0,l}|_{\mathcal{C}^n}}{|l|}$.
(2) If $|J + l_0/k_0| \le L$ then $|G_{k_0,l_0}(J)|_{\mathcal{C}^n} \le C\dfrac{|K_{k_0,l_0}|_{\mathcal{C}^{n+1}}}{|k_0|}$.
(3) If $|J + l_0/k_0| \ge 2L$ then $|G_{k_0,l_0}(J)|_{\mathcal{C}^n} \le C\dfrac{|K_{k_0,l_0}|_{\mathcal{C}^n}}{|k_0|\,(2L)^n}$.
(4) If $L \le |J + l_0/k_0| \le 2L$ then

$$|G_{k_0,l_0}(J)|_{\mathcal{C}^n} \le C\frac{|K_{k_0,l_0}(-l_0/k_0)|\,|\psi|_{\mathcal{C}^n} + |K_{k_0,l_0}|_{\mathcal{C}^n}}{|k_0|\,L^n}.$$

Then, $G(J, \varphi, s)$ is a trigonometric polynomial in (φ, s), and of class $\mathcal{C}^n$ with respect to J. □

8.3.2. *The main averaging result, Theorem 35.* Once we know how to solve any homological equation (48), we can proceed to obtain a suitable global normal form of our reduced Hamiltonian by applying repeatedly the procedure. The precise result is formulated in the following Theorem 35.

Theorem 35. *Let $k(J, \varphi, s; \varepsilon)$ be a $\mathcal{C}^n$ Hamiltonian, $n > 1$, and consider any $1 \le m < n$, independent of ε. Assume that*

$$k(J, \varphi, s; \varepsilon) = \frac{J^2}{2} + \varepsilon k^1(J, \varphi, s; \varepsilon). \tag{51}$$

Let $k_i(J,\varphi,s)$ $i=1,\dots,m$ be the coefficients in the Taylor expansion with respect to ε of $k^1(J,\varphi,s;\varepsilon)$, and assume that the $k_i(J,\varphi,s)$, $i=1,\dots,m$ are trigonometric polynomials in φ,s.

Then, there exists a finite set (of resonances)

$$\mathcal{R}=\mathcal{R}_1\cup\cdots\cup\mathcal{R}_m\subset\mathbb{Q}$$

(we will give rather explicit expressions for $\mathcal{R}_i$ involving the support of the Fourier transform of the k_i) such that:

Given L a number, independent of ε, such that the real intervals $[-l/k-2L,l/k+2L]$ for $-l/k\in\mathcal{R}$, are disjoint, there exists a symplectic change of variables, depending on time, $(\mathcal{B},\alpha,s)\mapsto(J,\varphi,s)$, periodic in φ and s, and of class $\mathcal{C}^{n-2m}$, 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

$$\bar{k}(\mathcal{B},\alpha,s;\varepsilon) \quad = \quad \bar{k}^0(\mathcal{B},\alpha,s;\varepsilon)+\varepsilon^{m+1}\bar{k}^1(\mathcal{B},\alpha,s;\varepsilon)$$

where the function $\bar{k}^0$ is of class $\mathcal{C}^{n-2m+2}$, and $\varepsilon^{m+1}\bar{k}^1$ is of class $\mathcal{C}^{n-2m}$ and they verify:

(1) *If $|\mathcal{B}+l/k|\geq 2L$ for any $(k,l)\in\mathbb{Z}^2$ such that $-l/k\in\mathcal{R}$, then*

$$\bar{k}^0(\mathcal{B},\alpha,s;\varepsilon)=\frac{1}{2}\mathcal{B}^2+\varepsilon\bar{k}^{0,0}(\mathcal{B};\varepsilon)$$

where $\bar{k}^{0,0}(\mathcal{B};\varepsilon)$ is a polynomial of degree $m-1$ in ε.

(2) *If $|\mathcal{B}+l_1/k_1|\leq L$ for some $(k_1,l_1)\in\mathbb{Z}^2$, and for some $1\leq i\leq m$ we have that $-l_1/k_1\in\mathcal{R}_i\setminus(\mathcal{R}_1\cup\cdots\cup\mathcal{R}_{i-1})$, then*

$$\bar{k}^0(\mathcal{B},\alpha,s;\varepsilon)=\frac{1}{2}\mathcal{B}^2+\varepsilon\bar{k}^{0,0}(\mathcal{B};\varepsilon)+\varepsilon^iU^{k_1,l_1}(k_1\alpha+l_1s;\varepsilon)$$

where $U^{k_1,l_1}(\theta;\varepsilon)$ is a polynomial in ε and a trigonometric polynomial in θ.

(3) *If $|\mathcal{B}+l_0/k_0|\leq L$ for some $(k_0,l_0)\in\mathbb{Z}^2$ and $-l_0/k_0\in\mathcal{R}_1$, then the function U^{k_0,l_0} defined in 2 is given by:*

$$U^{k_0,l_0}(\theta;\varepsilon)=\sum_{t=-N}^{N}\hat{k}^1_{tk_0,tl_0}(-l_0/k_0;0)e^{it\theta}+\mathrm{O}(\varepsilon) \tag{52}$$

where $\hat{k}^1_{k,l}(J;\varepsilon)$ are the Fourier coefficients of the function $k^1(J,\varphi,s;\varepsilon)$ with respect to (φ,s).

8.3.3. *Proof of Theorem 35.* The proof of this Theorem will follow by the repeated application of the following inductive Lemma 36, that we will prove using the method of Lie transforms [Car81, Mey91, Lla01].

The hypothesis of Lemma 36 will be that we have a Hamiltonian already in normal form outside of a set or resonances. The conclusions will be that, excluding an slightly larger set of resonances— which we will give rather explicitly—we can produce another Hamiltonian which is normalized to a higher order in ε.

Lemma 36. *Consider a Hamiltonian of the form:*

$$k_q(J,\varphi,s;\varepsilon) = k_q^0(J,\varphi,s;\varepsilon) + \varepsilon^{q+1}k_q^1(J,\varphi,s;\varepsilon) \tag{53}$$

where

1. $k_0^0(J,\varphi,s;\varepsilon) = \frac{J^2}{2}$ *and, if* $q \geq 1$, $k_q^0(J,\varphi,s;\varepsilon)$ *is a* $\mathcal{C}^{n+2-2q}$ *function that verifies:*

   There exist finite sets $\mathcal{R}_i \subset \mathbb{Q}$*, called resonances of order* i*,* $i = 1\cdots q$*, and a number* $L > 0$ *such that:*

   1.0. *The intervals* $\mathcal{I}_{-l/k} \equiv [-l/k - 2L, -l/k + 2L]$*,* $-l/k \in \mathcal{R}^{[\leq q]} \equiv \bigcup_{i=1,\dots,q} \mathcal{R}_i$ *are disjoint.*

   1.1. *If* $J \notin \bigcup_{-l/k\in\mathcal{R}^{[\leq q]}} \mathcal{I}_{-l/k}$*, then*

   $$k_q^0(J,\varphi,s;\varepsilon) = \frac{J^2}{2} + \varepsilon k_q^{0,0}(J;\varepsilon),$$

   where $\varepsilon k_q^{0,0}(J;\varepsilon)$ *is a polynomial of degree* q *in* ε*.*

   1.2. *If* $|J + l_1/k_1| \leq L$ *for some* $(k_1, l_1) \in \mathbb{Z}^2$ *such that*

   $$-l_1/k_1 \in \mathcal{R}_i \setminus (\mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{i-1}),$$

   for some $1 \leq i \leq q$*, then*

   $$k_q^0(J,\varphi,s;\varepsilon) = \frac{J^2}{2} + \varepsilon k_q^{0,0}(J;\varepsilon) + \varepsilon^i U_q^{k_1,l_1}(k_1\varphi + l_1 s;\varepsilon)$$

   where $U_q^{k_1,l_1}(\theta;\varepsilon)$ *is a polynomial in* ε *and a trigonometric polynomial in* $\theta \equiv k_1\varphi + l_1 s$*.*

2. $k_q^1(J,\varphi,s;\varepsilon)$ *is a* $\mathcal{C}^{n-2q}$ *function whose Taylor series coefficients with respect to* ε *are trigonometric polynomials in* (φ, s)*.*

Denote by $K = k_q^1(J,\varphi,s;0)$*, which is the term of the perturbation of order exactly* $q+1$*. Introduce also the set*

$$\mathcal{R}_{q+1} = \{-l/k,\ (k,l) \in \mathcal{N}(k_q^1(\cdot;0)),\ k \neq 0\}. \tag{54}$$

Choose $0 < \tilde{L} < L$ *such that the intervals* $[-l/k - 2\tilde{L}, -l/k + 2\tilde{L}]$ *are disjoint when* $-l/k \in \mathcal{R}^{[\leq q+1]}$*.*

Let $G(\mathcal{B},\alpha,s)$ be the $\mathcal{C}^{n-2q-1}$ function given by Lemma 34, verifying equation (48) with $K=k_q^1(J,\varphi,s;0)$ and with a distance $\tilde{L}$ away from the resonances.

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

$$(J,\varphi,s)=g(\mathcal{B},\alpha,s),$$

given by the time one flow of the Hamiltonian $\varepsilon^{q+1}G(\mathcal{B},\alpha,s)$, transforms the Hamiltonian $k_q(J,\varphi,s;\varepsilon)$ into a Hamiltonian

$$k_{q+1}(\mathcal{B},\alpha,s;\varepsilon)=k_{q+1}^0(\mathcal{B},\alpha,s;\varepsilon)+\varepsilon^{q+2}k_{q+1}^1(\mathcal{B},\alpha,s;\varepsilon),$$

with

$$k_{q+1}^0(\mathcal{B},\alpha,s;\varepsilon)=k_q^0(\mathcal{B},\alpha,s;\varepsilon)+\varepsilon^{q+1}\bar{k}_q^1(\mathcal{B},\alpha,s;0), \tag{55}$$

where $\bar{k}_q^1(\mathcal{B},\alpha,s;0)=\bar{K}(\mathcal{B},\alpha,s)$, given in Lemma 34, is a $\mathcal{C}^{n-2q}$ function.

Moreover, the Hamiltonian $k_{q+1}^0(\mathcal{B},\alpha,s;\varepsilon)$ verifies properties [1.0], [1.1], [1.2] up to order $q+1$ with $\tilde{L}$ replacing L.

Furthermore, $\varepsilon^{q+1}k_{q+1}^1(J,\varphi,s;\varepsilon)$ is a $\mathcal{C}^{n-2q-2}$ function whose Taylor series coefficients with respect to ε are trigonometric polynomials in (φ,s).

Proof.

We use Lemma 34 to solve equation (48), with $K=k_q^1(\mathcal{B},\alpha,s;0)$, which is a trigonometric polynomial in (α,s), obtaining a $\mathcal{C}^{n-2q-1}$ function $G(\mathcal{B},\alpha,s)$ and a $\mathcal{C}^{n-2q}$ function $\bar{k}_q^1(\mathcal{B},\alpha,s;0)$ which are also trigonometric polynomials in (α,s), and which verify the assumptions of Lemma 34 with the finite resonance set $\mathcal{R}_{q+1}$ defined in (54).

After applying the $\mathcal{C}^{n-2q-2}$ transformation g defined as the time one flow of the Hamiltonian $\varepsilon^{q+1}G$ to the extended autonomous Hamiltonian $A+k_q$, the new Hamiltonian is given by

$$\begin{aligned}
A+k_{q+1} &= (A+k_q)\circ g\\
&= (A+k_q^0)\circ g+\varepsilon^{q+1}k_q^1\circ g\\
&= A+k_q^0+\varepsilon^{q+1}\left(\left\{k_0^0+A,G\right\}+k_q^1\right)\\
&+ \varepsilon^{q+2}k_{q+1}^1,
\end{aligned}$$

where $k_0^0=\frac{J^2}{2}$. Using the homological equation verified by G, we get that the Hamiltonian k_{q+1}^0 in (55) verifies properties [1.0], [1.1], [1.2], up to order $q+1$, and using that $k_q^0=k_0^0+\mathrm{O}_{\mathcal{C}^{n+1-2q}}(\varepsilon)$, we obtain that

$$\begin{aligned}\varepsilon^{q+2}k^1_{q+1} &= (A+k^0_0)\circ g-(A+k^0_0)-\{(A+k^0_0),\varepsilon^{q+1}G\}\\ &+ \left(k^0_q-k^0_0\right)\circ g-\left(k^0_q-k^0_0\right)\\ &+ \varepsilon^{q+1}\left(k^1_q\circ g-k^1_q\right)+\varepsilon^{q+1}[k^1_q-k^1_q(\cdot;0)]\end{aligned}$$

is a bounded $\mathcal{C}^{n-2q-2}$ function because k^0_q is a $\mathcal{C}^{n-2q+2}$ function, k^1_q is a $\mathcal{C}^{n-2q}$ function, and G is analytic with respect to (φ, s).

Moreover, all the terms in the Taylor series of $k^1_{q+1}(\mathcal{B},\alpha,s;\varepsilon)$ with respect to ε, are obtained from a finite number of algebraic operations and Poisson brackets between the Taylor coefficients in ε of k^0_q, those of k^1_q and those of G.

Since all the Taylor coefficients of k^1_q are trigonometric polynomials in the angle variables, the G provided by Lemma 34 is a trigonometric polynomial in the angular variables.

Hence, we conclude that the Taylor coefficients of $k^1_{q+1}(\mathcal{B},\alpha,s;\varepsilon)$ with respect to ε are also trigonometric polynomials in the angle variables. □

Proof of Theorem 35. The proof is by induction in q. To begin the induction we apply the inductive Lemma 36 for $q=0$ to our Hamiltonian 51, with $k^0_0=J^2/2$ and $\varepsilon k^1_0(J,\varphi,s;\varepsilon)=\varepsilon k^1(J,\varphi,s;\varepsilon)$, which is a $\mathcal{C}^n$ function. The set $\mathcal{R}_1=\{J=-l/k,\ (k,l)\in\mathcal{N}(k^1(\cdot;0))\}$, by hypothesis, is a finite set.

Then the change of variables g provided by Lemma 36 is of class $\mathcal{C}^{n-2}$, and we obtain $k_1=k^0_1+\varepsilon^2k^1_1$, where k^0_1 is a $\mathcal{C}^n$ function, and $\varepsilon^2k^1_1$ is a $\mathcal{C}^{n-2}$ function, verifying properties [1.1], [1.2] of Lemma 36 with $q=1$. Applying Lemma 34 one obtains formula (52).

Once we have the normal form up to order q, Lemma 36 gives us the normal form up to order $q+1$ modifying the constant L to accommodate new resonances. We only need to observe that, if $q\geq 1$, the modifications to the Hamiltonian k^0_q are given in the terms of order $O(\varepsilon^{q+1})$, so the expression (52) remains invariant in all the process.

Applying the inductive Lemma 36 m times, we obtain Theorem 35. □

Remark 37. It is clear that we can apply Theorem 35 to our Hamiltonian $k(J,\varphi,s;\varepsilon)$ in (45) given in Proposition 30, with $n=r-2$, and, after m steps of averaging we will obtain a Hamiltonian $\bar{k}(\mathcal{B},\alpha,s;\varepsilon)=\bar{k}_0(\mathcal{B},\alpha,s;\varepsilon)+\varepsilon^{m+1}\bar{k}_1(\mathcal{B},\alpha,s;\varepsilon)$, where $\bar{k}_0(\mathcal{B},\alpha,s;\varepsilon)$ is of class $\mathcal{C}^{r-2m}$ and $\varepsilon^{m+1}\bar{k}_1(\mathcal{B},\alpha,s;\varepsilon)$ is of class $\mathcal{C}^{r-2m-2}$. □

8.4. **The non-resonant region (KAM theorem).** In this section we analyze the Hamiltonian in the non-resonant region $\mathcal{S}^L$ of Theorem 35. That is, we study the region where the action J is far from any resonance of order less or equal than m:

(56)
$$\mathcal{S}^L = \{(J, \varphi, s) \in \tilde{\Lambda}_\varepsilon : \ |J + l/k| \geq 2L, \ (k,l) \in \mathbb{Z}^2 \text{ for } -l/k \in \mathcal{R}\}$$

where $\mathcal{R} = \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_m$, and L are provided by Theorem 35.

By Theorem 35, in the averaged variables $(\mathcal{B}, \alpha, s)$ defined in the connected components of $\mathcal{S}^L$, we can write the Hamiltonian as

$$\bar{k}(\mathcal{B}, \alpha, s; \varepsilon) = \frac{\mathcal{B}^2}{2} + \varepsilon \bar{k}^{0,0}(\mathcal{B}; \varepsilon) + \varepsilon^{m+1} \bar{k}^1(\mathcal{B}, \alpha, s; \varepsilon). \tag{57}$$

The point that will be most important for us is that the first two terms of (57) correspond to a strongly integrable Hamiltonian, which moreover, satisfies twist conditions. The other term is extremely small. Hence, using quantitative versions of KAM theorem, we will conclude, in the following quantitative Theorem, that the non-resonant region contains KAM tori which are $O(\varepsilon^{(m+1)/2})$-closely spaced.

Theorem 38. *Let $\bar{k}$ be a Hamiltonian of the form* (57)*, where ε is sufficiently small and fixed. Assume that*

i) *$\bar{k}$ is a $C^{s+\beta}$ function and $\bar{k}^{0,0}$ is a $C^{s+\beta+2}$ function of the variables B, α, s,and that $||k^{0,0}||_{C^{s+\beta+2}}, ||k^1||_{C^{s+\beta}}$ are bounded independently of ε.*

ii) *We have $s \geq 5$, $0 < \beta < 1$.*

Then we can find $C > 0$, depending only on the properties of the Hamiltonian but independent of ε such that

a) *For every interval $\mathcal{I}$ of length $C\varepsilon^{(m+1)/2}$ there is an invariant torus $\mathcal{T}$ for the Hamiltonian flow contained in $\mathcal{I} \times \mathbb{T}^1 \times \mathbb{T}^1$.*

b) *The torus $\mathcal{T}$ is the graph of a $C^{s-2+\beta}$ function Ψ from the angle variables to the action-variables*
$$\mathcal{T} = \{(B, \alpha, s) \in \mathcal{I} \times \mathbb{T}^2; B = \Psi(\alpha, s)\}.$$

c) *We have, for some constant a,*
$$||\Psi - a||_{\mathcal{C}^{s-2+\beta}} \leq C\varepsilon^{(m+1)/2}.$$

d) *The motion on the torus is $C^{s-4+\beta}$ conjugate to a rigid translation of frequency $(\omega, 1)$, where ω is a Diophantine number of constant type with Markov constant less that $\varepsilon^{(m+1)/2}$ (see Definition 42).*

Remark 39. It is important to note that, for a fixed value of ε, as long as we fix $m > 2$, it is enough to consider a finite number of tori $\mathcal{T}_i$ to ensure that all the points in the non-resonant region $\mathcal{S}^L$ are $O(\varepsilon^{1+\delta})$-close, $\delta > 0$, to an invariant torus. Of course, this number goes to infinity when ε decreases to zero. □

The main technical tool to establish Theorem 38 is a quantitative version of KAM theorem (Theorem 45) which relates the Diophantine constants of the frequencies of tori with the size of the perturbations under which they can survive. We will also need to discuss the spacings between numbers with appropriate Diophantine properties in Section 8.4.1.

Theorem 45 will be useful later in the proof of Theorem 56. Theorem 45 is just an adaptation to our notation of the theorem on [Her83, p. 198]. At the end of Section 8.4.2, we will show how Theorem 38 can be deduced from Theorem 45 and Lemma 43.

Note that, since we are using systems that are one-degree of freedom systems depending periodically on time, we can take time-2π maps and reduce the problem to the study of a twist map of an annulus. It is easy to see that the invariant circles for the twist map correspond exactly to invariant tori for the Hamiltonian system.

The result of [Her83, p. 198], besides improving the differentiability requirements from other theorems, has the advantage that it contains very good results on the dependence on the Diophantine constants, which will lead to close spacing among the circles.

Remark 40. The theorem in [Her83] has the disadvantages that it does not generalize to higher dimensions and it requires that the rotation numbers are of constant type (see Definition 42), hence it cannot establish that the tori cover a set of positive measure. For our purposes, however, measure does not play any role. □

Remark 41. Since there are many versions of KAM theorem, differing in the differentiability required or smallness conditions, it is worth point out that optimizing these possibilities is not very crucial for us.

The smallness condition can always be adjusted by making more averaging transformations, which also amounts to require more differentiability. □

8.4.1. *Some results on Diophantine approximation.* In this section we recall some results on Diophantine approximations, notably we discuss how separated are the numbers with good Diophantine constants. When incorporated to a KAM theorem this will lead to estimates on the gaps between the KAM tori. Later on, it will lead to estimates on how close to a resonance we can get tori.

We start by recalling the standard definition:

Definition 42. *We say that a real number $\omega \in \mathcal{D}(\kappa, \nu)$ when we have:*

$$|\omega - \ell/k| \geq \kappa\, |k|^{-2-\nu} \quad \forall (k, \ell) \in \mathbb{Z}^2,\ k \neq 0. \tag{58}$$

We will refer to $\mathcal{D}(\kappa, \nu)$ as the Diophantine numbers of type κ, ν. When $\nu = 0$, we will refer to $\mathcal{D}(\kappa, 0)$ as constant type numbers of Markov constant κ.

We will denote $\mathcal{D}(\nu) = \cup_{\kappa>0}\mathcal{D}(\kappa, \nu)$ and refer to them as Diophantine numbers of type ν. When $\nu = 0$ we will call them constant type numbers.

The Definition 42 is the same as that used in [Her83, p. 158]. We note that the inequality (58) is equivalent to

$$|\omega k - \ell|^{-1} \leq C\kappa^{-1}\, |(k, l)|^{1+\nu} \quad \forall (k, \ell) \in \mathbb{Z}^2 \tag{59}$$

which is a form that generalizes to higher dimensions.

It is well known that $\mathcal{D}(\nu)$ is of full measure when $\nu > 0$. Constant type numbers are of zero measure but are dense. All quadratic irrationals are constant type numbers. More generally, it is equivalent to say that a number is constant type and to say that its continued fraction expansion is bounded. See Proposition 44 later.

The following statement makes it precise the idea that if κ is small, the numbers $\mathcal{D}(\kappa, \nu)$, $\nu \geq 0$, do not have big gaps among them. The result that we will use later is the case for $\nu = 0$.

Lemma 43. *Given $\nu \geq 0$, there exists a constant $K(\nu) > 0$ such that in every interval $\mathcal{I}$ of diameter $K(\nu)\kappa$ there is a number in $\mathcal{D}(\kappa, \nu)$.*

Proof.

There is a standard proof for $\nu > 0$ which also gives information on the measure and also generalizes for higher dimensions. We do it first.

We note that

$$\mathcal{D}(\kappa, \nu) = \mathbb{R} \setminus \bigcup_{(a,b)\in\mathbb{Z}^2, b\neq 0} B_{\kappa b^{-2-\nu}}(a/b)$$

Hence, given an interval $\mathcal{I}$ and denoting by $|\quad|$ the Lebesgue measure of a set, we have

$$
\begin{aligned}
|\mathcal{I} \cap \mathcal{D}(\kappa, \nu)| &\geq |\mathcal{I}| - 2 \sum_{(a,b)\in\mathbb{Z}^2, b\neq 0, B_{\kappa|b|^{-2-\nu}}(a/b)\cap\mathcal{I}\neq\emptyset} \kappa\,|b|^{-2-\nu} \\
&\geq |\mathcal{I}| - 2 \sum_{b\in\mathbb{Z}\setminus\{0\}} \left(|\mathcal{I}|\,\kappa\,|b|^{-1-\nu} + 2\kappa\,|b|^{-2-\nu} \right) \\
&\geq |\mathcal{I}|\,(1 - K_1\kappa) - \kappa K_2.
\end{aligned}
\tag{60}
$$

The second inequality in (60) follows from the observation that, once we fix b, the number of a such that $B_{\kappa b^{-2-\nu}}(a/b) \cap \mathcal{I} \neq \emptyset$ is bounded by $|\mathcal{I}|\,|b| + 2$. In the last inequality of (60), K_1, K_2 stand for positive constants that depend on ν.

We see that for κ small enough and for $|\mathcal{I}| \geq K\kappa$, the right hand side of (60) is positive, hence, the measure $\mathcal{I} \cap \mathcal{D}(\kappa, \nu)$ is positive, which establishes the claim.

The above derivation uses essentially that $\nu > 0$ since we use that $\sum_b b^{-1-\nu}$ converges. Indeed, it is well known that $\mathcal{D}(\kappa, 0)$ is of measure 0 for all κ.

For the case when $\nu = 0$, we will use the theory of continued fractions. We recall the following proposition whose proof can be found in [Her79, p. 64].

Proposition 44. *Given $L \in \mathbb{R}^+$, denote by*

$$
\mathcal{C}(L) = \{x = [a_1, a_2, \ldots, a_n, \ldots] : \ a_i \in \mathbb{N},\ 1 \leq a_i \leq L\}. \tag{61}
$$

With this notation, we have

$$
\mathcal{C}(L) = \mathcal{D}(1/(L+2), 0).
$$

Hence, to finish the proof of Lemma 43, it suffices to show that $\mathcal{I}_L$, the largest interval in $\mathbb{R} \setminus \mathcal{C}(L)$, satisfies

$$
|\mathcal{I}_L| \leq K/L. \tag{62}
$$

We claim that this interval $\mathcal{I}_L$, which is the largest gap among numbers in $\mathcal{C}(L)$, is precisely:

$$
\mathcal{I}_L = ([2, L, 1, L, 1, L, \ldots], [1, 1, L, 1, L, 1, \ldots]) \tag{63}
$$

Once we have the claim (63), the result in Lemma 43 follows because a direct calculation shows that

$$
[2, L, 1, L, 1, \ldots] = 1/2 - 1/L + \mathrm{O}(1/L^2)
$$

and

$$
[1, 1, L, 1, L, 1, \ldots] = 1/2 + 1/L + \mathrm{O}(1/L^2).
$$

Therefore $|\mathcal{I}_L| = 2/L + \mathrm{O}(1/L^2)$. Hence, using Proposition 44, the largest gap in $\mathcal{D}(1/(L+2), 0)$ is $2/L + \mathrm{O}(1/L^2)$. That is, we have established Lemma 43 for $\kappa = 1/(L+2)$, from which the general result follows.

Hence, the only thing left for the proof of Lemma 43 is to establish (63).

We recall that given two numbers

$$x = [a_1, a_2, \ldots, a_m, b_1, b_2, \ldots],$$
$$y = [a_1, a_2, \ldots, a_m, c_1, c_2, \ldots],$$

and $b_1 \neq c_1$, we have that $x > y$ if $b_1 > c_1$ and m is odd or if $b_1 < c_1$ and m is even.

The above observation allows to conclude immediately that the number $[2, L, 1, L, 1, L, \ldots]$ is the largest number in $\mathcal{C}(L)$ whose first entry in the continued fraction is 2. Similarly, the number $[1, 1, L, 1, L, 1, \ldots]$ is the smallest number in $\mathcal{C}(L)$ whose first entry is 1. This makes it clear that there are no points of $\mathcal{C}(L)$ inside $\mathcal{I}_L$.

The claim (63) follows from the following considerations.

(1) If the interval

$$\mathcal{I} = ([a_1, b_2, b_3, b_4, \ldots], [a_1, c_2, c_3, c_4, \ldots])$$

does not contain any point in $\mathcal{C}(L)$, then, neither does the interval

$$\hat{\mathcal{I}} = ([c_2, c_3, c_4, \ldots], [b_2, b_3, b_4, \ldots]).$$

Obviously, the interval $\hat{\mathcal{I}}$ is larger than the interval $\mathcal{I}$. Hence we conclude that the interval $\mathcal{I}_L$ has to have different first entries.

(2) If $M, N \in \mathbb{N}$ and $N > 1$, we have:

$$[M+1, b_2, b_3, b_4, \ldots] \in ([M+N, b_2, b_3, b_4, \ldots], [M, c_2, c_3, c_4, \ldots]).$$

Hence, $\mathcal{I}_L$ should be of the form

$$\mathcal{I}_L = ([M+1, b_2, b_3, b_4, \ldots], [M, c_2, c_3, c_4, \ldots]).$$

(3) The length of the interval $([2, b_2, b_3, b_4, \ldots], [1, c_2, c_3, c_4, \ldots])$ is larger than that of the interval $([M+1, b_2, b_3, b_4, \ldots], [M, c_2, c_3, c_4, \ldots])$ for $M > 1$.

Hence, $\mathcal{I}_L$ should be of the form

$$\mathcal{I}_L = ([2, b_2, b_3, b_4, \ldots], [1, c_2, c_3, c_4, \ldots]).$$

(4) If an interval of the form $([2, b_2, b_3, b_4, \ldots], [1, c_2, c_3, c_4, \ldots])$ does not contain any point in $\mathcal{C}(L)$, then, it is necessary that $b_2, b_3, \ldots$ are chosen in such a way that $[2, b_2, b_3, b_4, \ldots]$ is the

largest number in $\mathcal{C}(L)$ whose first entry in the continued fraction expansion is 2. Similarly, we also need that $c_2, c_3, c_4, \ldots$ are chosen in such a way that $[1, c_2, c_3, c_4, \ldots]$ is the smallest number in $\mathcal{C}(L)$ whose first entry in the continued fraction expansion is 1.

Hence, we conclude that $\mathcal{I}_L$ is of the form given in (63).

□

8.4.2. *The KAM Theorem for twist maps.* The following result is an easy consequence of the Theorem 5.4 stated in [Her83, p. 198] (see also the Theorem 5.6 in [Her83, p. 204]).

Theorem 45. *Let $F_0 : \mathbb{T} \times \mathbb{R}$ be an integrable symplectic mapping, that is:*

$$F_0(\theta, r) = (\theta + \Delta(r), r). \tag{64}$$

Assume that $F_0 \in C^{n+\beta}$, $n \geq 3$, $0 < \beta < 1$ and $\frac{d}{dr}\Delta(r) \geq M > 0$.

Then we can find a constant K depending only on n, β such that for any r_0 such that $\omega \equiv \Delta(r_0) \in \mathcal{D}(\kappa, 0)$, and for any F exact symplectic $C^{n+\beta}$ map of $\mathbb{T} \times \mathbb{R}$ verifying

i) $||F - F_0||_{C^{n+\beta}} \leq \delta$
ii) $\delta \leq K\kappa(\omega)^2$,

there exists an invariant circle $\mathcal{T}$ such that

a) *$\mathcal{T}$ is the graph of a function $\Psi : \mathbb{T} \to \mathbb{R}$*
$$\mathcal{T} = \{(\theta, \Psi(\theta) : \theta \in \mathbb{T}\}.$$
b) $||\Psi - r_0||_{C^{n-1+\beta}} \leq KM^{-1}\kappa^{-1}\delta.$
c) *The motion of F restricted to $\mathcal{T}$ has rotation number ω.*
d) *If we denote by g the map of the torus defined by*
$$F(\theta, \Psi(\theta)) = (g(\theta), \Psi(g(\theta))$$
and we denote by h the map that conjugates g to a rotation by ω, i.e.
$$g \circ h(\theta) = h(\theta + \omega)$$
normalized to $h(0) = 0$, we have
$$||h - Id||_{C^{n-2+\beta}} \leq KM^{-1}\kappa^{-1}\delta. \tag{65}$$

Moreover, if $n \geq 4$ we have for all the F in a $C^{n+\beta}$ neighborhood of F_0:

e) *The mappings that associate F to Ψ, h respectively, are Lipschitz when we give F the $C^{n+\beta}$ topology, Ψ the $C^{n-2+\beta}$ topology and h the $C^{n-3+\beta}$ topology.*

f) *Define* $\Gamma_F(\tilde{\Psi})$ *the graph transform of* $\tilde{\Psi}$ *by:*

$$F(\mathrm{Graph}(\tilde{\Psi})) = \mathrm{Graph}(\Gamma_F(\tilde{\Psi})).$$

Assume that for some $C^{n-1+\beta}$ *map we have*

$$||\tilde{\Psi} - \Gamma_F(\tilde{\Psi})||_{C^{n-1+\beta}} \leq \mu$$

and that $\tilde{\Psi}$ *is in a* $K\kappa^2$ *neighborhood of a constant.*

Then, there is a $C^{n-1+\beta}$ *function* Ψ^* *whose graph is an invariant circle for* F *such that*

$$||\tilde{\Psi} - \Psi^*||_{C^{n-1+\beta}} \leq K\kappa^{-1}\mu.$$

Remark 46. For the experts in KAM theory, we call attention to the fact that Theorem 45 allows to conclude more regularity for the graph than for the conjugating function. Most of the versions of the KAM theorem study the conjugating function, hence, the regularity established for the curve is the same as that for the conjugating function. (Some theorems that establish that the graphs are more regular than the conjugacy are [Sal86], [Pös82].) □

Proof.

We refer to [Her83] for the proof of the result. Here we only explain how the statement we have made follows from the statement in [Her83].

We note that the Theorem 5.4 of [Her83] is stated as a translated curve theorem for maps (not necessarily symplectic) of the form

$$F(\theta, r) = (\theta + r + \alpha, r + \varphi(\theta, r)). \tag{66}$$

The proof of the translated curve result in [Her83] does not use any geometric feature (e.g. that the mapping is exact symplectic) of the map.

Note that a twist mapping can be always put into the form (66) by a change of variables (possibly non-canonical) applying the implicit function theorem to define r by the first component of (66). The map of the form (66) obtained through this procedure, has the same regularity of the original map. This change of variables does not change the regularity of the invariant circles or of the conjugacy of the motion on them to rotations.

Observe also that an exact symplectic mapping has the intersection property (the image of a nontrivial circle by the map has to intersect itself), and the intersection property of a map is preserved under any continuous change of variables. For any map with the intersection property, a translated curve has to be invariant.

The proof of Theorem 45 presented in [Her83] is extremely simple since it is not based on a rapidly convergent Nash-Moser method but rather on a Schauder-Tychonoff fixed point Theorem (when $n \geq 3$) or on a contraction mapping principle (when $n \geq 4$).

The fact that the proof uses a contraction theorem for an appropriate map is the reason why one has properties e) and f). Even if we will not use it here (and, hence not state the results) we note that using the results on composition in [LO99], it is possible to show that the operator considered in [Her83] is locally C^ℓ when the F and the Ψ are given in appropriate topologies (and n is large enough) so that one gets that the invariant tori depend smoothly on F in appropriate topologies. This allows to justify formal expansions in parameters.

□

Proof of Theorem 38. Note that once that we have Theorem 45, Theorem 38 follows immediately if we consider the time-2π map of the flow generated by the Hamiltonian (57). Note that if $\bar{k}(\mathcal{B}, \alpha, s; \varepsilon)$ is $C^{s+\beta}$, the time-2π map is $C^{s-1+\beta}$. Moreover it is an exact symplectic twist map.

The time-2π map corresponding to the integrable part of the Hamiltonian in (57) gives rise to a $C^{s+1+\beta}$ strongly integrable symplectic map. Given the form of the integrable part, the mapping will satisfy the twist condition.

By the standard dependence on parameters of differential equations, we obtain that the $C^{s-1+\beta}$ norms of the difference of the time-2π maps can be bounded by ε^{m+1}.

Then, if we apply Theorem 45 to constant type numbers of Markov constant $\kappa = K\varepsilon^{(m+1)/2}$ with $n = s - 1$, and take into account Lemma 43 to control the spacing between such numbers, we obtain the statement of Theorem 38. □

By Remark 37, applying Theorem 38 to the averaged Hamiltonian (57), when $r-2m-2 \geq 6$, and going back to the variables (J, φ, s) using the change given by Theorem 35, we obtain the following result about the existence of invariant tori of Hamiltonian (45).

Proposition 47. *Assume $r \geq 2m + 8$. Then, for ε small enough, in any connected component of the non-resonant region $\mathcal{S}^L$ defined in* (56)*, there exists a finite set of values J_i such that:*

1) $\omega_i = J_i + \varepsilon \frac{\partial \bar{k}^{0,0}}{\partial J}(J_i, \varepsilon)$ *(see* (57)*) is a Diophantine number of constant type and Markov constant $K\varepsilon^{\frac{m+1}{2}}$.*
2) *There exists a torus $\mathcal{T}_i$, invariant by the flow of the Hamiltonian $k(J, \varphi, s; \varepsilon)$ given in* (45)*, such that:*

2.1) *The motion on the torus* $\mathcal{T}_i$ *is* $\mathcal{C}^1$*-conjugated to a rigid translation of frequencies* $(\omega_i, 1)$.

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

$$\mathcal{T}_i = \{(J, \varphi, s) \in \mathcal{S}^L,\ J = J_i + u_{\omega_i}(\varphi, s; \varepsilon)\},$$

where $u_{\omega_i}(\varphi, s; \varepsilon)$ *is a* $\mathcal{C}^{r-2m-4-\eta}$ *function, for any* $\eta > 0$, *and* $||u_{\omega_i}||_{\mathcal{C}^2} \leq \text{cte.}\,\varepsilon$.

3) *Denoting by*

$$\mathbf{B}(\mathcal{A}, \rho) = \{\tilde{x} \in \tilde{\Lambda}_\varepsilon :\ \operatorname{dist}(\mathcal{A}, \tilde{x}) \leq \rho\}, \tag{67}$$

for any $\mathcal{A} \subset \tilde{\Lambda}_\varepsilon$, *one has that*

$$\mathcal{S}^L \subset \bigcup_i \mathbf{B}(\mathcal{T}_i, K\varepsilon^{\frac{m+1}{2}}). \tag{68}$$

Remark 48. Proposition 47 gives the primary KAM tori $\mathcal{T}_i$ in the variables (J, φ, s), but we can obtain the tori in the original variables (I, φ, s) using the change given by Proposition 28, which is ε^2-close to the identity. The tori thus obtained are invariant for the flow (7) and are of the form

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

where the function U_{ω_i} verifies the same properties as u_{ω_i}. □

Remark 49. The importance of Proposition 47 is that in the non-resonant region $\mathcal{S}^L$ we can find primary KAM tori with extremely small gaps between them.

It is important to note that, for a fixed value of ε, it is enough to consider a finite number of tori $\mathcal{T}_i$ to ensure that the regions $\mathbf{B}(\mathcal{T}_i, K\varepsilon^{\frac{m+1}{2}})$ cover all the non-resonant region $\mathcal{S}^L$. Of course, this number goes to infinity when ε decreases to zero. □

8.5. **Analyzing the resonances.** In this section we study the invariant sets close to resonant regions.

The goal is that we can cover the whole region with invariant objects (either primary tori, secondary tori or periodic orbits with (un)stable manifolds) at a distance less than $O(\varepsilon^{3/2})$.

The case of resonances of order 3 and higher will be studied in Section 8.5.1. It will not be different from the non-resonant region and will be enough to apply KAM theorem 38 to obtain primary tori with the required gaps.

The case of resonances or order 1 and 2 is significantly more involved and it will be done in Section 8.5.2, Section 8.5.3.

8.5.1. *Resonances of order* 3 *and higher.* In this section we study the reduced Hamiltonian $k(J,\varphi,s;\varepsilon)$ given in (45), in the regions close to the resonances of order 3 or bigger. To this end, we fix $j \geq 3$ and define:

(69)
$$\mathcal{S}^{\mathcal{R}_j} = \bigcup_{-l_1/k_1 \in \mathcal{R}_j \setminus \mathcal{R}_1 \cup \cdots \cup \mathcal{R}_{j-1}} \{(J,\varphi,s) \in [-l_1/k_1 - L, -l_1/k_1 + L] \times \mathbb{T}^2\},$$

where L is the constant provided in Theorem 35.

The connected components of this region are the sets

$$\{(J,\varphi,s) \in [-l_1/k_1 - L, -l_1/k_1 + L] \times \mathbb{T}^2\}.$$

On them, after the averaging procedure given in Theorem 35, the Hamiltonian is a $\mathcal{C}^{r-2m-2}$ function of the form:

(70)
$$\frac{\mathcal{B}^2}{2} + \varepsilon \bar{k}^{0,0}(\mathcal{B};\varepsilon) + \varepsilon^j \left(U^{k_1,l_1}(k_1\alpha + l_1 s;\varepsilon) + \varepsilon^{m+1-j} \bar{k}^1(\mathcal{B},\alpha,s;\varepsilon) \right).$$

We can apply Theorem 38 to Hamiltonian (70) since it is of the form (57) with ε^j instead of ε^{m+1}. When we express the results in terms of the variables (J,φ,s) using the change given by Theorem 35, we obtain:

Proposition 50. *Assume $r \geq 2m+8$. Then for any $3 \leq j \leq m$, in any connected component of the resonant region $\mathcal{S}^{\mathcal{R}_j}$, there exists a finite set of values J_i such that:*

1) $\omega_i = J_i + \varepsilon \frac{\partial \bar{k}^{0,0}}{\partial J}(J_i,\varepsilon)$ *(see* (70)*) is a Diophantine number of constant type and Markov constant $K\varepsilon^{j/2}$.*
2) *There exists a torus $\mathcal{T}_i$ invariant by the flow of the Hamiltonian $k(J,\varphi,s;\varepsilon)$ given in* (45)*, such that:*
   2.1) *The motion on the torus $\mathcal{T}_i$ can be $\mathcal{C}^1$-conjugated to a rigid translation of frequencies $(\omega_i, 1)$.*
   2.2) *The torus $\mathcal{T}_i$ can be written as a graph of the variable J over the angle variables (φ, s):*

   $$\mathcal{T}_i = \{(J,\varphi,s) \in \mathcal{S}^{\mathcal{R}_j},\ J = J_i + u_{\omega_i}(\varphi,s;\varepsilon)\}$$

   where $u_{\omega_i}(\varphi,s;\varepsilon)$ is a $\mathcal{C}^{r-2m-4-\eta}$ function, for any $\eta > 0$, and $||u_{\omega_i}||_{\mathcal{C}^2} \leq \mathrm{cte.}\,\varepsilon$
3) *One has that*

$$\mathcal{S}^{\mathcal{R}_j} \subset \bigcup_i \mathbf{B}(\mathcal{T}_i, K\varepsilon^{j/2}) \tag{71}$$

where $\mathbf{B}(\mathcal{A},\delta)$ is defined in (67).

Remark 51. Proposition 50 gives the primary KAM tori $\mathcal{T}_i$ in the variables (J, φ, s), we can obtain the tori in the original variables (I, φ, s) using the change given by Proposition 28 which is ε^2-close to the identity. The tori thus obtained are of the form

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

where the function U_{ω_i} verifies the same properties as u_{ω_i}. □

8.5.2. *Preliminary analysis of resonances of order one or two.* In this section we study in more detail the resonant regions $\mathcal{S}^{\mathcal{R}_j}$, defined in (69), for $j = 1, 2$, where $\mathcal{R}_1$ and $\mathcal{R}_2$ are defined in (20), (21).

This is the place where the standard mechanism of [Arn64] breaks down. Hence, this section is a place where the mechanism of diffusion presented in this paper differs from that of [Arn64].

As it is well known, there are easy examples where the KAM tori in regions $\mathcal{S}^{\mathcal{R}_j}$ have gaps of order $\varepsilon^{j/2}$. When $j = 1, 2$, these gaps are bigger than or comparable to ε (which, as we will see later, is the size of the heteroclinic jumps). This is what has become known as the *large gap problem.*

What we will do in this and the next two sections is to show that, even if in the regions $\mathcal{S}^{\mathcal{R}_j}$ the primary KAM tori are rather scarce, we can find other geometric objects such as secondary KAM tori and stable and unstable manifolds of lower dimensional tori. These objects get rather close to the frontier of the gaps among the primary KAM tori.

From now on, we will work in one connected component of the domain $\mathcal{S}^{\mathcal{R}_j}$. More precisely, let us consider a resonance $-l_0/k_0 \in \mathcal{R}_j$, for $j = 1, 2$. In the component

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

the Hamiltonian $k(J, \varphi, s; \varepsilon)$ in (45) can be written in the averaged variables $(\mathcal{B}, \alpha, s)$ of Theorem 35 as:

$$\begin{aligned}\bar{k}(\mathcal{B}, \alpha, s; \varepsilon) = \frac{\mathcal{B}^2}{2} + \varepsilon \bar{k}^{0,0}(\mathcal{B}; \varepsilon) + \varepsilon^j U^{k_0, l_0}(k_0\alpha + l_0 s; \varepsilon) \\ + \varepsilon^{m+1} \bar{k}^1(\mathcal{B}, \alpha, s; \varepsilon),\end{aligned} \tag{72}$$

defined, if we take ε small enough, on

$$\bar{D} := \{(\mathcal{B}, \alpha, s) \in \mathbb{R} \times \mathbb{T}^2, \quad |\mathcal{B} + l_0/k_0| \le \bar{L}\},$$

where $\left|L - \bar{L}\right| \le \text{cte.}\, \varepsilon$, and where $\varepsilon \bar{k}^{0,0}$ is of class $\mathcal{C}^{r-2m}$, U^{k_0, l_0} is a trigonometric polynomial in (φ, s) and a polynomial in ε, and $\bar{k}^1$ is of class $\mathcal{C}^{r-2m-2}$ (see Remark 37).

Now, we can state explicitly the first part of the non-degeneracy conditions that constituted Hypothesis **H5** of Theorem 7:

H5' For any first or second order resonance, the function $U^{k_0,l_0}(\theta;0)$ has a global maximum which is non-degenerate.

Remark 52. By Theorem 35 applied to Hamiltonian (45) we have that, in the case of a first order resonance $-l_0/k_0 \in \mathcal{R}_1$

$$U^{k_0,l_0}(\theta;0) = \sum_{t=-N}^{N} \hat{k}^1_{tk_0,tl_0}(-l_0/k_0)e^{it\theta}$$

where $\hat{k}^1_{k,l}$ are the Fourier coefficients of $k_1(J,\varphi,s)$, the second term in the Taylor expansion with respect to ε, of $k(J,\varphi,s;\varepsilon)$, the reduced Hamiltonian, which can be computed using formula (47). In the case of a second order resonance one can also obtain, using proposition 30, explicit formulas for $U^{k_0,l_0}(\theta;0)$. Then, by Proposition 30, Hypothesis **H5'** can be checked by examining the term $h(p,q,I,\varphi,t;\varepsilon)$ in the original Hamiltonian (4).

The computations involved in these verifications are algebraic and quite explicit. In Section 13, we undertake the verification in a explicit example.

We note that the condition **H5'** is C^j, $j \geq 2$, open and dense in the space of polynomials $U^{k_0,l_0}(\theta,0)$.

A closer look to the calculation of the coefficients (see 47) shows that the coefficient $\hat{h}_{tk_0,tl_0}$ of the perturbation (9) enter linearly in the expression for $U^{k_0,l_0}(\theta,0)$. Hence, we conclude that, in the space of original Hamiltonians restricted to the manifold $\tilde{\Lambda}_\varepsilon$, the hypothesis **H5'** holds in a C^j $j \geq 2$ open and dense of Hamiltonians. $\square$

To study Hamiltonian (72) in the set $\bar{D}$, we consider the change of variables depending on time given by

$$b = k_0(\mathcal{B} + l_0/k_0), \quad \theta = k_0\alpha + l_0 s, \quad s = s. \tag{73}$$

The change (73) is not a true symplectic change of variables but it is conformally symplectic, hence the new system of differential equations verified by (b,θ,s) is also Hamiltonian, of Hamiltonian:

$$\bar{K}(b,\theta,s;\varepsilon) = \bar{K}^0(b;\varepsilon) + \varepsilon^j\bar{V}(\theta;\varepsilon) + \varepsilon^{m+1}\bar{K}^1(b,\theta,s;\varepsilon), \tag{74}$$

with

$$\begin{aligned} \bar{K}^0(b,\varepsilon) &= \frac{b^2}{2} + \varepsilon k_0^2 \bar{k}^{0,0}(-l_0/k_0 + b/k_0;\varepsilon), \\ \bar{V}(\theta;\varepsilon) &= k_0^2 U^{k_0,l_0}(\theta;\varepsilon), \\ \bar{K}^1(b,\theta,s;\varepsilon) &= k_0^2\bar{k}^1(-l_0/k_0 + b/k_0, \frac{\theta - l_0 s}{k_0}, s;\varepsilon). \end{aligned} \tag{75}$$

Note that $\bar{K}^0$ is of class $\mathcal{C}^{r-2m}$, $\bar{V}$ is analytic, and $\bar{K}^1$ is of class $\mathcal{C}^{r-2m-2}$.

Up to order ε^m, the Hamiltonian (74) is given by $\bar{K}_0(b;\varepsilon)+\varepsilon^j\bar{V}(\theta;\varepsilon)$ which is a one degree of freedom Hamiltonian. Moreover, it is close to a pendulum-like Hamiltonian

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

By hypothesis **H5'**, this Hamiltonian has a saddle at $(0,\theta_1)$, where θ_1 is the maximum of $U^{k_0,l_0}(\theta;0)$.

Using the implicit function theorem, one can see easily that the Hamiltonian $\bar{K}_0(b;\varepsilon) + \varepsilon^j\bar{V}(\theta;\varepsilon)$ given in (74) also has a saddle at $(b(\varepsilon),\theta_1(\varepsilon)) = (0,\theta_1)+O(\varepsilon)$. This saddle corresponds to a hyperbolic periodic orbit if we also include the s variable as the time coordinate. The function $b(\varepsilon)$ is of class $\mathcal{C}^{r-2m-1}$, and $\theta_1(\varepsilon)$ is analytic.

To analyze the behavior of this pendulum-like system, we will find it convenient to make the translation

$$y = b - b(\varepsilon), \qquad x = \theta - \theta_1(\varepsilon), \quad s = s, \tag{76}$$

obtaining the $\mathcal{C}^{r-2m-2}$ Hamiltonian

$$K(y, x, s; \varepsilon) = h^0(y; \varepsilon) + \varepsilon^j U(x; \varepsilon) + \varepsilon^{m+1} S(y, x, s; \varepsilon) \tag{77}$$

where

$$\begin{aligned} h^0(y;\varepsilon) &= \bar{K}^0(y + b(\varepsilon);\varepsilon) - \bar{K}^0(b(\varepsilon);\varepsilon) \\ U(x;\varepsilon) &= \bar{V}(x + \theta_1(\varepsilon);\varepsilon) - \bar{V}(\theta_1(\varepsilon);\varepsilon) \\ S(y,x;\varepsilon) &= \bar{K}^1(y + b(\varepsilon), x + \theta_1(\varepsilon), s;\varepsilon) \end{aligned} \tag{78}$$

where we have subtracted a constant term, the energy of the saddle, to normalize:

$$h^0(0;\varepsilon) = \frac{\partial h^0}{\partial y}(0;\varepsilon) = 0,\ U(0;\varepsilon) = \frac{\partial U}{\partial x}(0;\varepsilon) = 0,\ \frac{\partial^2 U}{\partial x^2}(0;\varepsilon) < 0, \tag{79}$$

and $x = 0$ is a global maximum of U.

Let us note for further reference that the expression of $h^0(y;\varepsilon)$ given in (78) shows that $h^0(y;\varepsilon)$ is a $\mathcal{C}^{r-2m-1}$ function verifying:

$$h^0(y;\varepsilon) = \frac{y^2}{2}\hat{h}(y;\varepsilon) = \frac{y^2}{2}(1 + \varepsilon\tilde{h}(y;\varepsilon)) \tag{80}$$

where $\varepsilon\tilde{h}$ is a $\mathcal{C}^{r-2m-2}$ function.

Notice that the Hamiltonian $K(y,x,s;\varepsilon)$ is a $2\pi k_0$-periodic in s function defined in the domain D_{k_0} given by:

$$D_{k_0} = \{(y,x,s) \in \mathbb{R} \times (\mathbb{R}/(2\pi k_0\mathbb{Z}))^2, \quad |y| \le k_0\bar{L}\}. \tag{81}$$

From now on, we work in the variables (y, x, s), and, when necessary, we will recover the original Hamiltonian by performing the changes of variables (73), (76).

The first important point is that, up to order ε^m, the Hamiltonian (77) is given by the one degree of freedom $\mathcal{C}^{r-2m-1}$ Hamiltonian:

$$K_0(y, x; \varepsilon) = h^0(y; \varepsilon) + \varepsilon^j U(x; \varepsilon), \tag{82}$$

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

Remark 53. Note that the $2\pi k_0$-periodic Hamiltonians (74) and (77) are, up to order ε^m, 2π periodic. This is the well known effect that is colloquially described as saying that the resonance has "k_0 eyes"(see Fig. 1). □

8.5.3. *Primary and secondary tori near the first and second order resonances.* As the Hamiltonian $K_0(y, x; \varepsilon)$ given in (82) is 2π periodic, the region D_{k_0} given in (81) can be seen as k_0 copies of the region

$$D = \{(y, x, s) \in \mathbb{R} \times \mathbb{T}^2, \quad |y| \leq k_0 \bar{L}\}. \tag{83}$$

where this Hamiltonian is well defined.

The region D—and then D_{k_0}—is filled by the energy surfaces of the Hamiltonian K_0, given in (82):

$$\mathcal{T}_E^0 = \{(y, x, s) \in [-k_0 \bar{L}, k_0 \bar{L}] \times \mathbb{T}^2 : K_0(y, x; \varepsilon) = E\}.$$

$\mathcal{T}_E^0$ will, of course, be invariant by the Hamiltonian flow of K_0.

The energy surfaces that correspond to values of $E > 0$, are two primary tori in the sense of Definition 2 in Section 2.1. These primary tori can be written as a graph of the variable y over the angular variables (x, s).

The energy surface corresponding to $E = 0$ consists of the saddle $(0, 0)$ and the homoclinic orbits to it. We will refer to $\mathcal{T}_0^0$ as the separatrix loop.

When $E < 0$ the invariant surfaces (which are contained inside the region bounded by the separatrix loop $\mathcal{T}_0^0$) are tori of different topology than the primary tori, since they are contractible to a point. They are secondary tori in the sense of Definition 2 and cover all the region inside the separatrix loop. When $E < 0$ is close to zero, the curve $\mathcal{T}_E^0$ will be very close to the separatrix loop. These circles that get very close to the separatrix loop will allow us to transverse the resonance.

The main point of this section will be to show that many of the invariant tori of $K_0(y, x; \varepsilon)$ (both primary and secondary) survive when we add the perturbation term $\varepsilon^{m+1} S(y, x, s; \varepsilon)$ in (77). We will pay special attention to getting some of these tori close—$O(\varepsilon^{3/2})$ will be enough for our purposes—to the separatrix loop of K_0. (See Fig. 1.)

To establish this, we will put the unperturbed Hamiltonian $K_0(y,x;\varepsilon)$ in action-angle variables and apply KAM theorem 45 to the time-$2\pi k_0$ map of the Hamiltonian (77).

As it is well known, systems with hyperbolic critical points and separatrices do not have global action-angle variables, because the angle variable is not defined in the separatrix, where the frequency of the motion becomes zero.

We will need to use different action-angle variables inside and outside the separatrix. The main difficulty arises because the changes of variables reducing to action-angle become singular as we approach the separatrix. Nevertheless, we will control how singular these changes become in terms of the distance to the separatrix loop. Hence, using that the perturbation S is small—it contains a factor ε^{m+1} and m is large—we will see that it remains small when written in the action angle-variables provided that we do not consider it too close to the separatrix. Then, we can apply the KAM theorem 38 and can obtain tori close to a high power of ε to the separatrix loop. For subsequent purposes getting tori which are at a distance $\varepsilon^{3/2}$ is enough.

Remark 54. Similar arguments for analytic systems were used in [Neĭ84]. In the analytic case considered there, the tori are exponentially close to the separatrix. Since our systems are only finitely differentiable we need to develop arguments. Since we can only perform a finite number of averaging transformations, we only obtain that there are invariant tori at a distance to the separatrix not larger than a power of the perturbation. □

Remark 55. We will find it convenient to consider three different regions in the domain D_{k_0} in which the behavior of the tori is very different.

To studying the tori in different regions in different ways is justified because of the fact that the tori have different quantitative properties.

Tori far away from the resonance can be considered extremely flat, whereas tori very close to the resonance seem to be bunched near the critical point. This can be understood by looking at the level sets of the Hamiltonian $H(y,x) = \frac{y^2}{2} + \cos x - 1$.

When $E >> 1$, we have that $\mathcal{T}_E$, $\mathcal{T}_{E+\delta}$ are graphs of functions which are at distance $O(\delta)$ at all the points. Nevertheless we see that $\mathcal{T}_\delta$ and $\mathcal{T}_0$ are two tori such that for $x = \pi/2$ the distance is $O(\delta)$, whereas at $x = 0$ the distance is $O(\sqrt{\delta})$.

For small E, the same situation persists. The tori $\mathcal{T}_E$ and $\mathcal{T}_{E+\delta}$ are at a distance $O(\delta)$ in the middle and at a distance $\frac{\delta}{\sqrt{E}}$ near the singularity $x = 0$.

It is clear that if we are considering the tori at distances $\delta = \varepsilon^{3/2}$ and are interested in effects of size ε, we have to distinguish the size of their energies. This will affect which terms in an expansion are dominant.

The choices we have made below are not the only possible ones, but the above considerations show that there are quantitatively different features in different regions. □

D_{k_0} will be divided in three regions. D_f is the region far from the separatrix, D_o close to the separatrices but out of the region bounded by the separatrix loop and D_{in} close to the separatrices but inside the separatrix loop.

The precise definitions we have found useful are:

$$D_f = \{(y,x,s) \in D_{k_0} :\ K_0(y,x;\varepsilon) = E,\ c_1\varepsilon^j \le E \le c_2\bar{L}\} \tag{84}$$
$$D_o = \{(y,x,s) \in D_{k_0} :\ K_0(y,x;\varepsilon) = F,\ c_3\varepsilon^\alpha \le F \le c_1\varepsilon^j\} \tag{85}$$
$$D_{in} = \{(y,x,s) \in D_{k_0} :\ K_0(y,x;\varepsilon) = G,\ -c_4\varepsilon^j \le G \le -c_3\varepsilon^\alpha\} \tag{86}$$

where α, for the time being, is arbitrary provided that $\alpha > j = 1, 2$, which is the order of the resonance, and c_i are suitable constants independent of ε. (In particular, $c_4 < c$, where $-c$ is the minimum of U, is taken small enough in order that D_{in} does not contain any other critical point of K_0.)

As we will see, α controls how close to the separatrices we claim to find tori. Roughly speaking, we can take α any number, provided that we are willing to perform enough steps of averaging (this amounts to take m big enough in (77), which can be done just by assuming that the original system is differentiable enough).

Even if we state Theorem 56 for arbitrarily $\alpha > j$ and m large enough with respect to α, we note that for the subsequent applications in this paper, it will suffice to take any value $\alpha > 1 + j/2$, $j = 1, 2$. Later, in Corollary 57, for the sake of definiteness, we will take $\alpha = 3/2 + j/2$, and $m = 26$.

The following Theorem 56 establishes the existence of primary tori in $D_f \cup D_o$ and secondary tori in D_{in}.

The three regions (84), (85) and (86) cover the resonant region except for a thin neighborhood of the separatrix given by

$$\{(y,x,s) \in D_{k_0} :\ |K_0(y,x;\varepsilon)| \le c_3\varepsilon^\alpha\}. \tag{87}$$

The region (87) not covered in the present analysis contains what in physical language is called the "chaotic sea". In Section 8.5.5 we will identify features, other than invariant tori, in (87) namely, periodic orbits and their (un)stable manifolds.

Recall that $K_0(y,x;\varepsilon)$ as given in (82) describes a pendulum. Since have assumed that we do not have any other critical points in the region we consider, all the orbits we will study are periodic orbits.

Since we will be applying KAM arguments, the values of the frequency $\omega(E)$ and its non-degeneracy properties will play an important role.

An orbit of K_0 of energy E has frequency

$$\omega(E) = \frac{2\pi}{T(E)}, \ T(E) = \int_{K_0^{-1}(E)} \frac{dx}{y}. \tag{88}$$

Theorem 56. *Consider the $\mathcal{C}^{r-2m-2}$ reduced Hamiltonian $K(y,x,s;\varepsilon)$ as in* (77)*, inside the region D_{k_0} given in* (81)*. Consider $\alpha > j$, for $j = 1, 2$, $m \geq \max(11j - 1, 14(\alpha - j) - 1 - j/2)$, and assume that $r \geq 2m + 8$. 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 $c_1\varepsilon^j \leq E_i \leq c_2\bar{L}$, such that:*
   - 1.1) *The frequencies $\omega(E_i)$ are Diophantine numbers of constant type and Markov constant $K\varepsilon^{\frac{m+1-6j}{2}}$ (see Definition 42).*
   - 1.2) *For any value E_i, there exist two primary invariant tori $\mathcal{T}^{\pm}_{E_i}$ of Hamiltonian* (77) *contained in D_f.*
   - 1.3) *The motion on the tori $\mathcal{T}^{\pm}_{E_i}$ is $\mathcal{C}^1$-conjugated to a rigid translation of frequencies $(\omega(E_i), 1)$.*
   - 1.4) *These tori can be written as:*

$$\mathcal{T}^{+}_{E_i} = \{(y,x,s) \in D_f, \ K_{E_i}(y,x,s;\varepsilon) = E_i, \ y > 0\},$$
$$\mathcal{T}^{-}_{E_i} = \{(y,x,s) \in D_f, \ K_{E_i}(y,x,s;\varepsilon) = E_i, \ y < 0\},$$

where $K_{E_i}(y,x,s;\varepsilon)$ is a $\mathcal{C}^{r-2m-5-\eta}$ function, for any $\eta > 0$, given by

$$K_{E_i}(y,x,s;\varepsilon) = K_0(y,x;\varepsilon) + \mathrm{O}_{\mathcal{C}^2}\left(\varepsilon^{\frac{m+1-11j}{2}}\right). \tag{89}$$

   - 1.5) *$D_f \subset \bigcup_i B\left(\mathcal{T}^{\pm}_{E_i}, \varepsilon^{\frac{m+1-7j}{2}}\right)$, where*

$$B\left(\mathcal{T}^{\pm}_{E}, \delta\right) = \{(y,x,s) \in D_{k_0}, \ |K_0(y,x;\varepsilon) - E| \leq \delta\}. \tag{90}$$

2) ***Primary tori close to resonance.*** *There exists a set of values $F_1 < \cdots < F_{l_F}$ verifying $c_3\varepsilon^{\alpha} \leq F_i \leq c_1\varepsilon^j$, such that:*

2.1) *The frequencies* $\omega(F_i)$ *are Diophantine numbers of constant type and Markov constant* $K\varepsilon^{\frac{m+1-6(\alpha-j)-j/2}{2}}$.
2.2) *For any value* F_i*, there exist two primary invariant tori* $\mathcal{T}_{F_i}^{\pm}$ *of Hamiltonian* (77) *contained in* D_o.
2.3) *The motion on the tori* $\mathcal{T}_{F_i}^{\pm}$ *is* $\mathcal{C}^1$*-conjugated to a rigid translation of frequencies* $(\omega(F_i), 1)$.
2.4) *These tori can be written as:*

$$\mathcal{T}_{F_i}^{+} = \{(y, x, s) \in D_o,\ K_{F_i}(y, x, s; \varepsilon) = F_i,\ y > 0\}$$
$$\mathcal{T}_{F_i}^{-} = \{(y, x, s) \in D_o,\ K_{F_i}(y, x, s; \varepsilon) = F_i,\ y < 0\}$$

where $K_{F_i}(y, x, s; \varepsilon)$ *is a* $\mathcal{C}^{r-2m-5-\eta}$ *function, for any* $\eta > 0$*, given by:*

$$K_{F_i}(y, x, s; \varepsilon) = K_0(y, x; \varepsilon) + \mathrm{O}_{\mathcal{C}^2}\left(\varepsilon^{\frac{m+1+j/2-14(\alpha-j)}{2}}\right). \tag{91}$$

2.5) $D_o \subset \bigcup_i B\left(\mathcal{T}_{F_i}^{\pm}, \varepsilon^{\frac{m+1+j/2-10(\alpha-j)}{2}}\right)$.

3) ***Secondary tori close to resonance.*** *There exists a set of values* $G_1 < \cdots < G_{l_G}$ *verifying* $-c_4\varepsilon^j \leq G_i \leq -c_3\varepsilon^\alpha$*, such that:*
3.1) *The frequencies* $\omega(G_i)$ *are Diophantine numbers of constant type and Markov constant* $K\varepsilon^{\frac{m+1-6(\alpha-j)-j/2}{2}}$.
3.2) *For any value* G_i*, there exists a secondary invariant torus* $\mathcal{T}_{G_i}$ *of Hamiltonian* (77) *contained in* D_{in}*, contractible to the set*

$$\{(0, a, s), a \in \mathbb{R},\ s \in \mathbb{R}/(2\pi k_0\mathbb{Z})\} \subset D_{in}.$$

3.3) *The motion on the torus* $\mathcal{T}_{G_i}$ *is* $\mathcal{C}^1$*-conjugated to a rigid translation of frequencies* $(\omega(G_i), 1)$.
3.4) *This torus can be written as:*

$$\mathcal{T}_{G_i} = \{(y, x, s) \in D_{in},\ K_{G_i}(y, x, s; \varepsilon) = G_i\}$$

where $K_{G_i}(y, x, s; \varepsilon)$ *is a* $\mathcal{C}^{r-2m-5-\eta}$ *function, for any* $\eta > 0$*, given by:*

$$K_{G_i}(y, x, s; \varepsilon) = K_0(y, x; \varepsilon) + \mathrm{O}_{\mathcal{C}^2}\left(\varepsilon^{\frac{m+1+j/2-14(\alpha-j)}{2}}\right). \tag{92}$$

3.5) $D_{in} \subset \bigcup_i B\left(\mathcal{T}_{G_i}, \varepsilon^{\frac{m+1+j/2-10(\alpha-j)}{2}}\right)$.

Theorem 56 gives the existence of invariant tori in D_{k_0}. In the following Corollary 57, we will make more explicit assertions about the proximity of these tori. We will also make precise assertions of their properties when expressed as graphs of functions from the angle variables to the action variable. The proof of Corollary 57 from Theorem 56 is just an easy application of the implicit function theorem. Nevertheless, we state it explicitly because the properties of the tori as graphs will be useful later in Section 9 when we study how the tori behave under the scattering map.

Corollary 57. *Consider the reduced Hamiltonian $K(y,x,s;\varepsilon)$ as in (77), inside the region D_{k_0} given in (81). Consider $\alpha = j/2+3/2$, for $j = 1,2$, and $m \geq 26$. Then, if $r \geq 2m+8$, the tori defined in Theorem 56 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):*

$$\mathcal{T}^{\pm}_{E_i} = \{(y,x,s) \in D_f,\ y = f^{\pm}_{E_i}(x,s;\varepsilon)\}.$$

(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):*

$$\mathcal{T}^{\pm}_{F_i} = \{(y,x,s) \in D_o,\ y = f^{\pm}_{F_i}(x,s;\varepsilon)\}.$$

(3) *There exists $\rho_0 > 0$ such that for any $0 < \rho_0 \leq \rho < \pi$, and for any value G_i, each of the components of*

$$\mathcal{T}_{G_i} \cap \{(y,x,s):\ x \in I_\rho\}, \quad I_\rho = \cup_{l=0}^{k_0-1}[2\pi l + \rho, 2\pi(l+1) - \rho]\},$$

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):

$$\mathcal{T}^{\pm,\rho}_{G_i} = \{(y,x,s) \in D_{in}, x \in I_\rho,\ y = f^{\pm}_{G_i}(x,s;\varepsilon)\}$$

(4) *There exists a constant K, independent of ε, such that*

$$\begin{aligned}
|E_i - E_{i+1}| &\leq K\varepsilon^{(m+1-7j)/2} \leq K\varepsilon^{\frac{3}{2}+\frac{j}{2}} \\
|F_i - F_{i+1}| &\leq K\varepsilon^{(m+1+j/2-10(\alpha-j))/2} \leq K\varepsilon^{\frac{3}{2}+\frac{j}{2}} \\
|G_i - G_{i+1}| &\leq K\varepsilon^{(m+1+j/2-10(\alpha-j))/2} \leq K\varepsilon^{\frac{3}{2}+\frac{j}{2}}
\end{aligned}$$

(5)

$$\begin{aligned}
|E_1 - F_{l_F}| &\leq K\varepsilon^{(m+1-7j)/2} \leq K\varepsilon^{\frac{3}{2}+\frac{j}{2}} \\
|F_1 - G_{l_G}| &\leq K\varepsilon^{\alpha} + \varepsilon^{(m+1+j/2-10(\alpha-j))/2} \leq K\varepsilon^{\frac{3}{2}+\frac{j}{2}}
\end{aligned}$$

(6) *All these functions $f_\upsilon = f^{\pm}_\upsilon$ are at least of class $\mathcal{C}^2$, and, denoting by D the derivatives with respect to x and s, for $\upsilon = E_i$, $i = 1,\ldots,l_E$, $\upsilon = F_i$, $i = 1,\ldots,l_F$, and $\upsilon = G_i$, $i = 1,\ldots,l_G$, they verify:*

(a) *There exists a function* $\mathcal{Y}(x, E)$, *given explicitly in* (97), *such that:*

$$\begin{aligned} |f_{E_i} - \mathcal{Y}(x, E_i)|_{\mathcal{C}^1} \leq & K\varepsilon^{(m+1-12j)/2} \leq K\varepsilon^{3/2} \\ |f_{F_i} - \mathcal{Y}(x, F_i)|_{\mathcal{C}^1} \leq & K\varepsilon^{(m+1-j/2-14(\alpha-j))/2} \leq K\varepsilon^{3/2} \\ |f_{G_i} - \mathcal{Y}(x, G_i)|_{\mathcal{C}^1} \leq & K\varepsilon^{(m+1-j/2-14(\alpha-j))/2} \leq K\varepsilon^{3/2} \end{aligned} \tag{93}$$

(b) $|Df_\upsilon| \leq K\varepsilon^{j/2}$, $\left|D^2 f_\upsilon\right| \leq K\varepsilon^{j/2}$.

(c) *For any two consecutive values* υ *and* $\bar{\upsilon}$ *we have:*

$$|f_\upsilon - f_{\bar{\upsilon}}|_{\mathcal{C}^1} \leq \frac{|\upsilon - \bar{\upsilon}|}{\varepsilon^{j/2}} \leq K\varepsilon^{3/2}. \tag{94}$$

Remark 58. If we go back to the original variables (I, φ, s) through the changes given by Proposition 28, Theorem 35 and the changes (73), and (76), we obtain that the tori inside the region $\mathcal{S}^{\mathcal{R}^j}$, $J = 1, 2$ are given by,

$$I = -k_0/l_0 + U_\upsilon^\pm(\varphi, s; \varepsilon)$$

where the functions $U_\upsilon^\pm(\varphi, s; \varepsilon)$ verify the same properties than the functions $f_\upsilon^\pm$. This is, of course, a reflection of the fact that the properties of proximity between the tori are geometric properties that are invariant under changes of variables. □

8.5.4. *Proof of Theorem 56 and Corollary 57.* As indicated, the main technique will be to estimate the singularities of the change to action angle variables in the curves that are close to the separatrix. As it is well known, these action angle variables are expressed as integrals of a form over the energy surfaces. The main source of problems is that the surfaces become singular as we approach the separatrix.

Remark 59. An exercise that the authors found useful is to go over the calculations of the action angle variables of the physical pendulum. In such a case, the integrals are explicitly known in terms of elliptic functions, so that the order of magnitude of the singularities can be checked. □

In the following Lemma 60, we want to fix ε sufficiently small, and $0 < \delta \leq 1$, and consider the foliation given by the level sets

$$h(y; \varepsilon) + \delta U(x; \varepsilon) = E, \tag{95}$$

and obtain that—excluding a small interval—we can consider them as graphs.

Furthermore, we want to obtain bounds on the singularities which are uniform in ε, δ.

The idea of the proof is very simple. We just need to realize that $h(y;\varepsilon) + \delta U(x;\varepsilon) \simeq \frac{y^2}{2} + \varepsilon^j U(x;\varepsilon)$, so that the main term of the solution of (95) is

$$y = \pm\sqrt{2(E - \delta U(x;\varepsilon))}. \tag{96}$$

Indeed, if we express the y in (95) as a function of (96) equation (95) becomes an equation which can be dealt by the implicit function theorem. Note that the energy levels of the pendulum, when expressed as a graph, have square root singularities at $E = 0$.

What we want to show is that, if we add extra terms to the Hamiltonian, the level sets can be still expressed as graphs and that the singularities are still the square root singularities. This will be important for us later since the singularities of the graph control the singularities of the action-angle variables.

As it is standard in dynamical systems, when dealing with perturbations of systems that involve singularities, it is more convenient to formulate the problem in a system of coordinates which incorporates the singularities. Hence we will express the results as a function of (96) so that the resulting functions are smooth.

Lemma 60. *Let $U(x;\varepsilon)$ and $h(y;\varepsilon)$ be two $\mathcal{C}^n$ functions, $n = r - 2m - 1$, defined in $(y, x) \in \mathbb{R} \times \mathbb{T}$. Assume:*

(1) *$U(x;\varepsilon)$ has a non-degenerate global maximum at the origin.*
(2) *$U(x;\varepsilon)$ verifies $-c \le U(x;\varepsilon) \le 0$, and $U(x;\varepsilon) = -ax^2 + O(x^3)$, as $x \to 0$.*
(3) *$h(y;\varepsilon)$ is of the form (80).*

Let $\delta \in [-\delta_0, \delta_0]$ and consider the equation (95). Then

(1) *For $|\varepsilon|$ small enough, and for $-c\delta \le E \le M$, where M is independent of ε and δ, equation (95) defines two functions $y = \mathcal{Y}_\pm(x, E)$ on $\mathcal{I} = \mathcal{I}(\varepsilon, \delta) := \{(x, E),\ x \in \mathbb{T},\ E - \delta U(x;\varepsilon) \ge 0\}$ such that:*

$$\mathcal{Y}_\pm(x, E) = \pm(1 + \varepsilon b)\ell(x, E) + \varepsilon\tilde{\mathcal{Y}}_\pm(\ell(x, E)), \tag{97}$$

where $\ell(x, E) = \sqrt{2(E - \delta U(x;\varepsilon))}$, and:

(a) *b is independent of δ, and $\tilde{\mathcal{Y}}_\pm(0) = \tilde{\mathcal{Y}}'_\pm(0) = 0$.*
(b) *$\varepsilon\tilde{\mathcal{Y}}_\pm$ is a $\mathcal{C}^n$ function and*

$$\begin{aligned} &\|\varepsilon\tilde{\mathcal{Y}}_\pm \circ \ell\|_{\mathcal{C}^s(\mathcal{I}_{E_0})} \le \text{cte.}\,\varepsilon, \quad s = 0, 1, \\ &\|\varepsilon\tilde{\mathcal{Y}}_\pm \circ \ell\|_{\mathcal{C}^s(\mathcal{I}_{E_0})} \le \frac{\text{cte.}\,\varepsilon}{E_0^{s-3/2}}, \quad 2 \le s \le n. \end{aligned} \tag{98}$$

where $\mathcal{I}_{E_0} := \{(x, E),\ x \in \mathbb{T},\ E \ge E_0 > 0\}$

(2) *There exists ρ independent of ε and δ, such that for $-\rho \leq x \leq \rho$, equation (95) defines two functions $x = \mathcal{X}_\pm(y, E)$ on $\mathcal{J} = \mathcal{J}(\varepsilon, \delta) := \{(y, E),\ U(\rho)\delta < E < 0,\ y \in \mathbb{R},\ h(y; \varepsilon) - E < -\delta U(\rho; \varepsilon)\}$ such that:*

$$x = \mathcal{X}_\pm(y; E) = m(y, E) + \tilde{\mathcal{X}}_\pm(m(y, E)), \tag{99}$$

where $m(y, E) = \sqrt{\frac{1}{a\delta}(h(y; \varepsilon) - E)}$, and:

(a) $\tilde{\mathcal{X}}_\pm(0) = \tilde{\mathcal{X}}'_\pm(0) = 0$.

(b) $\tilde{\mathcal{X}}_\pm(t)$ *is a $\mathcal{C}^n$ function and*

$$||\tilde{\mathcal{X}}_\pm \circ m||_{\mathcal{C}^s(\mathcal{J}_{E_0})} \leq \frac{\text{cte.}}{|E_0|^{s-3/2}}, \quad 2 \leq s \leq n, \tag{100}$$

where $\mathcal{J}_{E_0} := \{(y, E),\ U(\rho)\delta < E \leq E_0 < 0,\ y \in \mathbb{R},\ h(y; \varepsilon) - E < -\delta U(\rho; \varepsilon)\}$, and its $\mathcal{C}^0$ and $\mathcal{C}^1$ norms are bounded independently of δ and ε.

Proof. By (80) equation (95) is equivalent to

$$\frac{y^2}{2}(1 + \varepsilon \tilde{h}(y; \varepsilon)) = \frac{t^2}{2},$$

where $t = \pm\sqrt{2(E - \delta U(x; \varepsilon))} := \ell(x; E)$. Writing $y = tz$, equation becomes

$$\frac{z^2}{2} + \varepsilon z^2 \tilde{h}(tz; \varepsilon) = 1.$$

$z = 1$ is a solution when $\varepsilon = 0$. The implicit function theorem gives us the existence of a solution of the form

$$z = 1 + \varepsilon g(t; \varepsilon) = 1 + \varepsilon b(\varepsilon) + \varepsilon \tilde{g}(t; \varepsilon),$$

where $b(\varepsilon) = g(0; \varepsilon)$, so that $\varepsilon \tilde{g}$ is $\mathcal{C}^n$ function with a bounded $\mathcal{C}^n$ norm verifying $\tilde{g}(0, \varepsilon) = 0$. Therefore, the solution of equation (95) is given by

$$y = \pm\ell(x, E)(1 + \varepsilon b) + \varepsilon \ell(x, E)\tilde{g}(\ell(x, E); \varepsilon).$$

Writing $\tilde{y}(t) = t\tilde{g}(t; \varepsilon)$ we obtain (97). Bounds (98) follow from Faa-di-Bruno formulae.

The proof of part 2. of the Lemma is analogous.

□

The next Proposition 61 studies the action-angle variables (A, ψ) associated to the Hamiltonian $K_0(y, x; \varepsilon)$ in the domain D_f defined in (84).

In order to simplify the notation, we note that K_0 is independent of s and that it is 2π-periodic in x rather than $2k_0\pi$.

Therefore, we will consider the domain

$$D_f^* = \{(y, x) \in \mathbb{R} \times \mathbb{T},\ \exists l \in \{0, 1, , \cdots, k_0\},\quad (y, x + 2\pi l, s) \in D_f\}$$

in the variables y, x, which is 2π-periodic.

Roughly speaking, what we do is to restrict the action-angle variables in one of the "eyes of the resonance". The results obtained in one eye extend to the other eyes by the 2π-periodicity in x of K_0, and clearly are uniform for all s since s does not enter in the Hamiltonian.

Proposition 61. *Consider a Hamiltonian $K_0(y, x; \varepsilon)$ as in (82) of class $\mathcal{C}^{r-2m-1}$, with h^0 verifying (80), in the region D_f^*. Then, there exists a $\mathcal{C}^{r-2m-2}$ change of variables*

$$\chi_f : \tilde{D}_f \rightarrow D_f^* \tag{101}$$

$$(A, \psi) \mapsto (y, x) \tag{102}$$

with $\tilde{D}_f = \{(A; \psi) :\ \tilde{c}_1 \varepsilon^{j/2} \le A \le \tilde{c}_2 L^{j/2},\ \psi \in \mathbb{T}\}$, where $\tilde{c}_i$, $i = 1, 2$ are constants independent of ε, such that:

(1) $K_0 \circ \chi_f(A, \psi) = \mathcal{G}_f(A; \varepsilon)$.

(2) $||\chi_f||_{\mathcal{C}^s(\tilde{D}_f)} \le \frac{M_f}{\varepsilon^{sj}}$, $\quad ||\chi_f^{-1}||_{\mathcal{C}^s(D_f^*)} \le \frac{M_f}{\varepsilon^{sj}}$, $0 \le s \le r - 2m - 2$.

(3) $||\mathcal{G}_f||_{\mathcal{C}^3(\tilde{D}_f)} \le \frac{M_f}{\varepsilon^{j/2}}$ *and* $\left|\mathcal{G}_f''(A; \varepsilon)\right| \ge M_f$.

where M_f is a constant independent of ε.

Proof. In D_f^* we consider the curves $E = K_0(y, x; \varepsilon)$, with $E \ge c_1 \varepsilon^j$. Then the action variable is given by the well known formula:

$$A = \frac{S(E)}{2\pi} = \frac{S(E; \varepsilon)}{2\pi} = \frac{1}{2\pi} \oint_{K_0^{-1}(E)} y\, dx,$$

and ψ is the conjugate angle. The new Hamiltonian will be $\mathcal{G}_f(A; \varepsilon) = S^{-1}(A; \varepsilon)$.

To avoid an unpleasant typography, from now on in this section, we will omit the dependence on ε of many of the functions which appear during the proof.

Following standard practice in mechanics we find it useful to use variables (x, E) rather than (y, x). (Note that we are interested in the curves $E = K_0(y, x; \varepsilon)$.) We denote by Γ the map that to (y, x) associates (x, E). The mapping Γ satisfies $\Gamma(D_f^*) \subset \mathcal{J}_f$, where

$$\mathcal{J}_f = \{(x, E),\ x \in \mathbb{T},\ c_1 \varepsilon^j \le E \le c_2 L\}. \tag{103}$$

We also note that the mapping Γ is locally invertible.

Using part 1. of Lemma 60 with $\delta = \varepsilon^j$, and $n = r - 2m - 1$, we define in D_f^*:

$$\begin{aligned} E &= K_0(y, x; \varepsilon), \\ \tau(x, E) &= \int_0^x \frac{1}{\frac{\partial K_0}{\partial y}(\mathcal{Y}_\pm(u, E), u; \varepsilon)} du. \end{aligned}$$

where E is the energy of the orbit and τ is the time along the orbit of energy E (we have chosen the origin of time at $x = 0$). With this choice, we have that $T(E) = \tau(2\pi, E)$ is the period of the periodic orbit, and the action-angle variables are:

$$\begin{aligned} A &= \frac{S(E)}{2\pi} = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{Y}_\pm(x, E) dx, \\ \psi &= \frac{2\pi}{T(E)} \tau(x, E), \end{aligned} \tag{104}$$

where $S'(E) = T(E)$.

Although region D_f^* has two connected components, we will do all the details only when $y > 0$.

First, using implicit derivatives in equation (95) and formula (97), we have:

$$\begin{aligned} \tau(x, E) &= \int_0^x \frac{1}{\frac{\partial K_0}{\partial y}(\mathcal{Y}_+(u, E), u; \varepsilon)} du = \int_0^x \frac{\partial \mathcal{Y}_+}{\partial E}(u, E) du \\ &= \frac{1}{\sqrt{2}} \int_0^x \frac{1 + \varepsilon b}{\sqrt{E - \varepsilon^j U(u; \varepsilon)}} du + \varepsilon P_1(x, E), \end{aligned} \tag{105}$$

where the function $P_1(x, E)$ is given by $P_1(x, E) = \int_0^x \frac{\partial}{\partial E} \tilde{\mathcal{Y}}_+(\ell(u, E)) du$. Taking into account that in D_f^*, and so in $\mathcal{J}_f$ (see (103)), one has:

$$c_1 \varepsilon^j \le E \le E - \varepsilon^j U(x; \varepsilon) \le E + c\varepsilon^j \le \text{cte.}\, E, \tag{106}$$

bound (98) give us that

$$|P_1|_{\mathcal{C}^0(\mathcal{J}_f)} \le \text{cte.}\,, \ |P_1|_{\mathcal{C}^s(\mathcal{J}_f)} \le \frac{\text{cte.}}{\varepsilon^{j(s-1/2)}}, \ 1 \le s \le r - 2m - 2. \tag{107}$$

Differentiating (105) under the integral sign, and using (106) and (107), we obtain upper bounds for $\tau(x, E)$:

$$|\tau|_{\mathcal{C}^s(\mathcal{J}_f)} \le \frac{\text{cte.}}{\varepsilon^{j(s+1/2)}}, \ 0 \le s \le r - 2m - 2. \tag{108}$$

On the other hand, using that $\tau(2\pi, E) = S'(E)$, one obtains:

$$|S|_{\mathcal{C}^0(\mathcal{J}_f)} \le \text{cte.}\,, \ |S|_{\mathcal{C}^s(\mathcal{J}_f)} \le \frac{\text{cte.}}{\varepsilon^{j(s-1/2)}}, \ 1 \le s \le r - 2m - 2. \tag{109}$$

The next task is to obtain lower bounds for the first and second derivatives of S. We start by observing that the main term of $\tau(x, E)$

appears explicitly in (105), and we have developed in (107) bounds for the remainder. Using (106) we can obtain easily lower bounds for the main term.

$$\left|S'(E)\right| = |T(E)| \geq \frac{\text{cte.}}{E^{1/2}}, \tag{110}$$

$$\left|S''(E)\right| \geq \frac{\text{cte.}}{E^{3/2}}. \tag{111}$$

Once we have estimated the function $T(E)$, we can bound the conjugate angle ψ given in (104), where $\tau(x, E)$ is given by (105). First, using (109), (110) and Faa-di-Bruno formula for the derivative of the composition of functions, we obtain

$$\left|\frac{\partial^s}{\partial E^s}\left(\frac{1}{T(E)}\right)\right| \leq \frac{\text{cte.}}{E^{s-1/2}} \leq \frac{\text{cte.}}{\varepsilon^{j(s-1/2)}},$$

and then the Leibniz rule gives

$$\left|\frac{1}{T}P_1\right|_{\mathcal{C}^0(\mathcal{J}_f)} \leq \text{cte.}\,, \quad \left|\frac{1}{T}P_1\right|_{\mathcal{C}^s(\mathcal{J}_f)} \leq \frac{\text{cte.}}{\varepsilon^{j(s-1)}}, \ 1 \leq s \leq r - 2m - 2.$$

This gives, using again (106):

$$\left|\frac{1}{T}\tau\right|_{\mathcal{C}^s(\mathcal{J}_f)} \leq \frac{\text{cte.}}{\varepsilon^{js}}, \ 0 \leq s \leq r - 2m - 2. \tag{112}$$

Bound (112) together with (109) gives the upper bound claimed in item 2. of the Proposition for the $\mathcal{C}^s$ norm of the change $\chi_f^{(-1)}(y, x) = (\frac{S(E)}{2\pi}, \frac{2\pi}{T(E)}\tau(x, E))$, with $E = K_0(y, x; \varepsilon)$, for $0 \leq s \leq r - 2m - 2$ in D_f^*.

Since $\det D\chi_f^{(-1)}(y, x) = 1$, the $\mathcal{C}^s$ norm of χ_f satisfies the same bounds.

Moreover, taking into account that $\mathcal{G}_f = S^{-1}$, the lower bound for the second derivatives of $\mathcal{G}_f$ and the upper bound for the third derivative of $\mathcal{G}_f$ follow from (110-111). □

Now that once we have expressed K_0, given in (82), in action-angle variables in D_f^*, the proof of Theorem 56 will consist in applying Theorem 45 to the time-$2\pi k_0$ map of the full Hamiltonian (77), to show that it has primary invariant tori. Going back to the original variables (y, x, s), one obtains the result claimed in Theorem 56. We proceed to give the details.

Proof of part 1) of Theorem 56. The proof will consist in applying Theorem 45 to F, the time-$2\pi k_0$ map of the Hamiltonian

$$\tilde{K}(A, \psi, s; \varepsilon) = \mathcal{G}_f(A; \varepsilon) + \varepsilon^{m+1}\tilde{S}(A, \psi, s; \varepsilon), \tag{113}$$

where $\tilde{K}(A,\psi,s;\varepsilon) = K \circ \chi_f(A,\psi,s;\varepsilon)$, and $\tilde{S}(A,\psi,s;\varepsilon) = S \circ \chi_f(A,\psi,s;\varepsilon)$.

Since $\tilde{K}$, $\mathcal{G}$, $\tilde{S}$ are $\mathcal{C}^{r-2m-2}$, and we have assumed in Theorem 56 that $r-2m-2 \geq 6$, we have that $\tilde{K}$, $\mathcal{G}$, $\tilde{S}$ are $\mathcal{C}^6$.

We denote by F_0 the time-$2\pi k_0$ map of Hamiltonian $\mathcal{G}_f(A;\varepsilon)$ and we have that F and F_0 are $\mathcal{C}^5$ and we can bound

$$||F - F_0||_{\mathcal{C}^5} \leq \text{cte.}\, \varepsilon^{m+1}||S||_{\mathcal{C}^6} \leq \text{cte.}\, \varepsilon^{m+1-6j}.$$

Since $\mathcal{G}_f$ depends only on A, F_0 is an integrable map of the form $(A,\Psi) \to (A, \Psi + \Delta(A)$.

Furthermore, by item 3. of Proposition 61, we have

$$\frac{d}{dA}\Delta(A) = \frac{\partial^2}{\partial A^2}\mathcal{G}_f(A;\varepsilon) \geq M_f.$$

Hence, the mapping F_0 will be a twist mapping, and we can apply Theorem 45 with $\delta = \varepsilon^{m+1-6j}$, and we obtain, if $m > 6j-1$:

(1) There exist a set of values A_i, such that the Hamiltonian $K \circ \chi_f$ has invariant tori given by

$$A = A_i + \mathcal{A}_i(\psi, s;\varepsilon),$$

where $\mathcal{A}_i$ are $\mathcal{C}^{r-2m-5-\eta}$ functions, for any $\eta > 0$, and $||\mathcal{A}_i||_{\mathcal{C}^2} \leq \text{cte.}\, \varepsilon^{(m+1-6j)/2}$.

(2) The motion on these tori is $\mathcal{C}^1$-conjugate to a rigid translation of frequencies $(\omega(A_i), 1)$, where $\omega(A_i)$ is a Diophantine number of constant type and Markov constant $K\varepsilon^{\frac{m+1-6j}{2}}$, as stated in Definition 42.

(3) The union of neighborhoods of size $\varepsilon^{(m+1-6j)/2}$ of these tori cover all the region $\tilde{D}_f$.

Going back to the original variables $(y,x,s) = (\chi_f(A,\psi),s)$, and using that $K_0(y,x;\varepsilon) = E = \mathcal{G}(A;\varepsilon)$, and that $||\mathcal{G}_f||_{\mathcal{C}^3} \leq \frac{\text{cte.}}{\varepsilon^{j/2}}$, and $||\chi_f^{-1}||_{\mathcal{C}^2} \leq \frac{\text{cte.}}{\varepsilon^{2j}}$, one obtains the desired result if $m+1-11j > 0$, calling $E_i = \mathcal{G}(A_i;\varepsilon)$.

□

Now we turn to estimate the action-angle variables in the regions $D_{o,in}$. We use that, in $D_{o,in}$, the coordinate y is of size $\varepsilon^{j/2}$ and then, it is natural to perform the scaling $y = \varepsilon^{j/2}Y$, which leads directly to the following Lemma. This transformation is not symplectic, but conformally symplectic so that the equations are still in Hamiltonian form.

Lemma 62. *The scaling $y = \varepsilon^{j/2}Y$ transforms the Hamiltonian system of Hamiltonian $K(y,x,s;\varepsilon)$ given in* (77) *into a Hamiltonian* *system of* $\mathcal{C}^{r-2m-2}$ *Hamiltonian*

$$\begin{aligned}\mathcal{K}(Y,x,s;\varepsilon) &= \varepsilon^{j/2}(h_{\rm int}(Y;\varepsilon^{j/2}) + U(x;\varepsilon)) + \varepsilon^{m+1-j/2}S_1(Y,x,s;\varepsilon^{j/2}) \\ &= \varepsilon^{j/2}\mathcal{K}_{\rm int}(Y,x;\varepsilon^{j/2}) + \varepsilon^{m+1-j/2}S_1(Y,x,s;\varepsilon^{j/2}) \qquad (114)\end{aligned}$$

with

$$\begin{aligned} h_{\rm int}(Y;\varepsilon^{j/2}) &= \frac{Y^2}{2}\hat{h}(\varepsilon^{j/2}Y;\varepsilon), \\ S_1(Y,x,s;\varepsilon^{j/2}) &= S(\varepsilon^{j/2}Y,x,s;\varepsilon),\end{aligned}$$

where $\hat{h}(y,\varepsilon)$ *is given in* (80) *and, consequently,* $\mathcal{K}_{\rm int}$ *is a* $\mathcal{C}^{r-2m-1}$ *function.*

The scaling transforms the domains $D_{o,in}$ *given in* (85), (86) *into*

$$\begin{aligned}\mathcal{D}_o &= \{(Y,x,s) : \mathcal{K}_{\rm int}(Y,x;\varepsilon^{j/2}) = F/\varepsilon^j,\ c_3\varepsilon^\alpha \le F \le c_1\varepsilon^j\} \\ &= \{(Y,x,s) : \mathcal{K}_{\rm int}(Y,x;\varepsilon^{j/2}) = e,\ c_3\varepsilon^{\alpha-j} \le e \le c_1\}, \qquad (115) \\ \mathcal{D}_{in} &= \{(Y,x,s) : \mathcal{K}_{\rm int}(Y,x;\varepsilon^{j/2}) = G/\varepsilon^j,\ -c_4\varepsilon^j \le G \le -c_3\varepsilon^\alpha\} \\ &= \{(Y,x,s) : \mathcal{K}_{\rm int}(Y,x;\varepsilon^{j/2}) = e,\ -c_4 \le e \le -c_3\varepsilon^{\alpha-j}\}. \qquad (116)\end{aligned}$$

In the above formulas $(x,s) \in (\mathbb{R}/(2\pi k_0\mathbb{Z}))^2$. We define the action angle variables in

(117)

$$\mathcal{D}_o^* = \{(Y,x) \in \mathbb{R}\times\mathbb{T},\ \exists l \in \{0,1,,\cdots,k_0\},\ (Y,x+2\pi l,s) \in \mathcal{D}_o\}$$

by formulas (104) as in $\mathcal{D}_f^*$. The only change is that, $\mathcal{D}_o^*$, instead of (106) we have:

$$0 \le c_3\varepsilon^{\alpha-j} \le e \le e - U(x;\varepsilon) \le e + c \le c_1 + c. \qquad (118)$$

To bound accurately the action-angle variables for Hamiltonian $\mathcal{K}_{\rm int}(Y,x;\varepsilon^{j/2})$ in $\mathcal{D}_o^*$ we use the following:

Lemma 63. *There exist* A_1, A_2, *independent of* ε, e, *such that if* $0 < c_3\varepsilon^{\alpha-j} \le e \le c_1$ *then, for any* $x \in [0,2\pi]$

(1)

$$A_1\log(1/e) \le \int_0^x \frac{du}{\sqrt{e-U(u;\varepsilon)}} \le A_2\log(1/e).$$

(2) *If* $n \ge 2$

$$\frac{A_1}{e^{(n-1)/2}} \le \int_0^x \frac{du}{(e-U(u;\varepsilon))^{n/2}} \le \frac{A_2}{e^{(n-1)/2}}.$$

Proof. As we have by (79) that $U(x;\varepsilon) \simeq -ax^2 + O(x^3)$, as $x \to 0$, we know that, there exists some $\rho > 0$, independent of ε, and $a_i = a_i(\rho)$ such that

$$a_1x^2 \le -U(x;\varepsilon) \le a_2x^2 \quad \forall x \in [-\rho,\rho] \qquad (119)$$

Then, we compute

$$\begin{aligned}
&\int_0^x \frac{1}{(e-U(u;\varepsilon))^{n/2}}du \\
&= \int_0^\rho \frac{1}{(e-U(u;\varepsilon))^{n/2}}du + \int_\rho^{x-\rho} \frac{1}{(e-U(u;\varepsilon))^{n/2}}du \\
&\quad + \int_{x-\rho}^x \frac{1}{(e-U(u;\varepsilon))^{n/2}}du.
\end{aligned} \tag{120}$$

When $x < \rho$, we will not divide the integral, but we will analyze it as we analyze the first term in (120).

We will compute carefully the first integral in the right hand side of (120). Similar estimates are valid for the third integral when x is close to 2π, and the second is a bounded function even for $0 \le e \le c_1$.

To estimate the first integral in the right hand side of (120), we use (119), and the formula

$$\int_0^\rho \frac{du}{(e+a_i u^2)^{n/2}} = \frac{1}{\sqrt{a_i}e^{(n-1)/2}} \int_0^{\frac{\sqrt{a_i}}{\sqrt{e}}\rho} \frac{dt}{\sqrt{(1+t^2)^n}}$$

If $n = 1$, this integral diverges as $e \to 0$, but integrating explicitly we get

$$\frac{1}{\sqrt{a_i}} \int_0^{\frac{\sqrt{a_i}}{\sqrt{e}}\rho} \frac{1}{(1+t^2)^{1/2}} = \frac{1}{\sqrt{a_i}} \log\left(\frac{\sqrt{a_i}}{\sqrt{e}}\rho + \sqrt{1+\frac{a_i}{e}\rho^2}\right).$$

If $n \ge 1$, the integral is convergent for $e = 0$. This gives us the bounds of the Lemma 63. □

The next Proposition 64 is devoted to putting Hamiltonian $\mathcal{K}_{\mathrm{int}}(Y, x; \varepsilon^{j/2})$ into action-angle variables (A, ψ) in the regions $\mathcal{D}^*_{o,in}$, where $\mathcal{D}^*_0$ is defined in (117) and $\mathcal{D}^*_{in}$ is defined analogously.

Proposition 64. *Consider the $\mathcal{C}^{r-2m-1}$ Hamiltonian $\varepsilon^{j/2}\mathcal{K}_{\mathrm{int}}(Y, x; \varepsilon^{j/2})$ in the regions $\mathcal{D}^*_o$, $\mathcal{D}^*_{in}$. Then, for $\upsilon = o, in$, there exist $\mathcal{C}^{r-2m-2}$ changes of variables*

$$\begin{array}{rcl}
\chi_\upsilon : \tilde{\mathcal{D}}_\upsilon & \to & \mathcal{D}^*_\upsilon \\
(A, \psi) & \mapsto & (Y, x)
\end{array}$$

where $\tilde{D}_\upsilon = \{(A, \psi) : \tilde{c}^1_\upsilon \le A \le \tilde{c}^2_\upsilon\}$, and $\tilde{c}^l_\upsilon$, $l = 1, 2$, are suitable constants independent of ε, such that:

1. $\varepsilon^{j/2}\mathcal{K}_{\mathrm{int}} \circ \chi_\upsilon(A, \psi) = \varepsilon^{j/2}\mathcal{G}_\upsilon(A; \varepsilon^{j/2})$.
2. *For $0 \le s \le r - 2m - 2$ we have that:*

$$||\chi_\upsilon||_{\mathcal{C}^s(\tilde{\mathcal{D}}_\upsilon)} \le \frac{M_\upsilon}{\varepsilon^{s(\alpha-j)}}, \quad ||\chi_\upsilon^{-1}||_{\mathcal{C}^s(\mathcal{D}^*_\upsilon)} \le \frac{M_\upsilon}{\varepsilon^{s(\alpha-j)}}.$$

(3) $||\mathcal{G}_\upsilon||_{\mathcal{C}^3(\tilde{\mathcal{D}}_\upsilon)} \leq \frac{M_\upsilon}{\varepsilon^{2(\alpha-j)}}$, $\left|\mathcal{G}_\upsilon''(A, \varepsilon^{j/2})\right| \geq M_\upsilon$.
where M_υ *is a constant independent of* ε.

Proof. As in Proposition 61, we consider the curves $e = \mathcal{K}_{\rm int}(y, x; \varepsilon^{j/2})$, and using Lemma 60 with $\delta = 1$, we define in D_o^*:

$$\begin{aligned} e &= \mathcal{K}_{\rm int}(y, x; \varepsilon^{j/2}), \\ \tau(x, e) &= \int_0^x \frac{1}{\frac{\partial \mathcal{K}_{\rm int}}{\partial y}(\mathcal{Y}_\pm(u, e), u; \varepsilon^{j/2})} du. \end{aligned}$$

Then, $T(e) = \tau(2\pi, e)$ is the period of the periodic orbit, and the action-angle variables are given by:

$$\begin{aligned} A &= \frac{S(e)}{2\pi} = \frac{1}{2\pi} \int_0^{2\pi} \mathcal{Y}_\pm(x, e) dx, \\ \psi &= \frac{2\pi}{T(e)} \tau(x, e), \end{aligned}$$

where $S'(e) = T(e)$.

Although the region D_o^* has two connected components, we will give full details only for the component where $y > 0$. Let us remember that, when $(Y, s) \in \mathcal{D}_o^*$, the variables (x, e) are in $\mathcal{J}_o$:

$$\mathcal{J}_o = \{(x, e),\ x \in \mathbb{T},\ c_3\varepsilon^{\alpha-j} \leq e \leq c_1\} \tag{121}$$

First, using equation (97), we have, as in (105):

$$\tau(x, e) = \frac{1}{\sqrt{2}} \int_0^x \frac{1 + \varepsilon b}{\sqrt{e - U(u; \varepsilon)}} du + \varepsilon P_1(x, e), \tag{122}$$

where $P_1(x, e) = \int_0^x \frac{\partial}{\partial e} \tilde{\mathcal{Y}}_+(\ell(u, e)) du$. Bounds (98) and (118) give, in D_o, and then, in $\mathcal{J}_o$:

$$|P_1|_{\mathcal{C}^0(\mathcal{J}_o)} \leq \text{cte.}\,,\ |P_1|_{\mathcal{C}^s(\mathcal{J}_o)} \leq \frac{\text{cte.}}{\varepsilon^{(\alpha-j)(s-1/2)}},\ 1 \leq s \leq r - 2m - 2.$$

From these inequalities and Lemma 63, we obtain upper bounds for $\tau(x, e)$:
(123)

$$|\tau|_{\mathcal{C}^0(\mathcal{J}_o)} \leq \text{cte. } \log \frac{1}{\varepsilon^{\alpha-j}},\ |\tau|_{\mathcal{C}^s(\mathcal{J}_o)} \leq \frac{1}{\varepsilon^{(\alpha-j)s}},\ 1 \leq s \leq r - 2m - 2.$$

On the other hand, from $\tau(2\pi, e) = S'(e)$, we obtain:

$$|S|_{\mathcal{C}^0(\mathcal{J}_o)} \leq \text{cte.}\,,\ |S|_{\mathcal{C}^s(\mathcal{J}_o)} \leq \frac{\text{cte.}}{\varepsilon^{(\alpha-j)(s-1)}},\ 1 \leq s \leq r - 2m - 2, \tag{124}$$

and, using again Lemma 63, we also obtain lower bounds for the first and second derivatives of S:

$$|S'(e)| = |T(e)| \geq \text{cte.}\ \log(1/e), \tag{125}$$

$$|S''(e)| \geq \frac{\text{cte.}}{e}. \tag{126}$$

To bound $\psi = \frac{2\pi}{T(e)}\tau(x,e)$ we use (125), (124) and Faa-di-Bruno formulas for the composition of functions derivatives, obtaining:

$$\left|\frac{\partial^s}{\partial e^s}\frac{1}{T(e)}\right| \leq \frac{\text{cte.}}{\varepsilon^{(\alpha-j)s}},\ 0 \leq s \leq r-2m-2,$$

then, the Leibniz rule gives:

$$\left|\frac{1}{T}P_1\right|_{\mathcal{C}^0(\mathcal{J}_o)} \leq \text{cte.}\,,\ \left|\frac{1}{T}P_1\right|_{\mathcal{C}^s(\mathcal{J}_o)} \leq \frac{\text{cte.}}{\varepsilon^{(\alpha-j)(s-1/2)}},\ 1 \leq s \leq r-2m-2.$$

This gives, using again (118),

$$\left|\frac{1}{T}\tau\right|_{\mathcal{C}^s(\mathcal{J}_o)} \leq \frac{\text{cte.}}{\varepsilon^{(\alpha-j)s}},\ 0 \leq s \leq r-2m-2.$$

This bound together with (124) establishes the bound claimed in statement 2. of the proposition 64 for the $\mathcal{C}^s$ norm of the change of variables

$$\chi_v^{(-1)}(y,x) = (\frac{S(e)}{2\pi}, \frac{2\pi}{T(e)}\tau(x,e))$$

with $e = K_{\text{int}}(Y,x;\varepsilon^{j/2})$, $0 \leq s \leq r-2m-2$.

Since $\det D\chi_v^{(-1)}(Y,x) = 1$, the $\mathcal{C}^s$ norm of χ satisfies the same estimates.

Moreover, taking into account that $\mathcal{G}_f = S^{-1}$, the lower bound for the second derivatives of $\mathcal{G}_f$ and the upper bound for the third derivative of $\mathcal{G}_f$ follow from (125), (124), (126).

This finishes the proof of Proposition 64 for the region $\mathcal{D}_o^*$.

In the region $\mathcal{D}_{in}^*$ the action variable is defined as

$$A = \frac{S(e)}{2\pi} = \frac{1}{2\pi}\oint_{\mathcal{K}_{\text{int}}^{-1}(e)} Y\,dx, \tag{127}$$

and ψ is the conjugate angle.

The main complication to obtain the bounds of the change of variables is that the curves $\mathcal{K}_{\text{int}}^{(-1)}(e)$ when $e \leq 0$, cannot be parameterized as graphs of functions $Y = \mathcal{Y}(x,e)$ so that the computation of (127) will need to be divided in different pieces.

Choosing the origin of time in the section

$$\Sigma = \{(Y,x) : Y > 0, x = \pi\},$$

we define $t(Y,x)$ as the time that the trajectory starting in (Y,x) takes to arrive to the section Σ.

To understand the regularity properties of the function $t(Y,x)$ we find it convenient to perform different arguments in different regions of the (Y,x) plane.

Let $\rho > 0$ be the number that appears in Lemma 60, part 2.

If $(Y,x) \in \mathcal{D}^*_{in,\rho} \equiv \{(Y,x) \in \mathcal{K}^{(-1)}_{int}(e),\ Y > 0,\ \rho < x < 2\pi - \rho\}$, the function $t(Y,x)$ is of class $\mathcal{C}^{r-2m-2}$ and its $\mathcal{C}^{r-2m-2}$ norm is bounded independently of e, ε.

Outside of $\mathcal{D}^*_{in,\rho}$ the function $t(Y,x)$ is more complicated to analyze since the orbits starting outside of $\mathcal{D}^*_{in,\rho}$ can pass close to the critical point $(0,0)$ before reaching Σ, and the time $t(Y,x)$ goes to infinity when e goes to zero.

We will explain the case when $Y > 0$ and $0 < x < \rho$, and the other cases are easy modifications of this one.

If $x < \rho$, and $Y > 0$, we compute the time as

$$t(Y,x) = t_1(Y,x) + t_2(Y,x),$$

where t_1 is the time to arrive at the section given by $x = \rho$ (we denote by $Y(\rho)$ the corresponding Y coordinate: $\mathcal{K}_{int}(Y(\rho), x; \varepsilon^{j/2}) = e$), and t_2 is the time between the sections $x = \rho$ and Σ. Note that t_2 is $\mathcal{C}^{r-2m-2}$ with a bounded $\mathcal{C}^{r-2m-2}$ norm, so we only need to estimate t_1.

To bound t_1, we use Lemma 60 and the Hamiltonian equations obtaining that

$$\begin{aligned} t_1(Y,x) &= -\int_Y^{Y(\rho)} \frac{du}{\frac{\partial \mathcal{K}_{\mathrm{int}}}{\partial x}(u, \mathcal{X}_+(u,e); \varepsilon^{j/2})} \\ &= -\int_Y^{Y(\rho)} \frac{\partial \mathcal{X}_+}{\partial e}(u,e)du \\ &= \frac{1}{\sqrt{a}} \int_Y^{Y(\rho)} \frac{1}{\sqrt{h_{int}(u;\varepsilon^{j/2}) - e}} du + \tilde{P}_1(Y,e). \end{aligned}$$

This formula is very similar to formula (122) and analogous arguments to the ones used there establish that

$$|t_1|_{\mathcal{C}^0(\mathcal{J}_i)} \le \text{cte.}\ \log(1/e)),\ |t_1|_{\mathcal{C}^s(\mathcal{J}_i)} \le \frac{\text{cte.}}{\varepsilon^{(\alpha-j)s}},\ 1 \le s \le r - 2m - 2.$$

where $\mathcal{J}_i$ is defined analogously to $\mathcal{J}_o$ in (121) to be domain of the variables (x,e).

Analogous arguments can be used in the cases when $Y < 0$, or $x < \rho$, obtaining exactly the same kind of bounds in all D^*_{in}. Once we have defined the function $t(Y,x)$ in D^*_{in}, we have that the action-angle variables are given by

$$\begin{aligned} A &= \frac{S(e)}{2\pi} \\ \psi &= 2\pi\frac{t(Y,x)}{T(e)}, \end{aligned}$$

where $T(e) = S'(e)$. Proceeding as we did in D^*_o, we obtain the same kind of bounds in these two regions. This finishes the proof of Proposition 64. $\square$

Applying again Theorem 45 to the time-$2\pi k_0$ map of Hamiltonian (77) after it has been scaled and written in action-angle variables and going back to the original variables (y,x,s) we obtain the existence of primary tori in D^*_o and secondary tori in D^*_{in}.

Proof of parts 2) and 3) of Theorem 56. We will prove the results in both regions D^*_o, D^*_{in} at the same time by composing the Hamiltonian $\mathcal{K}$ with the two different $\mathcal{C}^{r-2m-2}$ changes χ_υ, $\upsilon = o, in$.

Again, the proof will consist in applying Theorem 45 to F, the time-$2\pi k_0$ map of the Hamiltonian

$$\tilde{\mathcal{K}}_\upsilon(A,\psi,s;\varepsilon) = \varepsilon^{j/2}\mathcal{G}_\upsilon(A;\varepsilon^{j/2}) + \varepsilon^{m+1-j/2}\tilde{S}_\upsilon(A,\psi,s;\varepsilon^{j/2}), \tag{128}$$

where $\tilde{K}_\upsilon = \mathcal{K}\circ\chi_\upsilon$ and $\tilde{S}_\upsilon = S_1\circ\chi_\upsilon$.

Since $\tilde{K}_\upsilon$, $\mathcal{G}_\upsilon$, $\tilde{S}_\upsilon$ are $\mathcal{C}^{r-2m-2}$, and we have assumed in Theorem 56 that $r-2m-2 \geq 6$, we have that $\tilde{K}_\upsilon$, $\mathcal{G}_\upsilon$, $\tilde{S}_\upsilon$ are $\mathcal{C}^6$.

We denote by F_0 the time-$2\pi k_0$ map of Hamiltonian $\varepsilon^{j/2}\mathcal{G}_\upsilon(A;\varepsilon^{j/2})$. We have that F and F_0 are $\mathcal{C}^5$ and using standard results on dependence of solutions of ode's on parameters

$$||F-F_0||_{\mathcal{C}^5} \leq \text{cte.}\,\varepsilon^{m+1-j/2}||\tilde{S}_\upsilon||_{\mathcal{C}^6} \leq \text{cte.}\,\varepsilon^{m+1-j/2-6(\alpha-j)}.$$

Since $\mathcal{G}_f$ depends only on A, F_0 is an integrable map of the form $(A,\Psi)\to(A,\Psi+\Delta(A)$.

Furthermore, by item 3. of Proposition 64, we have

$$\frac{d}{dA}\Delta(A) = \varepsilon^{j/2}\frac{\partial^2}{\partial A^2}\mathcal{G}_\upsilon(A;\varepsilon^{j/2}) \geq M_\upsilon\varepsilon^{j/2}.$$

Hence, the mapping F_0 will be a twist mapping, and we can apply Theorem 45 with $\delta = \varepsilon^{m+1-j/2-6(\alpha-j)}$, and we obtain, for $m > 6(\alpha-j)+3j/2-1$:

(1) There exist a set of values A_l, such that he Hamiltonian $\tilde{\mathcal{K}}$ has invariant tori given by

$$A = A_l + \mathcal{A}_l(\psi,s;\varepsilon^{j/2}),$$

where $\mathcal{A}_l$ are $\mathcal{C}^{r-2m-5-\eta}$ functions, for any $\eta > 0$, and $||\mathcal{A}_l||_{\mathcal{C}^2} \leq$ cte. $\varepsilon^{\frac{m+1-6(\alpha-j)-3j/2}{2}}$.

(2) The motion on these tori is $\mathcal{C}^1$-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 $K\varepsilon^{\frac{m+1-6(\alpha-j)-j/2}{2}}$, as stated in Definition 42.

(3) The union of neighborhoods of size cte. $\varepsilon^{\frac{m+1-6(\alpha-j)-3j/2}{2}}$ of these tori cover all the regions $\tilde{D}_\upsilon$.

Going back to the variables (Y, x, s), and using that $||\mathcal{G}_\upsilon||_{\mathcal{C}^3} \leq$ cte. $\varepsilon^{-2(\alpha-j)}$, and that $||\chi_\upsilon^{-1}||_{\mathcal{C}^2(\mathcal{D}_\upsilon^*)} \leq$ cte. $\varepsilon^{-2(\alpha-j)}$, we obtain, for $m > 14(\alpha-j) + 3j/2 - 1$, that the tori are given by

$$\mathcal{K}_{\text{int}}(Y, x; \varepsilon^{j/2}) = e_l + O_{\mathcal{C}^2}(\varepsilon^{\frac{m+1-3j/2-14(\alpha-j)}{2}}).$$

Multiplying this equality by ε^j and performing the scaling $y = \varepsilon^{j/2}Y$ we obtain that the tori are given by

$$K_0(y, x; \varepsilon) = \varepsilon^j e_l + \mathrm{O}_{\mathcal{C}^2}(\varepsilon^{\frac{m+1+j/2-14(\alpha-j)}{2}}).$$

Now, calling $F_l = \varepsilon^j e_l$, for $\upsilon = o$, and $G_l = \varepsilon^j e_l$, for $\upsilon = i$, we have the results claimed in Theorem 56. $\square$

Before proving Corollary 57, we analyze the equation

$$K_0(y, x; \varepsilon) = E + \nu g(y, x, s, E; \varepsilon). \tag{129}$$

Lemma 65. *Let us consider equation* (129) *where $K_0(y, x; \varepsilon)$ is given in* (82) *and* (80), *and $g(y, x, s, E; \varepsilon)$ is at least of class $\mathcal{C}^2$, with*

$$||g||_{\mathcal{C}^2} \leq \text{cte.} \tag{130}$$

for $|y| \leq c_2 L$ and

a) *$c_3\varepsilon^\alpha \leq E \leq c_2\bar{L}$ and $(x, s) \in \mathbb{T}^2$*

or

b) *$-c_4\varepsilon^j \leq E \leq 0$, $(x, s) \in \mathbb{T}^2$, and $x \in [\rho, 2\pi - \rho]\}$, where $0 < \rho < \pi$ is any number independent of ε, ν.*

Then, for $\alpha > j$, $j = 1, 2$, and for ε small enough, there exists some ν_0, independent of ε, such that if $\nu \leq \nu_0\varepsilon^\alpha$, equation (129) *defines a function $y = f^\pm(x, s, E; \varepsilon, \nu)$ of class $\mathcal{C}^2$ in the domains a) or b), such that:*

(1) *$f^\pm(x, s, E; \varepsilon, 0) = \mathcal{Y}_\pm(x, E)$, where $\mathcal{Y}_\pm(x, E)$ are the functions* (97) *introduced in Lemma 60, with $\delta = \varepsilon^j$.*

(2) *If we denote by $Df = \frac{\partial f}{\partial x}, \frac{\partial f}{\partial s}$, we have, for $f = f^\pm$:*

$$|f| \leq \text{cte.}, \quad |Df| \leq \varepsilon^{j/2}, \quad \left|D^2 f\right| \leq \varepsilon^{j/2}, \tag{131}$$

and:

$$\left|\frac{\partial f}{\partial E}\right| \le \varepsilon^{-j/2}, \quad \left|\frac{\partial Df}{\partial E}\right| \le \varepsilon^{-j/2}. \tag{132}$$

$$|f-\mathcal{Y}| \le \nu\varepsilon^{-j/2}, \quad |D(f-\mathcal{Y})| \le \nu\varepsilon^{-j/2}, \tag{133}$$

Proof. By part 1. of Lemma 60 with $\delta = \varepsilon^j$, for $y>0$, equation (129) is equivalent to the equation:

$$M(y,x,s,E;\nu) \equiv y - \mathcal{Y}_+(x,t) = 0, \quad t = E + \nu g(y,x,s,E;\varepsilon) \tag{134}$$

where $\mathcal{Y}_+(x,t)$, is the function (97).

Differentiating (97), and using (98), for cte. $\varepsilon^\alpha \le E -$ cte. $|\nu| \le |t| \le E +$ cte. $|\nu| \le$ cte. $\bar{L}$, one has in a) the following bounds for $\mathcal{Y}_+(x,E)$

$$\begin{array}{lll} \left|\dfrac{\partial \mathcal{Y}_+}{\partial x}\right| \le \text{cte.}\,\varepsilon^{j/2} & \left|\dfrac{\partial \mathcal{Y}_+}{\partial t}\right| \le \text{cte.}\,\varepsilon^{-j/2} & \\ \left|\dfrac{\partial^2 \mathcal{Y}_+}{\partial x^2}\right| \le \text{cte.}\,\varepsilon^{j/2} & \left|\dfrac{\partial^2 \mathcal{Y}_+}{\partial x,\partial t}\right| \le \text{cte.}\,\varepsilon^{-j/2} & \left|\dfrac{\partial^2 \mathcal{Y}_+}{\partial t^2}\right| \le \varepsilon^{-3j/2}. \end{array} \tag{135}$$

Using (135) and (130), one has:

$$\left|\frac{\partial M}{\partial y} - 1\right| \le \text{cte.}\,\nu\varepsilon^{-j/2}, \tag{136}$$

so that, the implicit function theorem applied to (134), gives the existence of $f^+(x,s,E;\varepsilon,\nu)$ if $\nu/\varepsilon^{j/2} \le \nu_0$, for some ν_0 small enough.

In order to bound the derivatives of f^+, we take implicit derivatives in equation (134). Then, taking into account (135), and using that $|\nu| < \nu_0\varepsilon^\alpha < \nu_0\varepsilon^j$, we see that:

$$\begin{array}{ll} \left|\dfrac{\partial M}{\partial \tau}\right| \le \text{cte.}\,\varepsilon^{j/2}, \ \tau = x,s, & \left|\dfrac{\partial M}{\partial \nu}\right| \le \text{cte.}\,\varepsilon^{-j/2}, \\ \left|\dfrac{\partial^2 M}{\partial \tau_1 \tau_2}\right| \le \text{cte.}\,\varepsilon^{j/2}, \ \tau_1,\tau_2 = y,x,s, & \left|\dfrac{\partial^2 M}{\partial \nu \partial \tau}\right| \le \text{cte.}\,\varepsilon^{-j/2}, \ \tau = y,x,s. \end{array}$$

These bounds and (136), give the desired bounds (131) and (133).

Moreover, using (135) we see that:

$$\begin{aligned} \left|\frac{\partial M}{\partial E}\right| &\le \text{cte.}\,\varepsilon^{-j/2}, \\ \left|\frac{\partial^2 M}{\partial \tau \partial E}\right| &\le \text{cte.}\,\varepsilon^{-j/2}, \ \tau = y,x,s. \end{aligned}$$

These inequalities and (136) give the desired bounds (132) in the domain a) for $y>0$.

An analogous proof gives the bounds in the domain a) for $y<0$.

For negative values of E—and consequently of t—, we note that $\mathcal{Y}(x,t)$ is the composition of a regular function with the function $\ell(x,t) = \sqrt{2}\sqrt{t - \varepsilon^j U(x,\varepsilon)}$.

We note that in the domain b)—recall that x is restricted to some interval $[\rho, 2\pi - \rho]$—the function $\ell(x,t)$, verifies bounds (135) and therefore so does $\mathcal{Y}_\pm$.

□

Proof of Corollary 57. From now on, we consider $\alpha = j/2 + 3/2$, $j = 1, 2$ and $m \geq 26$. We apply Lemma 65 to the implicit equations (89), (91) and (92), of the invariant tori given by Theorem 56.

The equation (89), that gives implicitly the tori $\mathcal{T}^{\pm}_{E_i}$ in $\mathcal{D}_f$, is a particular case of equation (129) taking $E = E_i$ and $\nu = \varepsilon^{(m+1-7j)/2}$. Analogously for equation (91) which gives implicitly the tori $\mathcal{T}^{\pm}_{F_i}$ in $\mathcal{D}_o$, taking $E = F_i$ and $\nu = \varepsilon^{(m+1+j/2-14(\alpha-j))/2}$ and equation (92), which gives the tori $\mathcal{T}_{G_i}$ in $\mathcal{D}_{in}$, if we take $E = G_i$ and $\nu = \varepsilon^{(m+1+j/2-14(\alpha-j))/2}$.

For $m \geq 26$ and $\alpha = 3/2 + j/2$, the condition $|\nu| \leq \varepsilon^\alpha$ of Lemma 65 is verified in the three cases of the previous paragraph. The results of Lemma 65 give us the items 1., 2., 3., and 6. a), b) of Corollary 57.

From 1.5, 2.5 and 3.5 of theorem 56 we obtain

$$
\begin{aligned}
|E_i - E_{i+1}| &\leq \varepsilon^{\frac{m+1-7j}{2}}, \\
|F_i - F_{i+1}| &\leq \varepsilon^{\frac{m+1+j/2-10(\alpha-j)}{2}}, \\
|G_i - G_{i+1}| &\leq \varepsilon^{\frac{m+1+j/2-10(\alpha-j)}{2}},
\end{aligned}
$$

Taking into account the definitions (84), (85), (86) of D_f, D_o, D_{in}, we get

$$
\begin{aligned}
|E_1 - F_{l_F}| &\leq \varepsilon^{\frac{m+1-7j}{2}}, \\
|F_1 - G_{l_G}| &\leq \varepsilon^\alpha + \varepsilon^{\frac{m+1+j/2-10(\alpha-j)}{2}}.
\end{aligned}
$$

Since $\alpha = j/2 + 3/2$ and $m \geq 26$, all these exponents are bigger than $\frac{3}{2} + \frac{j}{2}$. Using the bounds (132) of Lemma 65, and the inequalities above, we obtain the bounds of 6. c). □

8.5.5. *Existence of stable and unstable manifolds of periodic orbits.* We recall that in Section 8.5.3 we had shown that, in appropriate variables, the motion in the region $\mathcal{S}^{\mathcal{R}_j}$, $j = 1, 2$, is described by the $\mathcal{C}^{r-2m-2}$ Hamiltonian $K(y, x, s; \varepsilon)$ given in (77).

The main part of Hamiltonian K in (77) is the $\mathcal{C}^{r-2m-1}$ term $K_0(y, x; \varepsilon)$ in (82), which has a saddle at $(0, 0)$ with characteristic exponents $\pm\sqrt{c(\varepsilon)}\varepsilon^{j/2}$, where $c(\varepsilon) = -U''(0;\varepsilon)h_0''(0;\varepsilon)$, $c(0) \neq 0$ (see (79), (80)).

The stable and unstable manifolds of $(0,0)$ coincide along a separatrix contained in the level sets $K_0(y,x;\varepsilon)=0$. Hence, these manifolds are $\mathcal{C}^{r-2m-1}$, the same regularity as the Hamiltonian K_0.

When we consider also the variable s, the critical point $(0,0)$ becomes a $2\pi k_0$-periodic orbit given by $\lambda_0=\{(0,0)\}\times\mathbb{R}/(2\pi k_0\mathbb{Z})$. The orbit λ_0 is hyperbolic with characteristic exponents $\pm\sqrt{c(\varepsilon)}\varepsilon^{j/2}$, and the splitting among the stable and unstable directions is of $\mathrm{O}(\varepsilon^{j/2})$.

The following Proposition gives the existence of a hyperbolic periodic orbit for the full Hamiltonian (77) and provides quantitative estimates for the difference of this periodic orbit and its stable and unstable manifolds with those of the unperturbed system.

The main difficulty arises from the fact that the Liapunov exponents are close to zero, as well as the splitting between stable and unstable spaces.

Proposition 66. (1) *The Hamiltonian system of Hamiltonian* (77) *has a periodic orbit $\lambda(\varepsilon)$ which is ε^{m+1-j} close, in the $\mathcal{C}^{r-2m-4}$ sense, to the orbit λ_0. (In fact its y-coordinate is of order $\varepsilon^{m+1-j/2}$).*

(2) *Given $0<\rho<2\pi$, we can find $\mathcal{C}^{r-2m-3}$ functions*

$$\mathcal{Z}_i^{\mathrm{ws}},\ \mathcal{Z}_i^{\mathrm{wu}}:[-\rho,\rho]\times\mathbb{T}\to\mathbb{R}^{\pm}, \tag{137}$$

where $i=up,down$, so that the graph of these functions $y=\mathcal{Z}_i^{\mathrm{wu}}(x,s)$ is a subset of a connected component of $W^{\mathrm{ws}}_{\lambda(\varepsilon)}$, $W^{\mathrm{wu}}_{\lambda(\varepsilon)}$, the stable and unstable manifolds in $\tilde\Lambda_\varepsilon$ of the periodic orbit $\lambda(\varepsilon)$ of the Hamiltonian $K(y,x,s;\varepsilon)$.

(3) *The graphs $\mathcal{Z}_i^{\mathrm{wu}}$ are $\varepsilon^{m+1-j/2}$ close in C^{r-2m-4} norm to the set $K_0(x,y;\varepsilon)=0$. Moreover, they are given by:*

$$y=\mathcal{Z}_{up}^{\mathrm{wu}}(x,s)=\mathcal{Y}_+(x,0)+\mathrm{O}_{\mathcal{C}^1}(\varepsilon^{m+1-j}) \tag{138}$$

where $\mathcal{Y}_\pm$ are given in (97). *Analogous formulas hold for $\mathcal{Z}^{\mathrm{ws}}_{down}$, $\mathcal{Z}^{\mathrm{ws}}_{up}$, $\mathcal{Z}^{\mathrm{wu}}_{down}$.*

Proof. This reduces to Theorems A and C in the paper [FS90a].

We consider the scaling $y=\varepsilon^{j/2}Y$, and, using Lemma 62, we obtain Hamiltonian (114).

First, we write the perturbation term in (114), as

$$S_1(Y,x,s;\varepsilon)=\bar S_1(Y,x;\varepsilon)+\tilde S_1(Y,x,s;\varepsilon),$$

where $\bar S_1$ is the averaged Hamiltonian with respect to the periodic variable s.

It is clear that the Hamiltonian system of Hamiltonian $\varepsilon^{j/2}\mathcal{K}_{\mathrm{int}}+\varepsilon^{m+1-j/2}\bar S_1$ has a critical point which is ε^{m+1-j} close to the origin. Then, applying the results of Theorem C and Proposition 5.1

in [FS90a], we obtain the existence of the periodic orbit $\lambda(\varepsilon)$, which, after the scaling $y = \varepsilon^{j/2}Y$, has the x component of order ε^{m+1-j}, and the y component of order $\varepsilon^{m+1-j/2}$.

Hamiltonian (114) is $2\pi k_0$-periodic in the variable s. Then, one can consider the time $2\pi k_0$ map F_ε of this Hamiltonian, and this map has a fixed point P_ε, corresponding to the periodic orbit $\lambda(\varepsilon)$. We will consider also the time $2\pi k_0$ map F_0 of Hamiltonian $\varepsilon^{j/2}\mathcal{K}_{\rm int}$ given in (114).

These maps verify:

$$\begin{aligned} ||F_0 - \mathrm{Id}||_{\mathcal{C}^{r-2m-3}} &\leq \varepsilon^{j/2} \\ ||F_\varepsilon - F_0||_{\mathcal{C}^{r-2m-4}} &\leq \varepsilon^{m+1-j/2}, \end{aligned}$$

and the eigenvalues of the fixed points are $1 + O(\varepsilon^{j/2})$. Then, by Theorem A in [FS90a], we obtain the existence of $W^{\rm ws,wu}_{P_\varepsilon}$, which are of class $\mathcal{C}^{r-2m-3}$, and ε^{m+1-j} close, in the $\mathcal{C}^{r-2m-4}$ sense, to those of K_0. Moreover, in $[-\rho, \rho] \times \mathbb{T}$, using Lemma 60 and Lemma 65for $E = 0$, $\delta = \varepsilon^j$ and $\nu = \varepsilon^{m+1-j}$, they can be written as a graph of the variable Y over the variables (x, s) verifying (138).

Going back to the variables $(y, x, s) = (\varepsilon^{j/2}Y, x, s)$, we obtain Proposition 66 for $\mathcal{Z}_i^v$.

□

Remark 67. Proposition 66 gives the stable and unstable manifolds of the periodic orbit inside $\tilde\Lambda_\varepsilon$ in the variables (y, x, s). We can write them in the original variables (I, φ, s) using the changes given by Proposition 28, Theorem 35 and the changes (73) and (76).

In the variables (I, φ, s) these manifolds inside the region $\mathcal{S}^{\mathcal{R}^j}$ are given by

$$I = -k_0/l_0 + \bar{\mathcal{Z}}_i^v(\varphi, s; \varepsilon)$$

where the functions $\bar{\mathcal{Z}}_i^v(\varphi, s; \varepsilon)$ verify the same regularity properties as the functions $\mathcal{Z}_i^v$. □

Remark 68. Note that the results claimed in Proposition 66 are rather conservative and there are better results with more completed arguments. For example, applying the regularity theory of invariant manifolds in the variables (J, φ, s), we can conclude that the invariant manifolds are as smooth as the flow of the Hamiltonian $k(J, \varphi, s; \varepsilon)$ given in (45), that is, $\mathcal{C}^{r-3}$. Similar improvements can be made to the dependence on parameters.

Nevertheless, for our purposes, it is more convenient to discuss the manifolds in the averaged set of coordinates since this will allow us to compare them with primary and secondary KAM tori, which are better expressed in a system of coordinates with action-angle variables. (A reference where the smooth dependence of invariant manifolds, more general than the stable and unstable, is considered rather explicitly is [LW95], [CFdlL03].)

□

Remark 69. Note that we are not even attempting to study whether $W^{\rm ws}_{\lambda(\varepsilon)}$ and $W^{\rm wu}_{\lambda(\varepsilon)}$ intersect transversely in $\tilde{\Lambda}_\varepsilon$ or not. Even if such questions are crucial in other approaches to diffusion, they do not play any role in our method. (See also [FS90b, FS90a], for a discussion about the intersection of the stable and unstable manifolds in situations similar to the ones we are considering here.) □

Remark 70. It will be important to note that the two-dimensional invariant manifolds of periodic orbits $W^{\rm ws,wu}_{\lambda(\varepsilon)}$, produced by studying the dynamics on $\tilde{\Lambda}_\varepsilon$, have Lyapunov exponents that are $\varepsilon^{j/2}$ close to zero. Therefore, when we consider them in the full phase space, they are *slow* manifolds. The directions transverse to $\tilde{\Lambda}_\varepsilon$ have Lyapunov exponents of size $\tilde{\mu} = \mu + O(\varepsilon)$ (see Remark 23).

Hence, the manifolds $W^{\rm ws,wu}_{\lambda(\varepsilon)}$, when considered as invariant manifolds in the whole space are only *weak* (un)stable manifolds of the periodic orbit $\lambda(\varepsilon)$. We note that in [LW95] one can find a theory for these manifolds. The manifolds $W^{\rm ws,wu}_{\lambda(\varepsilon)}$ are not the (un)stable manifolds associated in the normal hyperbolicity theory.

For our purposes, it will be enough to define that if we consider the *total* (un)stable manifold of an invariant object $\mathcal{V}$ as:

(139)
$$W^{\rm tu,ts}_{\mathcal{V}} = \{\tilde{y} : \exists \tilde{x} \in \mathcal{V}, \operatorname{dist}(\Phi_{t,\varepsilon}(\tilde{y}), \Phi_{t,\varepsilon}(\tilde{x})) \to 0, \quad \text{as} \quad t \to \pm\infty\}$$

it contains both the weak and strong (un)stable manifolds of $\mathcal{V}$.

Indeed, this total stable manifold—without rate conditions—can be expressed as the strong stable manifold—in the sense of normal hyperbolicity [Fen74]—of the weak stable manifolds contained in $\tilde{\Lambda}_\varepsilon$.

Some examples of this construction of stable manifolds that will play a role later are:

When $\mathcal{V} = \mathcal{T}$ is a 2-dimensional torus in $\tilde{\Lambda}_\varepsilon$, its weak (un)stable manifolds are simply $\mathcal{T}$, and therefore $W^{\rm tu,ts}_{\mathcal{T}} = W^{\rm u,s}_{\mathcal{T}}$.

When $\mathcal{V} = \lambda(\varepsilon)$ is a periodic orbit in $\tilde{\Lambda}_\varepsilon$, then $\lambda(\varepsilon)$ is strictly contained in its weak (un)stable manifolds $W^{\rm wu,ws}_{\lambda(\varepsilon)}$, and then

$$W^{\rm tu,ts}_{\lambda(\varepsilon)} = W^{\rm u,s}_{W^{\rm wu,ws}_{\lambda(\varepsilon)}} = \bigcup_{x \in W^{\rm wu,ws}_{\lambda(\varepsilon)}} W^{\rm u,s}_x.$$

In subsequent chapters, when we discuss transition chains, we will use the total (un)stable manifolds of different invariant objects $\tilde{\Lambda}_\varepsilon$.
$\square$

9. The scattering map

The aim of this section is to define and compute the *scattering map* (also called *outer map*) $S : \tilde{\Lambda}_\varepsilon \longrightarrow \tilde{\Lambda}_\varepsilon$.

9.1. **Some generalities about the scattering map.** The following definitions and elementary facts come from [DLS00].

Definition 71. *Let $\Lambda \subset M$ be a normally hyperbolic invariant manifold for a flow Φ_t in a manifold M.*

Assume that $\gamma \subset W^{\mathrm{s}}_\Lambda \cap W^{\mathrm{u}}_\Lambda$ is a homoclinic manifold, and that the intersection of W^{s}_Λ and W^{u}_Λ is transverse *along γ, that is*

$$
\begin{aligned}
T_z W^{\mathrm{s}}_\Lambda + T_z W^{\mathrm{u}}_\Lambda &= T_z M, \quad \forall z \in \gamma. \\
T_z W^{\mathrm{s}}_\Lambda \cap T_z W^{\mathrm{u}}_\Lambda &= T_z \gamma, \quad \forall z \in \gamma.
\end{aligned}
\tag{140}
$$

Then, for any two points $x_\pm \in \Lambda$, we say that $x_+ = S(x_-)$, if there exists a point $z \in \gamma$ such that

$$\operatorname{dist}(\Phi_t(z), \Phi_t(x_\pm)) \to 0, \quad \textit{for } t \to \pm\infty.$$

Note that the definition of the scattering map S depends on the homoclinic manifold γ, but we will not include this in the notation since it will not lead to confusion.

Heuristically, the scattering map associates to the asymptotic behavior in the past of homoclinic orbits in γ, the asymptotic behavior in the future. Therefore, one should think about it as an accounting device describing the homoclinic excursions.

A feature that will be crucial for us is that one can study very comfortably the intersection of stable and unstable manifolds of sets in Λ using the scattering map independently of the topological type of the objects considered (we do not need that they are in a common system of coordinates). We will use this to discuss the existence of heteroclinic connections between primary and secondary tori or among tori of different dimensions. This is crucial for the mechanisms we introduce in this paper since we use these objects of different topologies and different dimensions to fill the gaps among KAM primary tori.

Note that the domain and the range of the scattering map associated to an intersection γ as above are

$$
\begin{aligned}
\operatorname{Dom}(S) &= \{x_- \in \Lambda, W^{\mathrm{u}}_{x_-} \subset \gamma\}, \\
\operatorname{Ran}(S) &= \{x_+ \in \Lambda, W^{\mathrm{s}}_{x_+} \subset \gamma\}.
\end{aligned}
$$

If γ verifies (140), then the dimension of Λ and γ are the same and we obtain that the domain and the range of the scattering map will be open (and non-empty) subsets. In the case considered in [DLS00], both the domain and the range of the scattering map were the whole manifold Λ. In this paper, however, we will have to pay more attention to the domains.

We recall from [DLS00] that the scattering map is locally well defined and as smooth as the invariant stable and unstable bundles of Λ. The reason is that if we take locally a section Σ of γ transversal to the flow, we can define $\Omega_\pm : \Lambda \to \Sigma$ by setting $\Omega_-(x)$ to be the (locally unique) point in $\Sigma \cup W^{\mathrm{u}}_x$ and $\Omega_+(x)$ to be the (locally unique) point in $\Sigma \cup W^{\mathrm{s}}_x$. These maps are clearly as smooth as the map $x \to W^{\mathrm{s}}_x$, $x \to W^{\mathrm{u}}_x$. The scattering map is given, locally by $S = \Omega_+ \circ \Omega_-^{-1}$.

The regularity of the maps $x \to W^{\mathrm{s,u}}_x$ is studied in great detail in [HPS77, Fen74]. In general, it depends on ratios of several rates of expansion. In our case, however, as we will see in Theorem 20, it is as smooth as the flow because the Lipschitz constant of the flow along the stable manifold is close to 1.

To check that the scattering map is globally defined, we just need to check that if we continue these local definitions around a loop there is no monodromy. This follows because if $W^{\mathrm{s}}_{x_+}$ intersects nontrivially $W^{\mathrm{s}}_{\tilde{x}_+}$ and $x_+, \tilde{x}_+ \in \Lambda$, we conclude that $x_+ = \tilde{x}_+$. (Of course, in this paper we may be considering situations in which $\mathrm{Dom}(S)$ is constructible and the issue of global definition of S does not arise.)

9.2. **The scattering map in our model: definition and computation.** In this section, we will show that, in the assumptions of Theorem 7, the manifold $\tilde{\Lambda}_\varepsilon$ constructed in Section 7 has a scattering map and we will compute the leading order.

The first difficulty to define the scattering map for $\tilde{\Lambda}_\varepsilon$ in our problem is that, for $\varepsilon = 0$, its stable and unstable manifolds coincide: $W^{\mathrm{s}}_{\tilde{\Lambda}} = W^{\mathrm{u}}_{\tilde{\Lambda}}$. We will show that under hypothesis **H4** of Theorem 7 the stable and unstable invariant manifolds of $\tilde{\Lambda}_\varepsilon$ have a transversal homoclinic intersection for $0 < |\varepsilon| < \varepsilon^*$. (This will just amount to verifying that in the situation considered here, the conditions in first order perturbation theory **H4** are enough to guarantee the existence of a transversal intersection associated to the critical point τ^*.)

The transversal intersection associated to the critical point τ^* will happen for the points in $H_- \subset \tilde{\Lambda}_\varepsilon$, where H_- is defined in (11). Hence, it will be possible to define a scattering map S with $H_- \subset$

$\mathrm{Dom}(S)$. We will denote by $H_+ = S(H_-)$, so we have that $H_+ \subset \mathrm{Ran}(S)$.

Both the verification of the existence of the scattering map and explicit computations will be accomplished by first order perturbation theory of the stable and unstable manifolds adapting the calculations carried out in Section 7.1.

It is convenient to take advantage of the exact symplectic structure of the problem. We will see that the existence of homoclinic intersections will be given, by first order perturbation theory, by the zeros of a directional derivative of a single function called the *Poincaré function* or the *Melnikov potential*, which was written explicitly in (10), and introduced in [DG00].

As we will see, the Poincaré function not only appears in the existence of intersections for $W^{\mathrm{s}}_{\tilde{\Lambda}_\varepsilon}$, $W^{\mathrm{u}}_{\tilde{\Lambda}_\varepsilon}$ but also in the calculation of the scattering map (see (147)). This, in turn, will help us with the calculations of the existence of transition chains in Section 10.3. The hypothesis **H4** and, hence, the other subsequent hypothesis **H5"**, **H5"'** are expressed in terms of the Poincaré function.

We recall from (10) that the definition of the Poincaré function is:

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

where $(p_0(t), q_0(t))$ is the parameterization (23) of the separatrix to the saddle point $(0,0)$ of the pendulum $P_\pm(p,q) = \pm(p^2/2 + V(q))$ with characteristic exponent $\mu = \sqrt{-V''(0)} > 0$.

The following Proposition 72 establishes the existence of transverse intersections for $W^{\mathrm{s}}(\tilde{\Lambda}_\varepsilon)$ and $W^{\mathrm{u}}(\tilde{\Lambda}_\varepsilon)$ along a manifold $\tilde{\gamma}_\varepsilon$ under the assumption **H4** of Theorem 7 and provides with perturbative formulas for the scattering map.

Proposition 72. *Given $(I,\varphi,s) \in [I_-, I_+] \times \mathbb{T} \times \mathbb{T}$, assume that the function $\tau \in \mathbb{R} \mapsto \mathcal{L}(I, \varphi - I\tau, s - \tau)$ has a non-degenerate critical point at $\tau = \tau^*(I,\varphi,s)$.*

Then, for ε small enough, there exists a locally unique point $\tilde{z}$ of the form

$$\tilde{z} = \tilde{z}(\tau, I, \varphi, s; \varepsilon) = (p_0(\tau) + \mathrm{O}(\varepsilon), q_0(\tau) + \mathrm{O}(\varepsilon), I, \varphi, s), \tag{141}$$

with $\tau = \tau^*(I,\varphi,s)$, such that $W^{\mathrm{s}}(\tilde{\Lambda}_\varepsilon) \pitchfork W^{\mathrm{u}}(\tilde{\Lambda}_\varepsilon)$ at $\tilde{z}$, that is,

$$\tilde{z} \in W^{\mathrm{s}}(\tilde{\Lambda}_\varepsilon) \cap W^{\mathrm{u}}(\tilde{\Lambda}_\varepsilon) \text{ and } T_{\tilde{z}} W^{\mathrm{s}}(\tilde{\Lambda}_\varepsilon) + T_{\tilde{z}} W^{\mathrm{u}}(\tilde{\Lambda}_\varepsilon) = T_{\tilde{z}}\mathcal{M},$$

where $\mathcal{M} = \mathbb{R} \times \mathbb{T} \times [I_-, I_+] \times \mathbb{T} \times \mathbb{T}$.

In particular, there exist unique points

$$\tilde{x}_\pm = \tilde{x}_\pm(I,\varphi,s;\varepsilon) = (0,0,I,\varphi,s) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon) \in \tilde{\Lambda}_\varepsilon$$

such that

$$|\Phi_{t,\varepsilon}(\tilde{z}) - \Phi_{t,\varepsilon}(\tilde{x}_\pm)| \le \text{cte.}\, e^{-\mu|t|/2} \quad \textit{for} \quad t \to \pm\infty. \tag{142}$$

Moreover, expressing the points $\tilde{x}_\pm = \tilde{\mathcal{F}}(I_\pm, \varphi_\pm, s_\pm; \varepsilon)$ *in terms of the parameterization* (30) *of* $\tilde{\Lambda}_\varepsilon$ *given in Theorem* 20*, the following formulas hold:*

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

and

$$I(\tilde{x}_+) - I(\tilde{x}_-) = \varepsilon \frac{\partial \mathcal{L}}{\partial \varphi}(I, \varphi - I\tau, s - \tau) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho}),$$

where τ *is given again by* $\tau = \tau^*(I,\varphi,s)$*, and* $\varrho > 0$*.*

Remark 73. We recall that **H4** in Theorem 7 is precisely that the hypothesis of Proposition 72 occurs with uniform constants for a non-empty open set $H_- \subset \tilde{\Lambda}_\varepsilon$.

In this case, the domain and the range of the scattering map will include H_- and $S(H_-)$ respectively.

This will make it somewhat easier to state and check the subsequent non-degeneracy properties. □

Proof. Let us consider the point

$$\tilde{z}_0 = \tilde{z}_0(\tau, I, \varphi, s) := (p_0(\tau), q_0(\tau), I, \varphi, s)$$

of the unperturbed homoclinic 4-dimensional manifold $\tilde{\gamma}$ given in (28). The line

$$N = \left\{ \tilde{z}_0 + u \left(\frac{\partial P_\pm}{\partial p}\left(p_0(\tau), q_0(\tau)\right), \frac{\partial P_\pm}{\partial q}\left(p_0(\tau), q_0(\tau)\right), 0, 0, 0 \right), u \in \mathbb{R} \right\}$$

is normal to $\tilde{\gamma}$ at $\tilde{z}_0$: $N \pitchfork_{\tilde{z}_0} \tilde{\gamma}$. Since $W^{\mathrm{s,u,loc}}_{\tilde{\Lambda}_\varepsilon}$ is C^1 ε-close to $\tilde{\gamma}$, it follows that $N \pitchfork W^{\mathrm{s,u,loc}}_{\tilde{\Lambda}_\varepsilon}$ at a locally unique $\tilde{z}^{\mathrm{s,u}} = \tilde{z}^{\mathrm{s,u}}(\tau, I, \varphi, s; \varepsilon) \in N$, ε-close to $\tilde{z}_0$. Notice that $\tilde{z}^{\mathrm{s,u}}$ has the form

$$\tilde{z}^{\mathrm{s,u}} = (p_0(\tau) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon), q_0(\tau) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon), I, \varphi, s).$$

Therefore, by the implicit function theorem, $W^{\mathrm{s}}(\tilde{\Lambda}_\varepsilon)$, $W^{\mathrm{u}}(\tilde{\Lambda}_\varepsilon)$ will have a non-empty intersection close to $\tilde{z}_0$ when

$$DP_\pm(p_0(\tau), q_0(\tau))(\tilde{z}^{\mathrm{u}} - \tilde{z}^{\mathrm{s}}) = 0, \tag{143}$$

where $DP_\pm(p,q)(\tilde{z}^{\mathrm{u}} - \tilde{z}^{\mathrm{s}})$ is a shorthand for $(DP_\pm(p,q), 0, 0, 0)(\tilde{z}^{\mathrm{u}} - \tilde{z}^{\mathrm{s}})$.

To provide a more easily computable formula for (143), we first notice that, for ε small enough, as a consequence of the normal hyperbolicity of $\tilde{\Lambda}_\varepsilon$, there exist unique points $\tilde{x}_\pm = \tilde{x}_\pm(I,\varphi,s;\varepsilon) \in \tilde{\Lambda}_\varepsilon$ such that $\Phi_{t,\varepsilon}(\tilde{z}^{\mathrm{s,u}}) - \Phi_{t,\varepsilon}(\tilde{x}_\pm) \to 0$ exponentially as in (142) for $t \to \pm\infty$, and that $\Phi_{t,\varepsilon}(\tilde{z}^{\mathrm{s,u}}) - \Phi_{t,0}(\tilde{z}_0) = \mathrm{O}_{\mathcal{C}^1}(\varepsilon)$, uniformly in ε for $t \to \pm\infty$. In particular, $\tilde{x}_\pm = (0,0,I,\varphi,s) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon)$ and

$$DP_\pm\left(p_0(\tau+t), q_0(\tau+t)\right)\left(\Phi_{t,\varepsilon}(\tilde{z}^{\mathrm{s,u}})\right) \longrightarrow 0 \text{ for } t \to \pm\infty.$$

We apply the fundamental theorem of calculus to

$$t \mapsto DP_\pm(p_0(\tau+t), q_0(\tau+t))\left(\Phi_{t,\varepsilon}(\tilde{z}^{\mathrm{s,u}})\right),$$

to get

$$\begin{aligned} DP_\pm(p_0(\tau), q_0(\tau))(\tilde{z}^{\mathrm{s,u}}) &= \varepsilon \int_{\pm\infty}^{0} \{P_\pm, h\}\left(\Phi_{\sigma,\varepsilon}(\tilde{z}^{\mathrm{s,u}})\right) d\sigma \\ &= \varepsilon \int_{\pm\infty}^{0} \{P_\pm, h\}\left(\Phi_{\sigma,0}(\tilde{z}_0)\right) d\sigma + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^2), \end{aligned}$$

where $\{P_\pm, h\} = \partial_q P_\pm\, \partial_p h - \partial_p P_\pm\, \partial_q h$ is the Poisson bracket of the functions $P_\pm$ and h. Subtracting the expressions above, we get

$$DP_\pm(p_0(\tau), q_0(\tau))(\tilde{z}^{\mathrm{u}} - \tilde{z}^{\mathrm{s}}) = \varepsilon \frac{\partial L}{\partial \tau}(\tau, I, \varphi, s) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^2)$$

where $L(\tau, I, \varphi, s)$ is given by

(144)

$$\begin{aligned} L(\tau, I, \varphi, s) = -\int_{-\infty}^{+\infty} \Big(& h(p_0(\tau+\sigma), q_0(\tau+\sigma), I, \varphi + I\sigma, s+\sigma) \\ & - h(0,0,I,\varphi + I\sigma, s+\sigma)\Big) d\sigma. \end{aligned}$$

From the expression of $L(\tau, I, \varphi, s)$, it is immediate to check that

$$L(\tau+t, I, \varphi + It, s+t) = L(\tau, I, \varphi, s) \text{ for any } t \in \mathbb{R}.$$

Therefore, writing $t = -\tau$, we arrive at

$$L(\tau, I, \varphi, s) = L(0, I, \varphi - I\tau, s-\tau) = \mathcal{L}(I, \varphi - I\tau, s-\tau),$$

where $\mathcal{L}$ is the Melnikov potential defined in (10). Then, equation (143) is equivalent to

$$\varepsilon \frac{\partial \mathcal{L}}{\partial \tau}(I, \varphi - I\tau, s-\tau) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^2) = 0.$$

By the implicit function Theorem, non-degenerate critical points τ of the function $\tau \mapsto \mathcal{L}(I, \varphi - I\tau, s-\tau)$ give rise, for ε small enough, to transverse intersections of the stable and unstable manifolds of $\tilde{\Lambda}_\varepsilon$ along points $\tilde{z} = \tilde{z}(\tau, I, \varphi, s; \varepsilon)$ of the form (141).

By the implicit function Theorem, we can find a function τ^* defined in an open set where $\tau^*(I;\varphi,s)$ is a critical point of $\tau \mapsto \mathcal{L}(I,\varphi - I\tau, s-\tau)$.

To finish the proof, we now consider the expression of the points $\tilde{x}_\pm = \tilde{\mathcal{F}}(I_\pm,\varphi_\pm,s_\pm;\varepsilon)$ in terms of the parameterization (30) of $\tilde{\Lambda}_\varepsilon$ given in Theorem 20. Since we already know the existence of $\tilde{z}$ given in (141) such that (142) holds, it is clear that

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

and it only remains to obtain the formula for $I(\tilde{x}_+) - I(\tilde{x}_-)$.

We apply now the Fundamental Theorem of Calculus to

$$t \mapsto I\left(\Phi_{t,\varepsilon}(\tilde{x}_\pm)\right) - I\left(\Phi_{t,\varepsilon}(\tilde{z})\right),$$

to get

$$I(\tilde{x}_\pm) - I(\tilde{z}) = \varepsilon \int_{\pm\infty}^{0} \left(\{I,h\}\left(\Phi_{\sigma,\varepsilon}(\tilde{x}_\pm)\right) - \{I,h\}\left(\Phi_{\sigma,\varepsilon}(\tilde{z})\right)\right) d\sigma,$$

where $\{I,h\} = \partial_\varphi I\,\partial_I h - \partial_I I\,\partial_\varphi h = -\partial_\varphi h$ is the Poisson bracket of the functions I and h. Subtracting the expressions above, we get

$$\begin{aligned} I(\tilde{x}_+) - I(\tilde{x}_-) = \varepsilon &\int_0^{+\infty} \frac{\partial h}{\partial\varphi}(\Phi_{\sigma,\varepsilon}(\tilde{z})) - \frac{\partial h}{\partial\varphi}(\Phi_{\sigma,\varepsilon}(\tilde{x}_+)) d\sigma \\ &+\varepsilon \int_{-\infty}^{0} \frac{\partial h}{\partial\varphi}(\Phi_{\sigma,\varepsilon}(\tilde{z})) - \frac{\partial h}{\partial\varphi}(\Phi_{\sigma,\varepsilon}(\tilde{x}_-)) d\sigma. \end{aligned}$$

We already know that

$$\Phi_{t,\varepsilon}(\tilde{z}) = \Phi_{t,0}(\tilde{z}_0) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon), \quad \forall t \in \mathbb{R}.$$

Taking c_2 sufficiently small (but independent of ε) and using Gronwall inequality we have, for $-c_2\,|\log\varepsilon| \le t \le c_2\,|\log\varepsilon|$

$$\Phi_{t,\varepsilon}(\tilde{x}_\pm) = \Phi_{t,0}(\tilde{x}_0) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho_1}),$$

for some $\varrho_1 > 0$.

From equation (142), we deduce that there exists a constant $c_1 > 0$,

$$\begin{aligned} &\left| \int_{\pm c_2|\log\varepsilon|}^{\pm\infty} \left(\frac{\partial h}{\partial\varphi}(\Phi_{\sigma,\varepsilon}(z)) - \frac{\partial h}{\partial\varphi}(\Phi_{\sigma,\varepsilon}(x_\pm)) \right) d\sigma \right| \\ &\qquad\qquad \le c_1 e^{-\mu c_2|\log\varepsilon|/2} = \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho_2}), \end{aligned}$$

and, since for the unperturbed system we have the same kind of behavior, we can conclude finally that, for some $\varrho > 0$

$$
\begin{aligned}
&I(\tilde{x}_+) - I(\tilde{x}_-)\\
&\quad = -\varepsilon \int_{-c_2|\log\varepsilon|}^{+c_2|\log\varepsilon|} \left(\frac{\partial h}{\partial\varphi}(q_0(\tau+\sigma), p_0(\tau+\sigma), I, \varphi + I\sigma, s+\sigma) \right.\\
&\qquad \left. - \frac{\partial h}{\partial\varphi}(0,0,I,\varphi+I\sigma, s+\sigma) \right) d\sigma + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})\\
&\quad = -\varepsilon \int_{-\infty}^{+\infty} \left(\frac{\partial h}{\partial\varphi}(q_0(\tau+\sigma), p_0(\tau+\sigma), I, \varphi + I\sigma, s+\sigma) \right.\\
&\qquad \left. - \frac{\partial h}{\partial\varphi}(0,0,I,\varphi+I\sigma, s+\sigma) \right) d\sigma + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})\\
&\quad = -\varepsilon \int_{-\infty}^{+\infty} \left(\frac{\partial h}{\partial\varphi}(q_0(r), p_0(r), I, \varphi + Ir - I\tau, s+r-\tau) \right.\\
&\qquad \left. - \frac{\partial h}{\partial\varphi}(0,0,I,\varphi+Ir-I\tau, s+r-\tau) \right) dr + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})\\
&\quad = -\varepsilon \frac{\partial \mathcal{L}}{\partial\varphi}(I, \varphi - I\tau, s-\tau) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho}).
\end{aligned}
$$

□

Remark 74. Observe 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

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

or, equivalently, there exists a function $\mathcal{L}^*(I,\theta)$, that we will call the reduced Poincaré function, defined by

$$\mathcal{L}(I, \varphi - I\tau^*(I,\varphi,s), s - \tau^*(I,\varphi,s)) := \mathcal{L}^*(I, \varphi - Is). \tag{145}$$

We note that, if we assume hypothesis **H4**, again by the definition of τ^*, for $(I,\varphi,s) \in H_-$ (see (11)), the function $\mathcal{L}^*$ verifies

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

for $x = I, \varphi, s$.

We denote by H_-^* the domain of $\mathcal{L}^*$. Clearly

(146)

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

□

Finally, from (145), the scattering map S on $\tilde{\Lambda}_\varepsilon$ can be computed

(147)

$$\begin{aligned} S : H_- \subset \tilde{\Lambda}_\varepsilon &\longrightarrow \tilde{\Lambda}_\varepsilon \\ (I, \varphi, s) &\longmapsto (I - \varepsilon \frac{\partial \mathcal{L}^*}{\partial \varphi}(I, \varphi - Is) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho}), \varphi + \mathrm{O}_{\mathcal{C}^1}(\varepsilon), s) \end{aligned}$$

10. Existence of transition chains

The goal of this section is to construct transition chains—see Definition 76 in Section 10.1—along the manifold $\tilde{\Lambda}_\varepsilon$. In particular, these transition chains will traverse the resonant regions. Later, in Section 11 we will show that, once we have a transition chain, there are orbits that follow it closely. Hence, there are orbits that traverse the resonant regions as claimed in Theorem 7.

Transition chains were an important ingredient of the method in [Arn64]. The main novelty here is that we introduce two new ingredients in the transition chains. We allow the transition chains to include secondary tori and periodic orbits, as well as the usual primary tori.

In Section 10.2, we give explicit conditions on the perturbations considered in (6) that imply that there are transition chains that overcome the large gap problem. These conditions are part of **H4** and **H5** in Theorem 7.

First, in Section 10.2 we show how the scattering map, constructed in Section 9, can be used as a tool to discuss heteroclinic intersections of objects of different topological types. Roughly —see Lemma 78— two submanifolds $\mathcal{V}_1$, $\mathcal{V}_2$ of $\tilde{\Lambda}_\varepsilon$ have a transverse heteroclinic connection, if $S(\mathcal{V}_1)$ is transversal to $\mathcal{V}_2$ as submanifolds of $\tilde{\Lambda}_\varepsilon$. This provides us with an alternative to the customary Melnikov calculations, which require common systems of coordinates for both objects.

Recall that in Section 9 we have obtained a explicit expression (147) for the scattering map. In Propositions 47, 50, and Corollary 57 we have obtained explicit expressions (depending on the proximity to resonances) for the KAM tori, both primary or secondary, in $\tilde{\Lambda}_\varepsilon$. In Proposition 66, we have studied the weak invariant manifolds in $\tilde{\Lambda}_\varepsilon$ of periodic orbits associated to double resonances. (See Figure 1 for a representation of the results obtained so far). Roughly speaking, all these objects are given very approximately by the level sets of the averaged Hamiltonian.

In Lemma 81 we will prove a general result that allows to verify transverse intersection in $\tilde{\Lambda}_\varepsilon$ of $S(\mathcal{V}_1)$ and $\mathcal{V}_2$, when $\mathcal{V}_1$, $\mathcal{V}_2$, submanifolds of $\tilde{\Lambda}_\varepsilon$, are given as level sets of some function. In Lemmas 82, 85, 88 we will apply Lemma 81 to the cases of primary or secondary tori or (un)stable manifolds of periodic orbits.

Finally, in Section 10.3 we show that all the gaps between the objects mentioned are close enough in order to obtain transition chains between them.

A pictorial representation of the results of this section is in Figure 2. For simplicity of the pictorial representation we have depicted only the case where the stronger version of **H4** discussed in Remark 10 holds.

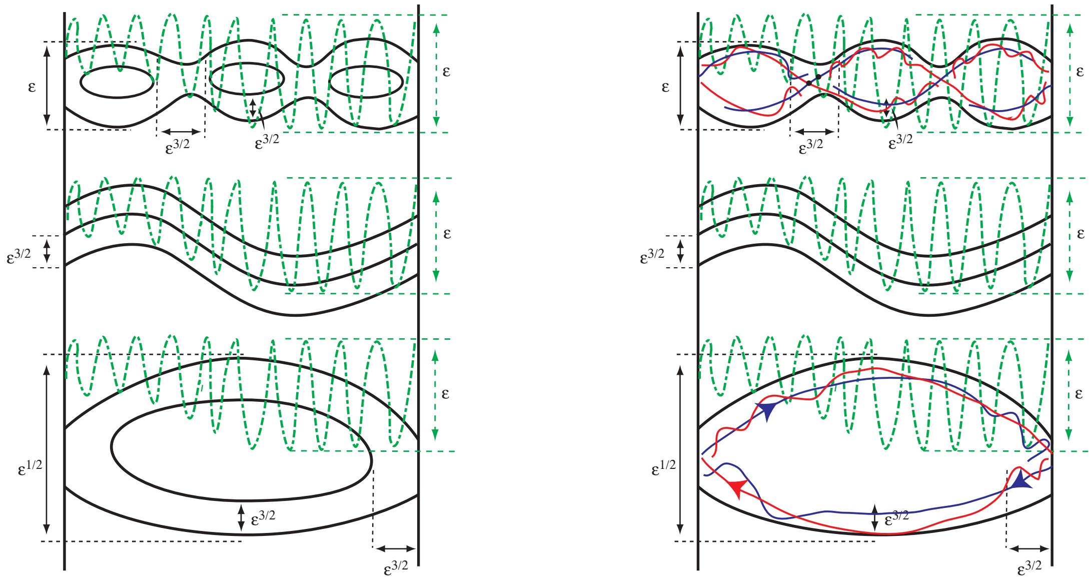

FIGURE 2. Schematic illustration of the main invariant objects in $\tilde{\Lambda}_\varepsilon$ studied in Section 8 as well as their images (dashed) under the scattering map.

Figure 2 represents a 2-dimensional surface of section of the three dimensional manifold $\tilde{\Lambda}_\varepsilon$. We have represented in continuous lines sections of some primary and secondary tori on the left and primary tori and (un)stable manifolds of periodic orbits on the right.

The dashed lines represent the image of some primary tori under the scattering map. They intersect transversally in $\tilde{\Lambda}_\varepsilon$ other primary tori, secondary tori or (un)stable manifolds of periodic orbits.

The transition chains will be constructed by taking chains of objects $\mathcal{V}_i$ such that

$$S(\mathcal{V}_i) \pitchfork_{\tilde{\Lambda}_\varepsilon} \mathcal{V}_{i+1}.$$

Since the scattering map (147) moves amounts $O(\varepsilon)$ in a direction somewhat different to the level sets—which have gaps $O(\varepsilon^{3/2})$—it is clear that one can construct transition chains from the bottom of the figure to the top.

Remark 75. We think that the incorporation of secondary tori and tori of lower dimension in the transition chains is quite natural. The resonant regions are devoid of primary KAM tori. Hence, if there are geometric mechanisms that produce diffusion across resonances, it is clear that they have to involve other geometric objects besides the customary primary KAM tori. Among all the objects that appear, the secondary tori and the periodic orbits are the most prominent, as witnessed by the fact that, very early in the numerical explorations, they were identified and given rather colorful names (e.g. islands, chaotic seas, X-points etc.) and discussed in great detail. □

10.1. **Transition chains.** We will find it convenient to define transition chains as:

Definition 76. *Let $\mathcal{H}$ be a vector field, and $N \in \mathbb{N} \cup \{\infty\}$.*

A transition chain consists of a sequence $\{\mathcal{T}_i\}_{i=0}^N$ of C^1 partially hyperbolic tori, invariant by $\mathcal{H}$ such that:

i) *The motion on each of the tori $\mathcal{T}_i$ is topologically conjugated to a transitive rotation on a torus.*
ii) *There exist orbits $\gamma_i \subset W^{\mathrm{tu}}_{\mathcal{T}_{i-1}} \cap W^{\mathrm{ts}}_{\mathcal{T}_i}$, where $W^{\mathrm{tu,ts}}_{\mathcal{T}}$ are defined in* (139)*.*
iii) *The intersection $W^{\mathrm{tu}}_{\mathcal{T}_{i-1}} \cap W^{\mathrm{ts}}_{\mathcal{T}_i}$ is transversal along the orbit γ_i.*

We emphasize that we are not assuming that the tori have the same dimension, or that they are homotopic.

As the tori are assumed to be partially hyperbolic, there exist δ_i such that

$$
\begin{aligned}
&\operatorname{dist}(\gamma_i(t), \mathcal{T}_{i-1}) \le C e^{-\delta_i |t|} \quad t \le 0, \\
&\operatorname{dist}(\gamma_i(t), \mathcal{T}_i) \le C e^{-\delta_i |t|} \quad t \ge 0.
\end{aligned} \tag{148}
$$

We are not assuming that the δ_i are bounded away from zero uniformly in i. In the case that $\mathcal{H}$ depends on some parameter ε, we are not assuming the δ_i are bounded away from zero uniformly in ε.

Remark 77. It is conceivable that the diffusion based on objects of different dimension has different quantitative properties.

One can argue heuristically that the lower the dimension of the orbit, the faster the transition time since the "ergodization time" is smaller the smaller the dimension of the tori. In this paper, however, we have not investigated these quantitative properties. □

10.2. **The scattering map and the transversality of heteroclinic intersections.** We first turn to proving the result about characterizing the existence of transverse heteroclinic intersections among two different invariant manifolds of $\tilde{\Lambda}_\varepsilon$ in terms of their transversality under the scattering map. This will be a general result that holds for normally hyperbolic invariant manifolds with transversal homoclinic intersections. We will prove it in this generality. For our applications, of course, the normally hyperbolic invariant manifold will be the $\tilde{\Lambda}_\varepsilon$ produced in Section 7.

Lemma 78. *Let $\tilde{\Lambda}$ be a normally hyperbolic invariant manifold. Assume that $W^{\mathrm{s}}_{\tilde{\Lambda}} \pitchfork W^{\mathrm{u}}_{\tilde{\Lambda}}$ along a manifold $\tilde{\gamma}$, and denote by S the scattering map associated to this intersection $\tilde{\gamma}$.*

Let $\mathcal{V}_1, \mathcal{V}_2 \subset \tilde{\Lambda}$ be C^1 submanifolds of $\tilde{\Lambda}$, and assume that $S(\mathcal{V}_1) \pitchfork_{\tilde{\Lambda}} \mathcal{V}_2$. (In particular $\mathcal{V}_1 \cap H_- \neq \emptyset$ and $\mathcal{V}_2 \cap H_+ \neq \emptyset$.)

Then, there exists a heteroclinic trajectory γ_1 such that $\gamma_1 \subset W^{\mathrm{u}}_{\mathcal{V}_1} \pitchfork W^{\mathrm{s}}_{\mathcal{V}_2}$.

Proof. By the definition of the scattering map, we have $W^{\mathrm{u}}_{\mathcal{V}_1} \cap \tilde{\gamma} = W^{\mathrm{s}}_{S(\mathcal{V}_1)} \cap \tilde{\gamma}$.

By the assumption of transversality of $S(\mathcal{V}_1)$ and $\mathcal{V}_2$ in $\tilde{\Lambda}$, we obtain $W^{\mathrm{s}}_{S(\mathcal{V}_1)} \pitchfork_{W^{\mathrm{s}}(\tilde{\Lambda})} W^{\mathrm{s}}_{\mathcal{V}_2}$, and therefore, $W^{\mathrm{s}}_{S(\mathcal{V}_1)} \pitchfork_{\tilde{\gamma}} W^{\mathrm{s}}_{\mathcal{V}_2}$. Hence,

$$W^{\mathrm{u}}_{\mathcal{V}_1} \pitchfork_{\tilde{\gamma}} W^{\mathrm{s}}_{\mathcal{V}_2}.$$

Since $W^{\mathrm{s}}_{\tilde{\Lambda}} \pitchfork W^{\mathrm{u}}_{\tilde{\Lambda}}$ along $\tilde{\gamma}$, we obtain the desired result. □

We now formulate and prove Lemma 81, that will allow us to verify the conditions of Lemma 78 in the case that the manifolds are close to level sets of a function. This Lemma will be useful for us since the objects we have considered before (the primary and secondary tori, the weak stable and unstable manifolds of periodic orbits) are close to being level sets of the averaged Hamiltonian as we have established in Section 8. (See specially Proposition 47, Proposition 50, Theorem 56 and Proposition 66.)

The precise application of Lemma 78 to the case that the manifolds are flat—primary KAM tori far from resonance—is done in Lemma 82, and for non-flat manifolds—primary and secondary tori close to the resonance or weak invariant manifolds of the lower dimensional tori—is done in Lemma 85 for the case of a resonance of order 1 and in Lemma 88 in the case of a resonance of order 2.

Lemma 81 is designed to deal in a unified way with the different type of tori that appear in our problem. That is, the rather flat tori that appear in the non-resonant region or in the resonant regions of order 3 or bigger, and the curved tori that appear near resonances of order 1 or 2.

As we have argued in Remark 55 these two types of tori have different quantitative properties and this leads to the fact that the leading terms in the asymptotics of the intersection could be different. Hence, several of the details and the formulas will be different.

Remark 79. The case of intersections of flat tori is significantly easier and can be dealt with other methods. Indeed, the study of intersections of flat tori is significantly easier that the study of [DLS00, Lemma 4.21]. In the case of [DLS00] the tori were flat but also presented a phase shift which does not appear in our case. However, in this paper, we will present the general approach that works in all cases.

□

Lemma 81 considers a foliation $\mathcal{F}_F$ whose leaves are the level sets of a function F:

$$L_E^F = \{(I, \varphi, s),\ F(I, \varphi, s; \varepsilon) = E\},\ E \in (E_1, E_2)$$

and that are also parameterized as

$$L_E^F = \{(I, \varphi, s)(I_-, I_+) \times \mathbb{T}^2,\ I = \lambda_E(\varphi, s; \varepsilon)\},$$

and it gives criteria to establish that their leaves intersect transversally their images under the scattering map S. Note that these images are contained in the leaves of the function $F \circ S^{-1}$.

Remark 80. Given two foliations $\mathcal{F}$, $\tilde{\mathcal{F}}$, which are $\mathcal{C}^1$-close, we say that $\mathcal{F}$ intersects transversally $\tilde{\mathcal{F}}$—denoted as $\mathcal{F} \pitchfork \tilde{\mathcal{F}}$—when given any leaf of $\mathcal{F}$, we can find another leaf of $\tilde{\mathcal{F}}$ for which there is a non-trivial intersection which is transversal. (There could be other non-transversal intersections)

Note that our use of "foliation intersects transversely "is at variance with standard use in differential topology where it is often taken to mean that, given a leaf of $\mathcal{F}$ and a leaf of $\tilde{\mathcal{F}}$, they either intersect transversally or do not intersect. □

To show the transversality between the foliations $\mathcal{F}_F$ and $\mathcal{F}_{F\circ S^{-1}}$ we only need to obtain lower bounds for the angle between the parameterized surface

$$S(L_E^F) = \{S(\lambda_E(\varphi, s; \varepsilon), \varphi, s)\}$$

and the implicit surface

$$L_{E'}^F = \{(I, \varphi, s),\ F(I, \varphi, s; \varepsilon) = E'\}.$$

More precisely, we will compute $\sin(\alpha)$, where α is the angle between the tangent planes to $L^F_{E'}$ and $S(L^F_E)$.

Recalling that the normal vector to the tangent plane to $L^F_{E'}$ is given by $\frac{\nabla_{I,\varphi,s}F}{|\nabla_{I,\varphi,s}F|}$, and that any vector of the tangent plane to $S(L^F_E)$ is written as $D_v(S\circ(\lambda_E,\mathrm{Id},\mathrm{Id})) = D\,(S\circ(\lambda_E,\mathrm{Id},\mathrm{Id}))\,v$, we obtain:

$$\sin(\alpha) = \max_{v\in\mathbb{R}^2} \frac{|D_vF\circ S\circ(\lambda_E,\mathrm{Id},\mathrm{Id})|}{|(\nabla_{I,\varphi,s}F)\circ(\lambda_E,\mathrm{Id},\mathrm{Id})|\,|D_vS\circ(\lambda_E,\mathrm{Id},\mathrm{Id})|}. \tag{149}$$

Hence, we will obtain lower bounds for the angles taking vectors v that make the computations in the right hand side of (149) simple.

Lemma 81. *Let*

$$F:\mathcal{A}=(I_1,I_2)\times\mathcal{J}\times(-\varepsilon_0,\varepsilon_0)\to\mathbb{R}$$

be a $\mathcal{C}^r$ function, $r\ge 2$, where $\mathcal{J}\subset\mathbb{T}^2$ is an open set.

Assume that for any $(I,\varphi,s)\in(I_1,I_2)\times\mathcal{J}$, the equation $F(I,\varphi,s;\varepsilon)=E$, for $E\in(E_1,E_2)=F(\mathcal{A})$ defines a smooth surface given as a graph

$$I=\lambda_E(\varphi,s;\varepsilon).$$

Let S be the scattering map which has been computed to first order in (147) *and assume that there exists an open set $\mathcal{J}'\subset\mathcal{J}$, $\mathcal{J}'\neq\emptyset$, and a constant $C>0$ independent of ε and E, such that for any $(\varphi,s)\in\mathcal{J}'$ one has,*

$$\frac{|\nabla_{\varphi,s}(F\circ S\circ(\lambda_E,\mathrm{Id},\mathrm{Id})(\varphi,s))|}{|\nabla_{I,\varphi,s}F(\lambda_E(\varphi,s;\varepsilon),\varphi,s)|}\ge C\varepsilon. \tag{150}$$

Then, if we denote by $\mathcal{F}_F$ the foliation given by

$$\{(I,\varphi,s),\ F(I,\varphi,s;\varepsilon)=E,\ E\in(E_1,E_2)\}=\cup_{E\in(E_1,E_2)}L^F_E,$$

we have:

$$\mathcal{F}_F\pitchfork\mathcal{F}_{F\circ S^{-1}}.$$

More precisely, there exists a constant C', independent of ε and E, such that the angle between the surfaces $S(L^F_E)$ and $L^F_{E'}$, at the intersection can be bounded from below by $C'\varepsilon$.

Proof.

To show the transversality between the foliations $\mathcal{F}_F$ and $\mathcal{F}_{F\circ S^{-1}}$, we only need to obtain lower bounds for the angle (149) between the parameterized surface $S(\lambda_E(\varphi,s;\varepsilon),\varphi,s)$ and the implicit surface $F(I,\varphi,s;\varepsilon)=E'$.

By formula (147), we have $S=\mathrm{Id}+\varepsilon S_1+O_{\mathcal{C}^1}(\varepsilon^{1+\varrho})$. Hence, there exists a constant $\bar{C}$, independent of ε such that

$$|D_vS\circ(\lambda_E,\mathrm{Id},\mathrm{Id})|\le\bar{C}\,|v|$$

and formula (149) is bounded if (150) is verified. Then, in order to obtain that the foliations intersect transversally we only need to assume condition (150), and we obtain that the angle of intersection is bounded from below by $C'\varepsilon$, where C' is some suitable constant. $\square$

Now, we apply Lemma 81 to study the non-resonant region $\mathcal{S}^L$ and the resonant regions $\mathcal{S}^{\mathcal{R}_j}$, for $j \geq 3$, where the tori, given in Propositions 47 and 50 —see also Remarks 48 and 51—, are rather "flat". That is, we can take $F(I, \varphi, s; \varepsilon) = I + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)$ in the previous Lemma 81, and $\lambda_E(\varphi, s; \varepsilon) = E + U_E(\varphi, s; \varepsilon)$, where $|U_E|_{\mathcal{C}^2} \leq \text{cte.}\,\varepsilon$ as given in Remarks 48 and 51.

Lemma 82. *Consider a foliation $\mathcal{F}_F$, contained in a connected component of the non resonant region $\mathcal{S}^L$ defined in* (56) *(or of the resonant regions $\mathcal{S}^{\mathcal{R}_j}$, defined in* (69)*, for $j \geq 3$).*

Assume that the function F is of the form $F(I, \varphi, s; \varepsilon) = I + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)$, so that equation $F(I, \varphi, s; \varepsilon) = E$ defines a smooth surface given as a graph $I = \lambda_E(\varphi, s; \varepsilon) = E + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)$, for $(\varphi, s) \in \mathbb{T}^2$.

Assume also that the reduced Poincaré function $\mathcal{L}^*$ defined in (145) *verifies, for any value of $(I, \varphi, s) \in H_- \cap \mathcal{S}^L$ (respectively for $(I, \varphi, s) \in H_- \cap \mathcal{S}^{\mathcal{R}_j}$) that the function*

$$\theta \mapsto \frac{\partial \mathcal{L}^*}{\partial \theta}(E, \theta) \tag{151}$$

is negative and non-constant for $\theta \in \mathcal{J}_E^*$.

Then, the foliations $\mathcal{F}_F$ and $\mathcal{F}_{F \circ S^{-1}}$ intersect transversally.

More precisely, there exist constants, C', C'', independent of ε and E, such that any surface $S(L_E^F)$ intersects at some point the surface $L_{E'}^F$ such that $0 < E' - E \leq C''\varepsilon$.

The angle between the surfaces $S(L_E^F)$ and $L_{E'}^F$ at the intersection can be bounded from below by $C'\varepsilon$.

Remark 83. We know, by Propositions 47 and 50, that the gaps between two consecutive tori are, at least, of order $\bar{I} - I \simeq \varepsilon^{3/2}$. Then, when we apply Lemma 82 to these tori we obtain that the image under the scattering map of a torus in this region given by $I = I_0 + \mathrm{O}(\varepsilon)$, intersect transversally another torus given by $I = \bar{I}_0 + \mathrm{O}(\varepsilon)$ with $\bar{I}_0 > I_0$, and $\bar{I}_0 - I_0 = \mathrm{O}(\varepsilon)$. $\square$

Proof. We apply Lemma 81, with $F(I,\varphi,s;\varepsilon) = I + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)$, and $\lambda_E(\varphi,s;\varepsilon) = E + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)$. To check inequality (150) we use formula (147), so that $S = \mathrm{Id} + \varepsilon S_1 + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})$. Hence, we compute:

$$\begin{aligned}
&F\circ S\circ(\lambda_E,\mathrm{Id},\mathrm{Id})(\varphi,s)\\
&\quad= F(\lambda_E(\varphi,s;\varepsilon),\varphi,s;\varepsilon)\\
&\qquad+\varepsilon\nabla_{I,\varphi,s}F(\lambda_E(\varphi,s;\varepsilon),\varphi,s;\varepsilon)\left(S_1(\lambda_E(\varphi,s;\varepsilon),\varphi,s;\varepsilon)+\mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho})\right)\\
&\qquad+\mathrm{O}_{\mathcal{C}^1}(\varepsilon^2)\\
&\quad= E-\varepsilon\frac{\partial\mathcal{L}^*}{\partial\theta}(\lambda_E(\varphi,s;\varepsilon),\varphi-\lambda_E(\varphi,s;\varepsilon)s)+\mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})\\
&\quad= E-\varepsilon\frac{\partial\mathcal{L}^*}{\partial\theta}(E,\varphi-Es)+\mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho}),
\end{aligned}$$

and $|\nabla_{I,\varphi,s}F(\lambda_E(\varphi,s;\varepsilon),\varphi,s;\varepsilon)| = 1 + \mathrm{O}_{\mathcal{C}^1}(\varepsilon)$. If for any value of (E,θ), the function $\frac{\partial\mathcal{L}^*}{\partial\theta}(E,\theta)$ is non-constant, there exists an interval $\bar{\mathcal{J}}_E \subset \mathcal{J}^*_E$ where

$$\left|\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}\right| \geq C > 0.$$

Then, hypothesis (150) is verified in $\bar{\mathcal{J}}_E$ and Lemma 81 applies. On the other hand, as $\frac{\partial\mathcal{L}^*}{\partial\theta}(E,\theta) < 0$, the surface $S(L^F_E)$ intersect surfaces $L^F_{E'}$, for $E' > E$, and $E' - E = \mathrm{O}(\varepsilon)$. □

In order to obtain an analogous result to the previous Lemma in the regions close to the resonances of first order, we will use the following technical Lemma:

Lemma 84. *Let $a(\theta)$, $b(\theta)$ be functions of class $\mathcal{C}^r$, $r \geq 0$, such that the function $a(\theta)/b(\theta)$ is not constant in some interval $\mathcal{J}$.*

Then, there exists a constant $\tilde{C} > 0$ and two intervals $\mathcal{J}_1$, $\mathcal{J}_2$ subsets of $\mathcal{J}$ such that, given any value $(x,y) \in \mathbb{R}^2$, one can choose θ which belongs to one of the two intervals $\mathcal{J}_1$ $\mathcal{J}_2$ and such that:

$$|a(\theta)x + b(\theta)y| \geq \tilde{C}(|x| + |y|) \tag{152}$$

Proof. If $a(\theta)/b(\theta)$ is not constant in $\mathcal{J}$ there exist at least two values $\theta_i \in \mathcal{J}$, $i = 1, 2$ (and intervals $\mathcal{J}_i$ around them) such that the matrix

$$A = \begin{pmatrix} a(\theta_1) & b(\theta_1) \\ a(\theta_2) & b(\theta_2) \end{pmatrix}$$

is invertible. Then, given any vector $z = (x,y)$, we have that

$$|Az| \geq \frac{|z|}{|A^{-1}|}$$

where $|\cdot|$ stands for the sup norm. Given z, we choose θ_i, for $i = 1$ or $i = 2$, such that $|Az| = |a(\theta_i)x + b(\theta_i)y|$, and taking into account that we are dealing with continuous functions we obtain that:

$$|a(\theta)x + b(\theta)y| \geq \frac{1}{|A^{-1}|}\,|z|$$

in $\mathcal{J}_i$. The proof finishes taking $\tilde{C} < \frac{1}{|A^{-1}|}$. □

Now, we will apply Lemma 81 to study the resonant regions $\mathcal{S}^{\mathcal{R}_j}$, $j = 1, 2$. In these regions, the tori and the stable and unstable manifolds of the periodic orbit, are not flat as we showed in Theorem 56 and Proposition 66. The fact that the tori are not flat has the consequence that the dominant effect of comparing a torus with the torus in the image of the scattering map will include some extra terms.

We recall that in Theorem 56 we showed that the invariant objects in the resonant regions $\mathcal{S}^{\mathcal{R}_j}$, $j = 1, 2$ are given very approximately by the level sets of the Hamiltonian $K_0(y, x; \varepsilon)$ in (82), when written in the averaged variables (y, x, s).

The relation between these variables and the original ones (I, φ, s) are the changes of variables given in Proposition 28, Theorem 35 and in (73), (76), and is given in first order by

$$y = k_0 I + l_0 + \mathrm{O}_{\mathcal{C}^2}(\varepsilon).$$

Then, using (80), we obtain that all these objects are given by the level sets of a function:

$$(153)\quad F(I, \varphi, s; \varepsilon) = \frac{(k_0 I + l_0 + \mathrm{O}_{\mathcal{C}^2}(\varepsilon))^2}{2}(1 + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)) + \mathrm{O}_{\mathcal{C}^2}(\varepsilon^j),$$

where $j = 1, 2$ is the order of the resonance.

Moreover, in Corollary 57 and Proposition 66 the tori and the (un)stable manifolds of periodic orbits were written as graphs in the variables (y, x, s) as functions $y = \mathcal{Y}_\pm(x, E) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{3/2})$, and in remarks 58 and 67 we have obtained the corresponding expressions in the original variables (I, φ, s).

The dominant terms in F and in the expressions of these invariant objects will be different wether the resonance is of order 1 or 2. The hypothesis of the following Lemmas 85 and 88 are tailored to apply to resonances of order 1 or 2.

Lemma 85. *Let us consider a foliation $\mathcal{F}_F$ in a connected component of the resonant region $\mathcal{S}^{\mathcal{R}_1}$, defined in* (69)*, where $\mathcal{R}_1$ is given in* (20)*. More precisely, consider a resonance $-l_0/k_0 \in \mathcal{R}_1$, and in the region*

$$\{(I, \varphi, s) \in [l_0/k_0 - L, l_0/k_0 + L] \times \mathbb{T}^2\},$$

the function F is of the form

$$F(I,\varphi,s;\varepsilon) = \frac{(k_0 I + l_0)^2}{2} + \mathrm{O}_{\mathcal{C}^2}(\varepsilon),$$

and, for some $0<\rho<\pi$, and for some range of energies $-c_4\varepsilon \le E \le c_2 L$, the equation $F(I,\varphi,s;\varepsilon)=E$, defines two smooth surfaces given as a graph $I=\lambda_E^{\pm}(\varphi,s;\varepsilon)$ with:

(154)

$$\begin{aligned}\lambda_E^{\pm}(\varphi,s;\varepsilon) &= -\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_{\pm}(\theta,E) + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)\\ &= -\frac{l_0}{k_0} \pm \frac{1}{k_0}(1+\varepsilon b)\ell(\theta,E) + \frac{\varepsilon}{k_0}\tilde{\mathcal{Y}}_{\pm}\left(\ell(\theta,E)\right) + \mathrm{O}_{\mathcal{C}^2}(\varepsilon),\end{aligned}$$

for $\rho \le \theta := k_0\varphi + l_0 s \le 2\pi - \rho$, where $\ell(\theta,E) = \sqrt{2(E-\varepsilon U(\theta;\varepsilon))}$ with $U(\theta;\varepsilon)$ defined in (78) *and $\tilde{\mathcal{Y}}_{\pm}$ is given in Lemma 60, with $\delta=\varepsilon$.*

Given the reduced Poincaré function $\mathcal{L}^*$, defined in (145), *for $(I,\varphi,s)\in H_-\cap \mathcal{S}^{\mathcal{R}_1}$, let us assume that the function*

$$\theta \mapsto \frac{\partial \mathcal{L}^*}{\partial\theta}(I,\theta) \tag{155}$$

is non-constant and negative for $\theta\in\mathcal{J}_I^*$.

Assume, moreover, the following hypothesis, which is **H5”** *in Theorem 7.*

The function

$$\theta \mapsto \frac{k_0 U'(\theta;0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) + 2U(\theta;0)\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0})}{2\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0})} \tag{156}$$

is non-constant.

Then, the foliations $\mathcal{F}_F$ and $\mathcal{F}_{F\circ S^{-1}}$ intersect transversally.

More precisely, there exist constants, C', C'', independent of ε and E, such that:

1. *Any surface $S(L_E^{F,-})$ intersects at some point the surface $L_{E'}^{F,-}$ such that $E'<E$ and $E-E' \le C''\varepsilon\max\left(|E|^{1/2},\varepsilon^{1/2}\right)$*
2. *Any surface $S(L_E^{F,+})$ intersects at some point the surface $L_{E'}^{F,+}$ such that $E'>E$ and $E'-E \le C''\varepsilon\max\left(|E|^{1/2},\varepsilon^{1/2}\right)$*

The angle between the surfaces $S(L_E^{F,\pm})$ and $L_{E'}^{F,\pm}$ at the intersection can be bounded from below by $C'\varepsilon$.

Remark 86. We know, by Theorem 56 that all the tori in the resonant region are given by formulas (154) for $E=E_i, F_i, G$ and verify that

$$\begin{aligned}
|E_i - E_{i+1}| &\le \varepsilon^{\frac{3}{2}+\frac{1}{2}} \le \max(\sqrt{E_i}, \varepsilon^{1/2}),\\
|E_{l_E} - F_1| &\le \varepsilon^{\frac{3}{2}+\frac{1}{2}} \le \max(\sqrt{E_{l_E}}, \varepsilon^{1/2}),\\
|F_i - F_{i+1}| &\le \varepsilon^{\frac{3}{2}+\frac{1}{2}} \le \max(\sqrt{F_i}, \varepsilon^{1/2})\\
|F_{l_F} - G| &\le \varepsilon^{\frac{3}{2}+\frac{1}{2}} \le \max(\sqrt{F_{l_F}}, \varepsilon^{1/2}).
\end{aligned}$$

and that by and Corollary 57 and Remark 58 that, when they are written as graphs, the distance between to consecutive tori is of order $\mathrm{O}(\varepsilon^{3/2})$.

Then, when we apply Lemma 85 to these tori we obtain that the image under the scattering map of a torus in this region intersect transversally another torus. $\square$

Proof. It suffices to apply Lemma 81 to the function

$$F(I, \varphi, s; \varepsilon) = \frac{(k_0 I + l_0)^2}{2} + \mathrm{O}_{\mathcal{C}^2}(\varepsilon),$$

and $\lambda_E^{\pm}$ given by (154).

Note that to study transversality of the foliations, we can consider E fixed since E is only a label for the leaves.

Since we consider E fixed, the analysis of dominant terms will only involve estimating derivatives with respect to the angle variables.

We will also see that the angle variables will enter in the dominant terms only trough $\theta = k_0\varphi + l_0 s$ (the resonant angle).

In order to check inequality (150) we use formula (147), so that $S = \mathrm{Id} + \varepsilon S_1 + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})$, and that, by (98), $\lambda_E^{\pm}$ is a bounded function, and its derivatives with respect to φ, s are of order $\mathrm{O}(\varepsilon^{1/2})$. Hence, we can compute:

$$
\begin{aligned}
F \circ S \circ & (\lambda_E^{\pm}, \mathrm{Id}, \mathrm{Id})(\varphi, s) \\
&= F(\lambda_E^{\pm}(\varphi, s; \varepsilon), \varphi, s; \varepsilon) + \varepsilon \nabla_{I,\varphi,s} F(\lambda_E^{\pm}(\varphi, s; \varepsilon), \varphi, s; \varepsilon) \cdot \\
&\big(S_1(\lambda_E^{\pm}(\varphi, s; \varepsilon), \varphi, s) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho})\big) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^2) \\
&= E - \varepsilon k_0 (k_0 \lambda_E^{\pm}(\varphi, s; \varepsilon) + l_0) \cdot \\
&\left(\frac{\partial \mathcal{L}^*}{\partial \theta} (\lambda_E^{\pm}(\varphi, s; \varepsilon), \varphi - \lambda_E^{\pm}(\varphi, s; \varepsilon) s) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho}) \right) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^2) \\
&= E - \varepsilon k_0 \mathcal{Y}_{\pm}(\theta, E) \cdot \\
&\left(\frac{\partial \mathcal{L}^*}{\partial \theta} \left(-\frac{l_0}{k_0} + \frac{1}{k_0} \mathcal{Y}_{\pm}(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \frac{1}{k_0} \mathcal{Y}_{\pm}(\theta, E) \right) s \right) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho}) \right) \\
&+ \mathrm{O}_{\mathcal{C}^1}(\varepsilon^2) \\
&= E \mp \varepsilon k_0 (1 + \varepsilon b) \ell(\theta, E) \cdot \\
&\left(\frac{\partial \mathcal{L}^*}{\partial \theta} \left(-\frac{l_0}{k_0} \pm \frac{(1 + \varepsilon b)}{k_0} \ell(\theta, E), \varphi - \left(-\frac{l_0}{k_0} \pm \frac{(1 + \varepsilon b)}{k_0} \ell(\theta, E) \right) s \right) \right. \\
&\left. + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho}) \right) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^2) \\
&E \mp \varepsilon \mathcal{M}(\theta; \varepsilon) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})
\end{aligned}
$$

Now, we compute the main part of the function $\mathcal{M}$

$$
\begin{aligned}
&\mathcal{M}(\theta; \varepsilon) = k_0 \ell(\theta, E) \cdot \\
&\frac{\partial \mathcal{L}^*}{\partial \theta} \left(-\frac{l_0}{k_0} + \frac{1}{k_0} \ell(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \frac{1}{k_0} \ell(\theta, E) \right) s \right)
\end{aligned}
$$

The analysis of the function $\mathcal{M}$ will be done differently whether we are close to the resonance or whether we are reasonably far because in these cases the dominant terms will be different. We choose any ν such that $0 \le \nu < 1$, and we will consider the following two cases:

Close to the resonance: $-c_4 \varepsilon \le E \le \tilde{c} \varepsilon^{\nu}$.

In this region, we have that $|\ell(\cdot, E)|_{\mathcal{C}^1} \le \varepsilon^{\nu/2}$, and then, by bounds (98), the dominant terms in $\mathcal{M}$ are

$$
\begin{aligned}
&\mathcal{M}(\theta; \varepsilon) = \\
&= k_0 \sqrt{2(E - \varepsilon U(\theta; 0))} \left(\frac{\partial \mathcal{L}^*}{\partial \theta} \left(-\frac{l_0}{k_0}, \frac{\theta}{k_0}\right) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\nu/2}) \right).
\end{aligned}
$$

To apply Lemma 81 we need to check (150), that is, it suffices to show that we can bound from below the derivative of this function divided by $k_0 \lambda_E^{\pm}(\varphi, s; \varepsilon) + l_0$. Computing this derivative, we obtain:

$$
\begin{aligned}
&\frac{\partial}{\partial\theta}\left(\mathcal{M}(\theta;\varepsilon)\right) = \\
=&\frac{k_0}{\sqrt{2(E-\varepsilon U(\theta;0))}}\Bigg(2E\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \\
&-\varepsilon\left[k_0U'(\theta;0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})+2U(\theta;0)\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right]\Bigg) \\
&+\sqrt{2(E-\varepsilon U(\theta;0))}\,\mathrm{O}_{\mathcal{C}^0}(\varepsilon^{\nu/2})+\mathrm{O}_{\mathcal{C}^0}(\varepsilon^{(1+\nu)/2}),
\end{aligned}
$$

and then,

$$
\begin{aligned}
&\frac{\frac{\partial}{\partial\theta}\left(\mathcal{M}(\theta;\varepsilon)\right)}{k_0\lambda_E^{\pm}(\varphi,s;\varepsilon)+l_0} = \\
&\frac{\pm k_0}{2(E-\varepsilon U(\theta;0))}\Bigg(2E\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \\
&-\varepsilon\left[k_0U'(\theta;0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})+2U(\theta;0)\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right]\Bigg) \\
&+\mathrm{O}_{\mathcal{C}^0}(\varepsilon^{\nu/2}),
\end{aligned}
$$

Now, we apply Lemma 84, with $x=-\varepsilon$, $y=E$, and

$$
\begin{aligned}
a(\theta) &= k_0U'(\theta;0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})+2U(\theta;0)\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \\
b(\theta) &= 2\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0}).
\end{aligned}
$$

Using Lemma 84, and that

$$|2(E-\varepsilon U(\theta,0))|\leq \text{cte.}\,(|E|+|\varepsilon|),$$

we obtain that for θ either in $\mathcal{J}_1$, $\mathcal{J}_2$, where $\mathcal{J}_i\subset\mathcal{J}^*_{-l_0/k_0}$, $i=1,2$

$$\left|\frac{\frac{\partial}{\partial\theta}\left(\mathcal{M}(\theta;\varepsilon)\right)}{k_0\lambda_E^{\pm}(\varphi,s;\varepsilon)+l_0}\right|\geq\tilde{C}.$$

Consequently, the angle of intersection can be bounded from below by $C'\varepsilon$, for some suitable constant independent of ε.

To identify more precisely which surfaces will intersect $S(L_E^{F,\pm})$, we only need to observe that the function $\mathcal{M}(\theta;\varepsilon)$ is given approximately by

$$k_0\sqrt{2(E-\varepsilon U(\theta;0))}\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0}),$$

and then it is a negative function. On the other hand we have that:

$$F \circ S \circ (\lambda_E^{\pm}, \mathrm{Id}, \mathrm{Id})(\varphi, s) \simeq E \mp \varepsilon \mathcal{M}(\theta; \varepsilon),$$

then the surface $S(L_E^{F,-})$ intersect surfaces $L_{E'}^{F,-}$ with $E' < E$. (Note that if $E' < E$ then $\lambda_E^-(\varphi, s; \varepsilon) < \lambda_{E'}^-(\varphi, s; \varepsilon)$). Moreover, one has that there exists some constant C'' such that

$$\max_{\theta \in \mathcal{J}_1 \cup \mathcal{J}_2} |\mathcal{M}(\varphi, s; \varepsilon)| \geq C'' \max(E^{1/2}, \varepsilon^{1/2}),$$

and applying Lemma 150 we obtain the desired result.

An analogous result is obtained for λ_E^+, with $E' > E$.

Far from the resonance: $\tilde{c}\varepsilon^\nu \leq E \leq c_2 L$.

This case is analogous to the non resonant region, because in this case,

$$\begin{aligned} \ell(\theta, s) &= \sqrt{2(E - \varepsilon U(\theta; \varepsilon))} = \sqrt{2E}\sqrt{1 - \frac{\varepsilon}{E} U(\theta; \varepsilon)} \\ &= \sqrt{2E}(1 + \mathrm{O}_{\mathcal{C}^2}(\varepsilon^{1-\nu})), \end{aligned}$$

consequently, the function $\mathcal{M}$ becomes

$$\begin{aligned} &\mathcal{M}(\theta; \varepsilon) \\ &= k_0 \sqrt{2E} \frac{\partial \mathcal{L}^*}{\partial \theta} \left(-\frac{l_0}{k_0} + \sqrt{2E}, \varphi - (-\frac{l_0}{k_0} + \sqrt{2E}) s \right) + \mathrm{O}_{\mathcal{C}^2}(\varepsilon^{1-\nu}). \end{aligned}$$

Then, if the function $\mathcal{L}^*(I, \theta)$ is not constant, we can apply Lemma 81, to get the desired result.

□

Remark 87. We observe that the transversality lemmas 82 and 85 use some non-degeneracy hypothesis. Hypothesis (151) and (155) refer to the reduced Poincaré function $\mathcal{L}^*$, and therefore to $\mathcal{L}$ (see (145)), and both are contained in Hypothesis **H4** of Theorem 7. The hypothesis that the function (156) is non-constant also involves the function $U(\theta; \varepsilon)$, which comes from the expression of the Hamiltonian reduced to $\tilde{\Lambda}_\varepsilon$ near the resonance $I = -\frac{l_0}{k_0}$. This Hypothesis is now explicitly stated, and is the so called **H5"** in Theorem 7. □

Now we turn to study resonances of order 2.

Lemma 88. *Consider a foliation $\mathcal{F}_F$ in a connected component of the resonant region $\mathcal{S}^{\mathcal{R}_2}$, defined in* (69), *where $\mathcal{R}_2$ is given in* (21). *More precisely, consider a resonance $-l_0/k_0 \in \mathcal{R}_2$, and in the region*

$$\{(I, \varphi, s) \in [l_0/k_0 - L, l_0/k_0 + L] \times \mathbb{T}^2\},$$

the function F is of the form

$$F(I,\varphi,s;\varepsilon)=\frac{y^2}{2}(1+\mathrm{O}_{\mathcal{C}^2}(\varepsilon))+\mathrm{O}_{\mathcal{C}^2}(\varepsilon^2),$$

where $y=y(I,\varphi,s;\varepsilon)$ is the variable defined in this resonant region through the changes of variables given in Proposition 28, Theorem 35 and in 73, 76, given, up to first order, by

$$\begin{aligned} y=&k_0I+l_0+\mathrm{O}_{\mathcal{C}^2}(\varepsilon),\\ x=&\theta+\mathrm{O}_{\mathcal{C}^2}(\varepsilon),\quad \theta=k_0\varphi+l_0s. \end{aligned}\tag{157}$$

For some $0<\rho<\pi$, and for some range of energies $-c_4\varepsilon^2\le E\le c_2L$, the equation $F(I,\varphi,s;\varepsilon)=E$, defines two smooth surfaces given as graphs $I=\lambda_E^{\pm}(\varphi,s;\varepsilon)$. These functions $\lambda_E^{\pm}$ are related by the change (157) *with $f_E^{\pm}(x,s;\varepsilon)$, which are the graphs of the surfaces in the variables (y,x,s):*

$$\lambda_E^{\pm}(\varphi,s;\varepsilon)=-\frac{l_0}{k_0}+\frac{1}{k_0}f_E^{\pm}(x,s;\varepsilon)+\mathrm{O}_{\mathcal{C}^2}(\varepsilon),\tag{158}$$

and

$$\begin{aligned} f_E^{\pm}(x,s;\varepsilon)&=\mathcal{Y}_{\pm}(x,E)+\mathrm{O}_{\mathcal{C}^2}(\varepsilon^{3/2})\\ &=\pm(1+\varepsilon b)\ell(x,E)+\varepsilon\tilde{\mathcal{Y}}_{\pm}\left(\ell(x,E)\right)+\mathrm{O}_{\mathcal{C}^2}(\varepsilon^{3/2}), \end{aligned}\tag{159}$$

for $\rho\le\theta\le 2\pi-\rho$, where $\ell(x,E)=\sqrt{2(E-\varepsilon^2U(x;\varepsilon))}$ with $U(x;\varepsilon)$ defined in (78) *and $\tilde{\mathcal{Y}}_{\pm}$ is given in Lemma 60 with $\delta=\varepsilon^2$.*

Given the reduced Poincaré function $\mathcal{L}^*$, defined in (145)*, for $(I,\varphi,s)\in H_-\cap\mathcal{S}^{\mathcal{R}_2}$, let us assume that the function*

$$\theta\mapsto\frac{\partial\mathcal{L}^*}{\partial\theta}(I,\theta)\tag{160}$$

is non-constant and negative for $\theta\in\mathcal{J}_I^*$.

Assume moreover the following hypothesis, which is **H5''** *in Theorem 7.*

There exists a constant C, independent of E and ε, and three intervals $\mathcal{J}_i\subset\mathcal{J}^*_{-l_0/k_0}$, $i=1,2,3$, such that, given any E,ε in this region ($-c_4\varepsilon^2\le E\le c_2L$), one can choose $i=1,2,3$ in such a way that

$$\begin{aligned} &\Big|\frac{k_0}{2(E-\varepsilon^2U(\theta;0))}\left(2E\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right.\\ &-\varepsilon^2\left[k_0U'(\theta,0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})+2U(\theta;0)\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right]\\ &\left.\pm\varepsilon\sqrt{2(E-\varepsilon^2U(\theta;0))}\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right)\Big|\ge C \end{aligned}\tag{161}$$

Then, the foliations $\mathcal{F}_F$ *and* $\mathcal{F}_{F\circ S^{-1}}$ *intersect transversally.*

More precisely, there exist constants, C', C'', *independent of* ε *and* E, *such that:*

(1) *Given a surface* $S(L_E^{F,-})$ *we have one of the following cases:*
   (a) $S(L_E^{F,-})$ *intersects at some point the surface* $L_{E'}^{F,-}$, *with* $E' < E$, *and* $E - E' \le C''\varepsilon \max\left(|E|^{1/2}, \varepsilon\right)$,
   (b) $S(L_E^{F,-})$ *intersects at some point the surface* $L_{E'}^{F,+}$, *with* $E' \ge E$, *and* $E' - E \le C''\varepsilon \max\left(|E|^{1/2}, \varepsilon\right)$.

(2) *Any surface* $S(L_E^{F,+})$ *intersects at some point the surface* $L_{E'}^{F,+}$, *with* $E' \ge E$, *and* $E' - E \le C''\varepsilon \max\left(|E|^{1/2}, \varepsilon\right)$.

In all the cases the angle between these surfaces at the intersection can be bounded from below by $C'\varepsilon$.

Remark 89. We know, by Theorem 56 that all the tori in the resonant region are given by formulas (154) for $E = E_i, F_i, G$ and verify that

$$\begin{aligned}
|E_i - E_{i+1}| &\le \varepsilon^{\frac{3}{2}+1} \le \max(\sqrt{E_i}, \varepsilon), \\
|E_{l_E} - F_1| &\le \varepsilon^{\frac{3}{2}+1} \le \max(\sqrt{E_{l_E}}, \varepsilon), \\
|F_i - F_{i+1}| &\le \varepsilon^{\frac{3}{2}+1} \le \max(\sqrt{F_i}, \varepsilon) \\
|F_{l_F} - G| &\le \varepsilon^{\frac{3}{2}+1} \le \max(\sqrt{F_{l_F}}, \varepsilon).
\end{aligned}$$

and that by and Corollary 57 and Remark 58 that, when they are written as graphs, the distance between to consecutive tori is of order $\mathrm{O}(\varepsilon^{3/2})$.

Then, when we apply Lemma 88 to these tori we obtain that the image under the scattering map of a torus in this region intersect transversally another torus. □

Proof. It suffices to apply Lemma 81 to the function

$$F(I, \varphi, s) = \frac{y^2}{2}(1 + \mathrm{O}(\varepsilon)) + \mathrm{O}_{\mathcal{C}^2}(\varepsilon^2),$$

and $\lambda_E^{\pm}$ given by (158) and (159).

In order to check inequality (150) we use formula (147), so that $S = \mathrm{Id} + \varepsilon S_1 + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\varrho})$, and that, by (98), $\lambda_E^{\pm}$ is a bounded function and its derivatives with respect to φ, s are of order $\mathrm{O}(\varepsilon)$.

We will also use that $\nabla_{I,\varphi,s} F(I, \varphi, s; \varepsilon) = y(I, \varphi, s; \varepsilon)(1 + \mathrm{O}_{\mathcal{C}^2}(\varepsilon))(k_0, 0, 0) + \mathrm{O}_{\mathcal{C}^2}(\varepsilon^2)$, and analogous estimates for the second derivatives of F.

Hence, we can compute, as in Lemma 85:

$$
\begin{aligned}
&F \circ S \circ (\lambda_E^{\pm}, \mathrm{Id}, \mathrm{Id})(\varphi, s) \\
&\quad = E - \varepsilon(k_0 + \mathrm{O}_{\mathcal{C}^2}(\varepsilon)) f_E^{\pm}(x, s; \varepsilon)\cdot \\
&\quad \left(\frac{\partial \mathcal{L}^*}{\partial \theta}(\lambda_E^{\pm}(\varphi, s; \varepsilon), \varphi - \lambda_E^{\pm}(\varphi, s; \varepsilon)s) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho})\right) \\
&\quad + k_0^2 \varepsilon^2 \left(\frac{\partial \mathcal{L}^*}{\partial \theta}(\lambda_E^{\pm}(\varphi, s; \varepsilon), \varphi - \lambda_E^{\pm}(\varphi, s; \varepsilon)s) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho})\right)^2 + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^3) \\
&\quad = E - \varepsilon k_0 \mathcal{Y}_{\pm}(\theta, E)\cdot \\
&\quad \left(\frac{\partial \mathcal{L}^*}{\partial \theta}\left(-\frac{l_0}{k_0} + \mathcal{Y}_{\pm}(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \mathcal{Y}_{\pm}(\theta, E)\right) s\right) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho})\right) \\
&\quad + k_0 \varepsilon^2 \left(\frac{\partial \mathcal{L}^*}{\partial \theta}\left(-\frac{l_0}{k_0} + \mathcal{Y}_{\pm}(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \mathcal{Y}_{\pm}(\theta, E)\right) s\right)\right)^2 \\
&\quad + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{5/2}, \varepsilon^{2+\varrho}) \\
&\quad = E - \varepsilon \mathcal{M}_{\pm}(\theta; \varepsilon) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{5/2}, \varepsilon^{2+\varrho})
\end{aligned}
$$

Now, we compute the main part of the functions $\mathcal{M}_{\pm}$

$$
\begin{aligned}
&\mathcal{M}_{\pm}(\theta; \varepsilon) = k_0 \mathcal{Y}_{\pm}(\theta, E) \cdot \\
&\left(\frac{\partial \mathcal{L}^*}{\partial \theta}\left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_{\pm}(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_{\pm}(\theta, E)\right) s\right)\right. \\
&\left. + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\varrho})\right) \\
&-\varepsilon k_0^2 \left(\frac{\partial \mathcal{L}^*}{\partial \theta}\left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_{\pm}(\theta, E), \varphi - \left(-\frac{l_0}{k_0} + \frac{1}{k_0}\mathcal{Y}_{\pm}(\theta, E)\right) s\right)\right)^2
\end{aligned}
$$

The analysis of the functions $\mathcal{M}_{\pm}$ will be done differently whether we are close to the resonance or whether we are reasonably far. We choose any ν such that $1 \leq \nu < 2$, and we will consider the following two cases:

Close to the resonance: $-c_4\varepsilon^2 \leq E \leq \tilde{c}\varepsilon^{\nu}$.
In this region, we have that $|\ell(\cdot, E)|_{\mathcal{C}^1} \leq \varepsilon^{\nu/2}$, and

$$\mathcal{Y}_{\pm}(\theta, E) = \pm(1 + \varepsilon b)\ell(\theta, E) + \varepsilon\tilde{\mathcal{Y}}_{\pm} \circ \ell(\theta, E).$$

By Lemma 60 we know that $\tilde{\mathcal{Y}}_{\pm}$ have $\mathcal{C}^2$ norm bounded independently of ε, E and $\tilde{\mathcal{Y}}(0) = \tilde{\mathcal{Y}}'(0) = 0$, so that

$$\left|\varepsilon\tilde{\mathcal{Y}} \circ \ell\right|_{\mathcal{C}^1} \leq \text{cte.}\, \varepsilon^{1+\nu}.$$

Then, the main terms in $\mathcal{M}_\pm$ are

$$\begin{aligned}
&\mathcal{M}_\pm(\theta;\varepsilon) = \\
&= \pm\, k_0\sqrt{2(E-\varepsilon^2 U(\theta;0))}\frac{\partial \mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) - \varepsilon k_0^2\left(\frac{\partial \mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right)^2 \\
&+|\ell|\,\mathrm{O}_{\mathcal{C}^1}(\varepsilon^{\nu/2},\varepsilon^{\varrho}) + \mathrm{O}_{\mathcal{C}^1}(\varepsilon^{1+\nu}).
\end{aligned}$$

To apply Lemma 81 we need to check (150), that is, it suffices to show that we can bound from below the derivative with respect to $\theta = k_0\varphi + l_0 s$ of this function divided by $k_0\lambda_E^\pm(\varphi,s;\varepsilon)+l_0$. Computing this derivative, we obtain:

$$\begin{aligned}
&\frac{\partial}{\partial\theta}\left(\mathcal{M}_\pm(\varphi,s;\varepsilon)\right) \\
&\quad=\frac{\pm k_0}{\sqrt{2(E-\varepsilon^2 U(\theta;0))}}\left(2E\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right. \\
&\quad\left.-\varepsilon^2\left[k_0 U'(\theta;0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) + 2U(\theta;0)\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right]\right) \\
&\quad-2\varepsilon k_0\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \\
&\quad+|\ell|\,\mathrm{O}_{\mathcal{C}^0}(\varepsilon^{\nu/2},\varepsilon^{\varrho}) + \mathrm{O}_{\mathcal{C}^0}(\varepsilon^{1+\nu}).
\end{aligned}$$

Then,

$$\begin{aligned}
&\frac{\frac{\partial}{\partial\theta}\left(\mathcal{M}(\theta;\varepsilon)\right)}{k_0\lambda_E^\pm(\varphi,s;\varepsilon)+l_0} \\
=\quad&\frac{k_0}{2(E-\varepsilon^2 U(\theta;0))}\left(2E\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right. \\
&-\varepsilon\left[k_0 U'(\theta;0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) + 2U(\theta;0)\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right] \\
&\left.\mp 2\varepsilon k_0\sqrt{2(E-\varepsilon^2 U(\theta;0))}\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\frac{\partial^2}{\partial\theta^2}\mathcal{L}^*(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right) \\
&+\mathrm{O}_{\mathcal{C}^0}(\varepsilon^{\nu/2},\varepsilon^{\varrho}).
\end{aligned}$$

By hypothesis (161) we know that the main term of this expression is bounded away from zero by a constant C for $\theta \in \mathcal{J}_i \subset \mathcal{J}^*_{-l_0/k_0}$, $i = 1,2,3$. Consequently, the angle of intersection can be bounded again from below by $C'\varepsilon$, for some suitable constant independent of ε.

In order to see which surfaces will intersect $S(L_E^{F,\pm})$, we observe that the function $\mathcal{M}_\pm(\theta;\varepsilon)$ is given approximately by

$$k_0\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\left(\pm\sqrt{2(E-\varepsilon^2U(\theta;0))}-\varepsilon k_0\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right)$$

and we will have different behaviors depending on the branch $S(L_E^{F,\pm})$.

If we focus in the case of $S(L_E^{F,-})$, which correspond to the lower branch, the function $\mathcal{M}_-$ can have different signs depending on the size of $\sqrt{2(E-\varepsilon^2U(\theta;0))}$.

If $\sqrt{2(E-\varepsilon^2U(\theta;0))}$ is bigger than $-\varepsilon k_0\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})$ then $\mathcal{M}_-(\theta;\varepsilon)>0$ and $S(L_E^{F,-})$ will intersect the surfaces $L_{E'}^{F,-}$, with $E'<E$. This happens for positive values of E, which correspond to the primary tori under the separatrix loop, and for negative values of E, which correspond to the secondary tori inside the separatrix loop.

If $\sqrt{2(E-\varepsilon^2U(\theta;0))}$ is smaller than $-\varepsilon k_0\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})$ then $\mathcal{M}_-(\theta;\varepsilon)<0$, we obtain that $S(L_E^{F,-})$ will intersect the surfaces with $E'>E$. This means that, due to the fact that the Melnikov function is big, this surface intersect the surfaces $L_{E'}^{F,+}$. Then, in the case of a resonance of second order it is possible that only with one application of S we cross the gap.

Once we have cross the separatrix loop, that is, when we consider $S(L_E^{F,+})$, we have that $\mathcal{M}_+(\theta;\varepsilon)$ is always negative, so we cross the surfaces $L_{E'}^{F,+}$, for $E'>E$.

In all these cases we have that there exists some constant C'' such that

$$\max_{\theta\in\mathcal{J}_1\cup\mathcal{J}_2\cup\mathcal{J}_3}|\mathcal{M}(\lambda_E(\varphi,s;\varepsilon),\varphi-\lambda_E(\varphi,s;\varepsilon)s)|\geq C''\max(E^{1/2},\varepsilon),$$

and applying Lemma 150 we obtain the desired result.

Far from the resonance: $\tilde{c}\varepsilon^\nu\leq E\leq c_2L$.

This case is analogous to the non resonant region, because in this case,

$$\begin{aligned}|\ell(\theta,s)| &= \sqrt{2(E-\varepsilon^2U(x;\varepsilon))}=\sqrt{2E}\sqrt{1-\frac{\varepsilon^2}{E}U(x;\varepsilon)}\\ &= \sqrt{2E}(1+\mathrm{O}_{\mathcal{C}^2}(\varepsilon^{2-\nu})),\end{aligned}$$

consequently, as $\sqrt{2E} \geq \varepsilon^{\nu/2} >> \varepsilon$, the functions $\mathcal{M}_\pm$ become

$$\mathcal{M}_\pm(\theta;\varepsilon) = k_0\sqrt{2E}\frac{\partial\mathcal{L}^*}{\partial\theta}\left(-\frac{l_0}{k_0}+\sqrt{2E}, \varphi-(-\frac{l_0}{k_0}+\sqrt{2E})s\right)+\mathrm{O}_{\mathcal{C}^2}(\varepsilon^{2-\nu}).$$

Then, if the function $\frac{\partial\mathcal{L}^*}{\partial\theta}(I,\theta)$ is not constant and negative as a function of θ, we can apply Lemma 81, to get the desired result.

□

Hypothesis **H5''** in (161) amounts to the existence of a lower bound for a function of the following type:

$$\frac{a(\theta)E+b(\theta)\varepsilon^2+c(\theta)\varepsilon\sqrt{E-\varepsilon^2U(\theta;0)}}{E-\varepsilon^2U(\theta;0)}$$

Concretely we have

(162)

$$\begin{aligned} a(\theta) &= k_0\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \\ b(\theta) &= -\frac{k_0}{2}\left[k_0U'(\theta;0)\frac{\partial\mathcal{L}^*}{\partial\theta}(-\frac{l_0}{k_0},\frac{\theta}{k_0})+2U(\theta;0)\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0})\right] \\ c(\theta) &= \frac{\pm k_0}{\sqrt{2}}\frac{\partial^2\mathcal{L}^*}{\partial\theta^2}(-\frac{l_0}{k_0},\frac{\theta}{k_0}) \end{aligned}$$

The following Lemma gives a computable sufficient condition that guaranties that hypothesis **H5''** is verified.

Lemma 90. *Let $a(\theta)$, $b(\theta)$, $c(\theta)$ be functions of class $\mathcal{C}^r$, $r \geq 0$, such that there exist θ_1, θ_2, θ_3 in some interval $\mathcal{J}$ verifying*

$$\begin{vmatrix} \tilde{a}(\theta_1) & \tilde{a}(\theta_2) & \tilde{a}(\theta_3) \\ \tilde{b}(\theta_1) & \tilde{b}(\theta_2) & \tilde{b}(\theta_3) \\ \tilde{c}(\theta_1) & \tilde{c}(\theta_2) & \tilde{c}(\theta_3) \end{vmatrix} \neq 0, \tag{163}$$

where

$$\begin{aligned} \tilde{a}(\theta) &= a(\theta)^2 \\ \tilde{b}(\theta) &= 2a(\theta)b(\theta)-c(\theta)^2 \\ \tilde{c}(\theta) &= b(\theta)^2-c(\theta)^2U(\theta;0) \end{aligned} \tag{164}$$

Then, there exists a constant $\tilde{C}$ and three intervals $\theta_i \in \mathcal{J}_i \subset \mathcal{J}$, $i = 1,2,3$, such that given any $(x,y) \in \mathbb{R}^2$, one can choose $i = 1,2,3$ and, for $\theta \in \mathcal{J}_i$

$$\left|a(\theta)x+b(\theta)y^2+c(\theta)y\sqrt{x-y^2U(\theta;0)}\right| \geq \tilde{C}\sqrt{x^2+y^4}. \tag{165}$$

Proof. We call $z = \frac{x}{y^2}$, and consider the function

$$f(\theta, z) = \frac{a(\theta)z + b(\theta) + c(\theta)\sqrt{z - U(\theta; 0)}}{\sqrt{1+z^2}} = \frac{g(\theta, z)}{\sqrt{1+z^2}}$$

Then, it is enough to prove that there exists a constant C and three intervals $\theta_i \in \mathcal{J}_i$, $i = 1, 2, 3$, such that given any value of $z \in [c_1, \infty)$, one can choose $i = 1, 2, 3$ and, for $\theta \in \mathcal{J}_i$

$$|f(z)| \geq \tilde{C} \tag{166}$$

One can check easily that condition (163) implies that there is no z such that the functions $g(\theta_i, z)$ vanish simultaneously. Then, given any value of z and calling

$$c(z) = \max(f(\theta_1, z), (f(\theta_2, z), (f(\theta_3, z))$$

we have that $c(z) > 0$. Moreover as $\lim_{z\to\infty} f(\theta_i, z) = a(\theta_i) \neq 0$, we have that $\min c(z) \geq \tilde{C} > 0$, then, for any value of $z \in [c_1, \infty)$, $c(z) \geq C$, and this concludes the proof.

□

10.3. **Existence of transition chains to different objects.** In this section, we prove the existence of heteroclinic trajectories between different kinds of invariant manifolds (primary tori, secondary tori, or periodic orbits) which fill the manifold $\tilde{\Lambda}_\varepsilon$ up to very small gaps of order $\varepsilon^{3/2}$ in the I direction. The precise formulation of the result is given in Proposition 92. In Theorem 93 we establish a result that takes advantage of the existence of several connections.

We recall that in Section 8, by analyzing the inner map, for any $|\varepsilon| \ll 1$ we have produced a collection $\{\mathcal{T}_i\}_{i=1}^{N(\varepsilon)}$ consisting of primary KAM tori, secondary KAM tori or periodic orbits such that primary tori, the secondary tori and the local stable and unstable manifolds of the periodic orbits are at a distance less or equal than $C\varepsilon^{3/2}$ in $\tilde{\Lambda}_\varepsilon$. We will assume without loss of generality that the $\mathcal{T}_i$ are different for different i.

We will refer to this collection as an $O(\varepsilon^{3/2})$ scaffolding. Note that these scaffoldings consist of objects of different dimensions (the periodic orbits are of dimension 1 and the the KAM tori are of dimension 2) and, of course, different topologies.

We will find it convenient to introduced the notion of leaves of the scaffolding. If $\mathcal{T}_i$ is a KAM tori, either primary or secondary, we denote the leaf corresponding to $\mathcal{T}_i$ by L_i. If $\mathcal{T}_i$ is a periodic orbit, we will associate two leaves of the scaffolding $L_i^{\mathrm{ws}} = W_{\mathcal{T}_i}^{\mathrm{ws,loc}}$, $L_i^{\mathrm{wu}} = W_{\mathcal{T}_i}^{\mathrm{wu,loc}}$

Remark 91. For many purposes, we will not need to distinguish between L_i^{ws}, L_i^{wu} since we have shown that they are at a distance ε^m in $\tilde{\Lambda}_\varepsilon$, which is much smaller than the effects we will be considering. Hence, when we say that $L_i \pitchfork_{\tilde{\Lambda}_\varepsilon} L_{i'}$ and $\mathcal{T}_i$ is a periodic orbit, we will mean that $L_i^{\mathrm{ws}} \pitchfork_{\tilde{\Lambda}_\varepsilon} L_{i'}$ and that $L_i^{\mathrm{wu}} \pitchfork_{\tilde{\Lambda}_\varepsilon} L_{i'}$ in the sense of intersection of the manifolds. □

Note that the leaves of the scaffolding are all of dimension 2 and we have already shown that they are very close to the level sets of the function given by the averaged hamiltonian. In particular, we can assume that the averaged energy evaluated in different leaves does not have any common value. This allows us to order the leaves by the values of the average energy on them. We can arrange that order in i is the order of increasing averaged energy an refer to them as an ordered scaffolding.

We also recall that in Section 9 we have computed and analyzed the scattering map. We have shown in Lemmas 82, 85 and 88 that, under the transversality conditions for a homoclinic intersection made in Theorem 7, any leaf of the scaffolding intersects transversally all the leaves of the scaffolding at a distance $O(\varepsilon)$ which is in the same direction.

It is worth emphasizing that the scaffolding depends only on the inner map and, therefore is independent of the the intersection that we are considering. In particular, the constant C is independent of the intersection. On the other hand, the scattering map and the intersections do depend on the homoclinic manifold considered.

Proposition 92. *Consider a Hamiltonian system of Hamiltonian* (6)*, and consider any interval* (I_-, I_+) *verifying hypothesis* **H1***,...,***H5** *stated in Theorem 7. Assume that* $0 < |\varepsilon| << 1$ *is given.*

Let $\{\mathcal{T}_i\}_{i=1}^{N(\varepsilon)}$ *be a* $C\varepsilon^{3/2}$ *ordered scaffolding as produced in Section* 8*.*

Then, $\{\mathcal{T}_i\}_{i=1}^{N(\varepsilon)}$ *is a transition chain.*

If in (13) *we take the positive alternative, we conclude that* $\{\mathcal{T}_{N(\varepsilon)+1-i}\}_{i=1}^{N(\varepsilon)}$.

The Proposition is an obvious consequence of the Lemmas 82, 85 which establish that, under the hypothesis of transversality in Theorem 7 and the proximity assumptions in Proposition 92 for $L_{\sigma(i)}, L_{\sigma(i+1)}$, we have that $S(L_{\sigma(i)}) \pitchfork_{\tilde{\Lambda}_\varepsilon} L_{\sigma(i+1)}$.

By Lemma 78 this means that

$$W^{\mathrm{u}}_{L_{\sigma(i)}} \pitchfork W^{\mathrm{s}}_{L_{\sigma(i+1)}}$$

When both $L_{\sigma(i)}$ $L_{\sigma(i+1)}$ are KAM tori, this implies immediately that $\mathcal{T}_{\sigma(i)}$ $\mathcal{T}_{\sigma(i+1)}$ are elements of a transition chain.

When either of $\mathcal{T}_{\sigma(i)}$ $\mathcal{T}_{\sigma(i+1)}$ are periodic orbits, we take into account the observations in Remark 91.

If $\mathcal{T}_{\sigma(i)}$ is a periodic orbit we observe that we have shown that $W^{\mathrm{wu}}_{\sigma(i)} \pitchfork_{\tilde{\Lambda}_\varepsilon} L_{\sigma(i+1)}$. If $\mathcal{T}_{\sigma(i)}$ is a periodic orbit we observe that we have shown that $W^{\mathrm{ws}}_{\sigma(i+1)} \pitchfork_{\tilde{\Lambda}_\varepsilon} L_{\sigma(i)}$. (In this paper, the case when both $\mathcal{T}_{\sigma(i)}$, $\mathcal{T}_{\sigma(i+1)}$ are periodic orbits does not appear, but the conclusions remain the same).

In either case, we obtain that $W^{\mathrm{u}}_{\mathcal{T}_{\sigma(i)}} \pitchfork W^{\mathrm{s}}_{\mathcal{T}_{\sigma(i+1)}}$. $\square$

The meaning of Proposition 92 is that we can use the objects produced in the scaffolding to produce transition chains. The only constraints is that the steps can not be longer that $O(\varepsilon)$ – the amount moved by a heteroclinic excursion – that the objects are ordered and that we do not step out of the region where we have the transversality conditions.

We emphasize that the conclusion

$$W^{\mathrm{u}}_{\mathcal{T}_i} \pitchfork W^{\mathrm{s}}_{\mathcal{T}_{i'}} \tag{167}$$

is independent of the scattering map that was used in the proof. Hence, if there are different intersections between $W^{\mathrm{s}}_{\tilde{\Lambda}_\varepsilon}$, $W^{\mathrm{u}}_{\tilde{\Lambda}_\varepsilon}$, we can construct more complicated transitions.

Recall—see remark 11—that there is a Poincaré function $\mathcal{L}(I, \varphi - I\tau, s-\tau)$, where $\mathcal{L}$ is given in (10), for each parameterization (23) chosen.

Moreover, the function $\tau \mapsto \mathcal{L}(I, \varphi - I\tau, s - \tau)$ will typically have several critical points $\tau^*(I, \varphi, s)$. (If it is not constant, we will have at least a minimum and a maximum.)
[1]

Furthermore, as observed in Remark 8, typically, given an homoclinic intersection, there will be two different open sets H_-^+, H_-^+ satisfying the hypothesis **H4** but taking the positive and negative signs in (13).

By Proposition 72, any of them gives rise to a homoclinic manifold to the manifold $\tilde{\Lambda}_\varepsilon$. Each of these homoclinic manifolds can be used to construct a scattering map S. Of course, the results of transversality of the heteroclinic intersections of Lemma 78, are valid for each of the scattering maps constructed.

[1]Moreover, as it is well known to dynamicists, the fact that one intersection exists implies that their images are also intersections and that there are other secondary intersections (See e.g. [Mos73] Theorem 3.8) which cannot be continued till the case $\varepsilon = 0$. Since in this paper, we are making statements that are valid for all ε sufficiently small, we will not use these secondary intersections.

In the following result, we formulate a result that takes into acount all these possibilities. Note that, in particular, we have that, when we have (16), we have a symbolic dynamics in the transition chains.

Theorem 93. *Consider a Hamiltonian system of Hamiltonian* (6) *satisfying* **H1**, **H2**, **H3** *in Theorem 7.*

Assume that the stable and unstable manifolds of $\tilde{\Lambda}_\varepsilon$ have K intersections which verify **H4**, **H5** *in Theorem 7.*

Denote by $I_k = (I_-^k, I_+^k)$, $k = 1, \dots K$ the intervals in **H4** *and by $\gamma_k = \pm$ be the signs taken by the function in* (13)*. Assume that $\bigcup_k I_k = I^* \equiv (I_-^*, I_+^*)$.*

There is a constant $C > 0$ and $\varepsilon^* > 0$ such that for $0 < |\varepsilon| < \varepsilon^*$ we have the following.

Let $\{\mathcal{T}_i\}_{i=1}^{N(\varepsilon)}$ be a $C\varepsilon^{3/2}$ scaffolding as produced in Section 8.

Given a map $\sigma : \mathbb{N} \to \{1, 2, \dots N(\varepsilon)\}$ such that:

- $$\mathcal{T}_{\sigma(i)} \subset (I_-^* + C\varepsilon, I_+^* - C\varepsilon) \times \mathbb{T}^2$$
- $$\mathrm{dist}_{\tilde{\Lambda}_\varepsilon}(L_{\sigma(i)}, L_{\sigma(i+1))} \le C\varepsilon$$
- $$L_{\sigma(i)}, L_{\sigma(i+1)} \subset I^k$$
  and the ordering of $L_{\sigma(i)}, L_{\sigma(i+1)}$ is compatible with the ordering of the interval.

Then, the sequence $\{\mathcal{T}_{\sigma(i)}\}_{i=1}^\infty$ is a transition chain.

Note that, in particular, if we have (16) we can take $I^1 = I^2$ and both equal to the interval, but we take a different choices of signs in (13). Hence, we can obtain many different transition chains.

In general, when it is possible to make a circuit along oriented intervals, we have an uncountable choice for transition chains.

11. Orbits shadowing the transition chains and proof of theorem 7

In this section, we just state the well known result that given a transition chain, we can find an orbit visiting all the elements of the chain.

There are many proofs of similar results in the literature. In our case, however, we have to make sure that the proofs remain valid for transition chains that incorporate objects with different topologies and indeed different dimensions. In our formulation also we allow the transition chains to have infinite many transition tori. This makes certain well known proofs in the literature not applicable. Hence, we present full details, following the exposition in [DLS00]. An exposition of related results along similar lines can be found in [FM03].

Lemma 94. *Let $\{\mathcal{T}_i\}_{i=0}^N$ be a transition chain.*

Given $(\delta_i)_{i=0,\dots,N}$ be a sequence of strictly positive numbers, we can find a point $\tilde{x} \in (\mathbb{R}\times\mathbb{T})^2 \times \mathbb{T}$, and a increasing sequence of numbers $0 = t_0 < \ldots < t_N$ such that

$$\tilde{\Phi}_{t_i,\varepsilon}(\tilde{x}) \in \mathbf{B}(\mathcal{T}_i, \delta_i)$$

where $\mathbf{B}(\mathcal{T}_i, \delta_i)$ is neighborhood of size δ_i of the torus $\mathcal{T}_i$.

A particular case of the results of [FM00] is:

Lemma 95. *Let f be a $\mathcal{C}^2$ symplectic mapping in a symplectic manifold. Assume that the map leaves invariant $\mathcal{C}^1$ torus $\mathcal{T}$ and that the motion on the torus is an irrational rotation. Let Γ be a manifold intersecting $W^{\mathrm{u}}_{\mathcal{T}}$ transversally. Then,*

$$W^{\mathrm{s}}_{\mathcal{T}} \subset \overline{\bigcup_{i>0} f^{-i}(\Gamma)}$$

We emphasize that since the proof in [FM00] only makes assumptions about the motion on the torus and the linearization around it, it applies independently of whether the torus is primary or secondary. Also, it is independent of the dimension.

In the case that the transition chain is finite, this establishes the desired result simply by the continuous dependence on the initial conditions.

The argument for infinite chains requires some elementary point set topology, which we reproduce verbatim from [DLS00].

Lemma 96. *Let $\{\mathcal{T}_i\}_{i=1}^\infty$ be a sequence of transition tori. Given $\{\varepsilon_i\}_{i=1}^\infty$ a sequence of strictly positive numbers, we can find a point P and a increasing sequence of numbers T_i such that*

$$\Phi_{T_i}(P) \in N_{\varepsilon_i}(\mathcal{T}_i)$$

where $N_{\varepsilon_i}(\mathcal{T}_i)$ is a neighborhood of size ε_i of the torus $\mathcal{T}_i$.

Proof. Let $x \in W^{\mathrm{s}}_{\mathcal{T}_1}$. We can find a closed ball B_1, centered on x, and such that

$$\Phi_{T_1}(B_1) \subset N_{\varepsilon_1}(\mathcal{T}_1). \tag{168}$$

By the Inclination Lemma 95

$$W^{\mathrm{s}}_{\mathcal{T}_2} \cap B_1 \neq \emptyset.$$

Hence, we can find a closed ball $B_2 \subset B_1$, centered in a point in $W^{\mathrm{s}}_{\mathcal{T}_2}$ such that, besides satisfying (168):

$$\Phi_{T_2}(B_2) \subset N_{\varepsilon_2}(\mathcal{T}_2).$$

Proceeding by induction, we can find a sequence of closed balls

$$\begin{aligned} &B_i \subset B_{i-1} \subset \cdots \subset B_1 \\ &\Phi_{T_j}(B_i) \subset N_{\varepsilon_j}(\mathcal{T}_j), \quad i \le j. \end{aligned}$$

Since the balls are compact, $\cap B_i \neq \emptyset$. A point P in the intersection satisfies the required property. □

End of the proof of Theorem 7. In order to prove the existence of an orbit $\tilde{x}(t) = \tilde{\Phi}_{t,\varepsilon}(\tilde{x})$ of system (6) verifying (14) we only need to apply Lemma 94 to the transition chains obtained in proposition 92.

Then, Theorem 7 is proved.

□

12. Conclusions and remarks

12.1. **The role of secondary tori and the speed of diffusion.** We have shown that the secondary whiskered tori as well as lower dimensional tori, i.e. periodic orbits, can be used in the construction of transition chains to overcome the *"large gap problem"* in the study of diffusion.

We think that this is a step toward reconciling the physical literature and intuition [Chi79, Ten82] (that holds that the main cause of instability and diffusion are resonances) and the mathematical literature, which hitherto emphasized methods that were based in the mechanism of [Arn64] which requires that KAM tori are close and, hence, that there are no resonances.

We think that it will be quite interesting to understand better the geometric objects that one can find in different types of resonances.

12.2. **Comparison with** [DLS00]**.** Even if this paper and [DLS00] are logically independent, there are several analogies in the general strategy and the tools of the present work and that in [DLS00].

Nevertheless, there are several important differences in the geometry of the systems considered here and that in [DLS00] (and in [Mat95, BT99]).

Let us emphasize the most salient geometric differences for the reader interested in comparing the two papers:

- The manifolds $W^{\mathrm{s}}_{\tilde{\Lambda}}$ and $W^{\mathrm{u}}_{\tilde{\Lambda}}$ considered in this paper do not intersect transversely in the unperturbed case. This is why we refer to the system here as a priori unstable, but we coined the name *a priori chaotic* for the system in [Mat95, DLS00, BT99, DLS01]. This has made it necessary for us to study the transverse intersection through a perturbation theory.

- The homoclinic connections of the model here do not include a phase shift. (See equation (29) and compare it with equation (2.1) of [DLS00].)

  In the case of [Mat95, DLS00, BT99], there was a phase shift and, hence, we needed to compare the KAM tori at different points. This required a more detailed KAM perturbation theory carried out in Lemma 4.16 of [DLS00]. Such considerations will not be needed in this paper, but could be reintroduced to analyze a more complicated model.
- The most important difference with [DLS00] is that in our problem there is no analogue of the scaling (Section 4.2 in [DLS00]) which makes the system, for high energy, to become fast with respect to the perturbation. Hence, the fact that gaps between tori in [DLS00] were very small has not been maintained in our case and we had deal with the so called *large gap problem*.

Of course, the fact that models with very different features can be understood with similar tools makes us quite hopeful that even more complicated models can be analyzed in similar ways.

12.3. **Heuristics on the genericity properties of the hypothesis and the phenomena.** The verification we have presented of the mechanism is geared towards the explicit verification of the mechanism in concrete systems.

This point of view is motivated because in applications or in mathematics, one has to deal with concrete systems – e.g. the solar system or the systems appearing in technology – which have very special properties.

Nevertheless, several colleagues have asked about the genericity of the mechanism that we have discussed here.

In this section, we will discuss briefly and informally the genericity hypothesis of Theorem 7 and of Theorem 93.

First note that the hypothesis on the pendulum hold for an open and dense set of penduli (indeed for all but a manifold of infinite codimension) in any $\mathcal{C}^r$ topology, $r \geq 1$.

Hence, we will considered the genericity of the hypothesis on h in the space of trigonometric polynomials once we fix the pendulum satisfying **H1**, **H2**. We note that the pendulum will have two connecting orbits that are geometrically different.

If we fix one of the connecting orbits, we note that the mapping the sends the perturbation to the Melnikov integral is linear and nontrivial. In this case, generic properties on the Melnikov potential are implied by generic assumptions on the perturbation. Moreover, since h is assumed to be a polynomial, we do not need to discuss in which C^r genericity properties hold.

We note that it is generic that there are open and non-empty sets for which we have **H4** which include the resonances in $\tilde{\Lambda}_\varepsilon$.

This follows because the existence if such an open set is clearly open in the space of h's. The density follows because the negation of the conclusion is that for all points in $\tilde{\Lambda}_\varepsilon$ in a resonance $\{k/l\} \times \mathbb{T}^2$, the function Γ has only degenerate critical points. This can be easily destroyed by arbitrarily small perturbations.

Hence, we conclude that the set of models of the form (6) that overcome the large problem in one resonance is open and dense.

We can also wonder about the genericity of the existence of orbits that transverse all the resonances. The following discussion, based simply on the counting of parameters, presents some plausibility arguments for what one would expect.

Notice that given a critical point for Γ we expect that it will be non-degenerate except for a set of I, φ, s of codimension 1. In these sets, we will have that it is generically possible to find sets $H^\pm_-$ where we have different signs in (13). Generically, these sets will have overlapping projections on the I direction, hence, it will be possible to obtain the existence of a symbolic dynamics.

If we consider two critical points, we expect that the sets of I, φ, s for which one of them are non-degenerate is a set of codimension 2.

Furthermore, if we consider systems for which the pendulum has two homoclinics, we have more critical points to study. Once we have 4 or more we expect that all the points in $\tilde{\Lambda}_\varepsilon$ correspond to a non-degenerate critical point.

Hence, we expect that the intervals corresponding to a positive sign in (13) overlap and their union covers all the resonant regions. Similarly, the intervals the negative signs in (13) will overlap and cover the resonances. Hence, we expect that, for the systems we have consider it will be generic to have trajectories that cross all the resonances.

We postpone a detailed verification of the results for generic systems. We hope that a generic verification of the system could be simpler and applicable without **H3**.

Note that all our results on existence of orbits changing the action are formulated for all ε with $|\varepsilon| \leq \varepsilon^*$ with $\varepsilon^* > 0$. The ε^*, however, depends on h and is only defined for h that satisfy some non-degeneracy conditions. As h approaches the degenerate h we have $\varepsilon^*(h)$ approaches zero.

If we do not want to consider parameters, we are lead to consider the set of εh for which we can establish diffusion across the biggaps. This set contains intervals in ε, but the size of the interval can go to zero as we approach the degenerate sets.

This is very similar to the *"cusp residual"* sets considered in [Mat02]. The main difference is that in cusp residual sets one does not require that there are whole intervals in the variables ε but allows to take away some exceptional values.

12.4. **The hypothesis of polynomial perturbations.** It seems quite possible that **H3** can be avoided at the price of just carrying out more delicate estimates. At the moment, we thought it better to avoid this complication. Nevertheless in the following, we outline a sketch of a heuristic argument for resonances of order 1. For the sake of clarity, we ignore constants.

We note that, because the scattering map moves by an amount of order ε, we only need to consider resonant regions whose size is bigger or equal than $\varepsilon^{1+\delta}$ for an arbitrary positive δ.

A more careful analysis shows that the size of a resonance of order 1 and denominator k can be bounded from above by $(\varepsilon|a_k|)^{1/2}$ where a_k is the coefficient of the Fourier coefficient of the perturbation.

If the perturbation is C^ℓ, we have that $|a_k| \le k^{-\ell}$. So that the size of a resonant region can be bounded from above by $\varepsilon^{1/2}|k|^{-\ell/2}$.

Hence, it is clear that it is only necessary to consider resonant regions with denominator $k \le k_{\max}$ where $k_{\max}$ is such that $\varepsilon^{1/2}|k_{\max}|^{-\ell/2} = \varepsilon^{1+\delta}$. That is,

$$|k_{\max}| = \varepsilon^{-1/\ell+2\delta/\ell}$$

We note that the minimum distance between two rational numbers of denominator smaller that $k_{\max}$ – and therefore, the minimum distance between the resonances that need to be considered is:

$$|k_{\max}|^{-2} = \varepsilon^{2/\ell+4\delta/\ell}$$

The maximum width of any of these resonant regions is $\varepsilon^{1/2}$.

Hence, we see that a condition that ensures that the resonant regions are separated is $\ell > 4$.

A similar analysis works for the resonances of order 2.

Once that we have that the resonant regions are separated, we can perform a very similar analysis than the one performed here. The main difference is that rather than choosing L as we have chosen here, one would need to take L depending on ε and k in such a way that L is much larger than the size of the resonant region. If ℓ is large enough, this can be accomplished e.g. by taking $L = \varepsilon^{1/4}|k|^{-\ell/4}$.

The dependence of the averaging results on L can be worked out. It seems that they can also be made small for sufficiently smooth systems. On the other hand, we remark that the intuition gathered in [Chi79], it seems that when the resonances overlap, the diffusion should be more intense.

12.5. **Involving other objects.** We also note that it should be possible to consider other orbits besides the whiskered tori we have used. In particular, given that in a generic system there are hyperbolic periodic orbits approximating the KAM tori (See e.g. [FL92]) and that the stable and unstable manifolds are close to the torus, it seems possible that, at least for generic systems one could construct hyperbolic orbits that connect.

It also seems plausible that one could adapt the method of [Lla02] using normally hyperbolic laminations to discuss the problem.

More tentatively, it looks plausible that Aubry-Mather Cantor sets and their invariant manifolds could be used as transition elements.

12.6. **Variational methods.** There is an uncanny relation between the geometric methods and variational methods. We think that it would be very useful to pursue the study of the parallels between the two very different methods. It seems plausible that if one proved connecting lemmas based on variational methods, one could also use Aubry-Mather sets in place of the tori. Of course, variational methods seem to have to use positivity or convexity properties that are not present in our methods.

We note that it is quite possible that one can use a mixed approach. Once one identifies the relevant geometric objects and produce heteroclinic connections among them, variational methods can produce very effective shadowing orbits [Mat93, CP02, RS02].

Some implementations of these mixed methods happen in [Bes96, BCV01]

12.7. **Diffusion times.** We notice that, during this paper, nothing is said about the diffusion time since the shadowing Lemma 94 is only based in topological methods and does not provide quantitative estimates about the ergodization time as other methods do ([Tre02, CG03, BB02]).

We also note the methods of [Lla02] using hyperbolic laminations, that yield very good times.

13. An example

Consider the Hamiltonian

$$H_\varepsilon(p,q,I,\varphi,t) = \pm\left(\frac{p^2}{2} + \cos q - 1\right) + \frac{I^2}{2} + \varepsilon \cos q \; g(\varphi,t), \tag{169}$$

where

$$g(\varphi,t) = \sum_{(k,l)\in\mathcal{N}} a_{k,l}\cos(k\varphi + lt) + b_{k,l}\sin(k\varphi + lt)$$

is a trigonometric polynomial in the angles φ, t ($\mathcal{N}$ is a finite set of indexes). The Hamiltonian of one degree of freedom $P_\pm(p,q) = \pm\left(p^2/2 + \cos q - 1\right)$ is the standard pendulum when we choose the $+$ sign, and its separatrix for positive p is given by (24)

$$q_0(t) = 4\arctan e^{\pm t}, \quad p_0(t) = 2/\cosh t.$$

An important feature of the Hamiltonian (169) is that the 3-dimensional hyperbolic invariant manifold

$$\tilde{\Lambda} = \{(0,0,I,\varphi,s) : (I,\varphi,s) \in \mathbb{R}\times\mathbb{T}^2\}$$

is *preserved* for $\varepsilon \neq 0$: $p = q = 0 \Rightarrow \dot{p} = \dot{q} = 0$. However, in contrast with the example in [Arn63b], the perturbation does not vanish on $\tilde{\Lambda}$. Indeed, restricted to $\tilde{\Lambda}$, the reduced Hamiltonian takes the form $I^2/2 + \varepsilon g(\varphi,t)$. Hence, 2-dimensional whiskered tori

$$\mathcal{T}_I^0 = \{(0,0,I,\varphi,s) : (\varphi,s) \in \mathbb{T}^2\}$$

are not preserved, and primary resonances take place at $I = -l/k$ for each $(k,l) \in \mathcal{N}, k \neq 0$ such that $a_{k,l} \neq 0$. Therefore, (169) presents the large gap problem.

The Melnikov potential (10) of the Hamiltonian (169) is given by

$$\mathcal{L}(I,\varphi,s) = \frac{1}{2}\int_{-\infty}^{\infty} p_0^2(\sigma) g(\varphi + I\sigma, s+\sigma)d\sigma,$$

and computing the integrals by the residue theorem, we obtain:

$$\mathcal{L}(I,\varphi,s) = \sum_{(k,l)\in\mathcal{N}} A_{k,l}(I)\cos(k\varphi + lt) + B_{k,l}(I)\sin(k\varphi + lt),$$

with

$$A_{k,l} = 2\pi\frac{(kI+l)}{\sinh\frac{\pi}{2}(kI+l)}a_{k,l}, \quad B_{k,l} = 2\pi\frac{(kI+l)}{\sinh\frac{\pi}{2}(kI+l)}b_{k,l}.$$

As we will verify in concrete examples, for a very general choice of the coefficients $a_{k,l}, b_{k,l}$ 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 a $\tau = \tau^*(I, \varphi, s)$ which verify hypothesis **H4**.

For instance, let us consider the case of a function g with only two harmonics

$$g(\varphi, t) = a_0 \cos(\varphi) + a_1 \cos(\varphi - t) \tag{170}$$

which gives rise to two "large gaps" associated to the two primary resonances (20) $I = 0, 1$, as well as to the "big gap" associated to the secondary resonance (21) $I = 1/2$. Assuming that

$$a_0 a_1 (a_0^2 - a_1^2) \neq 0 \tag{171}$$

and choosing, for instance, $[I_-, I_+] = [-1/2, 3/2]$, we are going to check that for all $(I, \varphi, s) \in [I_-, I_+] \times \mathbb{T}^2$, the function $\tau \in \mathbb{R} \mapsto \mathcal{L}(I, \varphi - I\tau, s - \tau)$ has non-degenerate critical points.

First, notice that

$$\mathcal{L}(\tau) := \mathcal{L}(I, \varphi - I\tau, s - \tau) = A_0 \cos(\varphi - I\tau) + A_1 \cos(\varphi - s - (I - 1)\tau),$$

with

$$A_0 = A_0(I) = 2\pi \frac{I}{\sinh \frac{\pi}{2} I} a_0, \quad A_1 = A_1(I) = 2\pi \frac{(I-1)}{\sinh \frac{\pi}{2}(I-1)} a_1,$$

so, fixed (I, φ, s), we only need to study the evolution of the Melnikov potential

$$\mathcal{L}(I, \psi_0, \psi_1) = A_0 \cos(\psi_0) + A_1 \cos(\psi_1), \tag{172}$$

expressed in the variables $(\psi_0, \psi_1) = (\varphi, \varphi - s) \in \mathbb{T}^2$, along the straight lines $R = R(I, \varphi, s)$ on the torus:

$$\psi_0(\tau) = \varphi - I\tau, \quad \psi_1(\tau) = \varphi - s - (I - 1)\tau. \tag{173}$$

As long as $a_0 a_1 \neq 0$ (and therefore $A_0 A_1 \neq 0$), the Melnikov potential (172) possesses four non-degenerate critical points: a maximum, a minimum and two saddles. 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 (173) that enters into some extremum basin is tangent to one of the convex closed level curves, giving rise to a non-degenerate extremum of $\mathcal{L}(\tau)$. So, degenerate extrema of $\mathcal{L}(\tau)$ can only exist for straight lines (173) that *never* enter inside such extremum basins. In particular, this could only happen for rational values of I, since an irrational value of I implies a dense straight line (173) (and infinite non-degenerate extrema for $\mathcal{L}(\tau)$).

A closer look at the Melnikov potential (172) shows that both extremum basins contain a vertical 1-dimensional torus (for $|A_0| \geq$

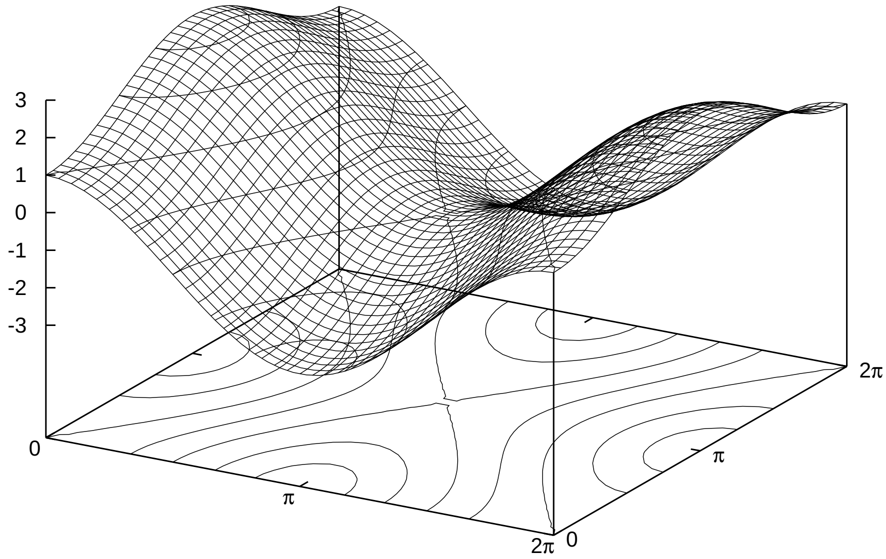

FIGURE 3. Graph and level curves of the Melnikov potential $\mathcal{L}(I, \psi_0, \psi_1) = A_0 \cos(\psi_0) + A_1 \cos(\psi_1)$ for $A_0 > -A_1 > 0$. There are four non-degenerate critical points $\psi_0 = 0, \pi$, $\psi_1 = 0, \pi$ and $\mathcal{L}(\pi, 0) = -A_0 + A_1 < \mathcal{L}(\pi, \pi) = -A_0 - A_1 < \mathcal{L}(0, 0) = A_0 + A_1 < \mathcal{L}(0, \pi) = A_0 - A_1$.

$|A_1|$) or a horizontal 1-dimensional torus (for $|A_0| \leq |A_1|$), except for a saddle point. See Figure 3 for a pictorial representation of the case $A_0 > -A_1 > 0$. In particular, this means that, except for the case of vertical lines (for $|A_0| \geq |A_1|$), horizontal lines (for $|A_0| \leq |A_1|$), or straight lines joining the two saddle points, every straight line R enters inside both extremum basins. So we already know that, except for the resonant values $I = 0, 1, 1/2$ corresponding to the slopes $\infty, 0, \pm 1$, respectively, of the straight lines (173), $\mathcal{L}(\tau)$ has non-degenerate extrema.

On the other hand, for $I = 0, 1, 1/2$ one can check directly that degenerate critical points of $\mathcal{L}(\tau)$ can only appear for $A_0(0) = 0, A_1(1) =$

$0, A_0(1/2) = \pm A_1(1/2)$, respectively, which is equivalent to $a_0 = 0, a_1 = 0, a_0 = \pm a_1$, respectively.

Therefore, as long as condition (171) for the coefficients a_0, a_1 is assumed, 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 non-constant along any straight line, it is clear that $\partial_\varphi \mathcal{L}^*$ is non-constant, so hypothesis **H4** is fulfilled.

The conditions **H5'**, **H5"**, **H5"'** can also be checked directly in the example (170) at the resonances $I = 0, 1, 1/2$, since we know explicitly the reduced Hamiltonian $I^2/2 + \varepsilon g(\varphi, t)$ given in Proposition 30 so that:

$$\varepsilon K_1(J, \varphi, s; \varepsilon) = \varepsilon g(\varphi, s).$$

For example, at $I = 0$ (which comes from the harmonic $k_0 = 1$, $l_0 = 1$), we have, by formula (52), $U^{1,0} = a_0 \cos\theta, \quad \theta = \varphi$ and Hypothesis **H5'** is verified because $U^{1,0}$ has a non-degenerate critical point at $\theta = 0$. By formulas (73), (76) and (78) we obtain the main term of the Hamiltonian in the resonant region close to $I = 0$:

$$U(x, 0) = a_0(\cos x - 1), \quad x = \theta.$$

On the other hand, for $I = 0$, we have:

$$\mathcal{L}(I, \varphi - I\tau, s - \tau) = A_0(0)\cos\varphi + A_1(0)\cos(\varphi - s + \tau)$$

where $A_0(0) = 4a_0$, $A_1(0) = \frac{2\pi}{\sinh(\pi/2)} a_1$. Therefore $\tau^*(0, \varphi, s) = \varphi - s$ or $\tau^*(0, \varphi, s) = \varphi - s + \pi$, so that the reduced Poincaré function (145) reads

$$\mathcal{L}^*(0, \theta) = A_0(0)\cos\theta \pm A_1(0), \quad \theta = \varphi,$$

whose extrema are non-degenerate by condition (171).

Then condition **H5"** is also satisfied because

$$a_0 \frac{-\sin^2\theta + 2(\cos\theta - 1)\cos\theta}{2\cos\theta}$$

is non-constant.

Analogous verifications can be carried out in the other resonances.

Hence, we conclude applying Theorem 7 and 93:

Proposition 97. *The Hamiltonian* (169) *with g as in* (170)*, where a_0, a_1 are such that they are not in the codimension* 1 *surface of equation $a_0 a_1 (a_0^2 - a_1^2) = 0$, and $|\varepsilon| \le \varepsilon^*(a_0, a_1)$, admits orbits following the mechanism described in this paper and such that $I(0) \le -1/2$, $I(T) \ge 3/2$ for some $T > 0$.*

We note that Proposition 97 is extremely conservative. We have only used the critical points close to the extrema of $\mathcal{L}$. It is clear that there are many other critical points.

The example (169) is somewhat non-generic because it has a symmetry that causes that the two homoclinic orbits of the pendulum give the same Melnikov function. Nevertheless, in spite of the fact that the system is very non-generic, we can verify easily the conditions of our theorem.

14. Acknowledgments

We thank J. Vaaler for several discussions and for help with the proof of Lemma 43.

This work has been supported by the *Comisión Conjunta Hispano Norteamericana de Cooperación Científica y Tecnológica*. The final version was prepared while R.L. was enjoying a *Cátedra de la Fundación FBBV*, and A.D. was visiting the *Centre de Recerca Matemàtica*, for whose hospitality he is very grateful. A.D. and T.S. have also been partially supported by the Catalan grant 2001SGR-70, the Spanish grant BFM2000-0805-C02 and the INTAS grant 00-221, and R.L. by NSF grants.

References

[AA67] V.I. Arnold and A. Avez. *Ergodic problems of classical mechanics.* Benjamin, New York, 1967.

[AKN88] V.I. Arnold, V.V. Kozlov, and A.I. Neishtadt. *Dynamical Systems III*, volume 3 of *Encyclopaedia Math. Sci.* Springer, Berlin, 1988.

[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. *Uspehi Mat. Nauk*, 18(6 (114)):91–192, 1963.

[Arn64] V.I. Arnold. Instability of dynamical systems with several degrees of freedom. *Sov. Math. Doklady*, 5:581–585, 1964.

[BB02] M. Berti and P. Bolle. A functional analysis approach to Arnold diffusion. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 19(4):395–450, 2002.

[BCV01] Ugo Bessi, Luigi Chierchia, and Enrico Valdinoci. Upper bounds on Arnold diffusion times via Mather theory. *J. Math. Pures Appl. (9)*, 80(1):105–129, 2001.

[Bes96] Ugo Bessi. An approach to Arnol′d's diffusion through the calculus of variations. *Nonlinear Anal.*, 26(6):1115–1135, 1996.

[BF02] I. Baldomà and E. Fontich. Exponentially small splitting of invariant manifolds of parabolic points. Preprint 02-166, `mp_arc@math.utexas.edu`, 2002.

[BT99] S. Bolotin and D. Treschev. Unbounded growth of energy in nonautonomous Hamiltonian systems. *Nonlinearity*, 12(2):365–388, 1999.

[Car81] John R. Cary. Lie transform perturbation theory for Hamiltonian systems. *Phys. Rep.*, 79(2):129–159, 1981.

[CFdlL03] Xavier Cabré, Ernest Fontich, and Rafael de la Llave. The parameterization method for invariant manifolds: I, II. *Indiana Univ. Math. J.*, 2003. To appear.

[CG94] L. Chierchia and G. Gallavotti. Drift and diffusion in phase space. *Ann. Inst. H. Poincaré Phys. Théor.*, 60(1):144, 1994.

[CG98] L. Chierchia and G. Gallavotti. Erratum drift and diffusion in phase space. *Ann. Inst. H. Poincaré Phys. Théor.*, 68:135, 1998.

[CG03] Jacky Cresson and Christophe Guillet. Periodic orbits and Arnold diffusion. *Discrete Contin. Dyn. Syst.*, 9(2):451–470, 2003.

[Chi79] B.V. Chirikov. A universal instability of many-dimensional oscillator systems. *Phys. Rep.*, 52(5):264–379, 1979.

[CLSV85] B. V. Chirikov, M. A. Lieberman, D. L. Shepelyansky, and F. M. Vivaldi. A theory of modulational diffusion. *Phys. D*, 14(3):289–304, 1985.

[CP02] Gonzalo Contreras and Gabriel P. Paternain. Connecting orbits between static classes for generic Lagrangian systems. *Topology*, 41(4):645–666, 2002.

[CSUZ89] A. A. Chernikov, R. Z. Sagdeev, D. A. Usikov, and G. M. Zaslavsky. *Weak chaos and structures*, volume 8 of *Soviet Scientific Reviews, Section C: Mathematical Physics Reviews*. Harwood Academic Publishers, Chur, 1989.

[CV89] B. V. Chirikov and V. V. Vecheslavov. How fast is the Arnold's diffusion?? Technical Report 98-72, Inst. Plasma Phys., Novosibirsk, 1989.

[DG96] A. Delshams and P. Gutiérrez. Effective stability and KAM theory. *J. Differential Equations*, 128(2):415–490, 1996.

[DG00] A. Delshams and P. Gutiérrez. Splitting potential and the Poincaré-Melnikov method for whiskered tori in Hamiltonian systems. *J. Nonlinear Sci.*, 10(4):433–476, 2000.

[DG01] A. Delshams and P. Gutiérrez. Homoclinic orbits to invariant tori in Hamiltonian systems. In Christopher K. R. T. Jones and Alexander I. Khibnik, editors, *Multiple-time-scale dynamical systems (Minneapolis, MN, 1997)*, pages 1–27. Springer, New York, 2001.

[DLC83] Raphaël Douady and Patrice Le Calvez. Exemple de point fixe elliptique non topologiquement stable en dimension 4. *C. R. Acad. Sci. Paris Sér. I Math.*, 296(21):895–898, 1983.

[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.

[DLS01] A. Delshams, R. de la Llave, and T.M. Seara. Orbits of unbounded energy in generic quasiperiodic perturbations of geodesic flows of certain manifolds, 2001. preprint.

[Dou88] R. Douady. Stabilité ou instabilité des points fixes elliptiques. *Ann. Sci. École Norm. Sup. (4)*, 21(1):1–46, 1988.

[EMR01] R. W. Easton, J. D. Meiss, and G. Roberts. Drift by coupling to an anti-integrable limit. *Phys. D*, 156(3-4):201–218, 2001.

[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.

[Fen74] N. Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[FL92] C. Falcolini and R. de la Llave. A rigorous partial justification of Greene's criterion. *J. Statist. Phys.*, 67(3-4):609–643, 1992.

[FM00] E. Fontich and P. Martín. Differentiable invariant manifolds for partially hyperbolic tori and a lambda lemma. *Nonlinearity*, 13(5):1561–1593, 2000.

[FM01] Ernest Fontich and Pau Martín. Arnold diffusion in perturbations of analytic integrable Hamiltonian systems. *Discrete Contin. Dynam. Systems*, 7(1):61–84, 2001.

[FM03] Ernest Fontich and Pau Martín. Hamiltonian systems with orbits covering densely submanifolds of small codimension. *Nonlinear Anal.*, 52(1):315–327, 2003.

[FS90a] E. Fontich and C. Simó. Invariant manifolds for near identity differentiable maps and splitting of separatrices. *Ergodic Theory Dynam. Systems*, 10(2):319–346, 1990.

[FS90b] E. Fontich and C. Simó. The splitting of separatrices for analytic diffeomorphisms. *Ergodic Theory Dynam. Systems*, 10(2):295–318, 1990.

[Gal94] G. Gallavotti. Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review. *Rev. Math. Phys.*, 6(3):343–411, 1994.

[Gal99] G. Gallavotti. Arnold's diffusion in isochronous systems. *Math. Phys. Anal. Geom.*, 1(4):295–312, 1998/99.

[Gra74] Samuel M. Graff. On the conservation of hyperbolic invariant tori for Hamiltonian systems. *J. Differential Equations*, 15:1–69, 1974.

[Hal97] G. Haller. Universal homoclinic bifurcations and chaos near double resonances. *J. Statist. Phys.*, 86(5-6):1011–1051, 1997.

[Hal99] G. Haller. *Chaos near resonance.* Springer-Verlag, New York, 1999.

[Her79] M.-R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Inst. Hautes Études Sci. Publ. Math.*, (49):5–233, 1979.

[Her83] M.-R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, volume 103 of *Astérisque*. Société Mathématique de France, Paris, 1983.

[HL00] A. Haro and R. de la Llave. New mechanisms for lack of equipartion of energy. *Phys. Rev. Lett.*, 89(7):1859–1862, 2000.

[HM82] P.J. Holmes and J.E. Marsden. Melnikov's method and Arnol'd diffusion for perturbations of integrable Hamiltonian systems. *J. Math. Phys.*, 23(4):669–675, 1982.

[HPS77] M.W. Hirsch, C.C. Pugh, and M. Shub. *Invariant manifolds*, volume 583 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1977.

[JVMU99] G. H. F. Diercksen J. Von Milczewski and T. Uzer. The arnold web in atomic physics. In *Hamiltonian Systems with Three or More Degrees of Freedom (S'Agaró, 1995)*, pages 499–503. Kluwer Acad. Publ., Dordrecht, 1999.

[Las93] Jacques Laskar. Frequency analysis for multi-dimensional systems. Global dynamics and diffusion. *Phys. D*, 67(1-3):257–281, 1993.

[LHRK02] Anna Litvak-Hinenzon and Vered Rom-Kedar. Resonant tori and instabilities in Hamiltonian systems. *Nonlinearity*, 15(4):1149–1177, 2002.

[Lla00] R. de la Llave. Persistence of hyperbolic manifolds, 2000.

[Lla01] R. 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.

[Lla02] R. Llave. Orbits of unbounded energy in perturbations of geodesic flows by periodic potentials. a simple construction. Preprint, 2002.

[LM88] P. Lochak and C. Meunier. *Multiphase Averaging for Classical Systems*, volume 72 of *Appl. Math. Sci.* Springer, New York, 1988.

[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.

[LT83] M. A. Lieberman and Jeffrey L. Tennyson. Chaotic motion along resonance layers in near-integrable Hamiltonian systems with three or more degrees of freedom. In *Long-time prediction in dynamics (Lakeway, Tex., 1981)*, pages 179–211. Wiley, New York, 1983.

[LW89] R. de la Llave and C.E. Wayne. Whiskered and lower dimensional tori in nearly integrable Hamiltonian systems. Preprint, 1989.

[LW95] R. de la Llave and C. E. Wayne. On Irwin's proof of the pseudostable manifold theorem. *Math. Z.*, 219(2):301–321, 1995.

[Mat93] John N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)*, 43(5):1349–1386, 1993.

[Mat95] J.N. Mather. Graduate course at Princeton, 95–96, and Lectures at Penn State, Spring 96, Paris, Summer 96, Austin, Fall 96, 1995.

[Mat02] J. N. Mather. Arnold diffusion I: Announcement of results. *Preprint*, 2002.

[Mei92] J. D. Meiss. Symplectic maps, variational principles, and transport. *Rev. Modern Phys.*, 64(3):795–848, 1992.

[Mey91] K. R. Meyer. Lie transform tutorial. II. In *Computer Aided Proofs in Analysis (Cincinnati, OH, 1989)*, pages 190–210. Springer, New York, 1991.

[Moe96] Richard Moeckel. Transition tori in the five-body problem. *J. Differential Equations*, 129(2):290–314, 1996.

[Moe02] Richard Moeckel. Generic drift on Cantor sets of annuli. In *Celestial mechanics (Evanston, IL, 1999)*, pages 163–171. Amer. Math. Soc., Providence, RI, 2002.

[Mos73] J. Moser. *Stable and random motions in dynamical systems*. Princeton University Press, Princeton, N. J., 1973. With special emphasis on celestial mechanics, Hermann Weyl Lectures, the Institute for Advanced Study, Princeton, N. J, Annals of Mathematics Studies, No. 77.

[Neĭ81] A. I. Neĭshtadt. Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions. *Prikl. Mat. Mekh.*, 45(6):1016–1025, 1981.

[Neĭ84] A. I. Neĭshtadt. The separation of motions in systems with rapidly rotating phase. *Prikl. Mat. Mekh.*, 48(2):197–204, 1984.

[Nie00] L. Niederman. Dynamics around simple resonant tori in nearly integrable Hamiltonian systems. *J. Differential Equations*, 161(1):1–41, 2000.

[Poi99] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste*, volume 1, 2, 3. Gauthier-Villars, Paris, 1892–1899.

[Pös82] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.*, 35(5):653–696, 1982.

[RS02] Paul H. Rabinowitz and Edward W. Stredulinsky. A variational shadowing method. In *Celestial mechanics (Evanston, IL, 1999)*, volume 292 of *Contemp. Math.*, pages 185–197. Amer. Math. Soc., Providence, RI, 2002.

[Sal86] D. Salamon. The Kolmogorov-Arnold-Moser theorem. *Zürich preprint*, 1986.

[Sor02] A. Sorrentino. *Sulle soluzioni quasi-periodiche di sistemi Hamiltoniani differenziabili*. PhD thesis, Univ. di Roma Tre, 2002.

[Sva80] N. V. Svanidze. Small perturbations of an integrable dynamical system with an integral invariant. *Trudy Mat. Inst. Steklov.*, 147:124–146, 204, 1980. English translation: *Proc. Steklov Inst. Math.*, 1981, no. 2.

[SZF95] Michael F. Shlesinger, George M. Zaslavsky, and Uriel Frisch, editors. *Lévy flights and related topics in physics*, Berlin, 1995. Springer-Verlag.

[Ten82] Jeffrey Tennyson. Resonance transport in near-integrable systems with many degrees of freedom. *Phys. D*, 5(1):123–135, 1982.

[TLL80] J. L. Tennyson, M. A. Lieberman, and A. J. Lichtenberg. Diffusion in near-integrable Hamiltonian systems with three degrees of freedom. In *Nonlinear dynamics and the beam-beam interaction (Sympos., Brookhaven Nat. Lab., New York, 1979)*, pages 272–301. Amer. Inst. Physics, New York, 1980.

[Tre89] D. V. Treshchëv. A mechanism for the destruction of resonance tori in Hamiltonian systems. *Mat. Sb.*, 180(10):1325–1346, 1439, 1989.

[Tre02] D. Treschev. Multidimensional symplectic separatrix maps. *J. Nonlinear Sci.*, 12(1):27–58, 2002.

[Val00] Enrico Valdinoci. Families of whiskered tori for a-priori stable/unstable Hamiltonian systems and construction of unstable orbits. *Math. Phys. Electron. J.*, 6:Paper 2, 31 pp. (electronic), 2000.

[Xia98] Zhihong Xia. Arnold diffusion: a variational construction. In *Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998)*, number Extra Vol. II, pages 867–877 (electronic), 1998.

[Zas02] G. M. Zaslavsky. Chaos, fractional kinetics, and anomalous transport. *Phys. Rep.*, 371(6):461–580, 2002.

[Zeh76] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems/II. *Comm. Pure Appl. Math.*, 29:49–111, 1976.

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain

E-mail address, AD: `Amadeu.Delshams@upc.es`

Department of Mathematics, University of Texas, Austin, TX 78712-1802

E-mail address, RL: `llave@math.utexas.edu`

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain

E-mail address, TS: `tere.m-seara@upc.es`