INSTABILITY OF HIGH DIMENSIONAL HAMILTONIAN SYSTEMS: MULTIPLE RESONANCES DO NOT IMPEDE DIFFUSION.

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

ABSTRACT. We consider models given by Hamiltonians of the form

$$H(I, \varphi, p, q, t; \varepsilon) = h(I) + \sum_{j=1}^{n} \pm \left(\frac{1}{2} p_j^2 + V_j(q_j) \right) + \varepsilon Q(I, \varphi, p, q, t; \varepsilon)$$

where $I \in \mathcal{I} \subset \mathbb{R}^d$, $\varphi \in \mathbb{T}^d$, $p, q \in \mathbb{R}^n$, $t \in \mathbb{T}^1$. These are higher dimensional analogues, both in the center and hyperbolic directions, of the models studied in [DLS03, DLS06a, GL06a, GL06b]. All these models present the *large gap problem*.

We show that, for $0 < \varepsilon \ll 1$, under regularity and explicit non-degeneracy conditions on the model, there are orbits whose action variables I perform rather arbitrary excursions in a domain of size $O(1)$. This domain includes resonance lines and, hence, large gaps among d-dimensional KAM tori.

The method of proof follows closely the strategy of [DLS03, DLS06a]. The main new phenomenon that appears when the dimension d of the center directions is larger than one, is the existence of multiple resonances. We show that, since these multiple resonances happen in sets of codimension greater than one in the space of actions I, they can be contoured. This corresponds to the mechanism called *diffusion across resonances* in the Physics literature.

The present paper, however, differs substantially from [DLS03, DLS06a]. On the technical details of the proofs, we have taken advantage of the theory of the scattering map [DLS08], not available when the above papers were written. We have analyzed the conditions imposed on the resonances in more detail.

More precisely, we have found that there is a simple condition on the Melnikov potential which allows us to conclude that the resonances are crossed. In particular, this condition does not depend on the resonances. So that the results are new even when applied to the models in [DLS03, DLS06a].

CONTENTS

1. Introduction

The goal of this paper is to present some explicit sufficient conditions for instability in models of the form

$$(1) \qquad H(I, \varphi, p, q, t; \varepsilon) = h(I) + P(p, q) + \varepsilon Q(I, \varphi, p, q, t; \varepsilon),$$

where

$$P(p, q) = \sum_{j=1}^{n} P_j(p_j, q_j), \quad P_j(p_j, q_j) = \pm \left(\frac{1}{2} p_j^2 + V_j(q_j) \right).$$

We will assume that $I \in \mathcal{I} \subset \mathbb{R}^d, \varphi \in \mathbb{T}^d, \mathcal{I}$ an open set, $p, q \in \mathbb{R}^n$. The symplectic form of the phase space is $\Omega = \sum_i dI_i \wedge d\varphi_i + \sum_j \mathrm{d}p_j \wedge dq_j$. We will assume that the dependence of Q on t is 1-periodic, so that $t \in \mathbb{T}^1$. Moreover, for simplicity, we will assume that Q is a trigonometric polynomial in the variables (φ, t). This is not a crucial assumption and can be eliminated. See remark 3 for details about this assumption.

We will show (see Theorem 5 for precise statements) that, under suitable regularity and non-degeneracy assumptions—that can be checked by studying the 3 jet in the ε variable of the perturbation εQ—there are orbits of the system in which the I variables can perform largely arbitrary excursions in a set $\mathcal{I}^* \subset \mathcal{I}$ of size of order 1 (that is, independent of ε as $\varepsilon \to 0$). An explicit example is shown in Section 4.

The main point is that the set $\mathcal{I}^*$ can include simple resonant surfaces, that is, resonances of multiplicity one. This makes the models considered here present the *large gap problem*. This problem, which will be discussed in more detail in Section 3.5, consists in the fact that the customary perturbation theory does not produce chains of whiskered KAM tori with transverse heteroclinic intersections. The reason is that a perturbation of size ε causes gaps in the set of whiskered KAM tori of order $\varepsilon^{1/2}$ near the (first order in ε) resonance surfaces of order 1. On the other hand, the effect of the perturbation on the stable and unstable manifolds of these whiskered tori is only of $O(\varepsilon)$. Hence, a naive perturbation theory cannot establish the existence of chains of transition tori traversing the resonance surface. The main goal of this paper is to describe a geometric mechanism, based on a the computation of a *scattering map* defined on the normally hyperbolic invariant manifold of the system, which is used to establish the existence of chains of transition tori that can traverse these resonance regions devoid of primary KAM tori, that is, tori that can be continued from invariant tori of the integrable system. The main idea of the mechanism proposed is to use secondary tori, that is, tori generated by the resonances that can not be continued from invariant tori of the integrable system. We will show that these secondary tori generated by the resonances fill the gaps created in the set of KAM primary tori.

The result for $d = 1$, $n = 1$ was already established in [DLS03, DLS06a]. The problem was reexamined in [DH09], where the hypothesis about Q being a trigonometric polynomial in (φ, t) was eliminated. The works [GL06a, GL06b] supplemented the methods of

[DLS03, DLS06a] with the use of the method of correctly-aligned windows. This method allowed to simplify the proof and to obtain explicit estimates on the time spent by the diffusion trajectories. In [GL06b], which considers n arbitrary, a particularly careful version of the method of correctly aligned windows produces essentially optimal estimates on time and allows to weaken the non-degeneracy assumptions in [DLS03, DLS06a].

The case $d \geq 2$ presents a difficulty that was not present in the case $d = 1$, namely, that there are points where the resonances have higher multiplicity (the multiplicity of a resonance is the dimension of its resonance module, see (33)).

The technique used in [DLS03, DLS06a] was to take advantage of the fact that, in the neighborhood of simple resonances, that is, resonances of multiplicity 1, it is possible to introduce a normal form which is integrable, and can be analyzed with great accuracy. Unfortunately, it is well known that multiple resonances, that is, resonances of multiplicity greater or equal than 2, lead to normal forms that are not integrable and need other techniques to be analyzed (see [AKN88, Hal99]). Recent progress in the analysis of double resonances can be found in [Mar13, KZ12, Che12]. We also note that [Tre12] establishes diffusion far from *strong resonances* for the case $n = 1$, $d \geq 2$, using the method of the separatrix map.

In this paper we adapt the methods of the previous papers [DLS03, DLS06a] to show instability under explicit conditions. The basic observation is that multiple resonances happen in subsets of codimension greater than 1 in $\mathcal{I}$. We will adapt the methods of [DLS03, DLS06a] to analyze the behavior of the system in regions of simple resonances and show that the diffusing trajectories can contour the multiple resonances.

As we will see in Theorem 5, we can choose largely arbitrary paths in the action space $\mathcal{I}$ (we just require that they do not pass through some higher codimension subsets of multiple resonances) and then, show that there are orbits whose I-projection follows these paths up to an error, which becomes arbitrarily small with ε. The sets that can be reached are of size $O(1)$ and they include simple resonances.

Similar definitions of diffusion along paths were also used in [CG94, CG98], but the methods of these and related papers [BB02, BBB03] only established the existence of diffusion in sets completely devoid of resonances (the so-called "gap bridging mechanism"). The orbits that we produce cross the codimension 1 resonant surfaces of multiplicity one. Similar phenomena have been observed in the heuristic literature

4

[Chi79, LT83]. In [Ten82], a similar mechanism of diffusion is called *diffusion across resonances*.

Of course, we are far from believing that the mechanism discussed in this paper is the only one to produce changes of order one in the actions. In particular, combinations of variational and geometric methods have been recently applied in Hamiltonians with $2 + 1/2$ degrees of freedom [Mat04, Che12, Zha11, KZ12, Ber10, Mar13, BCV01] which require the Hamiltonian to be positive definite, which is a non-generically (albeit open) property.

Also, the paper of [Ten82] suggests several other mechanisms that should be at play. It seems to be a very challenging problem to make rigorous the heuristic discussions on statistical and quantitative properties of different instability mechanisms in the heuristic literature [Chi79, LT83, Ten82]. Of course, the heuristic literature is convinced that double resonances help diffusion because they are one of the ingredients in some of the heuristic mechanisms (but not in others!). It is somewhat paradoxical that the rigorous mathematical theory has difficulty precisely at the places which heuristics considers favorable.

In this paper, we show rigorously that double resonances can be contoured. This can enhance the believe that there are several mechanisms.

Note however that if we consider (1) as a model of what happens in a resonance (the pendulum being the resonant variable), then the multiplicity of the resonances in the real system is one more than the order that appears in the model. Hence, what we call simple resonances in our Hamiltonian (modeling a resonance in a real model) would be double resonances in the real Hamiltonian.

Since the proof presented here is quite modular and has well defined milestones, we think that it is almost certain that other methods can be applied to improve some of our arguments. In particular, we expect that the method of correctly aligned windows can also give alternative proofs or to improve several steps of the proof. The field of instability has experienced a great deal of activity in recent years and there is a large variety of results that have been obtained or announced. For a more detailed survey of recent results, we refer to [DGLS08, PT07, Che08, Che10, Ber10].

Remark 1. It is customary in some literature to refer to models of the form (1) as a-priori-unstable models. We note, however, that this distinction only makes sense in considering analytic models depending only on one small parameter. The results we present here apply just as well when the potentials V_i in (1) are arbitrarily close to 0. In such

a case, we just need to choose ε very small (even exponentially small) relatively to the hyperbolicity properties of the V_i. Several papers in the literature, notably those dealing with generic results, which occur in typical or cusp-residual Hamiltonian, call these systems "a priori stable". In particular, one can use this method to produce systems that present instability but which are as close to integrable as desired. This procedure was pioneered in [Arn64].

Remark 2. Hamiltonian (1) can be considered as a simplified model of what happens in a neighborhood of a resonance of multiplicity n in a near integrable Hamiltonian. The averaging method [LM88, Hal97, Hal99, AKN88] shows that near a resonance of multiplicity j, one can reduce a near integrable Hamiltonian to a Hamiltonian of the form

$$(2) \qquad h(I) + \sum_{i=1}^{n} \frac{p_i^2}{2} + \varepsilon V(q_1, \ldots, q_n, I) + O(\varepsilon^2).$$

The assumption that the averaged system is given by uncoupled pendula is made often [HM82, Hal97]. It is a *generic* assumption for $n = 1$. Hence we expect that the mechanism presented here is typical in a neighborhood of a resonance. Of course, the hyperbolicity will be weak in systems close to integrable, but in families with two parameters, it would suffice to exclude wedges. See Remark 1.

For $n \geq 2$, the above model (2) is, in general, not integrable whereas the pendulum part of (1) is . Nevertheless, we point out that the only think we need for our analysis is that $\sum_{i=1}^{n} \frac{p_i^2}{2} + \varepsilon V(q_1, \ldots, q_n, I)$ admits transversal homoclinic orbits to a hyperbolic equilibrium point. Systems of the form (1) appear naturally in several physical models. A motivation to include this generality is that there is very little difference dealing with any n and it allows to emphasize that the geometric methods allow to deal with systems that are not positive definite.

2. Notation, assumptions and results

In this section, we will present an overview of the argument and formulate precisely most of the non-degeneracy assumptions we will assume. We will postpone the precise formulation of the most technical ones till we have developed the notation for them and motivated their explicit expressions.

The proof is divided in well defined steps and each of them can be accomplished using standard tools. We hope that the experts in these techniques can fill in the arguments better than the authors, so that for many possible readers, the heuristic discussion will be enough.

We have found it convenient to present the argument in an order slightly different from the one followed in [DLS03, DLS06a] so that, even if our hypotheses correspond closely to the assumptions of these papers, the numbers do not correspond. In some cases, we have chosen to present the result under slightly different assumptions than in [DLS03, DLS06a] to simplify the exposition. For the same reason—simplifying and shortening the exposition—some of the objects whose expansions were explicitly computed in [DLS06a] will now be given through existence theorems, so that the final conditions will become slightly less explicit. Nevertheless, since the procedures here are rather constructive, explicit formulas can be given through more detailed work. On the other hand, we note that the tool of the scattering map and its symplectic properties [DLS08], a tool which was not available when [DLS06a] was written, simplifies significantly the computations and thus, the conditions we obtain in this paper are simpler to verify and more generally applicable than those in [DLS06a].

The precise statement of the main result (Theorem 5) requires the definition of the resonance web, which depends on assumption **H3**. The statement of the last non-degeneracy conditions, **H6**, **H7**, **H8**, can only be made after the system has been analyzed near resonances. We note that these conditions are verifiable in concrete models with a finite calculation, as it is performed in the example (120).

2.1. **Some elementary notation: the extended flow, the time-one map.** We will always consider the extended flow $\tilde{\Phi}_{\varepsilon,t}(\tilde{x})$ which is obtained by supplementing the standard Hamilton equations with the equation $\dot{s} = 1$:

$$
\begin{aligned}
\dot{I} &= -\varepsilon \frac{\partial Q}{\partial \varphi}(I, \varphi, p, q, s; \varepsilon) \\
\dot{\varphi} &= \frac{\partial h}{\partial I}(I) + \varepsilon \frac{\partial Q}{\partial I}(I, \varphi, p, q, s; \varepsilon) \\
\dot{p} &= -\frac{\partial P}{\partial q}(p, q) - \varepsilon \frac{\partial Q}{\partial q}(I, \varphi, p, q, s; \varepsilon) \\
\dot{q} &= \frac{\partial P}{\partial p}(p, q) + \varepsilon \frac{\partial Q}{\partial p}(I, \varphi, p, q, s; \varepsilon) \\
\dot{s} &= 1
\end{aligned}
$$

(3)

To the extended differential equations (3) corresponds the extended phase space $\tilde{M} := \mathcal{I} \times \mathbb{T}^d \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{T}$ associated to the variables $\tilde{x} = (I, \varphi, p, q, s)$ respectively.

Some of our calculations are made easier by considering the time-1 map of the flow. We will use the notation f_ε to denote the time one map starting at the initial condition $t = 0$.

2.2. The first elementary assumptions: regularity, hyperbolicity of the pendula and non-degeneracy of the integrable part.

We will be making the following assumptions:

- **H1** We will assume that the functions h, V_j, Q are C^r in their corresponding domains with $r \geq r_0$ sufficiently large.

- **H2** We will assume that the potentials V_j have non-degenerate local maxima each of which gives rise, at least, to a homoclinic orbit of the pendulum P_j.

Without loss of generality and to simplify the notation, we will assume that the maxima of the potentials V_j happen at $q_j = 0$. That is, we will assume that $V_j'(0) = 0$, $V_j''(0) = -\alpha_j^2$ with $\alpha_j \geq \alpha > 0$, $j = 1, \ldots, n$.

We will denote by $(p_j^*(t), q_j^*(t))$ a parameterization by the natural time of the homoclinic orbit we have chosen. That is,

$$
(4) \qquad
\begin{aligned}
&\frac{d}{dt} q_j^*(t) = p_j^*(t); \\
&\frac{d}{dt} p_j^*(t) = -V_j'(q_j^*(t)); \\
&\lim_{t \to \pm\infty} (p_j^*(t), q_j^*(t)) = (0, 0).
\end{aligned}
$$

When the variables q_j have the physical interpretation of angles it is natural to assume that the V_j are periodic. In such a case, the limit in (4) is understood modulus the period of the potential. The method of proof only requires the existence of the homoclinic orbits to hyperbolic saddles. Hence it applies to the coupling of an integrable system in the (I, φ) variables to a chaotic system in the (p, q) variables.

We note that the choice of a parameterization of the homoclinic orbit in the full space involves the choice of n origins of time in each of the homoclinic orbits. Subsequent hypotheses will be independent of these choices. The possibility of choosing the origin of the parameterizations of each of the homoclinic orbits independently will play an important role in our discussion of the Poincaré function in (9).

Once we have chosen a homoclinic orbit to the origin in any pendulum, we obtain a homoclinic connection in the space of the pendula. We will denote by $\mathcal{U} \subset \mathbb{R}^n \times \mathbb{R}^n$ a neighborhood of the homoclinic connection chosen in the p, q space.

In this paper, we will assume that the equilibrium points are hyperbolic. It would be interesting to extend the result to degenerate maxima (leading to weakly hyperbolic points). This case has been proposed in the literature [DMM09].

From now on during the paper, we will consider $\mathcal{I}^* \subset \mathcal{I}$, and we consider the compact set

$$\text{(5)} \qquad \mathcal{D} := \mathcal{I}^* \times \mathbb{T}^d \times \mathcal{U} \times \mathbb{T}^1$$

to be the domain of our problem. So, all the hypotheses refer to this domain.

- **H3** The mapping $I \to \omega(I) := \frac{\partial}{\partial I} h(I)$ is a diffeomorphism from $\mathcal{I}^*$ to its image.

2.3. **Assumption on the structure of the perturbation.** We will furthermore assume:

- **H4** The function Q in (1) is a trigonometric polynomial on (φ, t).

That is, we can write

$$\text{(6)} \qquad Q(I, \varphi, p, q, t; \varepsilon) = \sum_{(k,l) \in \mathcal{N}_Q} Q_{k,l}(I, p, q; \varepsilon) e^{2\pi i(k \cdot \varphi + lt)}$$

with $\mathcal{N}_Q \subset \mathbb{Z}^d \times \mathbb{Z}$ a *finite set*, with $Q_{k,l} \not\equiv 0$ in $\mathcal{I}^* \times \mathcal{U}$, if $(k, l) \in \mathcal{N}_Q$.

Hypothesis **H4** clearly does not belong in the problem and we hope to eliminate it in future treatments. Since the main goal of this paper is to deal with the issue of multiple resonances, we have thought it convenient to make the result as easy to read as possible, even if we do not achieve the largest possible level of generality.

Remark 3. Hypothesis **H4** appeared in [DLS06a] for the case $d = 1$, $n = 1$. In that case, the trigonometric polynomial hypothesis has been eliminated in [DH09, GL06b] under generic assumptions. The paper [GL06b] eliminated the trigonometric polynomial hypothesis for $d = 1$, n arbitrary (one can argue that, possibly, the orbits produced are not the same as the orbits in the previous papers). Similar improvements are clearly possible in the higher dimensional case $d > 1$.

Remark 4. The methods we will use here can reach the same conclusions under slightly weaker hypotheses.

We only need that, for some big enough but finite $m \leq r$ the sets of integer indexes

$$\text{(7)} \qquad \{(k, l) \in \mathcal{N}_Q; \quad \frac{\partial^i}{\partial \varepsilon^i} Q_{k,l}(I, p = 0, q = 0; \varepsilon = 0) \neq 0\}$$

for $i \leq m$ are finite. This happens, in fact, for some models in celestial mechanics [FGKR11]. Nevertheless, since the writing in this case becomes more cumbersome, we will only claim the weaker result.

2.4. **Non degeneracy assumptions.** The first non-degeneracy assumption concerns the averaged Hamiltonian near simple resonances, and are stated in Section 3.5.3.

- **H5** Consider the set of integer indexes $\mathcal{N}^{[\leq 2]} = \mathcal{N}_1 \cup \mathcal{N}_2 \subset \mathbb{Z}^{d+1}$ where $\mathcal{N}_1$ is the support of the Fourier series of the perturbation $Q(I, \varphi, p, q, t; 0)$, $\mathcal{N}_2 = (\mathcal{N}_1 + \mathcal{N}_1) \cup \bar{\mathcal{N}}$, where $\bar{\mathcal{N}}$ is the support of the Fourier series of $\frac{\partial Q}{\partial \varepsilon}(I, \varphi, p, q, t; 0)$.

  Then we assume that, for any $(k, l) \neq (0, 0) \in \mathcal{N}^{[\leq 2]}$, the set

$$
(8) \qquad \{I \in \mathcal{I}^*, Dh(I)k + l = 0, \quad k^\top D^2 h(I)k = 0\}
$$

  is empty or a manifold of codimension at least two in $\mathcal{I}^*$.

  In the case the map $\tilde{h}(I_0, I) = I_0 + h(I)$ is a quasi convex function the set (8) is an empty set for any $(k, l) \neq (0, 0) \in \mathbb{Z}^d \times \mathbb{Z}$ and a fortiori for any $(k, l) \neq (0, 0) \in \mathcal{N}^{[\leq 2]}$. Therefore Hypothesis **H5** is true for any perturbation Q in this case.

- **H6** Assume that the perturbation Q satisfies some non-degeneracy conditions stated in Section 3.5.3 in the connected domain $\mathcal{I}^* \times \mathbb{T}^{d+1}$ related to the averaged Hamiltonian.

The following non-degeneracy assumptions concern the so called Poincaré function (or Melnikov potential) associated to the homoclinic connection (p^*, q^*) chosen in assumption **H2**:

$$
(9)
$$

$$
L(\tau, I, \varphi, s) = -\int_{-\infty}^{\infty} \left[Q(I, \varphi + \omega(I)\sigma, p^*(\tau + \sigma), q^*(\tau + \sigma), s + \sigma; 0) \right.
$$

$$
\left. - Q(I, \varphi + \omega(I)\sigma, 0, 0, s + \sigma; 0) \right] d\sigma
$$

where

$$
\begin{aligned}
\tau &= (\tau_1, \ldots, \tau_n) \\
p^*(\tau + \sigma) &= (p_1^*(\tau_1 + \sigma), \ldots, p_n^*(\tau_n + \sigma)) \\
q^*(\tau + \sigma) &= (q_1^*(\tau_1 + \sigma), \ldots, q_n^*(\tau_n + \sigma))
\end{aligned}
$$

- **H7** Assume that, for any value of $I \in \mathcal{I}^*$, there exists a non-empty set $\mathcal{J}_I \subset \mathbb{T}^{d+1}$, with the property that when $(I, \varphi, s) \in H_-$, where

$$
(10) \qquad H_- = \bigcup_{I \in \mathcal{I}^*} \{I\} \times \mathcal{J}_I \subset \mathcal{I}^* \times \mathbb{T}^{d+1},
$$

the system of equations

$$(11) \qquad \frac{\partial}{\partial \tau} L(\tau, I, \varphi, s) = 0$$

admits a non degenerate solution $\tau = \tau^*(I, \varphi, s)$ with τ^* a smooth function.

- **H8** Define the auxiliary functions (related to the scattering map that will be introduced in Section 3.6)

$$(12) \qquad \mathcal{L}(I, \varphi, s) = L(\tau^*(I, \varphi, s), I, \varphi, s), \quad \mathcal{L}^*(I, \theta) = \mathcal{L}(I, \theta, 0)$$

Assume that the *reduced Poincaré function* $\mathcal{L}^*(I, \varphi - \omega(I)s)$ satisfies some non-degeneracy conditions stated in Section 3.8 in the domain H_- (see (98), (114)). Nevertheless we anticipate that an informal description of the hypothesis will be discussed after the main theorem 5.

We also note that the hypothesis **H8** is simplified by the very simple hypothesis:

- **H8':** $\forall I \in \mathcal{I}^*$, the reduced Poincaré function $\mathcal{L}^*(I, \theta)$ defined in (12) has non degenerate critical points.

2.5. Statement of the main result.
The main result of this paper is the following:

Theorem 5. *Let H be a Hamiltonian of the form* (1) *satisfying the elementary assumptions* **H1, H2**, *the regularity assumption* **H3**, *the simplifying assumption* **H4** *and the non-degeneracy assumptions* **H5, H6, H7, H8**.

Then, for every $\delta > 0$, there exists $\varepsilon_0 > 0$, such that for every $0 < |\varepsilon| < \varepsilon_0$, given $I_\pm \in \mathcal{I}^*$, there exists an orbit $\tilde{x}(t)$ of (1) *and $T > 0$, such that:*

$$(13) \qquad \begin{aligned} |I(\tilde{x}(0)) - I_-| &\leq C\delta \\ |I(\tilde{x}(T)) - I_+| &\leq C\delta. \end{aligned}$$

Actually, we will show that given a largely arbitrary path $\gamma(s) \subset \mathcal{I}^*$, we can find orbits $\tilde{x}(t)$ such that $I(\tilde{x}(t))$ is δ-close to $\gamma(\Psi(t))$ for some reparameterization Ψ. We postpone the precise statement till we have developed the notation. See Theorem 26.

The set $\mathcal{I}^*$ will be described precisely in the course of the proof. The set is determined by the non-degeneracy assumptions **H5**, **H6**, **H7** and **H8**. Given any concrete system, the assumptions can be verified from the finite jet in ε of H. Therefore, these conditions hold in regions of order 1 of $\mathcal{I}$.

The main restriction to obtain the set $\mathcal{I}^*$ from the domain of definition $\mathcal{I}$ is given by assumption **H7**, which guarantees the transversality of the intersection of the stable and unstable manifolds of the perturbed normally hyperbolic manifold Λ_ε obtained in section 3.1. Once **H7** is imposed, one needs to eliminate some sets of codimension two from it to obtain $\mathcal{I}^*$:

- **H8** Eliminates the values of I for which the scattering map is not transversal to the inner map. More precisely, the invariant KAM tori of the inner map are transverse to their images under the scattering map. See Section 3.8.

  In fact, the different conditions given in Section 3.8 which constitute hypothesis **H8** can be replaced by the sufficient condition **H8'**.

- **H5** and **H6** Eliminate the region in the resonances where the leading term of the averaged system is degenerate. That is, given the codimension 1 resonant surfaces, we have to eliminate the place when some function vanishes. See Section 3.5.3.

Since $\mathcal{I}^*$ can contain multiple resonances appearing up to finite order averaging theory, and our mechanism is based on avoiding these multiple resonances, we choose $\delta > 0$ and contour these codimension two sets up to a distance of order $O(\delta)$, obtaining a reduced domain $\mathcal{I}_\delta \subset \mathcal{I}^*$. The precise definition of the sets to eliminate is deferred to Section 3.3. Roughly, we eliminate the double resonances in which one of the resonances are or order 1 or 2 (double resonances in which both resonances are of order higher than 2 are allowed in $\mathcal{I}_\delta$).

Note that, since by **H4**, the perturbation is a trigonometric polynomial and we assume hypothesis **H3**, we only need to eliminate the intersection of finite number of codimension one manifolds.

We note that all the conditions **H5-H8** are generic: C^2 open in the space of Hamiltonians and hold except in sets of infinite codimension. Note that all these hypothesis are transversality conditions among objects that are independent. That is, we require transversality conditions among objects that depend on the perturbation restricted to different places. See the details later.

The only hypothesis that is not generic in the set-up is assumption **H4**. It seems clear that this assumption **H4** can be eliminated using the techniques developed in [DH09]. However, we have preferred to maintain it to simplify the exposition. Roughly, the idea is that, for every $\varepsilon > 0$ one can approximate the perturbation by a trigonometric polynomial. If the trigonometric polynomial verifies the hypotheses of Theorem 5, one can obtain the existence of wandering paths and

information about their robustness. As it turns out, if the perturbation is smooth enough, one can show that truncation error is much smaller than the robustness allowed by the mechanism. Of course, there are quite a number of details to be verified and we will not endeavor to do them now.

Once we have defined the set $\mathcal{I}_\delta \subset \mathcal{I}^*$, we will show that KAM tori–either primary or secondary–are closely spaced on it. We will show that, given any KAM torus with I coordinates in $\mathcal{I}_\delta$, it has transversal heteroclinic connections with all the other KAM tori in a small neighborhood. Applying the shadowing lemma, we can find orbits whose actions follow almost arbitrary paths inside $\mathcal{I}_\delta$. This is, of course, slightly stronger than the conclusion (13) of Theorem 5.

It is important to note that codimension 2 objects do not separate the regions and can be contoured so that they do not obstruct the change along the paths.

3. Proof of Theorem 5

3.1. **First step: The use of normal hyperbolicity.** We note that for $\varepsilon = 0$, the manifold

$$\Lambda_0 = \{p = 0, q = 0, \ I \in \mathcal{I}^*, \varphi \in \mathbb{T}^d\}$$

is locally invariant under f_0, the time-1 map of the flow.

In this paper, we will not only work with the time one map but also with the flow $\tilde{\Phi}_{0,t}$ of system (3), so we will work in the extended phase space $\tilde{\mathcal{M}} = \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{I} \times \mathbb{T}^d \times \mathbb{T}$.

In the extended phase space, we consider the invariant manifold

$$\tilde{\Lambda}_0 = \{p = 0, q = 0, \ I \in \mathcal{I}^*, \varphi \in \mathbb{T}^d, s \in \mathbb{T}^1\}$$

which is a normally hyperbolic invariant manifold under the flow $\tilde{\Phi}_{0,t}$ in the sense of [Fen72, Fen74, HPS77, Pes04]. That is, for every $\tilde{x} \in \tilde{\Lambda}_0$, there is a decomposition

$$(14) \qquad\qquad T_{\tilde{x}}\tilde{\mathcal{M}} = E_{\tilde{x}}^s \oplus E_{\tilde{x}}^u \oplus T_{\tilde{x}}\tilde{\Lambda}_0$$

and numbers $C > 0$, $0 < \beta < \alpha$, such that the decomposition (14) is characterized by:

$$v \in E_{\tilde{x}}^s \iff |D\tilde{\Phi}_{0,t}(\tilde{x})v| \le Ce^{-\alpha t}|v| \quad \forall t \ge 0$$

$$(15) \qquad v \in E_{\tilde{x}}^u \iff |D\tilde{\Phi}_{0,t}(\tilde{x})v| \le Ce^{-\alpha|t|}|v| \quad \forall t \le 0$$

$$v \in T_{\tilde{x}}\tilde{\Lambda}_0 \iff |D\tilde{\Phi}_{0,t}(\tilde{x})v| \le Ce^{\beta|t|}|v| \quad \forall t \in \mathbb{R}$$

It is clear that the stable and unstable spaces $E_{\tilde{x}}^{s,u}$ are the direct sum of the stable and unstable spaces at the critical point of each of the

pendula P_j, and α is given in assumption **H2**. Furthermore, we can take for β any number satisfying $0 < \beta << \alpha$.

The standard theory of persistence of normally hyperbolic invariant manifolds[Fen72, Fen74, HPS77, Pes04] implies that, for $|\varepsilon| < \varepsilon_0$ there is a locally invariant normally hyperbolic manifold $\tilde{\Lambda}_\varepsilon$ verifying (14) and (15) for the perturbed flow $\tilde{\Phi}_{\varepsilon,t}$ (with $\alpha_\varepsilon, \beta_\varepsilon, C_\varepsilon$ close to α, β, C respectively). The theory in [Fen72, Fen74, HPS77, Pes04] guarantees that $\tilde{\Lambda}_\varepsilon$ is a somewhat smooth family of manifolds but the degree of smoothness can be limited by ratios of normal and tangential exponents α and β. In our case, since the motion on the manifold for $\varepsilon = 0$ is integrable and therefore the Liapunov exponents are zero, we can guarantee that for $|\varepsilon|$ small enough, the family $\tilde{\Lambda}_\varepsilon$ will be a C^{r-1} family if $\tilde{\Phi}_{\varepsilon,t}$ is a C^r family.

Moreover, as it was shown in [DLS08, Theorem 24] there is a naturally defined symplectic parametrization k_ε, such that the perturbed manifold can be written as $\tilde{\Lambda}_\varepsilon = k_\varepsilon(\tilde{\Lambda}_0)$, using as the reference manifold the unperturbed manifold $\tilde{\Lambda}_0$, and choosing $k_0 = \mathrm{Id}$.

Using this symplectic parameterization, one can show that the reduced flow $\tilde{\phi}_{\varepsilon,t}$ on $\tilde{\Lambda}_0$, characterized by $k_\varepsilon \circ \tilde{\phi}_{\varepsilon,t} = \tilde{\Phi}_{\varepsilon,t} \circ k_\varepsilon$ is a Hamiltonian flow. The following proposition makes explicit its Hamiltonian.

Proposition 6. *The reduced flow $\tilde{\phi}_{\varepsilon,t}$ on $\tilde{\Lambda}_0$ defined through $k_\varepsilon \circ \tilde{\phi}_{\varepsilon,t} = \tilde{\Phi}_{\varepsilon,t} \circ k_\varepsilon$ is generated by a C^{r-1} time dependent Hamiltonian vector field with Hamiltonian of the form*

$$(16) \qquad K_\varepsilon(I, \varphi, s) = h(I) + \sum_{i=1}^{N} \varepsilon^i K^i(I, \varphi, s) + O_{C^{r-N-2}}(\varepsilon^{N+1}),$$

where each of the terms K^i is a trigonometric polynomial in the φ, s variables.

Moreover, K^i is an algebraic expression in terms of $\nabla^\ell Q(I, \varphi, p = 0, q = 0, s; \varepsilon = 0)$, for $\ell = 0, \ldots, i-1$. In particular, $K^1(I, \varphi, s) = Q(I, \varphi, 0, 0, s; 0)$.

3.2. **Analyzing the flow restricted to the invariant manifold.** The goal of next sections is to study the objects in $\tilde{\Lambda}_\varepsilon$ invariant by the flow $\tilde{\Phi}_{\varepsilon,t}$. Using Proposition 6, this is equivalent to studying the objects in $\tilde{\Lambda}_0$ invariant under the Hamiltonian flow $\tilde{\phi}_{\varepsilon,t}$ of Hamiltonian $K_\varepsilon(I, \varphi, s)$ given in (16).

The main tool used to obtain invariant objects in $\tilde{\Lambda}_\varepsilon$ will be averaging theory [DLS06a] and KAM theorem applied to the Hamiltonian $K_\varepsilon(I, \varphi, s)$. We see that, after we add some extra variable I_0 conjugated to the variable $s \in \mathbb{T}$, to make it symplectic and autonomous,

we are lead to considering a Hamiltonian of the form:

$$(17) \qquad \tilde{K}_\varepsilon = I_0 + K_\varepsilon = \tilde{h}(\tilde{I}) + \sum_{i=1}^{N} \varepsilon^i K^i(I, \tilde{\varphi}) + O_{C^{r-2-N}}(\varepsilon^{N+1})$$

were we have introduced the notation $\tilde{\varphi} = (\varphi, s)$, $\tilde{I} = (I, I_0)$, and $\tilde{h}(\tilde{I}) = I_0 + h(I)$. We recall that Proposition 6 tells us that, by choosing $|\varepsilon|$ small enough and assuming the regularity r of the Hamiltonian $H(p, q, I, \varphi, s; \varepsilon)$ in (1) large enough, we can take N as large as we want and the regularity of the remainder in (17) is as large as we want.

Furthermore, in our case, using the assumption **H4** it follows that the K^i are trigonometric polynomials in the angle variables $\tilde{\varphi}$ with Fourier coefficients that depend on I, but not on I_0:

$$(18) \qquad K^i(I, \tilde{\varphi}) = \sum_{(k,l) \in \mathcal{N}_i} K_{k,l}^i(I) e^{2\pi i(k\varphi + ls)}$$

$\mathcal{N}_i$ being finite sets. Very explicit formulas for the coefficients $K_{k,l}^i(I)$ are given in [DLS06a].

The fact that the perturbation terms do not depend on the I_0 variable is a reflection of the fact that I_0 is just a variable introduced to keep the time rotating at unit speed. This is independent of the perturbations.

3.3. **The averaging method.** In this section we recall the basis of the averaging method for time periodic perturbations. The averaging method is a rather standard tool in Hamiltonian dynamical systems and has an extensive literature. Modern surveys are [LM88, AKN88].

The basic idea of the averaging method is to make symplectic changes of variables carefully chosen so that the resulting Hamiltonian presents a particularly simple form. There are many averaging theories depending on what is the simple form to be achieved and what is the method used to keep track of the simplifying transformations. In this paper, we will follow [DLS06a] and use the method of Lie transforms. The averaged Hamiltonians we will use here are different from those used in [DLS06a] to accommodate the fact that resonances for systems with two or more degrees of freedom are manifolds in the action space whereas for one degree of freedom, resonances are just points. In this paper, we also consider more general unperturbed Hamiltonians $h(I)$—in [DLS06a], the unperturbed Hamiltonian was just quadratic—but, under hypothesis **H3**, this makes little difference.

3.3.1. *Some generalities on the averaging method.* We will follow the method of Lie transforms [Car81, Mey91], considering transformations obtained as the time-1 map of a Hamiltonian $G(I, \tilde{\varphi})$ in $\mathbb{R}^d \times \mathbb{T}^{d+1}$.

Given a Hamiltonian function $G(I, \tilde{\varphi})$ in the extended phase space, we denote by $\exp(G)$ the time-1 map of the Hamiltonian flow generated by G.

The main technical result we will use about the time-1 map is a direct consequence of Taylor's expansions and the regularity of the solutions of an ordinary differential equation as well as the expression of the derivatives of Hamiltonian functions in terms of Poisson brackets [Thi97]. So that the following Lemma 7 is just a Taylor expansion along trajectories.

Lemma 7. *Let $A \subset B \subset \mathbb{R}^{d+1}$ be compact sets and $\ell \geq 2$. There exists a constant $C = C(l, k)$, such that, given $G \in C^\ell(B \times \mathbb{T}^{d+1})$ satisfying*

$$\|G\|_{C^1(B \times \mathbb{T}^{d+1})} < d(A, \mathbb{R}^{d+1} \setminus B)$$

so that the Hamiltonian flow associated to G starting in $A \times \mathbb{T}^{d+1}$ stays in the interior of $B \times \mathbb{T}^{d+1}$, we have:

    a) $\exp G \in C^{\ell-1}(A \times \mathbb{T}^{d+1})$
    b) $\|\exp(G) - \mathrm{Id}\|_{C^{\ell-1}(A \times \mathbb{T}^{d+1})} \leq C\|G\|_{C^\ell(B \times \mathbb{T}^{d+1})}$
    c) *given $H \in C^\ell(B \times \mathbb{T}^{d+1})$ then:*

$$\|H \circ \exp(G) - H - \{H, G\}\|_{C^{\ell-2}(A \times \mathbb{T}^{d+1})}$$
$$\leq C\|H\|_{C^\ell(B \times \mathbb{T}^{d+1})}\|G\|_{C^\ell(B \times \mathbb{T}^{d+1})}^2$$

    where $\{\cdot, \cdot\}$ denotes the Poisson bracket in $\mathbb{R}^{d+1} \times \mathbb{T}^{d+1}$.
    d) *More generally, if $\ell > k + 1$, there is an asymptotic expansion*

$$\|H \circ \exp(G) - H - \{H, G\} - \frac{1}{2}\{\{H, G\}, G\}$$
$$- \ldots - \frac{1}{k!}\{\{\ldots \{\{H, G\}, G\}, G\} \ldots \}\|_{C^{\ell-k-1}(A \times \mathbb{T}^{d+1})}$$
$$\leq C\|H\|_{C^\ell(B \times \mathbb{T}^{d+1})}\|(G\|_{C^\ell(B \times \mathbb{T}^{d+1})})^{k+1}$$

As a consequence of Lemma 7, we obtain the following algorithm, which is the main formal step of the general averaging method and which allows computations to high order.

Algorithm 8. Given a sufficiently smooth Hamiltonian averaged up to order $N - 1 \geq 0$

(19)
$$\tilde{K}_\varepsilon^{N-1} = \tilde{h} + \sum_{i=1}^{N-1} \varepsilon^i \bar{K}^i + \varepsilon^N K^N + O(\varepsilon^{N+1})$$

Assume that we can find sufficiently smooth $\bar{K}^N$, G^N solving:

(20)
$$\bar{K}^N = K^N + \{\tilde{h}, G^N\}$$

Then

$$\tilde{K}_\varepsilon^N = K_\varepsilon^{N-1} \circ \exp(\varepsilon^N G^N) = \tilde{h} + \sum_{i=1}^{N-1} \varepsilon^i \bar{K}^i + \varepsilon^N \bar{K}^N + O(\varepsilon^{N+1})$$

$$= \tilde{h} + \sum_{i=1}^{N} \varepsilon^i \bar{K}^i + \varepsilon^{N+1} K^{N+1} + O(\varepsilon^{N+2})$$

In general there are many choices for $\bar{K}^N$ and G^N. In the following subsections, we will specify the choices that we make for our case and establish estimates for the transformation G^N and the new Hamiltonian $\tilde{K}_\varepsilon^N$. In particular, we will have estimates for the averaged Hamiltonian $\bar{K}^N$.

3.3.2. *One step of averaging: The infinitesimal equations. Resonances in one averaging step.* Our Hamiltonian $\tilde{K}_\varepsilon$ given in (17) is of the form

$$\tilde{K}_\varepsilon = I_0 + K_\varepsilon$$

where K_ε is given in (16) and only depends on $(I, \tilde{\varphi})$. Then, in our case, we will take $\tilde{K}_\varepsilon^N$ in (19), with $\tilde{h} = I_0 + h(I)$ and $\tilde{K}_\varepsilon^N - \tilde{h}$ only depends on $(I, \tilde{\varphi})$. So that the function G will depend only on $(I, \tilde{\varphi})$.

The fact that $\tilde{h}$ in (19) is given by $\tilde{h} = I_0 + h(I)$, will allow us to treat the averaging equation (20) using Fourier series. We find it convenient to divide the phase space into different regions and perform different averaging procedures in each region.

At every step of the iteration algorithm 8, given $K(I, \tilde{\varphi})$, we have to solve equation (20) for the unknowns $\bar{K}(I, \tilde{\varphi}), G(I, \tilde{\varphi})$:

(21)
$$\bar{K} = K + \{\tilde{h}, G\}$$

with $\tilde{h}(\tilde{I}) = I_0 + h(I)$. Writing

$$K = \sum_{\tilde{k} \in \mathcal{N} \subset \mathbb{Z}^{d+1}} K_{\tilde{k}}(I) e^{2\pi i \tilde{k} \cdot \tilde{\varphi}} = \sum_{(k,l) \in \mathcal{N} \subset \mathbb{Z}^{d+1}} K_{k,l}(I) e^{2\pi i (k \cdot \varphi + ls)},$$

it is clear, because $\tilde{h}$ depends only on $\tilde{I}$, that the Poisson bracket with $\tilde{h}$ is diagonal in Fourier series. Hence, it is natural to search for $\bar{K}, G$

such that their Fourier series are supported in $\mathcal{N}$, the support of the Fourier series of K. Hence we look for:

$$G = \sum_{(k,l) \in \mathcal{N} \subset \mathbb{Z}^{d+1}} G_{k,l}(I) e^{2\pi i (k \cdot \varphi + l s)}$$

$$\bar{K} = \sum_{(k,l) \in \mathcal{N} \subset \mathbb{Z}^{d+1}} \bar{K}_{k,l}(I) e^{2\pi i (k \cdot \varphi + l s)}.$$

Using that

$$\{\tilde{h}, G\} = -2\pi i \sum_{(k,l) \in \mathcal{N} \subset \mathbb{Z}^{d+1}} (\omega(I) \cdot k + l) \, G_{k,l}(I) e^{2\pi i (k \cdot \varphi + l s)}$$

where $\omega(I) = \frac{\partial h}{\partial I}(I) = \nabla h(I)$, equation (21) becomes a set of equations for the Fourier coefficients:

(22) $$\bar{K}_{k,l}(I) - K_{k,l}(I) = -2\pi i (\omega(I) \cdot k + l) G_{k,l}(I)$$

The solution of (22) is obtained by choosing $\bar{K}_{k,l}$ and then, setting

(23) $$G_{k,l}(I) = \frac{K_{k,l}(I) - \bar{K}_{k,l}(I)}{2\pi i (\omega(I) \cdot k + l)}.$$

It is clear that the solution (23) requires special treatment when

(24) $$\omega(I) \cdot k + l = 0, \ (k, l) \in \mathcal{N}.$$

This motivates next definition.

Definition 9. *Given a Hamiltonian $h(I)$ we define a resonance as the set*

(25) $$\mathcal{R}_{k,l} = \{I \in \mathcal{I}, \ \omega(I) \cdot k + l = 0\} = \omega^{-1} \left(\{\Omega \in \mathbb{R}^d, \ \Omega \cdot k + l = 0\} \right).$$

where $\omega(I) = \frac{\partial h}{\partial I}(I) = \nabla h(I)$.

Let us observe that, by Hypothesis **H3**, for $I \in \mathcal{I}^*$, the resonances are smooth surfaces for $k \neq 0$, as smooth as the map $\omega = \frac{\partial h}{\partial I}$.

Moreover, $\mathcal{R}_{mk,ml} = \mathcal{R}_{k,l}$ for any $m \in \mathbb{Z}$, therefore any two of these sets $\mathcal{R}_{k,l}$ and $\mathcal{R}_{\bar{k}, \bar{l}}$ either:

- Are identical, if and only if $(k, l) = (m\bar{k}, m\bar{l})$ for some $m \in \mathbb{Z}$
- Do not intersect
- Intersect transversally in a manifold of codimension two without boundary

More generally, if we consider different resonances, the intersection

$$\mathcal{R}_{k_1,l_1} \cap \cdots \cap \mathcal{R}_{k_j,l_j} = \omega^{-1} \left(\{\Omega \in \mathbb{R}^d, \ k_i \cdot \Omega + l_i = 0, i = 1, \ldots, j\} \right)$$

will be a manifold of dimension m, where m is the dimension of the $\mathbb{Z}$-module $\mathcal{M}[(k_1, l_1), \ldots, (k_j, l_j)]$ generated by $(k_1, l_1), \ldots, (k_j, l_j)$.

Given $\mathcal{N} \subset \mathbb{Z}^{d+1}$, the support of the Fourier series of K, we introduce the notation

$$(26) \quad \tilde{\mathcal{N}} = \{(k,l) \in \mathcal{N} \mid \nexists\, m \neq 1 \in \mathbb{Z}, (\tilde{k}, \tilde{l}) \in \mathcal{N}, (k,l) = m(\tilde{k}, \tilde{l})\}$$

Notice that, with this notation if $\mathcal{R}_{k_1, l_1} = \mathcal{R}_{k_2, l_2}$, with $(k_i, l_i) \in \tilde{\mathcal{N}}$, $i = 1,2$ one has that $(k_1, l_1) = (k_2, l_2)$ and then

$$\cup_{(k,l) \in \mathcal{N}} \mathcal{R}_{k,l} = \cup_{(k,l) \in \tilde{\mathcal{N}}} \mathcal{R}_{k,l}.$$

Since resonances are sets of codimension 1, it is natural to give in them a system of $d-1$ coordinates. In the next lemma we will consider the function

$$\Gamma_{k,l} : \mathcal{I}^* \to \mathcal{R}_{k,l}$$

which is a projection onto the resonance $\mathcal{R}_{k,l}$ along a transversal bundle to it. The standard choice is the *orthogonal* projection, in such a way that the projection $\Gamma_{k,l}(I)$ is the closest point to I in $\mathcal{R}_{k,l}$. The orthogonal projection is well defined in a tubular neighborhood with respect to the normal bundle of $\mathcal{R}_{k,l}$.

There is another simpler choice, which appears naturally when one realizes that the dynamics close to resonances is generically "pendulum-like" (see equation (51)). It is just projecting along the bundle $\mathcal{R}_{k,l} + \langle k \rangle$, that is, defining $\Gamma_{k,l}(I)$ as the intersection of the straight line $\{I + tk, t \in \mathbb{R}\}$ with the resonance $\mathcal{R}_{k,l}$. This projection will be not well defined close to points $I \in \mathcal{R}_{k,l} \subset \mathcal{I}^*$ such that the direction given by the vector k is tangent to $\mathcal{R}_{k,l}$, that is, close to points I satisfying:

$$(27) \qquad \begin{aligned} Dh(I) \cdot k + l &= 0 \\ k^\top D^2 h(I) k &= 0 \end{aligned}$$

Nevertheless, under hypothesis **H5**, for *secular resonances*, that is, for $\mathcal{R}_{k,l}$ such that $(k,l) \in \mathcal{N}^{[\leq 2]}$, these points are a codimension two set (as one would expect from naive parameter counting because (27) are two conditions). In the averaging procedure used in this paper, we just need to deal carefully with secular resonances to obtain a "pendulum like" Hamiltonian near them. So, our strategy will be to use this projection near the secular resonances and the standard orthogonal one near the rest of resonances we encounter in the higher averaging steps.

We define $\mathcal{J}_\delta \subset \mathcal{I}^*$ as the set of points of $\mathcal{I}^*$ taking out a neighborhood of size δ of the points verifying (27) for some $(k,l) \in \mathcal{N}^{[\leq 2]}$. In the resulting set $\mathcal{J}_\delta$ we will have a well defined projection in the k direction for any secular resonance:

$$(28) \text{ If } I^* \in \mathcal{J}_\delta \cap \mathcal{R}_{k_0, l_0} \text{ then } k_0^\top D^2 h(I^*) k_0 \neq 0 \text{ for } (k_0, l_0) \in \mathcal{N}^{[\leq 2]}$$

Of course, if our Hamiltonian h is quasiconvex, since the set satisfying (27) is empty, we obtain that $\mathcal{J}_\delta = \mathcal{I}^*$.

For future reference, let us write the characterization of the projection $\Gamma_{k,l}(I)$ along the k-direction:

$$(29) \qquad I^* = \Gamma_{k,l}(I) \iff I - I^* \in \langle k \rangle \text{ and } \frac{\partial h}{\partial I}(I^*) \cdot k + l = 0.$$

Let us observe that for $I \in \mathcal{J}_\delta$ in a neighborhood of a secular resonance $\mathcal{R}_{k,l}$, $(k,l) \in \mathcal{N}^{[\leq 2]}$, there exists a constant $C \geq 1$ such that

$$(30) \qquad \text{dist}(I, \mathcal{R}_{k,l}) \leq \text{dist}(I, \Gamma_{k,l}(I)) \leq C \, \text{dist}(I, \mathcal{R}_{k,l}).$$

Indeed, C can be chosen as any constant satisfying

$$C > \frac{\|k\| \, \|D^2 h(I)k\|}{\|k^\top D^2 h(I)k\|}, \quad \forall I \in \mathcal{J}_\delta, \; (l,k) \in \mathcal{N}^{[\leq 2]}$$

In the case of the orthogonal projection the value of the constant C (30) is 1. We emphasize that, in the averaging procedure, the resonant sets are determined by the integrable Hamiltonian $h(I)$ and not by the perturbation. Of course, given a concrete system, many resonances do not play any role and only the resonances excited by the perturbation play a role, as we will see in next lemma.

As we indicate before, the solution of the homological equation involves choices of which terms are eliminated and which terms are kept in the averaged Hamiltonian. The following lemma indicates the choice we will follow. We will use in it a general projection $\Gamma_{k,l}$, just assuming that it verifies (30).

Lemma 10. *Let*

$$K(I, \varphi, s) = \sum_{(k,l) \in \mathcal{N}} K_{k,l}(I) e^{2\pi i (k\varphi + ls)}$$

be a Hamiltonian, with $\mathcal{N} = \mathcal{N}(K) \subset \mathbb{Z}^{d+1}$ a finite set. Assume that K is of class $\mathcal{C}^\ell$ with respect to $I \in \mathcal{J}_\delta \subset \mathcal{I}^* \subset \mathbb{R}^d$ and consider the resonant set $\mathcal{R}_\mathcal{N} = \{I \in \mathcal{J}_\delta, \; \omega(I) \cdot k + l = 0, \; (k,l) \in \mathcal{N}\} = \cup_{(k,l) \in \mathcal{N}} \mathcal{R}_{k,l} \subset \mathcal{J}_\delta$.

Choose $0 < L < 1$ small enough such that for any $(k,l), (\bar{k}, \bar{l}) \in \mathcal{N}$, either $\mathcal{R}_{k,l} = \mathcal{R}_{\bar{k}, \bar{l}}$ or the tubular neighborhood of $\mathcal{R}_{k,l}$ of radius L does not contain $\mathcal{R}_{\bar{k}, \bar{l}}$.

Assume that we have a projection $\Gamma_{k,l} : \mathcal{J}_\delta \to \mathcal{R}_{k,l}$ such that it verifies (30).

Then, there exist $G(I, \varphi, s)$ of class $\mathcal{C}^{\ell-1}$ with respect to I, and $\bar{K}$ of class $\mathcal{C}^\ell$ with respect to I, with $\mathcal{N}(G)$ and $\mathcal{N}(\bar{K})$ finite sets. Moreover $\mathcal{N}(G), \mathcal{N}(\bar{K}) \subset \mathcal{N}(K)$.

The functions $G(I, \varphi, s)$ and $\bar{K}$ solve the homological equation (21) in $\mathcal{J}_\delta$ and satisfy:

a) If $d(I, \mathcal{R}_\mathcal{N}) \geq 2L$, then
$$\bar{K}(I, \varphi, s) = K_{0,0}(I).$$

b) If $d(I, \mathcal{R}_{k_i, l_i}) \leq L$ for $i = 1, \ldots j$, then

$$
\begin{aligned}
\bar{K}(I, \varphi, s) &= K_{0,0}(I) + \sum_{i=1}^{j} \left(\sum_{\nu=-N_i}^{N_i} K_{\nu k_i, \nu l_i}(\Gamma_{k_i, l_i}(I)) e^{2\pi i \nu (k_i \varphi + l_i s)} \right) \\
&= K_{0,0}(I) + U_{k_1, l_1, \ldots, k_j, l_j}(I, k_1 \varphi + l_1 s, \ldots, k_j \varphi + l_j s),
\end{aligned}
$$

where $0 < N_i < \infty$ are such that if $(\nu k_i, \nu l_i) \in \mathcal{N}$, then $|\nu| \leq N_i$.

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

d) The function G verifies $\|G\|_{\mathcal{C}^{\ell-1}} \leq \dfrac{C}{L^{\ell+1}} \|K\|_{\mathcal{C}^\ell}$.

Remark 11. We observe that in the formula of $\bar{K}$ the angles can be redundant because some of the angles included in the sum can be combination of others. One can be more precise by considering the module generated by $(k_1, l_1), \ldots, (k_j, l_j)$ and the dimension of this module gives us the number of independent angles among $k_1 \varphi + l_1 s, \ldots, k_j \varphi + l_j s$. However, in this paper, this will not be needed. Our strategy later will be to reduce the domain in such a way that we will not need to deal with multiple resonances.

Proof. If we write the homological equation (21) in Fourier coefficients, we obtain equation (23). Our first choice is $\bar{K}_{k,l}(I) = G_{k,l}(I) = 0$, if $(k, l) \notin \mathcal{N}$. For $(k, l) \in \mathcal{N}$, we solve equation (23) choosing:

(1) If $(0, 0) \in \mathcal{N}$ we take $\bar{K}_{0,0}(I) = K_{0,0}(I)$.
(2) If $(0, l) \in \mathcal{N}$, $l \neq 0$, $\bar{K}_{0,l}(I) = 0$.
(3) If $(k, l) \in \mathcal{N}$, $k \neq 0$, we choose $\bar{K}_{k,l}(I)$ as:

$$\bar{K}_{k,l}(I) = K_{k,l}(\Gamma_{k,l}(I)) \psi \left(\frac{1}{L} (d(I, \mathcal{R}_{k,l})) \right)$$

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 $d(I, \mathcal{R}_{k,l}) \leq L$ then $\bar{K}_{k,l}(I) = K_{k,l}(\Gamma_{k,l}(I))$,
(b) if $d(I, \mathcal{R}_{k,l}) \geq 2L$ then $\bar{K}_{k,l}(I) = 0$.

Once we have defined $\bar{K} = \sum_{(k,l) \in \mathcal{N}} \bar{K}_{k,l} e^{i(k\varphi + ls)}$, it is clear that it has the form announced in a) and b), is a $\mathcal{C}^\ell$ function with respect to I, and that it verifies the bounds c), where the constant C only depends on

21

the cut-off $\mathcal{C}^\infty$ function ψ, the functions $\Gamma_{k,l}$, the degree of the Fourier polynomials and ℓ.

Now, we choose G that satisfies equation (22):

(1) $G_{0,0}(I) = 0$,

(2) For $(0,l) \in \mathcal{N}$, $l \neq 0$, $G_{0,l}(I) = \dfrac{K_{0,l}(I)}{2\pi i l}$,

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

    (a) If $\omega(I) \cdot k + l \neq 0$ then $G_{k,l}(I) = \dfrac{K_{k,l}(I) - \bar{K}_{k,l}(I)}{2\pi i (\omega(I) \cdot k + l)}$,

    (b) If $\omega(I) \cdot k + l = 0$ and we are using the standard orthogonal projection, then $G_{k,l}(I) = \dfrac{\nabla K_{k,l}(\Gamma_{k,l}(I)) \cdot D^2 h(I) k}{2\pi i \|D^2 h(I) k\|^2}$.

    (c) If $\omega(I) \cdot k + l = 0$ and we are using the k-projection, then
$$G_{k,l}(I) = \frac{D K_{k,l}(I) k}{2\pi i \, \|k^\top D^2 h(I) k\|}.$$

To bound the function G we first bound its Fourier coefficients $G_{k,l}(I)$:

(1) For $(0,l) \in \mathcal{N}$, $l \neq 0$, $\|G_{0,l}\|_{\mathcal{C}^{\ell-1}} \leq C \|K_{0,l}\|_{\mathcal{C}^{\ell-1}}$.

(2) Given $(k_0, l_0) \in \mathcal{N}$, $k_0 \neq 0$, by the definition of $\bar{K}$ and G, we have:

    (a) On $\{I \in \mathcal{I}, \ d(I, \mathcal{R}_{k_0,l_0}) \leq L\}$, we have $\|G_{k_0,l_0}\|_{\mathcal{C}^{\ell-1}} \leq C \dfrac{\|K_{k_0,l_0}\|_{\mathcal{C}^\ell}}{|k_0|}$.

    (b) On $\{I \in \mathcal{I}, \ d(I, \mathcal{R}_{k_0,l_0}) \geq 2L\}$, we have $\|G_{k_0,l_0}\|_{\mathcal{C}^{\ell-1}} \leq C \dfrac{\|K_{k_0,l_0}\|_{\mathcal{C}^\ell}}{(|k_0| L)^\ell}$.

    (c) On $\{I \in \mathcal{I}, \ L \leq d(I, \mathcal{R}_{k_0,l_0}) \leq 2L\}$, we have

$$\|G_{k_0,l_0}\|_{\mathcal{C}^{\ell-1}} \leq C \frac{\|K_{k_0,l_0}\|_{\mathcal{C}^{\ell-1}}}{(|k_0| L)^{\ell-1}}.$$

Therefore, $G(I, \varphi, s)$ is a trigonometric polynomial in (φ, s), and of class $\mathcal{C}^{\ell-1}$ with respect to I and satisfies the bounds in d). $\qquad \square$

Remark 12. We note that the above estimates use the fact that the function K is a trigonometric polynomial in (φ, s) and the constant C in the bounds of G depend on the degree of the polynomial.

One can follow the same procedure when the function is not a polynomial by estimating the Fourier coefficients using Cauchy bounds:

$$\|K_{k,l}\| \leq C \|K\|_{\mathcal{C}^r} (|k| + |l|)^{-r}$$

and then using the estimates we presented. The only difference is that for general functions we obtain estimates:

$$\|G\|_{\mathcal{C}^{\ell-d-1}} \leq \frac{C}{L^{\ell+1}} \|K\|_{\mathcal{C}^\ell}.$$

where C depends only on the dimension of the space.

As we will use Lemma 10 in the averaging algorithm 8, in the next definition we want to emphasize that resonances only play a role when they are present in the Fourier transform of the Hamiltonian at the step N, that is, when the numerator in the the equation (23) for $G_{k,l}(I)$ is not zero.

Note that the denominator in the expression (23) depends on the unperturbed system, but the numerator depends on the perturbation. The places where the denominator vanishes are the resonances. Clearly the resonances do not matter unless the numerator is not zero.

Definition 13. *Given a resonance $\mathcal{R}_{k,l}$ as defined in definition 9, we say that it is activated at order N if N is the smallest value such that $(k,l) \in \mathcal{N}_N$, were $\mathcal{N}_N$ is the support of the Fourier transform of the term of order ε^N, after applying $N-1$ steps of the averaging algorithm 8. That is, $\mathcal{N}_N$ is the support of the Fourier transform of K^N in (19).*

We denote the set of resonances activated at order N by

$$(31) \qquad\qquad \mathcal{R}^N = \cup_{(k,l)\in\mathcal{N}_N} \mathcal{R}_{k,l}$$

and we introduce the resonances open up to order N as the set

$$\mathcal{R}^{[\leq N]} = \cup_{i=1}^N \mathcal{R}^i = \cup_{(k,l)\in\mathcal{N}^{[\leq N]}} \mathcal{R}_{k,l}$$

where

$$\mathcal{N}^{[\leq N]} = \cup_{i=1}^N \mathcal{N}_i$$

The resonances open up to order 2, $\mathcal{R}^{[\leq 2]}$, are called secular resonances.

We note that if a resonance has been activated at order q, we have that, in a neighborhood of that resonance, the system is reduced to integrable up to order ε^q. If the term of order ε^q in the averaged system does not vanish, averaging to higher order does not change the leading order term anymore. We will only need the cases $q = 1, 2$.

Definition 14. *Given a frequency ω and an order of averaging N, we define its active resonances up to order N:*

$$(32) \qquad\qquad \mathcal{A}(\omega, N) = \{(k,l) \in \mathcal{N}^{[\leq N]}, \omega \cdot k + l = 0\}$$

and $\mathbf{m}(\omega, N)$, *the multiplicity of ω up to order N, as the dimension of the $\mathbb{Z}$-module generated by $\mathcal{A}(\omega, N)$:*

(33)
$$\mathbf{m}(\omega, N) = \dim(\mathcal{M}[\mathcal{A}(\omega, N)]).$$

Of course $\mathbf{m}(\omega, N+1) \geq \mathbf{m}(\omega, N)$ and inequality can be strict.

The relevance of the concept of multiplicity comes because, as we emphasize in Remark 11, the main result of the averaging method is that in a neighborhood of a point $I \in \mathcal{J}_\delta$ such that the frequency $\omega(I)$ is of multiplicity $\mathbf{m}(\omega, N)$ up to order N, there exists a change of variables that reduces the Hamiltonian K in Lemma 10 to a function of I and $\mathbf{m}(\omega, N)$ angles up to an error of order $O(\varepsilon^{N+1})$.

Using the method of Lemma 10 in the Algorithm 8 we obtain straightforwardly Lemma 15, that is the main iterative step in the averaging procedure.

The hypothesis of next Lemma 15 are that we have a Hamiltonian averaged up to order q. The conclusions are that we can produce another Hamiltonian which is averaged up to a higher order $q+1$ in ε.

Lemma 15. *Consider a Hamiltonian of the form:*

(34)
$$K_q(I, \varphi, s; \varepsilon) = K_q^0(I, \varphi, s; \varepsilon) + \varepsilon^{q+1} K_q^1(I, \varphi, s; \varepsilon).$$

Assume that K_q is of class $\mathcal{C}^\ell$ with respect to $I \in \mathcal{J}_\delta \subset \mathcal{I}^* \subset \mathbb{R}^d$.

Consider the finite collection of sets $\mathcal{R}^s \subset \mathcal{J}_\delta$, called resonances activated at order s, $s = 1, \ldots, q$, and $\mathcal{R}^{[\leq q]} = \cup_{1 \leq s \leq q} \mathcal{R}^s$ the set of resonances open up to order q.

Consider a number $L < 1$ small enough such that:

L1 *for any $\mathcal{R}_{k,l}, \mathcal{R}_{\tilde{k},\tilde{l}} \in \mathcal{R}^{[\leq q]}$, either $\mathcal{R}_{k,l} = \mathcal{R}_{\tilde{k},\tilde{l}}$ or the tubular neighborhood of $\mathcal{R}_{k,l}$ of radius L does not contain $\mathcal{R}_{\tilde{k},\tilde{l}}$.*

And assume that:

- $K_q^0(I, \varphi, s; \varepsilon)$ *satisfies:*
  - *If $q = 0$, $K_0^0(I, \varphi, s; \varepsilon) = I_0 + h(I)$.*
  - *If $q \geq 1$, $K_q^0(I, \varphi, s; \varepsilon)$ is a $\mathcal{C}^{n+2-2q}$ function that verifies:*
  1.1. *If $d(I, \mathcal{R}^{[\leq q]}) \geq 2L$, then*

    $$K_q^0(I, \varphi, s; \varepsilon) = I_0 + h(I) + \varepsilon K_q^{0,0}(I; \varepsilon),$$

    where $\varepsilon K_q^{0,0}(I; \varepsilon)$ is a polynomial of degree q in ε.
  1.2. *If $d(I, \mathcal{R}^{[\leq q]}) \leq L$, then we can find at least one (may be more) $0 \leq j \leq q$ such that $d(I, \mathcal{R}^j) \leq L$, and therefore at least one k_i^j, l_i^j such that*

    $$\mathcal{R}_{k_i^j, l_i^j} \in \mathcal{R}^s, \quad i = 1, \ldots, n_j,$$

    and that $d(I, \mathcal{R}_{k_i^j, l_i^j}) \leq L$.

Then,

(35)
$$\bar{K}_q^0(I, \varphi, s; \varepsilon) = h(I) + \varepsilon \bar{K}^{0,0}(I; \varepsilon)$$
$$+ \sum_{0 \le j \le q} \varepsilon^j U^{j,q}(I, k_1^j \varphi + l_1^j s, \dots, k_{n_j}^j \varphi + l_{n_j}^j s; \varepsilon)$$

where the functions $U^{j,q}(I, \theta_1^j, \dots, \theta_{n_j}^j; \varepsilon)$ are polynomials in ε and trigonometric polynomials in the angle variables θ_i^j, $i = 1, \dots, n_j$, with support of the Fourier transform with respect to the (φ, s) contained in $\mathcal{N}_1 \cup \cdots \cup \mathcal{N}_q$.
Moreover, for $j = 1$ the function $U^{1,q}(I, \theta_1^1, \dots, \theta_{n_1}^1; \varepsilon)$ is given by:

$$U^{1,q} = \sum_{i=1}^{n_1} \left(\sum_{p=-N_i}^{N_i} K^1_{pk_i^1, pl_i^1}(\Gamma_{k_i^1, l_i^1}(I)) e^{2\pi i p(k_i^1 \varphi + l_i^1 s)} \right)$$
$$+ \mathrm{O}(\varepsilon)$$

where $K_{k,l}^1(I)$ are the Fourier coefficients of $K_0^1(I, \varphi, s; 0)$ with respect to the angle variables (φ, s).

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

Denote by

$$K = K_q^1(I, \varphi, s; 0) = \sum_{(k,l) \in \mathcal{N}_{q+1}} K_q^{k,l}(I) e^{2\pi i(k\varphi + ls)},$$

where $\mathcal{N}_{q+1}$ is assumed to be a finite set. K is the term of the perturbation of order exactly $q + 1$. Introduce also the set of resonances activated at order $q + 1$:

(36)
$$\mathcal{R}^{q+1} = \cup_{(k,l) \in \mathcal{N}^{q+1}} \mathcal{R}_{k,l} \setminus \mathcal{R}^{[\le q]}$$

Choose $0 < \tilde{L} < L$ such that $L)$ holds for $\mathcal{R}^{[\le q+1]}$.

  Let $G(I, \varphi, s)$ be the $\mathcal{C}^{n-2q-1}$ function whose Fourier coefficients $G_{k,l}$ verify equation (22), for $(k, l) \in \mathcal{N}_{q+1}$, with $K = K_q^1(I, \varphi, s; 0)$.

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

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

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

(37)
$$K_{q+1}^0(\mathcal{B}, \alpha, s; \varepsilon) = K_q^0(\mathcal{B}, \alpha, s; \varepsilon) + \varepsilon^{q+1} \bar{K}_q^1(\mathcal{B}, \alpha, s; 0),$$

where $\bar{K}_q^1(\mathcal{B}, \alpha, s; 0) = \bar{K}(\mathcal{B}, \alpha, s)$, given in Lemma 10, is a $\mathcal{C}^{n-2q}$ function whose Fourier coefficients solve equations (22).

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, $\varepsilon^{q+2} K_{q+1}^1(B, \alpha, s; \varepsilon)$ is a $\mathcal{C}^{n-2q-2}$ function whose Taylor series coefficients with respect to ε are trigonometric polynomials in (α, s).

Once we know how to solve any homological equation (22), 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 16, which is a straightforward generalization of Theorem 8.9 in [DLS06a].

Theorem 16. *Let* $\tilde{K}(\tilde{I}, \varphi, s; \varepsilon)$ *be a* $\mathcal{C}^n$ *Hamiltonian,* $n > 1$, *for* $I \in \mathcal{J}_\delta \subset \mathcal{I}^* \subset \mathbb{R}^d$ *and consider any* $1 \le m < n$, *independent of* ε. *Assume that*

$$(38) \qquad \tilde{K}(\tilde{I}, \varphi, s; \varepsilon) = I_0 + h(I) + \varepsilon K(I, \varphi, s; \varepsilon).$$

Let $K_i(I, \varphi, s)$ $i = 1, , \ldots, m$ *be the coefficients in the Taylor expansion with respect to* ε *of* $K(I, \varphi, s; \varepsilon)$, *and assume that the* $K_i(I, \varphi, s)$, $i = 1, \ldots, m$ *are trigonometric polynomials in* φ, s.

Consider the finite collection of sets $\mathcal{R}^i \subset \mathcal{I}^*$, *called resonances activated at order* i, $i = 1 \ldots m$, *following definition 13, as well as the resonances open up to order* m: $\mathcal{R}^{[\le m]} = \cup_{1=1,\ldots,m} \mathcal{R}^i$.

Consider a number $0 < L < 1$ *small enough such that:*

   L1 *for any* $\mathcal{R}_{k,l}, \mathcal{R}_{\tilde{k},\tilde{l}} \in \mathcal{R}^{[\le m]}$, *either* $\mathcal{R}_{k,l} = \mathcal{R}_{\tilde{k},\tilde{l}}$ *or the tubular neighborhood of* $\mathcal{R}_{k,l}$ *of radius* L *does not contain* $\mathcal{R}_{\tilde{k},\tilde{l}}$.

Then, there exists a symplectic change of variables, depending on time, $(I, \varphi, s) \mapsto (\mathcal{B}, \alpha, s)$, *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* $\tilde{K}(\tilde{I}, \varphi, s; \varepsilon)$ *into a Hamiltonian system of Hamiltonian*

$$B_0 + \bar{K}(\mathcal{B}, \alpha, s; \varepsilon) = B_0 + \bar{K}^0(\mathcal{B}, \alpha, s; \varepsilon) + \varepsilon^{m+1} \bar{K}^1(\mathcal{B}, \alpha, s; \varepsilon)$$

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* $d(\mathcal{B}, \mathcal{R}^{[\le m]}) \ge 2L$, *then*

$$\bar{K}^0(\mathcal{B}, \alpha, s; \varepsilon) = h(\mathcal{B}) + \varepsilon \bar{K}^{0,0}(\mathcal{B}; \varepsilon)$$

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

(2) *If $d(\mathcal{B}, \mathcal{R}^{[\leq m]}) \leq L$, then we can find at least one (may be more) $0 \leq j \leq m$ such that $d(\mathcal{B}, \mathcal{R}^j) \leq L$, and therefore at least one k_i^j, l_i^j such that*

$$\mathcal{R}_{k_i^j, l_i^j} \in \mathcal{R}^j, \quad i = 1, \ldots n_j,$$

and that $d(\mathcal{B}, \mathcal{R}_{k_i^j, l_i^j}) \leq L$.
   Then,

(39)
$$
\begin{aligned}
\bar{K}^0(\mathcal{B}, \alpha, s; \varepsilon) &= h(\mathcal{B}) + \varepsilon \bar{K}^{0,0}(\mathcal{B}; \varepsilon) \\
&+ \sum_{0 \leq j \leq m} \varepsilon^j U^{j,m}(\mathcal{B}, k_1^j \alpha + l_1^j s, \ldots, k_{n_j}^j \alpha + l_{n_j}^j s; \varepsilon)
\end{aligned}
$$

where the functions $U^{j,m}(\mathcal{B}, \theta_1^j, \ldots, \theta_{n_j}^j; \varepsilon)$ are polynomials in ε and trigonometric polynomials in the angle variables θ_i^j, $i = 1, \ldots, n_j$, with support of the Fourier transform with respect to the (φ, s) contained in $\mathcal{N}_1 \cup \cdots \cup \mathcal{N}_m$.
   Moreover, if $j = 1$, the function $U^{1,m}(I, \theta_1^1, \ldots, \theta_{n_1}^1; \varepsilon)$ is given by:

$$
U^{1,m} = \sum_{i=1}^{n_1} \left(\sum_{p=-N_i}^{N_i} K^1_{pk_i^1, pl_i^1}(\Gamma_{k_i^1, l_i^1}(\mathcal{B})) e^{2\pi i p(k_i^1 \alpha + l_i^1 s)} \right)
$$
$$
+ \mathrm{O}(\varepsilon)
$$

where $K^1_{k,l}(\mathcal{B})$ are the Fourier coefficients of the $K(\mathcal{B}, \alpha, s; 0)$ with respect to the angle variables (α, s).

Note that in this Theorem 16 we have not claimed anything in the regions at a distance between L and $2L$ of the resonance set $\mathcal{R}$. This is not a problem because, by remembering that L is arbitrary, we can obtain the same results using $L/2$ in place of L.

Hence, the analysis that we will carry out in each of the different pieces applies to the whole space.

3.4. **Averaging close to simple resonances.** Theorem 16 reduces the original system to a non-integrable system of a partially simple form by a change of variables. The number of angles that enter in the averaged Hamiltonian for a given I depends on the resonances which are close to I. The key of the following reasoning is to understand the geometry of the set of points I for which the averaged system involves only one angle. The set of points that we have found useful to omit are the points in secular resonances (i.e. in the resonances which appear in averaging to order 1 or 2) which are also part of another resonance activated when averaging up to order m.

So, next step is to define a region $\mathcal{I}_\delta \subset \mathcal{J}_\delta$ where we take out the intersection of the secular resonances with any other resonances which appear in the process of averaging up to order m.

To this end we consider

$$\mathbf{B} = \mathcal{R}^{[\leq 2]} \cap \{I \in \mathcal{J}_\delta, \ \mathbf{m}(\omega(I), m) \geq 2\}$$

which is a finite union of surfaces of codimension two or higher in $\mathcal{J}_\delta$, and we consider $\mathbf{B}_\delta$ a δ-neighborhood of these surfaces. Reducing $L = L(\delta)$ if necessary, the set $\mathcal{I}_\delta = \mathcal{J}_\delta \setminus \mathbf{B}_\delta$ verifies the following property:

L2 If $I \in \mathcal{I}_\delta$ there is at most one $\mathcal{R}_{k,l} \in \mathcal{R}^{[\leq 2]}$ such that $d(I, \mathcal{R}_{k,l}) \leq L$.

Theorem 16 in the domain $\mathcal{I}_\delta$ reads:

Theorem 17. *Let $\tilde{K}(\tilde{I}, \varphi, s; \varepsilon)$ be the $\mathcal{C}^n$ Hamiltonian of Theorem 16, $n > 1$, and consider any $1 \leq m < n$, independent of ε.*

Consider the finite collection of sets $\mathcal{R}^i \subset \mathcal{I}^*$, called resonances activated at order i, $i = 1, \ldots m$, given in definition 13.

Let $0 < \delta < 1$ be any number and consider $0 < L < 1$ verifying L1 and L2 in $\mathcal{I}_\delta$.

Then, the symplectic change of variables given in Theorem 16,

$$(I, \varphi, s) \mapsto (\mathcal{B}, \alpha, s),$$

transforms the Hamiltonian system associated to $\tilde{K}(I, \varphi, s; \varepsilon)$ in $\mathcal{I}_\delta$ into a Hamiltonian system of Hamiltonian

(40) $\quad B_0 + \bar{K}(\mathcal{B}, \alpha, s; \varepsilon) = B_0 + \bar{K}^0(\mathcal{B}, \alpha, s; \varepsilon) + \varepsilon^{m+1} \bar{K}^1(\mathcal{B}, \alpha, s; \varepsilon)$

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 $\mathcal{B} \in \mathcal{I}_\delta$, satisfies $d(\mathcal{B}, \mathcal{R}^{[\leq 2]}) \geq 2L$, then*

$$\bar{K}^0(\mathcal{B}, \alpha, s; \varepsilon) = h(\mathcal{B}) + \varepsilon \bar{K}^{0,0}(\mathcal{B}; \varepsilon) + O(\varepsilon^3)$$

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

(2) *If $\mathcal{B} \in \mathcal{I}_\delta$, satisfies $d(\mathcal{B}, \mathcal{R}^{[\leq 2]}) \leq L$, there exists a unique resonance activated at order one or two*

$$\mathcal{R}_{k_0, l_0} \in \mathcal{R}^j, \quad j = 1, 2$$

such that $d(\mathcal{B}, \mathcal{R}_{k_0, l_0}) \leq L$. Then

(41) $\quad \begin{aligned} \bar{K}^0(\mathcal{B}, \varphi, s; \varepsilon) & = h(\mathcal{B}) + \varepsilon \bar{K}^{0,0}(\mathcal{B}; \varepsilon) \\ & + \varepsilon^j U^{k_0, l_0}(\Gamma_{k_0, l_0}(\mathcal{B}), k_0 \cdot \alpha + l_0 s; \varepsilon) \end{aligned}$

where the functions $U^{k_0, l_0}(\Gamma_{k_0, l_0}(\mathcal{B}), \theta; \varepsilon)$ are polynomial in ε and trigonometric polynomial in the angle variable $\theta = k_0 \cdot \alpha + l_0 s$.

Moreover, if $\mathcal{R}_{k_0, l_0} \subset \mathcal{R}^1$ the function U^{k_0, l_0} is given by:

$$U^{k_0, l_0} = \sum_{p=-N_1}^{N_1} K^1_{pk_0, pl_0}(\Gamma_{k_0, l_0}(\mathcal{B})) e^{2\pi i p(k_0 \cdot \alpha + l_0 s)} + O(\varepsilon)$$

where $K^1_{k, l}(\mathcal{B})$ are the Fourier coefficients of $K(\mathcal{B}, \alpha, s; 0)$ with respect to the angle variables (α, s).

The next goal is to study in more detail the behavior of the system predicted by the averaged Hamiltonian.

The main remark is that, near simple resonances, the averaged system contains only one angle and, therefore, it is integrable. This allows us to analyze explicitly its dynamics. Its turns out that, for the problem at hand, we only need to study the resonances of order 1 or 2, which are called "secular resonances" by astronomers.

3.5. **Geometric properties of the orbits of the averaged Hamiltonian.** In this section, we study the invariant tori of the averaged system obtained in Theorem 17, that is, the system given by Hamiltonian $\bar{K}^0$ in (40). Later, in section 3.5.2, 3.5.4, we show that, under some non-degeneracy conditions, some of these tori are also present in the original system. This is, basically, the KAM theorem.

In Sections 3.5.1, 3.5.3, we will see that the phase space $\mathcal{I}_\delta \times \mathbb{T}^{d+1}$ is foliated by (quasi)-periodic solutions of the averaged system. Nevertheless, the topology of the solutions is very different in the non-resonant regions and in the resonant regions. We define the *non-resonant region* as the set:

$$(42) \qquad \mathcal{S}^L = \{(I, \varphi, s) \in \mathcal{I}_\delta \times \mathbb{T}^{d+1}, \ d(I, \mathcal{R}^{[\leq 2]}) \geq 2L\}$$

In particular, $\mathcal{S}^L$ includes the intersection of $\mathcal{I}_\delta$ with all the resonances activated at order higher than 2. This region $\mathcal{S}^L$ will be covered, up to very small gaps of order $O(\varepsilon^{3/2})$, by KAM tori.

In the resonant regions of $\mathcal{I}_\delta \times \mathbb{T}^{d+1} \backslash \mathcal{S}^L$, we will obtain tori which are contractible to tori of lower dimension and, therefore, are not homotopic to a torus present in the unperturbed system. We call *secondary* KAM tori the invariant tori which have different topological type from the tori of the unperturbed system. We use the name *primary* tori for the invariant tori which are homotopic to those of the unperturbed system. Primary tori are those usually considered in the perturbative versions of KAM theorem for quasi-integrable systems.

The importance of the secondary tori is that they dovetail precisely into the gaps between the set of KAM primary tori created by the resonances, so that it will be possible to construct a web of KAM tori,

primary and secondary, which are $\varepsilon^{3/2}$-close. For systems with $2+1/2$ degrees of freedom this was introduced in [DLS03, DLS06a]. See also [GL06a].

In hypothesis **H6** given in 3.5.3, we formulate precisely one non-degeneracy assumption on the averaged system which allows us to apply the KAM theorem and conclude that some of the solutions found in the averaged system $\bar{K}^0$ (including secondary tori) are indeed present in the full Hamiltonian (40), and therefore in the original system (16). Of course, since the averaged system is computable from the original model, the non-degeneracy conditions on the averaged system amount to some non-degeneracy conditions on the original system.

3.5.1. *The invariant tori of the averaged system in the non-resonant region of $\mathcal{I}_\delta$.* By item (1) in Theorem 17, in the non-resonant region $\mathcal{S}^L$ defined in (42), the full averaged hamiltonian (40) reads

$$\mathcal{B}_0 + h(\mathcal{B}) + \varepsilon \bar{K}^{0,0}(\mathcal{B};\varepsilon) + O(\varepsilon^3).$$

For the truncated Hamiltonian

(43) $$\mathcal{B}_0 + h(\mathcal{B}) + \varepsilon \bar{K}^{0,0}(\mathcal{B};\varepsilon)$$

the tori are given as the level sets of the averaged action variables

$$\mathcal{B}_0 = c_0, \mathcal{B}_1 = c_1, \ldots, \mathcal{B}_d = c_d$$

where the equation $\mathcal{B}_0 = c_0$ is a reflection of the fact that the Hamiltonian (43) is autonomous.

When written in the original variables of the time dependent Hamiltonian (16), these tori in $\mathcal{S}^L \subset \mathcal{I}_\delta \times \mathbb{T}^{d+1}$, as shown in Theorem 16 and Theorem 17, are given by the equations:

$$F_1(I, \varphi, s; \varepsilon) = c_1, \ldots, F_d(I, \varphi, s; \varepsilon) = c_d$$

where

(44) $$F_1(I, \varphi, s; \varepsilon) = I_1 + O(\varepsilon), \ldots, F_d(I, \varphi, s; \varepsilon) = I_d + O(\varepsilon).$$

3.5.2. *The invariant tori in the non-resonant region of $\mathcal{I}_\delta$: KAM Theorem.* We note that in the non-resonant region $\mathcal{S}^L$, we have managed to transform the system into an integrable system up to an error which is ε^3 when measured in the C^{r-3} norm.

Furthermore, we point out that the averaged part has a frequency map which is a diffeomorphism (it is an $O(\varepsilon)$ perturbation of the diffeomorphism $I \mapsto \frac{\partial h}{\partial I}$ in a smooth norm).

If r is sufficiently large (so that $r - 3$ is larger than $2d + 3$) we can apply a KAM theorem [Pös82] and conclude that there are invariant

tori which cover the non-resonant region $\mathcal{S}^L$ except for a set of measure smaller than $O(\varepsilon^{3/2})$.

Theorem 18. *Under the conditions of theorem 5, there exists ε_0 such that, for $0 < |\varepsilon| < \varepsilon_0$, the region $\mathcal{S}^L$ can be covered by $O(\varepsilon^{3/2})$ neighborhoods of invariant objects under the Hamiltonian flow of the Hamiltonian $K_\varepsilon(I, \varphi, s)$ in (16). Moreover:*

- *These invariant objects are given by the level sets $F = E$, for $|E - E'| \leq \varepsilon^{\frac{3}{2}}$.*
- *The $\mathcal{C}^2$ function $F : \mathbb{R}^d \times \mathbb{T}^d \times \mathbb{T} \to \mathbb{R}$ is given by (44).*
- *These invariant objects are regular primary KAM $d + 1$-tori*

Therefore, in the non-resonant region, each torus has several tori which are much closer than $O(\varepsilon)$ to it. This is what in [CG98] was called the *"gap bridging mechanism"*.

Remark 19. For experts, we note that there are different KAM theorems in the literature, which differ in some subtle features; a systematic comparison can be found in [Lla01]. The main difference in the literature is whether one step of averaging requires to solve one cohomology equation or two. The methods which use only one cohomology equation (e.g. the method in [Arn63, Pös82, Sal04]), called *first order methods*, establish that the gaps between tori are bounded by the error to the power $1/2$. Those that use two cohomology equations (e.g. the methods in [Kol54, Mos66, Zeh75, Zeh76a]), called *second order methods*, lead to gaps which are bounded by the error to power $1/4$. These quantitative estimates for the Newton method are found in [Zeh76b] Very explicit verifications of the quantitative estimates of the method of [Arn63] appear in [Neĭ81]. Simple examples show that the exponent $1/2$ cannot be increased.

In our case, either method could be applied. If we wanted to just refer to the second order methods to obtain gaps of order $O(\varepsilon^{1/4})$, it would have been enough, to obtain gaps of order $O(\varepsilon^{3/2})$, to define the non-resonant region as the region where one can average to order $m = 5$ instead of $m = 2$ (in fact, it would be enough $m = 4$ if we allow gaps of order $O(\varepsilon^{5/4})$).

Another technical point is that some results in the literature lose more derivatives. This is totally irrelevant for us since it only affects the number of derivatives that we need to assume in the original Hamiltonian.

3.5.3. *Primary and secondary invariant tori of the averaged system in the resonant region.* We define the *secular resonant region* as:

$$(45) \qquad \mathcal{S}^{[\leq 2]} = \{(I, \varphi, s) \in \mathcal{I}_\delta \times \mathbb{T}^{d+1}, \ d(I, \mathcal{R}^{[\leq 2]}) \leq L\}$$

In this region we will perform an elementary change of variables that makes it clear that, close to a resonance activated at order one or two, the Hamiltonian is a function of $d - 1$ actions and one resonant angle.

The first observation is that the region $\mathcal{S}^{[\leq 2]}$ is the union of the regions $\mathcal{R}_{k,l}^{L}$, which consist on tubular neighborhoods of size L of the resonances $\mathcal{R}_{k,l}$ defined in Definition 9.

For points $(\mathcal{B}, \alpha, s) \in \mathcal{R}_{k_0, l_0}^{L}$, where

$$\mathcal{R}_{k_0, l_0} \subset \mathcal{R}^{j} \subset \mathcal{R}^{[\leq 2]}, \quad j = 1, 2$$

the averaged Hamiltonian is

$$(46) \qquad \mathcal{B}_0 + \bar{K}^0(\mathcal{B}, \alpha, s; \varepsilon)$$

where, by Theorem 17:

(47)
$$\bar{K}^0(\mathcal{B}, \alpha, s; \varepsilon) = h(\mathcal{B}) + \varepsilon \bar{K}^{0,0}(\mathcal{B}; \varepsilon) + \varepsilon^j U^{k_0, l_0}(\Gamma_{k_0, l_0}(\mathcal{B}), k_0 \cdot \alpha + l_0 s; \varepsilon)$$

is given in (41).

Assuming $k_0^m \neq 0$ for some $1 \leq m \leq d$, to understand the geometry of the averaged Hamiltonian we first perform the following change of angles

$$(48) \qquad \tilde{\theta} = M \tilde{\alpha},$$

where $\tilde{\theta} = (s, \theta)$, and $\tilde{\alpha} = (s, \alpha)$, and M is the $d + 1 \times d + 1$ matrix:

$$M = \begin{pmatrix} 1 & \cdots & \\ \cdot & \cdot & \\ l_0 & k_0^{\top} & \\ \cdot & \cdot & \\ & & 1 \end{pmatrix}$$

Let us observe that this change is just to take as a new angle the resonant angle:

$$\theta_i = \alpha_i, \quad i \neq m, \quad \theta_m = k_0 \cdot \alpha + l_0 s$$

To make the change symplectic we perform the change in actions:

$$(49) \qquad \tilde{J} = M^{-\top} \tilde{\mathcal{B}}.$$

The change $\tilde{\mathcal{B}} = M^{\top} \tilde{J}$ is equivalent to $\mathcal{B} = N^{\top} J$ where

$$N = \begin{pmatrix} 1 & \cdots & \\ \cdot & \cdot & \\ & k_0^{\top} & \\ \cdot & \cdot & \\ & & 1 \end{pmatrix}$$

and $\mathcal{B}_0 = J_0 + l_0 J_m$, which is just the change in action which reflects the fact that we are doing a time dependent change of angles.

In components, this change is simply:

$$\mathcal{B}_m = k_0^m J_m, \quad \mathcal{B}_i = J_i + k_0^i J_m, \quad i \neq m$$

or, analogously

$$(50) \qquad J_m = \frac{\mathcal{B}_m}{k_0^m}, \quad J_i = \mathcal{B}_i - \frac{\mathcal{B}_m}{k_0^m} k_0^i, \quad i \neq m$$

With this change, the averaged Hamiltonian (46) is given by:

$$(51) \quad J_0 + l_0 J_m + h(N^\top J) + \varepsilon \bar{K}^{0,0}(N^\top J; \varepsilon) + \varepsilon^j U^{k_0, l_0}(\Gamma_{k_0, l_0}(N^\top J), \theta_m; \varepsilon)$$

which, in the region $\mathcal{R}_{k_0, l_0}^L \subset \mathcal{I}_\delta \times \mathbb{T}^{d+1}$, corresponds to the autonomous Hamiltonian:

$$(52) \quad l_0 J_m + h(N^\top J) + \varepsilon \bar{K}^{0,0}(N^\top J; \varepsilon) + \varepsilon^j U^{k_0, l_0}(\Gamma_{k_0, l_0}(N^\top J), \theta_m; \varepsilon).$$

Working in the variables (J, θ, s) makes easier to identify the invariant tori. The invariant tori of Hamiltonian (52) will be given by prescribing the values of the $d-1$ action variables J_i, for $i \neq m$, and the value of the Hamiltonian (52), which is a constant of motion.

Abusing slightly the notation, let us write $J = (\hat{J}, J_m)$, $\theta = (\hat{\theta}, \theta_m)$, with

$$\hat{J} = (J_1, \ldots, J_{m-1}, J_{m+1}, \ldots, J_d), \ \hat{\theta} = (\theta_1, \ldots, \theta_{m-1}, \theta_{m+1}, \ldots, \theta_d).$$

Given a value of $\mathcal{B} = N^\top J = (\hat{\mathcal{B}}, \mathcal{B}_m)$, we want to compute its projection $\Gamma_{k_0, l_0}(\mathcal{B}) = \mathcal{B}_{k_0, l_0}^* = N^\top J^*$.

By the k_0-characterization (29) of the projection Γ_{k_0, l_0} we have that

$$\mathcal{B} - \mathcal{B}_{k_0, l_0}^* \in \langle k_0 \rangle \text{ and } Dh(\mathcal{B}_{k_0, l_0}^*) \cdot k_0 + l_0 = 0.$$

Note that using the form of the change (50), and since $\mathcal{B} = N^\top J$ and $\mathcal{B}_{k_0, l_0}^* = N^\top J^*$, we have that $J_i = J_i^*$, if $i \neq m$ and that $\mathcal{B} - \mathcal{B}_{k_0, l_0}^* = (J_m - J_m^*) k_0$.

Then, given $\mathcal{B} = N^\top J$, with $J = (\hat{J}, J_m)$, one can characterize the projection $\Gamma_{k_0, l_0}(\mathcal{B})$ as:

We compute:

$$(53) \qquad \hat{J} = \hat{\mathcal{B}} - \frac{\mathcal{B}_m}{k_0^m} \hat{k}_0, \quad J_m = \frac{\mathcal{B}_m}{k_0^m}$$

and therefore

$$\mathcal{B} = (\hat{J}, 0) + \frac{\mathcal{B}_m}{k_0^m} k_0,$$

one can obtain the projection $\mathcal{B}_{k_0,l_0}^* = N^\top J^*$, with $J^* = (\hat{J}, J_m^*)$ in terms of $\hat{J}$:

$$(54) \qquad \mathcal{B}_{k_0,l_0}^* = \mathcal{B}^*(\hat{J}) = (\hat{J}, 0) + \frac{\mathcal{B}_m^*}{k_0^m} k_0$$

where $\mathcal{B}_m^* = k_0^m J_m^*$ and $J_m^* = J_m^*(\hat{J})$ is obtained solving

$$(55) \qquad Dh(N^\top(\hat{J}, J_m^*)) \cdot k_0 + l_0 = 0.$$

Let us emphasize that J_m^* and therefore J^* and $\mathcal{B}_{k_0,l_0}^*$ are uniquely determined through (55) for a given $\mathcal{B} = N^\top J$, and therefore for a given $\hat{J}$, assuming $\mathrm{dist}(\mathcal{B}, \mathcal{R}_{k_0,l_0}) < L$ and L small enough.

Let us observe that for values $\mathcal{B}$ such that $(\mathcal{B}, \alpha, s) \in \mathcal{I}_\delta$, one has that the corresponding values of $\hat{J}$ vary in a compact set that we will denote by $\hat{\mathcal{J}}$.

For $\hat{J} \in \hat{\mathcal{J}}$, we will denote:

$$(56) \qquad U^{k_0,l_0,*}(\theta_m; \hat{J}, \varepsilon) = U^{k_0,l_0}(\mathcal{B}_{k_0,l_0}^*(\hat{J}), \theta_m; \varepsilon)$$

Using this notation, system (52) can be written as:

$$(57) \quad l_0 J_m + h(N^\top(\hat{J}, J_m)) + \varepsilon \bar{K}^{0,0}(N^\top(\hat{J}, J_m); \varepsilon) + \varepsilon^j U^{k_0,l_0,*}(\theta_m; \hat{J}, \varepsilon).$$

The next step is to use Hypotheses **H5** and **H6** to obtain a change of variables

$$(J_m, \theta_m) \to (y, x)$$

such that Hamiltonian (57) in these new variables becomes:

$$(58) \qquad \mathcal{K}_0(y, x; \hat{J}, \varepsilon) = a(\hat{J}, \varepsilon) \frac{y^2}{2}(1 + O(y)) + \varepsilon^j U(x; \hat{J}, \varepsilon).$$

We first proceed to formulate hypothesis **H5** which states precisely that the leading part of the kinetic energy is quadratic.

Taylor expanding the function $h(N^\top J)$ around the resonant point $N^\top J^*$ and using (55), we obtain:

$$
\begin{aligned}
l_0 J_m + h(N^\top J) &= l_0 J_m + h(N^\top J^* + k_0(J_m - J_m^*)) \\
&= l_0 J_m^* + h\left(N^\top J^*\right) \\
(59) \qquad\qquad &\quad + \frac{1}{2}(J_m - J_m^*)^2 k_0^\top D^2 h\left(N^\top J^*\right) k_0 \\
&\quad + O((J_m - J_m^*)^3)
\end{aligned}
$$

where we have used that

$$Dh(\mathcal{B}_{k_0,l_0}^*(\hat{J}))k_0 + l_0 = 0, \quad \forall \hat{J} \in \hat{\mathcal{J}}.$$

by the definition of $\mathcal{B}^*_{k_0,l_0}(\hat{J}) = N^\top J^*$ in (54). Therefore, introducing

$$(60) \quad a(\hat{J}) = a(\hat{J})_{k_0,l_0} := k_0^\top D^2 h \left(N^\top J^* \right) k_0 = k_0^\top D^2 h \left(\mathcal{B}^*_{k_0,l_0}(\hat{J}) \right) k_0,$$

the hypothesis **H5** is:

    H5 $a(\hat{J}) \neq 0$, for any $(k_0, l_0) \in \mathcal{R}^{[\leq 2]}$ and $\hat{J} \in \hat{\mathcal{J}}$.

With the notation (60), equation (59) becomes
(61)

$$l_0 J_m + h(N^\top J) = l_0 J_m^* + h(N^\top(\hat{J}, J_m^*)) + \frac{a(\hat{J})}{2}(J_m - J_m^*)^2 + O((J_m - J_m^*)^3).$$

Since the actions $\hat{J}$ are $d-1$ first integrals of the averaged Hamiltonian (52), we have that the dynamics in the (J_m, θ_m) variables is that of a nonlinear oscillator with potential $U^{k_0,l_0,*}(\theta_m; \hat{J}, \varepsilon)$. We can think of the variables $\hat{J}$ as parameters in the non-linear oscillator.

We now introduce Hypothesis **H6**, which is the other non-degeneracy assumption which will make precise the heuristic notion that *"the averaged system near secular resonances looks like a pendulum"*.

Assumption **H6** formulates precisely that the potential of the truncated averaged Hamiltonian (47) at the resonance (see also (57)), if $a(\hat{J}) > 0$, has a unique non-degenerate maximum or, if $a(\hat{J}) < 0$, has a unique non-degenerate minimum.

    H6 For any $\mathcal{B} = N^\top J \in \mathcal{S}^{[\leq 2]} \subset \mathcal{I}_\delta$, consider the value (k_0, l_0) such that $d(\mathcal{B}, \mathcal{R}_{k_0,l_0}) \leq L$ and its k_0-projection $\Gamma_{k_0,l_0}(\mathcal{B}) = \mathcal{B}^*_{k_0,l_0}(\hat{J})$. By hypothesis **H5** we know that $a(\hat{J}) = a(\hat{J})_{k_0,l_0} \neq 0$.

        If $a(\hat{J}) > 0$, we assume that there is a unique non-degenerate maximum of the potential of Hamiltonian (47)

$$U^{k_0,l_0}(\Gamma_{k_0,l_0}(\mathcal{B}), \theta_m; 0) = U^{k_0,l_0}(\mathcal{B}^*_{k_0,l_0}(\hat{J}), \theta_m; 0)$$

        with respect to θ_m, which is uniformly non-degenerate with respect to $\mathcal{B} \in \mathcal{S}^{[\leq 2]}$. If $a(\hat{J}) < 0$, we assume instead that there is a unique non-degenerate minimum with the same uniformity conditions.

        That is, there is a unique θ_m^* such that

$$a(\hat{J})U^{k_0,l_0}(\Gamma_{k_0,l_0}(\mathcal{B}), \theta_m^*; 0) = \max_{\theta_m} a(\hat{J})U^{k_0,l_0}(\Gamma_{k_0,l_0}(\mathcal{B}), \theta_m; 0),$$

(62)

$$a(\hat{J})\frac{\partial^2}{\partial \theta_m^2}U^{k_0,l_0}(\Gamma_{k_0,l_0}(\mathcal{B}), \theta_m^*; 0) \leq \beta < 0.$$

Remark 20. We call attention to the fact that, as $\Gamma_{k_0,l_0}(\mathcal{B}) \in \mathcal{R}_{k_0,l_0} \subset \mathcal{R}^{[\leq 2]}$ the conditions **H5**, **H6** need to be verified only on the codimension one set $\mathcal{R}^{[\leq 2]} \cap \mathcal{I}_\delta \times \mathbb{T}^{d+1}$ formed by the secular resonances in $\mathcal{I}_\delta \times \mathbb{T}^{d+1}$.

Note that the assumptions **H5**, **H6** are $\mathcal{C}^r$ open conditions in the space of Hamiltonians. If assumptions **H5**, **H6** are verified for a family, they will also be verified for all families close to it in a $\mathcal{C}^r$ topology with $r \geq 2$ sufficiently large so that we can carry out the averaging procedure. Therefore, the conditions also hold in a $\mathcal{C}^r$ set of original Hamiltonians.

Using the notation introduced in (56) hypothesis **H6** can be written as:

(63)
$$a(\hat{J})U^{k_0,l_0,*}(\theta_m^*; \hat{J}, 0) = \max_{\theta_m} a(\hat{J})U^{k_0,l_0,*}(\theta_m; \hat{J}, 0),$$
$$a(\hat{J})\frac{\partial^2}{\partial\theta_m^2}U^{k_0,l_0,*}(\theta_m^*; \hat{J}, 0) \leq \beta < 0.$$

From now on we will assume that $a(\hat{J}) > 0$. Moreover, it is uniform respect to $\hat{J} \in \hat{\mathcal{J}}$. Therefore hypothesis **H6** can be written as:

(64)
$$U^{k_0,l_0,*}(\theta_m^*; \hat{J}, 0) = \max_{\theta_m} U^{k_0,l_0,*}(\theta_m; \hat{J}, 0),$$
$$\frac{\partial^2}{\partial\theta_m^2}U^{k_0,l_0,*}(\theta_m^*; \hat{J}, 0) \leq \beta < 0.$$

The case $a(\hat{J}) < 0$ can be done analogously.

Assumptions **H5**, **H6** imply that, as a function of (J_m, θ_m), for any value of $\hat{J}$, the Hamiltonian

$$l_0 J_m + h(N^\top(\hat{J}, J_m)) + \varepsilon^j U^{k_0,l_0,*}(\theta_m; \hat{J}, 0)$$

has a saddle point at $(J_m^*(\hat{J}), \theta_m^*(\hat{J}))$, which gives rise to a saddle equilibrium point for the associated Hamiltonian system.

We note that, because of uniformity of the second derivative of the potential in (64) and the hypothesis **H5**, we obtain that the point $(J_m^*(\hat{J}), \theta_m^*(\hat{J}))$ is uniformly hyperbolic. For a given $\varepsilon > 0$, the Liapunov exponents are bounded away from zero for any $\hat{J} \in \hat{\mathcal{J}}$ uniformly, and the angle between its stable and unstable directions is also bounded away from zero.

Therefore, by the implicit function Theorem, for $|\varepsilon| < \varepsilon_0$, the Hamiltonian system associated to Hamiltonian (57) in the phase space of (J_m, θ_m), has a saddle equilibrium point

(65) $\qquad (\tilde{J}_m(\hat{J}, \varepsilon), \tilde{\theta}_m(\hat{J}, \varepsilon)) = (J_m^*(\hat{J}), \theta_m^*(\hat{J})) + O(\varepsilon).$

for any $\hat{J} \in \hat{\mathcal{J}}$.

To make the pendulum-like structure of the system given by Hamiltonian (57) more apparent and to analyze the behavior, we will find it convenient to make the translation

(66) $\qquad y = J_m - \tilde{J}_m(\hat{J}, \varepsilon), \qquad x = \theta_m - \tilde{\theta}_m(\hat{J}, \varepsilon), \quad s = s,$

and we obtain the $\mathcal{C}^{r-2m-2}$ Hamiltonian

(67) $\qquad \mathcal{K}_0(y, x; \hat{J}, \varepsilon) = h_0(y; \hat{J}, \varepsilon) + \varepsilon^j U(x; \hat{J}, \varepsilon)$

where

$$
\begin{aligned}
h_0(y; \varepsilon\hat{J},) = {}& l_0 y + h\left(N^\top(\hat{J}, \tilde{J}_m + y)\right) \\
& - h\left(N^\top(\hat{J}, \tilde{J}_m)\right) \\
& + \varepsilon \bar{K}^{0,0}\left(N^\top(\hat{J}, \tilde{J}_m + y); \varepsilon\right) \\
& - \varepsilon \bar{K}^{0,0}\left(N^\top(\hat{J}, \tilde{J}_m); \varepsilon\right)
\end{aligned}
$$

(68)

$$U(x; \hat{J}, \varepsilon) = U^{k_0, l_0, *}(\tilde{\theta}_m + x; \hat{J}, \varepsilon) - U^{k_0, l_0, *}(\tilde{\theta}_m; \hat{J}, \varepsilon)$$

where we have subtracted a constant term to the averaged Hamiltonian (57), the energy of the saddle $(\tilde{J}_m, \tilde{\theta}_m) = (\tilde{J}_m(\hat{J}, \varepsilon), \tilde{\theta}_m(\hat{J}, \varepsilon))$, to normalize:

$$h_0(0; \hat{J}, \varepsilon) = \frac{\partial h_0}{\partial y}(0; \hat{J}, \varepsilon) = 0,$$

(69) $\qquad \dfrac{\partial^2 h_0}{\partial y^2}(0; \hat{J}, \varepsilon) = a(\hat{J}, \varepsilon) = a(\hat{J}) + O(\varepsilon) \neq 0,$

$$U(0; \hat{J}, \varepsilon) = \frac{\partial U}{\partial x}(0; \hat{J}, \varepsilon) = 0, \ \frac{\partial^2 U}{\partial x^2}(0; \hat{J}, \varepsilon) \leq \beta < 0,$$

therefore the averaged Hamiltonian (67) can be written as:

(70) $\qquad \mathcal{K}_0(y, x; \hat{J}, \varepsilon) = a(\hat{J}, \varepsilon)\dfrac{y^2}{2}(1 + O(y)) + \varepsilon^j U(x; \hat{J}, \varepsilon)$

and $(0, 0)$ is a saddle point of Hamiltonian (67), with energy level $\mathcal{K}_0(0, 0; \hat{J}, \varepsilon) = 0$.

Once we have the averaged Hamiltonian written in the form (58), we discuss the geometry and the dynamics on the sets obtained by fixing the energy level.

The main observation is that if we fix $\hat{J} = \hat{c}$, there is a critical value $c_m^*(\hat{c}, \varepsilon) = 0 = \mathcal{K}_0(0, 0; \hat{J}, \varepsilon)$ for c_m at which the topology and

37

the dynamics of the level sets of the Hamiltonian $\mathcal{K}_0(y, x; \hat{J}, \varepsilon)$, and therefore of the Hamiltonian (57), change.

Now we describe the invariant sets of Hamiltonian $\mathcal{K}_0(y, x, \hat{J}; \varepsilon)$ given in (67) in the region:

$$(71) \qquad D = \{(\hat{J}, \hat{\theta}, y, x, s) \in \hat{\mathcal{J}} \times \mathbb{T}^{d-1} \times \mathbb{R} \times \mathbb{T}^2, \quad |y| \leq \bar{L}\}.$$

for some $0 < \bar{L} < L$, where $\hat{\theta} = (\theta_1, \ldots, \theta_{m-1}, \theta_{m+1}, \ldots, \theta_d)$.

Given any value $\hat{J} = \hat{c}$ we consider in the (y, x) space, the level set $\mathcal{K}_0(y, x; \hat{J}, \varepsilon) = c_m$:

- When $c_m > 0$ but close enough to zero, the level set in the (y, x) annulus is composed by two non-contractible circles.
- When $c_m < 0$ but close enough to zero, the level set in the (y, x) annulus is a circle which, however, is contractible to a point.
- When $c_m = 0$, the level set is the union of two separatrices and the hyperbolic critical point $(0, 0)$.

Therefore, the region D is filled by the the level sets of the constants of motion, that is, the energy surfaces of the Hamiltonian $\mathcal{K}_0$, and the corresponding $\hat{J}$:

(72)
$$\mathcal{T}_c^0 = \{(\hat{J}, \hat{\theta}, y, x, s) \in \mathbb{R}^{d-1} \times \mathbb{T}^{d-1} \times \mathbb{R} \times \mathbb{T}^2 : \mathcal{K}_0(y, x; \hat{J}, \varepsilon) = c_m, \hat{J} = \hat{c}\}.$$

$\mathcal{T}_c^0$ will, of course, be invariant by the Hamiltonian flow of $\mathcal{K}_0$.

The sets $\mathcal{T}_c^0$ consist on:

- When $c_m > 0$ but close enough to zero, the level set $\mathcal{T}_c^0$ is composed by two primary tori (non-contractible tori of dimension $d + 1$).
- When $c_m < 0$ but close enough to zero, the level set $\mathcal{T}_c^0$ is a secondary torus (torus of dimension $d + 1$ contractible to a d-dimensional torus).
- The level set corresponding to $c_m = 0$ consists of one whiskered torus and its coincident whiskers: the hyperbolic torus $\mathbb{T}^d \times \{(0, 0)\}$ and the homoclinic orbits to it. We will refer to $\mathcal{T}_0^0$ as the separatrix loop.

Formula (72) gives an implicit equation for the tori $\mathcal{T}_c^0$. To compute the images of these tori under the scattering map in section 3.8, it will be convenient to have the explicit equation of these tori.

Lemma 21. *There exists $\rho > 0$, such that the two primary tori (components of the secondary tori) $\mathcal{T}_c^0$ of Hamiltonian $\mathcal{K}_0$ can be written as graphs of the variables $(\hat{J}, y)$ over the angle variables $(\hat{\theta}, x)$, for*

$\rho \leq x \leq 2\pi - \rho$:

$$(73) \quad \mathcal{T}_c^0 = \{(\hat{J}, \hat{\theta}, y, x, s) \in \mathbb{R}^{d-1} \times \mathbb{T}^{d-1} \times \mathbb{R} \times [\rho, 2\pi - \rho] \times \mathbb{T} :$$
$$\hat{J} = \hat{c}, \ y = \pm \tilde{\mathcal{Y}}(x; c, \varepsilon)\}.$$

where the function $\tilde{\mathcal{Y}}(x; c, \varepsilon)$ has different expressions depending of the value c_m:

(1) *If $0 < c_m \leq \varepsilon^j$:*

$$(74) \qquad \tilde{\mathcal{Y}}(x; c, \varepsilon) = \tilde{\ell}(x; c, \varepsilon)(1 + O(\varepsilon^{\frac{j}{2}}))$$

(2) *If $\varepsilon^j \leq c_m < 1$, for $c_m = d_m \varepsilon^\alpha$, with $0 < \alpha < j$:*

$$(75) \qquad \tilde{\mathcal{Y}}(x; c, \varepsilon) = \tilde{\ell}(x; c, \varepsilon)(1 + O(\varepsilon^{\frac{\alpha}{2}}))$$

(3) *If $c_m = O(1)$:*

$$(76) \qquad \tilde{\mathcal{Y}}(x; c, \varepsilon) = h_0^{(-1)}(c_m)(1 + O(\varepsilon^j))$$

where the function where h_0 is given in (69) *and $\tilde{\ell}$ is given by:*

$$\tilde{\ell}(x; c, \varepsilon) = \sqrt{\frac{2}{a(\hat{c}, \varepsilon)}(c_m - \varepsilon^j U(x; \hat{c}, \varepsilon))}$$

Once we know the structure of the level sets of the averaged Hamiltonian (57) in terms of the variables $(\hat{J}, \hat{\theta}, y, x)$, we can write the equations of these sets in the original variables of the problem.

First, in terms of the variables $(\mathcal{B}, \alpha, s) \in \mathbb{R}^d \times \mathbb{T}^{d+1}$, using (50), (53), (66) and (67) and (68), the tori $\mathcal{T}_E^0$, with $E = (\hat{E}, \tilde{E}_m)$, are given by:

$$\hat{\mathcal{B}} - \frac{\mathcal{B}_m}{k_0^m} \hat{k}_0 = \hat{E}$$

$$(77) \quad \frac{\mathcal{B}_m}{k_0^m} l_0 + h(\mathcal{B}) + \varepsilon K^{0,0}(\mathcal{B}; \varepsilon) + \varepsilon^j U^{k_0, l_0, *}(k_0 \alpha + l_0 s; \hat{E}, \varepsilon) = \tilde{E}_m.$$

$\mathcal{T}_E^0$ will, of course, be invariant by the Hamiltonian flow of the averaged system (47).

Remark 22. The equations (77) are a natural consequence of the fact that the non-autonomous Hamiltonian (47) has as first integrals the functions that are at the left hand side of (77).

Let us observe that the relation between the constants c and E is given by

$$\hat{E} = \hat{c}$$
$$\tilde{E}_m = c_m + \tilde{E}_m^*(\hat{E}) = c_m + l_0 \tilde{J}_m + h(N^\top(\hat{E}, \tilde{J}_m))$$
$$+ \varepsilon \bar{K}^{0,0}(N^\top(\hat{E}, \tilde{J}_m)) + \varepsilon^j U^{k_0, l_0, *}(\tilde{\theta}_m; \hat{E}, \varepsilon)$$

and the critical value where the topology of the invariant tori change is now $\tilde{E}_m^* = \tilde{E}_m^*(\hat{E})$, which is the energy level of the saddle $(\tilde{J}_m, \tilde{\theta}_m) = (\tilde{J}_m(\hat{E}, \varepsilon), \tilde{\theta}_m(\hat{E}, \varepsilon))$ and corresponds to taking the critical value $c_m = 0$, $\hat{c} = \hat{E}$, in (72), of Hamiltonian $\mathcal{K}_0$ in (67) with variables (66).

It is important to note that equations (77) can also be written, using (50), (66) and (53), as:

$$\hat{\mathcal{B}} - \frac{\mathcal{B}_m}{k_0^m}\hat{k}_0 = \hat{E}$$

(78)
$$a(\hat{E}, \varepsilon)\frac{y^2}{2}(1 + O(y)) + \varepsilon^j U^{k_0, l_0, *}(k_0\alpha + l_0 s; \hat{E}, \varepsilon) = E_m$$

where
$$y = \frac{\mathcal{B}_m}{k_0^m} - \tilde{J}_m(\hat{E}, \varepsilon) = \frac{\mathcal{B}_m - \mathcal{B}_m^*(\hat{E})}{k_0^m} + O(\varepsilon).$$

The value of the critical value E_m^* where the topology of the tori changes is given by
$$E_m^* = \varepsilon^j U^{k_0, l_0, *}(\tilde{\theta}_m(\hat{E}, \varepsilon); \hat{E}, \varepsilon),$$

which is just the value of the potential at the saddle point. Again, using the changes (50), (66) and formula (73) of Lemma 21, we can obtain explicit formulae for these tori:

$$\begin{aligned}
\hat{\mathcal{B}} &= \hat{E} + \frac{\mathcal{B}_m}{k_0^m}\hat{k}_0 \\
\mathcal{B}_m &= k_0^m \tilde{J}_m(\hat{E}, \varepsilon) \pm k_0^m \mathcal{Y}(k_0\alpha + l_0 s; E, \varepsilon) \\
&= k_0^m J_m^*(\hat{E}) \pm k_0^m \mathcal{Y}(k_0\alpha + l_0 s; E, \varepsilon) + O(\varepsilon) \\
&= \mathcal{B}_m^*(\hat{E}) \pm k_0^m \mathcal{Y}(k_0\alpha + l_0 s; E, \varepsilon) + O(\varepsilon)
\end{aligned}$$

where
$$\mathcal{Y}(\theta_m; E, \varepsilon) = \tilde{\mathcal{Y}}(\theta_m - \tilde{\theta}_m(\hat{E}, \varepsilon); E; \varepsilon).$$

Going back to the original variables $(I, \varphi, s) \in \mathbb{R}^d \times \mathbb{T}^{d+1}$ we can write the implicit equations for these tori $\mathcal{T}_E^0$ as:

$$\hat{I} - \frac{I_m}{k_0^m}\hat{k}_0 + O(\varepsilon) = \hat{E},$$

$$\frac{I_m}{k_0^m}l_0 + h(I) + \varepsilon K^{0,0}(I; \varepsilon) + \varepsilon^j U^{k_0, l_0, *}(k_0\varphi + l_0 s; \hat{E}, \varepsilon) + O(\varepsilon) = \tilde{E}_m$$

That can also be also written as:

(79)
$$\hat{I} - \frac{I_m}{k_0^m}\hat{k}_0 + O(\varepsilon) = \hat{E},$$

$$a(\hat{E}, \varepsilon)\frac{y^2}{2}(1 + O(y)) + \varepsilon^j U^{k_0, l_0, *}(k_0\varphi + l_0 s; \hat{E}, \varepsilon) + O(\varepsilon^{j+1}) = E_m$$

where:

$$(80) \qquad y = \frac{I_m - \mathcal{B}_m^*(\hat{E})}{k_0^m} + O(\varepsilon), \quad E_m^* = \varepsilon^j U^{k_0, l_0, *}(\tilde{\theta}_m(\hat{E}, \varepsilon); \hat{E}, \varepsilon).$$

Finally, these tori can be also written explicitly as:

$$(81) \qquad \hat{I} = \hat{E} + \frac{I_m}{k_0^m} \hat{k}_0 + O(\varepsilon)$$

$$I_m = \mathcal{B}_m^*(\hat{E}) \pm k_0^m \mathcal{Y}(k_0 \varphi + l_0 s; E, \varepsilon) + O(\varepsilon)$$

Let us observe that the function $\mathcal{Y}$ verifies, according to Lemma 21:
- If $0 < |E_m - E_m^*| \leq \varepsilon^j$:

$$(82) \qquad \mathcal{Y}(\theta_m; E, \varepsilon) = \ell(\theta_m; E, \varepsilon)(1 + O(\varepsilon^{\frac{j}{2}}))$$

- If $\varepsilon^j < |E_m - E_m^*| < 1$, writing $|E_m - E_m^*| = d_m \varepsilon^\gamma$, with $0 < \gamma < j$:

$$(83) \qquad \mathcal{Y}(\theta_m; E, \varepsilon) = \ell(\theta_m; E, \varepsilon)(1 + O(\varepsilon^{\frac{\gamma}{2}}))$$

- If $|E_m - E_m^*| = O(1)$:

$$(84) \qquad \mathcal{Y}(\theta_m; E, \varepsilon) = h_0^{(-1)}(E_m)(1 + O(\varepsilon^j))$$

where the function ℓ is given by (see (68):

$$(85) \qquad \ell(\theta_m; E, \varepsilon) = \sqrt{\frac{2}{a(\hat{E}, \varepsilon)}(E_m - \varepsilon^j U^{k_0, l_0, *}(\theta_m; \hat{E}, \varepsilon))}.$$

3.5.4. *Primary and secondary tori near the secular resonances: KAM Theorem.* If we apply the changes of variables (48), (50) and (66) to Hamiltonian (40) we obtain:

$$(86) \qquad \mathcal{K}(\hat{J}, \hat{\theta}, y, x, s; \varepsilon) = \mathcal{K}_0(\hat{J}, y, x; \varepsilon) + \varepsilon^{m+1} S(\hat{J}, \hat{\theta}, y, x, s; \varepsilon).$$

First, we change to action–angle variables of the integrable part $\mathcal{K}_0$. The two only difficulties are that the action angle variables become singular near the level set $\mathcal{K}_0 = 0$, which is the separatrix of the torus $\{(x, y) = (0, 0)\}$, and also the twist condition becomes singular. In fact, the twist goes to ∞ when one approaches the separatrix and this is favorable to application of the KAM theorem [Her83, DLS06a]) because this theorem only requires lower bounds on the twist and a larger twist improves the quantitative assumptions of the theorem.

If the number of averaging steps m is large enough we can ensure that there exist KAM tori (both primary and secondary) that cover the whole resonant region up to distances $O(\varepsilon^{3/2})$ and which are ε^3 close to the level sets $\mathcal{K}_0 = c_m$ of the averaged Hamiltonian $\mathcal{K}_0$. Of course, we could get even higher powers in the density by averaging more times.

- We select a region $|c_m - c_m^*| \leq \varepsilon^\alpha$ surrounding the separatrix $\mathcal{K}_0 = c_m^* = 0$. In this region (the chaotic zone) we will not perform any further analysis. We will just remark that it is small. In particular, if $\alpha = \frac{3}{2} + \frac{j}{2}$, the level sets of energy c_m and c_m^* are at a distance $\varepsilon^{3/2}$ if $(c_m - c_m^*) = \pm\varepsilon^\alpha$.
- In the complementary region: $|c_m - c_m^*| \geq \varepsilon^\alpha$ we change to action angle variables adapted to the level sets, $F = c = (\hat{c}, c_m)$, of the function $F = (\hat{J}, \mathcal{K}_0)$ defined in (72). We note that one of the components of F is precisely $\mathcal{K}_0$, the integrable part of the averaged Hamiltonian (86). The action variables can be obtained geometrically integrating the canonical form over the loops in a torus [Arn78, AM78].

  It is well known that the singularities of the action variable are only a power of $(c_m - c_m^*)$. Therefore, since the size of the remainder in Hamiltonian (86) is $O(\varepsilon^{m+1})$, when expressed in the action angle variables in the region $|c_m - c_m^*| \geq \varepsilon^\alpha$ the smallness of the remainder will be $O(\varepsilon^{m+1-A\alpha})$, for some value $A > 0$.
- The KAM theorem in action-angle variables [Pös01] gives tori which are at a distance $\varepsilon^{m+1-A\alpha}$ of the level sets of the action variables. The gaps between these tori are $O(\varepsilon^{(m+1-A\alpha)/2})$.
- Coming back to variables $(y, x, \hat{J}, \hat{\theta})$ we obtain tori at a distance between them of order $O(\varepsilon^{(m+1-A\alpha)/2-\alpha A})$, and that are at a distance $O(\varepsilon^{m+1-2A\alpha})$ from the level sets of $(\hat{J}, \mathcal{K}_0)$.

  We note that, if we fix $\alpha = \frac{3}{2} + \frac{j}{2}$, taking into account that A is a fixed number, we obtain that, taking m large enough we can ensure that we have tori for $|c - c'| < \varepsilon^{\frac{3}{2} + \frac{j}{2}}$, that is, the gaps are smaller than $\varepsilon^{3/2}$ as claimed.
- Going back to the original variables through changes (53), (66) (which are close to the identity), we obtain the result in next theorem 23.

Theorem 23. *Under the conditions of theorem 5, there exists ε_0 such that, for $0 < |\varepsilon| < \varepsilon_0$, the secular resonant region $\mathcal{S}^{[\leq 2]}$ can be covered by $O(\varepsilon^{3/2})$ neighborhoods of invariant objects under the Hamiltonian flow of the Hamiltonian $K_\varepsilon(I, \varphi, s)$ in (16). Moreover:*

- *These invariant objects are given by the level sets $F = E = (\hat{E}, E_m)$, for $|E - E'| \leq \varepsilon^{\frac{3}{2} + \frac{j}{2}}$ and where $\varepsilon^{\frac{3}{2} + \frac{j}{2}} \leq |E_m - E_m^*| \leq 1$ with E_m^* given in (80).*
- *The $\mathcal{C}^2$ function $F : \mathbb{R}^d \times \mathbb{T}^d \times \mathbb{T} \to \mathbb{R}$ is given by (79).*

- *These invariant objects are either regular primary KAM $d + 1$-tori, secondary $d+1$-KAM tori of class 1 (i.e. $d+1$-dimensional invariant tori which are contractible to a $\mathbb{T}^d$) or invariant manifolds of d-dimensional whiskered invariant tori.*

3.6. Second step: The generation of a homoclinic manifold and computation of the scattering map.

Let us observe that by hypothesis **H2**, for $\varepsilon = 0$, the manifold $\tilde{\Lambda}_0$ has stable and unstable manifolds which coincide along a homoclinic manifold

$$\tilde{\Gamma}_0 = W^s(\tilde{\Lambda}_0) = W^u(\tilde{\Lambda}_0)$$

with

$$\tilde{\Gamma}_0 = \{(p^*(\tau), q^*(\tau), I, \varphi, s), \ (I, \varphi, s, \tau) \in \mathcal{I}^* \times \mathbb{T}^{d+1} \times \mathbb{R}^n\}$$

The first result in this section is that, if system (3) satisfies the non-degeneracy assumption **H7**, then for all $0 < |\varepsilon| < \varepsilon_0$, $W^s(\tilde{\Lambda}_\varepsilon)$, $W^u(\tilde{\Lambda}_\varepsilon)$, the stable and unstable manifolds of the normally hyperbolic invariant manifold $\tilde{\Lambda}_\varepsilon$ introduced in Section 3.1, have a transversal intersection along a homoclinic manifold $\tilde{\Gamma}_\varepsilon$. Then, following [DLS08] we use this intersection to define the scattering map in $H_- \subset \tilde{\Lambda}_\varepsilon$.

We will use a notation very similar to that of [DLS06a] and, indeed refer to this paper for a series of detailed calculations. The proof of the next proposition is identical to proposition 9.2 in [DLS06a].

Proposition 24. *Assume that hypothesis **H7** is fulfilled. Then, given $(I, \varphi, s) \in H_- \subset \mathcal{I}^* \times \mathbb{T}^{d+1}$, for ε small enough, there exists a locally unique point $\tilde{z}^*$ of the form*
(87)
$$\tilde{z}^*(I, \varphi, s; \varepsilon) = (p^*(\tau^*(I, \varphi, s)) + \mathrm{O}(\varepsilon), q^*(\tau^*(I, \varphi, s)) + \mathrm{O}(\varepsilon), I, \varphi, s)$$

such that $W^s(\tilde{\Lambda}_\varepsilon) \pitchfork W^u(\tilde{\Lambda}_\varepsilon)$ at $\tilde{z}^*$, that is,

$$\tilde{z}^* \in W^s(\tilde{\Lambda}_\varepsilon) \cap W^u(\tilde{\Lambda}_\varepsilon) \ \text{and} \ T_{\tilde{z}^*}W^s(\tilde{\Lambda}_\varepsilon) + T_{\tilde{z}^*}W^u(\tilde{\Lambda}_\varepsilon) = T_{\tilde{z}^*}\tilde{\mathcal{M}},$$

where $\tilde{\mathcal{M}} = \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{I} \times \mathbb{T}^d \times \mathbb{T}$.

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

(88)
$$\left| \tilde{\Phi}_{\varepsilon,t}(\tilde{z}^*) - \tilde{\Phi}_{\varepsilon,t}(\tilde{x}_\pm) \right| \leq \text{cte.} \, e^{-\alpha|t|/2} \ \text{for} \ t \to \pm\infty.$$

Moreover, expressing the points $\tilde{x}_\pm = k_\varepsilon(I_\pm, \varphi_\pm, s_\pm; \varepsilon)$ in terms of the symplectic parametrization of $\tilde{\Lambda}_\varepsilon$ introduced in section 3.1, 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

(89) $$I(\tilde{x}_+) - I(\tilde{x}_-) = \varepsilon \frac{\partial \mathcal{L}^*}{\partial \varphi}(I, \varphi - \omega(I)s) + O_{\mathcal{C}^1}(\varepsilon^{1+\varrho}),$$

where $\mathcal{L}^*(I, \theta)$ is given in (12), and $\varrho > 0$.

3.7. **The scattering map.** From now on we will take the homoclinic manifold

$$\tilde{\Gamma}_\varepsilon = \{\tilde{z}^*(I, \varphi, s; \varepsilon), \ (I, \varphi, s) \in H_- \subset \mathcal{I}^* \times \mathbb{T}^{d+1}\}$$

given by proposition 24. Following [DLS08] we will call $\tilde{\Gamma}_\varepsilon$ an homoclinic channel and we can define the scattering map s_ε on $\tilde{\Lambda}_\varepsilon$, also called outer map, associated to $\tilde{\Gamma}_\varepsilon$.

Following [DLS06a, DLS08] the scattering map is defined as follows: for any two points $\tilde{x}_\pm \in \tilde{\Lambda}_\varepsilon$, we say that $\tilde{x}_+ = s_\varepsilon(\tilde{x}_-)$, if there exists a point $\tilde{z} \in \tilde{\Gamma}_\varepsilon$ such that

$$\text{dist}(\tilde{\Phi}_{\varepsilon,t}(z), \tilde{\Phi}_{\varepsilon,t}(x_\pm)) \to 0, \quad \text{for } t \to \pm\infty.$$

Since the unperturbed system is a product system, it is clear that, independently of what is the homoclinic manifold, the stable manifold of one point in $\tilde{\Lambda}_0$ is the same as its unstable manifold. Therefore, $s_0 = \text{Id}$.

As shown in [DLS08] the scattering map is an exact symplectic map and depends smoothly on parameters because the homoclinic manifold depends smoothly on parameters even through $\varepsilon = 0$.

It is well known (see [LMM86]) that a family of exact symplectic mappings s_ε is conveniently described using a generator $\mathcal{S}_\varepsilon$ and the associated Hamiltonian S_ε:

$$\frac{d}{d\varepsilon} s_\varepsilon = \mathcal{S}_\varepsilon \circ s_\varepsilon; \qquad \imath_{\mathcal{S}_\varepsilon} \omega = dS_\varepsilon.$$

Indeed in [DLS08] it is shown that the Hamiltonian S_ε is given, up to first order in ε by the function $-\mathcal{L}^*(I, \varphi - \omega(I)s)$, where the reduced Melnikov potential $\mathcal{L}^*(I, \theta)$ is given in (12):

$$S_\varepsilon = -\mathcal{L}^* + \varepsilon S_1 + O(\varepsilon^2)$$

Therefore, the scattering map can be written, using the coordinates (I, φ, s) as:

(90) $$\begin{aligned} s_\varepsilon(I, \varphi, s) &= (I + \varepsilon \partial_\theta \mathcal{L}^*(I, \varphi - \omega(I)s) + O(\varepsilon^2), \\ &\quad \varphi - \varepsilon \partial_I \mathcal{L}^*(I, \varphi - \omega(I)s) + O(\varepsilon^2), s) \end{aligned}$$

and, for any fixed $s \in \mathbb{T}$, up the first order in ε it is given, in the coordinates (I, φ), as the time $-\varepsilon$ map of the Hamiltonian flow of Hamiltonian $\mathcal{L}^*(I, \theta)$ evaluated at $(I, \varphi - \omega(I)s)$.

The fundamental property to have instability will be to check, for any fixed s, that the tori invariant for the inner flow in $\tilde{\Lambda}_\varepsilon$ are not invariant by the perturbed scattering map s_ε. Therefore, we will pay attention at how the scattering map moves the tori $\mathcal{T}_E$ given in (44), (79), (80) using the results in [DH09].

3.8. Interaction between the inner flow and the scattering map and hypothesis H6.
We have already shown in theorems 18 and 23, that the KAM tori $\mathcal{T}_E$ (both primary and secondary) are the level sets of an $\mathbb{R}^d$-valued function F_ε. Indeed we have approximate expressions for it in (44), (79), (80) (see also (81)).

The scattering map transports the level sets of F_ε into other manifolds, which are the level sets of $F_\varepsilon \circ s_\varepsilon^{-1}$.

The key observation relies on Lemma 10.4 in [DLS06a] (see also [DLS00]), which states that, given two invariant manifolds for the inner flow $\Sigma_i \subset \Lambda$, $i = 1, 2$, if Σ_1 intersects transversally $s_\varepsilon(\Sigma_2)$ in $\tilde{\Lambda}_\varepsilon$, then $W^u_{\Sigma_2} \pitchfork W^s_{\Sigma_1}$.

Our next goal will be to make explicit the conditions to ensure that the scattering map creates heteroclinic intersections between the KAM tori, primary or secondary, created in sections 3.5.2 and 3.5.4.

Fix $s \in \mathbb{T}$. The tori $\mathcal{T}_{E'}$ and $s_\varepsilon(\mathcal{T}_E)$ intersect if there exists a point $\tilde{x} = (I, \varphi, s)$ such that:

$$
\begin{aligned}
F_\varepsilon(I, \varphi, s; \varepsilon) &= E \\
F_\varepsilon \circ s_\varepsilon(I, \varphi, s; \varepsilon) &= E'.
\end{aligned}
\tag{91}
$$

Let us observe that the first equation is the implicit equation for the torus $\mathcal{T}_E$. Instead, we can use its explicit equation $I = \lambda_E(\varphi, s; \varepsilon)$ to eliminate the first d equations. The manifolds $\mathcal{T}_{E'}$ and $s_\varepsilon(\mathcal{T}_E)$ intersect if there exists (φ, s) such that:

$$
F_\varepsilon \circ s_\varepsilon(\lambda_E(\varphi, s; \varepsilon), \varphi, s; \varepsilon) = E'
\tag{92}
$$

and the intersection will be transversal if

$$
\det D(F_\varepsilon \circ s_\varepsilon(\lambda_E(\varphi, s; \varepsilon), \varphi, s; \varepsilon)) \neq 0,
\tag{93}
$$

where $D = D_\varphi$.

Using formula (159) in [DH09], we know that, given a function F:

$$
F \circ s_\varepsilon = F - \varepsilon\{F, \mathcal{L}^*\} + \frac{\varepsilon^2}{2}(\{\{F, \mathcal{L}^*\}, \mathcal{L}^*\} + \{F, S_1\}) + O(\varepsilon^3).
\tag{94}
$$

Therefore equation (92) reads:

$$
-\{F_\varepsilon, \mathcal{L}^*\}(\lambda_E(\varphi, s; \varepsilon), \varphi, s; \varepsilon) + O(\varepsilon) = \frac{E' - E}{\varepsilon}
\tag{95}
$$

and we will have intersection as long as $\frac{E'-E}{\varepsilon}$ is small enough, close to the non-degenerate zeros of

$$\{F_\varepsilon, \mathcal{L}^*\}(\lambda_E(\varphi, s; \varepsilon), \varphi, s; \varepsilon) = 0$$

therefore the transversality condition (93) is equivalent to

$$(96) \qquad D(\{F_\varepsilon, \mathcal{L}^*\})(\lambda_E(\varphi, s; \varepsilon), \varphi, s; \varepsilon) \neq 0$$

for a point in each of the level sets of F_ε.

3.8.1. *The non-degeneracy condition* **H8** *in the non-resonant region.* The non-resonant region $\mathcal{S}^L$ (see (42)) is of $O(1)$ and is covered by $\varepsilon^{\frac{3}{2}}$ neighborhoods of tori which are given by the level sets of the function:

$$F_\varepsilon(I, \varphi, s) = I + O(\varepsilon) = E.$$

Therefore $\lambda_E(\varphi, s; \varepsilon) = E + O(\varepsilon)$, and equation (95) reads:

$$(97) \qquad \frac{\partial \mathcal{L}^*}{\partial \theta}(E, \varphi - \omega(E)s) + O(\varepsilon) = \frac{E' - E}{\varepsilon}$$

Moreover, by the KAM theorem 18 given in section 3.5.2 we have tori for $|E - E'| \leq c\varepsilon^{3/2}$, and therefore equation (97) has solutions for φ, which are non-degenerate if condition (96) is verified, which in our case, becomes:

$$\det \left| \frac{\partial^2 \mathcal{L}^*}{\partial \theta^2}(E, \varphi - \omega(E)s) \right| \neq 0,$$

and is guaranteed if

$$(98) \qquad \det \left| \frac{\partial^2 \mathcal{L}^*}{\partial \theta^2}(I, \varphi - \omega(I)s) \right| \neq 0$$

is satisfied for $(I, \varphi, s) \in H_- \subset \mathcal{I}^* \times \mathbb{T}^{d+1}$, and is one of the non-degeneracy conditions included in Hypothesis **H8**.

3.8.2. *Heteroclinic orbits close to homoclinic ones in the non-resonant region.* If condition (98) is verified in the region $\mathcal{S}^L$ we can guarantee the existence of heteroclinic connections between neighboring KAM tori in this region. If we look for heteroclinic connections close to homoclinic ones, one can obtain a more explicit sufficient condition for equations (97) to have a solution. The main idea is to solve equations (97), using the implicit function theorem. The small parameter will be

$$\delta = \frac{E' - E}{\varepsilon} + \varepsilon$$

and then equation (97) read:

$$\frac{\partial \mathcal{L}^*}{\partial \theta}(E, \varphi - \omega(E)s) = O(\delta)$$

Therefore, a non-degeneracy condition which guarantees that equation (97) have solutions close to the solutions of:

$$\frac{\partial \mathcal{L}^*}{\partial \theta}(E, \varphi - \omega(E)s) = 0$$

is that the function $\mathcal{L}^*$ has non-degenerate critical points, that is:

$$(99) \quad \frac{\partial \mathcal{L}^*}{\partial \theta}(E, \varphi - \omega(E)s) = 0, \implies \quad \det \left| \frac{\partial^2 \mathcal{L}^*}{\partial \theta^2}(E, \varphi - \omega(E)s) \right| \neq 0,$$

in the region $\mathcal{S}^L \cap H_-$. Equation (99) is part of Hypothesis **H8**.

3.8.3. *The non-degeneracy condition* **H8** *in the resonant region.* Now, we study the intersection equation (91) in the secular resonant region (45) $\mathcal{S}^{[\leq 2]} \cap H_-$, to ensure that the image under the scattering map of a primary or secondary torus intersects other nearby tori. We denote by F_ε again the function whose level sets gives the tori. We recall that the secular resonant region $\mathcal{S}^{[\leq 2]}$ is the union of the tubular neighborhoods $\mathcal{R}^L_{k,l}$ of the secular resonances $\mathcal{R}_{k,l}$, for $(k,l) \in \mathcal{N}^{[\leq 2]}$.

If $\mathcal{R}_{k_0,l_0}$ is a resonance of order j, $j = 1, 2$, in the region $\mathcal{R}^L_{k_0,l_0}$, according to (79), (80) and the KAM theorem 23, the invariant tori are given by the level sets of a function:

$$F_\varepsilon = (\hat{F}, F_m) = E = (\hat{E}, E_m),$$

for $\varepsilon^{\frac{3}{2}+\frac{j}{2}} \leq |E_m - E_m^*| \leq 1$, where $E_m^* = \varepsilon^j U^{k_0,l_0,*}(\tilde{\theta}_m(\hat{E}, \varepsilon); \hat{E}, \varepsilon)$ (see [DLS06a]), with $U^{k_0,l_0,*}$ given in (56),

$$(100) \qquad \hat{F}(I, \varphi, s; \varepsilon) = \hat{I} - \frac{I_m}{k_0^m}\hat{k}_0 + O(\varepsilon),$$

and
(101)

$$F_m(I, \varphi, s; \varepsilon) = a(\hat{E}, \varepsilon)\frac{y^2}{2}(1+O(y)) + \varepsilon^j U^{k_0,l_0,*}(k_0\varphi + l_0 s; \hat{E}, \varepsilon) + O(\varepsilon^{j+1})$$

with $a(\hat{E}, \varepsilon)$ given in (69), (60), and

$$y = \frac{I_m - \mathcal{B}_m^*(\hat{E})}{k_0^m} + O(\varepsilon)$$

where $\mathcal{B}_m^*(\hat{E}) = \Gamma_{k_0,l_0}(I)$ is the k_0-projection in the resonance $\mathcal{R}_{k_0,l_0}$.

Moreover, by the KAM theorem 23, we know that there exist tori $F_\varepsilon = E$, $F_\varepsilon = E'$ for $|E - E'| \leq \varepsilon^{\frac{3}{2}+\frac{j}{2}}$.

The way to solve equation (91) is slightly different for a resonance or order one or for a resonance of order two. We give all the details in the

case of a first order resonance. The case of a resonance of order two can be done with minor modifications, as it is explained in Remark 25.

For $F_\varepsilon = (\hat{F}, F_m)$, the set of equations (91) reads

$$\hat{F}(I, \varphi, s; \varepsilon) = \hat{E}$$
$$F_m(I, \varphi, s; \varepsilon) = E_m$$
$$\hat{F}(I, \varphi, s; \varepsilon) - \varepsilon\{\hat{F}, \mathcal{L}^*\}(I, \varphi, s; \varepsilon) + O(\varepsilon^2) = \hat{E}'$$
$$F_m(I, \varphi, s; \varepsilon) - \varepsilon\{F_m, \mathcal{L}^*\}(I, \varphi, s; \varepsilon) + O(\varepsilon^2) = E_m'$$

and, using (100), (101), these equations are equivalent to

(102)
$$\begin{aligned}
\hat{F}(I, \varphi, s; \varepsilon) &= \hat{E} \\
F_m(I, \varphi, s; \varepsilon) &= E_m \\
\partial_{\hat{\theta}}\mathcal{L}^*(I, \varphi - \omega(I)s) - \frac{1}{k_0^m}\partial_{\theta_m}\mathcal{L}^*(I, \varphi - \omega(I)s)\hat{k}_0 & \\
+O(\varepsilon) &= \frac{\hat{E}' - \hat{E}}{\varepsilon} \\
a(\hat{E}, \varepsilon)\frac{I_m - \mathcal{B}_m^*(\hat{E})}{(k_0^m)^2}\partial_{\theta_m}\mathcal{L}^*(I, \varphi - \omega(I)s) & \\
+O(\varepsilon) &= \frac{E_m' - E_m}{\varepsilon}.
\end{aligned}$$

From the first two equations we obtain, using (81) for $E_m - E_m^* = O(\varepsilon^\gamma)$, $0 < \gamma \le \frac{3}{2} + \frac{j}{2} = 2$, we obtain:

$$I = (\hat{E}, 0) + \frac{I_m}{k_0^m}k_0 + O(\varepsilon)$$

$$I_m = \mathcal{B}_m^*(\hat{E}) \pm k_0^m \mathcal{Y}(k_0\varphi + l_0 s; E; \varepsilon) + O(\varepsilon)$$

where $E = (\hat{E}, E_m)$ and

(103) $$\mathcal{Y}(k_0\varphi + l_0 s; E; \varepsilon) = \ell(k_0\varphi + l_0 s; E, \varepsilon)(1 + O(\varepsilon^{\frac{\gamma}{2}}))$$

and the function ℓ is given in (85):

$$\ell(k_0\varphi + l_0 s; E; \varepsilon) = \sqrt{\frac{2}{a(\hat{E}, \varepsilon)}(E_m - \varepsilon U^{k_0, l_0, *}(k_0\varphi + l_0 s; \hat{E}, \varepsilon))}$$

which gives, using that $E_m - E_m^* = O(\varepsilon^\gamma)$:

$$I = (\hat{E}, 0) + \frac{\mathcal{B}_m^*(\hat{E})}{k_0^m}k_0 = I^*(\hat{E}) + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}})$$

48

and therefore, the two last equations of (102) read:

$$\partial_{\hat{\theta}}\mathcal{L}^* - \frac{1}{k_0^m}\partial_{\theta_m}\mathcal{L}^*\hat{k}_0 + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}}) = -\frac{\hat{E} - \hat{E}'}{\varepsilon}$$

(104)
$$\pm\frac{a(\hat{E};\varepsilon)}{k_0^m}\left[\mathcal{Y}(k_0\varphi + l_0 s; E; \varepsilon) + O(\varepsilon)\right] \times$$

$$\times\left(\partial_{\theta_m}\mathcal{L}^* + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}})\right) = -\frac{\tilde{E}_m - E'_m}{\varepsilon}$$

where, to shorten the notation, we have just written

$$\mathcal{L}^* = \mathcal{L}^*(I^*(\hat{E}), \varphi - \omega(I^*(\hat{E}))s).$$

Using equation (103) and that, by (69), $a(\hat{E}, \varepsilon) = a(\hat{E}) + O(\varepsilon)$, equations (104) are equivalent to:
(105)

$$\partial_{\hat{\theta}}\mathcal{L}^* - \frac{1}{k_0^m}\partial_{\theta_m}\mathcal{L}^*\hat{k}_0 + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}}) = \frac{\hat{E}' - \hat{E}}{\varepsilon}$$

$$\pm\frac{\left(a(\hat{E}) + O(\varepsilon)\right)}{k_0^m}\left[\ell(k_0\varphi + l_0 s; E; \varepsilon)(1 + O(\varepsilon^{\frac{\gamma}{2}})) + O(\varepsilon)\right]$$

$$\times\left(\partial_{\theta_m}\mathcal{L}^* + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}}))\right) = \frac{\tilde{E}'_m - E_m}{\varepsilon}.$$

We will see that we will have a solution of equations (105) for φ if $|E' - E| \leq O(\varepsilon^{\frac{3}{2}+\frac{j}{2}}) = O(\varepsilon^2)$.

It will be useful to work in the variables $\theta = (\hat{\theta}, \theta_m) = (\hat{\varphi}, k_0\varphi + l_0 s)$. Observe that conversely:

(106)
$$\hat{\varphi} = \hat{\theta}, \quad \varphi_m = \frac{\theta_m - \hat{k}_0\hat{\varphi} - l_0 s}{k_0^m}$$

Now, if we define the auxiliary function:
(107)
$$\mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) = \mathcal{L}^*(I^*(\hat{E}), \varphi - \omega(I^*(\hat{E}))s)$$
$$= \mathcal{L}^*(I^*(\hat{E}), \hat{\varphi} - \hat{\omega}(I^*(\hat{E}))s, \varphi_m - \omega_m(I^*(\hat{E}))s)$$

Using that $\omega(I^*(\hat{E})) \cdot k_0 + l_0 = 0$, we obtain:

$$\mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) =$$

$$\mathcal{L}^*(I^*(\hat{E}), \hat{\theta} - \hat{\omega}(I^*(\hat{E}))s, \frac{\theta_m - (\hat{\theta} - \hat{\omega}(I^*(\hat{E}))s)\hat{k}_0}{k_0^m}),$$

and then, taking derivatives with respect to $\hat{\theta}$ and θ_m:

$$\frac{\partial}{\partial\hat{\theta}}\mathcal{L}^*_{k_0,l_0}(\hat{\varphi}, k_0\varphi + l_0 s, s; \hat{E}) = \frac{\partial}{\partial\hat{\theta}}\mathcal{L}^*(I^*(\hat{E}), \varphi - \omega(I^*(\hat{E}))s)$$

$$- \frac{1}{k_0^m}\frac{\partial}{\partial\theta_m}\mathcal{L}^*(I^*(\hat{E}), \varphi - \omega(I^*(\hat{E}))s)\ \hat{k}_0$$

$$\frac{\partial}{\partial\theta_m}\mathcal{L}^*_{k_0,l_0}(\hat{\varphi}, k_0\varphi + l_0 s, s; \hat{E}) = \frac{1}{k_0^m}\frac{\partial}{\partial\theta_m}\mathcal{L}^*(I^*(\hat{E}), \varphi - \omega(I^*(\hat{E}))s)$$

Therefore equations (105) become:
(108)

$$\frac{\partial}{\partial\hat{\theta}}\mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}}) = \frac{\hat{E}' - \hat{E}}{\varepsilon}$$

$$\pm\frac{(a(\hat{E}) + O(\varepsilon))}{k_0^m}\left[\ell(\theta_m; E; \varepsilon)(1 + O(\varepsilon^{\frac{\gamma}{2}})) + O(\varepsilon)\right] \times$$

$$\times\left(\frac{\partial}{\partial\theta_m}\mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}})\right) = \frac{\tilde{E}'_m - E_m}{\varepsilon}$$

which are the generalization to higher dimensions of the function $\mathcal{M}$ in [DLS06a, page 108].

Before looking for the solutions of these equations, me make a further simplification. First observe that there exist primary and secondary tori close to the separatrix of the averaged Hamiltonian for energies

$$|E_m - E_m^*| > \varepsilon^{\frac{3}{2}+\frac{j}{2}} = O(\varepsilon^2), \quad E_m^* = \varepsilon U^{k_0,l_0,*}(\tilde{\theta}(\hat{E}, \varepsilon); \hat{E}, \varepsilon)),$$

therefore it makes sense to scale $E_m = \varepsilon e_m$ in the function ℓ of equations (108) obtaining:
(109)

$$\frac{\partial}{\partial\hat{\theta}}\mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}}) = \frac{\hat{E}' - \hat{E}}{\varepsilon}$$

$$\pm\frac{(a(\hat{E}) + O(\varepsilon))}{k_0^m}\left[\bar{\ell}(\theta_m; E; \varepsilon)(1 + O(\varepsilon^{\frac{\gamma}{2}}) + O(\varepsilon^{\frac{1}{2}}))\right] \times$$

$$\times\left(\frac{\partial}{\partial\theta_m}\mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}})\right) = \frac{\tilde{E}'_m - E_m}{\varepsilon^{1+\frac{1}{2}}}$$

and the function $\bar{\ell}$ is a scaled version of the one given in (85):

$$\bar{\ell}(\theta_m; \hat{E}, e_m; \varepsilon) = \sqrt{\frac{2}{a(\hat{E}, \varepsilon)}(e_m - U^{k_0,l_0,*}(\theta_m; \hat{E}, \varepsilon))}.$$

The function $\bar{\ell} = O(1)$, but the important observation is that if we take $\rho > 0$ and we exclude a small region around the critical point $\tilde{\theta}_m$,

50

(that is, for $\rho < \theta_m - \tilde{\theta}_m \leq 2\pi - \rho$) $\bar{\ell}$ never vanishes. In fact one has:

(110) $\qquad \bar{\ell}(\theta_m; \hat{E}, e_m; \varepsilon) \geq d > 0, \quad \text{for} \quad \rho < \theta_m - \tilde{\theta}_m \leq 2\pi - \rho.$

To have non-degenerate solutions of equations (109) it suffices to assume:

(111) $\qquad \det D_\theta \left(\begin{array}{c} \frac{\partial}{\partial \hat{\theta}} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; E) \\ \pm \frac{a(\hat{E})}{k_0^m} \bar{\ell}(\theta_m; E; \varepsilon) \frac{\partial}{\partial \theta_m} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; E) \end{array} \right) \neq 0.$

Making explicit the derivatives in (111) and separating in blocks corresponding to $\hat{\theta}$ and θ_m, one obtains:

(112) $\qquad \pm \frac{a(\hat{E})}{k_0^m} \left| \begin{array}{cc} \frac{\partial^2}{\partial \hat{\theta}^2} \mathcal{L}^*_{k_0,l_0} & \frac{\partial^2}{\partial \hat{\theta} \partial \theta_m} \mathcal{L}^*_{k_0,l_0} \\ \frac{\partial}{\partial \hat{\theta}} (\bar{\ell} \frac{\partial}{\partial \theta_m} \mathcal{L}^*_{k_0,l_0}) & \frac{\partial}{\partial \theta_m} (\bar{\ell} \frac{\partial}{\partial \theta_m} \mathcal{L}^*_{k_0,l_0}) \end{array} \right| (\hat{\theta}, \theta_m, s; E, \varepsilon) \neq 0$

which gives:
(113)

$\left| \begin{array}{cc} \frac{\partial^2}{\partial \hat{\theta}^2} \mathcal{L}^*_{k_0,l_0} & \frac{\partial^2}{\partial \hat{\theta} \partial \theta_m} \mathcal{L}^*_{k_0,l_0} \\ \bar{\ell} \frac{\partial^2}{\partial \hat{\theta} \partial \theta_m} \mathcal{L}^*_{k_0,l_0} & \bar{\ell} \frac{\partial^2}{\partial \theta_m^2} \mathcal{L}^*_{k_0,l_0} - \frac{(U^{k_0,l_0,*})'}{a(\hat{E}) \bar{\ell}} \frac{\partial}{\partial \theta_m} \mathcal{L}^*_{k_0,l_0} \end{array} \right| (\hat{\theta}, \theta_m, s; E, \varepsilon) \neq 0,$

which, using that neither $\bar{\ell}$ nor $a(\hat{E})$ vanish, is equivalent to:

$$\left(2(e_m - U^{k_0,l_0,*})[\frac{\partial^2}{\partial \hat{\theta}^2} \mathcal{L}^*_{k_0,l_0} \frac{\partial^2}{\partial \theta_m^2} \mathcal{L}^*_{k_0,l_0} - (\frac{\partial^2}{\partial \hat{\theta} \partial \theta_m} \mathcal{L}^*_{k_0,l_0})^2] - \right.$$

(114) $\qquad \left. \frac{\partial^2}{\partial \hat{\theta}^2} \mathcal{L}^*_{k_0,l_0} \frac{\partial}{\partial \theta_m} \mathcal{L}^*_{k_0,l_0} (U^{k_0,l_0,*})' \right) (\hat{\theta}, \theta_m, s; E, \varepsilon) \neq 0$

This inequality (or (113)) constitutes part of hypothesis **H8**, and is the generalization of the non-degeneracy conditions $H5$' and $H5$" in [DLS06a]. We call attention to the fact that (113) takes a value for $\varepsilon = 0$.

An equivalent formulation for this non-degeneracy conditions can be written using the symplectic structure of the system.

Introducing the poisson brackets:

$$\{\hat{F}, \cdot\} = \frac{\partial}{\partial \hat{\theta}}$$

$$\{F_m, \cdot\} = \bar{\ell} \frac{\partial}{\partial \theta_m}$$

we see that equations (108) read:

$$\{\hat{F}, \mathcal{L}^*_{k_0,l_0}\} + O(\varepsilon, \varepsilon^{\frac{\gamma}{2}}) = \frac{\hat{E} - \hat{E}'}{\varepsilon}$$

$$\{F_m, \mathcal{L}^*_{k_0,l_0}\} + O(\varepsilon^{\frac{1}{2}}, \varepsilon^{\frac{\gamma}{2}}) = \frac{\hat{E}_m - \hat{E}'_m}{\varepsilon}$$

and the non-degeneracy condition (113) becomes:

$$\left| \begin{array}{cc} \{\hat{F}, \{\hat{F}, \mathcal{L}^*_{k_0,l_0}\}\} & \{F_m, \{\hat{F}, \mathcal{L}^*_{k_0,l_0}\}\} \\ \{\hat{F}, \{\{F_m, \mathcal{L}^*_{k_0,l_0}\}\} & \{\{F_m, \{\{F_m, \mathcal{L}^*_{k_0,l_0}\}\} \end{array} \right| (\hat{\theta}, \theta_m, s; E, \varepsilon) \neq 0.$$

3.8.4. *Heteroclinic connections between primary tori and secondary tori close to homoclinic connections.* If condition (114) is verified in the region $\mathcal{R}^{[\leq L]}_{k_0,l_0}$ we can guarantee the existence of heteroclinic connections between the primary and secondary tori in this region. If we look for this heteroclinic connections close to homoclinic ones, one can obtain a more explicit sufficient condition to have a solution of equations (109).

The main idea is to solve equations (109) using the implicit function theorem. The small parameters will be

$$\hat{\delta} = \frac{\hat{E} - \hat{E}'}{\varepsilon} + O(\varepsilon, \varepsilon^{\gamma/2}), \quad \delta_m = \frac{E_m - E'_m}{\varepsilon^{\frac{1}{2}+1}} + O(\varepsilon^{\frac{1}{2}}, \varepsilon^{\gamma/2}),$$

and then equations (109) read:

$$\frac{\partial}{\partial\hat{\theta}} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) = O(\hat{\delta})$$

$$\frac{a(\hat{E})}{k_0^m} \bar{\ell}(\theta_m; E; \varepsilon) \frac{\partial}{\partial\theta_m} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) = O(\delta_m)$$

and, using that, by (110), the function $\bar{\ell}$ never vanishes neither does $a(\hat{E})$, they are equivalent to:

(115) $$\frac{\partial}{\partial\hat{\theta}} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) = O(\hat{\delta})$$

(116) $$\frac{\partial}{\partial\theta_m} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) = O(\delta_m)$$

Therefore, the non-degeneracy condition which guarantees that these equations have solutions close to the solutions of:

$$\frac{\partial}{\partial\hat{\theta}} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) = 0$$

(117) $$\frac{\partial}{\partial\theta_m} \mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m, s; \hat{E}) = 0$$

is simply:

$$
(118) \qquad \det \left| \frac{\partial^2 \mathcal{L}^*_{k_0,l_0}}{\partial \theta^2}(\theta, s; \hat{E}) \right| \neq 0.
$$

holding in the region $\mathcal{R}^{[\leq L]}_{k_0,l_0}$. This is part of the non-degeneracy conditions which constitute Hypothesis **H8**.

In summarizing, the Hypothesis **H8** consists in assuming inequalities (98), (99), (113), (118).

It is important to note that the function $\mathcal{L}^*_{k_0,l_0}(\hat{\theta}, \theta_m)$ is the Poincaré function $\mathcal{L}^*(I^*(\hat{E}), \varphi - \omega(I^*(\hat{E}))s)$ after the linear change of variables (106). Therefore condition (118) is equivalent to condition (99) that ensures that the Poincaré function has non-degenerate critical points.

Any of the non-degeneracy simplified conditions (117), (118), or equivalently (99), constitute Hypothesis **H8'** stated after Theorem 5, since they are sufficient conditions to ensure that the surface $\mathcal{T}'_E$ intersects transversally $s_\varepsilon(\mathcal{T}_E)$ for $|E - E'| = O(\varepsilon^{\frac{3}{2}+\frac{1}{2}})$.

Remark 25. In the case of a second order resonance, one needs to take into account the terms of order ε^2 in equation (94). Nevertheless, if one looks for heteroclinic solutions close to homoclinic ones some easy computations show that these heteroclinic connections exist if equations (117) have non-degenerate zeros, and this is also guaranteed by condition (118).

3.9. Constructing chains of invariant tori. Contouring the resonances of higher multiplicity. Formulation of the symbolic dynamics.

In this section, we will see how to put together the information we have gathered on the scattering map and the KAM tori, and show that we can construct largely arbitrary motions in action space. In particular, we can go around double resonances and other effects of codimension 2.

We will prove the following result which clearly implies Theorem 5 since it has the same hypothesis and clearly stronger conclusions.

Theorem 26. *Let H_ε be a family of the form* (1). *Assume that H_ε satisfies all the hypothesis* **H1**–**H8**. *In particular, it is C^r for $r \geq r_0$.*

Let m_0 be a sufficiently large number. Fix $\delta > 0$ sufficiently small and consider the set $\mathcal{I}_\delta \subset \mathcal{I}^* \subset \mathcal{I}$ defined before theorem 17 to verify condition L2. Then, there exists $\varepsilon_0 > 0$ such that for all $|\varepsilon| \leq \varepsilon_0$, given any C^1 path $\gamma : [0,1] \to \mathcal{I}_\delta$ in $\mathcal{I}_\delta$ there exists $x_\varepsilon(t)$ a trajectory of the flow generated by H and a time reparameterization Ψ_ε (i.e. a diffeomorphism $\Psi_\varepsilon : \mathbb{R}^+ \to [0,1]$) in such a way that

$$
(119) \qquad |I(x_\varepsilon(t)) - \gamma(\Psi_\varepsilon(t))| \leq C\varepsilon^{1/2}
$$

Of course, Theorem 26 immediately implies Theorem 5. Clearly, the hypothesis of both theorems are the same and, for δ sufficiently small so that given any two points $I_-, I_+ \in \mathcal{I}^*$, we can get a path contained in $\mathcal{I}_\delta$ which starts at a distance less than δ from the I_- and ends at a distance less than δ from I_+. Applying Theorem 26 to this path we obtain the statement of Theorem 5 for $\delta + C\varepsilon_0^{1/2}$.

As a corollary of the proof of Theorem 26, we obtain that it is possible to construct orbits that are δ dense on invariant manifold Λ_ε for ε small enough. These orbits also include excursions on the stable and unstable manifolds. So that they are dense in a larger domain. Some constructions of models with orbits dense on submanifolds appear also in [FM03].

Remark 27. Note that we do not prescribe first the path and then state conditions on the perturbations. We have identified conditions on the Hamiltonian that give the simultaneous existence of trajectories that follow any path in $\mathcal{I}_\delta$.

Remark 28. As we will see, the estimate in (119), is rather pessimistic for most of the paths. Indeed, except when the path is close to the resonant region we can have a bound $C\varepsilon$ in (119).

3.9.1. *Proof of Theorem 26.* The proof of Theorem 26 will consist in recalling all the information that we have been gathering to construct a transition chain of whiskered tori that follows the indicated path. Then, it will suffice to invoke an obstruction argument that establishes that given

a transition chain of whiskered tori (i.e. a sequence of whiskered tori $\mathcal{T}_i$ such that $W^u_{\mathcal{T}_i} \pitchfork W^s_{\mathcal{T}_{i+1}}$), there is an orbit that follows the path.

Recall that we have shown that there is a normally hyperbolic invariant manifold $\tilde{\Lambda}_\varepsilon$.

We have shown that under the nondegeneracy assumptions. **H5**, **H6** we can define a scattering map in the region $\mathcal{I}^*$, which is of a size independent of ε. In this region, we could define the scattering map and give explicit formulas for its leading behavior.

Independently of the scattering map, we have developed averaging theory and obtained information about a geography of the resonances that appear when

averaging. It is important to note that the geography of resonances depends only on the integrable flow. The perturbations activate some of them at the order that we consider.

We recall that the set $\mathcal{I}_\delta$ was obtained by removing from the set $\mathcal{I}^*$ (defined through hypotheses **H3**—**H8**) all the points at a distance less than δ from either of

    a) The set of double resonances activated at order smaller than m_0 one of whose resonances is a secular resonance (i.e. a resonance of order 1,2, see (33) and Definition 13).

    b) The set of points in (8) for which the secular resonance is degenerate.

Note that by assumption **H3** and **H8**, the sets involved in a), b) above are the union of a finite number of codimension 2 manifolds—b) will be empty for quasi-convex Hamiltonians—. Hence, for sufficiently small δ, the set $\mathcal{I}_\delta$ will be connected. Note also that $\mathcal{I}_\delta$ is independent of ε and that has a size of order 1.

Recall that in Section 3.2 we have shown that the region $\mathcal{I}_\delta$ can be covered by a collection of KAM tori which are $\varepsilon^{3/2}$ close to each other (as mentioned in Remark 19, we could have obtained a larger power of ε simply by averaging more times, which requires to remove some more double resonances and assume more derivatives in the model). We will refer to this collection as the *scaffolding* since the motions we construct consist on jumping from one element of the scaffolding to the next by the scattering map and moving along the element for a while.

We have shown that, under the hypothesis **H8**, we have that the image under the scattering map of any of the tori constructed in Section 3.5 intersects transversally all the other tori which are at a distance smaller that a quantity $O(\varepsilon)$.

That is, if $\mathcal{T}$ is an invariant torus in $\tilde{\Lambda}_\varepsilon$ – hence a whiskered torus in the whole phase space – we have $\mathcal{T} \pitchfork_{\tilde{\Lambda}_\varepsilon} \mathcal{T}'$ for all other tori $\mathcal{T}'$ at a distance smaller than $C\varepsilon$.

Given a C^1 path as in the conclusion of Theorem 26, we can find a sequence $\{\mathcal{T}_i\}_{i=0}^\infty$ of tori at a distance $O(\varepsilon)$ from each other and from the path γ (recall that we have shown that these tori are at a distance not more than $O(\varepsilon^{3/2})$.) These tori satisfy

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

By Lemma 10.4 in [DLS00], we obtain that these invariant tori in $\tilde{\Lambda}_\varepsilon$ – hence whiskered tori in the full phase space – satisfy

$$W^u_{\mathcal{T}_i} \pitchfork_{\tilde{\Lambda}_\varepsilon} W^s_{\mathcal{T}_{i+1}}.$$

That is, they constitute a transition chain.

In these circumstances, there are theorems that show that there are orbits that follow the transition whole transition chain. One theorem

particularly well suited for our purposes is that of [FM00] (See also some extensions [FM00, DLS00, DLS06b].) Of course, there are many versions of these results in the literature, but some of them include the extra assumption that the Birkhoff normal form of the tori does not contain some terms, or that the system is C^∞ or that the transition chain is finite. The paper [FM00] does not have any of these limitations and also does not need any assumptions on the topology of the embedding of the torus. It applies just as well to chains in which some of the tori are primary and other that are secondary.

4. AN EXAMPLE

In this section, we present an explicit example where one can check it verifies the conditions **H1** to **H8**. Consider the Hamiltonian:

$$(120) \quad H(I_1, I_2, \varphi_1, \varphi_2, p, q, t, \varepsilon) = \pm \left(\frac{p^2}{2} + \cos q - 1 \right) + h(I_1, I_2)$$
$$+ \varepsilon \cos q \; g(\varphi_1, \varphi_2, t)$$

where

$$h(I_1, I_2) = \Omega_1 \frac{I_1^2}{2} + \Omega_2 \frac{I_2^2}{2},$$

and

$$g(\varphi_1, \varphi_2, t) = a_1 \cos \varphi_1 + a_2 \cos \varphi_2 + a_3 \cos(\varphi_1 + \varphi_2 - t).$$

Proposition 29. *Assume that a_0, a_1, a_2, Ω_1, Ω_2, $\Omega_1 + \Omega_2$, $4\Omega_1 + \Omega_2$ and $\Omega_1 + 4\Omega_2$ are non zero. Then Hamiltonian* (120) *verifies hypotheses* **H1** *to* **H8** *of Theorem 5.*

As we will see, the proof of this proposition is very explicit and we can give a rather precise description of the geometric objects involved in the construction. This proof also shows that there are other heteroclinic connections which could be used to construct unstable orbits. These other choices would lead, through similar calculations, to other regions of parameters where Theorem 5 applies.

Proof. The first observation is that g is a trigonometric polynomial in the angles φ_1, φ_2, t, so it is clear that Hamiltonian (120) satisfies hypotheses **H1** to **H4**. The Hamiltonian of one degree of freedom $P_\pm(p, q) = \pm(p^2/2 + \cos q - 1)$ is the standard pendulum when we choose the $+$ sign, and its separatrix for positive p is given by:

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

An important feature of the Hamiltonian (120) is that the 5-dimensional hyperbolic invariant manifold

$$\tilde{\Lambda} = \{(0, 0, I_1, I_2, \varphi_1, \varphi_2, s) : (I_1, I_2, \varphi_1, \varphi_2, s) \in \mathbb{R}^2 \times \mathbb{T}^3\}$$

is *preserved* for $\varepsilon \neq 0$: $p = q = 0 \Rightarrow \dot{p} = \dot{q} = 0$. However, in contrast with the example in [Arn64], the perturbation does not vanish on $\tilde{\Lambda}$. Indeed, the dynamics on $\tilde{\Lambda}$ is provided simply by the restriction of $H_{|\tilde{\Lambda}}$, which is a 2 and a half degrees of freedom Hamiltonian taking the form

$$h(I_1, I_2) + \varepsilon g(\varphi_1, \varphi_2, t).$$

However, for any $I = (I_1, I_2)$, the 3-dimensional whiskered tori

$$\mathcal{T}_I^0 = \{(0, 0, I, \varphi_1, \varphi_2, s) : (\varphi_1, \varphi_2, s) \in \mathbb{T}^3\}$$

are not preserved if $a_i \neq 0$, and the resonances activated at order one are given by the equations $\omega_i = 0$, $i = 1, 2, 3$, where we introduce the notation

$$\omega_1 = \Omega_1 I_1, \quad \omega_2 = \Omega_2, \quad \omega_3 = \Omega_1 I_1 + \Omega_2 I_2 - 1.$$

Therefore, (120) presents the large gap problem, because it has "large gaps" associated to any of these resonances activated at order one (and also to the resonances activated at order two that will be introduced later on).

The Melnikov potential (9) associated to the Hamiltonian (120) is given by

$$L(\tau, I, \varphi_1, \varphi_2, s) = \frac{1}{2} \int_{-\infty}^{\infty} p_0^2(\tau + \sigma) g(\varphi_1 + \Omega_1 I_1 \sigma, \varphi_2 + \Omega_2 I_2 \sigma, s + \sigma) d\sigma,$$

and computing the integrals by the residue theorem, we obtain

$$L(\tau, I, \varphi_1, \varphi_2, s) = \sum_{i=1}^{3} A_i \cos(\varphi_i - \omega_i \tau)$$

where we introduce $\varphi_3 := \varphi_1 + \varphi_2 - s$, and

$$A_i = A_i(\omega_i) = \frac{2\pi\omega_i}{\sinh(\pi\omega_i/2)} a_i, \qquad i = 1, 2, 3.$$

Since $\tau \in \mathbb{R}$, it can be written as

$$L(\tau, I, \varphi_1, \varphi_2, s) = \mathcal{L}(I, \varphi_1 - \omega_1\tau, \varphi_2 - \omega_2\tau, \varphi_3 - \omega_3\tau)$$

with

$$\mathcal{L}(I, \varphi) = \sum_{i=1}^{3} A_i \cos \varphi_i.$$

57

Therefore

$$\frac{\partial L}{\partial \tau}(\tau, I, \varphi_1, \varphi_2, s) = \sum_{i=1}^{3} \omega_i A_i \cos(\varphi_i - \omega_i \tau)$$

so that, given $(I, \varphi_1, \varphi_2, s)$, the condition

$$\frac{\partial L}{\partial \tau}(\tau, I, \varphi_1, \varphi_2, s) = 0$$

is equivalent to the search of critical points τ^* of the map

(121) $$\tau \in \mathbb{R} \mapsto L(\tau, I, \varphi_1, \varphi_2, s) = \sum_{i=1}^{3} A_i \cos(\varphi_i - \omega_i \tau)$$

that is, of the function $\mathcal{L}$ restricted to the straight line in $\mathbb{T}^3$:

(122) $$\mathbf{R} = \mathbf{R}(I, \varphi) = \{\varphi - \omega \tau, \ \tau \in \mathbb{R}\}.$$

Fixing $I \in \mathbb{R}^2$, the 8 critical points of

$$\varphi \in \mathbb{T}^3 \mapsto \mathcal{L}(I, \varphi)$$

satisfy $\tau^* = 0$, as well as the points (I, φ) in

$$\mathcal{C}(I) = \{\varphi \in \mathbb{T}^3, \sum_{i=1}^{3} \omega_i A_i \cos \varphi_i = 0, \ \sum_{i=1}^{3} \omega_i^2 A_i \sin \varphi_i \neq 0\}.$$

As a consequence, the search for critical points of the map (121) is equivalent to the search for intersections between the straight line $\mathbf{R}(I, \varphi)$ and the set $\mathcal{C}(I)$.

For Hamiltonian (120) the equation of $\mathcal{C}(I)$ is simply

(123) $$\omega_1 A_1 \sin \varphi_1 + \omega_2 A_2 \sin \varphi_2 + \omega_3 A_3 \sin \varphi_3 = 0$$

which is just $D\mathcal{L}(\varphi)\omega = 0$ and defines, locally, the equation of a regular surface in the angles $\varphi = (\varphi_1, \varphi_2, \varphi_3 = \varphi_1 + \varphi_2 - s)$ as long as $\omega\top D^2\mathcal{L}(\varphi)\omega \neq 0$ holds. We notice that for any $I \in \mathbb{R}^2$, the points φ_M^* and φ_m^* where the Melnikov potential $\mathcal{L}$ reaches its maximum and minimum ($\varphi_i^* = 0$ or π) belong to the set $\mathcal{C}(I)$, so there exist at least two zones contained in $\mathcal{C}(I)$ where this set behaves as a local regular surface $\mathcal{C}_M(I)$, $\mathcal{C}_m(I)$, respectively, which will be called *crests* in analogy with the case when φ is two-dimensional (see [DH09]).

Once the set $\mathcal{C}(I)$ is known to be formed at least by the two crests $\mathcal{C}_m(I)$ and $\mathcal{C}_M(I)$, it is clear that, for any φ, there exist several possible intersections of the straight line $\mathbf{R}(I, \varphi)$ given in (122) with the crests $\mathcal{C}_m(I)$ and $\mathcal{C}_M(I)$, parameterized by several values τ^* of the parameters τ which give rise to several scattering maps.

From now on, we will choose *only one* of these intersections, the "first one" with the crest $\mathcal{C}_M(I)$. Given $(I, \varphi_1, \varphi_2, s)$, we define $\tau^*(I, \varphi_1, \varphi_2, s)$ $= \tau_M^*(I, \varphi)$ as the real number τ with minimum absolute value $|\tau|$ among all τ satisfying:

$$\varphi - \omega\tau \in \mathcal{C}_M(I).$$

To determine a domain of definition of τ^* in the variables (I, φ), it suffices to check that the straight line $\mathbf{R}(I, \varphi)$ intersects transversally $\mathcal{C}_M(I)$, that is, that $\omega\top D^2\mathcal{L}(\varphi)\omega \neq 0$ which is exactly the inequality satisfied by $\mathcal{C}(I)$ and a fortiori by $\mathcal{C}_M(I)$.

We can now choose the domain of definition $H_- = H_M$, where τ^* is continuous simply by taking H_- as an appropriate neighborhood of φ_M^*, so that hypothesis **H7** is fulfilled.

Recall that the reduced Poincaré function defined in (12) is

$$\mathcal{L}^*(I, \theta) = L(\tau^*(I, \theta, 0), I, \theta, 0) = \mathcal{L}(I, \theta - \omega\tau^*(I, \theta)).$$

Given (I, θ), $\mathcal{L}^*(I, \theta)$ is the value of $\mathcal{L}$ on $\mathbf{R}(I, \theta) \cap \mathcal{C}_M(I)$ and it is constant along $\mathbf{R}(I, \theta)$ so $\mathcal{L}^*(I, \theta)$ is well defined on $\mathcal{C}_M = \cup_{I \in \mathcal{I}^*}\mathcal{C}_M(I)$.

Recall that the scattering map written in coordinates (I, φ, s) takes the form (90), which, in coordinates $(I, \theta = \varphi - \omega(I)s)$ becomes

$$s_\varepsilon(I, \theta) = (I + \varepsilon\partial_\theta\mathcal{L}^*(I, \theta) + O(\varepsilon^2), \ \theta - \varepsilon\partial_I\mathcal{L}^*(I, \theta) + O(\varepsilon^2), s).$$

We will check the hypotheses in the non-resonant region and in the resonances activated up to order two. There are three resonances activated at order one in this model

$$\begin{aligned}
\mathbf{R}_1 &= \mathcal{R}_{1,0,0} = \{(I_1, I_2), \ I_1 = 0\} \\
\mathbf{R}_2 &= \mathcal{R}_{0,1,0} = \{(I_1, I_2), \ I_2 = 0\} \\
\mathbf{R}_3 &= \mathcal{R}_{1,1,-1} = \{(I_1, I_2), \ \Omega_1 I_1 + \Omega_2 I_2 = 1\}
\end{aligned}$$

and four more activated at order two:

$$\begin{aligned}
\mathbf{R}_4 &= \mathcal{R}_{1,0,-1} = \{(I_1, I_2), \ \Omega_1 I_1 = 1\} \\
\mathbf{R}_5 &= \mathcal{R}_{0,1,-1} = \{(I_1, I_2), \ \Omega_2 I_2 = 1\} \\
\mathbf{R}_6 &= \mathcal{R}_{2,1,-1} = \{(I_1, I_2), \ 2\Omega_1 I_1 + \Omega_2 I_2 = 1\} \\
\mathbf{R}_7 &= \mathcal{R}_{1,2,-1} = \{(I_1, I_2), \ \Omega_1 I_1 + 2\Omega_2 I_2 = 1\}
\end{aligned}$$

For (I_1, I_2) in the non-resonant region, the condition to have heteroclinic orbits between the KAM tori are given by (99). In the resonant regions, one has to check (117) and (118). In our example, one can easily check that both conditions are implied by the conditions

$$D\mathcal{L}(\theta)\omega = 0, \quad \omega^\top D^2\mathcal{L}(\theta)\omega \neq 0$$

defining $\mathcal{C}(I)$, which are a fortiori satisfied by the crest $\mathcal{C}_M(I)$.

To check conditions **H5**, **H6** and **H8** in $\mathbf{R}_i$ we simply need to impose that Ω_1, Ω_2, $\Omega_1 + \Omega_2$, $4\Omega_1 + \Omega_2$ and $\Omega_1 + 4\Omega_2$ are non zero. Moreover, the potential at the resonance $\mathbf{R}_2$ is given by

$$U^1(0, I_2, \varphi_1) = a_1 \cos \varphi_1$$

and therefore hypothesis **H6** is also verified. The study of the potential in the other resonances $\mathbf{R}_i$ is analogous.

Acknowledgements

References

[AKN88]   V. Arnold, V. Kozlov and A. Neishtadt. *Dynamical Systems III*, volume 3 of *Encyclopaedia Math. Sci.* Springer, Berlin, 1988.

[AM78]   R. Abraham and J. E. Marsden. *Foundations of mechanics.* Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978. ISBN 0-8053-0102-X. Second edition, revised and enlarged, With the assistance of Tudor Raţiu and Richard Cushman.

[Arn63]   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.

[Arn64]   V. Arnold. Instability of dynamical systems with several degrees of freedom. *Sov. Math. Doklady*, 5:581–585, 1964.

[Arn78]   V. I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, New York, 1978. ISBN 0-387-90314-3. Translated from the Russian by K. Vogtmann and A. Weinstein, Graduate Texts in Mathematics, 60.

[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.

[BBB03]   M. Berti, L. Biasco and P. Bolle. Drift in phase space: a new variational mechanism with optimal diffusion time. *J. Math. Pures Appl. (9)*, 82(6):613–664, 2003.

[BCV01]   U. Bessi, L. Chierchia and E. Valdinoci. Upper bounds on Arnold diffusion times via Mather theory. *J. Math. Pures Appl. (9)*, 80(1):105–129, 2001.

[Ber10]   P. Bernard. Arnold's diffusion: from the *a priori* unstable to the *a priori* stable case. In *Proceedings of the International Congress of Mathematicians. Volume III*, pages 1680–1700. Hindustan Book Agency, New Delhi, 2010.

[Car81]   J. R. Cary. Lie transform perturbation theory for Hamiltonian systems. *Phys. Rep.*, 79(2):129–159, 1981.

[CG94]   L. Chierchia and G. Gallavotti. Drift and diffusion in phase space. *Ann. Inst. H. Poincaré Phys. Théor.*, 60(1):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.

[Che08]   C.-Q. Cheng. Variational methods for the problem of Arnold diffusion. In *Hamiltonian dynamical systems and applications*, NATO Sci. Peace Secur. Ser. B Phys. Biophys., pages 337–365. Springer, Dordrecht, 2008.

[Che10]   C.-Q. Cheng. Variational construction of diffusion orbits for positive definite Lagrangians. In *Proceedings of the International Congress of Mathematicians. Volume III*, pages 1714–1728. Hindustan Book Agency, New Delhi, 2010.

[Che12]   C.-Q. Cheng. Arnold diffusion in nearly integrable Hamiltonian systems. *ArXiv e-prints*, July 2012.

[Chi79]   B. Chirikov. A universal instability of many-dimensional oscillator systems. *Phys. Rep.*, 52(5):264–379, 1979.

[DGLS08]   A. Delshams, M. Gidea, R. de la Llave and T. M. Seara. Geometric approaches to the problem of instability in Hamiltonian systems. An informal presentation. In *Hamiltonian dynamical systems and applications*, NATO Sci. Peace Secur. Ser. B Phys. Biophys., pages 285–336. Springer, Dordrecht, 2008.

[DH09]   A. Delshams and G. Huguet. Geography of resonances and Arnold diffusion in a priori unstable Hamiltonian systems. *Nonlinearity*, 22(8):1997–2077, 2009.

[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 $\mathbf{T}^2$. *Comm. Math. Phys.*, 209(2):353–392, 2000.

[DLS03]   A. Delshams, R. de la Llave and T. M. Seara. A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: announcement of results. *Electron. Res. Announc. Amer. Math. Soc.*, 9:125–134 (electronic), 2003.

[DLS06a]   A. Delshams, R. de la Llave and T. M. Seara. A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. *Mem. Amer. Math. Soc.*, 179(844):viii+141, 2006.

[DLS06b]   A. Delshams, R. de la Llave and T. M. Seara. Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows. *Adv. Math.*, 202(1):64–188, 2006.

[DLS08]   A. Delshams, R. de la Llave and T. M. Seara. Geometric properties of the scattering map of a normally hyperbolic invariant manifold. *Adv. Math.*, 217(3):1096–1153, 2008.

[DMM09]   P. Du Toit, I. Mezić and J. Marsden. Coupled oscillator models with no scale separation. *Phys. D*, 238(5):490–501, 2009.

[Fen72]   N. Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana Univ. Math. J.*, 21:193–226, 1971/1972.

[Fen74]   N. Fenichel. Asymptotic stability with rate conditions. *Indiana Univ. Math. J.*, 23:1109–1137, 1973/74.

[FGKR11] J. Fejoz, M. Guardia, V. Kaloshin and P. Roldán. Diffusion along mean motion resonance in the restricted planar three-body problem. *ArXiv e-prints*, 2011.

[FM00]   E. Fontich and P. Martín. Differentiable invariant manifolds for partially hyperbolic tori and a lambda lemma. *Nonlinearity*, 13(5):1561–1593, 2000.

[FM03]   E. Fontich and P. Martín. Hamiltonian systems with orbits covering densely submanifolds of small codimension. *Nonlinear Anal.*, 52(1):315–327, 2003.

[GL06a]   M. Gidea and R. de la Llave. Arnold diffusion with optimal time in the large gap problem, 2006. Preprint.

[GL06b]   M. Gidea and R. de la Llave. Topological methods in the instability problem of Hamiltonian systems. *Discrete Contin. Dyn. Syst.*, 14(2):295–328, 2006.

[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. ISBN 0-387-98697-9.

[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.

[HM82]   P. Holmes and J. Marsden. Melnikov's method and Arnold diffusion for perturbations of integrable Hamiltonian systems. *J. Math. Phys.*, 23(4):669–675, 1982.

[HPS77]   M. Hirsch, C. Pugh and M. Shub. *Invariant manifolds*, volume 583 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1977.

[Kol54]   A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954. English translation in *Stochastic Behavior in Classical and Quantum Hamiltonian Systems (Volta Memorial Conf., Como, 1977)*, Lecture Notes in Phys., 93, pages 51–56. Springer, Berlin, 1979.

[KZ12]   V. Kaloshin and K. Zhang. A strong form of Arnold diffusion for two and a half degrees of freedom. *ArXiv e-prints*, December 2012.

[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.

[LM88]   P. Lochak and C. Meunier. *Multiphase Averaging for Classical Systems*, volume 72 of *Appl. Math. Sci.* Springer, New York, 1988.

[LMM86]   R. de la Llave, J. M. Marco and R. Moriyón. Canonical perturbation theory of Anosov systems and regularity results for the Livšic cohomology equation. *Ann. of Math. (2)*, 123(3):537–611, 1986.

[LT83]   M. A. Lieberman and J. L. Tennyson. Chaotic motion along resonance layers in near-integrable Hamiltonian systems with three or more degrees of freedom. In J. C. W. Horton and L. E. Reichl, editors, *Long-time prediction in dynamics (Lakeway, Tex., 1981)*, pages 179–211. Wiley, New York, 1983.

[Mar13]   J.-P. Marco. Generic hyperbolic properties of classical systems on the torus, 2013. Preprint.

[Mat04]   J. N. Mather. Arnol′d diffusion. I. Announcement of results. *J. Math. Sci. (N. Y.)*, 124(5):5275–5289, 2004.

[Mey91]   K. R. Meyer. Lie transform tutorial. II. In K. R. Meyer and D. S. Schmidt, editors, *Computer aided proofs in analysis (Cincinnati, OH, 1989)*, volume 28 of *IMA Vol. Math. Appl.*, pages 190–210. Springer, New York, 1991.

[Mos66]   J. Moser. A rapidly convergent iteration method and non-linear partial differential equations. I. *Ann. Scuola Norm. Sup. Pisa (3)*, 20:265–315, 1966.

[Neĭ81]   A. I. Neĭshtadt. Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions. *Prikl. Mat. Mekh.*, 45(6):1016–1025, 1981.

[Pes04]   Y. B. Pesin. *Lectures on partial hyperbolicity and stable ergodicity.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2004. ISBN 3-03719-003-5.

[Pös82]   J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.*, 35(5):653–696, 1982.

[Pös01]   J. Pöschel. A lecture on the classical KAM theorem. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 707–732. Amer. Math. Soc., Providence, RI, 2001.

[PT07]   G. N. Piftankin and D. V. Treshchëv. Separatrix maps in Hamiltonian systems. *Russian Math. Surveys*, 62(2):219–322, 2007.

[Sal04]   D. A. Salamon. The Kolmogorov-Arnold-Moser theorem. *Math. Phys. Electron. J.*, 10:Paper 3, 37 pp. (electronic), 2004.

[Ten82]   J. Tennyson. Resonance transport in near-integrable systems with many degrees of freedom. *Phys. D*, 5(1):123–135, 1982.

[Thi97]   W. Thirring. *Classical mathematical physics.* Springer-Verlag, New York, third edition, 1997. ISBN 0-387-94843-0. Dynamical systems and field theories, Translated from the German by Evans M. Harrell, II.

[Tre12]   D. Treschev. Arnold diffusion far from strong resonances in multidimensional *a priori* unstable Hamiltonian systems. *Nonlinearity*, 25(9):2717–2757, 2012.

[Zeh75]   E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems/I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

[Zeh76a]   E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems/II. *Comm. Pure Appl. Math.*, 29:49–111, 1976.

[Zeh76b]   E. Zehnder. Moser's implicit function theorem in the framework of analytic smoothing. *Math. Ann.*, 219(2):105–121, 1976.

[Zha11] K. Zhang. Speed of Arnold diffusion for analytic Hamiltonian systems. *Invent. Math.*, 186(2):255–290, 2011.

E-mail address: `Amadeu.Delshams@upc.es`

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain

School of Mathematics, Georgia Inst. of Technology, 686 Cherry St., Atlanta, GA 30332
E-mail address: `rafael.delallave@math.gatech.edu`

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain
E-mail address: `Tere.M-Seara@upc.es`