Available online at www.sciencedirect.com
SCIENCE DIRECT
ADVANCES IN Mathematics
Advances in Mathematics 202 (2006) 64–188
www.elsevier.com/locate/aim

Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows

Amadeu Delshams[a], Rafael de la Llave[b,*], Tere M. Seara[a]

[a]*Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain*

[b]*Department of Mathematics and ICES, University of Texas, Austin, TX 78712-0257, USA*

Received 15 July 2004; accepted 18 March 2005

Communicated by C. Fefferman
Available online 11 May 2005

Abstract

We show that certain mechanical systems, including a geodesic flow in any dimension plus a quasi-periodic perturbation by a potential, have orbits of unbounded energy.

The assumptions we make in the case of geodesic flows are:

(a) The metric and the external perturbation are smooth enough.
(b) The geodesic flow has a hyperbolic periodic orbit such that its stable and unstable manifolds have a tranverse homoclinic intersection.
(c) The frequency of the external perturbation is Diophantine.
(d) The external potential satisfies a generic condition depending on the periodic orbit considered in (b).

The assumptions on the metric are $\mathcal{C}^2$ open and are known to be dense on many manifolds. The assumptions on the potential fail only in infinite codimension spaces of potentials.

The proof is based on geometric considerations of invariant manifolds and their intersections. The main tools include the scattering map of normally hyperbolic invariant manifolds, as well as standard perturbation theories (averaging, KAM and Melnikov techniques).

We do not need to assume that the metric is Riemannian and we obtain results for Finsler or Lorentz metrics. Indeed, there is a formulation for Hamiltonian systems satisfying scaling

*Corresponding author. Fax: +1 512 471 9038.

E-mail addresses: Amadeu.Delshams@upc.edu (A. Delshams), llave@math.utexas.edu (R. de la Llave), Tere.M-Seara@upc.edu (T.M. Seara).

0001-8708/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.aim.2005.03.005

hypotheses. We do not need to make assumptions on the global topology of the manifold nor on its dimension.
© 2005 Elsevier Inc. All rights reserved.

MSC: 37J40; 37J25; 53D25; 70H08

Keywords: Geodesic flows; Scattering map; Melnikov theory; Homoclinic intersections; KAM theory; Arnold diffusion

Contents

1. Introduction

The goal of this paper is to give a proof, using geometric perturbation methods, of a generalization of a result proved by the authors in [DLS00], which provided a geometric version of a result of [Mat96]. (Another geometric version of the results of [Mat96] was developed in [BT99].)

More precisely, we will show that the mechanical system consisting of a geodesic flow in a manifold plus a time quasi-periodic potential possesses orbits of unbounded energy, provided that the geodesic flow and the potential satisfy some mild non-degeneracy assumptions. We refer to Theorem 1.3 for a precise formulation of the results on geodesic flows.

The only feature of the geodesic flow that we use in the proof is that the metric is homogeneous in the momenta so that, for high enough energy, we can consider the potential as a small and slow perturbation of the geodesic flow. In contrast with many variational results, we do not need that the Hamiltonian is convex in the momenta.

Hence, we obtain results for geodesic flows not only in Riemannian metrics but also when the metric is Finsler or Lorentz or when the system has a magnetic field.

Similarly, we do not use the homology of the manifold. Provided that the geodesic flow has a periodic orbit with a transverse homoclinic intersection, all our analysis is carried out in a neighborhood of the orbit and the connection. In particular, our results apply just as well to geodesic flows on the sphere.

Even if many of the methods of this paper are similar to [DLS00], there are important differences both conceptual and technical. In any case, we have striven to make this paper self-contained so that it can be read independently from [DLS00]. (A more detailed comparison of this paper with [DLS00] and with other papers can be found in Section 1.2.)

Indeed, we have attempted to make this paper not only self-contained but also pedagogical and have included many details that can be found in standard references and indeed documented results that are in the folklore. Since the main result is proved by an assembly of diverse techniques, which are developed in different places, we hope that this would be useful for the readers.

In this paper we will deal with a n-dimensional manifold M, and we will consider a $\mathcal{C}^r$ metric g on it (r sufficiently large).

We recall that a geodesic "λ" is a curve "λ": $\mathbb{R} \to M$, parameterized by arc length which is a critical point for length between any two points. It is also possible to consider a dynamical system given by the geodesic flow in $\mathbf{S}_1 M$, the unit tangent bundle of M. We denote the parameterized curve in $\mathbf{S}_1 M$ corresponding to the geodesic "λ" as $\lambda(t)$, and we denote by

$$\hat{\lambda} = \mathrm{Range}(\lambda) \subset \mathbf{S}_1 M.$$

Note that we can change the origin in the parameterization of the geodesic arbitrarily. We will assume that this origin is chosen once and for all. Once this choice is made, other choices of coordinates that we will introduce later will become unique. See (6).

We will assume that the metric g verifies:

H1. There exists a closed geodesic "Λ" such that its corresponding periodic orbit $\hat{\Lambda}$ under the geodesic flow is hyperbolic.

H2. There exists another geodesic "γ" such that $\hat{\gamma}$ is a transversal homoclinic orbit to $\hat{\Lambda}$.

That is, $\hat{\gamma}$ is contained in the intersection of the stable and unstable manifolds of $\hat{\Lambda}$, $W^{\mathrm{s}}_{\hat{\Lambda}}$, $W^{\mathrm{u}}_{\hat{\Lambda}}$, in the unit tangent bundle.

Moreover, we assume that the intersection of the stable and unstable manifolds of $\hat{\Lambda}$ is transversal along $\hat{\gamma}$. That is

$$T_{\gamma(t)} W^{\mathrm{s}}_{\hat{\Lambda}} + T_{\gamma(t)} W^{\mathrm{u}}_{\hat{\Lambda}} = T_{\gamma(t)} \mathbf{S}_1 M, \quad t \in \mathbb{R}. \tag{1}$$

We will assume, without loss of generality and just to avoid typographical clutter that the period of "λ" is 1. This, clearly, can be achieved by choosing the units of time, which does not affect any of the subsequent discussions.

The abundance of systems satisfying hypotheses H1, H2 is described in Section 2. We just note that they can be found arbitrarily close to integrable metrics (e.g. the standard metrics in the torus or in the sphere).

Remark 1.1. A consequence of the hyperbolicity of $\hat{\Lambda}$ assumed in H1 is that the orbits that tend to $\hat{\Lambda}$ actually approach a specific orbit contained in $\hat{\Lambda}$. (This is a particular case of the well-known fact in normal hyperbolicity theory that $W^{\mathrm{s}}_{\hat{\Lambda}} = \cup_{x \in \hat{\Lambda}} W^{\mathrm{s}}_x$. See Appendix B for an account of the theory of normally hyperbolic invariant manifolds.)

That is, if "γ", "Λ" are as in H2, there exist real numbers a_+, a_-, such that

$$\operatorname{dist}(\text{“}\Lambda\text{”}(s + a_\pm), \text{“}\gamma\text{”}(s)) \to 0 \quad \text{as} \quad s \to \pm\infty. \tag{2}$$

We recall that hyperbolicity of $\hat{\Lambda}$ implies that there exist $C > 0$, $\beta_0 > 0$ such that

$$\operatorname{dist}(\text{“}\Lambda\text{”}(s + a_\pm), \text{“}\gamma\text{”}(s)) \leqslant C e^{-\beta_0 |s|}, \quad \text{as} \quad s \to \pm\infty.$$

Then, the orbits of $\hat{\Lambda}$ and $\hat{\gamma}$ approach exponentially fast, both in the future and in the past. Standard perturbation theory for ordinary differential equations (see e.g. [CL55]) shows that the asymptotic phase shift $\Delta := a_+ - a_-$ exists and is unique modulo an integer multiple of the period of "Λ".

In this paper, we study the effects of adding a quasi-periodic potential U to the geodesic flow. We will need that the frequency of the perturbation satisfies Diophantine conditions, so we recall the standard definition.

Definition 1.2. We say that $\nu \in \mathbb{R}^d$ is Diophantine when there exist $\kappa > 0$ and $\tau \geqslant d - 1$ such that

$$|\nu \cdot k| \geqslant \kappa |k|^{-\tau} \qquad \forall k \in \mathbb{Z}^d \setminus \{0\}. \tag{3}$$

We have collected some more information on Diophantine numbers in Appendix A, Section A.1.

Note that when $d = 1$, all real numbers different from zero satisfy condition (3) with $\tau = 0$. When ν is the frequency of a quasi-periodic motion, $d = 1$ corresponds to a periodic motion. Hence, the periodic potentials considered in [DLS00] are particular cases of the quasi-periodic potentials considered in this paper.

The main result of this paper for geodesic flows is:

Theorem 1.3. *Let $\nu \in \mathbb{R}^d$ be Diophantine, $r \in \mathbb{N}$ be sufficiently large (depending on τ, the Diophantine exponent of ν).*

Let g be a C^r metric on a compact manifold M, verifying hypotheses H1, H2, *and $U : M \times \mathbb{T}^d \to \mathbb{R}$ a generic C^r function.*

Consider the time dependent Lagrangian

$$L(q, \dot{q}, \nu t) = \frac{1}{2} g^q(\dot{q}, \dot{q}) - U(q, \nu t), \tag{4}$$

where g^q denotes the metric in $\mathbf{T}_q M$.

Then, the Euler–Lagrange equation of L has a solution $q(t)$ whose energy

$$E(t) = \frac{1}{2} g^q(\dot{q}(t), \dot{q}(t)) + U(q(t), \nu t),$$

tends to infinity as $t \to \infty$.

We will deduce Theorem 1.3 from a more general result, Theorem 3.4, stated in the Hamiltonian formulation. In turn, Theorem 3.4 will be deduced from the more general Theorem 4.27 which establishes the existence of orbits whose energy changes in largely arbitrary ways. In particular, we will establish the existence of uncountably many orbits whose energy is unbounded.

Remark 1.4. Even if, for the moment, we have only claimed the result for a generic potential, the genericity condition for U will be described very explicitly in the statement of Theorem 3.4. It amounts to assuming that the *Poincaré* function $\mathcal{L}$ (introduced in (16)), associated to the homoclinic intersection and the potential, is not constant. This Poincaré function is, roughly, a combination of integrals of U along the geodesics "Λ", "γ" satisfying H1, H2. If we fix "Λ", "γ", this condition is true for all U's except those in a set of infinite codimension.

Notice that once a system has some geodesics satisfying H1, H2, it has infinitely many (e.g. apply Smale's horseshoe theorem). If the hypothesis on the Poincaré function $\mathcal{L}$ is verified just for one pair Λ, γ, the existence of orbits with unbounded energy will follow. Hence, it is extremely rare to have potentials that fail to satisfy this hypothesis. We conjecture that given a geodesic flow satisfying H1, H2, *all non-degenerate* analytic potentials (i.e. all potentials U such that $\partial_q \partial_\theta U(q, \theta) \not\equiv 0$) satisfy it.

Remark 1.5. Let us emphasize that the conclusion of Theorem 1.3 is the existence of orbits whose energy changes significantly. This is different from the results in

[Gal97,Gal99], which also consider quasi-periodic perturbations. In [Gal97,Gal99], the variables that experience changes of order 1 are actions introduced in the Hamiltonian formalism which are not present in the Lagrangian formalism. We will present a more detailed comparison in Remark 4.24.

Remark 1.6. We can give explicit bounds for the value of r for which the method presented here works. The proof that we present here shows that the argument works if $r > \max(53 + 4\tau, 28 + 5\tau)$, where $\tau \geqslant d - 1$ is the Diophantine exponent of v given in Definition (1.2), but we do not claim this to be the minimum value for the result to be true or for the techniques presented here to work.

Theorem 1.3 is a generalization of Theorem 1.1 of [DLS00]. In that paper, periodic perturbations of a geodesic flow in a two-torus were considered. The hypotheses of [DLS00] are contained in the hypotheses of this paper.

The existence of orbits with unbounded energy in perturbations of geodesic flows of $\mathbb{T}^2$ by periodic potentials had been established in [Mat96] using variational methods. We also note that [BT99] presents a geometric mechanism for existence of unbounded orbits different from the one presented in [DLS00] and in this paper. Other mechanisms that rely more on hyperbolicity can be found in [Lla02,Tre02a,Tre02b]. An example where the diffusion is generated by oscillations of an adiabatic invariant is presented somewhat heuristically in [IdlLNV02].

One motivation to study external quasi-periodic perturbations, is that they are a natural step towards more realistic models in which one can find orbits with unbounded energy. Models in which the energy is affected by a quasi-periodic perturbation are natural models of the solar system (see [GJSM01a,GJSM01b,GLMS01,GSLM01,SGJM95]).

The method of proof that we present in this paper is related to the method of proof used in [DLS00]. A description of the proof used in the present paper is given in Section 1.1 and the comparison with that of [DLS00] in Section 1.2.

The orbits that we produce can be described heuristically in the same way that the orbits of [DLS00] and the orbits of [Mat96]. They are orbits that remain "parked" near the periodic geodesic, but when the phases of the external perturbation are such that a homoclinic excursion will lead to a gain of energy, they perform it. Of course, the details of the proof are very different for quasi-periodic perturbations and for periodic ones.

Remark 1.7. We do not know whether a variational proof of existence of unbounded orbits in the models considered here could be obtained. The main obstacles seem to be the consideration of quasi-periodic perturbations, the use of Hamiltonians which are not positive definite and the fact that the manifold we consider may not have any non-trivial homology. Of course, there are quite a number of variants of variational methods and it seems possible that some of them could be adapted to the models at hand. In particular, we call attention to [Itu96] which contains a study of variational methods for non-autonomous systems with rather general time dependence.

Remark 1.8. In the mechanism of [BT99], the orbits stay parked not near periodic orbits but near whiskered tori with one hyperbolic direction. Hence, for two-dimensional geodesic flows, the orbits of [DLS00] and those of [BT99] admit a similar geometric description. Nevertheless, for higher-dimensional manifolds they are quite different. For example, there are quite detailed studies [Kli78,Kli83] of the abundance of metrics with periodic orbits and homoclinic intersections. On the other hand, for dimensions greater than 2, we are not aware of studies of abundance of whiskered tori with 1-D stable or unstable manifolds. We think that it would be interesting to obtain a method that unifies the methods of [BT99,DLS00].

In the mechanism of [Lla02], there are no periodic or quasi-periodic orbits that play an explicit role. The mechanism of [Lla02] does not require that the perturbation is Diophantine.

Remark 1.9. Note that the mechanism for unbounded growth of energy presented here can be considered somewhat related to the classical Arnol'd diffusion. Indeed, we obtain the growth of energy by establishing the existence of a transition chain of whiskered tori with unbounded energy. (See the precise definitions of whiskered tori and transition chains and what we call transition path later in Definitions 4.28 and 4.32.)

We note that the unperturbed system (the geodesic flow) already has transversal homoclinic orbits. This is different from most of the situations considered in Arnol'd diffusion in which the unperturbed system is considered to be integrable; either in the sense of having action angle variables—called *a priori stable* in [CG94,CG98]—or in the sense of having conserved quantities with separatrices—called *a priori unstable* in [CG94]—. In [DLS00] the systems in which the unperturbed part has already transversal homoclinic intersections are called *a priori chaotic*.

The field of diffusion has received a great deal of attention recently. We mention the recent papers [BB02,BBB03,CQC03,DdlLMS03,DdlLS03,EMR01,Moe02,MS02,Tre02a, Tre02b,Tre02c,Tre04] as well as the announcements [Mat2,Xia98].

1.1. Summary of the method

The proof will be organized in a sequence of steps. We emphasize that the steps are quite independent of each other and that most of them are just extensions of well-known techniques (normal hyperbolicity, averaging theory, KAM theory, Melnikov method, shadowing method). Besides extending and unifying the above-mentioned standard techniques, a tool that will be very useful for us is the "*scattering map*" associated to a normally hyperbolic invariant manifold with homoclinic connections. This tool was introduced in [DLS00].

We emphasize that the strategy has a rather simple geometric interpretation and that in the geometric language, the progression of the argument is very clear. Of course, rigorous proofs require a variety of techniques and are, therefore, long, but we hope that the fact that the steps are largely independent can make it possible to read it modularly and even skip some sections which may look obvious to the expert. Needless to say, most of the steps used in the present proof can be accomplished using techniques different from the ones we present here. We hope that the modularity will encourage alternative approaches.

We hope that the strategy presented here (and the attendant toolkit we used to implement it) can be applied to a variety of problems. Indeed, in [DdlLMS03,DdlLS03], we have used geometric methods to construct orbits that overcome the *large gap problem*.

In the rest of this section, we give a description of the method and in the next section, we will highlight some of the main technical differences with [DLS00].

In Section 2, we study some cases where hypotheses H1 and H2 are verified. For example, they are verified for all closed surfaces of genus bigger or equal than 2, and are generic for all compact surfaces. They are also known to happen in higher dimensional manifolds. We think that it is reasonable to conjecture that they hold for generic metrics in all manifolds.

In Section 3 we highlight some geometric features of the geodesic flow defined in the extended phase space $\mathbf{T}^*M \times \mathbb{T}^d$ when using the Hamiltonian formalism. We show that H1 and H2 can be formulated as the existence of a $(d+2)$-dimensional normally hyperbolic invariant manifold $\tilde{\Lambda}$ filled by $(d+1)$-dimensional tori, and the existence of a $(d+2)$-dimensional homoclinic manifold $\tilde{\gamma}$ to $\tilde{\Lambda}$.

One important geometric observation is that the geodesic flow restricted to $\tilde{\Lambda}$ is the product of a one degree of freedom Hamiltonian and a Kronecker flow in a d-dimensional torus, and therefore, it is integrable.

In Section 3.7 we introduce *the scattering map*, $\tilde{S} : \tilde{\Lambda} \to \tilde{\Lambda}$ associated to $\tilde{\gamma}$, which is one of the main tools we will use to find homoclinic and heteroclinic connections. We construct the scattering map $\tilde{S}$ associated to $\tilde{\gamma}$ as follows: Given an orbit $\gamma(t)$ in $\tilde{\gamma}$, $\tilde{S}$ associates to the orbit in $\tilde{\Lambda}$ asymptotic to $\gamma(t)$ in the past, the orbit in $\tilde{\Lambda}$ asymptotic to $\gamma(t)$ in the future.

The scattering map is a geometrically natural way to describe homoclinic or heteroclinic transitions between invariant objects of a normally hyperbolic invariant manifold. The geometrical naturalness of the method will become useful when we carry out the perturbative calculations to establish existence of heteroclinic intersections.

In Section 4 we study the effects of the quasi-periodic potential in all the invariant objects for the geodesic flow that we had considered before.

First of all, to exhibit the perturbative character of the problem for high energy, we introduce, in Section 4.2, scaled variables. This leads to the fact that, for high enough energy, the potential can be considered as a small and slow perturbation of the geodesic flow. More concretely, setting $\varepsilon = \frac{1}{\sqrt{E^*}}$ for some large value of the energy E^*, the potential can be considered, in a neighborhood of the surface of energy E^*, as a perturbation of the geodesic flow of size $O(\varepsilon^2)$ and of frequency $O(\varepsilon)$. We note that, for the subsequent analysis, the fact that the perturbation is slow plays a much more significant role than its smallness.

In Section 4.3 we use the theory of normally hyperbolic invariant manifolds to obtain the persistence of a $(d+2)$-dimensional manifold $\tilde{\Lambda}_\varepsilon$, close to $\tilde{\Lambda}$.

We also formulate the persistence of the stable and unstable manifolds to $\tilde{\Lambda}_\varepsilon$ and their homoclinic intersections along a manifold $\tilde{\gamma}_\varepsilon$, close to $\tilde{\gamma}$. That is, all the features that we highlighted in Section 3 persist. As a consequence, it is possible to define still a scattering map for the perturbed manifold $\tilde{\Lambda}_\varepsilon$.

In Section 4.3.4 we consider the Hamiltonian flow restricted to the perturbed invariant manifold $\tilde{\Lambda}_\varepsilon$. We first observe that the flow restricted to $\tilde{\Lambda}_\varepsilon$ is a slow perturbation of an integrable system.

Therefore, in Section 4.3.5, taking advantage of the fact that our system is very differentiable, we use an averaging method to high order to show that, in some canonical variables, the flow is an extremely small quasi-periodic perturbation of an integrable flow. In Section 4.3.6, we apply KAM theory to this averaged flow and we prove that the manifold $\tilde{\Lambda}_\varepsilon$ contains an abundance of KAM tori, with extremely small gaps between them (the gaps can be bounded from above by a large power of the inverse of the energy).

In Section 4.4 we compute perturbatively the scattering map on the perturbed manifold $\tilde{\Lambda}_\varepsilon$. In particular, we show that the computation of heteroclinic orbits to invariant tori by means of the scattering map reduces to a variant of the Melnikov method introduced in [Tre94] and already developed in [DLS00]. We work out the relation of the scattering map formalism and the more commonly used language of Melnikov functions and Melnikov potentials. One advantage of the scattering map method is that the scattering map is defined on the whole manifold $\tilde{\Lambda}_\varepsilon$ and does not use that the objects considered can be reduced to a common system of coordinates or have a particular dynamics. This allows to compute connections among invariant objects of a different nature. This advantage is particularly crucial in [DdlLS03,DdlLMS03].

We conclude in Lemma 4.33 that, under the hypothesis that some explicitly computable function (the Poincaré function (16)) is non-constant, there exist transition paths joining tori with arbitrarily large energies.

The non-triviality of the Poincaré function is the generic hypothesis on U alluded in Theorem 1.3. It will be clear that, once we select the Λ, γ that verify hypotheses H1, H2, the condition on the potential is verified by a $\mathcal{C}^s$ open and dense set of potentials with $s \geqslant 1$.

We note for experts that in this paper we formulate our results in terms of *transition paths* (see Definition 4.32) rather than in terms of transition chains as it is commonly done in the literature. While transition chains are just a sequence of whiskered tori with heteroclinic connections between consecutive ones, transition paths specify the sequence of tori and the heteroclinic connections between consecutive ones. Hence, when we construct orbits shadowing transition paths, we not only specify which tori they visit but also which paths they use to move from one torus to the next. This makes explicit a much more detailed control.

In Section 4.5.3 we establish the existence of transition paths involving tori whose energy goes to infinity. As a matter of fact, we will establish the existence of a sequence of tori, whose energy goes to infinity so that there are connections going from each one of them to the next and to the precedent. This makes it possible to construct transition paths whose energy changes in largely arbitrary ways.

Once we have the existence of a transition path between invariant tori whose energy goes to infinity, in Section 4.5.4, we show that, given a—possibly infinite—transition path, there exist orbits which follow the transition path with arbitrarily chosen accuracy. In particular, the energy remains always close to that of the invariant tori and the connections in the transition chain.

Theorem 4.27 is a rather precise statement about the existence of orbits whose energy performs largely arbitrary excursions, in particular, of an uncountable number of orbits whose energy tends to infinity.

Even if the KAM theorem we use, namely Theorem 4.8, is very close to results that can be found in the literature—see in particular [Zha00] and [BHTB90]—we have not found a statement that includes explicitly the quantitative results we need. Hence, we have developed Appendix A to establish Theorem 4.8 by modifying slightly [Zeh76a,Zeh76b]. Undoubtedly, the modifications presented will be well known to experts but we hope that the explicit presentation could be useful for some readers and make this paper self-contained.

Also for the sake of completeness, we present in Appendix B proofs of the result on persistence of normally hyperbolic invariant manifolds that we use. The statement that we need presents some peculiarities which are not covered in standard treatment, such as dealing with non-compact manifolds or with unbounded vector fields. Nevertheless, using the special structure of the problem, it is possible to give a very simple and quantitative proof.

Remark 1.10. We emphasize that, as pointed out in [AA67], the argument of existence of shadowing orbits we use in the proof of Theorem 4.27 is purely topological, once one has an appropriate λ-lemma which uses $\mathcal{C}^2$ properties of the map. We follow the version of the argument in [DLS00] which relies on the λ-lemma of [FM00,FM03].

The topological argument presented here does not produce any estimates on the times of transition. Hence Theorem 1.3 does not make any claim on the times required for the energy to grow. It seems quite possible to us that one could modify this paper using the more quantitative connecting arguments developed in—among others—[BB02,BBB03,BCV01,CG03,Tre02a]. It also seems that at these stage, one could also use some connecting argument based in the *broken geodesics* method [Bes96,RS02], or on topological methods [Eas78,Eas89,GR03,GR04]. This may be the subject of future work.

It seems almost certain that the model presented here contains other mechanisms of diffusion with quantitatively different properties.

1.2. Comparison of the method of this paper with the methods of [DLS00] and other papers

Even if at the level of a superficial description the method of proof of this paper is similar to that of [DLS00], there are several important differences which we now describe. Of course, this section does not contain any result used later in this paper and can be skipped safely.

We call attention to two technical improvements that can have further applications. Notably, we include a more precise description of transition chains (transition paths) and shadowing lemmas and a more efficient perturbative calculation of the scattering map.

1.2.1. Removing assumptions on the underlying manifold

We have removed completely the assumption that the manifold is $\mathbb{T}^2$.

Indeed, in [DLS00], this assumption was not used much, only to obtain a system of coordinates.

The geometric method just needs to analyze the behavior of the system in a neighborhood of the periodic orbit and the homoclinic excursion to it.

Therefore, independently of the ambient dimension and the properties of the manifold M, it suffices to perform a largely two-dimensional analysis on the $(d+2)$-dimensional manifolds $\tilde{\Lambda}$, $\tilde{\gamma}$, formed by the periodic and the homoclinic geodesics for all sufficiently large values of the energy plus the d quasi-periodic variables.

1.2.2. KAM theorem

In this paper, we have to resort to a more sophisticated KAM theorem (4.8) instead of the standard KAM theorem for twist maps as was done in [DLS00]. In contrast with the KAM theorem for twist maps, we could not find a proof of the theorem we need in the literature. So, we present a proof in Appendix A. The proof is significantly less optimized (in aspects such as the differentiability assumptions and the size of the gaps) than the KAM theorem for twist maps. When applied to the case considered in [DLS00], the present proof requires more differentiability on the metric and on the potential than the result in [DLS00].

1.2.3. Asymptotic expansion of the KAM tori

Related to that, we note that we have also changed the method of obtaining an approximate description of the $(d+1)$-dimensional invariant tori produced by KAM theorem 4.8.

We recall that the problem of obtaining approximate expresions for KAM tori with a fixed frequency is solved by the standard Linsdstedt series. In our case, however, one of the frequencies is also become small at the same time that the perturbation becomes small. (in particular, the frequency depends on the perturbation parameter ε). Hence, if one considers all the possible Fourier modes for a range of ε, one has zero divisors.

In [DLS00], we just considered Fourier modes of size $\varepsilon^{-1/2}$. In this way, the smallerst denominators are $\varepsilon^{1/2}$. Using the decay of coefficients due to the differentiability, it is possible to bound—in a space of less regularity—the error of the approximate solution by a power of ε. The power is large if we allow a larger loss of differentiability.

In this paper, in Section 4.3.7, we use just that, in the averaged coordinates, the tori are almost flat, but we make back the change of variables explicitly. This has the advantage that it makes more explicit the fact that tori of similar frequency fit together.

We also note that the problem of obtaining Lindstedt series when the frequencies are in different scales has been extensively studied in [Gal94,GGM99,GGM00]. Unfortunately, the methods of these papers are not suitable for our case, since they rely on the perturbations being analytic. We will present full details of an elementary geometric method which yields the results we need.

1.2.4. A quasi-periodic Melnikov method

The Melnikov method that has been developed in [DLS00] has been adapted to the quasi-periodic case. This requires to develop again the theory of Melnikov functions. Since there are d-dimensional phases in the problem, the properties of the Melnikov vector function will be more complicated that those of the periodic case. Nevertheless we show that this Melnikov vector function is the gradient of a Melnikov potential. This potential unifies the existence of heteroclinic orbits and the gain of energy since both can be obtained by taking directional derivatives of the Melnikov potential.

1.2.5. Geometry of the intersections in extended phase spaces

We call attention to the fact that the geometry of the intersections of manifolds in the quasi-periodic case is very different from that of the periodic case. Indeed, it contains a geometric surprise. In contrast with the periodic case, naive dimension counting does not predict the intersection of the stable and unstable invariant manifolds we consider in this paper. The structure given by the fact that the system is a quasi-periodic perturbation is very important.

We will introduce angle variables that give the phases of the external perturbations and, to keep the symplectic structure, we will need to introduce external actions conjugated to these angles. These action variables have little dynamical meaning. We will be dealing with a Hamiltonian system of $n+d$ degrees of freedom and the phase space will be $(2n+2d)$-dimensional. We will find in it a family of $(d+1)$-dimensional tori with heteroclinic connections between them. A naive dimension counting of dimensions would suggest that the dimension of the family of tori with heteroclinic intersections is $d+1$. Nevertheless, taking into account that the d extra action variables are not dynamical variables, we will see that the dimension of the families is just one-dimensional, indeed the parameter is the energy.

1.2.6. Transition paths

Our study of diffusing orbits is based on the study of "transition paths" (see Definition 4.32) which is more precise than the usual study based on transition chains.

We recall that a transition chain specifies a sequence of whiskered tori so that the unstable manifold of one intersects transversally the stable manifold of the next.

A transition path specifies a transition chain and the heteroclinic orbits connecting one torus and the next. This is more precise than just specifying the transition chain since two whiskered tori have several (infinitely many in the cases we consider) connecting orbits between them.

We will prove that, given a transition path, there is a true orbit that stays arbitrarily close to it.

This has some useful consequences. In the cases we consider, it is possible to arrange transition paths for which the energy changes only a small amount during the connecting path. Hence, we obtain that the energy of the shadowing orbit is very close to the sequence of energies of the tori in the chain.

This later result seems to have been known to experts (and was used e.g. in [BT99] and [DLS00]) but, as it was pointed out by M. Sevryuk, it was never explicitly written in the literature even if it was commonly used and its proof was folklore.

1.2.7. Comparison with [BT99]

The main difference with [BT99] is that, rather than basing our transition path on whiskered invariant tori of codimension one, we base it on the remnants of the periodic orbits under quasi-periodic perturbation. This allows us to carry out a good part of the argument using the theory of normally hyperbolic invariant manifolds. The method of [BT99] does not need to use the theory of normally hyperbolic invariant manifolds. For geodesic flows in two-dimensional manifolds, hyperbolic periodic orbits are the same as whiskered tori of codimension 1. Hence, for the case considered in [Mat96], the orbits produced in this paper and [DLS00] are the same as those produced in [BT99]. Nevertheless, for higher dimensional geodesic flows, they are very different. In particular, the results on abundance of our hypotheses H1, H2 for higher-dimensional systems are very different from the results on abundance of the corresponding hypotheses for [BT99].

2. On the abundance of hypotheses H1, H2

There are many cases where the hypotheses H1 and H2 are known to hold. In this section, we summarize some of these cases, as a motivation to them.

2.1. Riemannian surfaces

For two-dimensional compact boundaryless manifolds, the situation is very clear.

First of all, the theory of [Hed32], based on [Mor24] (see [Ban88] for a modern exposition of the classical theory and several developments) and supplemented by some remarks in [Mat96], shows that H1, H2 are $\mathcal{C}^r$ generic for metrics on $\mathbb{T}^2$, if $r \geqslant 2$.

On the sphere $\mathbf{S}^2$, one can also construct [Don88] some particular examples that contain a great abundance of horseshoes. A recent paper [CBP02] shows that on $\mathbf{S}^2$ there is a $\mathcal{C}^2$ dense and open set of $\mathcal{C}^\infty$ metrics whose geodesic flows contain a hyperbolic orbit with a transverse homoclinic intersection. See also [KW02].

On surfaces of genus bigger or equal than 2, the following argument of [Kat82] shows that hyperbolic orbits with homoclinic connections exist for *all* $\mathcal{C}^{2+\delta}$ metrics, $\delta > 0$.

Lemma 2.1. *Let M^2 be a surface of genus 2 or higher. Then, any $\mathcal{C}^{2+\delta}$ metric, $\delta > 0$, on M^2 has hyperbolic geodesics with transversal homoclinic connections.*

We just reproduce the relevant steps of the proof since in the paper [Kat82] the argument is used to reach a slightly different conclusion.

Positive topological entropy. We note that the fundamental group of a surface of genus $g \geqslant 2$ has exponential growth. That is, given a set $\alpha_1, \ldots, \alpha_N$ of generators of $\Pi_1(M^2)$, we have that the number of different words $\alpha_{i_1} \cdot \cdots \cdot \alpha_{i_L}$ of length L is at least $e^{\sigma L}$ for some $\sigma > 0$.

By Tonneli's theorem (see [KH95, Theorem 9.5.10, p. 371]), in each of these homotopy classes, we can find a shortest periodic geodesic $\beta_{\alpha_{i_1}\cdots\alpha_{i_L}}$ which minimizes the length among all the closed curves in this homotopy class.

It is also true that we can find an $\varepsilon_0 > 0$ such that the distance between two of these minimizers in different homotopy classes is bigger than ε_0. The reason is that, since the minimizers satisfy differential equations, they are uniformly differentiable. If two uniformly differentiable curves of finite length are sufficiently close (depending only on the modulus of continuity of the derivative), they are homotopically equivalent.

Since for some $S < \infty$ we can find a closed curve of length smaller that LS in the homotopy class—take as a test function a closed curve which is a concatenation of the generators of the fundamental group—we conclude that

$$|\beta_{\alpha_{i_1}\cdots\alpha_{i_L}}| \leqslant LS.$$

From these considerations, it follows that the topological entropy of the geodesic flow in a manifold with genus greater than 1 is strictly positive. (See also [KH95, Theorem 9.6.7, p. 374].)

Alternative arguments for positive topological entropy of geodesic flows in this situation can be found in [Din71,Kat82]. A systematic study of the relations between topology of the manifold and entropy (and other dynamical properties of the geodesic flow) can be found in [Pat99].

Invariant measure with some positive Lyapunov exponents. From the variational principle, (see [KH95, Theorem 4.5.3, p. 181]) we have that there is an invariant measure μ whose measure-theoretic entropy is arbitrarily close to the topological entropy, in particular, positive.

By Ruelle's inequality [Rue78] we have that the measure μ has to have some positive Lyapunov exponents. (In contrast to most of the results of Pesin theory, Ruelle's inequality only needs that the flow is $\mathcal{C}^1$.)

Invariant measure with no zero Lyapunov exponents. Since the flow is symplectic, the existence of a strictly positive Lyapunov exponent implies the existence of a strictly negative Lyapunov exponent.

Since the geodesic flow of a two-dimensional manifold takes place in the unit tangent bundle, which is three dimensional—this is the only place of the argument where we use that the manifold is two dimensional—the only zero Lyapunov exponent will correspond with the motion along the flow.

Existence of horseshoes. The theory of measures with no zero Lyapunov exponents [Kat80] implies the existence of horseshoes. Indeed, the topological entropy of these horseshoes approximates the metric entropy of the measure. (Simpler versions of this result can be found in [FL92,Pol93].)

Even if [FL92,Kat80,Pol93] are only written for diffeomorphisms, it is not difficult to adapt the result for flows. See the remarks on [Kat82], where a similar adaptation is used.

This part of the proof as written in the literature uses that the flow is $\mathcal{C}^{1+\delta}$. We do not know if this could be lowered to just requiring $\mathcal{C}^1$ regularity.

2.2. Hamiltonian systems, higher-dimensional manifolds, Finsler metrics

The main theorem of this paper—Theorem 3.4—is formulated in the generality of Hamiltonian systems of the form

$$H(p,q,t) = H_0(p,q) + U(q,\nu t),$$

where H_0 is homogeneous of degree two in the momenta, and plays the role of the geodesic flow. It is therefore useful to discuss the abundance of the analogous hypotheses H1′, H2′ in the class of Hamiltonian systems.

In the context of Hamiltonian systems, once one has periodic orbits with elliptic directions, by studying the associated Poincaré map in a neighborhood of the corresponding fixed point in the center manifold, one can find periodic points with transversal homoclinic intersections in $\mathcal{C}^r$ generic Hamiltonian systems for several r. See [New77,Rob70a,Rob70b,Tak70,Zeh73].

Since geodesic flows are more restrictive than Hamiltonian systems, the arguments showing genericity for Hamiltonian systems do not apply straightforwardly to geodesic flows. Nevertheless, it is not unreasonable to conjecture that many of the above properties for Hamiltonian systems have analogues for geodesic flows of Riemannian metrics.

In particular, we think it is reasonable to conjecture (as it is widely believed by experts) that the existence of a periodic orbit with a transverse homoclinic is $\mathcal{C}^r$ dense for $r = 2$ in the space of smooth Riemannian metrics for any manifold. One can also make a similar conjecture for $r > 2$, but the proof seems to be out of reach since the known techniques seem to require generalizing closing lemmas and the like, which appear to be quite difficult to obtain for higher regularities.

There is some progress in the direction of the proof of the above conjecture. The fact that intersections of stable and unstable manifolds can be made transverse by arbitrary small perturbations is established for surfaces in [Don95] and for any manifold in [BW02] following an unpublished argument of Petroll [Pet96]. A detailed account of the argument of [Pet96] can be found in Appendix A of [CBP02].

Well-known surveys of results on the existence and abundance of closed geodesics in Riemannian manifolds are [Kli78,Kli83].

In a less systematic direction, it is not difficult to produce examples of systems satisfying H1, H2 by considering perturbations of metrics with integrable geodesic flows in spheres or tori. Similarly, taking products of manifolds that satisfy the hypotheses, we obtain manifolds that satisfy them.

For the case of Finsler metrics, we remark that many of the results on abundance of periodic orbits with homoclinic intersections for Hamiltonians can be extended straightforwardly to geodesic flows of Finsler metrics. See for example [ISM00].

2.3. Hedlund examples

In [Hed32], one can find examples of metrics in $\mathbb{T}^d$, $d \geqslant 3$, where there are very few class-A minimizing geodesics.

These examples consist in modifying the metric so that there are periodic orbits along the generators of the homology so that the metric is much weaker in tubes along them than outside. As a consequence, to obtain an orbit with a certain homology, it is advantageous to just move along the generators except for at most $d-1$ jumps than to move close to the direction in the homology, so that there are no class-A minimizers except along the minimizers.

The paper [Lev97] shows that, when $d = 3$, one one can construct Hedlund examples in such a way that they admit symbolic dynamics. That is, one can prescribe sequences of the minimizers and get an orbit that visits the terms in the sequence in the given order. Furthermore, the orbits thus produced have the property that any two of them with sequences that agree in the future, actually converge exponentially. This is, of course, very reminiscent of what one expects of hyperbolic orbits.

Indeed, the examples described in [Lev97] have the property that the periodic orbits associated to the minimizers are hyperbolic. The constructions in [Lev97] have as a corollary that the stable manifolds of the periodic orbits associated to the special directions have intersections. By modifying the metric slightly, it is possible to make the intersections transversal, so that they verify the hypothesis H1.

3. Hamiltonian formalism of the unperturbed problem

In this section, we introduce a Hamiltonian formalism for the problem and formulate the hypotheses H1′, H2′ which are the Hamiltonian counterparts of the hypotheses H1, H2 for the geodesic flow. We find the Hamiltonian formalism more convenient than the Lagrangian one because the geometric tools we use (averaging, KAM, etc.) are customarily formulated in Hamiltonian language and normal hyperbolicity results are formulated for first-order differential equations.

The rest of the arguments in this paper will be formulated in terms of Hamiltonian formalism and will only use H1′, H2′ and that the main part of the Hamiltonian and the perturbation behave differently under scaling. We formulate the main result of the Hamiltonian formalism as Theorem 3.4. It clearly implies Theorem 1.3. It is also clear that Theorem 3.4 applies as well to Lorentz or to Finsler metrics.

3.1. Hamiltonian formalism and notation

The Hamiltonian phase space of the geodesic flow is $\mathbf{T}^*M$. We will denote the local coordinates in M by q and the cotangent directions by p.

As it is well known, the phase space, being a cotangent bundle, admits a canonical exact symplectic form

$$\Omega = d\alpha. \tag{5}$$

In local coordinates, $\alpha = \sum_i p_i dq_i$, $\Omega = \sum_i dp_i \wedge dq_i$.

With respect to the form Ω, the geodesic flow is Hamiltonian and the Hamiltonian function is

$$H_0(p,q) = \tfrac{1}{2} g_q(p,p),$$

where g_q is the metric in $\mathbf{T}^*M$. We will denote by Φ_t this geodesic flow.

Since the energy H_0 is preserved and it is not degenerate, for each E, the energy level $\Sigma_E = \{(p,q) \;\; H_0(p,q) = E\}$, is a $(2n-1)$-dimensional manifold invariant under the geodesic flow.

Given an arbitrary geodesic "λ": $\mathbb{R} \to M$ we will denote

$$\lambda_E(t) = (\lambda_E^p(t), \lambda_E^q(t))$$

the orbit of the geodesic flow that lies in the energy surface Σ_E, and such that λ_E^q, the projection over q, runs along the range of "λ".

Moreover, we fix the origin of time in λ_E so that it corresponds to the origin of the parameterization in "λ". More precisely,

$$H_0(\lambda_E(t)) = E$$

and

$$\text{Range}(\text{“}\lambda\text{”}) = \text{Range}(\lambda_E^q), \qquad \text{“}\lambda\text{”}(0) = \lambda_E^q(0). \tag{6}$$

It is easy to check that the above conditions determine uniquely the orbit of the geodesic flow in the cotangent bundle corresponding to a geodesic "λ".

We use $\hat{\lambda}_E$ to denote the range of the orbit $\lambda_E(t)$. That is, $\hat{\lambda}_E$ is the lift to the hypersurface of energy E of the geodesic "λ".

Note that the orbits of the geodesic flow rescale with energy as

$$\left(\lambda_E^p(t), \lambda_E^q(t)\right) = \left(\sqrt{2E}\,\lambda_{1/2}^p\left(\sqrt{2E}t\right), \lambda_{1/2}^q\left(\sqrt{2E}t\right)\right). \tag{7}$$

Since $\Lambda_{1/2}$ has period 1 with our conventions that the geodesic "Λ" is normalized to have length 1, then Λ_E has period $1/\sqrt{2E}$.

3.2. Normal hyperbolicity properties

Now, we start to discuss normal hyperbolicity properties of certain objects invariant under the geodesic flow. This will lead to the fact that these objects have analogues in the system with the potential perturbation included. Standard references for normal hyperbolicity theory are [Fen72,Fen74,HP70,HPS77]. For the sake of making this paper more self-contained we have presented proofs of the results we use for our system in Appendix B.

Notation 3.1. We follow standard practice in the theory of normally hyperbolic invariant manifolds and call stable and unstable manifolds different objects than those called stable and unstable manifolds in topological dynamics.

In topological dynamics, the stable set of an invariant object is the set of points whose orbit converges to the object. In the theory of normally hyperbolic invariant manifolds, the stable manifold is the set of points whose orbit converges to an orbit of the normally hyperbolic invariant manifold at an exponential rate with an exponent larger than the bound on the tangential exponents. (See Appendix A of [DLS00] for some discussion on this point and references to the original literature.)

Similarly, homoclinic and heteroclinic orbits refers to orbits in the intersections of the stable and unstable manifolds in the sense of the theory of normally hyperbolic invariant manifolds, that is, including explicit exponential rates. We will not use homoclinic or heteroclinic just in the sense of convergence.

Notation 3.2. Given a manifold M and two submanifolds N_1, N_2, we say that N_1 intersects transversally N_2 (denoted as $N_1 \pitchfork N_2$) when there is a point $x \in N_1 \cap N_2$ and

$$T_x N_1 + T_x N_2 = T_x M. \tag{8}$$

Since the manifolds $N_{1,2}$ can be considered as submanifolds of several manifolds, when there is risk of confusion, we will write $N_1 \pitchfork_M N_2$ to denote that they intersect transversally when considered as submanifolds of M.

We recall that the standard usage in transversality theory is to say that the intersection between N_1 and N_2 is transversal either if $N_1 \cap N_2 = \emptyset$ or when all the points of the intersection satisfy (8). We will maintain the difference between these two usages of the word by emphasizing that the sentences where *the intersection is* allow for empty intersection whereas the sentences *manifolds intersect transversally* imply that the manifolds intersect. Indeed, in the overwhelming majority of the usage in this paper, we have that the intersection is not empty.

The hypotheses H1, H2 of the geodesic flow when formulated in the Hamiltonian formalism for the Hamiltonian H_0 translate into:

H1$'$. For any $E > 0$, there exists a periodic orbit $\Lambda_E(t)$, as in (7), of the Hamiltonian H_0 whose range $\hat{\Lambda}_E$ is a normally hyperbolic invariant manifold in the energy surface

$$\Sigma_E := \{(p, q) \in \mathbf{T}^* M\,,\ H_0(p, q) = E\}. \tag{9}$$

H2$'$. The stable and unstable manifolds $W^{\mathrm{s,u}}_{\hat{\Lambda}_E}$ of $\hat{\Lambda}_E$ are n-dimensional, and there exists a homoclinic orbit $\gamma_E(t)$. That is, the range of γ_E satisfies

$$\hat{\gamma}_E \subset \left(W^{\mathrm{s}}_{\hat{\Lambda}_E} \setminus \hat{\Lambda}_E\right) \cap \left(W^{\mathrm{u}}_{\hat{\Lambda}_E} \setminus \hat{\Lambda}_E\right).$$

Moreover, this intersection is transversal as intersection of invariant manifolds in the energy surface Σ_E along $\hat{\gamma}_E$.

As a consequence of the hyperbolicity of $\hat{\Lambda}_{1/2}$, we have that, analogously to (2), for some $a_\pm \in \mathbb{R}$,

$$\operatorname{dist}(\Lambda_{1/2}(t + a_\pm), \gamma_{1/2}(t)) \to 0 \quad \text{as} \quad t \to \pm\infty. \tag{10}$$

Indeed, as we already noticed in Remark 1.1, the hyperbolicity of $\hat{\Lambda}_{1/2}$ implies that there exist $C > 0$ and an *exponential rate* $\beta_0 > 0$, such that

$$\operatorname{dist}(\Lambda_{1/2}(t + a_\pm), \gamma_{1/2}(t)) \leqslant C e^{-\beta_0 |t|}, \quad \text{as} \quad t \to \pm\infty. \tag{11}$$

Moreover, it is clear that there exists a number $A > 0$ such that

$$\max\{\operatorname{dist}_{W^{\mathrm{s}}_{\hat{\Lambda}_{1/2}}}(\hat{\gamma}_{1/2}, \hat{\Lambda}_{1/2}),\ \operatorname{dist}_{W^{\mathrm{u}}_{\hat{\Lambda}_{1/2}}}(\hat{\gamma}_{1/2}, \hat{\Lambda}_{1/2})\} \leqslant A, \tag{12}$$

where given a submanifold W we denote by $\operatorname{dist}_W(x, y)$ the distance along the submanifold.

As a consequence of (12) and the rescaling properties (7), we also have that

$$\max\{\operatorname{dist}_{W^{\mathrm{s}}_{\hat{\Lambda}_E}}(\hat{\gamma}_E, \hat{\Lambda}_E),\ \operatorname{dist}_{W^{\mathrm{u}}_{\hat{\Lambda}_E}}(\hat{\gamma}_E, \hat{\Lambda}_E)\} \leqslant A\sqrt{E}. \tag{13}$$

As a consequence of (13), there exist compact subsets $K^{\mathrm{s}}_E \subset W^{\mathrm{s}}_{\hat{\lambda}_E}$, $K^{\mathrm{u}}_E \subset W^{\mathrm{u}}_{\hat{\lambda}_E}$, such that

$$\hat{\gamma}_E \subset (K^{\mathrm{s}}_E \cap W^{\mathrm{u}}_{\hat{\lambda}_E}) \bigcup (K^{\mathrm{u}}_E \cap W^{\mathrm{s}}_{\hat{\lambda}_E}). \tag{14}$$

This property will play a role in Section 4.4 when we study perturbation theory of the stable and unstable manifolds and their homoclinic intersections for a finite range of energies, say, $E \in [1/2, 2]$.

Since $\hat{\gamma}_E$ is one-dimensional, $W^{\mathrm{s}}_{\hat{\Lambda}_E}$,$W^{\mathrm{u}}_{\hat{\Lambda}_E}$ are n-dimensional and the ambient manifold Σ_E is $(2n-1)$-dimensional, we have

$$T_x \hat{\gamma}_E = T_x W^{\mathrm{s}}_{\hat{\Lambda}_E} \cap T_x W^{\mathrm{u}}_{\hat{\Lambda}_E}$$

for all the points $x \in \hat{\gamma}_E$.

Hence, the transversality assumption gives that $\hat{\gamma}_E$ is the locally unique intersection between the stable and unstable manifolds of $\hat{\Lambda}_E$.

For the Hamiltonian H_0, the energy is preserved and therefore the dynamics can be analyzed on each energy surface, but when we consider the external quasi-periodic potential depending on time, the energy will change, and then, it will be useful to consider any fixed value $E_0 > 0$, and introduce the manifold $\Lambda = \bigcup_{E \geqslant E_0} \hat{\Lambda}_E$ for all values of the energy larger than E_0. In subsequent lemmas, we will assume that E_0 is large enough.

The following Proposition is an obvious description of the situation.

Proposition 3.3. *Define $\Lambda = \bigcup_{E \geqslant E_0} \hat{\Lambda}_E$. It is a 2-dimensional normally hyperbolic invariant manifold with boundary satisfying*:

- *Λ is diffeomorphic to $[E_0, \infty) \times \mathbb{T}^1$.*
- *The canonical symplectic form Ω on $\mathbf{T}^*M$ restricted to Λ is non-degenerate.*
- *The form $\Omega|_\Lambda$ is invariant under the flow Φ_t of the Hamiltonian $H_0(p, q)$.*
- *The stable and unstable manifolds of Λ, W^{s}_Λ and W^{u}_Λ, are $(n+1)$-dimensional manifolds diffeomorphic to $[E_0, \infty) \times \mathbb{T}^1 \times \mathbb{R}^{n-1}$.*
- *W^{s}_Λ and W^{u}_Λ intersect transversally along γ, defined by*

$$\gamma = \bigcup_{E \geqslant E_0} \hat{\gamma}_E \subset \left(W^{\mathrm{s}}_\Lambda \setminus \Lambda\right) \cap \left(W^{\mathrm{u}}_\Lambda \setminus \Lambda\right)$$

which is diffeomorphic to $[E_0, \infty) \times \mathbb{R}$.

3.3. Statement of results for Hamiltonian systems

In this section, we will state our main result in the Hamiltonian language, Theorem 3.4, which clearly implies Theorem 1.3.

Theorem 3.4. *Let M be a compact manifold. Let H_0 be a Hamiltonian in $\mathbf{T}^*M$ which satisfies*:

(i) *H_0 is homogeneous of degree 2 in the momenta, that is $H_0(\rho p, q) = \rho^2 H_0(p, q)$ for $\rho \in \mathbb{R}^+$.*
(ii) *The Hamiltonian system generated by H_0 satisfies* H1′, H2′.

Let $\nu \in \mathbb{R}^d$ be a Diophantine number as in (1.2). *Let $U : M \times \mathbb{T}^d \to \mathbb{R}$ be a function and consider the time-dependent Hamiltonian system*

$$H(p, q, t) = H_0(p, q) + U(q, \nu t). \tag{15}$$

Assume furthermore that:

(iii) *The functions U and H_0 are $\mathcal{C}^r$, where r is sufficiently large depending on the dimension d and the Diophantine exponent of ν.*

(iv) *Consider the Poincaré function* $\mathcal{L} : \mathbb{T}^d \to \mathbb{R}$ *associated to the orbits* $\Lambda_{1/2}$ *and* $\gamma_{1/2}$ *verifying* (10) *and* (11):

$$\mathcal{L}(\theta) = \lim_{T_1, T_2 \to \infty} \left[\int_{-T_1}^{T_2} dt\, \widetilde{U}(\gamma^q_{1/2}(t), \theta) - \int_{-T_1+a_-}^{T_2+a_+} dt\, \widetilde{U}(\Lambda^q_{1/2}(t), \theta) \right], \tag{16}$$

where the functions $\overline{U}(\theta)$, *and* $\widetilde{U}(q, \theta)$ *are defined by*:

$$\overline{U}(\theta) = \int_0^1 U(\Lambda^q_{1/2}(\varphi), \theta)\, d\varphi, \qquad \widetilde{U}(q, \theta) = U(q, \theta) - \overline{U}(\theta). \tag{17}$$

Assume that the Poincaré function (16) *is non-constant.*

Then the Hamiltonian system generated by H in (15) *has orbits whose energy tends to infinity.*

The proof of Theorem 3.4 will be accomplished in the rest of the paper. Indeed, we will establish the more general result Theorem 4.27, which implies Theorem 3.4.

We note that the hypotheses we make do not involve neither convexity properties of the Hamiltonian H_0 nor that it is a quadratic form. Therefore, they apply to Lorentz metrics or to Finsler metrics. Similarly, we do not need to assume that the orbits Λ_E, γ_E of hypotheses H1′, H2′ are minimizers.

We note that the Poincaré function (16) depends only on the geometric geodesics "Λ" $= \Lambda^q_{1/2}$, "γ" $= \gamma^q_{1/2}$, which live in the manifold M.

Although we will not discuss this in detail, we will not even need H_0 to be quadratic. It would suffice that H_0 is homogeneous in p of positive order or sums of terms each of which is homogeneous of positive order.

3.4. A remark on normally hyperbolic invariant manifolds in product systems

We will use the following elementary result in the next section.

Proposition 3.5. *Let* Λ *be a normally hyperbolic invariant manifold under a flow* Φ_t *on a manifold M. Let* $\beta > \beta_\Lambda \geqslant 0$ *be the exponential expansion rates corresponding to the normal hyperbolicity of* Λ *as in* (11). *Let N be another manifold with a flow* φ_t *with exponential expansion rates less or equal than* β_Λ. *Consider the flow* $\tilde{\Phi}_t := (\Phi_t, \varphi_t)$ *on the manifold* $M \times N$.

Then the manifold $\tilde{\Lambda} := \Lambda \times N$ *is a normally hyperbolic invariant manifold for the flow* $\tilde{\Phi}_t$.

Moreover, $W^{\mathrm{s}}_\Lambda \times N = W^{\mathrm{s}}_{\tilde{\Lambda}}$ *is the stable manifold of* $\tilde{\Lambda}$ *for the extended flow* $\tilde{\Phi}_t$.

For $x \in \Lambda$, $y \in N$, *we have that* $W^{\mathrm{s}}_{(x,y)} = W^{\mathrm{s}}_x \times N$ *is the stable manifold of the point* (x, y).

The same results hold for the unstable manifold.

In the applications we have in mind for this paper, the flow φ_t will be either a rotation on a d-dimensional torus or the identity.

Proof. The proof of Proposition 3.5 is an obvious consequence of the definition of normally hyperbolic invariant manifolds. Note that $T_{(x,y)}(M \times N) = T_x M \oplus T_y N$ and $T_{(x,y)}(\Lambda \times N) = T_x \Lambda \oplus T_y N$.

Hence the decomposition

$$T_x M = T_x \Lambda \oplus E_x^{\mathrm{s}} \oplus E_x^{\mathrm{u}} \tag{18}$$

assumed to exist in the normal hyperbolicity of Λ, leads to a decomposition

$$T_{(x,y)}(M \times N) = T_{(x,y)}(\Lambda \times N) \oplus E_x^{\mathrm{s}} \oplus E_x^{\mathrm{u}}. \tag{19}$$

It is easy to see that if decomposition (18) is invariant under Φ_t, then decomposition (19) is invariant under $\tilde{\Phi}_t$. Moreover, using the fact that the exponential expansion rates of the vector field φ_t are smaller than β_Λ it is immediate that

$$||D\tilde{\Phi}_t(x, y)|_{T_{(x,y)}\Lambda \times N}|| = ||D\Phi_t(x)|_{T_x\Lambda} \oplus D\varphi_t(y)|| \leqslant C e^{\beta_\Lambda |t|}. \qquad \square$$

3.5. Extended phase spaces

To study quasi-periodic systems it is customary to make the system autonomous by introducing an extra variable $\theta \in \mathbb{T}^d$, which moves at a constant frequency v. Then, the phase space will be the $(2n + d)$-dimensional manifold $\mathbf{T}^*M \times \mathbb{T}^d$, which we will call the *extended phase space*.

We will denote by $\tilde{\Phi}_t(p, q, \theta) = (\Phi_t(p, q), \theta + vt)$ the flow on the extended phase space corresponding to the unperturbed Hamiltonian. Since in the unperturbed system, the variable θ does not affect the other variables, the flow $\tilde{\Phi}_t$ is the Cartesian product of the flow on $\mathbf{T}^*M$ and the rotation at constant velocity v on $\mathbb{T}^d$.

Following Proposition 3.5, we will introduce the notation $\tilde{\Lambda} = \Lambda \times \mathbb{T}^d$, and analogously $\tilde{\gamma} = \gamma \times \mathbb{T}^d$. Then, applying Proposition 3.5 to the results of Proposition 3.3 we obtain:

Proposition 3.6. *Under assumptions* H1′ *and* H2′ *we have*, *for any value of* $E > 0$:

- $\mathcal{T}_E := \hat{\Lambda}_E \times \mathbb{T}^d$ *is a* $(d+1)$*-dimensional whiskered invariant torus* (*sees Definition* 4.28). *Its stable and unstable manifolds* $W^{\mathrm{s}}_{\mathcal{T}_E} = W^{\mathrm{s}}_{\hat{\Lambda}_E} \times \mathbb{T}^d$, $W^{\mathrm{u}}_{\mathcal{T}_E} = W^{\mathrm{u}}_{\hat{\Lambda}_E} \times \mathbb{T}^d$, *are* $(n+d)$*-dimensional manifolds.*
- *The manifolds* $W^{\mathrm{s}}_{\mathcal{T}_E}$ *and* $W^{\mathrm{u}}_{\mathcal{T}_E}$ *intersect along* $\hat{\gamma}_E \times \mathbb{T}^d$. *This intersection is a transversal intersection in* $\Sigma_E \times \mathbb{T}^d$, *where* Σ_E *is the energy surface introduced in* (9).

That is, for all $\tilde{x} \in \hat{\gamma}_E \times \mathbb{T}^d$, *we have*:

$$T_{\tilde{x}} W^{\mathrm{s}}_{\mathcal{T}_E} + T_{\tilde{x}} W^{\mathrm{u}}_{\mathcal{T}_E} = T_{\tilde{x}}(\Sigma_E \times \mathbb{T}^d).$$

Moreover, for any $E_0 > 0$:

- *The manifold* $\tilde{\Lambda} := \cup_{E \geqslant E_0} \mathcal{T}_E = \Lambda \times \mathbb{T}^d$ *is a* $(d+2)$*-dimensional manifold.*
- $\tilde{\Lambda}$ *is a normally hyperbolic invariant manifold for the extended flow* $\tilde{\Phi}_t$.
- *The* (*un*)*stable manifolds of* $\tilde{\Lambda}$ *are*

$$W^{\mathrm{u,s}}_{\tilde{\Lambda}} = W^{\mathrm{u,s}}_{\Lambda} \times \mathbb{T}^d. \tag{20}$$

  In particular, they are $(n+d+1)$*-dimensional manifolds.*
- *The manifolds* $W^{\mathrm{s}}_{\tilde{\Lambda}}$ *and* $W^{\mathrm{u}}_{\tilde{\Lambda}}$ *intersect transversally in the extended phase space* $\mathbf{T}^*M \times \mathbb{T}^d$ *along*

$$\tilde{\gamma} := \gamma \times \mathbb{T}^d = \bigcup_{E \geqslant E_0} \hat{\gamma}_E \times \mathbb{T}^d \subset \left(W^{\mathrm{s}}_{\tilde{\Lambda}} \setminus \tilde{\Lambda}\right) \cap \left(W^{\mathrm{u}}_{\tilde{\Lambda}} \setminus \tilde{\Lambda}\right).$$

This extended phase space is obviously not symplectic. In order to keep the symplectic character we add d real variables (actions) $A = (A_1, \ldots, A_d)$ symplectically conjugated to θ, which do not change with time.

Then, the symplectically extended phase space, which we will call the *full symplectic space*, is $\mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$, which is $(2n+2d)$-dimensional. The symplectic form in the full symplectic space is

$$\Omega^* = \Omega + \sum_{i=1}^{d} dA_i \wedge d\theta_i.$$

The full symplectic form Ω^* is exact:

$$\Omega^* = d\left(\alpha + \sum_{i=1}^{d} A_i d\theta_i\right) = d\alpha^*. \tag{21}$$

We will denote by $\Phi^*_t(p, q, A, \theta) = (\Phi_t(p, q), A, \theta + vt)$ the full symplectic flow. This flow is Hamiltonian with respect to the form Ω^*, and the Hamiltonian function is

$$H^*_0(p, q, A, \theta) := v \cdot A + H_0(p, q). \tag{22}$$

Since $A_1, \ldots, A_d$ are conserved, the restriction of the flow Φ^*_t to each of the manifolds obtained by fixing the values of all the actions is identical to the flow of H_0 in the extended phase space. These A variables do not have a dynamical meaning since their value does not affect the dynamics of the variables (p, q, θ).

Proposition 3.7. *We have the following geometric properties in the full symplectic space* $\mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$:

- *For any fixed* $E > 0$, $B \in \mathbb{R}^d$, *the set*

$$\Sigma^*_{E,B} := \{(p, q, A, \theta)\,,\ H_0(q, p) = E,\, A = B\}$$

  is a $(2n - 1 + d)$*-dimensional manifold. It is invariant under the full symplectic flow* Φ^*_t.
- *The set*

$$\mathcal{T}^*_{E,B} = \hat{\Lambda}_E \times \{B\} \times \mathbb{T}^d$$

  is a $(d + 1)$*-dimensional whiskered invariant torus contained in* $\Sigma^*_{E,B}$.
- *The torus* $\mathcal{T}^*_{E,B}$ *has stable and unstable manifolds* $W^{\mathrm{s}}_{\mathcal{T}^*_{E,B}} = W^{\mathrm{s}}_{\hat{\Lambda}_E} \times \{B\} \times \mathbb{T}^d$, $W^{\mathrm{u}}_{\mathcal{T}^*_{E,B}} = W^{\mathrm{u}}_{\hat{\Lambda}_E} \times \{B\} \times \mathbb{T}^d$, *which are* $(n + d)$*-dimensional manifolds.*
- *The manifolds* $W^{\mathrm{s}}_{\mathcal{T}^*_{E,B}}$, $W^{\mathrm{u}}_{\mathcal{T}^*_{E,B}}$ *intersect along* $\hat{\gamma}_E \times \{B\} \times \mathbb{T}^d$. *This intersection is a transversal intersection in* $\Sigma^*_{E,B}$. *That is, for all* $x^* \in \hat{\gamma}_E \times \{B\} \times \mathbb{T}^d$, *we have*:

$$T_{x^*} W^{\mathrm{s}}_{\mathcal{T}^*_{E,B}} + T_{x^*} W^{\mathrm{u}}_{\mathcal{T}^*_{E,B}} = T_{x^*} \Sigma^*_{E,B}. \tag{23}$$

Moreover, for any $E_0 > 0$,

- $\Lambda^* = \cup_{E \geqslant E_0, B} \mathcal{T}^*_{E,B}$ *is a* $(2d + 2)$*-dimensional manifold.*
- Λ^* *is a normally hyperbolic invariant manifold for the full symplectic flow* Φ^*_t *in* $\mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$.
- *The (un)stable manifolds of* Λ^* *are*

$$W^{\mathrm{u,s}}_{\Lambda^*} = W^{\mathrm{u,s}}_{\Lambda} \times \mathbb{R}^d \times \mathbb{T}^d.$$

  In particular, they are $(n + 1 + 2d)$*-dimensional manifolds.*
- *The manifolds* $W^{\mathrm{s}}_{\Lambda^*}$ *and* $W^{\mathrm{u}}_{\Lambda^*}$ *intersect transversally in the full symplectic space along*

$$\gamma^* := \gamma \times \mathbb{R}^d \times \mathbb{T}^d \subset \left(W^{\mathrm{s}}_{\Lambda^*} \setminus \Lambda^*\right) \cap \left(W^{\mathrm{u}}_{\Lambda^*} \setminus \Lambda^*\right).$$

3.6. A coordinate system on $\tilde{\Lambda}$, and Λ^*

Now we want to describe a coordinate system in $\tilde{\Lambda}$ (and Λ^*) that is not only defined on $\tilde{\Lambda}$ but also on a neighborhood of it.

Recall that the only difference between $\tilde{\Lambda}$ and Λ^* is that Λ^* includes the variables A that restore the symplectic character of the problem.

One real valued coordinate in $\tilde{\Lambda}$ is $J = \sqrt{2H_0} \geqslant \sqrt{2E_0}$. For the conjugate angle coordinate, we will take $\varphi \in \mathbb{T}^1$, which is determined by $dJ \wedge d\varphi = \Omega|_{\Lambda}$, and $\varphi = 0$ corresponds to the origin of the parameterization in the curve "Λ" chosen at the begining, (6) (of course, the choice of the origin of the parameterization is arbitrary and there are many other choices which will work.) The other d angles are the global coordinates θ of the quasi-periodic perturbation.

If we are in Λ^* we will take also the conjugate momentum to θ which are the global real coordinates A. Hence $\alpha|_{\Lambda^*} = J\,d\varphi + A\,d\theta$. If we express the full symplectic flow in Λ^* in these variables, it is an integrable Hamiltonian flow of Hamiltonian $v \cdot A + \frac{1}{2}J^2$ and the equations of motion are

$$\dot{J} = 0, \quad \dot{\varphi} = J, \quad \dot{A} = 0, \quad \dot{\theta} = v.$$

By formula (7), the geodesic Λ_E of hypothesis H1$'$ is given in coordinates (J, φ) by

$$\Lambda_E = \{(J, \varphi) : \ J = \sqrt{2E}\,, \varphi \in \mathbb{T}\},$$

and the flow Φ_t on it associated to $\Lambda_E(t + \varphi_0/\sqrt{2E}) = \Lambda_{1/2}(\sqrt{2E}t + \varphi_0)$, for any $\varphi_0 \in \mathbb{R}$ is $J = \sqrt{2E}$, $\varphi = \sqrt{2E}t + \varphi_0$.

Associated to Λ_E there is a family of $(d+1)$-dimensional tori in Λ^* given by

$$\mathcal{T}^*_{E,B} = \{(J, \varphi, A, \theta) : \ J = \sqrt{2E},\ A = B,\ \varphi \in \mathbb{T},\ \theta \in \mathbb{T}^d\}, \tag{24}$$

for any B.

The full symplectic flow Φ^*_t on the tori $\mathcal{T}^*_{E,B}$ is given by

$$J = \sqrt{2E}, \quad \varphi = \sqrt{2E}t + \varphi_0, \quad A = B, \quad \theta = vt + \theta_0, \quad \varphi_0 \in \mathbb{T},\ \theta_0 \in \mathbb{T}^d.$$

All these tori $\mathcal{T}^*_{E,B}$ project, in $\tilde{\Lambda}$, to the same torus

$$\mathcal{T}_E = \{(J, \varphi, \theta) : \ J = \sqrt{2E},\ \varphi \in \mathbb{T},\ \theta \in \mathbb{T}^d\}, \tag{25}$$

and the extended flow $\tilde{\Phi}_t$ on the torus $\mathcal{T}_E$ is given by

$$J = \sqrt{2E}, \quad \varphi = \sqrt{2E}t + \varphi_0, \quad \theta = vt + \theta_0, \quad \varphi_0 \in \mathbb{T},\ \theta_0 \in \mathbb{T}^d.$$

3.7. The scattering (outer) map

Once we have seen in Proposition 3.6 that our system possess a normally hyperbolic invariant manifold $\tilde{\Lambda}$ and a transversal homoclinic manifold $\tilde{\gamma}$ associated to it, we are going to study the heteroclinic or homoclinic connections in $\tilde{\gamma}$ between the invariant objects (specially the $(d+1)$-dimensional tori $\mathcal{T}_E$) that fill $\tilde{\Lambda}$. To this end, we introduce a map $\tilde{S} : \tilde{\Lambda} \to \tilde{\Lambda}$ that we call the "*scattering map*" or the "*outer map*". This scattering map $\tilde{S}$ will transform the asymptotic point at $-\infty$ of a homoclinic orbit to $\tilde{\Lambda}$ into its asymptotic point at $+\infty$. More concretely, we define the scattering (or outer) map $\tilde{S} : \tilde{\Lambda} \mapsto \tilde{\Lambda}$ associated to $\tilde{\gamma}$ by

$$\tilde{x}_+ = \tilde{S}(\tilde{x}_-) \iff W^{\mathrm{s}}_{\tilde{x}_+} \cap W^{\mathrm{u}}_{\tilde{x}_-} \cap \tilde{\gamma} \neq \emptyset. \tag{26}$$

Since the manifolds $W^{\mathrm{s}}_{\tilde{\Lambda}}$, and $W^{\mathrm{u}}_{\tilde{\Lambda}}$ are characterized by the exponential convergence with a uniform exponential expansion rate β, the condition in (26) is equivalent to $\exists \tilde{z} \in \tilde{\gamma} \subset \mathbf{T}^*M \times \mathbb{T}^d$, such that

$$\operatorname{dist}\left(\tilde{\Phi}_t(\tilde{x}_\pm), \tilde{\Phi}_t(\tilde{z})\right) \leqslant Ce^{-\beta|t|}, \quad \text{as} \quad t \to \pm\infty.$$

for some constant $C > 0$.

Using the coordinates (J, φ, θ) on $\tilde{\Lambda}$ introduced in Section 3.6, the scattering map can be computed explicitly. Let us recall that, by hypothesis H2$'$, inequality (11) is fulfilled, and the rescaling properties (7) imply

$$\operatorname{dist}\left(\Lambda_E\left(t + \frac{\varphi_0 + a_\pm}{\sqrt{2E}}\right), \gamma_E\left(t + \frac{\varphi_0}{\sqrt{2E}}\right)\right) \leqslant C\sqrt{2E}e^{-\beta_0\sqrt{2E}|t|} \quad \text{as} \quad t \to \pm\infty. \tag{27}$$

Hence, the points $\tilde{x}_\pm = \left(\Lambda_E\left((\varphi_0 + a_\pm)/\sqrt{2E}\right), \theta_0\right) \in \mathcal{T}_E$ are asymptotically connected through $\tilde{z} = \left(\gamma_E\left(\varphi_0/\sqrt{2E}\right), \theta_0\right)$, and the orbit $\tilde{\Phi}_t(\tilde{z})$ is an homoclinic orbit to the torus $\mathcal{T}_E$.

Then, the map $\tilde{S}$ is expressed as

$$\tilde{S}(J, \varphi + a_-, \theta) = (J, \varphi + a_+, \theta),$$

or more simply, calling $\Delta = a_+ - a_-$ the phase shift in the φ-coordinate:

$$\begin{aligned} \tilde{S} : \tilde{\Lambda} &\to \tilde{\Lambda}, \\ (J, \varphi, \theta) &\mapsto (J, \varphi + \Delta, \theta). \end{aligned} \tag{28}$$

Note that the phase shift Δ is uniquely defined in spite of the fact that the point $\tilde{z}$ is not unique and that the $a_\pm$ are defined only up to the simultaneous addition of an integer. (For more details about the definition of $\tilde{S}$ see [DLS00].)

Let us note also that, as $\tilde{S}(\mathcal{T}_E) = \mathcal{T}_E$, any torus $\mathcal{T}_E$ has only homoclinic orbits and no orbits heteroclinic to another torus. Of course, this will change when we add perturbations.

It is useful to have an analogous definition of the scattering map in the full symplectic space. Since the actions A do not change, we just have that the unperturbed scattering map is given by

$$\begin{aligned} S^* : \Lambda^* &\to \Lambda^*, \\ (J, \varphi, A, \theta) &\mapsto (J, \varphi + \Delta, A, \theta). \end{aligned} \tag{29}$$

The scattering map is formulated in terms of the intersections of stable and unstable manifolds. The fact that these manifolds persist under changes in the dynamical system and depend smoothly on parameters will allow us to compute the scattering map perturbatively for high enough energies in Section 4.4.

Remark 3.8. As it is obvious from the definition, the map $\tilde{S}$ depends on the $\tilde{\gamma}$ chosen and, therefore, of the geodesic "γ" verifying hypothesis H2. Indeed, different choices of $\tilde{\gamma}$ lead to different scattering maps.

Nevertheless, in the rest of the paper, $\tilde{\gamma}$ will be fixed and we will impose conditions on the external potential depending on $\tilde{\gamma}$. Hence, we will omit the $\tilde{\gamma}$ from the notation for the scattering map.

Remark 3.9. By the implicit function theorem and the transversality of the intersection between the stable and the unstable manifolds of $\tilde{\Lambda}$ along $\tilde{\gamma}$, it is clear that the scattering map is locally well defined.

The fact that if we continue $\tilde{S}$ along a closed loop in $\tilde{\Lambda}$ we obtain the same map follows from Remark 1.1.

Remark 3.10. The scattering map can be defined in other situations [DLS00,DdlLS03,DdlLMS03], as long as there exists a normally hyperbolic invariant manifold $\tilde{\gamma}$ associated to it (see [DdlLMS04] for more details). If the stable and the unstable manifolds $W^{\mathrm{s,u}}_{\tilde{\Lambda}}$ fail to intersect transversally at some points of $\tilde{\gamma}$, the domain and the range of the scattering map are subsets of $\tilde{\Lambda}$, which we denote by H_-, H_+. These sets can be characterized by

$$H_\pm = \{\tilde{x} \in \tilde{\Lambda},\ \exists z \in \tilde{\gamma},\ \mathrm{dist}\left(\tilde{\Phi}_t(\tilde{x}), \tilde{\Phi}_t(\tilde{z})\right) \leqslant C e^{-\beta|t|}, \quad \text{as} \quad t \to \pm\infty\},$$

or, equivalently

$$W^{\mathrm{s}}_{H_+} \cap W^{\mathrm{u}}_{H_-} \cap \tilde{\gamma} \neq \emptyset.$$

Remark 3.11. From dynamical systems theory, we know that the existence of a trans versal homoclinic connection implies the existence of infinitely many others, usually called secondary homoclinic connections. If the external potential verifies the hypotheses of Theorem 3.4 for any of these connections, there will be orbits of unbounded energy.

One can be even more precise. It is true that one can use the existence of several homoclinic connections and therefore several scattering maps to produce orbits that gain energy jumping through more than one of them at different times. We will not develop this idea in this paper since in the models we consider it does not weaken the sufficient conditions we obtain for the existence of unbounded orbits. Nevertheless, in the models considered in [DdlLS03,DdlLMS03], it leads to weaker conditions for the main results.

4. The problem with external potential

4.1. Introduction and overview

The goal of Section 4.2 is to show that, for high energy, the external potential is a small (and slow) perturbation of the extended flow $\tilde{\Phi}_t$ introduced in Section 3.5. As a first consequence, in Section 4.3, we will see that the manifold $\tilde{\Lambda}$, which is normally hyperbolic for the unperturbed flow, will persist for high energy when we consider the system with the external potential.

Furthermore, in Sections 4.3.5 and 4.3.6 we will show that this manifold is almost filled by $(d+1)$-dimensional invariant tori, which are close to the unperturbed ones $\mathcal{T}_E$ (see (25)). The gaps between those tori will be smaller than a negative power of the energy. Thanks to the fact that the perturbation is slow, provided that the Hamiltonian is differentiable enough, we can take the power to be as large as we want.

We note that the whiskered invariant tori $\mathcal{T}_E$ are full dimensional KAM tori when considered in the manifold $\tilde{\Lambda}$. Equivalently $\mathcal{T}^*_{E,B}$ given in (24) are full dimensional in Λ^* (as was already noticed in [Moe96], see also [Sor02] for further developments).

The transverse intersection of the stable and unstable manifolds of $\tilde{\Lambda}$ along $\tilde{\gamma}$ will persist for high enough energies. This will allow us to define $\tilde{S}$ and to compute it perturbatively in Section 4.4. Moreover, we will show in Lemma 4.33 that, under the non-degeneracy condition given in Theorem 3.4, the image of a torus by the perturbed scattering map intersects transversally other tori which are close enough. This will give us in Section 4.5 the existence of transversal heteroclinic connections between some perturbed tori. These heteroclinic orbits will give us an infinite transition path of tori with increasing energies.

The existence of orbits which follow finite transition paths is quite well known in the field of Arnol'd diffusion. In Section 4.5.4 we will present an argument that can deal with infinite paths and we will also pay attention to the behavior of the energy during the transitions. The argument follows closely the presentation in [DLS00].

Moreover, by performing the analysis with more care, we will show that there are orbits which visit the tori of the chain in an almost arbitrary order and hence the energy can make almost arbitrary excursions. We also remark that the energy of the orbits is well approximated by the energy of the visited tori, so that this symbolic description gives a very accurate picture of the evolution of the energy. In particular, there are orbits whose energy grows to infinity.

4.2. The scaled problem

To make precise the idea that the external potential is a slow and small perturbation of the geodesic flow for high energy, we scale the variables and the time. Thus, we pick a (large) number $E^* > 0$ and introduce

$$\varepsilon = 1/\sqrt{E^*}. \tag{30}$$

The Hamiltonian corresponding to (15) in the full symplectic space is

$$H^*(p, q, A, \theta) = v \cdot A + H_0(p, q) + U(q, \theta). \tag{31}$$

If we denote $\varepsilon p = \bar{p}$, $\varepsilon A = \bar{A}$ and consider the symplectic form $\bar{\Omega}^*$, given in local coordinates by $\bar{\Omega}^* = d\bar{p} \wedge dq + d\bar{A} \wedge d\theta = \varepsilon \Omega^*$, we see that $q, \bar{p}$ and $\theta, \bar{A}$ are conjugate variables with respect to $\bar{\Omega}^*$. We also introduce a new time $\bar{t} = t/\varepsilon$, and then the original equations for the Hamiltonian H^*:

$$\begin{aligned}
\frac{dp}{dt} &= -\frac{\partial H_0}{\partial q}(p, q) - \frac{\partial U}{\partial q}(q, \theta),\\
\frac{dq}{dt} &= \frac{\partial H_0}{\partial p}(p, q),\\
\frac{dA}{dt} &= -\frac{\partial U}{\partial \theta}(q, \theta),\\
\frac{d\theta}{dt} &= v
\end{aligned}$$

are equivalent to

$$\begin{aligned}
\frac{d\bar{p}}{d\bar{t}} &= -\frac{\partial H_0}{\partial q}(\bar{p}, q) - \varepsilon^2 \frac{\partial U}{\partial q}(q, \theta),\\
\frac{dq}{d\bar{t}} &= \frac{\partial H_0}{\partial p}(\bar{p}, q),\\
\frac{d\bar{A}}{d\bar{t}} &= -\varepsilon^2 \frac{\partial U}{\partial \theta}(q, \theta),\\
\frac{d\theta}{d\bar{t}} &= \varepsilon v,
\end{aligned} \tag{32}$$

which are Hamiltonian equations with respect to the symplectic form $\bar{\Omega}^*$, for the time $\bar{t}$, corresponding to the Hamiltonian

$$\begin{aligned}\bar{H}^*_\varepsilon(\bar{p}, q, \bar{A}, \theta) &:= \varepsilon\nu \cdot \bar{A} + H_0(\bar{p}, q) + \varepsilon^2 U(q, \theta)\\ &= \varepsilon\nu \cdot \bar{A} + \bar{H}_\varepsilon(\bar{p}, q, \bar{A}, \theta).\end{aligned}$$

From (32) we have $\theta = \theta_0 + \varepsilon\nu t$. Hence the flow of the Hamiltonian $\bar{H}^*_\varepsilon$ is a small and slow perturbation of the flow of Hamiltonian H_0.

Since $\bar{H}^*_\varepsilon(\bar{p}, q, \bar{A}, \theta) = \varepsilon^2 H^*(p, q, A, \theta)$ and

$$\bar{H}_\varepsilon(\bar{p}, q, \theta) = \varepsilon^2 H(p, q, \theta), \tag{33}$$

we introduce the notation

$$\overline{E} = \varepsilon^2 E = E/E_*. \tag{34}$$

For the first stages of our analysis, it suffices to analyze a fixed range in scaled energies $\bar{E}$ which we will fix arbitrarily to be $[1/2, 2]$. Our goal will be to establish that, for large enough E^*, given two KAM tori $\mathcal{T}_a, \mathcal{T}_b$ with energy $\bar{H}_\varepsilon$ close to $\overline{E} = 1/2$ and $\overline{E} = 2$, we can find a sequence of tori $\{\mathcal{T}_i\}_{i=1}^N$ such that $W^{\mathrm{u}}_{\mathcal{T}_i} \pitchfork W^{\mathrm{s}}_{\mathcal{T}_{i+1}}$, $\mathcal{T}_1 = \mathcal{T}_a$, $\mathcal{T}_N = \mathcal{T}_b$.

Once we have the existence of these finite paths, using that the result is true for arbitrarily large E^*, we will obtain, in Lemma 4.35, that we can get transition paths that transverse all the sufficiently large energies.

From now on, and until further notice, we will drop the bar from the rescaled variables since we will not work for a while on the original variables. Then, our Hamiltonian will be

$$\begin{aligned}H^*_\varepsilon(p, q, A, \theta) &= \varepsilon\nu \cdot A + H_0(p, q) + \varepsilon^2 U(q, \theta)\\ &= \varepsilon\nu \cdot A + H_\varepsilon(p, q, A, \theta).\end{aligned} \tag{35}$$

We will refer to the original variables as the physical variables if there is need to distinguish between them and the rescaled ones.

It is important to note that as S^*_0 was defined through geometric considerations it does not change when rescaled. Hence in the rescaled variables, we also have, as in (29)

$$S^*_0(J, \varphi, A, \theta) = (J, \varphi + \Delta, A, \theta). \tag{36}$$

Similarly, we can study the hyperbolic properties of $\tilde{\Lambda}$ (or Λ^*) under the rescaled flow. It is easy to note that the stable and unstable bundles do not change under rescaling of time, and that the exponential expansion rates $\beta_0\sqrt{2E}$ in (27) get multiplied by $\varepsilon = \frac{1}{\sqrt{2E^*}}$ becoming $\beta_0\sqrt{2\bar{E}}$. Hence, in the scaled variables, the exponential expansion rates are bounded between β_0 and $2\beta_0$.

4.3. The perturbed invariant manifold and study of the inner motion

The main goal of this Section is to show that, for high enough energies, the manifold $\tilde{\Lambda}$ of Proposition 3.6—and the manifold Λ^* of Proposition 3.7—persists, and that it is almost covered by KAM tori leaving only extremely small gaps.

Using that, for high energies, the potential is a small perturbation of the geodesic flow the theory of normally hyperbolic invariant manifolds implies that the manifold $\tilde{\Lambda}$ persists. In Theorem 4.1 we state the main consequences of the theory. Since the proof of persistence of the invariant manifold presents some peculiarities with respect to standard presentations, we have included a detailed presentation in Appendix B.

In Section 4.3.2 we will study the symplectic geometry of the perturbed manifold. In Section 4.3.3 we will introduce a system of coordinates on the perturbed manifold which will allow us to exhibit the motion on the perturbed manifold as a slow perturbation of an integrable system. In Section 4.3.5 we will take advantage of the fact that the perturbation induced by the external potential is slow and we will perform several steps of averaging. After averaging, the system will be an extremely small perturbation of an integrable system. In Section 4.3.6 we apply the KAM theorem to the averaged system and conclude that the perturbed manifold is almost covered by KAM tori except for very small gaps. Even if the quantitative version of the KAM theorem we use is rather straightforward and, we presume, well known to experts, we have included in Appendix A a proof based on [Zeh75,Zeh76a,Zeh76b], since we could not find a proof in the literature that covered the desired result.

In Section 4.3.7 we obtain some approximate expressions for the KAM tori. This will be used later in the subsequent discussion of existence of heteroclinic intersections. Note that the perturbation moves the tori, so that to discuss whether their asymptotic manifolds intersect, we need to take into account also the displacement of the tori.

4.3.1. Persistence of the invariant manifold

Using the hyperbolicity properties of the manifold $\tilde{\Lambda}$ for the extended flow $\tilde{\Phi}_t$ described in Section 3.2, we will apply some results of hyperbolic perturbation theory.

We note that in Hamiltonian (35), ε enters in two different ways, on one hand it is a perturbation parameter in the Hamiltonian and on the other hand $\varepsilon\nu$ is the frequency of the perturbing potential. To distinguish these two different rôles of ε, we find it more convenient to introduce the autonomous flow

$$\begin{aligned}
\dot{p} &= -\frac{\partial H_0}{\partial q}(p,q) - \delta\frac{\partial H_1}{\partial q}(p,q,\theta),\\
\dot{q} &= \frac{\partial H_0}{\partial p}(p,q) + \delta\frac{\partial H_1}{\partial p}(p,q,\theta),
\end{aligned}$$

$$\dot{\theta} = \frac{v}{\tau} \tag{37}$$

defined on the extended phase space $\mathbf{T}^*M \times \mathbb{T}^d$. This problem is equivalent to our original one if we set $\delta = \varepsilon^2$, $\tau = 1/\varepsilon$, and $H_1(p, q, \theta) = U(q, \theta)$. We will prove results for δ small enough, which are uniform on τ.

We denote the flow of (37) by

$$\tilde{\Phi}_{t,\tau,\delta}(p, q, \theta) = \left(\Gamma_{t,\tau,\delta}(p, q, \theta), \theta + \frac{v}{\tau}t\right). \tag{38}$$

Setting $\delta = 0$ in (37) we have, by Propositions 3.3 and 3.6, that

$$\tilde{\Lambda} = \Lambda \times \mathbb{T}^d \simeq [E_0, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \simeq [J_0, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d,$$

where $J_0 = \sqrt{2E_0}$ (see Section 3.6), is a manifold locally invariant for the flow.

The following theorem on persistence of normally hyperbolic invariant manifolds is proved in Appendix B. The proof requires some adaptation of the standard proof to deal with the non-compactness of $\tilde{\Lambda}$ and that the vector fields are not bounded. The main point is the fact that, by (7), the exponential rates of $\tilde{\Lambda}$ scale with the energy.

Theorem 4.1. *Assume that we have a system of equations as in* (37), *where the Hamiltonian $H = H_0 + \delta H_1$ is $\mathcal{C}^r$, $2 \leqslant r < \infty$, and H_0 satisfies the hypotheses of Theorem* 3.4. *Then, there exists a $\delta^* > 0$ and a $K > 0$, depending only on the $\mathcal{C}^2$ norm of H_0, H_1, and the $\mathcal{C}^1$ properties of the unperturbed manifold $\tilde{\Lambda}$, such that for $|\delta| < \delta^*$, there is a $\mathcal{C}^{r-1}$ function*

$$\tilde{\mathcal{F}} : [J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \times (-\delta^*, \delta^*) \longrightarrow \mathbf{T}^*M \times \mathbb{T}^d$$

which is of the form $\tilde{\mathcal{F}} = (\mathcal{F}, \mathrm{Id}_{\mathbb{T}^d})$, with

$$\mathcal{F} : [J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \times (-\delta^*, \delta^*) \longrightarrow \mathbf{T}^*M$$

such that the manifold

$$\begin{aligned} \tilde{\Lambda}_{\tau,\delta} &= \tilde{\mathcal{F}}\left([J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \times \{\delta\}\right) \\ &= \mathcal{F}\left([J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \times \{\delta\}\right) \times \mathbb{T}^d \end{aligned} \tag{39}$$

is locally invariant for the flow of (37) *and the manifold $\tilde{\Lambda}_{\tau,\delta}$ is δ-close to $\tilde{\Lambda}_{\tau,0} = \tilde{\Lambda}$ in the $\mathcal{C}^{r-2}$ sense. Moreover, $\tilde{\Lambda}_{\tau,\delta}$ is a normally hyperbolic invariant manifold.*

Furthermore, $\mathcal{F}$ verifies

$$\mathcal{F}(J, \varphi, \theta, 0) = \Lambda_E\left(\frac{\varphi}{\sqrt{2E}}\right), \quad J = \sqrt{2E}. \tag{40}$$

We can find a $\mathcal{C}^{r-1}$ function

$$\tilde{\mathcal{F}}^{\mathrm{s}} : [J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \times [0, \infty)^{(n-1)} \times (-\delta^*, \delta^*) \longrightarrow \mathbf{T}^*M \times \mathbb{T}^d$$

of the form $\tilde{\mathcal{F}}^{\mathrm{s}} = (\mathcal{F}^{\mathrm{s}}, \mathrm{Id}_{\mathbb{T}^d})$, with

$$\mathcal{F}^{\mathrm{s}} : [J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \times [0, \infty)^{(n-1)} \times (-\delta^*, \delta^*) \longrightarrow \mathbf{T}^*M$$

such that the (local) stable invariant manifold of $\tilde{\Lambda}_{\tau,\delta}$ takes the form

$$W^{\mathrm{s,loc}}_{\tilde{\Lambda}_{\tau,\delta}} = \tilde{\mathcal{F}}^{\mathrm{s}}\left([J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{T}^d \times [0, \infty)^{(n-1)} \times \{\delta\}\right). \tag{41}$$

If $\tilde{x} = \tilde{\mathcal{F}}(J, \varphi, \theta, \delta) = (\mathcal{F}(J, \varphi, \theta, \delta), \theta) \in \tilde{\Lambda}_{\tau,\delta}$, then

$$W^{\mathrm{s,loc}}_{\tilde{x}} = \tilde{\mathcal{F}}^{\mathrm{s}}(\{J\} \times \{\varphi\} \times \{\theta\} \times [0, \infty)^{(n-1)} \times \{\delta\}).$$

Therefore $W^{\mathrm{s,loc}}_{\tilde{\Lambda}_{\tau,\delta}}$ is δ-close to $W^{\mathrm{s,loc}}_{\tilde{\Lambda}}$ in the $\mathcal{C}^{r-2}$ sense.

Analogous results hold for the (local) unstable manifold.

To maintain a symplectic structure, it is convenient to apply the normal hyperbolicity theory to the full Hamiltonian $H^*_{\tau,\delta}$.

Theorem 4.2. *If we consider the Hamiltonian equations associated to the Hamiltonian $H^*_{\tau,\delta} = \frac{\nu}{\tau} \cdot A + H_0 + \delta H_1$, then there exists $\delta^* > 0$ and $K > 0$, depending only on the $\mathcal{C}^2$ norm of H_0, H_1, and the $\mathcal{C}^1$ properties of the unperturbed manifold Λ^*, such that there is a $\mathcal{C}^{r-1}$ function*

$$\mathcal{F}^* : [J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{R}^d \times \mathbb{T}^d \times (-\delta^*, \delta^*) \longrightarrow \mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$$

such that

$$\Lambda^*_{\tau,\delta} = \mathcal{F}^*\left([J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{R}^d \times \mathbb{T}^d \times \{\delta\}\right) \tag{42}$$

is locally invariant for the Hamiltonian flow associated to $H^*_{\tau,\delta}$. Therefore, $\Lambda^*_{\tau,\delta}$ is δ-close to $\Lambda^*_{\tau,0} = \Lambda^*$ in the $\mathcal{C}^{r-2}$ sense.

As the Hamiltonian $H^*_{\tau,\delta}$ *is quasi-periodic, we have that* $\mathcal{F}^* = (\mathcal{F}, \mathcal{A}, \mathrm{Id}_{\mathbb{T}^d})$, *with* $\mathcal{A}(J, \varphi, B, \theta, 0) = B$, *and that, if* $x^* \in \Lambda^*_{\tau,\delta}$ *then we write in coordinates*

$$x^* = \mathcal{F}^*(J, \varphi, B, \theta, \delta) = (\mathcal{F}(J, \varphi, \theta, \delta), \mathcal{A}(J, \varphi, B, \theta, \delta), \theta),$$

and

$$\tilde{x} = \tilde{\mathcal{F}}(J, \varphi, \theta, \delta) = (\mathcal{F}(J, \varphi, \theta, \delta), \theta) \in \tilde{\Lambda}_{\tau,\delta}.$$

Then, the manifold $\tilde{\Lambda}_{\tau,\delta}$ *is the projection in the extended phase space of the manifold* $\Lambda^*_{\tau,\delta}$.

Moreover, $\Lambda^*_{\tau,\delta}$ *is a normally hyperbolic invariant manifold. We can find a* $\mathcal{C}^{r-1}$ *function*

$$\mathcal{F}^{*\,\mathrm{s}} : [J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{R}^d \times \mathbb{T}^d \times [0, \infty)^{(n-1)} \times (-\delta^*, \delta^*) \longrightarrow \mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$$

such that its (*local*) *stable invariant manifold takes the form*

$$W^{\mathrm{s,loc}}_{\Lambda^*_{\tau,\delta}} = \mathcal{F}^{*\,\mathrm{s}}\left([J_0 + K\delta, \infty) \times \mathbb{T}^1 \times \mathbb{R}^d \times \mathbb{T}^d \times [0, \infty)^{(n-1)} \times \{\delta\}\right). \tag{43}$$

If $x^* = \mathcal{F}^*(J, \varphi, B, \theta, \delta) = (\mathcal{F}(J, \varphi, \theta, \delta), \mathcal{A}(J, \varphi, B, \theta, \delta), \theta) \in \Lambda^*_{\tau,\delta}$, *then*

$$W^{\mathrm{s,loc}}_{x^*} = \mathcal{F}^{*\,\mathrm{s}}(\{J\} \times \{\varphi\} \times \{B\} \times \{\theta\} \times [0, \infty)^{(n-1)} \times \{\delta\}).$$

Therefore $W^{\mathrm{s,loc}}_{\Lambda^*_{\tau,\delta}}$ *is* δ*-close to* $W^{\mathrm{s,loc}}_{\Lambda^*}$ *in the* $\mathcal{C}^{r-2}$ *sense. Analogous results hold for the* (*local*) *unstable manifold.*

Remark 4.3. We emphasize that, since δ^* depends only on the $\mathcal{C}^2$ norm of H_0 and H_1, and H_1 depends periodically on θ, the value of δ^* can be chosen independently of τ. In particular, we obtain results for $\delta = \varepsilon^2$, $\tau = 1/\varepsilon$.

Remark 4.4. Since $W^{\mathrm{s}}_{\tilde{\Lambda}}$, $W^{\mathrm{u}}_{\tilde{\Lambda}}$ are transversal at $\tilde{\gamma} \subset W^{\mathrm{s}}_{\tilde{\Lambda}} \pitchfork W^{\mathrm{u}}_{\tilde{\Lambda}}$, we see that there exists a locally unique $\tilde{\gamma}_{\tau,\delta}$ which is δ-close to $\tilde{\gamma}$ in the $\mathcal{C}^{r-2}$ sense, such that $\tilde{\gamma}_{\tau,\delta} \subset W^{\mathrm{s}}_{\tilde{\Lambda}_{\tau,\delta}} \pitchfork W^{\mathrm{u}}_{\tilde{\Lambda}_{\tau,\delta}}$.

Notation 4.5. From now on, we are going to fix our attention on the case $\delta = \varepsilon^2$ and $\tau = 1/\varepsilon$, and we will call $\tilde{\Lambda}_\varepsilon = \tilde{\Lambda}_{1/\varepsilon,\varepsilon^2}$, $\tilde{\gamma}_\varepsilon = \tilde{\gamma}_{1/\varepsilon,\varepsilon^2}$, $\tilde{\Phi}_{t,\varepsilon} = \tilde{\Phi}_{t,1/\varepsilon,\varepsilon^2}$. The same simplified notation for the full space with Λ^*_ε,γ^*_ε, $\Phi^*_{t,\varepsilon}$.

4.3.2. Symplectic geometry on the invariant manifold Λ^*_ε

In order to introduce a Hamiltonian flow in the perturbed invariant manifold Λ^*_ε we start by investigating first its symplectic geometry and the symplectic properties of the full symplectic flow restricted to it. In subsequent sections, we will find how these properties can be expressed in a convenient system of coordinates.

By Theorem 4.2, we have that $\Lambda^*_\varepsilon = \cup_{(\theta,B)\in\mathbb{T}^d\times\mathbb{R}^d}\Lambda^{*,\theta,B}_\varepsilon \times \{\theta\}$, where

$$\Lambda^{*,\theta,B}_\varepsilon = (\mathcal{F},\mathcal{A})\left([J_0+K\varepsilon^2,\infty)\times\mathbb{T}^1\times\{\theta\}\times\{\varepsilon^2\}\right).$$

Note that all the two-dimensional manifolds $\Lambda^{*,\theta,B}_\varepsilon$ are ε^2-close to $\Lambda^\theta_0\times\{B\}=\Lambda\times\{B\}$ in the $\mathcal{C}^{r-2}$ sense. The parameters θ, B both d-dimensional. The manifold $\Lambda^{*,\theta,B}_\varepsilon$ is $2+2d$ dimensional.

For every θ, we define $\Omega^\theta_\varepsilon$ to be the two-form obtained by restricting the symplectic form Ω to $\Lambda^\theta_\varepsilon$.

Since $\Lambda^\theta_\varepsilon$ is ε^2-close to Λ and $\Omega|_\Lambda$ is non-degenerate, we obtain that $\Omega^\theta_\varepsilon$ is non-degenerate. Hence $\Omega^\theta_\varepsilon$ is an exact symplectic form on $\Lambda^\theta_\varepsilon$.

Indeed, if we denote by $\alpha^\theta_\varepsilon=\alpha|_{\Lambda^\theta_\varepsilon}$, where α is the primitive of Ω introduced in (5) and by d^θ_ε the exterior differential on $\Lambda^\theta_\varepsilon$, we have

$$\Omega^\theta_\varepsilon = d^\theta_\varepsilon\alpha^\theta_\varepsilon.$$

Since the coordinate θ moves at constant velocity εv, we have that the extended flow satisfies

$$\tilde\Phi_{t,\varepsilon}\left(\Lambda^\theta_\varepsilon,\theta\right) = (\Lambda^{\theta+\varepsilon vt}_\varepsilon, \theta+\varepsilon vt). \tag{44}$$

Since the flow $\Phi^*_{t,\varepsilon}$ preserves the form Ω^* and moreover is exact, we have

$$\begin{aligned}(\Phi^*_{t,\varepsilon})_*\Omega^* &= \Omega^*,\\ (\Phi^*_{t,\varepsilon})_*\alpha^* &= \alpha^* + dS_{t,\varepsilon}.\end{aligned}$$

Since by (38) $\tilde\Phi_{t,\varepsilon} = (\Gamma_{t,\varepsilon}(\cdot,\theta),\theta+\varepsilon vt)$, it follows by restriction that $\Gamma_{t,\varepsilon}(\cdot,\theta)$ are exact symplectic transformations from $\Lambda^\theta_\varepsilon$ to $\Lambda^{\theta+\varepsilon vt}_\varepsilon$ endowed with the exact symplectic forms $\Omega^\theta_\varepsilon$.

4.3.3. A system of coordinates on Λ^*_ε

Theorem 4.2 gives a system of coordinates on the manifold Λ^*_ε by pushing the system of coordinates (J,φ,B,θ) on Λ^* through the function $\mathcal{F}^*$. Since the mapping $\mathcal{F}^*$ is not unique, we can take advantage of the non-uniqueness to obtain a parameterization with extra features that will make subsequent analysis more convenient. We will choose the system of coordinates in such a way that the standard representation of the symplectic form holds. This will allow us to use the customary formulas of perturbation theory. (Of course, if we had developed the perturbation theory in a more coordinate-independent form, this preparation of the system of coordinates would not be necessary.)

Proposition 4.6. *For $r > 2$, it is possible to find a $\mathcal{C}^{r-2}$ family $\mathcal{F}^*$ satisfying the conclusions of Theorem 4.2 in such a way that, moreover*

$$\Omega_\varepsilon^\theta = dJ \wedge d\varphi,$$

where J, φ are the push forward by $\mathcal{F}$ of the coordinates on $[J_0 + \infty) \times \mathbb{T}^1$.

Proof. We note that the $\mathcal{F}$ claimed in Theorem 4.2 is such that $\Omega_\varepsilon^\theta$ is $O(\varepsilon^2)$ close to $dJ \wedge d\varphi$ in the $\mathcal{C}^{r-2}$ sense. Then, applying a global version of Darboux theorem [Wei77] we obtain a $\mathcal{C}^{r-2}$ change of coordinates $C_\varepsilon^\theta : \Lambda_\varepsilon^\theta \to \Lambda$, such that in the new coordinates, that we will also denote by J, φ, we have: $\Omega_\varepsilon^\theta = dJ \wedge d\varphi$.

Moreover, it is possible to arrange that the change of variables is $\mathcal{C}^{r-2}$ in the variables (J, φ, θ). This amounts to a parameterized version of the Darboux theorem [BLW96] that shows that these transformations depend smoothly on parameters. The case that $\theta \in \mathbb{T}^1$ can be found in [DLS00] and a more explicit construction is in [DdlLMS03]

The idea of the proof—we refer to [BLW96] for full details—is that, using the standard deformation method for the proof of Darboux theorem, we obtain a family of symplectic forms Ω_δ such that Ω_0 is the original form and Ω_1 is the target form. The equation for f_δ given by $f_\delta^* \Omega_0 = \Omega_\delta$ is equivalent to the fact that f_δ satisfies a differential equation with independent variable δ. It is straightforward to check that when we take Ω_1 depending smoothly on another parameter μ the differential equation we derive depends differentiably on μ. The desired regularity with respect to parameters is a straightforward consequence of the results on dependence on parameters for ODEs. $\square$

Using the parameterization of Λ_ε^* given by the family $\mathcal{F}^* = (\mathcal{F}, \mathcal{A}, \mathrm{Id}_{\mathbb{T}^d})$ provided by Proposition 4.6, we have a system of coordinates (J, φ, B, θ) on Λ_ε^*. The flow $\Psi_{t,\varepsilon}^*$ obtained by restricting the full symplectic flow $\Phi_{t,\varepsilon}^*$, is given, in these coordinates, by the relation

$$\mathcal{F}^*(\Psi_{t,\varepsilon}^*(J, \varphi, B, \theta), \varepsilon^2) = \Phi_{t,\varepsilon}^*(\mathcal{F}^*(J, \varphi, B, \theta, \varepsilon^2)).$$

Since the transformations are exact, we have

$$(\Psi_{t,\varepsilon}^*)_* \alpha^* = \alpha^* + dS_{t,\varepsilon}.$$

Hence, the flow $\Psi_{t,\varepsilon}^*$ is a Hamiltonian flow. Moreover $\Psi_{t,\varepsilon}^*$ is a quasi-periodic Hamiltonian flow with $\dot\theta = \varepsilon\nu$, and this allows us to define $\tilde\Psi_{t,\varepsilon}$ on $\tilde\Lambda_\varepsilon$. We will think of $\tilde\Psi_{t,\varepsilon}$

as the representation in the coordinates (J, φ, θ) of the extended flow $\tilde{\Phi}_{t,\varepsilon}$ restricted to the invariant manifold $\tilde{\Lambda}_\varepsilon$.

4.3.4. The Hamiltonian flow in Λ^*_ε

Now we start to compute more explicitly the Hamiltonian of the flow restricted to the invariant manifold Λ^*_ε.

We will write $\varepsilon v \cdot B + k_\varepsilon(J, \varphi, \theta)$ to denote the Hamiltonian generating $\Psi^*_{t,\varepsilon}$ with respect to the standard symplectic form.

Since we have generated the system of coordinates by changes of variables that transform the symplectic form in the standard one, it is easy to see that $\varepsilon v \cdot B + k_\varepsilon(J, \varphi, \theta)$ will be the push-forward by $\mathcal{F}^*$, of the Hamiltonian $\varepsilon v \cdot A + H_\varepsilon(p, q, \theta)$ given in (35). In particular, $\varepsilon v \cdot B + k_\varepsilon(J, \varphi, \theta)$ is $\mathcal{C}^{r-2}$ and $\Psi^*_{t,\varepsilon}$ is a $\mathcal{C}^{r-3}$ flow, quasi-periodic, and it is a small perturbation of size $O(\varepsilon^2)$ of the constant flow $\dot{J} = 0$, $\dot{\varphi} = J$, $\dot{B} = 0$, $\dot{\theta} = \varepsilon v$ of the unperturbed Hamiltonian $\varepsilon v \cdot B + \frac{1}{2}J^2$ in Λ^*.

Therefore, $k_\varepsilon(J, \varphi, \theta) = \frac{1}{2}J^2 + \varepsilon^2 k^1_\varepsilon(J, \varphi, \theta)$, so that the Hamiltonian in Λ^*_ε is given by

$$\varepsilon v \cdot B + k_\varepsilon(J, \varphi, \theta) = \varepsilon v \cdot B + \frac{1}{2}J^2 + \varepsilon^2 k^1_\varepsilon(J, \varphi, \theta). \tag{45}$$

As we saw in Section 3.6, the unperturbed Hamiltonian has a $(d + 1)$-parametric family of invariant tori $\mathcal{T}^*_{E,B}$ given in (24) which fill the invariant manifold Λ^*. Looking at the extended phase space we saw that all the tori obtained by taking different values of B project into the same torus $\mathcal{T}_E$ in $\tilde{\Lambda}$ given in (25), obtaining a one-parameter family of tori in $\tilde{\Lambda}$.

Our next goal is to study $\tilde{\Psi}_{t,\varepsilon}$, the extended flow of the perturbed Hamiltonian (45) restricted to $\tilde{\Lambda}_\varepsilon$, and show that some of the tori $\mathcal{T}_E$ persist and that the gaps between them are very small.

A direct application of KAM theory to (45) will establish only the existence of tori with gaps significantly bigger than what would be desirable for our later purposes. Therefore, we will take advantage of the fact that the perturbation is slow in the angles θ and we will apply the averaging method to eliminate the fast angle φ. Applying the KAM theorem to the averaged system we will establish the existence of tori that are much closer that what a direct application of KAM Theory to (45) would yield. Note that the fact that one the perturbation is slow with respect to the unperturbed system implies that the only resonaces possible happen only in very high orders of the perturbation theory. In the following Sections 4.3.5 and 4.3.6 we will implement the averaging procedure and the KAM theorem.

4.3.5. Averaging theory

The following result is a version of the classical averaging theorem that allows to change variables on a system perturbed by a slowly evolving term and reduce it to a system under a much smaller perturbation.

Theorem 4.7. *Let $\varepsilon\nu \cdot B + k_\varepsilon(J, \varphi, \theta)$ as in* (45) *be a $\mathcal{C}^n$ family of Hamiltonians,* 1*-periodic on φ and on $\theta = (\theta_1, \ldots, \theta_d)$.*

Then, for any $0 < m < n$, there exists a canonical change of variables $(J, \varphi, B, \theta) \mapsto (\hat{J}, \hat{\varphi}, \hat{B}, \theta)$, 1*-periodic in φ and θ, which is ε^2-close to the identity in the $\mathcal{C}^{n-m}$ topology, such that it transforms the Hamiltonian system of Hamiltonian* (45) *into a Hamiltonian system of Hamiltonian $\varepsilon\nu \cdot \hat{B} + K_\varepsilon(\hat{J}, \hat{\varphi}, \theta)$. This new Hamiltonian is a $\mathcal{C}^{n-m}$ function with*

$$K_\varepsilon(\hat{J}, \hat{\varphi}, \theta) = K_\varepsilon^0(\hat{J}, \theta) + \varepsilon^{m+1} K_\varepsilon^1(\hat{J}, \hat{\varphi}, \theta),$$

where $K_\varepsilon^0(\hat{J}, \theta) = \frac{1}{2}\hat{J}^2 + O_{\mathcal{C}^{n-m}}(\varepsilon^2)$, and the notation $O_{\mathcal{C}^l}(\varepsilon)$ means a function whose $\mathcal{C}^l$ norm is $O(\varepsilon)$.

Proof. It is standard, and it is carried out in detail for the case of periodic perturbations in [DLS00, Theorem 4.6]. See also [AKN88,LM88].

The only change needed to transform the write up of the one-dimensional theorem in [DLS00] to the case considered here is that the one-dimensional variable εs of the case of periodic perturbations has to be replaced by the d-dimensional variable θ. Note that in both cases the perturbation is slow because $\dot{\theta} = \varepsilon\nu$ and, for the periodic case $(\varepsilon s)\dot{=}\varepsilon$. All the arguments and calculations in [DLS00] go through without any other change.

It is important to note that the changes of variables are found to eliminate the variable φ—which continues to be one dimensional. Therefore, the homological equation that appear in the steps of the averaging procedure can be solved by quadratures. (In particular, there is no need to solve any small divisors equation.) In other words, the phases θ which change from those in [DLS00] enter only as parameters in all the transformations required in the proof. □

4.3.6. K.A.M. theory

By Theorem 4.7, with $n = r - 2$, $0 < m < r - 2$, Hamiltonian (45) is given, in these new averaged variables, by

$$\varepsilon\nu \cdot \hat{B} + K_\varepsilon(\hat{J}, \hat{\varphi}, \theta) = \varepsilon\nu \cdot \hat{B} + K_\varepsilon^0(\hat{J}, \theta) + \varepsilon^{m+1} K_\varepsilon^1(\hat{J}, \hat{\varphi}, \theta), \tag{46}$$

where

$$K_\varepsilon^0(\hat{J}, \theta) = \tfrac{1}{2}\hat{J}^2 + O_{\mathcal{C}^{r-2-m}}(\varepsilon^2) \tag{47}$$

is a $\mathcal{C}^{r-2-m}$ function.

We note that the first part of (46), namely $\varepsilon\nu \cdot \hat{B} + K_\varepsilon^0(\hat{J}, \theta)$, is integrable, but it is not written in the classical action-angle variables. The next step will be to apply the KAM theorem to Hamiltonian (46). To this end, it is convenient to introduce action-angle variables $(\hat{J}, \psi, \mathcal{B}, \theta)$ in place of the variables $(\hat{J}, \hat{\varphi}, \hat{B}, \theta)$, that is, the variables $\hat{J}$ and θ remain unchanged.

We introduce the generating function $\psi \hat{J} + \theta \hat{B} + \varepsilon \chi(\hat{J}, \theta; \varepsilon)$, where $\chi(\hat{J}, \theta; \varepsilon)$ is the solution of the classical small divisors equation:

$$\varepsilon^2 v \cdot \partial_\theta \chi(\hat{J}, \theta; \varepsilon) = K^0_\varepsilon(\hat{J}, \theta) - < K^0_\varepsilon(\hat{J}, \cdot) > \tag{48}$$

with

$$< K^0_\varepsilon(\hat{J}, \cdot) >= \int_{\mathbb{T}^d} K^0_\varepsilon(\hat{J}, \theta) d\theta.$$

Note that, by (47), $K^0_\varepsilon(\hat{J}, \theta) - < K^0_\varepsilon(\hat{J}, \cdot) >= O_{\mathcal{C}^{r-2-m}}(\varepsilon^2)$.

Using the results of [Rus75] on solutions of small divisors equations, reproduced in Lemma A.23 for the analytic case, the solution χ of the homological equation has a $\mathcal{C}^{r-2-m-\tau}$ norm bounded independently of ε, where τ is the diophantine exponent of v (see (3)).

The change of variables generated by χ is

$$\begin{aligned} \hat{J} &= \hat{J}, \\ \mathcal{B} &= \hat{B} + \varepsilon \frac{\partial \chi}{\partial \theta}(\hat{J}, \theta; \varepsilon), \\ \psi &= \hat{\varphi} - \varepsilon \frac{\partial \chi}{\partial \hat{J}}(\hat{J}, \theta; \varepsilon), \\ \theta &= \theta. \end{aligned} \tag{49}$$

Note that this change of variables is $O_{\mathcal{C}^{r-3-m-\tau}}(\varepsilon)$ close to the identity.

With the change of variables (49) Hamiltonian (46) transforms into

$$\varepsilon v \cdot \mathcal{B} + < K^0_\varepsilon(\hat{J}, \cdot) > + \varepsilon^{m+1} \bar{K}^1_\varepsilon(\hat{J}, \psi, \theta), \tag{50}$$

where $\varepsilon^{m+1} \bar{K}^1_\varepsilon$ is a $\mathcal{C}^{r-3-m-\tau}$ function.

We note that our Hamiltonian is finitely differentiable. The KAM theorem which we will use is patterned after the KAM theorem in [Zeh76a], which requires that the unperturbed Hamiltonian is analytic. As suggested in [Zeh76a, p. 70], it is natural to separate from the integrable part a polynomial part. We apply this suggestion to our case.

For any value $\hat{J} = I_0$ the unperturbed Hamiltonian $\varepsilon v \cdot \mathcal{B} + < K^0_\varepsilon(\hat{J}, \cdot) >$ leaves invariant a torus of frequency $(\omega, \varepsilon v)$ in the extended phase space, where, by (47)

$$\omega = \omega(I_0) = \frac{\partial}{\partial \hat{J}} < K^0_\varepsilon(\hat{J}, \cdot) >_{|\hat{J}=I_0} = I_0 + O_{\mathcal{C}^{r-3-m-\tau}}(\varepsilon^2). \tag{51}$$

If we change the origin of $\hat{J}$ through $I = \hat{J} - I_0$, to translate the torus we are interested in to the origin, the unperturbed Hamiltonian can be written, around this torus, as

$$\mathcal{H}_0(I, \mathcal{B}; \varepsilon) + I^{2m+2} R_2(I; \varepsilon),$$

where

$$\mathcal{H}_0(I, \mathcal{B}; \varepsilon) = \varepsilon v \cdot \mathcal{B} + c_0 + \omega I + Q_0 I^2 + c_3 I^3 + \cdots + c_{2m+1} I^{2m+1}, \tag{52}$$

contains the $2m+1$-degree Taylor polynomial of $< K^0_\varepsilon(I_0 + I, \cdot) >$ around $I = 0$,

$$Q_0 = \frac{\partial^2}{\partial \hat{J}^2} < K^0_\varepsilon(\hat{J}, \cdot) >_{|\hat{J}=I_0} = 1 + O_{\mathcal{C}^{r-4-m-\tau}}(\varepsilon^2),$$

and $I^{2m+2} R_2(I; \varepsilon)$ is a $\mathcal{C}^{r-3m-4}$ function.

In these variables, Hamiltonian (46) reads as a perturbation of (52):

$$\mathcal{H}(I, \psi, \mathcal{B}, \theta; \varepsilon) = \mathcal{H}_0(I, \mathcal{B}; \varepsilon) + \varepsilon^{m+1} R_1(I, \psi, \mathcal{B}, \theta; \varepsilon) + I^{2m+2} R_2(I; \varepsilon), \tag{53}$$

where $\varepsilon^{m+1} R_1(I, \psi, \mathcal{B}, \theta; \varepsilon)$ is a $\mathcal{C}^{r-m-3-\tau}$ function.

We are now in a position to apply a KAM theorem. We will show that, for ε small enough, Hamiltonian (53) has an invariant torus with frequency $(\omega, \varepsilon v)$, provided this frequency satisfies some Diophantine conditions. We will see that the gaps between such Diophantine frequencies are of size $O(\varepsilon^{(m+1)/4})$.

In terms of the original variables (J, φ, B, θ) inside Λ^*_ε, Hamiltonian (45) will have invariant tori with gaps between them also of size $O(\varepsilon^{(m+1)/4})$.

Even if the component ω of the frequency $(\omega, \varepsilon v)$ of the torus is to be chosen depending on the initial conditions, the last components are determined by those of the external forcing. Hence, we will not only require that the external frequency εv is Diophantine, but also that the frequency ω is Diophantine relative to εv. Precise definitions about relatively Diophantine numbers and on their abundance are collected in Definition A.1, Propositions A.2 and A.3.

Theorem 4.8. *Let* $v \in \mathcal{D}_d(\kappa, \tau)$, $\omega \in \mathcal{D}_n(\varepsilon v, \tilde{\kappa}, \tilde{\tau})$ (*see Definition* A.1), *with* $0 < \tilde{\kappa} \leqslant \varepsilon\kappa$, $\tilde{\tau} \geqslant \tau > 0$.

Denote $\gamma = 2\tilde{\tau} + 1$ *and assume* $l \geqslant 2\gamma + 3 = 4\tilde{\tau} + 5$.

Let $\mathcal{H}_0$ *be a polynomial of the form*:

$$\mathcal{H}_0(I, \mathcal{B}) = c_0 + \varepsilon v \cdot \mathcal{B} + \omega \cdot I + \tfrac{1}{2} I^\top Q_0 I + R_0(I), \tag{54}$$

where $R_0(I) = O(|I|^3)$ *and*

$$B_s = \{(I, \psi, \mathcal{B}, \theta) \in \mathbb{R}^n \times \mathbb{T}^n \times \mathbb{R}^d \times \mathbb{T}^d, \ |I| \leqslant s, \ |\mathcal{B}| \leqslant s\}, \ s > 0$$

and assume that

$$e := \|\mathcal{H}_0\|_{\mathcal{C}^l(B_s)} < \infty,$$
$$\rho := \left\| < Q_0 >^{-1} \right\| < \infty.$$

(*The constant* ρ *will be referred as the twist constant.*)

Then, there is a constant $C > 0$ *depending on* $e, \rho, \tilde{\tau}, \tau, l$*—but not on* $\tilde{\kappa}$ *or on* κ*—such that, given any* $\mathcal{C}^l$*function* $\mathcal{H}$ *of the form*

$$\mathcal{H}(I, \psi, \mathcal{B}, \theta) = \varepsilon v \cdot \mathcal{B} + E(I, \psi, \theta) \tag{55}$$

satisfying

$$\|\mathcal{H} - \mathcal{H}_0\|_{\mathcal{C}^l(B_s)} \leqslant C\rho^{-2}\tilde{\kappa}^4, \tag{56}$$

where s *is large enough so that* $s - C\tilde{\kappa}^2 > 0$, *we have*

(1) *There is a symplectic diffeomorphism* $F \in \mathcal{S}_1 \cap \mathcal{C}^{l-\gamma-1}(B_{\tilde{s}})$, *where* $\tilde{s} = s - C\tilde{\kappa}^2$ *and* $\mathcal{S}_1$ *is a subset of the space of canonical diffeomorphisms specified in Definition* A.9.

(2) *The diffeomorphism* F *is of the form*

$$F(I, \psi, \mathcal{B}, \theta) = (f(I, \psi, \mathcal{B}, \theta), \theta). \tag{57}$$

(3) *The transformed Hamiltonian is of the form*

$$\mathcal{H} \circ F(I, \psi, \mathcal{B}, \theta) = c + \varepsilon v \cdot \mathcal{B} + \omega \cdot I + \tfrac{1}{2} I^\top Q(\psi, \theta) I + R(I, \psi, \theta), \tag{58}$$

where $R = O(I^3)$.

(4) *The following estimates hold*:

$$\|F - \mathrm{Id}\|_{\mathcal{C}^{l-\gamma-1}(B_{\tilde{s}})} \leqslant C\tilde{\kappa}^{-2}\rho \|\mathcal{H} - \mathcal{H}_0\|_{\mathcal{C}^l(B_s)}, \tag{59}$$

$$\|Q - Q_0\|_{\mathcal{C}^{l-\gamma-1}(B_{\tilde{s}})} \leqslant C\tilde{\kappa}^{-2}\rho \|\mathcal{H} - \mathcal{H}_0\|_{\mathcal{C}^l(B_s)}. \tag{60}$$

The proof of Theorem 4.8 is given in Appendix A.

Remark 4.9. Theorem 4.8 is not optimal in two aspects. First, it requires a loss in the derivatives heavier than needed and consequently higher differentiability assumptions. Second, the values obtained for the exponent of $\tilde{\kappa}$ in (56), (59) and (60) are twice the optimal ones. In Remark A.25 we discuss in more detail the optimality. For the purposes of this paper, the only effect of lack of optimality is that we assume more differentiability than needed and, of course, that the initial energy of the orbits is higher than needed.

Applying Theorem 4.8 to Hamiltonian (53), we obtain invariant tori for Hamiltonian (46) and consequently for Hamiltonian (35), for

$$\bar{l} := \min(r - m - 3 - \tau, r - 3m - 4) \geqslant l_0, \quad l_0 = 4\tilde{\tau} + 5, \tag{61}$$

or equivalently for $r \geqslant \max(m + 3 + \tau + l_0, 3m + 4 + l_0)$.

There are two properties of the KAM tori for (46) that will be important for future analysis. One is that the tori are close (in a $\mathcal{C}^1$ topology) to the tori obtained fixing the value of the actions to appropriate values (59). The second property is that the tori leave very small gaps between them. The precise meaning of "leaving small gaps" is that in the extended phase space we can find tori whose $\mathcal{C}^1$ distance will be bounded by a (high) power of ε.

It is clear that, to accomplish this proximity, we have to consider an increasing number of tori for smaller ε. That is, we will need to consider tori of frequencies $(\omega, \varepsilon\nu)$ with worse Diophantine properties as we consider $\varepsilon \to 0$. Since we will also need to obtain explicit approximations of these tori, it will be important to keep track of the size of the Diophantine constants allowed.

Proposition 4.10. *Assume that r is big enough so that* (61) *holds for $\tilde{\tau} = \tau + 2$, ε is sufficiently small and $m > 3$.*

The KAM tori produced applying Theorem 4.8 *to Hamiltonian* (46) *are $O_{\mathcal{C}^{\bar{l}-6-2\tau}}(\varepsilon^{(m+1)/2})$-close to surfaces of the form $\{I_0\} \times \mathbb{T}^{1+d}$.*

Given any of these KAM tori, we can find another KAM torus which is $O_{\mathcal{C}^{\bar{l}-6-2\tau}}(\varepsilon^{(m+1)/4})$-close to it. In particular, it is $O_{\mathcal{C}^1}(\varepsilon^{(m+1)/4})$-close to it.

Remark 4.11. Due to the fact that the actions A are not dynamical variables, the torus of frequency $(\omega, \varepsilon\nu)$ is not unique in the full symplectic space. We can always obtain a family of tori $\mathcal{T}^*_{\omega,B}(\varepsilon)$ with the same frequency $(\omega, \varepsilon\nu)$ by translating by B in the A-direction the torus produced by Theorem 4.8.

Of course, all these tori $\mathcal{T}^*_{\omega,B}(\varepsilon) \subset \Lambda^*_\varepsilon$ project in the same torus $\mathcal{T}_\omega(\varepsilon) \subset \tilde{\Lambda}_\varepsilon$ when we consider the extended phase space.

Therefore, when we speak about gaps between tori we will refer to the gaps between the tori $\mathcal{T}_\omega(\varepsilon)$ of the extended system.

Note also that the diffusion we establish in this paper takes place also in the extended phase space.

Proof. Recall that the external frequency v verifies that $v \in \mathcal{D}(\kappa, \tau)$.

Given $\varepsilon > 0$ small enough take $\tilde{\tau} = \tau + 2$ so that $\gamma = 2\tau + 5$, $\tilde{\kappa} = C\varepsilon^{(m+1)/4}$ and $s = \varepsilon^{1/2}$. Since $m > 3$ we have $\tilde{\kappa} \leqslant \varepsilon\kappa$ and $s - c\tilde{\kappa}^2 > 0$. Consider I_0 such that its frequency $\omega = \omega(I_0)$ as given in (51), verifies $\omega \in \mathcal{D}_1(\varepsilon v, \tilde{\kappa}, \tilde{\tau})$.

Therefore, the change of variables (49) transforms Hamiltonian (46) into (50).

After the change of origin $I = \hat{J} - I_0$, Hamiltonian (50) takes the form (53). All the conditions of Theorem 4.8, specially (56) are satisfied, so there exists an invariant torus of frequency $(\omega, \varepsilon v)$ of the $\mathcal{C}^{\bar{l}}$ Hamiltonian (53).

This invariant torus is the image of the torus $\{0\} \times \mathbb{T}^{1+d}$ by the diffeomorphism F produced in Theorem 4.8. By (59) and (61), we have

$$||F - \mathrm{Id}\,||_{\mathcal{C}^{\bar{l}-6-2\tau}} \leqslant C\varepsilon^{-(m+1)/2}\varepsilon^{(m+1)} = C\varepsilon^{(m+1)/2}.$$

Since change (49) is $O_{\mathcal{C}^{r-m-3-\tau}}(\varepsilon)$-close to the identity and does not change the action $\hat{J}$, we obtain an invariant torus $\mathcal{T}_\omega(\varepsilon)$ of Hamiltonian (46) which is $O_{\mathcal{C}^{\bar{l}-6-2\tau}}(\varepsilon^{(m+1)/2})$-close to $\{I_0\} \times \mathbb{T}^{1+d}$.

Proposition A.3 shows that any ball of $\mathbb{R}$ of radius r, with r bigger than $\tilde{\kappa} = C\varepsilon^{(m+1)/4}$ contains an $\omega \in \mathcal{D}_1(\varepsilon v; C\varepsilon^{(m+1)/4}, \tilde{\tau})$. By Eq. (51) the mapping that to frequency ω associates the action I_0 is a diffeomorphism with uniform Lipschitz constant of order one. Therefore, in any ball of radius bigger than $C_1\tilde{\kappa}$ in the space of actions, we can find a point so that the frequency of the unperturbed part in (46) satisfies the conditions of Theorem 4.8.

As the tori are $O_{\mathcal{C}^{\bar{l}-6-2\tau}}(\varepsilon^{(m+1)/2})$-close to the surfaces $\hat{J} = I_0$, the gaps between them in the $\mathcal{C}^1$ topology are of order $O_{\mathcal{C}^{\bar{l}-6-2\tau}}(\varepsilon^{(m+1)/4})$. Since $\bar{l} - 6 - 2\tau > 2\tau + 6 > 1$, the gaps between them in the C^1 topology are of order $\varepsilon^{(m+1)/4}$. $\square$

Remark 4.12. By Theorem 4.7, the change of variables transforming Hamiltonian (45) in (46) is ε^2-close to the identity in the $\mathcal{C}^{r-3-\tau-m}$ topology. Therefore, Hamiltonian (45) will have KAM tori that are $O_{\mathcal{C}^1}(\varepsilon^2)$-close to surfaces of the form $J_0 \times \mathbb{T}^{1+d}$ in the extended phase space, but the gaps between them are still of order $O_{\mathcal{C}^1}(\varepsilon^{(m+1)/4})$.

Remark 4.13. The KAM tori we have produced are, obviously, invariant for the full Hamiltonian (35). Nevertheless, considered as subsets of the full system they are not maximal tori (see [Moe96,Sor02]). They inherit the stable and unstable directions from the normally hyperbolic invariant manifold $\tilde{\Lambda}_\varepsilon$.

Remark 4.14. Note that the tori produced by KAM theory are of codimension 1 inside $\tilde{\Lambda}_\varepsilon$. Therefore we can consider the manifolds with boundary inside $\tilde{\Lambda}_\varepsilon$ trapped between KAM tori. Any of these submanifolds will be a normally hyperbolic invariant manifold for the extended flow, and then the results of hyperbolic perturbation theory of Theorem 4.1 give in this case results of uniqueness for the stable and unstable manifolds as is explained in observation 4 after Theorem A.14 in [DLS00].

4.3.7. Some approximate expressions for the KAM tori

Theorem 4.8, Proposition 4.10 and Remark 4.12 provide the existence of KAM invariant tori of Hamiltonian (45), which is the restriction of Hamiltonian (35) H^*_ε to Λ^*_ε.

It will be useful for further analysis to obtain some explicit approximations for these invariant tori in the coordinate system given by the phases (φ, θ), the value of the Hamiltonian H_ε, and the actions A.

The reason we want approximations for the KAM tori in these coordinates is that, to decide whether the image of a torus by the scattering map crosses another torus, we need to take into account that not only the scattering map changes from its unperturbed value but also that the position of the tori change under the perturbation.

In order to compare the effect of the perturbation of the tori with the scattering map it will be useful for us that the approximation for the invariant tori satisfy a functional equation. The functional equation will be very similar to the equations satisfied by the first order approximation of the change in the scattering map and it will allow us to combine them into a function that measures whether tori intersect. This procedure is a generalization of the one used in [DLS00,Tre94].

The following proposition is a consequence of the fact that the KAM tori are close to flat in a $\mathcal{C}^1$ norm.

Proposition 4.15. *Under the assumptions of Proposition* 4.10, *let* $\omega \in \mathcal{D}_1(\varepsilon\nu; \tilde{\kappa}, \tilde{\tau})$ *be one of the frequencies for which we can apply Theorem* 4.8 *and let us call* $l = \bar{l} - 6 - 2\tau$ *and assume* (61).

Then, in the coordinate system $(H_\varepsilon, A, \varphi, \theta)$, *where* H_ε, A *are given in* (35) *and* φ *is the angle variable introduced in Section* 3.6, *we can write the invariant torus* $\mathcal{T}^*_{\omega,B}(\varepsilon) \subset \Lambda^*_\varepsilon$ *of frequencies* $(\omega, \varepsilon\nu)$ *as the graph of functions from the angle coordinates to the energy and the action variables*:

$$\begin{aligned} H_\varepsilon &= G_\omega(\varphi, \theta; \varepsilon), \\ A &= P_{\omega,B}(\varphi, \theta; \varepsilon). \end{aligned} \tag{62}$$

Moreover, the functions G_ω, $P_{\omega,B}$ *can be written as*:

$$\begin{aligned} G_\omega(\varphi, \theta; \varepsilon) &= \omega^2/2 + \varepsilon^2 \overline{U}(\theta) + \varepsilon^3 \nu \cdot \partial_\theta \tilde{h}(\varphi, \theta; \varepsilon) + O_{\mathcal{C}^{l-\tau}}(\varepsilon^4), \\ P_{\omega,B}(\varphi, \theta; \varepsilon) &= B - \varepsilon \partial_\theta \bar{h}(\theta) - \varepsilon^2 \partial_\theta \tilde{h}(\varphi, \theta; \varepsilon) + O_{\mathcal{C}^{l-\tau}}(\varepsilon^3). \end{aligned} \tag{63}$$

(*Even if* $\tilde{h}, \bar{h}$ *depend on* ω, *we omit the* ω *from the notation for typographical reasons.*)

The functions $\tilde{h}(\varphi, \theta; \varepsilon)$ *and* $\bar{h}(\theta)$ *are* 1-*periodic functions which satisfy*

$$\begin{aligned} &\omega \partial_\varphi \tilde{h}(\varphi, \theta; \varepsilon) + \varepsilon \nu \cdot \partial_\theta \tilde{h}(\varphi, \theta; \varepsilon) = \widetilde{U}(\Lambda^q_{1/2}(\varphi), \theta), \\ &\nu \cdot \partial_\theta \bar{h}(\theta) = \overline{U}(\theta), \end{aligned} \tag{64}$$

where the functions $\overline{U}(\theta)$ *and* $\widetilde{U}(q,\theta)$ *are defined in* (17), *and* $\left\|\tilde{h}(\cdot,\cdot;\varepsilon)\right\|_{\mathcal{C}^{l-\tau}}$ *is bounded uniformly in* ε.

Remark 4.16. From expression (63), it is clear that any trajectory in the invariant torus $\mathcal{T}^*_{\omega,B}(\varepsilon)$ experiences an oscillation of order ε^2 in the energy H_ε, due to the averaged term $\varepsilon^2\bar{U}(\theta)$, which is the same for all the invariant tori.

On the other hand, the same trajectory experiences an oscillation of order ε^3 in the averaged energy $H_\varepsilon - \varepsilon^2\bar{U}(\theta)$, which could have been chosen as alternative variable to describe the tori. Of course, if we had averaged more times, the tori would have been flatter.

Proof. First notice that, since the total extended energy H^*_ε in (35) is conserved, we have that $G_\omega + \varepsilon v \cdot P_{\omega,B}$ is independent of φ, θ.

We will derive the equations satisfied by the functions G_ω, $P_{\omega,B}$. These equations will be just expressions of the derivatives of H_ε and A with respect to time along the trajectories of the flow.

The KAM theorem 4.8, Proposition 4.10, the change of variables (49), the Averaging Theorem 4.7, and Theorem 4.1 about the persistence of normally hyperbolic invariant manifolds, provide us with a parameterization $\mathcal{K}_{\omega,B}(\psi,\theta)$ of the invariant torus $\mathcal{T}^*_{\omega,B}(\varepsilon)$ of Hamiltonian (35) in terms of two variables $\psi \in \mathbb{T}^1, \theta \in \mathbb{T}^d$ (θ will be the phase of the perturbation and will remain unchanged) so that the evolution equations inside the torus are equivalent to $\dot{\psi} = \omega, \dot{\theta} = \varepsilon v$.

By Proposition 4.10, the KAM parameterization is $O_{\mathcal{C}^l}(\varepsilon^{(m+1)/2})$-close to the identity when expressed in the coordinates $(I,\psi,\mathcal{B},\theta)$ of Hamiltonian (53). Moreover, the change of variables in the averaging method given in Theorem 4.7 is $O_{\mathcal{C}^{r-2-m}}(\varepsilon^2)$ close to the identity. Therefore, using also change (49), the function that gives the variable φ as a function of ψ, θ satisfies

$$\varphi = \varphi_\omega(\psi,\theta) = \psi + \varepsilon\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta;\varepsilon) + O_{\mathcal{C}^l}(\varepsilon^2), \quad \hat{J} = J_\omega(\psi,\theta) = \omega + O_{\mathcal{C}^l}(\varepsilon^2). \quad (65)$$

The phase θ of the external perturbation is not changed in any of the changes of coordinates that we introduce.

Continuing one step further, using the parameterization on Λ^*_ε given by Theorem 4.2, and using (40) and (7), we obtain

$$\begin{aligned}
\mathcal{K}_{\omega,B}(\psi,\theta) &= (p(\psi,\theta), q(\psi,\theta), A(\psi,\theta), \theta)\\
&= \left(\Lambda_{\frac{\omega^2}{2}}\left(\frac{1}{\omega}\left(\psi + \varepsilon\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta;\varepsilon)\right)\right), B, \theta\right) + O_{\mathcal{C}^l}(\varepsilon^2)\\
&= \left(\omega\Lambda^p_{\frac{1}{2}}\left(\psi + \varepsilon\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta)\right), \Lambda^q_{\frac{1}{2}}\left(\psi + \varepsilon\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta;\varepsilon)\right), B, \theta\right)\\
&\quad + O_{\mathcal{C}^l}(\varepsilon^2).
\end{aligned} \quad (66)$$

We write

$$G_\omega(\psi,\theta;\varepsilon) = H_\varepsilon(p(\psi,\theta), q(\psi,\theta),\theta), \quad P_{\omega,B}(\psi,\theta;\varepsilon) = A(p(\psi,\theta), q(\psi,\theta),\theta),$$

and observe that $G_\omega(\psi,\theta;\varepsilon) = \frac{\omega^2}{2} + O_{\mathcal{C}^l}(\varepsilon^2)$ and $P_{\omega,B}(\psi,\theta;\varepsilon) = B + O_{\mathcal{C}^l}(\varepsilon^2)$.

Moreover

$$\begin{aligned}\frac{d}{dt}(G_\omega\circ\Phi^*_{t,\varepsilon})_{|t=0} &= \varepsilon^3 \nu\cdot\partial_\theta U(q(\psi,\theta),\theta),\\ \frac{d}{dt}(P_{\omega,B}\circ\Phi^*_{t,\varepsilon})_{|t=0} &= -\varepsilon^2 \partial_\theta U(q(\psi,\theta),\theta),\end{aligned}$$

where $\Phi^*_{t,\varepsilon}$ is the flow associated to the Hamiltonian H^*_ε in (35).

Taking into account (66), we obtain

$$\begin{aligned}\frac{d}{dt}(G_\omega\circ\Phi^*_{t,\varepsilon})_{|t=0} &= \varepsilon^3 \nu\cdot\partial_\theta U\left(\Lambda^q_{1/2}\left(\psi+\varepsilon\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta;\varepsilon)\right),\theta\right) + O_{\mathcal{C}^l}(\varepsilon^5),\\ \frac{d}{dt}(P_{\omega,B}\circ\Phi^*_{t,\varepsilon})_{|t=0} &= -\varepsilon^2 \partial_\theta U\left(\Lambda^q_{1/2}\left(\psi+\varepsilon\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta;\varepsilon)\right),\theta\right) + O_{\mathcal{C}^l}(\varepsilon^4).\end{aligned}$$

On the other hand, using the reduced equations $\dot\psi = \omega$, $\dot\theta = \varepsilon\nu$, we have that

$$\frac{d}{dt}(G_\omega\circ\Phi^*_{t,\varepsilon})_{|t=0} = \omega\frac{\partial G_\omega}{\partial\psi}(\psi,\theta;\varepsilon) + \varepsilon\nu\frac{\partial G_\omega}{\partial\theta}(\psi,\theta;\varepsilon),$$

and an analogous equation for $P_{\omega,B}$.

So we can obtain G_ω as a solution of the functional equation

$$\omega\frac{\partial G_\omega}{\partial\psi}(\psi,\theta) + \varepsilon\nu\frac{\partial G_\omega}{\partial\theta}(\psi,\theta) = \varepsilon^3\nu\cdot\partial_\theta U\left(\Lambda^q_{1/2}\left(\psi+\varepsilon\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta;\varepsilon)\right),\theta\right) + O_{\mathcal{C}^l}(\varepsilon^5).$$

Expanding the right-hand side of the equation by Taylor's formula we can look for G_ω as $G_\omega = \frac{\omega^2}{2} + g_1 + g_2 + g_3$, where

$$\omega\frac{\partial g_j}{\partial\psi} + \varepsilon\nu\frac{\partial g_j}{\partial\theta} = \Gamma_j, \tag{67}$$

with

$$\Gamma_1(\psi,\theta) = \varepsilon^3 v\cdot\partial_\theta U(\Lambda^q_{1/2}(\psi),\theta),$$

$$\Gamma_2(\psi,\theta) = \varepsilon^4 v\cdot\partial_\psi\left(\partial_\theta U(\Lambda^q_{1/2}(\psi),\theta)\right)\frac{\partial\chi}{\partial\hat{J}}(\omega,\theta;\varepsilon),$$

$$\Gamma_3(\psi,\theta) = O_{\mathcal{C}^l}(\varepsilon^5).$$

The classical way to solve the homological equation

$$\omega\frac{\partial g}{\partial\psi}+\varepsilon v\frac{\partial g}{\partial\theta}=\Gamma \tag{68}$$

is expanding Γ and g in Fourier series $g(\psi,\theta)=\sum g_{k,l}e^{i(k\psi+l\cdot\theta)}$, obtaining

$$g_{k,l}=\frac{\Gamma_{k,l}}{i(\omega k+\varepsilon v\cdot l)}. \tag{69}$$

From expression (69) it follows that for $k\neq 0$ the coefficient $g_{k,l}$ is of the same order as the coefficient $\Gamma_{k,l}$ (see Remark 4.14 in [DLS00]).

In particular, when $\int_{\mathbb{T}}\Gamma(\psi,\theta)d\psi=0$, we have that there exists some constant $C>0$ such that $\|g\|_{\mathcal{C}^l}\leqslant C\,\|\Gamma\|_{\mathcal{C}^l}$. This reasoning leads to $\|g_2\|_{\mathcal{C}^l}\leqslant C\,\|\Gamma_2\|_{\mathcal{C}^l}=O_{\mathcal{C}^l}(\varepsilon^4)$.

For $k=0$, Eq. (69) gives $g_{0,l}=\frac{\Gamma_{0,l}}{i\varepsilon(v\cdot l)}$.

So, in the general case, using the results of [Rus75] (see also Lemma A.23) the solution of Eq. (68) verifies that there exists some constant $C>0$ such that $\|g\|_{\mathcal{C}^{l-\tau}}\leqslant C\frac{1}{\varepsilon}\|\Gamma\|_{\mathcal{C}^l}$, where τ is the Diophantine exponent of v. This reasoning leads to $\|g_3\|_{\mathcal{C}^{l-\tau}}\leqslant C\frac{1}{\varepsilon}\|\Gamma_3\|_{\mathcal{C}^l}=O_{\mathcal{C}^{l-\tau}}(\varepsilon^4)$.

g_1 can be found explicitly solving Eq. (67) for $j=1$, and using the decomposition introduced in (17), so that

$$\partial_\theta U(\Lambda_{1/2}(\psi),\theta)=\partial_\theta\overline{U}(\theta)+\partial_\theta\tilde{U}(\Lambda_{1/2}(\psi),\theta),$$

which gives immediately

$$g_1=\varepsilon^2\overline{U}(\theta)+\varepsilon^3\tilde{g}_1(\psi,\theta;\varepsilon),$$

where $\tilde{g}_1$ verifies the same equation as g_1 with the $\tilde{U}$ instead of U. Moreover $\tilde{g}_1=v\partial_\theta\tilde{h}$, if we choose $\tilde{h}$ to be the solution of the homological equation (68) with $\Gamma=\tilde{U}(\Lambda_{1/2}(\psi),\theta)$.

Putting together the expression for $g-1$, as well as the bounds for g_2 and g_3, we obtain the approximation formula (63) for G_ω in terms of the variables (ψ,θ). Using

(65) we see that we can change the arguments of the right-hand side from the ψ, θ to φ, θ without any change in the explicit terms of formula (63).

An analogous reasoning leads to the equation for the components $A = P_{\omega,B}$ of the torus.

4.4. The perturbed scattering map

4.4.1. Introduction and overview

The goal of this Section is to define and to compute the scattering map S^*_ε and to use it to characterize intersections of stable and unstable manifolds of the different invariant objects in Λ^*_ε for the perturbed flow.

We recall that, according to Theorem 4.2 and Remark 4.4, when we consider the perturbed full symplectic flow of (35) in the full symplectic space, we can find Λ^*_ε, $W^{\mathrm{s,u}}(\Lambda^*_\varepsilon)$, γ^*_ε, continuing those of the unperturbed system. Then, given $x^*_+, x^*_- \in \Lambda^*_\varepsilon$, we say, as in Eq. (26), that $x^*_+ = S^*_\varepsilon(x^*_-)$ when

$$W^{\mathrm{s}}_{x^*_+} \cap W^{\mathrm{u}}_{x^*_-} \cap \gamma^*_\varepsilon \neq \emptyset, \tag{70}$$

that is, there exists $z^* \in \gamma^*_\varepsilon$ such that

$$\operatorname{dist}\left(\Phi^*_{t,\varepsilon}(x^*_\pm), \Phi^*_{t,\varepsilon}(z^*)\right) \leqslant C\, e^{-\beta|t|} \qquad \text{for} \qquad \pm t \geqslant 0. \tag{71}$$

for some $\beta > 0$ that will be close to β_0 in (11).

If we write $x^*_\pm = (x_\pm, A_\pm, \theta_\pm,)$, $z^* = (z, A_{z^*}, \theta_{z^*})$, since the flow of (35) satisfies $\dot\theta = \varepsilon\nu$, we see that (71) implies $\theta_+ = \theta_- = \theta_{z^*}$, which we will henceforth denote by θ.

In order to perform explicit calculations, we will compute the scattering map for a finite range of energies $E \in [1/2, 2]$ in the scaled variables. This will allow us to use uniform continuity arguments. This will be enough for our purposes since at this stage we only want to study transition chains for this range of energies.

We note that since E^* chosen in (30) for the change to scaled variables was arbitrary, we can construct the scattering map for all the manifold Λ^*_ε.

We will express the map S^*_ε in the full space in terms of the explicit coordinates (J, φ, B, θ) for Λ^*_ε that we have introduced in Section 4.3.3. We recall that the coordinates B are deformations of the actions A, obtained through the parameterization $\mathcal{F}^*$ of Theorem 4.2, chosen so that the symplectic form in Λ^*_ε (restriction of the ambient one) has the standard form.

In these coordinates, if we consider $x^*_+ = S^*_\varepsilon(x^*_-)$ connected through a point z^* verifying (71), we have

$$x^*_\pm = (\mathcal{F}(J_\pm, \varphi_\pm, \theta, \varepsilon^2), \mathcal{A}(J_\pm, \varphi_\pm, B_\pm, \theta, \varepsilon^2), \theta)$$

and, using the regular dependence on parameters of the stable and unstable manifolds on compact sets and (14), we obtain that

$$\begin{aligned} \varphi_\pm &= \varphi_0 + a_\pm + O(\varepsilon^2), \\ J_\pm &= J_0 + O(\varepsilon^2), \\ B_\pm &= B_0 + O(\varepsilon^2), \end{aligned} \tag{72}$$

for some $\varphi_0 \in \mathbb{R}$, $J_0 \in \mathbb{R}$, $B_0 \in \mathbb{R}^d$, and where $a_\pm$ were introduced in hypotheses H2 and H2′, in formulas (2) and (10). Moreover, we have, by (40)

$$\begin{aligned} \Phi^*_{t,\varepsilon}(x^*_\pm) &= \left(\Lambda_E\left(t + \frac{\varphi_0 + a_\pm}{\sqrt{2E}}\right) + O_{\mathcal{C}^1}(\varepsilon^2),\, B_0 + O_{\mathcal{C}^1}(\varepsilon^2),\, \theta + \varepsilon v t\right) \\ \Phi^*_{t,\varepsilon}(z^*) &= \left(\gamma_E\left(t + \frac{\varphi_0}{\sqrt{2E}}\right) + O_{\mathcal{C}^1}(\varepsilon^2),\, B_0 + O_{\mathcal{C}^1}(\varepsilon^2),\, \theta + \varepsilon v t\right) \end{aligned} \tag{73}$$

with $E = J_0^2/2$.

In the formulas (73), the $O_{\mathcal{C}^1}(\varepsilon^2)$ is uniform for $t \in \mathbb{R}$. This follows from the dependence on parameters in hyperbolicity theory and (14).

Our next goal is to give quantitative conditions for the existence of heteroclinic connections between the KAM tori obtained in Section 4.3.6. To this end, we will obtain enough quantitative information that allows us to conclude that the heteroclinic intersections are transversal and to estimate the energy of the tori connected through these intersections.

That is, we will state sufficient conditions that ensure that, given two different KAM tori $\mathcal{T}_1^*$, $\mathcal{T}_2^*$ in Λ^*_ε, the unstable manifold of $\mathcal{T}_1^*$ intersects transversally the stable manifold of $\mathcal{T}_2^*$ in the full space. The intersections of the stable and unstable manifolds of different KAM tori are referred to as heteroclinic connections. As we will see in Lemma 4.33, due to the fact that the stable and unstable manifolds of Λ^*_ε intersect transversally along γ^*_ε, it will be enough to see that $S^*_\varepsilon(\mathcal{T}_1^*)$ is transversal to $\mathcal{T}_2^*$ in Λ^*_ε.

To characterize the intersection between $S^*_\varepsilon(\mathcal{T}_1^*)$ and $\mathcal{T}_2^*$ we will use the fact that $(H_\varepsilon, \varphi, A, \theta)$ constitutes a good system of coordinates in the manifold Λ^*_ε. Moreover, by Proposition 4.15 we know that, in this system of coordinates, the KAM tori can be expressed as graphs of functions. Indeed, in the coordinates above, the KAM tori will be $O_{\mathcal{C}^1}(\varepsilon)$-close to surfaces of the form $H_\varepsilon = c$, $A_i = c_i$.

One reason why the system of coordinates given by $(H_\varepsilon, \varphi, A, \theta)$ is useful for us is that not only it is defined in the manifold Λ^*_ε but also on its homoclinic orbits. This will allow us to compute the change of the coordinates in a homoclinic excursion to Λ^*_ε and compare it with the change of coordinates of an orbit who stays in a KAM torus. In this way we will be able to decide whether a homoclinic connection connects two KAM tori.

As it is well known in Melnikov theory, the change of the coordinates in a homoclinic excursion can be computed by applying the fundamental theorem of calculus along the

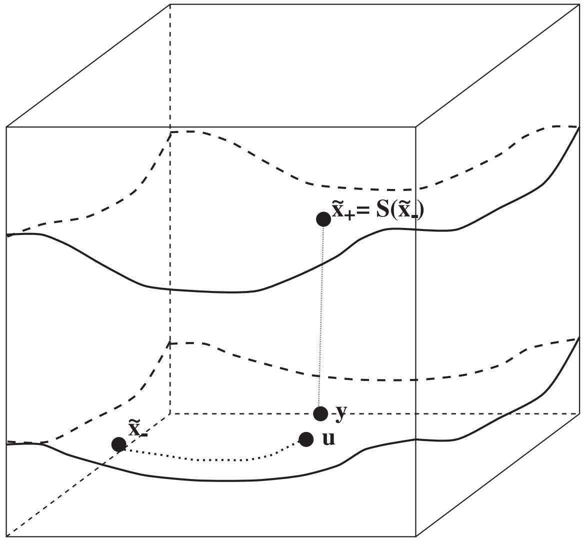

Fig. 1. Illustration of the perturbed tori and the outer map.

homoclinic orbit. In turn, the perturbed trajectory will be computed approximately using the smooth dependence on parameters for the invariant manifolds.

It is also important to note the regular dependence on the parameter ε of all the objects we are studying, such as invariant tori and their manifolds, and consequently of the scattering map S^*_ε. This will allows us to compute this map perturbatively.

We will detect whether x^*_- and $x^*_+ = S^*_\varepsilon(x^*_-)$ lie on two different KAM tori by computing $H_\varepsilon(x^*_+) - H_\varepsilon(y)$ and $A(x^*_+) - A(y)$, where y is the projection of x^*_+ on the KAM torus containing x^*_-, i.e., y has the same angle coordinates (φ_+, θ) as x^*_+. See Fig. 1.

Since we are looking for heteroclinic connections between tori, such tori have to be in the same energy level of the full space $\mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$ of the autonomous Hamiltonian $H^*_\varepsilon(p, q, A, \theta)$ in (35). Then, it is not restrictive to set $H^*_\varepsilon(p, q, A, \theta) = \varepsilon v \cdot A + H_\varepsilon(p, q, \theta) = 0$. Once we fix the total energy level, the quantities (H_ε, A) are not independent and it is natural to work only with the coordinates A since we have that

$$H_\varepsilon(x^*_+) - H_\varepsilon(y) = -\varepsilon v \cdot (A(x^*_+) - A(y)). \tag{74}$$

Therefore, the main goal in the discussion of intersections of tori under the scattering map is to compute

$$\Delta A \equiv A(x^*_+) - A(y), \tag{75}$$

where, as mentioned before, y is the point in the KAM torus which contains x^*_- and that has the same angle coordinates (φ_+, θ) as x^*_+.

We will compute ΔA as

$$\Delta A = [A(x_+^*) - A(x_-^*)] + [A(x_-^*) - A(y)]. \tag{76}$$

The first term of (76) will be computed in Lemma 4.18 by means of a classical calculation, that goes back to Poincaré. The idea is that since x_+^* and x_-^* are connected through an orbit, we can use the fundamental theorem of calculus and obtain $A(x_+^*) - A(x_-^*)$ by integrating the derivative of A along this orbit and taking appropriate limits. Moreover, by the regular dependence on parameters of normal hyperbolicity, the connecting orbit is approximately the unperturbed homoclinic orbit γ and, to compute the leading contribution, it suffices to integrate along the unperturbed orbit γ. Hence, we can explicitly write the main term of the first term of (76).

The term $A(x_-^*) - A(y)$ in (76) can be computed using the explicit expansions of KAM tori from Proposition 4.15. In Lemma 4.19 we obtain the main term of $A(x_+^*) - A(y)$ as an explicit vector function (82), that we will call Melnikov vector following [DR97].

Thanks to the Hamiltonian structure of the problem, the Melnikov vector is the gradient of a scalar function (88) called in [DR97,DG00] the Melnikov potential.

Remark 4.17. The fact that the main term of the perturbation is a gradient has deep topological implications for the number of zeros of the Melnikov vector, and hence, the number of homoclinic intersections. However, in this paper, we will use it mainly as a tool to simplify the calculations and to make connections with the variational formulation in [Mat96], which involved similar expressions.

Similarly, we note that the main term of $H_\varepsilon(x_+^*) - H_\varepsilon(y)$, which we will compute in (93), and which we call Melnikov function, is the directional derivative of the Melnikov potential in the direction given by v.

We emphasize that the calculation of Melnikov functions up to this point is quite general and does not use the fact that the system is quadratic in the momenta perturbed by a potential, or that the perturbation is slow. Taking this into account in Lemma 4.26 we establish that $\mathcal{L}$ in (90) (called Poincaré function) gives the leading term of the expansion of the Melnikov potential for large energies.

We further show that if the function $\mathcal{L}$ is non-constant there are heteroclinic connections among all sufficiently large energies forming a transition path. Once we have transition paths for all large values of the energy, we will be able to construct orbits that follow them in Lemma 4.37.

4.4.2. Calculation of the scattering map

In this section, we start computing perturbatively the scattering map. By formula (74) and the discussion after it, it is enough to compute the change of the action coordinates A. Since we will only be computing the scattering map to order ε, we will carry the calculations of the change in A only to that order.

Lemma 4.18. *Let x_-^* and x_+^* be two points on Λ_ε^* such that $x_+^* = S_\varepsilon^*(x_-^*)$. Then*

$$
\begin{aligned}
A(x_+^*) - A(x_-^*) = \varepsilon^2 \lim_{T_1, T_2 \to \infty} \Bigg[& - \int_{-T_1}^{T_2} dt\, \partial_\theta \widetilde{U} \left(\gamma_E^q \left(t + \frac{\varphi_0}{\sqrt{2E}} \right), \theta + \varepsilon v t \right) \\
& + \int_{-T_1}^{0} dt\, \partial_\theta \widetilde{U} \left(\Lambda_E^q \left(t + \frac{\varphi_0 + a_-}{\sqrt{2E}} \right), \theta + \varepsilon v t \right) \\
& + \int_{0}^{T_2} dt\, \partial_\theta \widetilde{U} \left(\Lambda_E^q \left(t + \frac{\varphi_0 + a_+}{\sqrt{2E}} \right), \theta + \varepsilon v t \right) \Bigg] \\
& + O_{\mathcal{C}^1}(\varepsilon^4),
\end{aligned}
\tag{77}
$$

where, using (72),

$$
\begin{aligned}
x_\pm^* &= \left(\mathcal{F}(J_\pm, \varphi_\pm, \theta, \varepsilon^2), \mathcal{A}(J_\pm, \varphi_\pm, B_\pm, \theta, \varepsilon^2), \theta \right) \\
&= \left(\Lambda_E \left(\frac{\varphi_0 + a_\pm}{\sqrt{2E}} \right) + O_{\mathcal{C}^1}(\varepsilon^2), B_0 + O_{\mathcal{C}^1}(\varepsilon^2), \theta \right)
\end{aligned}
$$

for some $\varphi_0 \in \mathbb{R}$, $E > 0$, $B_0 \in \mathbb{R}$, where $\mathcal{F}$ is introduced in Theorem 4.1, *and $\widetilde{U}$, introduced in* (17), *is the forcing potential minus its average on the periodic orbit $\Lambda_{1/2}$.*

Proof. Recall that if $\lambda^*(t) = (\lambda^p(t), \lambda^q(t), A(\lambda^*(t)), \alpha + \varepsilon v t)$ is a trajectory of the Hamiltonian (35) then

$$
\frac{d}{dt} A(\lambda^*(t)) = -\varepsilon^2 \partial_\theta U(\lambda^q(t), \alpha + \varepsilon v t).
$$

Therefore, for two trajectories $\lambda^*(t)$ and $\mu^*(t)$ of (37), we have, by the fundamental theorem of calculus,

$$
\begin{aligned}
A(\lambda^*(T)) - A(\mu^*(T)) = & \, A(\lambda^*(0)) - A(\mu^*(0)) \\
& - \varepsilon^2 \int_0^T dt\, \partial_\theta U(\lambda^q(t), \alpha + \varepsilon v t) \\
& + \varepsilon^2 \int_0^T dt\, \partial_\theta U(\mu^q(t), \beta + \varepsilon v t).
\end{aligned}
\tag{78}
$$

Since $x_+^* = S_\varepsilon^*(x_-^*)$, we know that there exists $z^* \in \mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$, such that the trajectories $\gamma_\varepsilon^*(t) = \Phi_{t,\varepsilon}^*(z^*)$ and $\Lambda_{\pm,\varepsilon}^*(t) = \Phi_{t,\varepsilon}^*(x_\pm^*)$ verify (71).

Then, taking limits $T \to \pm\infty$ in (78) for $\lambda^*(t) = \Lambda^*_{\pm,\varepsilon}(t)$ and $\mu^*(t) = \gamma^*_\varepsilon(t)$ as appropriate, we obtain

$$\begin{aligned}
0 &= A(x_+^*) - A(z^*) \\
&\quad - \lim_{T_2\to\infty} \varepsilon^2 \int_0^{T_2} dt\ \left(\partial_\theta U\left(\Lambda^q_{+,\varepsilon}(t), \theta+\varepsilon\nu t\right) - \partial_\theta U\left(\gamma^q_\varepsilon(t), \theta+\varepsilon\nu t\right)\right), \\
0 &= A(x_-^*) - A(z^*) \\
&\quad - \lim_{T_1\to\infty} \varepsilon^2 \int_0^{-T_1} dt\ \left(\partial_\theta U\left(\Lambda^q_{-,\varepsilon}(t), \theta+\varepsilon\nu t\right) - \partial_\theta U\left(\gamma^q_\varepsilon(t), \theta+\varepsilon\nu t\right)\right).
\end{aligned}$$

Subtracting these two equations we obtain

$$\begin{aligned}
&A(x_+^*) - A(x_-^*) \\
&\quad = +\varepsilon^2 \lim_{T_1,T_2\to\infty} \left[\int_0^{T_2} dt\ \left(\partial_\theta U\left(\Lambda^q_{+,\varepsilon}(t), \theta+\varepsilon\nu t\right) - \partial_\theta U\left(\gamma^q_\varepsilon(t), \theta+\varepsilon\nu t\right)\right)\right. \\
&\qquad \left. + \int_{-T_1}^{0} dt\ \left(\partial_\theta U\left(\Lambda^q_{-,\varepsilon}(t), \theta+\varepsilon\nu t\right) - \partial_\theta U\left(\gamma^q_\varepsilon(t), \theta+\varepsilon\nu t\right)\right)\right].
\end{aligned} \tag{79}$$

By (71), the limits in (79) are reached uniformly in ε, and the dependence of the trajectories on ε is uniform on compact intervals of time. Hence, at the expense only of introducing an error of higher order in ε, we can substitute in (79) for $\Lambda_{\pm,\varepsilon}$ and γ_ε the unperturbed orbits given by (73).

We note that the right-hand side of (79) is linear in U. Hence if we use the decomposition $U(q,\theta) = \overline{U}(\theta) + \widetilde{U}(q,\theta)$ given in (17), the computation of the right-hand side of the terms in (79) containing $\overline{U}$ are zero, so we obtain (77). □

4.4.3. Intersection of KAM tori with the image of KAM tori under the scattering map

The goal of this section is to decide whether the image under the scattering map of a KAM torus intersects another (or perhaps the same) KAM torus.

To carry out this calculation, we will work in the coordinates given by $(H_\varepsilon, A, \varphi, \theta)$. In these coordinates, the torus is approximately given by the graph of a function from the angles into the action variables computed in Proposition 4.15.

Lemma 4.19. *Let y be a point with the phases (φ_+, θ) of x_+^* and which lies on the invariant torus for the perturbed flow which contains x_-^*, where*

$$\begin{aligned}
x_\pm^* &= \left(\mathcal{F}(J_\pm, \varphi_\pm, \theta, \varepsilon^2), \mathcal{A}(J_\pm, \varphi_\pm, B_\pm, \theta, \varepsilon^2), \theta\right) \\
&= \left(\Lambda_E\left(\frac{\varphi_0 + a_\pm}{\sqrt{2E}}\right) + O_{\mathcal{C}^1}(\varepsilon^2), B_0 + O_{\mathcal{C}^1}(\varepsilon^2), \theta\right).
\end{aligned}$$

Then

$$
\begin{aligned}
A(x_+^*) - A(y) = \varepsilon^2 \lim_{T_1,T_2\to\infty} \Bigg[&- \int_{-T_1}^{T_2} dt\, \partial_\theta \widetilde{U}\left(\gamma_E^q\left(t+\frac{\varphi_0}{\sqrt{2E}}\right), \theta+\varepsilon v t\right) \\
&+\partial_\theta \tilde{h}\left(\varphi_0 + a_+ + \sqrt{2E}\, T_2, \theta+\varepsilon v T_2; \varepsilon\right) \\
&-\partial_\theta \tilde{h}\left(\varphi_0 + a_- - \sqrt{2E}\, T_1, \theta-\varepsilon v T_1; \varepsilon\right) \Bigg] \\
&+O_{\mathcal{C}^1}(\varepsilon^3),
\end{aligned} \tag{80}
$$

where $\tilde{h}(\varphi, \theta; \varepsilon)$ *verifies Eq.* (64), *associated to the invariant torus of the perturbed flow which contains* x_-^*.

Proof. We use Lemma 4.18 for $A(x_+^*) - A(x_-^*)$ and formulas (62) and (63), $l-\tau > 1$, of Proposition 4.15 for $A(x_-^*) - A(y)$:

$$
\begin{aligned}
A(x_+^*) - A(y) &= A(x_+^*) - A(x_-^*) + A(x_-^*) - A(y) \\
&= \varepsilon^2 \lim_{T_1,T_2\to\infty} \Bigg[- \int_{-T_1}^{T_2} dt\, \partial_\theta \widetilde{U}\left(\gamma_E^q\left(t+\frac{\varphi_0}{\sqrt{2E}}\right), \theta+\varepsilon v t\right) \\
&\quad + \int_{-T_1}^{0} dt\, \partial_\theta \widetilde{U}\left(\Lambda_E^q\left(t+\frac{\varphi_0+a_-}{\sqrt{2E}}\right), \theta+\varepsilon v t\right) \\
&\quad + \int_{0}^{T_2} dt\, \partial_\theta \widetilde{U}\left(\Lambda_E^q\left(t+\frac{\varphi_0+a_+}{\sqrt{2E}}\right), \theta+\varepsilon v t\right) \\
&\quad -\partial_\theta \tilde{h}(\varphi_0 + a_-, \theta; \varepsilon) + \partial_\theta \tilde{h}(\varphi_0 + a_+, \theta; \varepsilon) \Bigg] \\
&\quad +O_{\mathcal{C}^1}(\varepsilon^3).
\end{aligned} \tag{81}
$$

Now, calling $A_-(t) = \partial_\theta \tilde{h}\left(\varphi_0 + a_- + \sqrt{2E}t, \theta+\varepsilon v t; \varepsilon\right)$, we have, using the functional equation (64) verified by $\tilde{h}$ and the scaling (7):

$$
\begin{aligned}
\frac{d}{dt} A_-(t) = &\sqrt{2E}\, \partial_\varphi \partial_\theta \tilde{h}\left(\varphi_0 + a_- + \sqrt{2E}t, \theta+\varepsilon v t; \varepsilon\right) \\
&+\varepsilon v \partial_\theta \partial_\theta \tilde{h}\left(\varphi_0 + a_- + \sqrt{2E}t, \theta+\varepsilon v t; \varepsilon\right)
\end{aligned}
$$

$$= \partial_\theta \widetilde{U}\left(\Lambda^q_{1/2}\left(\sqrt{2E}t + \varphi_0 + a_-\right), \theta + \varepsilon v t\right)$$
$$= \partial_\theta \widetilde{U}\left(\Lambda^q_E\left(t + \frac{\varphi_0 + a_-}{\sqrt{2E}}\right), \theta + \varepsilon v t\right)$$

and a similar identity holds for $A_+(t) = \partial_\theta \tilde{h}\left(\varphi_0 + a_+ + \sqrt{2E}t, \theta + \varepsilon v t; \varepsilon\right)$, which verifies

$$\frac{d}{dt}A_+(t) = \partial_\theta \widetilde{U}\left(\Lambda^q_E\left(t + \frac{\varphi_0 + a_+}{\sqrt{2E}}\right), \theta + \varepsilon v t\right)$$

Then, using the fundamental theorem of Calculus, we have for any T:

$$A_\pm(T) - A_\pm(0) = \int_0^T dt\, \partial_\theta \widetilde{U}\left(\Lambda^q_E\left(t + \frac{\varphi_0 + a_\pm}{\sqrt{2E}}\right), \theta + \varepsilon v t\right)$$

and using the above identities to express the second and third integrals in (81) with $T = -T_1, T_2$ we obtain formula (80). $\square$

Remark 4.20. The vector function M giving the main term of (80) in Lemma 4.19 is

$$\begin{aligned} &M(\varphi_0, \theta, E; \varepsilon) \\ &= \lim_{T_1,T_2\to\infty}\left[-\int_{-T_1}^{T_2} dt\, \partial_\theta \widetilde{U}\left(\gamma^q_E\left(t + \frac{\varphi_0}{\sqrt{2E}}\right), \theta + \varepsilon v t\right)\right. \\ &\quad + \partial_\theta \tilde{h}\left(\varphi_0 + a_+ + \sqrt{2E}T_2, \theta + \varepsilon v T_2; \varepsilon\right) \\ &\quad \left. - \partial_\theta \tilde{h}\left(\varphi_0 + a_- - \sqrt{2E}T_1, \theta - \varepsilon v T_1; \varepsilon\right)\right] \end{aligned} \tag{82}$$

is usually called (see [DR97]) the Melnikov vector associated to the perturbed torus in Λ^*_ε which arises, when $\varepsilon \neq 0$, from the torus $\mathcal{T}^*_{E,B_0}$ given in (24).

As, by Lemma 4.19, we have that

$$A(x^*_+) - A(y) = \varepsilon^2 M(\varphi_0, \theta, E; \varepsilon) + O_{\mathcal{C}^1}(\varepsilon^3), \tag{83}$$

M is the leading term of the vector we will use to study the existence of heteroclinic intersections among tori.

Remark 4.21. Let us note that the Melnikov vector (82) has terms of two different kinds: A term which is an integral of $\tilde{U}$ over orbits of the unperturbed system and the terms given by the gradient of the function $\tilde{h}$. Both of these terms have a clear dynamical meaning. The term $\partial_\theta \tilde{h}$ measures the displacement of the invariant torus under the perturbation, whereas the integral term represents the changes induced in the stable manifolds of the unperturbed torus.

Clearly, when one wants to establish the existence of intersections between two tori labeled by different frequencies, one has to take into account the change of the tori and the changes of the invariant manifolds. Since we are considering only the leading order in ε of the changes, it suffices to add both effects.

Unfortunately, in the literature on the Melnikov method it is a rather extended mistake to omit the term corresponding to the displacement of the torus. For instance, it is omitted in (2.16) of [HM82, p. 671]. This leads to somewhat paradoxical results, such as the fact that the integrals that appear in the Melnikov vector are not convergent but conditionally convergent. For example, in [Wig90, Proposition 4.1.29, p. 412 ff.] it is realized—but not acted upon—that these omitted terms lead to consideration of improper integrals with oscillatory integrands and that they only converge along subsequences of times. Of course, the the value of the integral depends on the arbitrary choice such as the sequence on which it is evaluated. Hence it is somewhat suspicious that one can draw dynamical conclusions about their zeros.

A more careful analysis of the subsequences allowed for the convergence of these conditionally convergent integrals can be found in [Rob88]. A correct treatment of these problems including the correction due to the change of the unperturbed torus can be found in [Tre94] and also in [DG00,DLS00].

We note however, that if one is interested only in making statements that include that some phenomena happen generically, the exact form of the formula is not used, only that the formula exists. Hence, the results which conclude that phenomena happen generically are not affected. Of course, those that obtain conclusions for specific systems may be affected.

Remark 4.22. Even if we will not be concerned with homoclinic intersections of tori, we note that the non-degenerate zeros of the vector function $M(\varphi_0, \theta, E; \varepsilon)$ lead to homoclinic intersections to the perturbed torus $\mathcal{T}^*_{\omega,B_0}(\varepsilon)$ in the full space $\mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$.

From (83), it is straightforward that

$$H_\varepsilon(x_+^*) - H_\varepsilon(y) = -\varepsilon^3 v \cdot M(\varphi_0, \theta, E; \varepsilon) + O_{\mathcal{C}^1}(\varepsilon^4). \tag{84}$$

Hence, we arrive to a rather explicit expression of the image of the KAM tori under the scattering map.

We will introduce the notation

$$M_{\varepsilon,E}(\varphi_0, \theta) = M(\varphi_0, \theta, E, \varepsilon). \tag{85}$$

This is useful when we want to think of M as a function of the φ_0, θ and think of the variables E, ε as parameters. As we will see this notation plays a role in (86) in Proposition 4.23 when we want to express images of tori as graphs.

Next, proposition is an immediate consequence of (83) and (84) letting (φ, θ) varying in $\mathbb{T} \times \mathbb{T}^d$

Proposition 4.23. *Consider a KAM torus $\mathcal{T}^*_{\omega,B}(\varepsilon) \subset \Lambda^*_\varepsilon$ represented in Proposition* 4.15 *as the graph of functions G_ω, $P_{\omega,B}$, giving respectively H_ε, A as a function of the angle coordinates φ, θ.*

Then, we can also represent $S^*_\varepsilon(\mathcal{T}^*_{\omega,B_0}(\varepsilon))$, the image of $\mathcal{T}^*_{\omega,B_0}(\varepsilon)$ under the scattering map S^*_ε, as the graph of two functions $\hat{G}_\omega$, $\hat{P}_\omega$ that satisfy

$$\hat{G}_\omega = G_\omega + \varepsilon^3 v \cdot M_{\varepsilon,E} + O_{\mathcal{C}^1}(\varepsilon^4), \tag{86}$$

$$\hat{P}_{\omega,B} = P_{\omega,B} - \varepsilon^2 M_{\varepsilon,E} + O_{\mathcal{C}^1}(\varepsilon^3). \tag{87}$$

where $M_{\varepsilon,E}$ is as introduced in (85).

Remark 4.24. There are significant differences in the geometry of the problem in the extended phase space $\mathbf{T}^*M \times \mathbb{T}^d$ and the full symplectic phase space $\mathbf{T}^*M \times \mathbb{R}^d \times \mathbb{T}^d$.

We now that the symplectically extended equations (32) are a skew product. That is, the equations of motion for the extended phase space for the dynamical variables p, q, θ do not involve the action A. On the other hand, the motion of the A can be obtained from the solution of the other variables by a quadrature.

The fact that the system is a skew product causes that the dimension counting for the intersections has to be somewhat different when one considers the system in the dynamical variables and in the full symplectically extended variables.

Roughly speaking, every feature in the dynamical variables p, q, θ lifts to a full family of features in the full symplectically extended variables. In particular, a KAM or a whiskered torus in the dynamical variables lifts to a full family of KAM or whiskered tori. Note that heteroclinic connections among tori in the full extended phase space may correspond to homoclinic connections for tori in the extended space.

There are important consequences of the fact that tori in the extended system lift to families of tori in the full symplectically extended system. In the full symplectically extended system the lifts of KAM or whiskered tori in the extended space do not have gaps in the direction of the actions A. Of course, they have gaps in the direction of the dynamical variables and consequently in the energy H_ε.

Note that, since total energy $H^*_\varepsilon(p, q, \theta, A) = v \cdot A + H_\varepsilon(p, q, \theta)$ is conserved, changes in the energy H_ε imply changes in A but changes in A do not imply necessarily changes in H_ε since they could be in the direction orthogonal to v.

Hence, the diffusion in the action variables A is significantly different from the diffusion in the energy. The problem of diffusion in the A variables is much closer to the model considered in [Arn64], where the tori also form a continuum.

The problem of diffusion in the action variables A in a priori unstable isochronous systems has been studied in [Gal97,Gal99].

The results in this paper do not apply to the models considered in [Gal97,Gal99]. We assume that the systems that we study in this paper are a priori chaotic and satisfy some scaling hypothesis rather than just assuming that they are a priori unstable. In

[DdlLS03,DdlLMS03], we have extended the methods here to suggest a mechanism for diffusion in the dynamical variables for a priori unstable anysochronous systems and verified it in some models.

4.4.4. The Melnikov potential

In this subsection we express the Melnikov vector (82) as a gradient and compute its leading term as $\varepsilon \to 0$. The representation of the Melnikov function as a gradient will allow us to simplify the expression for high energy, or equivalently, for ε small enough, and show that just one condition is enough to guarantee the existence of transition chains consisting of infinitely many tori whose energy tends to infinity.

Following [DG00,DR97] we will show that the Melnikov vector M can be obtained taking appropriate derivatives of a function L in (88) called the Melnikov potential.

We define the Melnikov potential as

$$\begin{aligned} L(\varphi_0, \theta, E; \varepsilon) = \lim_{T_1, T_2 \to \infty} \Bigg[& - \int_{-T_1}^{T_2} dt\, \widetilde{U} \left(\gamma_E^q \left(t + \frac{\varphi_0}{\sqrt{2E}} \right), \theta + \varepsilon v t \right) \\ & + \tilde{h} \left(\varphi_0 + a_+ + \sqrt{2E}\, T_2, \theta + \varepsilon v T_2; \varepsilon \right) \\ & - \tilde{h} \left(\varphi_0 + a_- - \sqrt{2E}\, T_1, \theta - \varepsilon v T_1; \varepsilon \right) \Bigg], \end{aligned} \tag{88}$$

where $\tilde{h}$ verifies (64).

As can be verified exchanging the derivatives and the limit $T_1, T_2 \to \infty$ (justified by the extremely fast convergence of the terms inside the limit) we have:

Proposition 4.25. *The Melnikov potential defined in* (88) *satisfies the following properties*:

(1) $L(\varphi_0, \theta, E; \varepsilon)$ *is* 1-*periodic in* θ.
(2) *For any* $u \in \mathbb{R}$ *one has*:

$$L \left(\varphi_0 + \sqrt{2E} u, \theta + \varepsilon v u, E; \varepsilon \right) = L(\varphi_0, \theta, E; \varepsilon) \tag{89}$$

and therefore $\sqrt{2E} \frac{\partial L}{\partial \varphi_0} + \varepsilon v \cdot \frac{\partial L}{\partial \theta} = 0$.
Taking $u = -\varphi_0/\sqrt{2E}$ *in* (89), *we have*:

$$L(\varphi_0, \theta, E; \varepsilon) = L \left(0, \theta - \varepsilon v \varphi_0/\sqrt{2E}, E; \varepsilon \right).$$

In other words, L *is a quasi-periodic function of* φ_0 *with frequency* $\varepsilon v/\sqrt{2E}$.
(3) $M(\varphi_0, \theta, E; \varepsilon) = \partial_\theta L(\varphi_0, \theta, E; \varepsilon)$.

In the following lemma we are going to give an approximation of the Melnikov potential $L(\varphi_0, \theta, E; \varepsilon)$ in terms of a function $\mathcal{L}(\theta)$, which we will call Poincaré function.

Lemma 4.26. *Under the assumptions of Proposition* 4.15

$$L(\varphi_0, \theta, E; \varepsilon) = \frac{1}{\sqrt{2E}}\mathcal{L}\left(\theta - \varepsilon v \frac{\varphi_0}{\sqrt{2E}}\right) + O_{\mathcal{C}^2}(\varepsilon), \tag{90}$$

where

$$\begin{aligned}\mathcal{L}(\theta) = -\lim_{T_1, T_2 \to \infty} \Big[&\int_{-T_1}^{+T_2} dt\, \widetilde{U}(\gamma^q_{1/2}(t), \theta) \\ &- \int_{-T_1+a_-}^{+T_2+a_+} dt\, \widetilde{U}(\Lambda^q_{1/2}(t), \theta)\Big]\end{aligned}$$

is called the Poincaré function and was already defined in (16).

Proof. In order to obtain the first-order terms in the Melnikov potential we write (88), with $\varphi_0 = 0$, as

$$\begin{aligned} L(0, \theta, E; \varepsilon) = \lim_{T_1, T_2 \to \infty} \Bigg[& - \int_{-T_1}^{T_2} dt\, \widetilde{U}\left(\gamma^q_E(t), \theta + \varepsilon v t\right) \\ & + \tilde{h}\left(a_+ + \sqrt{2E}\, T_2, \theta + \varepsilon v T_2; \varepsilon\right) - \tilde{h}\left(a_+, \theta; \varepsilon\right) \\ & - \tilde{h}\left(a_- - \sqrt{2E}\, T_1, \theta - \varepsilon v T_1; \varepsilon\right) + \tilde{h}\left(a_-, \theta; \varepsilon\right) \\ & + \tilde{h}\left(a_+, \theta; \varepsilon\right) - \tilde{h}\left(a_- + \Delta, \theta + \varepsilon v \Delta/\sqrt{2E}; \varepsilon\right) \\ & + \tilde{h}\left(a_- + \Delta, \theta + \varepsilon v \Delta/\sqrt{2E}; \varepsilon\right) - \tilde{h}(a_-, \theta; \varepsilon) \Bigg]. \end{aligned} \tag{91}$$

The terms in the fourth line of expression (91) are of order $O_{\mathcal{C}^2}(\varepsilon)$ due to the fact that $\|\tilde{h}(\cdot, \cdot; \varepsilon)\|_{\mathcal{C}^{l-\tau}}$ is bounded and $l - \tau > 2$, and $a_- + \Delta = a_+$.

To obtain integral expressions for the terms in the other lines of Eq. (91), it suffices to use the fundamental theorem of Calculus and the functional equation (64) verified by $\tilde{h}$. Thus,

$$L(0, \theta, E; \varepsilon)$$

$$= \lim_{(T_1, T_2) \to \infty} \Bigg[- \int_{-T_1}^{0} dt\, \widetilde{U}\left(\gamma^q_E(t), \theta + \varepsilon v t\right) - \widetilde{U}\left(\Lambda^q_{1/2}\left(a_- + \sqrt{2E} t\right), \theta + \varepsilon v t\right)$$

$$
-\int_0^{T_2} dt\, \widetilde{U}\left(\gamma_E^q(t), \theta+\varepsilon vt\right) - \widetilde{U}\left(\Lambda_{1/2}^q\left(a_+ + \sqrt{2E}t\right), \theta+\varepsilon vt\right)
$$

$$
\left. + \int_0^{\Delta/\sqrt{2E}} dt\, \widetilde{U}\left(\Lambda_{1/2}^q\left(a_- + \sqrt{2E}t\right), \theta+\varepsilon vt\right)\right] + O_{\mathcal{C}^2}(\varepsilon),
$$

or equivalently, by the change of variable $u = \sqrt{2E}t$, and using the scaling properties (7) for γ_E:

$$
\begin{aligned}
L(0, \theta, E; \varepsilon) = & -\frac{1}{\sqrt{2E}} \\
& \times \lim_{T_1,T_2\to\infty}\left[\int_{-T_1\sqrt{2E}}^{0} du\, \widetilde{U}\left(\gamma_{1/2}^q(u), \theta+\varepsilon v\frac{u}{\sqrt{2E}}\right)\right. \\
& -\widetilde{U}\left(\Lambda_{1/2}^q(a_-+u), \theta+\varepsilon v\frac{u}{\sqrt{2E}}\right) \\
& +\int_0^{T_2\sqrt{2E}} du\, \widetilde{U}\left(\gamma_{1/2}^q(u), \theta+\varepsilon v\frac{u}{\sqrt{2E}}\right) \\
& -\widetilde{U}\left(\Lambda_{1/2}^q(a_++u), \theta+\varepsilon v\frac{u}{\sqrt{2E}}\right) \\
& \left.-\int_0^{\Delta} du\, \widetilde{U}\left(\Lambda_{1/2}^q(a_-+u), \theta+\varepsilon v\frac{u}{\sqrt{2E}}\right)\right] + O_{\mathcal{C}^2}(\varepsilon),
\end{aligned}
$$

and taking the dominant terms in ε,

$$
\begin{aligned}
& L(0, \theta, E; \varepsilon) \\
& = -\frac{1}{\sqrt{2E}} \lim_{T_1,T_2\to\infty}\left[\int_{-T_1\sqrt{2E}}^{0} du\, \widetilde{U}\left(\gamma_{1/2}^q(u), \theta\right) - \widetilde{U}\left(\Lambda_{1/2}^q(a_-+u), \theta\right)\right. \\
& \quad + \int_0^{T_2\sqrt{2E}} du\, \widetilde{U}\left(\gamma_{1/2}^q(u), \theta\right) - \widetilde{U}\left(\Lambda_{1/2}^q(a_++u), \theta\right) \\
& \quad \left. - \int_0^{\Delta} du\, \widetilde{U}\left(\Lambda_{1/2}^q(a_-+u), \theta\right)\right] + \frac{1}{\sqrt{2E}}R(\theta,\varepsilon) + O_{\mathcal{C}^2}(\varepsilon) \\
& = \frac{1}{\sqrt{2E}}\mathcal{L}(\theta) + \frac{1}{\sqrt{2E}}R(\theta,\varepsilon) + O_{\mathcal{C}^2}(\varepsilon),
\end{aligned}
$$

where $R(\theta, \varepsilon)$ is defined so that the above is an identity. Note that it only involves the difference of integrals whose integrands have second arguments that differ by $\varepsilon v\frac{u}{\sqrt{2E}}$. This is a reflection of the fact that the potential is a slow perturbation.

We bound $R(\theta, \varepsilon)$ using properties (71) and the fact that $\widetilde{U}(q, \theta)$ is a periodic function with respect to its second variable θ, as

$$
\begin{aligned}
|R(\theta, \varepsilon)| &\leqslant K\varepsilon \left(\int_{-\infty}^{+\infty} du \ \ |u|\, e^{-\beta|u|} + \int_0^{\Delta} |u| \ du \right) \\
&\leqslant C\varepsilon.
\end{aligned}
$$

Similarly, one can bound the first and second derivatives of $R(\theta, \varepsilon)$ because one can take derivatives under the integral sign (the convergence of the integrand is exponentially fast) and then, similar cancellations than those used above, establish

$$
L(0, \theta, E; \varepsilon) = \frac{1}{\sqrt{2E}} \mathcal{L}(\theta) + O_{\mathcal{C}^2}(\varepsilon),
$$

Replacing θ by $\theta - \varepsilon v \frac{\varphi_0}{\sqrt{2E}}$, and using property 2 of Proposition 4.25, we obtain the desired conclusions in Lemma 4.26. $\square$

From item (3) of Proposition 4.25 and Eq. (90) is clear that

$$
M(\varphi_0, \theta, E; \varepsilon) = \frac{1}{\sqrt{2E}} \frac{\partial}{\partial \theta} \left[\mathcal{L} \left(\theta - \varepsilon v \frac{\varphi_0}{\sqrt{2E}} \right) \right] + O_{\mathcal{C}^1}(\varepsilon). \tag{92}
$$

By formula (84) we have

$$
\begin{aligned}
H_\varepsilon(x_+^*) - H_\varepsilon(y) &= -\varepsilon^3 v \cdot M(\varphi_0, \theta, E; \varepsilon) + O_{\mathcal{C}^1}(\varepsilon^4) \\
&= -\frac{\varepsilon^3}{\sqrt{2E}} v \cdot \partial_\theta \mathcal{L} \left(\theta - \varepsilon v \frac{\varphi_0}{\sqrt{2E}} \right) + O_{\mathcal{C}^1}(\varepsilon^4).
\end{aligned} \tag{93}
$$

To establish the existence of heteroclinic orbits in the extended phase space, it suffices that the leading term in (93) is non-constant as a function of θ and φ_0, that is, the function

$$
\theta \in \mathbb{T}^d \mapsto v \cdot \frac{\partial \mathcal{L}}{\partial \theta}(\theta) \in \mathbb{R} \tag{94}
$$

is not identically zero, where $\mathcal{L}$ is the Poincaré function defined in (16). This is equivalent to hypothesis (iv) in Theorem 3.4. Indeed, given a smooth function $f : \mathbb{T}^d \to \mathbb{R}$ and a rational independent vector $v \in \mathbb{R}^d$, the equations $v \cdot \partial_\theta f \equiv 0$ and $f \equiv$ cte. are equivalent, as can be easily checked, for example, writing these equations in Fourier coefficients.

Roughly speaking, the fact that the Poincaré function (16) is non-trivial will imply that all the tori with sufficiently high energy get moved by the scattering map by an amount that can be bounded from below by ε^3.

We point out that the work that we have done simplifying the Melnikov potential, leads to a condition which is independent of the energy.

Note also that the condition that $\mathcal{L}$ is non-constant is satisfied except for very exceptional potentials. We also note that the Poincaré function $\mathcal{L}$ in (16) depends on the periodic orbit $\Lambda_{1/2}$ and the homoclinic connection $\gamma_{1/2}$ chosen. If there is one such orbit $\gamma_{1/2}$, the theory of dynamical systems shows that there are infinitely many. To establish the existence of orbits with unbounded energy, it suffices to verify that the potential U satisfies hypothesis (iv) in Theorem 3.4 for one Λ, γ. Hence, for the systems that we consider, the potentials which do not lead to diffusion are extremely rare.

4.5. Existence of transition chains and orbits of unbounded energy

The goal of this Section is to provide a proof of Theorem 4.27, which establishes the existence of orbits of Hamiltonian (15) whose energy follows largely arbitrary excursions (including tending to ∞).

Theorem 4.27 clearly implies Theorem 3.4, which in turn implies Theorem 1.3.

Note that Theorem 4.27 is expressed in the original unscaled variables (just as in Theorems 1.3 and 3.4).

The proof of Theorem 4.27 will consist in showing the existence of largely arbitrary transition paths (see Definition 4.32) and, then, showing the existence of orbits that shadow the transition paths.

The proof of the existence of transition paths, being a local problem, will be carried out in the scaled variables, but the shadowing, being a more global problem, will be carried out in the physical variables.

Theorem 4.27. *Given* $E_0 > 0$, *and any continuous function* $\mathcal{E} : [0, \infty) \to [E_0, \infty)$, *and any* $K > 0$, *consider the set*

$$\mathcal{I}_K^{\mathcal{E}(s)} = \{E \in [E_0, \infty) \ : \quad |E - \mathcal{E}(s)| \leqslant K\}. \tag{95}$$

Under the assumptions of Theorem 3.4 *we can find* E_0 *sufficiently large and* $K > 0$ *such that for any function* $\mathcal{E}$ *as above*, *there exists a monotone function* $T : [0, \infty) \to [0, \infty)$ *and an orbit* $x(t)$ *of the Hamiltonian* H *given in* (15) *such that*

$$H(x(T(s)), T(s)) \in \mathcal{I}_K^{\mathcal{E}(s)}.$$

The meaning of Theorem 4.27 is that we can find orbits that follow somewhat arbitrary changes in the energy—given by the arbitrarily prescribed function $\mathcal{E}(s)$—up to a certain error—given by K—. We make no assertions in this paper about the time it takes to accomplish these changes (this is the role of the function T which reparameterizes the time). Nevertheless, we can guarantee that the energy of the orbits does not stray very much from the desired goal.

Since we are allowing reparameterizations, the only features that are relevant of the function $\mathcal{E}$ are the values of its local maxima and minima.

We emphasize that K can be chosen for all functions $\mathcal{E}$ and for all E_0 sufficiently large. It only depends on the Hamiltonian.

4.5.1. Whiskered tori

In this section we collect some definitions of the objects we are considering.

A class of objects that play an important role in our argument are whiskered tori and transition paths among them. Their role in diffusion was emphasized already in [Arn64]. See also [AA67].

In this section we collect some rather standard definitions and rather standard facts.

Definition 4.28. We say that $\mathcal{T}$ is a whiskered torus for a n degrees of freedom Hamiltonian on a $2n$-dimensional symplectic manifold M when:

(a) $\mathcal{T}$ is a $\mathcal{C}^1$ embedded $\mathbb{T}^k$, $1\leqslant k < n$: $\mathcal{T} = v(\mathbb{T}^k)$.
(b) $\mathcal{T}$ is invariant under the Hamiltonian flow Φ_t.
(c) It is possible to choose v, the embedding of $\mathbb{T}^k$ above, in such a way that the Hamiltonian flow is an irrational rotation, that is

$$\Phi_t \circ v(\theta) = v(\theta + \omega t), \tag{96}$$

where $\omega \in \mathbb{R}^k$ satisfies

$$\omega \cdot \ell \neq 0 \quad \forall \ell \in \mathbb{Z}^k - \{0\}.$$

(d) The symplectic form restricted to $\mathcal{T}$ vanishes. (This property is usually referred to saying that the manifold $\mathcal{T}$ is isotropic.)
(e) There exist $n-k$ dimensional bundles $E^{\mathrm{s}}_{v(\theta)}, E^{\mathrm{u}}_{v(\theta)} \subset T_{v(\theta)}M$ such that for some $C > 0$, $0 < \lambda < 1$, we have

$$\begin{aligned} &D\Phi_t(v(\theta))E^{\mathrm{s,u}}_{v(\theta)} = E^{\mathrm{s,u}}_{v(\theta+\omega t)},\\ &||D\Phi_t(v(\theta))|_{E^{\mathrm{s}}_{v(\theta)}}|| \leqslant C\lambda^t, \quad t\geqslant 0,\\ &||D\Phi_t(v(\theta))|_{E^{\mathrm{u}}_{v(\theta)}}|| \leqslant C\lambda^{|t|}, \quad t\leqslant 0. \end{aligned}$$

Remark 4.29. A well-known argument shows that, when the symplectic form is exact—as is the case in the cotangent bundles we are considering in this paper—property (c) implies property (d). Hence, for an exact symplectic manifold, it is only necessary to assume (c). □

We note that whiskered tori are not normally hyperbolic invariant manifolds. Indeed note that since the symplectic form vanishes on $\mathcal{T}$, we have that the k-dimensional space $N_{v(\theta)}$ symplectically conjugate to $T_{v(\theta)}\mathcal{T}$ satisfies $N_{v(\theta)} \cap T_{v(\theta)}\mathcal{T} = \{0\}$. Moreover, by the preservation of the symplectic form, the bundle N is invariant under $D\Phi_t$ and it is impossible to have for some $C > 0$, $0 < \lambda < 1$ inequalities $||D\Phi_t|_N|| \leqslant C\lambda^{|t|}$ for either all $t \geqslant 0$ or all $t \leqslant 0$.

Therefore it cannot be true that $T_{v(\theta)}M$ is spanned by $T_{v(\theta)}\mathcal{T}$ and directions $E^{\mathrm{s,u}}_{v(\theta)}$ that contract either in the future or in the past as required by the definition of normally hyperbolic manifolds.

Nevertheless, we note that

$$\dim(T_{v(\theta)}\mathcal{T}) + \dim(E^{\mathrm{s}}_{v(\theta)}) + \dim(E^{\mathrm{u}}_{v(\theta)}) = k + (n-k) + (n-k) = 2n - k.$$

Hence, the whiskered tori contain as many hyperbolic directions as allowed by the geometric properties of supporting a rotation and being isotropic. Therefore, in Hamiltonian mechanics literature, where these properties are quite common, whiskered tori are sometimes called *hyperbolic tori*. Since in this paper we use theorems for normally hyperbolic invariant manifolds, many of which are not true for whiskered tori, we keep the distinction between normally hyperbolic invariant manifolds and whiskered ones.

Even if the whiskered tori are not normally hyperbolic, it is true that there exists a gap in the spectrum of the linearization acting on the tangent bundle. Standard results in the theory of normally hyperbolic invariant manifolds (see e.g. [Fen74]) show that there exist immersed manifolds $W^{\mathrm{s}}_{\mathcal{T}}$, $W^{\mathrm{u}}_{\mathcal{T}}$ which are characterized by

$$\begin{aligned} W^{\mathrm{s}}_{\mathcal{T}} &= \{x \ : \ \mathrm{dist}(\Phi_t(x), \mathcal{T}) \leqslant C_x(1-\delta)^t \quad t \geqslant 0\} \\ &= \{x \ : \ \mathrm{dist}(\Phi_t(x), \mathcal{T}) \leqslant C_x(\lambda+\delta)^t \quad t \geqslant 0\}, \end{aligned}$$

$$\begin{aligned} W^{\mathrm{u}}_{\mathcal{T}} &= \{x \ : \ \mathrm{dist}(\Phi_t(x), \mathcal{T}) \leqslant C_x(1-\delta)^{|t|} \quad t \leqslant 0\} \\ &= \{x \ : \ \mathrm{dist}(\Phi_t(x), \mathcal{T}) \leqslant C_x(\lambda+\delta)^{|t|} \quad t \leqslant 0\}. \end{aligned}$$

Note that the above equations show that, as soon as the convergence to the manifold is exponential with any rate, it is exponential with a rate close to that given by λ. See [Fen74].

If the Hamiltonian flow Φ_t is $\mathcal{C}^r$ $r \in \mathbb{N} + [0, \mathrm{Lip}] \cup \{\infty, \omega\}$, the manifolds $W^{\mathrm{s,u}}_{\mathcal{T}}$ are $\mathcal{C}^r$.

When $r \in \mathbb{N} + [0, \mathrm{Lip}]$ the above regularity conclusion follows by the standard theory of normally hyperbolic invariant manifolds noting that the motion on the bundles $N + T\mathcal{T}$ is bounded by arbitrarily small exponentials while the stable and unstable bundles are exponentially contracting.

The cases $r = \infty, \omega$ do not follow from the standard arguments on normally hyperbolic invariant manifolds. They use essentially that the motion on the manifold is a rotation. The proof for $r = \infty, \omega$ can be found in [dlLW04].

Another important fact is that

$$W_{\mathcal{T}}^{\mathrm{s,u}} = \cup_{x\in\mathcal{T}} W_x^{\mathrm{s,u}},$$

where

$$\begin{aligned} W_x^{\mathrm{s}} &= \{y \ : \ \mathrm{dist}(\Phi_t(x), \Phi_t(y)) \leqslant C(1-\delta)^t, \quad t \geqslant 0\} \\ &= \{y \ : \ \mathrm{dist}(\Phi_t(x), \Phi_t(y)) \leqslant C(\lambda+\delta)^t, \quad t \geqslant 0\}, \end{aligned}$$

$$\begin{aligned} W_x^{\mathrm{u}} &= \{y \ : \ \mathrm{dist}(\Phi_t(x), \Phi_t(y)) \leqslant C(1-\delta)^{|t|}, \quad t \leqslant 0\} \\ &= \{y \ : \ \mathrm{dist}(\Phi_t(x), \Phi_t(y)) \leqslant C(\lambda+\delta)^{|t|}, \quad t \leqslant 0\}. \end{aligned}$$

The standard theory of [Fen74,HPS77] shows that if the flow Φ_t is $\mathcal{C}^r$, $r \in \mathbb{N} + [0, \mathrm{Lip}] \cup \{\infty, \omega\}$, the manifolds W_x^s, W_x^u are $\mathcal{C}^r$. This part of the argument does not use many properties of the motion on the whiskered tori and holds in much greater generality. The results about regularity of $W_{\mathcal{T}}^s$ mentioned above can be understood as regularity statements about the map $x \mapsto W_x^s$, which is rather regular for whiskered tori but not for general invariant manifolds with stable and unstable bundles.

For our applications, it is important to mention that the notion of whiskered torus can be adapted without difficulty to quasi-periodic vector fields.

Given a vector field of the form

$$\begin{aligned} \dot{x} &= F(x, \theta), \\ \dot{\theta} &= \nu, \end{aligned} \tag{97}$$

where $x \in M$, $\theta \in \mathbb{T}^d$, we obtain a flow $\tilde{\Phi}_t(x, \theta) = (\Gamma_t(x, \theta), \theta + t\nu)$. A whiskered torus is an embedding $\tilde{\upsilon} : \mathbb{T}^k \times \mathbb{T}^d \to M \times \mathbb{T}^d$ of the form

$$\tilde{\upsilon}(\varphi, \theta) = (\upsilon(\varphi, \theta), \theta) \tag{98}$$

satisfying the properties included in Definition 4.28 for the extended flow (97).

We note that the equation of invariance for the extended flow required in Definition 4.28 is

$$\tilde{\Phi}_t(\tilde{\upsilon}(\varphi, \theta)) = \tilde{\upsilon}(\varphi + t\omega, \theta + t\nu).$$

This amounts, in terms of the maps of the manifold to

$$\Gamma_t(\upsilon(\varphi, \theta), \theta) = \upsilon(\varphi + t\omega, \theta + t\nu).$$

When we consider symplectic flows, we note that the mapping $x \mapsto \Gamma_t(x, \theta)$ is symplectic. In case that we study symplectic structures, it is natural to consider not only the extended flows (97) but also the full symplectic flows which require adding action variables $A \in \mathbb{R}^d$ conjugate to the angles $\theta \in \mathbb{T}^d$ with respect to the standard symplectic form $\sum_{j=1}^d dA_i \wedge d\theta_i$.

We note that if the symplectic form Ω in the manifold M is exact, i.e. $\Omega = d\alpha$, then the symplectic form Ω^* in the full symplectic manifold $M^* \equiv M \times \mathbb{R}^d \times \mathbb{T}^*$ is exact: $\Omega^* = d(\alpha + \sum_{j=1}^d A_i d\theta_i)$.

A quasi-periodic flow corresponds to a Hamiltonian $\mathcal{H}(x, \theta, A)$ of the form $\mathcal{H}(x, \theta, A) = H(x, \theta) + \nu A$. Note that the equations of motion are:

$$\begin{aligned} \dot{x} &= J\nabla_x H(x, \theta), \\ \dot{\theta} &= \nu, \\ \dot{A} &= -\nabla_\theta H(x, \theta). \end{aligned} \tag{99}$$

Note that the equations of motion of the variables x, θ determine the motion of A just by quadrature. It is clear that from a quasi-periodic torus in the full symplectic system (99), just by ignoring the variable A, we obtain a whiskered torus in the extended system (97).

On the other hand, given a whiskered torus in the extended system (97), we can wonder if it is possible to find a similar embedding in system (99). This is not true for Hamiltonian systems in general, since it can happen that the last equation of (99) has secular terms. Similarly, when integrating a quasi-periodic function, if the number is not Diophantine, it could happen that the integral is not quasi-periodic. See, for example [Fur61].

Nevertheless, the following proposition shows that if the torus is Diophantine and the system is exact and differentiable enough (depending on the exponent of the Diophantine number) then, both definitions of whiskered tori in quasiperiodic systems agree. We note that all the quasi-periodic solutions that we will consider, since they are produced invoking the KAM theorem, satisfy the hypothesis.

Proposition 4.30. *If the form Ω is exact, given a whiskered torus with embedding $\tilde{v}$* (98) *in the extended system* (97), *we can find a function $a : \mathbb{T}^k \times \mathbb{T}^d \to \mathbb{R}^d$ such that the mapping $v^* : \mathbb{T}^k \times \mathbb{T}^d \to M \times \mathbb{T}^d \times \mathbb{R}^d$ defined by*

$$v^*(\varphi, \theta) = \big(\tilde{v}(\varphi, \theta), a(\varphi, \theta)\big)$$

is a whiskered torus in the sense of Definition 4.28 *for the full symplectic system* (99).

Proof. Given $F : \mathbb{T}^d \to \mathbb{R}$, we recall that the equation

$$\dot{A} = F(\omega t)$$

has a formal quasiperiodic solution $A(t) = G(\omega t)+ < F > t$, where $< F >$ denotes the average of the periodic function F, and $G : \mathbb{T}^d \to \mathbb{R}$ is determined, in Fourier coefficients by

$$\hat{G}_k = 2\pi < \omega, k >^{-1} \hat{F}_k$$

The above formal solution is a genuine solution when ω is Diophantine and F is differentiable enough (see the results from [Rus75] summarized in Lemma A.23).

It then follows that Eqs. (99) evaluated on a quasiperiodic solition have a solution in which A is of the form $A = \beta t + G(\omega t)$ for some constant $\beta \in \mathbb{R}$ and then, all the other variables change quasi-periodically.

The desired result will be established when we show that $\beta = 0$.

We note that if we consider the evolution of the whole embedded torus it is just translated by βt along the A direction. Using that the flow is exact, the integral of the symplectic form Ω^* along a loop of the form given by letting only θ_i vary while keeping all the other variables, should be independent of t. On the other hand, a direct calculation shows that the change of the integral of α along this loop changes by β_i. This allows us to conclude that $\beta = 0$. $\square$

We note that the KAM tori $\mathcal{T}_\omega(\varepsilon)$ inside $\tilde{\Lambda}_\varepsilon$ produced in Section 4.3.6 are whiskered tori for a quasi-periodic Hamiltonian system. Analogously, the KAM tori $\mathcal{T}^*_{\omega,B}(\varepsilon)$ are whiskered tori of the full symplectic system.

4.5.2. Transition paths

The idea of transition chain goes back to Arnold [Arn64] who introduced a transition chain as a sequence of whiskered tori $\{\mathcal{T}_i\}_{i=K_-}^{K_+}$ such that

$$W^{\mathrm{s}}_{\mathcal{T}_{i+1}} \pitchfork W^{\mathrm{u}}_{\mathcal{T}_i}.$$

In this paper, we formulate explicitly the more precise concept of *transition path*. See Definition 4.32. A transition path specifies not only the sequence of tori but also a choice of connecting orbits. Including the connecting orbits allows us to formulate more precise results. We will prove that, given a connecting path, there are shadowing orbits which never move at a distance from the path. (We believe that this result was more or less folklore, but M. Sevryuk pointed to us that it was not formulated in the literature.) The fact that the diffusing orbit stays at a bounded distance from the transition path provides control of the energy of the diffusing orbit for all times.

Also, since usually there are many connecting orbits, this will give rise to different transition paths and, hence, different diffusing orbits.

Definition 4.31. We say that γ is a connecting orbit between two whiskered tori $\mathcal{T}^+$, $\mathcal{T}^-$ when

(a) $\gamma \subset W^{\mathrm{s}}_{\mathcal{T}^+} \cap W^{\mathrm{u}}_{\mathcal{T}^-}$. That is

$$\lim_{t\to\pm\infty} \mathrm{dist}(\gamma(t), \mathcal{T}_\pm) = 0.$$

(b) The intersection of $W^{\mathrm{s}}_{\mathcal{T}^+}$ and $W^{\mathrm{u}}_{\mathcal{T}^-}$ is transversal along γ. That is

$$T_{\gamma(t)} W^{\mathrm{s}}_{\mathcal{T}^+} \oplus T_{\gamma(t)} W^{\mathrm{u}}_{\mathcal{T}^-} = T_{\gamma(t)} M.$$

Sometimes, one speaks about connecting orbits without assuming that the intersection is transversal. In this paper, all the connecting orbits that we consider are transversal.

Definition 4.32. We say that a sequence of whiskered tori and connecting orbits $\{\mathcal{T}_i\}_{i=K_-}^{K_+}$, $\{\gamma_{i+1}\}_{i=K_-}^{K_+-1}$(possibly infinite, i.e. $K_- \in \mathbb{Z} \cup \{-\infty\}$, $K_+ \in \mathbb{Z} \cup \{\infty\}$, $K_+ > K_-$) is a transition path when

$$\gamma_{i+1} \subset W^{\mathrm{s}}_{\mathcal{T}_{i+1}} \pitchfork W^{\mathrm{u}}_{\mathcal{T}_i}$$

4.5.3. Existence of transition paths for systems as in Theorem 3.4

We recall that our study of Hamiltonian flows satisfying H1′, H2′ has produced three types of results:

- Studying the dynamics inside $\tilde{\Lambda}_\varepsilon$, we have shown in Proposition 4.10 that if the system is differentiable enough, the manifold $\tilde{\Lambda}_\varepsilon$ is covered with KAM tori that leave gaps between them smaller that $O_{\mathcal{C}^1}(\varepsilon^{\frac{m+1}{4}})$. If we assume that m, the number of steps of averaging taken in Theorem 4.7 is large enough, $m \geqslant 12$, which can be done provided that the system is smooth enough, the size of the gaps is at most $\varepsilon^{3+1/4}$.
- We have obtained in Proposition 4.15 an expression (63) for the energy H_ε introduced in (35) of the KAM tori. In particular, for any invariant torus $\mathcal{T}$, the energy H_ε experiences an oscillation of order ε^2:

$$\max H_\varepsilon(\mathcal{T}) - \min H_\varepsilon(\mathcal{T}) = O(\varepsilon^2). \tag{100}$$

- Studying the scattering map, we have shown in Eq. (93) that, provided that the Poincaré function (16) is non-constant, the homoclinic excursions connect tori whose distance is $O(\varepsilon^3)$.

In the next lemmas, we put together the information to show that we can construct transition paths whose energy changes more or less arbitrarily.

Lemma 4.33. *Consider the Hamiltonian system of Hamiltonian* (35) *and assume that the Poincaré function* (16) *is non-constant.*

Then, under the assumptions of Proposition 4.15 *with* $m \geqslant 12$, *there exist intervals* $I^\pm = [\alpha_1^\pm, \alpha_2^\pm]$, *with* $\alpha_i^- < 0$, $\alpha_i^+ > 0$, $i = 1, 2$, *such that, for* ε *sufficiently small, given a KAM torus* $\mathcal{T}_\omega(\varepsilon) \subset \tilde{\Lambda}_\varepsilon$ *of frequency* $(\omega, \varepsilon\nu)$ *in the extended phase space, there exist KAM tori* $\mathcal{T}_{\omega^+}(\varepsilon)$, $\mathcal{T}_{\omega^-}(\varepsilon)$ *produced by KAM theorem* 4.8 (*see Remark* 4.11) *which* *verify*

$$E^{\pm} - E := \frac{(\omega^{\pm})^2}{2} - \frac{\omega^2}{2} \in \varepsilon^3 I^{\pm},$$

with connecting orbits

$$\gamma^{+} \subset W^{\mathrm{s}}_{\mathcal{T}_{\omega^{+}}(\varepsilon)} \pitchfork W^{\mathrm{u}}_{\mathcal{T}_{\omega}(\varepsilon)}, \quad \gamma^{-} \subset W^{\mathrm{s}}_{\mathcal{T}_{\omega}(\varepsilon)} \pitchfork W^{\mathrm{u}}_{\mathcal{T}_{\omega^{-}}(\varepsilon)}.$$

Moreover, $\gamma^{\pm} \subset \tilde{\gamma}_{\varepsilon}$, *the intersection manifold we used in the construction of the scattering map*, *and they verify*

$$\begin{aligned} \operatorname{dist}(\gamma^{\pm}, \gamma_E) &= O_{\mathcal{C}^1}(\varepsilon^2), \\ \left| H_{\varepsilon}(\gamma^{\pm}) - E \right| &= O_{\mathcal{C}^1}(\varepsilon^2). \end{aligned} \tag{101}$$

Proof. A very similar lemma for periodic perturbations is established in [DLS00, Lemma 4.21].

The only difference between the argument in [DLS00] and the present paper is that now, as the perturbation is quasi-periodic, the Poincaré function $\mathcal{L}$ depends on a variable $\theta \in \mathbb{T}^d$ rather than on a variable on the circle.

We know that, by Proposition 4.23, the effect of the scattering map acting on an invariant torus is to add a term $O(\varepsilon^3)$ to the function G_{ω}—given by Proposition 4.15—whose graph in the coordinates $(H_{\varepsilon}, \varphi, \theta)$ represents the invariant torus.

Under the assumption that the Poincaré function $\mathcal{L}$ is non-constant, this term is non-trivial.

More concretely, if a torus $\mathcal{T}_{\omega}(\varepsilon)$ is represented as the graph of the function G_{ω}, then, by Proposition 4.23, $\tilde{S}_{\varepsilon}(\mathcal{T}_{\omega})$ will be represented as the graph of the function

$$\begin{aligned} \hat{G}_{\omega} &= G_{\omega} + \varepsilon^3 v \cdot M_{\varepsilon,E}(\varphi, \theta) + O_{\mathcal{C}^1}(\varepsilon^4) \\ &= G_{\omega} + \frac{\varepsilon^3}{\sqrt{2E}} v \cdot \partial_{\theta} \mathcal{L}\left(\theta - \varepsilon v \frac{\varphi}{\sqrt{2E}}\right) + O_{\mathcal{C}^1}(\varepsilon^4). \end{aligned}$$

By the assumption that $\mathcal{L}(\theta)$ is not trivial, we conclude that there is a non-empty open set $\mathcal{U} \subset \mathbb{T}^d$ and a constant $C_0 > 0$, such that $\forall \theta \in \mathcal{U}$ we have that $\det |\partial_{\theta\theta} \mathcal{L}(\theta)| \geqslant C_0$.

We choose $I^{\pm} = [\alpha_1^{\pm}, \alpha_2^{\pm}] \subset \frac{v}{\sqrt{2E}} \cdot \partial_{\theta} \mathcal{L}(\mathcal{U})$. For any other torus $\mathcal{T}_{\tilde{\omega}}$ such that $\tilde{E} - E = \frac{\tilde{\omega}^2}{2} - \frac{\omega^2}{2} \in \varepsilon^3 I^{+}$, we have that for any $(\varphi, \theta) \in \mathbb{T} \times \mathbb{T}^d$:

$$\hat{G}_{\omega}(\varphi, \theta) - G_{\tilde{\omega}}(\varphi, \theta) = \frac{\omega^2}{2} - \frac{\tilde{\omega}^2}{2} + \frac{\varepsilon^3}{\sqrt{2E}} v \cdot \partial_{\theta} \mathcal{L}\left(\theta - \varepsilon v \frac{\varphi}{\sqrt{2E}}\right) + O_{\mathcal{C}^1}(\varepsilon^4).$$

By definition of I^+ there exists (φ^*, θ^*) for which $\hat{G}_\omega(\varphi^*, \theta^*) - G_{\tilde{\omega}}(\varphi^*, \theta^*) = 0$ and $\theta^* - \varepsilon v \frac{\varphi^*}{\sqrt{2E}} \in \mathcal{U}$, so $\tilde{S}_\varepsilon(\mathcal{T}_\omega(\varepsilon))$ intersects $\mathcal{T}_{\tilde{\omega}}$.

Since

$$\partial_\theta \hat{G}_\omega(\varphi^*, \theta^*) - \partial_\theta G_{\tilde{\omega}}(\varphi^*, \theta^*) = -\frac{\varepsilon^3}{\sqrt{2E}} v \cdot \partial_{\theta\theta}\mathcal{L}\left(\theta^* - \varepsilon v \frac{\varphi^*}{\sqrt{2E}}\right) + O_{\mathcal{C}^0}(\varepsilon^4),$$

and $\partial_{\theta\theta}\mathcal{L}$ is non-degenerate in $\mathcal{U}$, we conclude that the intersection between $\tilde{S}_\varepsilon(\mathcal{T}_\omega(\varepsilon))$ and $\mathcal{T}_{\tilde{\omega}}$ is transversal in $\tilde{\Lambda}_\varepsilon$.

The existence of tori $\mathcal{T}_{\omega^+(\varepsilon)}$ whose frequency verifies $\frac{(\omega^+)^2}{2} - \frac{\omega^2}{2} \in \varepsilon^3 I^+$ is guaranteed by KAM theorem 4.8 and Proposition 4.10 if $m \geqslant 12$.

Hence, we have that there exists $\mathcal{T}_{\omega^+}(\varepsilon)$ such that

$$\tilde{S}_\varepsilon(\mathcal{T}_\omega(\varepsilon)) \pitchfork_{\tilde{\Lambda}_\varepsilon} \mathcal{T}_{\omega^+}(\varepsilon). \tag{102}$$

Now we want to argue that $W^{\mathrm{s}}_{\mathcal{T}_{\omega^+}(\varepsilon)} \pitchfork W^{\mathrm{u}}_{\mathcal{T}_\omega(\varepsilon)}$.

The argument is exactly the same as the argument in [DLS00, p.382]. First, by (102):

$$\left(W^{\mathrm{s}}_{\tilde{S}_\varepsilon(\mathcal{T}_\omega)(\varepsilon)} \cap \tilde{\gamma}_\varepsilon\right) \pitchfork_{\tilde{\gamma}_\varepsilon} \left(W^{\mathrm{s}}_{\mathcal{T}_{\omega^+}(\varepsilon)} \cap \tilde{\gamma}_\varepsilon\right)$$

Second, by the definition of the scattering map, $W^{\mathrm{u}}_{\mathcal{T}_\omega(\varepsilon)} \cap \tilde{\gamma}_\varepsilon = W^{\mathrm{s}}_{\tilde{S}_\varepsilon(\mathcal{T}_\omega(\varepsilon))} \cap \tilde{\gamma}_\varepsilon$. Therefore,

$$\left(W^{\mathrm{u}}_{\mathcal{T}_\omega(\varepsilon)} \cap \tilde{\gamma}_\varepsilon\right) \pitchfork_{\tilde{\gamma}_\varepsilon} \left(W^{\mathrm{s}}_{\mathcal{T}_{\omega^+}(\varepsilon)} \cap \tilde{\gamma}_\varepsilon\right).$$

Since $W^{\mathrm{s}}_{\tilde{\Lambda}_\varepsilon}$ intersects $W^{\mathrm{u}}_{\tilde{\Lambda}_\varepsilon}$ along $\tilde{\gamma}_\varepsilon$, we obtain that $W^{\mathrm{u}}_{\mathcal{T}_\omega(\varepsilon)} \pitchfork W^{\mathrm{s}}_{\mathcal{T}_{\omega^+}(\varepsilon)}$.

An analogous argument gives the existence of $\mathcal{T}_{\omega^-}(\varepsilon)$ such that that $W^{\mathrm{s}}_{\mathcal{T}_\omega(\varepsilon)} \pitchfork W^{\mathrm{u}}_{\mathcal{T}_{\omega^-}(\varepsilon)}$.

Concerning quantitative estimates, we know, by (100), that the energy H_ε on the torus $\mathcal{T}_\omega(\varepsilon)$ experiences an oscillation of $O(\varepsilon^2)$ around $E = \frac{\omega^2}{2}$. By (73) any homoclinic orbit to $\mathcal{T}_\omega(\varepsilon)$ experiences an oscillation of the same order. Since $E^\pm - E = \frac{(\omega^\pm)^2}{2} - \frac{\omega^2}{2} \in \varepsilon^3 I^\pm$, estimates (101) hold. $\square$

So far, we had been working in the scaled variables introduced in Section 4.2, where we introduced the variable $\varepsilon = \frac{1}{\sqrt{E^*}}$. It is important to remark that all the results we have obtained—the KAM theorem and the abundance of heteroclinic intersections—hold for sufficiently large energy E^*.

Now, we formulate the results in terms of the original variables for which the Hamiltonian takes the form (15) in Theorem 3.4. This translation is straightforward by recalling that $E^{\mathrm{scaled}} = E^{\mathrm{phys}}\varepsilon^2$ (see (34)).

As indicated at the beginning of Section 4.2 all the results in this section have been formulated for scaled variables, but we wrote E instead of $\bar{E}$ for E^{scaled} for typographical reasons.

First we note that by Proposition 4.15 and (33), KAM theorem provides the existence of invariant tori given by $H = E^{\mathrm{phys}} + \bar{U}(\theta) + O((E^{\mathrm{phys}})^{-\frac{1}{2}})$, where E^{phys} can be chosen in a set whose graphs are smaller than $\Delta E^{\mathrm{phys}} = O((E^{\mathrm{phys}})^{-\frac{m+1}{8}+1})$.

Lemma 4.33 expressed in the physical variables reads:

Lemma 4.34. *Consider the Hamiltonian system of Hamiltonian* (15) *and assume that the Poincaré function* (16) *is non-constant. Assume that* $r \geqslant \max(m + 5\tau + 16, 3m + 4\tau + 17)$, $m \geqslant 12$.

Then there exist intervals $I^{\pm} = [\alpha_1^{\pm}, \alpha_2^{\pm}]$, *with* $\alpha_i^{-} < 0$, $\alpha_i^{+} > 0$, $i = 1, 2$, *such that, for sufficiently large energies E—or for sufficiently large frequencies* $\omega = \sqrt{2E}$*—, given a KAM torus* $\mathcal{T}_\omega$ *of frequency* (ω, v) *in the extended phase space, there exist KAM tori* $\mathcal{T}_{\omega^+}$, $\mathcal{T}_{\omega^-}$ *produced by KAM theorem* 4.8 *which verify*

$$E^{\pm} - E := \frac{(\omega^{\pm})^2}{2} - \frac{\omega^2}{2} \in \frac{1}{\sqrt{E}} I^{\pm}, \tag{103}$$

with connecting orbits

$$\gamma^{+} \subset W^{\mathrm{u}}_{\mathcal{T}_\omega} \pitchfork W^{\mathrm{s}}_{\mathcal{T}_{\omega^+}}, \quad \gamma^{-} \subset W^{\mathrm{s}}_{\mathcal{T}_\omega} \pitchfork W^{\mathrm{u}}_{\mathcal{T}_{\omega^-}},$$

verifying

$$\left| H(\gamma^{\pm}) - E \right| \leqslant \bar{C}. \tag{104}$$

By applying repeatedly Lemma 4.34, we can obtain the following lemma.

Lemma 4.35. *Given* $E_0 > 0$, *and any continuous function* $\mathcal{E} : [0, \infty) \to [E_0, \infty)$ *and given any* $K > 0$, *consider the set*

$$\mathcal{I}_K^{\mathcal{E}(s)} = \{E \in [E_0, \infty) \ : \ |E - \mathcal{E}(s)| \leqslant K\}. \tag{105}$$

Under the assumptions of Theorem 3.4 *we can find* E_0 *sufficiently large and* $K > 0$ *such that for any function* $\mathcal{E}$ *as above, there exists a monotone sequence* $s_i < s_{i+1}$ *and a transition path* $\{\mathcal{T}_i, \gamma_{i+1}\}_{i=1,\infty}$, *such that*

(1) $|\mathcal{E}(s_i) - \mathcal{E}(s)| \leqslant 1$, *for* $s \in [s_i, s_{i+1}]$.
(2) $H(\mathcal{T}_i) \in \mathcal{I}_K^{\mathcal{E}(s_i)}$.
(3) $H(\gamma_{i+1}) \in \mathcal{I}_K^{\mathcal{E}(s_i)}$.

4.5.4. Existence of orbits shadowing transition paths

Note that Lemma 4.35 is very close to implying Theorem 4.27. It follows immediately from Lemma 4.35 that we obtain Theorem 4.27 if we replace in it "*orbit*" with *transition path*.

The next task, accomplished in this section, is to show that there is a true orbit that follows the transition path thus constructed.

There are many papers in the literature based on very different methods, for example, variational methods, methods based on normal forms, etc. to obtain true orbits through "pseudo-orbits".

In what follows, we will present an elementary point set topology argument that applies to infinite transition paths. We recall that for us, a transition path as defined in 4.32, includes not only the tori but also the connecting orbits between them. Hence, our argument shows that the orbit not only visits the prescribed tori but also that it does not deviate from the prescribed connections. The argument presented here has the advantage that it does not use either differentiability beyond $\mathcal{C}^2$ or the symplectic structure. It also works without much change for finite or infinite chains. Unfortunately, it does not provide any information on the time needed to perform an excursion. The argument is inspired by the arguments in [AA67].

The first step is to use a Lambda-lemma for whiskered tori. The following result is a particular case of the results of [FM00]. (Related results appear in [Cre97]. See also [FM03].) The result of [FM00] does not need that the map is symplectic and applies to whiskered tori in a general map. Nevertheless, since we have made all our definitions to be used in our applications where there is a symplectic structure, we state the result for quasi-periodic symplectic maps and quasi-periodic whiskered tori as defined in Section 4.5.1.

Lemma 4.36. *Let f be a $\mathcal{C}^2$ quasi-periodic symplectic mapping in a $2n+d$ manifold. Assume that the map leaves invariant a $\mathcal{C}^1$ $d+1$-dimensional whiskered torus $\mathcal{T}$ and that the motion on the torus is an irrational rotation. Let Γ be a l-dimensional manifold, $l \geqslant n$, intersecting $W^{\mathrm{s}}_{\mathcal{T}}$ transversally.*

Then there exists a neighborhood $\mathcal{V}$ of $\mathcal{T}$ such that for any $\mathcal{U} \subset \mathcal{V}$, then

$$\mathcal{U} \cap W^{\mathrm{u}}_{\mathcal{T}} \subset \overline{\bigcup_{i>0} \Gamma_i},$$

where $\Gamma_1 = \Gamma \cap \mathcal{U}$, $\Gamma_i = f(\Gamma_{i-1}) \cap \mathcal{U}$.

Of course, an analogous result for flows follows by taking time-1 maps.

An immediate consequence of this is that any finite transition path can be shadowed by a true orbit. The argument for infinite paths requires some elementary point set topology. The following result and its proof are similar to the ones in [DLS00]. They are very inspired by the arguments in [AA67]. The only difference is that here we also control the shadowing along the heteroclinic connections.

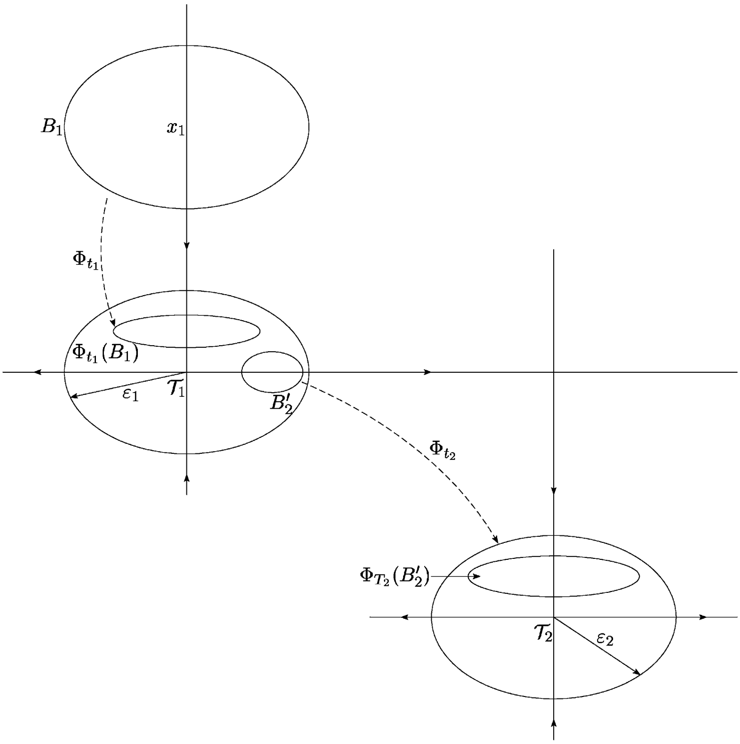

Fig. 2. Fist part of the inductive step in the proof of Lemma 4.37.

Lemma 4.37. *Let* $\{\mathcal{T}_i, \gamma_{i+1}\}_{i=1}^{\infty}$ *be a transition path. Given* $\{\varepsilon_i\}_{i=1}^{\infty}$ *a sequence of strictly positive numbers, we can find a point* P *and an increasing sequence of numbers* T_i, *with* $T_0 = 0$, *such that*

$$\Phi_{T_i}(P) \in N_{\varepsilon_i}(\mathcal{T}_i), \quad \Phi_t(P) \in N_{\varepsilon_i}(\gamma_i), \quad t \in [T_{i-1}, T_i], \ i \geqslant 1,$$

where $N_{\varepsilon_i}(\mathcal{T}_i)$, $N_{\varepsilon_i}(\gamma_i)$, *denote, respectively, neighborhoods of size* ε_i *of the torus* $\mathcal{T}_i$ *and the heteroclinic connection* γ_i.

An illustration of the proof is given in Figs. 2–5.

Proof. We will assume without loss of generality that ε_i are small enough so that $N_{\varepsilon_i}(\mathcal{T}_i) \subset \mathcal{V}_i$ where the Inclination Lemma 4.36 applies.

Let $x_1 \in W^{\mathrm{s}}_{\mathcal{T}_1}$. We can find a closed ball B_1 centered on x_1 such that, for some $t_1 \geqslant 0$,

$$\Phi_{t_1}(B_1) \subset N_{\varepsilon_1}(\mathcal{T}_1). \tag{106}$$

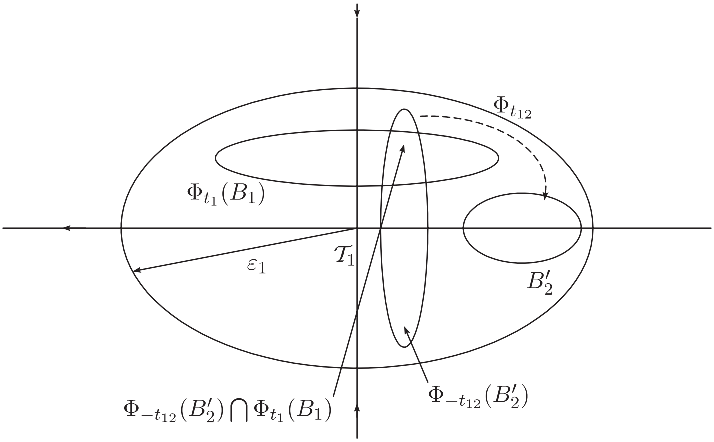

Fig. 3. Second part of the inductive step in the proof of Lemma 4.37.

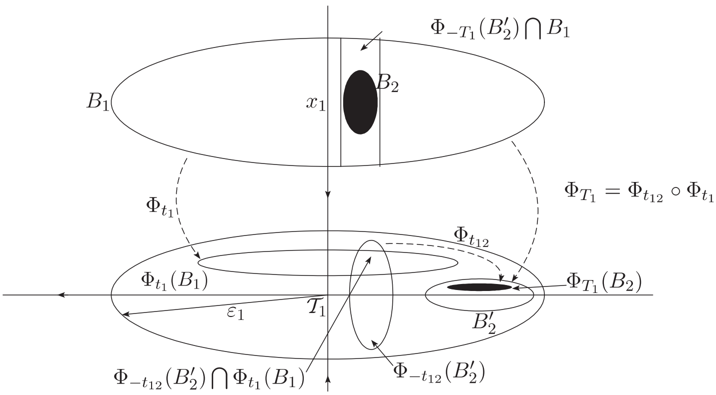

Fig. 4. Third part of the inductive step in the proof of Lemma 4.37.

Let $x_2 \in N_{\varepsilon_1}(\mathcal{T}_1) \cap \gamma_2$. There exists $t_2 > 0$, and a ball $B_2' \subset N_{\varepsilon_1}(\mathcal{T}_1)$ centered at x_2 such that

$$\Phi_{t_2}(B_2') \subset N_{\varepsilon_2}(\mathcal{T}_2), \qquad \Phi_t(B_2') \subset N_{\varepsilon_2}(\gamma_2), \quad t \in [0, t_2].$$

By the Inclination Lemma 4.36 applied inside $N_{\varepsilon_1}(\mathcal{T}_1)$, there exists $t_{12} \geqslant 0$ such that

$$\Phi_{-t_{12}}(B_2') \cap \Phi_{t_1}(B_1) \neq \emptyset, \qquad \Phi_{-t}(B_2') \subset N_{\varepsilon_1}(\mathcal{T}_1), \quad t \in [0, t_{12}].$$

Hence, introducing $T_1 = t_1 + t_{12}$, $T_2 = T_1 + t_2$, we can find a closed ball $B_2 \subset B_1$, centered in a point of $\Phi_{-T_1}(B_2') \cap W^{\mathrm{s}}_{\mathcal{T}_2}$ such that satisfies

$$\Phi_{T_1}(B_2) \subset \Phi_{T_1}(B_1) \subset N_{\varepsilon_1}(\mathcal{T}_1),$$

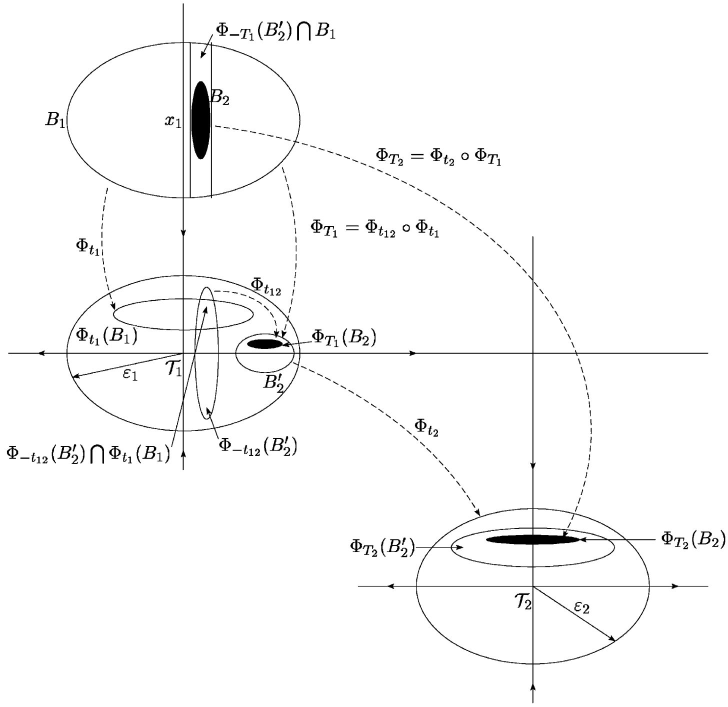

Fig. 5. The whole inductive step in the proof of Lemma 4.37.

as well as

$$\Phi_{T_2}(B_2) \subset N_{\varepsilon_2}(\mathcal{T}_2), \qquad \Phi_t(B_2) \subset N_{\varepsilon_2}(\gamma_2), \quad t \in [T_1, T_2].$$

Proceeding by induction, we can find a sequence of closed balls $B_i \subset B_{i-1} \subset \cdots \subset B_1$ such that

$$\Phi_{T_i}(B_i) \subset N_{\varepsilon_i}(\mathcal{T}_i), \qquad \Phi_t(B_i) \subset N_{\varepsilon_i}(\gamma_i), \quad t \in [T_{i-1}, T_i].$$

Since the balls B_i are compact, $\cap B_i \neq \emptyset$. A point P in the intersection satisfies the required property. $\square$

4.5.5. Proof of Theorem 4.27

The proof of Theorem 4.27 follows applying the results of Lemma 4.37 to the transition chain provided by Lemma 4.35.

Acknowledgements

We thank K. Burns, and specially G. Contreras and G. Paternain, for information about geodesic flows. We also thank M. Sevryuk for suggesting that we include the details for the proof of Theorem 4.27, which was considered well known in [BT99,DLS00]. This work has been supported by the *Comisión Conjunta Hispano Norteamericana de Cooperación Científica y Tecnológica* and MCyT-FEDER Grant BFM2003-9504. Visits of R.L. to Barcelona which were crucial for this work have been supported by a *Cátedra de la Fundación FBBV* and by *Investigador Visitant* of I. C. R. E. A. A.D. and T.S. have also been partially supported by the Catalan Grant 2001SGR-70 and R.L. by NSF Grants.

Appendix A. Proof of Theorem 4.8

In this section we present a proof of Theorem 4.8. The proof is a quite straightforward adaptation of standard techniques and it is, undoubtedly, trivial for the experts. Nevertheless, we could not find a place in the literature where the needed theorem was presented exactly in the way we needed. The main goal for us is to obtain upper estimates of the size of the gaps between the tori which survive. We hope that this appendix will help to make the paper more self-contained. The main theorem of this appendix, Theorem A.15, will be proved for an arbitrary number of degrees of freedom even if for the applications in the paper, we only need the one degree of freedom case.

There are several papers that come very close to proving the result in the form we need it. In particular, we call attention to Theorem 4.1 of [BHTB90] and the discussions in pages 50 and 135 which indeed come very close to the statement of Theorem A.15. Also [Zha00] contains a proof for a KAM theorem for quasi-periodic perturbations of analytic one degree of freedom maps.

We will follow the method started in [Mos66a,Mos66b] and prove a quantitative Theorem A.15 in analytic regularities.

A corresponding result Theorem A.26 for finite regularity will be deduced using a characterization of differentiable functions by approximations with analytic functions.

To use this approximation method it is important that the analytic Theorem A.15 is formulated in the form that existence of an approximate torus with some quantitative properties implies the existence of a true torus and that, moreover, the distance of the approximate torus to the true one can be bounded by the error in the approximation.

We will follow the method of proof and the presentation of [Zeh75,Zeh76a,Zeh76b] but we will change some of the details. In particular, we will pay attention to the dependence of the tori on the diophantine constants. Also, to make the paper more self-contained, rather than invoking a general implicit function theorem, we will present the detailed iterative procedure. This will have the advantage that the assumption that the unperturbed Hamiltonian is analytic can be eliminated.

Of course, a similar result could have been established in many other ways following different schemes of proof. The reader is referred to [dlL01] for a comparative discussion of different proofs of KAM theorem. All the methods discussed in [dlL01] can be adapted to give a proof of Theorem A.15.

The advantage for us of the presentation of [Zeh76a] with respect to other references, is that it is easy to adapt and that it yields estimates which are suitable for our applications—even if not optimal—. We will discuss in more detail the issue of optimal estimates in Remark A.25.

A.1. Diophantine properties

In this section we collect some definitions on Diophantine numbers and their abundance. Most of the definitions and proofs are quite standard except for the Diophantine properties with respect to an external frequency (for which the standard methods also work).

Definition A.1. We introduce the following sets of Diophantine numbers:

$$\mathcal{D}_d(\kappa,\tau) = \left\{ \nu \in \mathbb{R}^d \ : |\nu\cdot k| \geqslant \kappa\,|k|^{-\tau}, \quad k \in \mathbb{Z}^d \setminus \{0\} \right\}. \tag{107}$$

When $\nu \in \mathcal{D}_d(\kappa,\tau)$ and $\tilde\kappa \leqslant \kappa$, $\tilde\tau \geqslant \tau$, we define the numbers which are Diophantine with respect to ν

$$\mathcal{D}_n(\nu;\tilde\kappa,\tilde\tau) = \left\{ \omega \in \mathbb{R}^n : (\omega,\nu) \in \mathcal{D}_{n+d}(\tilde\kappa,\tilde\tau) \right\}. \tag{108}$$

The following is quite well known. We present a rather explicit argument to keep track of the dependence on the constants, which will be important later. The argument is far from optimal in many aspects, specially in the case $n = 1$, but will be sufficient for our applications.

Proposition A.2. *When $\tau > d-1$, the set*

$$\mathcal{D}_d(\tau) := \bigcup_{\kappa>0} \mathcal{D}_d(\kappa,\tau)$$

has full measure in $\mathbb{R}^d$.

Moreover, if B_r is any ball in $\mathbb{R}^d$ of radius r we have

$$|\mathcal{D}_d(\kappa,\tau) \cap B_r| \geqslant |B_r| - C_d \kappa r^{d-1}.$$

In particular, when $r \geqslant \kappa C_d$, we have:

$$\mathcal{D}_d(\kappa,\tau) \cap B_r \neq \emptyset,$$

where C_d is a positive number which depends only on d.

Similarly, given $\nu \in \mathcal{D}_d(\kappa, \tau)$, $\tilde{\tau} > n + d$, $\tilde{\tau} \geqslant \tau$ and $\tilde{\kappa} \leqslant \kappa$ then the set

$$\mathcal{D}_n(\nu; \tilde{\tau}) := \bigcup_{0<\tilde{\kappa}\leqslant\kappa} \mathcal{D}_n(\nu; \tilde{\kappa}, \tilde{\tau})$$

has full measure in $\mathbb{R}^n$.

Moreover, if B_r is any ball in $\mathbb{R}^n$ of radius r we have

$$|\mathcal{D}_n(\nu; \tilde{\kappa}, \tilde{\tau}) \cap B_r| \geqslant |B_r| - C_{n,d}\tilde{\kappa} r^{n-1}$$

In particular, when $r \geqslant \tilde{\kappa} C_{n,d}$

$$\mathcal{D}_n(\nu; \tilde{\kappa}, \tilde{\tau}) \cap B_r \neq \emptyset,$$

where $C_{n,d}$ is a positive number which depends only on n and on d.

Proof. The proof is quite standard. Consider the sets

$$\mathcal{N}_k(d, \kappa, \tau) = \left\{ \nu \in \mathbb{R}^d : |\nu \cdot k| < \kappa |k|^{-\tau} \right\},$$
$$\mathcal{N}_{\ell,k}(n, \nu, \tilde{\kappa}, \tilde{\tau}) = \left\{ \omega \in \mathbb{R}^n : |\omega \cdot \ell + \nu \cdot k| < \tilde{\kappa}\, |(\ell, k)|^{-\tilde{\tau}} \right\},$$

for $0 \neq k \in \mathbb{Z}^d$ and $(0, 0) \neq (\ell, k) \in \mathbb{Z}^n \times \mathbb{Z}^d$, respectively, where one of the inequalities required by the definitions of Diophantine numbers (107), (108) fail.

Note that

$$\mathcal{D}_d(\kappa, \tau) = \mathbb{R}^d \Big\backslash \bigcup_{k\in\mathbb{Z}^d\backslash\{0\}} \mathcal{N}_k(d, \kappa, \tau),$$
$$\mathcal{D}_n(\nu; \tilde{\kappa}, \tilde{\tau}) = \mathbb{R}^n \Big\backslash \bigcup_{(\ell,k)\in(\mathbb{Z}^n\times\mathbb{Z}^d)\backslash\{(0,0)\}} \mathcal{N}_{\ell,k}(n, \nu, \tilde{\kappa}, \tilde{\tau}). \tag{109}$$

Hence, to get lower bounds for the measure of the sets of Diophantine numbers, it suffices to obtain upper bounds for the measures of the sets $\mathcal{N}_k$, $\mathcal{N}_{\ell,k}$ and add them.

Geometrically, the sets $\mathcal{N}_k(d, \kappa, \tau)$, are slabs in $\mathbb{R}^d$, bounded by parallel planes with normals k and at a distance $\kappa\, |k|^{-\tau-1}$. Similarly, when $\ell \neq 0$, the sets $\mathcal{N}_{\ell,k}(n, \nu, \tilde{\kappa}, \tilde{\tau})$ are slabs in $\mathbb{R}^n$ bounded between planes with normals ℓ and at a distance $\tilde{\kappa}\, |(\ell, k)|^{-\tilde{\tau}}\, |\ell|^{-1}$. When $\ell = 0$, the condition defining $\mathcal{N}_{0,k}$ does not depend on ω. We can see that $\mathcal{N}_{0,k}(n, \nu, \tilde{\kappa}, \tilde{\tau})$ is empty provided that $\nu \in \mathcal{D}_d(\kappa, \tau)$ and that $\tilde{\kappa} \leqslant \kappa$, $\tilde{\tau} \geqslant \tau$.

Since the measure of the intersection of a ball of radius r in $\mathbb{R}^d$ with a slab of width w is bounded by $\tilde{C}_d r^{d-1} w$, we can estimate the measure in the excluded sets in (109) to obtain

$$
\begin{aligned}
|B_r \setminus \mathcal{D}_d(\kappa,\tau)| &\leqslant \sum_{k\in\mathbb{Z}^d\setminus\{0\}} |B_r \cap \mathcal{N}_k(d,\kappa,\tau)| \\
&\leqslant \sum_{k\in\mathbb{Z}^d\setminus\{0\}} \tilde{C}_d r^{d-1}\kappa |k|^{-\tau-1} \\
&\leqslant \sum_{m\geqslant 1} \bar{C}_d r^{d-1}\kappa\, m^{-\tau-1+d-1}. \qquad (110)
\end{aligned}
$$

When $\tau > d-1$, from the summation in the right hand side we obtain

$$
|B_r \setminus \mathcal{D}_d(\kappa,\tau)| \leqslant C_d r^{d-1}\kappa. \qquad (111)
$$

When $C_d\kappa \leqslant r$, the right-hand side of (111) is smaller than the volume of the ball of radius r. Hence, the intersection $B_r \cap \mathcal{D}_d(\kappa,\tau)$ has positive measure and, therefore, is non-empty.

When we take the union over $\kappa > 0$ we obtain that $\mathcal{D}_d(\tau)$ is a set of full measure since the excluded set goes to zero with κ.

The argument for the estimates of the measure of $B_r \setminus \mathcal{D}_n(v;\tilde{\kappa},\tilde{\tau})$, where B_r is now a ball in $\mathbb{R}^n$, is very similar to the argument just presented.

Recall that, under the assumptions of the Proposition A.2, we have

$$
\mathcal{N}_{0,k}(n,v,\tilde{\kappa},\tilde{\tau}) = \emptyset.
$$

Using that $\tilde{\tau} > n+d$ we obtain,

$$
\begin{aligned}
|B_r \setminus \mathcal{D}_n(v;\tilde{\kappa},\tilde{\tau})| &\leqslant \sum_{(\ell,k)\in\mathbb{Z}^n\times\mathbb{Z}^d\setminus\{(0,0)\}} \left|B_r \cap \mathcal{N}_{\ell,\kappa}(n,v,\tilde{\kappa},\tilde{\tau})\right| \\
&\leqslant \sum_{(\ell,k)\in(\mathbb{Z}^n\setminus\{0\})\times\mathbb{Z}^d} \tilde{C}_n r^{n-1}\tilde{\kappa}\, |(\ell,k)|^{-\tilde{\tau}}\, |\ell|^{-1} \\
&\leqslant \sum_{(\ell,k)\in(\mathbb{Z}^n\setminus\{0\})\times\mathbb{Z}^d} \tilde{C}_n r^{n-1}\tilde{\kappa}\, |(\ell,k)|^{-\tilde{\tau}} \\
&\leqslant \bar{C}_n r^{n-1}\tilde{\kappa} \sum_{m\geqslant 1} m^{-\tilde{\tau}+n+d-1} \\
&\leqslant C_{n,d} r^{n-1}\tilde{\kappa}. \qquad (112)
\end{aligned}
$$

Hence, when $C_{n,d}\tilde{\kappa} \leqslant r$, we can make the right-hand side of the above inequality smaller than the measure of the ball of radius r in $\mathbb{R}^n$, obtaining that B_r intersects $\mathcal{D}_n(v,\tilde{\kappa},\tilde{\tau})$.

Again, when we take the union over $\tilde{\kappa} > 0$ we obtain that $\mathcal{D}_n(\nu, \tilde{\tau})$ is a set of full measure. □

For the applications of KAM theorems we have in mind, it will be important to investigate the dependence of the sets of Definition A.1 and the constants in Proposition A.2 when we multiply the frequency ν by a number. The following is obvious from the definitions, but we record it.

Proposition A.3. *With the Definitions* A.1 *we have that for any* $\varepsilon > 0$:

$$\nu \in \mathcal{D}_d(\kappa, \tau) \iff \varepsilon\nu \in \mathcal{D}_d(\varepsilon\kappa, \tau).$$

If $\nu \in D_d(\kappa, \tau)$ *and* $\tilde{\kappa} \leqslant \varepsilon\kappa$, *then for any ball* B_r *in* $\mathbb{R}^n$ *of radius* $r \geqslant \tilde{\kappa} C_{n,d}$, *we have for* $\tilde{\tau} \geqslant \tau$, $\tilde{\tau} > n + d$:

$$\mathcal{D}_n(\varepsilon\nu, \tilde{\kappa}, \tilde{\tau}) \cap B_r \neq \emptyset.$$

Moreover,

$$|B_r \setminus \mathcal{D}_n(\varepsilon\nu, \tilde{\kappa}, \tilde{\tau})| \leqslant C_{n,d} r^{n-1} \tilde{\kappa}. \tag{113}$$

A.2. Spaces of functions

The following is an adaptation of the definitions for spaces of analytic functions in [Zeh76a, p.57]. The only change is that we have included the extra variables corresponding to the quasi-periodic motion.

Definition A.4. For $\sigma > 0$, $r > 0$ we denote

$$\begin{aligned} U_\sigma &:= \{x \in \mathbb{C}^{n+d} : |\mathrm{Im}(x_i)| < \sigma, \quad 1 \leqslant i \leqslant n + d\}, \\ U_{\sigma,r} &:= \{(x, y) \in \mathbb{C}^{n+d} \times \mathbb{C}^{n+d} : (x, y) \in U_\sigma, |y| < r - \sigma\}. \end{aligned}$$

We define $\mathcal{H}_\sigma$, $\mathcal{H}_{\sigma,r}$ the Banach spaces of analytic functions on U_σ, $U_{\sigma,r}$ which are bounded, periodic in the x variables and real for real arguments, endowed with the norms

$$\begin{aligned} \|f\|_\sigma &:= \sup_{x \in U_\sigma} |f(x)|, \\ \|g\|_{\sigma,r} &:= \sup_{(x,y) \in U_{\sigma,r}} |g(x, y)|. \end{aligned} \tag{114}$$

In some estimates we will find it convenient to use the notation

$$\|f\|_{\sigma,\mathcal{C}^2} := \sup_{x\in U_\sigma} \max(|f(x)|, |Df(x)|, \left|D^2 f(x)\right|),$$

$$\|g\|_{\sigma,r,\mathcal{C}^2} := \sup_{(x,y)\in U_{\sigma,r}} \max(|g(x,y)|, |Df(x,y)|, \left|D^2 f(x,y)\right|). \tag{115}$$

The method of [Mos66a,Mos66b,Zeh75,Zeh76a] deduces results for finite differentiability out of results for analytic maps. A key ingredient is a characterization of functions with finite regularity by their quantitative approximation properties by analytic functions. The following result goes back to [Mos66a,Mos66b]. A shorter proof appears in [Zeh75,Zeh76a].

Definition A.5. Let $l \in \mathbb{N}$ and $\alpha \in (0,1)$.

We denote by $\mathcal{C}^{l+\alpha}(\mathbb{T}^d \times B_r)$ the space of functions $f : \mathbb{T}^d \times B_r \to \mathbb{R}$ which are l times continuously differentiable and whose derivatives of order l are Hölder with exponent α. The space $\mathcal{C}^{l+\alpha}(\mathbb{T}^d \times B_r)$ is endowed with the norm

$$\|f\|_{\mathcal{C}^{l+\alpha}(\mathbb{T}^d\times B_r)} = \max_{0\leqslant i\leqslant l}\left(\sup_x |D^i f(x)|, \sup_{x\neq y} |D^l f(x) - D^l f(y)| \cdot |x-y|^{-\alpha}\right). \tag{116}$$

Endowed with the norm (116), $\mathcal{C}^{l+\alpha}(\mathbb{T}^d \times B_r)$ is a Banach space.

Proposition A.6. *A function f is in $\mathcal{C}^{l+\alpha}(\mathbb{T}^d \times B_r)$ if and only if there exists a sequence of analytic functions $\{f^n\}$ such that, for some $\eta > 1$, $r > 0$, $\delta > 0$*

$$\begin{aligned} &f^n \in \mathcal{H}_{\delta\eta^{-n},r}, \\ &f^n \overset{\mathcal{C}^0(\mathbb{T}^d\times B_r)}{\longrightarrow} f, \quad \|f^n - f\|_{\mathcal{C}^0(\mathbb{T}^d\times B_r)} \leqslant C\,\delta^{l+\alpha}\eta^{-n(l+\alpha)} \\ &\left\|f^n - f^{n+1}\right\|_{\delta\eta^{-(n+1)},r} \leqslant C\,\delta^{l+\alpha}\eta^{-(n+1)(l+\alpha)} \end{aligned} \tag{117}$$

Moreover

(a) *The best constant C in* (117) *is equivalent to* $\|f\|_{\mathcal{C}^{l+\alpha}}(\mathbb{T}^d \times B_r)$.
(b) *It is possible to choose the approximations in such a way that*

$$||f^0||_{\delta,r} \leqslant C(\delta)\, \|f\|_{\mathcal{C}^{l+\alpha}}(\mathbb{T}^d \times B_r). \tag{118}$$

For details of the proof of Proposition A.6, we refer to [Zeh75]. (Earlier proofs can be found in [Mos66a,Mos66b].) We just indicate that the fact that the limit of a sequence f^n as above is differentiable is a consequence of Cauchy inequalities for functions. Conversely, given a differentiable function, the approximating functions f^n can be obtained by convoluting with an analytic kernel. Since the convolution with the kernel is a linear operator, a more precise formulation of Proposition A.6 can be obtained using the language of smoothing operators.

Remark A.7. It is clear that if (117) holds for some $\delta > 0$, $\eta > 1$, it holds for any other such δ, η. Often, one can find them stated just for $\eta = 2$, $\delta = 1$.

Remark A.8. We note that for the characterization Proposition A.6 to hold, it is important that $\alpha \neq 0, 1$. In those cases, conditions (117) define other spaces of functions than the usual $\mathcal{C}^r$ spaces. The paper [Zeh75] uses the notation $\hat{\mathcal{C}}^r$ to indicate these spaces. In [Ste70], these are called Λ_r spaces.

A.3. Some canonical transformations

We will consider phase spaces $M = \mathbb{T}^n \times \mathbb{T}^d \times U$ where $U \subset \mathbb{R}^n \times \mathbb{R}^d$ is an open set. We will consider M endowed with the exact symplectic form $\Omega = \sum_{i=1}^n dI_i \wedge d\psi_i + \sum_{j=1}^d dA_j \wedge d\theta_j$ and we will use the notation $x = (\psi, \theta)$, $y = (I, A)$.

We will find it convenient to select a particular class of canonical transformations which we will denote by $\mathcal{S}_1$. The following is an adaptation of the definition of [Zeh76a, p. 54] to incorporate an extra variable which is moving quasi-periodically.

Definition A.9. Consider the symplectic manifold $M = \mathbb{T}^n \times \mathbb{T}^d \times U \subset \mathbb{T}^n \times \mathbb{T}^d \times (\mathbb{R}^n \times \mathbb{R}^d)$ and denote by $\mathcal{S}_1$ the set of symplectic diffeomorphisms S on M of the form

$$S(x, y) = \big(a(x), [(Da(x))^\top]^{-1}(-\partial_x b(x) + y)\big), \tag{119}$$

where a is a diffeomorphism of $\mathbb{T}^n \times \mathbb{T}^d$ of the form

$$a(\psi, \theta) = (\psi + \alpha(\psi, \theta), \theta) \tag{120}$$

with $\alpha : \mathbb{T}^n \times \mathbb{T}^d \to \mathbb{R}^n$ and $Da(x)^\top$ means the transpose of $Da(x)$. The function $b : \mathbb{R}^n \times \mathbb{T}^d \to \mathbb{R}$ is of the form

$$b(\psi, \theta) = \lambda \cdot \psi + \beta(\psi, \theta), \tag{121}$$

where $\beta : \mathbb{T}^n \times \mathbb{T}^d \to \mathbb{R}$ and $\lambda \in \mathbb{R}^n$.

The transformations of the form (119) are determined by (λ, α, β). Hence, if there is a possibility of confusion, we will use the notation $S_{\lambda,\alpha,\beta}$.

We note that the functions $S_{\lambda,\alpha,\beta}$ are affine in the action variables. Hence to estimate their analyticity properties it suffices to estimate the analyticity properties of the functions a, b which are functions of the angle variables. Using the notation of analytic spaces introduced before we obtain

$$||S_{\lambda,\alpha,\beta}||_{\sigma,r} \leqslant rA + B,$$
$$||S_{\lambda,\alpha,\beta} - \mathrm{Id}\,||_{\sigma,r} \leqslant rC + D.$$

Explicit expressions of A, B, C, D in terms of can be readily obtained in terms of λ, α, β.

Hence, in our work, we will just need to estimate the λ, α, β and the estimates in the action variables will be automatic. This is an important advantage of the present method compared with the study of canonical transformations based on generating functions or on Lie transforms.

We introduce the notation

$$a(x) = x + \hat{a}(x),$$

where

$$\hat{a}(x) = (\alpha(x), 0)$$

is a periodic function of x. We will also use the notation

$$Da(x)^{-\top} = [Da(x)^{\top}]^{-1}.$$

Remark A.10. In mechanics, one often says that the map S defined in (119) is associated to the generating function

$$G(x, y') = a(x)^{\top} y' + b(x) = (y')^{\top} a(x) + b(x).$$

That is, it is equivalent to say that $S(x, y) = (x', y')$ than to say that x, y, x', y' satisfy

$$x' = \partial_{y'} G(x, y'),$$
$$y = \partial_x G(x, y').$$

The fact that the transformations can be associated to a generating function establishes that the transformations in $\mathcal{S}_1$ are canonical. Of course, this can also be verified by a direct calculation.

Note also that it is possible to think of the transformations in $\mathcal{S}_1$ as families of canonical transformations of the (I, ψ) variables which are indexed by the variables θ.

Remark A.11. We also note that the transformations of $\mathcal{S}_1$ are formally a subgroup of the group of symplectic transformations with respect to the composition.

We will not use this property much, but we note that it has the consequence that the transformation produced in Theorem A.15 will be a transformation of the form (119).

Remark A.12. The transformations in $\mathcal{S}_1$ are always symplectic but they are not exact symplectic if $\lambda \neq 0$ because the constant vector λ is essentially a translation in the I coordinates which makes the transformations not exact if $\lambda \neq 0$.

A.3.1. Transformations close to identity

In the proof, we will pay special attention to the transformations in $\mathcal{S}_1$ which are close to the identity. That is, to the situations where α, β, λ are small.

The distance to the identity can be estimated by

$$\left\| S_{\lambda,\alpha,\beta} - \mathrm{Id} \right\|_{\mathcal{C}^r} \leqslant C \max(|\lambda|, \|\alpha\|_{\mathcal{C}^{r+1}}, \|\beta\|_{\mathcal{C}^{r+1}}).$$

provided that $\max(|\lambda|, \|\alpha\|_{\mathcal{C}^{r+1}}, \|\beta\|_{\mathcal{C}^{r+1}})$ is small. (Note that when α and β are both large, the term $(Da(x))^{-\top}(-\partial_x b(x))$ is quadratic).

Similarly, in analytic spaces we have, for $\sigma > \delta > 0$:

$$\left\| S_{\lambda,\alpha,\beta} - \mathrm{Id} \right\|_{\sigma-\delta,r_0} \leqslant C \max(|\lambda|, \delta^{-1} \|\alpha\|_{\sigma,r_0}, \delta^{-1} \|\beta\|_{\sigma,r_0}) \tag{122}$$

provided $\max(|\lambda|, \delta^{-1} \|\alpha\|_{\sigma,r_0}, \delta^{-1} \|\beta\|_{\sigma,r_0})$ is small. For simplicity, we will use the—obviously wasteful—estimate

$$\max(|\lambda|, \delta^{-1} \|\alpha\|_{\sigma,r_0}, \delta^{-1} \|\beta\|_{\sigma,r_0}) \leqslant \delta^{-1} \max(|\lambda|, \|\alpha\|_{\sigma,r_0}, \|\beta\|_{\sigma,r_0})$$

in (122) to obtain expressions which are easier to handle.

The leading approximation to the identity corresponds to expanding S and ignoring the terms which are quadratic in α, β, λ.

Formally, we have that the infinitesimal approximations are:

$$(S_{\lambda,\alpha,\beta} - \mathrm{Id})(x, y) \approx (\alpha(x), -D\alpha(x)^{\top} y + \partial_x \beta(x) + \lambda).$$

We will denote

$$\hat{S}_{\lambda,\alpha,\beta}(x, y) = (\alpha(x), -D\alpha(x)^{\top} y + \partial_x \beta(x) + \lambda).$$

Note that we have, by Cauchy estimates

$$\left\| \hat{S}_{\lambda,\alpha,\beta} \right\|_{\sigma-\delta,r} \leqslant C\delta^{-1} \max(|\lambda|, \|\alpha\|_{\sigma}, \|\beta\|_{\sigma}).$$

A more precise formulation of the calculations on infinitesimal transformation is:

Proposition A.13. *With the notations above and assuming that*

$$\delta^{-1} \max(|\lambda| \,, \|\alpha\|_\sigma \,, \|\beta\|_\sigma)$$

is sufficiently small, we have

$$\left\| S_{\lambda,\alpha,\beta} - \mathrm{Id} - \hat{S}_{\lambda,\alpha,\beta} \right\|_{\sigma-\delta,r} \leqslant C\delta^{-2} \max(|\lambda| \,, \|\alpha\|_\sigma \,, \|\beta\|_\sigma)^2. \tag{123}$$

Similarly, assuming that the composition $H \circ S$ makes sense we have

$$\left\| H \circ S_{\lambda,\alpha,\beta} - H - DH\hat{S}_{\lambda,\alpha,\beta} \right\|_{\sigma-\delta,r-\delta} \leqslant C\delta^{-2} \max(|\lambda| \,, \|\alpha\|_\sigma \,, \|\beta\|_\sigma)^2, \tag{124}$$

where C depends only on $\|H\|_{\sigma,r,\mathcal{C}^2}$.

Proof. We have, using the Neumann series that

$$\begin{aligned} \left\| Da^{-\top} - \mathrm{Id} - D\alpha^\top \right\|_{\sigma-\delta,r} &\leqslant C \left\| D\alpha^{-\top} \right\|^2_{\sigma-\delta,r} \\ &\leqslant C\delta^{-2} \|\alpha\|^2_{\sigma,r} . \end{aligned}$$

Eq. (123) follows immediately using the Banach algebra properties adding and subtracting terms.

The proof of (124) is to observe that by Taylor's theorem we have

$$|H \circ S(x, y) - H(x, y) - DH(x, y)\hat{S}(x, y)| = C \sup_{\tilde{x},\tilde{y}} |D^2 H(\tilde{x}, \tilde{y})| \cdot |\hat{S}(x, y)|^2,$$

where $(\tilde{x}, \tilde{y})$ is a point intermediate between x, y, $S(x, y)$.

Then, we can apply the Banach algebras property, Cauchy estimates and the previous estimates. □

Remark A.14. It will be important to note that if H has the form (55) and S has the form (119), then $H \circ S$ has the form (55). That is, when we apply quasi-periodic transformations of frequency v to a quasi-periodic Hamiltonian of the same frequency v, we obtain a quasi-periodic Hamiltonian of the same frequency. Hence, when we consider quasi-periodic Hamiltonians as in (55) we can use the transformations in $\mathcal{S}_1$ in the same way that [Zeh75] uses the transformations which are denoted there as $\mathcal{S}_1$.

A.4. Proof of an analytic version of the KAM theorem

In this section, we establish a quantitative theorem for analytic invariant tori. Then, in Section A.5, we will deduce the result for finitely differentiable perturbations from the analytic one by using the characterization of finitely differentiable functions by their approximation properties by analytic ones. We will follow [Zeh75,Zeh76a,Zeh76b] rather closely but rather than reducing the proof to an implicit function theorem we will present details on the iterative scheme.

The quantitative version of the theorem for analytic regularity we will establish is:

Theorem A.15. *Let* $v \in \mathcal{D}_d(\kappa, \tau)$, $\omega \in \mathcal{D}_n(v, \tilde{\kappa}, \tilde{\tau})$, *for some* $0 < \tilde{\kappa} \leqslant \kappa$, $\tilde{\tau} \geqslant \tau > 0$. Denote $\gamma = 2\tilde{\tau} + 1$.

Let $H_0 \in \mathcal{H}_{\sigma_0,1}$, *for some* $\sigma_0 > 0$, *be of the form*

$$H_0(\psi, \theta, I, A) = c_0 + v \cdot A + \omega \cdot I + \tfrac{1}{2} I^\top Q_0(\psi, \theta) I + R_0(\psi, \theta, I), \tag{125}$$

where $R_0(\psi, \theta, I) = O(I^3)$.

Denote

$$< Q_0 >= \int_{\mathbb{T}^{n+d}} Q_0(\psi, \theta)\, d\psi\, d\theta,$$

and assume that

$$e := \|H_0\|_{\sigma_0, 1, \mathcal{C}^2} < \infty,$$

$$\rho := \left\| < Q_0 >^{-1} \right\| < \infty.$$

Then, there are constants $C, \sigma^* > 0$, *depending on* $e, \tau, \tilde{\tau}, \rho$*—but not on* $\tilde{\kappa}$*—such that, for all* $\sigma \leqslant \sigma^*$, *given any function* $H \in \mathcal{H}_{\sigma,1}$ *of the form*

$$H(\psi, \theta, I, A) = v \cdot A + E(\psi, \theta, I) \tag{126}$$

satisfying

$$\|H - H_0\|_{\sigma,1} \leqslant C \tilde{\kappa}^4 \sigma^{2\gamma+3} \rho^{-2}, \tag{127}$$

then, there is a symplectic diffeomorphism $F \in \mathcal{S}_1 \cap \mathcal{H}_{\sigma/4,\infty}$ *such that*
(A) *We have*

$$H \circ F(\psi, \theta, I, A) = c + v \cdot A + \omega \cdot I + \tfrac{1}{2} I^\top Q(\psi, \theta) I + R(\psi, \theta, I), \tag{128}$$

where $R = O(I^3)$.

(B) *$F = F_{\lambda,\alpha,\beta}$ is close to the identity*

$$\max\{|\lambda|\,, \|\alpha\|_{\sigma/4}\,, \|\beta\|_{\sigma/4}\} \leqslant C\rho\tilde{\kappa}^{-2}\sigma^{-\gamma}\,\|H - H_0\|_{\sigma,1}\,. \tag{129}$$

Hence, in particular,

$$\|F - \mathrm{Id}\|_{\sigma/6,1,\mathcal{C}^1} \leqslant C\rho\tilde{\kappa}^{-2}\sigma^{-\gamma-1}\,\|H - H_0\|_{\sigma,1}\,. \tag{130}$$

(C)
$$\|Q - Q_0\|_{\sigma/4} \leqslant C\rho\tilde{\kappa}^{-2}\sigma^{-\gamma}\,\|H - H_0\|_{\sigma,1}\,. \tag{131}$$

Remark A.16. Theorem A.15 is the same as Theorem 2.1 of [Zeh76a], except for a few differences. We note:

(1) We are assuming a special quasi-periodic form for the Hamiltonian and for the perturbation (see (126)) and we are obtaining that the transformation F has a special form. (It is the identity in the θ variable.)
   As we will see later, this does not affect much the proof. The proof is based in an iterative procedure and we just have to check that the transformations at each step can be taken as identities in the θ variable and that the Hamiltonians at each step have the form (126).
(2) We are making explicit the dependence on conclusions (129)–(131) on the constant $\tilde{\kappa}$ of the Diophantine properties of (ω, v). This explicit dependence, namely $\tilde{\kappa}^{-2}$, is included in the main lemma of [Zeh76b].
   Note, in particular, that it suffices that the initial error (127) is smaller than $\tilde{\kappa}^4$. Then, we obtain in (130) that the distance of the correction to the identity can be bounded by the initial error multiplied by $\tilde{\kappa}^{-2}$.
(3) We are also making explicit the dependence of the constants on the twist parameter ρ. This improvement is also included in [Zeh76b] although it will not be used in this paper.

Making explicit the dependence on $\tilde{\kappa}$ is important for our purposes since it makes it possible to discuss the proximity of the tori that survive. The dependence of the result on $\tilde{\kappa}$ is also crucial to obtain lower bounds on the the measure occupied by the tori. In this paper, we will not consider measure theoretic properties of the tori produced, since for us it is enough to bound from above the size of the gaps.

Remark A.17. Following [Zeh76a], we have assumed that the analyticity domain of the unperturbed hamiltonian in the action variables is 1.

This is an assumption that can always be arranged. If we make the change of variables for $0 < \lambda < 1$:

$$\begin{aligned}
&\tilde{I} = \lambda^{-1} I, \tilde{A} = \lambda^{-1} A,\\
&\tilde{\psi} = \psi, \tilde{\theta} = \theta,\\
&\tilde{t} = \lambda^{-1} t,
\end{aligned}$$

the resulting equations are also Hamiltonian, and the new Hamiltonian is given by

$$\tilde{H}_0(\tilde{\psi}, \tilde{\theta}, \tilde{I}, \tilde{A}) = \lambda^{-1} H_0(\tilde{\psi}, \tilde{\theta}, \lambda\tilde{I}, \lambda\tilde{A})$$

but then the domain of analyticity of the hamiltonian in the $\tilde{I}, \tilde{A}$ variables is λ^{-1} times larger than in the original variables.

Note also that $\tilde{H}_0$ is also of the same form assumed in (125). The frequencies are unchanged, the twist constant changes according to $\tilde{\rho} = \lambda^{-1}\rho$. We also note that $||\tilde{H}_0||_{\sigma,1,\mathcal{C}^2} = ||H_0||_{\sigma,\lambda,\mathcal{C}^2}$ so that provided that we have inequalities (127) for the scaled quantities, we can apply the result.

Remark A.18. Note that the symplectic diffeomorphism $F \in \mathcal{S}_1$ claimed in Theorem A.15 solving Eq. (128) is non-unique. If we compose F on the left with an arbitrary rotation in the angle variable ψ, it will also be a solution of (128).

The estimates claimed in the rest of the theorem are not true for all the solutions of (128) but only for the one produced through the procedure described in the proof.

Remark A.19. Theorem A.15 implies in particular the persistence of quasi-periodic solutions in systems that are close to integrable. It suffices to use as approximate solutions the quasiperiodic solutions of the integrable system.

Remark A.20. In spite of the non-uniqueness for F observed in Remark A.18, under the conditions we have assumed (notably $< Q_0 >$ invertible) the torus with the assigned frequency is locally unique in a neighborhood of $\mathbb{T}^{n+d} \times \{0\}$.

This follows as in [Zeh75,Zeh76a] because we will show that, if we impose an extra normalization, the Newton method admits a left inverse. That is, the only non-uniqueness of the solution of (58) is that described in Remark A.18.

Remark A.21. The uniqueness of the torus with respect to the frequency, allows to define a mapping $\mathcal{T}$ that to a number $\omega \in \mathcal{D}_n(v, \tilde{\kappa}, \tilde{\tau})$ associates the torus $\mathcal{T}_\omega$ with frequency (ω, v).

The space of tori can considered as a Banach manifold modeled on a space of analytic functions. (If a torus corresponds to $\{(\psi, \theta, I) : I = 0\}$, then the close enough tori can be written as the graph of a function $I = T(\psi, \theta)$, where $T : \mathbb{T}^{n+d} \to \mathbb{R}^n$.)

The mapping $\mathcal{T}$ is Whitney differentiable. This is well known for regular KAM theorems from [CG82,Pos82,Sva80]. It is also a consequence of the proof based on abstract implicit function theorems. This is shown in [LV01].

We note Theorems A.15 and A.26 imply immediately that the dependence is Lipschitz since we can take a torus as an approximate solution for a torus with a slightly different frequency. The error is proportional to the difference of frequencies.

Remark A.22. The results of Theorem A.15 can be improved significantly when v is one dimensional—i.e. periodic perturbations—. This case is treated in [Zeh76a], where it is shown that this case implies analogous results for maps.

As shown in [Zeh76a], in the case of periodic perturbations, it is possible to assume much less regularity than what follows from applying Theorem A.15.

These improvements are based on the observation—which we have also used in Section 4.3.5 and especially in Theorem 4.7—that the cohomology equations in one variable can be solved by integrating and one does not need to deal neither with Fourier series nor small divisors.

Proof. We follow the strategy of proof in [Zeh76a] but carry out the estimates keeping explicit the dependence on the constants $\tilde{\kappa}$, ρ which appear in the hypotheses of Theorem A.15. Similar estimates are done in the proof of the main Lemma of [Zeh76b, p.108].

We recall that the method of proof in [Zeh76a] is to develop an iterative procedure. A step of the iterative procedure consists in applying a canonical transformation in $\mathcal{S}_1$ that eliminates the leading part of the terms in the Hamiltonian that prevent it from being in the desired normal form (128).

As it is well known the choice of the convenient transformations involves solving small divisors equations and, hence, at each step of the iterative procedure we will obtain control only in a slightly smaller domain. This loss of domain can be controlled if the error is small because of the quadratic convergence of the method.

The key observation in adapting the proof of [Zeh76a] to our case is that the changes of variables used to reduce the perturbation can be taken to be in $\mathcal{S}_1$ when the perturbations are of the quasiperiodic form (126).

Since the changes of variables in $\mathcal{S}_1$ preserve the form (126), we see that the iterative procedure in [Zeh76a] can be carried out using only Hamiltonians of the form (126) and transformations in $\mathcal{S}_1$.

We will also need to check that the transformations at each stage are derived through equations which are analogue to the infinitesimal equations (2.27)–(2.29) [Zeh76a, p. 63].

We will check that if we add to the objects in [Zeh76a] the extra variables θ and impose quasi-periodic dependence, we obtain quasi-periodic dependence on the solutions. Moreover, we will see that the estimates obtained in the iterative step are very similar to those in [Zeh76a] and in [Zeh76b]. After that, the changes from the above papers will be quite minimal. The estimates that establish convergence and quantitative properties of the solutions will not require any changes from those in the above papers. We give details.

The proof will consist of an iterative procedure which produces transformations that reduce progressively the error. At any step of the iterative procedure, we consider a Hamiltonian of the form

$$\begin{aligned} H(I,\psi,A,\theta) = c + v\cdot A + \omega\cdot I + E_0(\psi,\theta) + E_1(\psi,\theta)\cdot I \\ + \tfrac{1}{2} I^\top Q(\psi,\theta) I + R(I,\psi,\theta) \end{aligned} \tag{132}$$

with $R = O(I^3)$.

We note that the scalar function $c + E_0$ and the vector function E_1 are the Taylor expansions in the variables I of order 0 and 1 respectively of functions as in (55). The c is chosen so that the average of E_0 vanishes. (Note that, since Hamiltonians are defined up to additive constants, we could also have just ignored the constant terms.)

We want to find a transformation S in $\mathcal{S}_1$ as in Definition A.9 in such a way that the *errors* $\tilde{E}_0$, $\tilde{E}_1$ of the function $\tilde{H} \equiv H \circ F$ are *quadratic*: As usual in KAM theory, this means estimated in a smaller domain by the square of the original errors times a function of the domain loss which does not grow too fast. For example, a negative power. That is, we will find the (λ, α, β) which, according to Definition A.9, determine the transformation $S = S_{\lambda,\alpha,\beta}$. We will also need to determine $\tilde{Q}$, the new quadratic term in Hamiltonian (55). $\square$

A.4.1. Estimates for a step of the iterative procedure

As in the proof of Lemma 2.1 of [Zeh76a], taking into account that $S \simeq \mathrm{Id} + \hat{S}$ (see (123)) we compute $H \circ S \approx H + DH\hat{S}$, keeping only the terms which are linear in $E_0, E_1, \alpha, \beta, \lambda$, and imposing that the new terms of order $0, 1$ in I vanish. It will be possible to obtain estimates for the new error terms by using estimates for the λ, α, β in terms of E_0, E_1 and using (124).

By requiring, as indicated above, that the leading terms in the error vanish, we derive the following equations for λ, α, β, which are complete analogues of equations (2.27), (2.28), of [Zeh76a].

$$\begin{aligned} &\partial\beta + \omega \cdot \lambda + c - \tilde{c} = -E_0, \\ &-\partial\alpha + Q(\lambda + \partial_\psi \beta) = -E_1, \end{aligned} \tag{133}$$

where ∂ is the differential operator acting on periodic functions given by

$$\partial := \omega \cdot \partial_\psi + \nu \cdot \partial_\theta. \tag{134}$$

The above equations (133) are the well known small divisor equations that appear in KAM theory. Very sharp estimates for their solutions appear in [Rus75] (see also [Rus76b,Rus76a]). We summarize them now.

Lemma A.23. *Assume that $\omega \in \mathcal{D}_n(\nu, \tilde{\kappa}, \tilde{\tau})$.*

Let η be an analytic function on $\mathbb{T}^{n+d}$ with zero average.

Then, we can find a unique φ solving

$$\partial\varphi = \eta$$

and having zero average.

Moreover, if $\eta \in \mathcal{H}_\sigma$, we have for all $0 < \sigma' < \sigma$:

$$\|\varphi\|_{\sigma'} \leqslant C\tilde{\kappa}(\sigma - \sigma')^{-\tilde{\tau}} \|\eta\|_\sigma, \tag{135}$$

where C is a constant which depends only on $\tilde{\tau}$, $n + d$.

Explicit expressions for C in (135) appear in [Rus75]. We also note that if we do not impose that the solution φ has zero average, we obtain that the solution of (135) is unique up to an additive constant.

Proofs of (135) with a factor $(\sigma-\sigma')^{-(\tilde{\tau}+n+d)}$ instead of $(\sigma-\sigma')^{-\tilde{\tau}}$ can be obtained rather elementarily by noting that the Fourier coefficients of the solution are given by $\hat{\varphi}_k = (2\pi i k\tilde{\omega})^{-1}\hat{\eta}_k$, where $\tilde{\omega} = (\omega, v)$. It suffices to use Cauchy estimates for the Fourier coefficients and the inequalities in the definition of Diophantine properties. The proof in [Rus75] obtains sharper results by observing that the bounds in the definition of Diophantine properties cannot be saturated very often.

Coming back to the analysis of the iterative step in the proof of KAM theorem, we note that system (133) has an upper triangular structure which allows to reduce its analysis to an application of Lemma A.23.

We first solve $\partial\beta = -E_0$. We recall that E_0 was assumed to have zero average. Hence, applying Lemma A.23 we obtain β with zero average.

To solve the second equation of (133), we choose first λ such that

$$< Q > \lambda = - < E_1 > - < Q\partial_\psi\beta > .$$

With this choice of λ, the function $Q(\lambda + \partial_\psi\beta) + E_1$ has zero average, and we can apply Lemma A.23 to obtain α.

Finally, we take $\tilde{c}$ such that

$$c - \tilde{c} + \omega \cdot \lambda = 0$$

to complete the solution of the first equation of (133). (Of course, this step could be omitted by noting that Hamiltonians can be defined up to additive constants.)

Hence, we obtain that it is possible to solve Eqs. (133) and that the solutions obtained satisfy the estimates:

En la siguiente formula habia un exponente mas esto afecta al exponente γ despues.

$$|\lambda|\,,\, \|\alpha\|_{\sigma'}\,,\, \|\beta\|_{\sigma'} \leqslant C\rho\tilde{\kappa}^{-2}(\sigma-\sigma')^{-2\tilde{\tau}-1}(||E_0||_\sigma + ||E_1||_\sigma), \tag{136}$$

From the estimates on α, β, λ we have using Proposition A.13

$$\|S - \mathrm{Id}\|_{\sigma',1} \leqslant C\rho\tilde{\kappa}^{-2}(\sigma-\sigma')^{-\gamma-1}(||E_0||_\sigma + ||E_1||_\sigma), \tag{137}$$

where $\gamma = 2\tilde{\tau} + 1$.

Eq. (2.29) of [Zeh76a], which determines the new Q, is simply an algebraic equation which does not require any change. We present the details.

We just note that we can bound

$$\begin{aligned}\left\| < Q > - < \tilde{Q} > \right\| &\leqslant C \max(|\lambda| \, , \|\alpha\|_{\sigma'} \, , \|\beta\|_{\sigma'}) \\ &\leqslant C\rho\tilde{\kappa}^{-2}(\sigma-\sigma')^{-\gamma}(||E_0||_\sigma + ||E_1||_\sigma). \qquad (138)\end{aligned}$$

In particular, if $< Q >$ is invertible and the perturbation is small enough, then $< \tilde{Q} >$ will be invertible and we can bound

$$\begin{aligned}|\rho - \tilde{\rho}| &\leqslant C \max(|\lambda| \, |, |\alpha||_{\sigma'}, ||\beta||_{\sigma'}) \\ &\leqslant C\rho\tilde{\kappa}^{-2}(\sigma-\sigma')^{-\gamma}(||E_0||_\sigma + ||E_1||_\sigma). \qquad (139)\end{aligned}$$

Note that this are the same estimates as in [Zeh76a] except that we have made explicit the dependence on the small divisor conditions and on ρ.

Provided that $\delta = \sigma - \sigma'$ is larger than the RHS of (137) we can consider $H \circ S$ defined in the domain $U_{\sigma',1-\delta}$. This sufficient condition will ensure that we can define the iterative step. More explicitly, using the estimates that we have for $\|S - \mathrm{Id}\|$, we have that a sufficient condition for the possibility of carrying out the iterative step is

$$\sigma - \sigma' > C\rho\tilde{\kappa}^{-2}(\sigma-\sigma')^{-\gamma-1}(||E_0||_\sigma + ||E_1||_\sigma). \qquad (140)$$

In the rest of the proof we will check that

(1) Condition (140) can be verified for successive steps.
    Note that condition (140) is verified provided that the error terms are much smaller than the loss of domain. Hence, the inductive assumption (140) will be verified when the errors decrease much faster than the loss of domain.
(2) The conditions on ρ and on the second derivative of the Hamiltonian do not deteriorate and, for example, can be assumed to be bounded by twice their original values.
    We note that the change of the Hamiltonian can be controlled by the difference of the transformation to the identity, which in turn can be bounded by the size of the errors. Again, we note that if we obtain good bounds on the size of decrease of the errors, then, the deterioration of these parameters during the proof is small.
(3) The transformations converge in a smaller domain.

So, the crux of the proof is to estimate that the error decreases fast enough provided that the constants ρ and $\|H^n\|_{\sigma_n,1-\delta_n}$ are twice the original values.

We note that the α, β, λ are chosen so that, in the linear approximation, the new error terms are exactly zero. Hence, the new error terms of $H \circ S$ can be estimated by the error of the linear approximation as in Proposition A.13. More precisely, we note that the new error terms, $\tilde{E}_0$, $\tilde{E}_1$, are obtained by evaluating respectively $H \circ S$

and $\partial_I H \circ S$ at $I = 0$. We note that, by the construction of S, $H - DH\hat{S}|_{I=0} = 0$, $\partial_I H - DH\hat{S}|_{I=0} = 0$. Hence

$$
\begin{aligned}
H \circ S|_{I=0} &= (H \circ S - H - DH\hat{S})|_{I=0},\\
\partial_I H \circ S|_{I=0} &= \partial_I (H \circ S - H - DH\hat{S})|_{I=0}
\end{aligned}
$$

Applying Proposition A.13 and Cauchy bounds in the last equation we obtain

$$
\begin{aligned}
\left\|\tilde{E}_0\right\|_{\sigma'} &\leqslant C(\sigma-\sigma')^{-2}\max(|\lambda|, \|\alpha\|_\sigma, \|\beta\|_\sigma)^2\\
&\leqslant C\rho^2\tilde{\kappa}^{-4}(\sigma-\sigma')^{-2\gamma-2}(||E_0||_\sigma + ||E_1||_\sigma)^2,
\end{aligned}
$$

$$
\begin{aligned}
\left\|\tilde{E}_1\right\|_{\sigma'} &\leqslant C(\sigma-\sigma')^{-3}\max(|\lambda|, \|\alpha\|_\sigma, \|\beta\|_\sigma)^2,\\
&\leqslant C\rho^2\tilde{\kappa}^{-4}(\sigma-\sigma')^{-2\gamma-3}(||E_0||_\sigma + ||E_1||_\sigma)^2. \qquad (141)
\end{aligned}
$$

Hence, we have established that, after applying one step, the error $\varepsilon = ||E_0||_\sigma + ||E_1||_\sigma$ can be bounded by a quadratic estimate in a slightly smaller domain. Namely

$$
\tilde{\varepsilon} \leqslant C\rho^2\tilde{\kappa}^{-4}(\sigma-\sigma')^{-q}\varepsilon^2, \qquad (142)
$$

where $q = 2\gamma + 3$, and $\tilde{\varepsilon} = ||\tilde{E}_0||_{\sigma'} + ||\tilde{E}_1||_{\sigma'}$.

We also note that we have bounds (137) on how close to the identity is the transformation S that reduces the error.

Finally, we now show that estimates on the second derivative of the new Hamiltonian are not much worse than that of the original one (provided that the original error is small enough).

Proposition A.24. *Provided that*

$$
\begin{aligned}
&\|S - \mathrm{Id}\|_{\sigma-\delta,\eta-\delta} \leqslant K\delta^{-\gamma-1}\varepsilon < \delta,\\
&\eta - \delta > 0, \qquad (143)
\end{aligned}
$$

we have

$$
S(U_{\sigma-\delta,\eta-\delta}) \subset U_{\sigma,\eta} \qquad (144)
$$

and

$$\begin{aligned}||H\circ S||_{\sigma-\delta,\eta-\delta,\mathcal{C}^2} &\leqslant ||H||_{\sigma,\eta,\mathcal{C}^2}(1+K\delta^{-2}||\hat{S}||_{\sigma-(7/8)\delta,\eta-(7/8)\delta})\\ &\leqslant ||H||_{\sigma,\eta,\mathcal{C}^2}(1+K\delta^{-\gamma-3}\varepsilon).\end{aligned}$$

Proof. Since we have that real values are mapped into real values, we have (144).

To estimate the derivatives, we note that

$$\begin{aligned}D^2(H\circ S) &= D^2H\circ S DS^{\otimes 2} + DH\circ S D^2 S\\ &= D^2H\circ S(\mathrm{Id}+D\hat{S})^{\otimes 2} + DH\circ S D^2 S.\end{aligned}$$

Using (144), we can bound the supremum in the indicated domain of the second derivative.

The increase of the bounds for the first derivative are much easier and are left for the reader. □

A.4.2. Repeating the iterative step and convergence

After the work carried out so far, we could apply an implicit function theorem or the Main Lemma of [Zeh76b] to obtain the desired result, Theorem A.15. Nevertheless, for the sake or completeness, we present the details, which are not too complicated.

The argument is to show that, provided that ε_0 is small enough, and that we can perform n iterations, then, the error has reduced so much that we can perform the next iteration.

Assuming that we can perform n iterative steps—which amounts to the fact that we have (140) in all the steps—and that all the ρ that appear are bounded by $2\rho_0$, we obtain

$$\begin{aligned}\varepsilon_{n+1} &\leqslant (C\tilde{\kappa}^{-4}\rho_0^2\sigma_0^{-q})^{1+2+2^2+\cdots+2^n}2^{q[n+2(n-1)+2^2(n-2)+\cdots+2^n]}(\varepsilon_0)^{2^{n+1}}\\ &\leqslant (C\sigma_0^{-q}\rho_0^2\tilde{\kappa}^{-4}2^{2q}\varepsilon_0)^{2^{n+1}} \qquad (145)\end{aligned}$$

which we can see converges to zero extremely fast if the term in parenthesis is strictly smaller than 1. We have used the elementary bounds

$$2^{q[n+2(n-1)+2^2(n-2)+\cdots+2^n]}\leqslant 2^{q2^{n+1}[n2^{-n+1}+(n-1}2^{-n+2}\cdots 2^{-1}]\leqslant 2^{2q2^{n+1}}$$

In summary, if

$$C\rho_0^2 2^{2q}\sigma_0^{-q}\tilde{\kappa}^{-4}\varepsilon_0 < 1, \qquad (146)$$

and we can perform the iteration n times, we have bounds (145). These bounds will verify the inductive bounds (140)

Now, the only thing that remains to do, to verify that we can perform a next iteration, is that the estimates in ρ and on $\|H\|_{\sigma',\eta',\mathcal{C}^2}$ are not much worse than the initial ones.

We observe also that by (139) we have

$$\rho_n \leqslant \rho_0 + \sum_{j}^{n-1} C\rho_0^2\tilde{\kappa}^{-2}\sigma_0^{-\gamma}2^{j\gamma)}\varepsilon_j$$

and, by Proposition (A.24)

$$\begin{aligned}||H \circ F^n||_{\sigma_n,\eta_n,\mathcal{C}^2} &\leqslant ||H||_{\sigma_0,\eta_0,\mathcal{C}^2} \prod_{i=0}^{n-1}(1 + K\sigma_0^{-\gamma-3}2^{i(\gamma+3)}\varepsilon_i)\\ &\leqslant ||H||_{\sigma_0,\eta_0,\mathcal{C}^2} \prod_{i=0}^{n-1}(1 + K\sigma_0^{-\gamma-3}2^{i(\gamma+3)}(C\sigma_0^{-q}\rho_0^2\tilde{\kappa}^{-4}2^{2q}\varepsilon_0)^{2^i}).\end{aligned}$$

We will, therefore assume that ε_0 is small enough, so that $C2^{2q}\rho_0^2\sigma_0^{-q}\tilde{\kappa}^{-4}\varepsilon_0 < 1/20$. The bounds imply that

$$||H \circ S||_{\sigma_n,\eta_n,\mathcal{C}^2} \leqslant 2||H \circ S||_{\sigma_0,\eta_0,\mathcal{C}^2}.$$

These bounds and (145) imply that conditions (140) are satisfied and we can take one step more.

Furthermore, since the distance of the changes of variables S_n to the identity is controlled by ε_n, we can estimate the changes of variables and obtain the result claimed in Theorem A.15. We leave these details to the reader or refer to [Zeh76a]. The main observation is that since the changes at the n step are converging to the identity very fast, the distance of the total change to the identity is comparable to that of the first step.

Remark A.25. The dependence on the Diophantine constants $\tilde{\kappa}$, $\tilde{\tau}$ and the loss of derivatives in Theorem A.15 are not optimal.

One can prove the same result as Theorem A.15 for $\|H - H_0\| \leqslant C\tilde{\kappa}^a\sigma_0^{2\gamma+2}$, with $a = 2$, $\gamma = \tilde{\tau} + 1$ instead of $a = 4$ and $\gamma = 2\tilde{\tau} + 1$.

These improved estimates, to the best of our knowledge, require longer proofs and more severe adaptations. Since the version presented here is enough for the purposes of the present paper, we decided to present only the proofs that seemed to us simpler to adapt.

Nevertheless, for the sake of completeness, we just indicate some possibilities that lead to better results.

Ref. [dlL04] contains better results on the loss of derivatives by using weighted norms which are adapted to the upper triangular equations (133).

There is an alternative proof of KAM theorem based on a strategy started in [Kol54] and implemented fully in [Arn63]. The optimal value of a for the standard KAM theorem for analytic Hamiltonians is obtained in [Nei81]. The fact that $a = 2$ cannot be improved is easy to verify in examples.

Some modern proofs of the standard KAM theorem which lead to optimal estimates on the measure and on the loss of derivatives are [Pos82,Pos01,Sal86] It seems that any of the proofs mentioned above can be adapted to the situation presented here and obtain the results with better exponents in the loss of derivatives and in the Diophantine constant.

We refer to [dlL01] for a comparison between different proofs of KAM theorem.

A.5. Proof of a differentiable version of the KAM theorem

In this section, we use Theorem A.15, the analytic version of the KAM theorem, to prove Theorem A.26 below, which is a differentiable version of the KAM theorem.

Following [Mos66a,Mos66b,Zeh76a], the proof consists in applying Theorem A.15 to a sequence of approximations given by Proposition A.6. The transformations required will converge fast. Hence, by an application of Proposition A.6 (indeed, just Cauchy estimates), we will conclude that there is a finitely differentiable transformation F that casts the Hamiltonian H into (128).

Given the work that we have already done, the result follows from an invocation to the abstract theorem in [Zeh75] as modified in [Zeh76b]. (The main difference between [Zeh76b] and [Zeh75] is that [Zeh76b] keeps track of the dependence of the smallness conditions on the Diophantine constants. This is crucial for our quantitative arguments.) We will present the details of the proof—following the steps suggested by the abstract theorem in our concrete case—to make this paper more self-contained.

Theorem A.26. *Let $\nu \in \mathcal{D}_d(\kappa, \tau)$, $\omega \in \mathcal{D}_n(\nu, \tilde{\kappa}, \tilde{\tau})$, for some $0 < \tilde{\kappa} \leqslant \kappa$, $\tilde{\tau} \geqslant \tau > 0$. Denote $\gamma = 2\tilde{\tau} + 1$.*

Let $H_0 \in \mathcal{H}_{\sigma_0,1}$, for some $\sigma_0 > 0$, be of the form

$$H_0(\psi, \theta, I, A) = c_0 + \nu \cdot A + \omega \cdot I + \tfrac{1}{2} I^\top Q_0(\psi, \theta) I + R_0(\psi, \theta, I), \tag{147}$$

where $R_0(\psi, \theta, I) = O(I^3)$.

Denote

$$< Q_0 >= \int_{\mathbb{T}^{n+d}} Q_0(\psi, \theta)\, d\psi\, d\theta,$$

and assume that

$$e := \|H_0\|_{\sigma_0,1,\mathcal{C}^2} < \infty,$$

$$\rho := \left\| < Q_0 >^{-1} \right\| < \infty.$$

Then, there is a constant $C_0 > 0$, *depending on* $e, \tau, \tilde{\tau}, \rho$*—but not on* $\tilde{\kappa}$*—such that given any function* $H \in \mathcal{C}^l(\mathbb{T}^{n+d} \times B_1)$, $l \geqslant 2\gamma + 3 = 4\tilde{\tau} + 5$, $l \notin \mathbb{N}$, *of the form*

$$H(\psi, \theta, I, A) = v \cdot A + E(\psi, \theta, I) \tag{148}$$

satisfying

$$\|H - H_0\|_{C^l(\mathbb{T}^{n+d}\times B_1)} \leqslant C_0 \tilde{\kappa}^4 \rho^{-2}, \tag{149}$$

then, there is a symplectic diffeomorphism $F \in \mathcal{S}_1 \cap \mathcal{C}^{l-\gamma-1}(\mathbb{T}^{n+d} \times B_1)$ *such that*

(A) *We have*

$$H \circ F(\psi, \theta, I, A) = c + v \cdot A + \omega \cdot I + \tfrac{1}{2} I^\top Q(\psi, \theta) I + R(\psi, \theta, I) \tag{150}$$

where $R = O(I^3)$.

(B) *F is close to the identity,*

$$||F - \mathrm{Id}\,||_{\mathcal{C}^{l-\gamma-1}(\mathbb{T}^{n+d}\times B_1)} \leqslant C\rho\tilde{\kappa}^{-2} ||H - H_0||_{\mathcal{C}^l(\mathbb{T}^{n+d}\times B_1)} \tag{151}$$

(C)
$$\|Q - Q_0\|_{\mathcal{C}^{l-\gamma-1}(\mathbb{T}^{n+d}\times B_1)} \leqslant C\rho\tilde{\kappa}^{-2}\, \|H - H_0\|_{\mathcal{C}^l(\mathbb{T}^{n+d}\times B_1)}\,. \tag{152}$$

Remark A.27. Note that we have followed [Zeh75,Zeh76a,Zeh76b] in assuming that the unperturbed Hamiltonian is analytic rather than just finite differentiable.

This assumption can be removed from Theorem A.26 by applying Lemma A.6 not only to the perturbation but also to the unperturbed one.

This would have been somewhat useful in our case since the unperturbed Hamiltonian is indeed finitely differentiable. However, we have decided not to include this improvement to remain reasonably close to the papers mentioned above. In our case, it is easy to modify a suggestion in [Zeh76a, p. 80] and use a Taylor approximation of the unperturbed problem so that we can consider an analytic unperturbed problem and apply Theorem A.26. See Section 4.3.6 for more details.

A.5.1. Proof of Theorem A.26

We apply Proposition A.6 with $\delta = 1$, $\eta = 8$, $r = 1$ to

$$\Delta \equiv H - H_0 \tag{153}$$

and obtain $\{\Delta^n\}_{n=0}^{\infty}$ with $\Delta^n \in \mathcal{H}_{8^{-n},1}$. We have, by (117) and (118) in Proposition A.6, that

$$\begin{aligned} \left\| \Delta^n \right\|_{8^{-n},1} &\leqslant C \left\| H - H_0 \right\|_{C^l} \\ &\leqslant C C_0 \tilde{\kappa}^4 \rho^{-2}. \end{aligned} \tag{154}$$

Moreover,

$$\begin{aligned} \left\| \Delta^{n+1} - \Delta^n \right\|_{8^{-(n+1)},1} &\leqslant C 8^{-l(n+1)} \left\| H - H_0 \right\|_{C^l} \\ &\leqslant C 8^{-l(n+1)} C_0 \tilde{\kappa}^4 \rho^{-2}. \end{aligned} \tag{155}$$

We introduce the notation $H^n = \Delta^n + H_0$. The proof of Theorem A.26 consists in showing inductively that there is a symplectic change of variables F^n such that $H^n \circ F^n$ is of the form (150).

By the definition of H^n, we have

$$H^{n+1} \circ F^n = (\Delta^{n+1} - \Delta^n) \circ F^n + H^n \circ F^n \tag{156}$$

Hence, $H^{n+1} \circ F^n$ is a small perturbation of $H^n \circ F^n$, which can be considered as an unperturbed system in Theorem A.15. Then, we apply Theorem A.15 to (156), and obtain a transformation $\hat{F}^n$, which is close to the identity and such that $H^{n+1} \circ F^n \circ \hat{F}^n$ is of the form (150). Setting $F^{n+1} = F^n \circ \hat{F}^n$, finishes the induction step.

Moreover, using the estimates of $\left\| \hat{F}^n - \mathrm{Id} \right\|_{(1/6)\cdot 8^{-(n+1)},1,\mathcal{C}^1}$ provided in (130) in the conclusions of Theorem A.15, we will establish bounds for $\left\| F^n - F^{n+1} \right\|_{(1/8)\cdot 8^{-(n+1)},1,\mathcal{C}^0}$, which, using the easy part of Proposition A.6, will yield that the limiting transformation satisfies the regularity claimed in Theorem A.26. Bounds (130) also allow to show that the smallness assumptions needed to apply Theorem A.15 are satisfied. They will also be useful to establish that all the compositions we have indicated are defined in the appropriate domains.

Let us now carry out the detailed estimates.

In the first step of the induction we find F^0 such that $H^0 \circ F^0 = (\Delta^0 + H_0) \circ F^0$ is of the form (150).

To apply Theorem A.15 to H^0, the only condition that we need to verify is the smallness condition on $\left\| \Delta^0 \right\|_{1,1}$, which, given (154), is implied by the smallness assumption on $\left\| H - H_0 \right\|_{\mathcal{C}^l}$ in Theorem A.26.

Applying Theorem A.15, we obtain that there is $F^0 \in \mathcal{S}_1 \cap \mathcal{H}_{1/4,\infty}$ such that $H^0 \circ F^0$ is of the form (150). Moreover,

$$\begin{aligned} \left\| F^0 - \mathrm{Id} \right\|_{1/6,1,\mathcal{C}^1} &\leqslant C\rho\tilde{\kappa}^{-2} \left\| \Delta^0 \right\|_{1,1} \\ &\leqslant C\rho\tilde{\kappa}^{-2} \left\| H - H_0 \right\|_{\mathcal{C}^l} . \end{aligned}$$

Because of the previous bound, and assuming that its right-hand side is smaller than $(1/4) \cdot 8^{-1}$, which is implied by smallness of $\|H - H_0\|_{\mathcal{C}^l}$, we have for the domains introduced in Definition A.4:

$$F^0(U_{(3/4)\cdot 8^{-1},1}) \subset U_{8^{-1},1}.$$

Hence, we can define $H^1 \circ F^0 = [(\Delta^1 - \Delta^0) + H_0] \circ F^0$ on $U_{(3/4)\cdot 8^{-1},1}$ and we can bound

$$\begin{aligned} \left\| (\Delta^1 - \Delta^0) \circ F^0 \right\|_{(3/4)\cdot 8^{-1},1} &\leqslant \left\| \Delta^1 - \Delta^0 \right\|_{8^{-1},1} \\ &\leqslant C\, 8^{-l} \left\| H - H_0 \right\|_{\mathcal{C}^l} \end{aligned}$$

Applying conclusions (131), we obtain that, denoting by Q^1 the quadratic part of H^1, we have:

$$\left\| Q^1 - Q^0 \right\|_{1/4} \leqslant C\rho\tilde{\kappa}^{-2} ||\Delta^0||_{1,1}.$$

This finishes the first step in the induction. Now, we present the estimates for the general inductive step.

We will assume inductively that the twist constant ρ in all the steps is not bigger than 2-times the initial constant.

We will assume inductively that we have applied the inductive procedure n times and defined F^n in such a way that $H^n \circ F^n$ is of the form (150)

We will assume inductively that the functions $\hat{F}^n \in \mathcal{S}_1$ satisfy

$$\left\| \hat{F}^i - \mathrm{Id} \right\|_{(1/6)\cdot 0.9\cdot 8^{-i},1,\mathcal{C}^1} \leqslant C\tilde{\kappa}^2 \rho 8^{-i(l-(\gamma+1))} \left\| H - H_0 \right\|_{\mathcal{C}^l} \tag{157}$$

for $i = 0, \ldots, n$.

Assumption (157) implies, when $\|H - H_0\|_{\mathcal{C}^l}$ is small enough that

$$\left\| F^n - \mathrm{Id} \right\|_{(1/6)\cdot 8^{-n},1,\mathcal{C}^1} \leqslant 1/10 \tag{158}$$

A consequence of (158) is that

$$F^n(U_{0.9\cdot 8^{-(n+1)},1}) \subset U_{8^{-(n+1)},1}$$

Therefore, we have that the compositions $\Delta^{n+1} \circ F^n$, $\Delta^n \circ F^n$ are defined on $U_{0.9\cdot 8^{-(n+1)},1}$ and we have

$$\begin{aligned} \left\| (\Delta^{n+1} - \Delta^n) \circ F^n \right\|_{0.9\cdot 8^{-(n+1)},1} &\leqslant \left\| (\Delta^{n+1} - \Delta^n) \circ F^n \right\|_{8^{-(n+1)},1} \\ &\leqslant C 8^{-(n+1)l} \left\| H - H_0 \right\|_{C^l}. \end{aligned}$$

Under the condition $l \geqslant 2\gamma + 3$, and that $\|H - H_0\|_{C^l}$ is sufficiently small (note that the smallness condition of $\|H - H_0\|_{C^l}$ is independent of n), condition (127) is satisfied and we can apply Theorem A.15 to obtain $\hat{F}^{n+1}$ as indicated before. Out of the conclusions of Theorem A.15 we obtain

$$\left\| \hat{F}^{n+1} - \mathrm{Id} \right\|_{(1/6)\cdot 0.9\cdot 8^{-(n+1)},\mathcal{C}^1} \leqslant C\tilde{\kappa}^2 \rho 8^{-(n+1)(l-(\gamma+1))} \left\| H - H_0 \right\|_{C^l}. \tag{159}$$

Eq. (159), is, of course, the inductive assumption (157) for $n+1$.

Another consequence of the application of Theorem A.15 is:

$$\left\| Q^{n+1} - Q^n \right\|_{0.9\cdot 8^{-(n+1)}} \leqslant C\tilde{\kappa}^2 \rho 8^{-(n+1)(l-(\gamma+1)} \left\| H - H_0 \right\|_{C^l}.$$

Therefore, if $\|H - H_0\|_{C^l}$ is sufficiently small, then the twist condition at all steps is less than twice the initial value.

Finally, we show that the F^n's produced converge in $\mathcal{C}^1$ to a transformation which is differentiable as claimed.

We estimate

$$\begin{aligned} \left\| F^{n+1} - F^n \right\|_{(1/6)\cdot 8^{-(n+1)},1,\mathcal{C}^1} &= \left\| F^n \circ \hat{F}^n - F^n \circ \mathrm{Id} \right\|_{(1/6)\cdot 8^{-(n+1)},1,\mathcal{C}^1} \\ &\leqslant \left\| F^n \right\|_{(1/3)\cdot 8^{-(n+1)},1,\mathcal{C}^1} \cdot 6 \cdot 8^{n+1} \\ &\quad \cdot \left\| \hat{F}^n - \mathrm{Id} \right\|_{(1/6)\cdot 8^{-(n+1)},1,\mathcal{C}^1} \\ &\leqslant (1/10 + 1) C\tilde{\kappa}^2 \rho 8^{-(n+1)(l-(\gamma+2))} \left\| H - H_0 \right\|_{C^l} \end{aligned} \tag{160}$$

The first equality above is just the definition of $F^{n+1} = F^n \circ \hat{F}^n$. The second inequality is the intermediate value theorem and Cauchy inequalities for F^n. The third inequality is just the inductive assumption on the size of $\hat{F}^n$. Note that in the second line, we have used that $\hat{F}^n(U_{(1/6)\cdot 8^{-(n+1)},0.9}) \subset U_{8^{-n},1}$ which follow from the fact that $\hat{F}^n$ is close to the identity and linear in the actions.

Applying Cauchy inequalities—which is just the easy implication of Proposition A.6—we obtain that, provided that $0 < l - \gamma - 1$, F^n converges in C^0. Inequalities (160) and Proposition A.6 imply that the limiting transformation is $\mathcal{C}^{l-\gamma-1}$ as claimed and the bounds claimed on $\|F - \mathrm{Id}\|_{\mathcal{C}^l}$.

Appendix B. Proof of Theorem 4.1

In this appendix we present a proof of Theorem 4.1

The theorem follows from more general results in the theory of normally hyperbolic manifolds which undoubtedly are well known to the experts. Nevertheless, the result presents some peculiarities such as having the manifold not compact, which are not often addressed in the literature (it is, however, a standard remark that compactness only enters to assume that the vector fields are uniformly bounded and uniformly continuous, so that if one assumes that instead of compactness one has the uniform continuity, the standard proof goes through). Even if the vector field in our problem is not uniformly bounded, using the scaling properties we will see that the Poincaré map with respect to a conveniently chosen section is uniformly bounded together with its derivatives up to of order $r-1$. The derivatives of order $r-1$ are uniformly continuous.

In our case, taking advantage of the peculiarities of the system, it is easy to give a proof that is simpler than the standard arguments (notably, the extension arguments which are cumbersome in the general case, can be done rather easily for our model using that we have a global system of coordinates). We can also avoid the use of adapted metrics, bundle maps, some linear operators become constants, etc.

We present the proof for the sake of making this paper more self-contained. We point out for experts that after the system of coordinates (Section B.1) and the use of Poincaré map (Section B.3) the rest are small modifications of fairly standard methods.

One important difference between our case and the general results on normally hyperbolic invariant manifolds is that, in our case, the manifolds we produce are invariant and not just locally invariant as one concludes from the general theory of normally hyperbolic invariant manifolds. The invariance of the manifold $\tilde{\Lambda}$ happens because the KAM circles, being codimension 1, separate the manifold. The region bounded between KAM tori is invariant. The fact that the manifold is invariant has deep consequences. For example, the definitions of (un)stable manifolds, which are quite cumbersome for locally invariant manifolds become very transparent. Also, the invariant manifold is unique, which eliminates some ambiguities.

B.1. System of coordinates

In a neighborhood of the unperturbed manifold $\tilde{\Lambda}$ we consider the coordinates $(\varphi, I, \theta, s, u) \in \mathbb{T} \times \mathbb{R} \times \mathbb{T}^d \times \mathbb{R}^{d_s} \times \mathbb{R}^{d_u}$ which correspond, respectively, to the phase of the periodic orbit Λ_E, $I = \sqrt{2E}$, the phases of the external perturbation and the stable and unstable directions of the unstable orbits, normalized in such a way that the expansion rate is constant around the periodic orbit. These coordinates are constructed in Floquet theory.

Even in Theorem 4.1 de dimension of the stable and unstable directions are $d_s = d_u = n-1$ we will do the proof in the general case. We will use the scaled variables introduced in Section 4.2 but work on the whole manifold (we will introduce later several cut-offs).

We will also add the extra variable $\delta \in B_{\delta^*} \subset \mathbb{R}^p$, which does not evolve in time. This standard device is a way to introduce parameters in the problem. We will prove the existence of a smooth manifold parameterized by $\varphi, \theta, I, \delta$. From this, it follows that for each δ, there is a smooth manifold and that they can be joined in a smooth family.

Due to the rescaling properties (7) the equations of the geodesic flow in a neighborhood of $\tilde{\Lambda}$ expressed in this system of coordinates are:

$$\begin{aligned}
\dot{\varphi} &= I + I\tilde{N}_\varphi(\varphi, \theta, I, \delta, s, u),\\
\dot{I} &= 0 + I\tilde{N}_I(\varphi, \theta, I, \delta, s, u),\\
\dot{s} &= IA_s s + I\tilde{N}_s(\varphi, \theta, I, \delta, s, u),\\
\dot{u} &= IA_u u + I\tilde{N}_u(\varphi, \theta, I, \delta, s, u),\\
\dot{\theta} &= v/\tau,\\
\dot{\delta} &= 0,
\end{aligned} \tag{161}$$

where A_s, A_u are constant matrices, the functions N are $\mathcal{C}^{r-1}$ and the $r-1$ derivatives are uniformly continuous. Moreover

$$\begin{aligned}
\tilde{N}(\varphi, \theta, \delta, I, s = 0, u = 0) &= 0,\\
D_s\tilde{N}(\varphi, \theta, \delta, I, s = 0, u = 0) &= 0,\\
D_u\tilde{N}(\varphi, \theta, \delta, I, s = 0, u = 0) &= 0.
\end{aligned} \tag{162}$$

We note that (162) is just an expression of the fact that, by definition, we have segregated the terms of order 0 and 1 in the Taylor expansion.

The fact that the terms of order 0 and 1 take the indicated expression—in particular that A_s, A_u are constant matrices—is a consequence of the scaling properties (7) of the geodesic flow. For us, the more important consequence of the scaling is that the vector field can be written as I multiplied by a smooth function, which is reflected in (161).

The perturbed equations (37) written in the scaled coordinates take the form

$$\begin{aligned}
\dot{\varphi} &= I + I\tilde{N}_{\varphi}(\varphi, \theta, I, s, u, \delta) + \delta\tilde{P}_{\varphi}(\varphi, \theta, I, s, u, \delta),\\
\dot{I} &= I\tilde{N}_{I}(\varphi, \theta, I, s, u, \delta) + \delta\tilde{P}_{I}(\varphi, \theta, I, s, u, \delta),\\
\dot{s} &= IA_s s + I\tilde{N}_{s}(\varphi, \theta, I, s, u, \delta) + \delta\tilde{P}_{s}(\varphi, \theta, I, s, u, \delta),\\
\dot{u} &= IA_u u + I\tilde{N}_{u}(\varphi, \theta, I, s, u, \delta) + \delta\tilde{P}_{u}(\varphi, \theta, I, s, u, \delta),\\
\dot{\theta} &= \nu/\tau,\\
\dot{\delta} &= 0
\end{aligned} \tag{163}$$

in the domain

$$\widehat{\mathcal{D}} = \{I \in [1, \infty), \varphi \in \mathbb{T}, \theta \in \mathbb{T}^d,\ s^2 + u^2 \leqslant \rho^2\}. \tag{164}$$

(We recall that we have taken $\delta = \varepsilon^2$, $\varepsilon = E_0^{-1/2}$, so that in the original coordinates, this corresponds to the domain $I_{\mathrm{phys}} \geqslant \sqrt{2E^*}$.)

The function P in (163) as well as its derivatives of order up to $r-1$ are uniformly bounded and uniformly continuous in the domain considered since they correspond to derivatives of the scaled potential. The fact that the derivatives of P and N are bounded will be important since it will allow us to work with a non-compact manifold. We also note that in our case, the derivatives of order $r-1$ of $\tilde{N}$ and $\tilde{P}$ are uniformly continuous.

B.2. Extensions of the equations

As usual in the theory of invariant manifolds, we will proceed to extend Eqs. (161) to the domain

$$\mathcal{D} = \{\varphi \in \mathbb{T},\ I \in \mathbb{R},\ s \in \mathbb{R}^{d_s},\ u \in \mathbb{R}^{d_u}\ , \theta \in \mathbb{T}^d\} \tag{165}$$

in such a way that they agree with the original unperturbed equations (161) in a small neighborhood of $\{s = 0,\ u = 0\}$.

Note that we are extending the equations to the coordinate space, not to the geometric phase space. In our case, the identification between coordinates and points of the phase space happens only in a neighborhood of the unperturbed invariant manifold.

We will show that the extended perturbed equations have a manifold invariant in a small neighborhood of the set $\{s = 0,\ u = 0\}$.

It follows that a manifold which is invariant for the extended system is locally invariant for the original system (that is, the true evolution of a point in the manifold, remains in the manifold or takes it to a place where the extended equations do not agree with the original equations). Later, in Section B.7, we will show that indeed the manifold is invariant and that the second alternative above does not happen.

Similarly, the perturbed equations will be extended to the set $\mathcal{D}$ in such a way that they agree with the original equation (163) in the set $\widehat{\mathcal{D}}$ in (164).

We will show that the extension of (163) has an invariant manifold contained in $\{s^2 + u^2 \leqslant \sigma(\delta)\}$ where $\sigma(\delta) \underset{\delta\to 0}{\to} 0$. Therefore, for $|\delta|$ small enough we obtain that $\sigma(\varepsilon) < \rho$ and the invariant manifold of the extended system is a locally invariant manifold for the original system.

Let $\psi : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^\infty$ function such that

$$\psi(x) = 1 \text{ when } |x| \leqslant 1,$$
$$\psi(x) = 0 \text{ when } |x| \geqslant 2.$$

We first modify the vector field (161) so that

$$\dot{\varphi} = \psi(2|I|) + (1 - \psi(2|I|)) \left(|I| + I\tilde{N}_\varphi(\ ) \right). \tag{166}$$

Clearly, this new vector field agrees with the original one (161) in $\{|I| \geqslant 1\}$.

The reason to introduce this change is that in the subsequent proof we will consider φ as a fast variable with respect to θ. This, clearly runs into problems when $I = 0$. When we introduce the above change of vector field, the variable φ will indeed by uniformly fast in all the domain and we can simplify the exposition by using the same techniques over all the domain.

We furthermore introduce

$$N(\varphi, \theta, I, \delta, s, u) = \tilde{N}(\varphi, \theta, I, \delta, s, u)\psi\left(\frac{s^2 + u^2}{\rho^2}\right)(1 - \psi(|I|))$$

and

$$P(\varphi, \theta, I, \delta, s, u) = \tilde{P}(\varphi, \theta, I, \delta, s, u)\psi\left(\frac{s^2 + u^2}{\rho^2}\right)(1 - \psi(|I|)).$$

Clearly the extended system can be defined in $\mathcal{D}$ and agrees with the original one in $\hat{\mathcal{D}}$. By (162) we have that the $\mathcal{C}^{r-1}$ norm of N is small.

B.3. The Poincaré map

The extended vector field corresponding to (161) is unbounded in $\mathcal{D}$. This makes it hard to adapt the general results in the literature on normally hyperbolic invariant manifolds directly. (We note, however, that the unboundedness comes in the good terms, since the unboundedness of the vector field makes the hyperbolicity become stronger, so that a more general proof than the one we present here is possible.)

In our case, however, since the unboundedness is of a very special form, we can take advantage of this and reduce it to a more standard problem.

We observe that for system (161) the set $\{\varphi = 0\}$ is a global section.

The return map to this section of the unperturbed system (161) is given by

$$\begin{aligned} I' &= I + \hat{N}_I(\theta, I, \delta, s, u), \\ s' &= e^{A_s} s + \hat{N}_s(\theta, I, \delta, s, u), \\ u' &= e^{A_u} u + \hat{N}_u(\theta, I, \delta, s, u), \\ \theta' &= \theta + \varepsilon \nu \Gamma(\theta, I, \delta, s, u), \\ \delta' &= \delta, \end{aligned}$$

where Γ is the $\mathcal{C}^{r-1}$ function measuring the time it takes to go from the invariant section $\{\varphi = 0\}$ to it again.

Clearly, for large I, $\Gamma(\theta, I, \delta, s, u)$ is approximately $1/|I|$. On the other hand, $\Gamma(\theta, I, \delta, s, u)$ remains bounded for any value of I because we introduced change (166).

Recalling the scaling properties of the vector field, we will argue that the map $\hat{N} = (\hat{N}_I, \hat{N}_s, \hat{N}_u)$ is uniformly bounded with its derivatives of order up to $r-1$. Heuristically, the reason is that, as I increases, the time of return to the section decreases as I^{-1}, which balances exactly the growth of the vector field (which except for the explicit factor I is uniformly bounded). Indeed, the scaling transformations in Section 4.2 transform the return map to $\varphi = 0$ in a domain $I \in (I^*, 2I^*)$ into the return map of any other such domain. Hence, the $\mathcal{C}^{r-1}$ norm of all these return maps are bounded uniformly in I^*. Also, the $r-1$ derivatives are uniformly continuous.

We also have

$$\begin{aligned} \hat{N}(\theta, I, \delta, 0, 0) &= 0, \\ D_s \hat{N}(\theta, I, \delta, 0, 0) &= 0, \\ D_u \hat{N}(\theta, I, \delta, 0, 0) &= 0, \\ \Gamma(\theta, I, \delta, 0, 0) &= \frac{1}{I}, \\ D_s \Gamma(\theta, I, \delta, 0, 0) &= 0, \\ D_u \Gamma(\theta, I, \delta, 0, 0) &= 0. \end{aligned}$$

The Poincaré map of the perturbed flow (163) has the form

$$\begin{aligned} I' &= I + \hat{N}_I(\theta, I, \delta, s, u) + \delta \hat{P}_I(\theta, I, \delta, s, u), \\ s' &= e^{A_s} s + \hat{N}_s(\theta, I, \delta, s, u) + \delta \hat{P}_s(\theta, I, \delta, s, u), \\ u' &= e^{A_u} u + \hat{N}_u(\theta, I, \delta, s, u) + \delta \hat{P}_u(\theta, I, \delta, s, u), \\ \theta' &= \theta + \varepsilon v \Gamma(\theta, I, \delta, s, u) + \delta \hat{P}_{\theta(\theta, I, \delta, s, u)}, \end{aligned} \tag{167}$$

where $\|e^{A_s}\|, \|e^{-A_u}\| \leqslant \lambda < 1$.

Henceforth, we will denote the return map by F.

B.4. Equations for the invariant manifolds

As it is standard in invariant manifold theory we will write the invariant manifold as the graph of a function that gives s, u as a function of θ, I, δ.

We introduce the notation $x = (\theta, I, \delta)$.

Let W be a function that to a point $x \in \mathbb{T}^d \times \mathbb{R} \times B_{\delta^*}$ associates $(W_s(x), W_u(x))$ which are coordinates in the s, u directions.

Given a point

$$(x, W_s(x), W_u(x)) = (x, W(x)) \tag{168}$$

in the graph of W, its image under the map in (167) is in the graph of W if and only if, after applying (167) to (168), we obtain that the s, u components can be expressed as the function W evaluated on the $x = (\theta, I, \delta)$ components of $F(x, W(x))$.

This amounts to

$$\begin{aligned} W_s(R(x)) &= e^{A_s} W_s(x) + \hat{N}_s(x, W_s(x), W_u(x)) + \delta \hat{P}_s(x, W_s(x), W_u(x)), \\ W_u(R(x)) &= e^{A_u} W_u(x) + \hat{N}_u(x, W_s(x), W_u(x)) + \delta \hat{P}_u(x, W_s(x), W_u(x)), \end{aligned} \tag{169}$$

where the function $R : \mathbb{T}^d \times \mathbb{R} \times B_{\delta^*} \to \mathbb{T}^d \times \mathbb{R} \times B_{\delta^*}$ is given by

$$\begin{aligned} R(x) &= R(\theta, I, \delta) \\ &= (\theta + v/\tau \Gamma(\theta, I, \delta, W_s(\theta, I, \delta), W_u(\theta, I, \delta)) \\ &\quad + \delta \hat{P}_\theta(\theta, I, \delta, W_s(\theta, I, \delta), W_u(\theta, I, \delta)), \\ &\quad I + \hat{N}_I(\theta, I, \delta, W_s(\theta, I, \delta), W_u(\theta, I, \delta)) \\ &\quad + \delta \hat{P}_I(\theta, I, \delta, W_s(\theta, I, \delta), W_u(\theta, I, \delta)), \delta). \end{aligned} \tag{170}$$

Note that R depends on W, but we will omit the dependence from the notation for simplicity, except in some points where the dependence is particularly important.

We will rearrange (169) and (170) in such a way that they can be reduced to a fixed point theorem.

We will formulate the fixed point theorem in the space of functions $W : \mathbb{T}^d \times \mathbb{R} \times \mathbb{R}^p \mapsto \mathbb{R}^{d_s} \times \mathbb{R}^{d_u}$ such that

$$\|W\|_{\mathcal{C}^{r-1}} \leqslant \sigma(\delta, \|\hat{N}\|_{\mathcal{C}^{r-1}}),$$

where $\sigma = \sigma(\delta, \|\hat{N}\|_{\mathcal{C}^{r-1}})$ is a function which will be described along the proof. We note that σ tends to zero when its arguments tend to zero.

We first note that if $\left\|\hat{N}\right\|_{\mathcal{C}^1}$, δ, $\|W\|_{\mathcal{C}^1}$ are sufficiently small, R is a $\mathcal{C}^1$-small perturbation of the identity, hence, globally invertible.

Eqs. (169) can be re-written, composing the first equation with R^{-1} on the left and applying algebraic manipulation on the second as

$$\begin{aligned} W_s(x) &= e^{A_s} W_s \circ R^{-1}(x) + \hat{N}_s(R^{-1}x, W \circ R^{-1}(x)) \\ &\quad + \delta \hat{P}_s(R^{-1}(x), W \circ R^{-1}(x)), \\ W_u(x) &= e^{-A_u}[W_u \circ R(x) - \hat{N}_u(x, W(x)) - \delta \hat{P}_u(x, W(x)]. \end{aligned} \tag{171}$$

B.5. Solution of the fixed point equations

The Eqs. (171) are the fixed point formulation of Eqs. (169) which express the fact that the graph of W is invariant.

We will consider the equations as a fixed point problem for the operator $\mathcal{T}$ that to a function W associates the right-hand side of (171). We will denote by $\mathcal{T}_s[W]$ the right-hand side of the first equation in (171) and by $\mathcal{T}_u[W]$ the right-hand side of the second.

Remark B.1. For the experts we note that the operator $\mathcal{T}$ that we have introduced above is not the graph transform operator. It is rather an operator that has the same fixed points as the graph transform.

The graph transform operator $\mathcal{G}$ is defined by

$$F(\text{Graph}(W)) = \text{Graph}(\mathcal{G}(W)).$$

When the functions W map into stable and unstable components, the graph transform is not a contraction. We will, however, find the graph transform useful when discussing stable and unstable manifolds.

See Remark B.4 for a discussion of the proof of the existence and regularity for $\tilde{\Lambda}$ using the graph transform.

The analysis of Eqs. (171) will be done using the uniform $\mathcal{C}^{r-1}$ spaces in $\mathbb{T}^d \times \mathbb{R} \times B_{\delta^*}$. These Banach spaces are the spaces of $r-1$ times continuously differentiable functions with the norm

$$\|f\|_{\mathcal{C}^{r-1}} = \sup_x \max_{0 \leqslant i \leqslant r-1} \|D^i f(x)\| \ .$$

Note that for these spaces we not only require that the derivatives are continuous but also that they are uniformly bounded. This is a significant restriction for functions defined in domains which are not compact.

Following the standard procedure in the study of invariant manifolds, we will show

(a) That the operator $\mathcal{T}$ indeed is well defined on

$$\mathcal{B} \equiv \{\|W\|_{\mathcal{C}^{r-1}} \leqslant \sigma(\varepsilon, \|\hat{N}\|_{\mathcal{C}^{r-1}})\}.$$

(b) That $\mathcal{T}(\mathcal{B}) \subset \mathcal{B}$.

(c) That $\mathcal{T}$ is a contraction on $\mathcal{B}$ in the $\mathcal{C}^0$ norm topology. That is

$$\|\mathcal{T}(W) - \mathcal{T}(\tilde{W})\|_{\mathcal{C}^0} \leqslant K\|W - \tilde{W}\|_{\mathcal{C}^0}, \quad K < 1$$

for $W, \tilde{W} \in \mathcal{B}$.

Then, it follows that $\mathcal{T}$ has a fixed point in the $\mathcal{C}^0$ closure of $\mathcal{B}$.

From Ascoli-Arzelà theorem it follows that $\mathcal{C}^0$-closure of $\mathcal{B}$ is contained in $\mathcal{C}^{r-2+\mathrm{Lip}}$, which is what we had claimed.

Remark B.2. A much more general characterization of the closure (in topologies much weaker than $\mathcal{C}^0$) of $\mathcal{B}$ can be found in [LI73]. The characterization in [LI73] also works in infinite dimensional Banach spaces.

We refer to [dlLO99] for a compendium of properties of the composition operator in $\mathcal{C}^\ell$ spaces which we will find useful.

In particular, we will use repeatedly that

$$\|f \circ g\|_{\mathcal{C}^{r-1}} \leqslant C\|f\|_{\mathcal{C}^{r-1}}(1 + \|g\|_{\mathcal{C}^{r-1}})^{r-1}, \tag{172}$$

where C is a constant which only depends on r.

The proof of (172) is not very hard. It suffices to note that if we take derivatives of $f \circ g$ of order $j \leqslant r-1$, we obtain a polynomial expression in derivatives of f, g. All the terms contain a derivative of f of order not larger than j (a fortiori not larger than $r-1$). A term contains not more than $r-1$ factors all of which are derivatives of g of order not larger than $r-1$. We use the estimate that any product of k derivatives of g of order up to $r-1$ can be estimated by

$$\left\| D^{l_1} g \cdots D^{l_k} \right\|_{\mathcal{C}^0} \leqslant C(1 + \|g\|_{\mathcal{C}^{r-1}})^{r-1}. \tag{173}$$

Of course, estimate (173) is very conservative and, when for instance we know that $\|g\|_{\mathcal{C}^{r-1}} \leqslant 1$ we can obtain more precise estimates.

Since we will often be considering composition with functions which are close to the identity, it is also useful to remark that, if $\|g - \mathrm{Id}\|_{\mathcal{C}^{r-1}} \leqslant 1/2$ we have

$$\|f \circ g\|_{\mathcal{C}^{r-1}} \leqslant C \, \|f\|_{\mathcal{C}^{r-1}} \, (1 + \|g - \mathrm{Id}\|_{\mathcal{C}^{r-1}})^{r-1}. \tag{174}$$

The proof of (174) follows the same lines than the proof of (172). We just write $Dg = D(g - \mathrm{Id}) + \mathrm{Id}$ and $D^l g = D^l(g - \mathrm{Id})$ for $2 \leqslant l \leqslant r-1$. Then, we note that $D^j f \circ g$ can be expressed in the same way as before as a polynomial in derivatives of the desired form.

B.5.1. The operator $\mathcal{T}$ is well defined

The fact that the operator $\mathcal{T}$ is well defined has been essentially accomplished.

From the estimates of the composition of two $\mathcal{C}^{r-1}$ functions we obtain that

$$\begin{aligned} \|\hat{N}(\cdot, W(\cdot))\|_{\mathcal{C}^{r-1}} &\leqslant C\|\hat{N}\|_{\mathcal{C}^{r-1}}(1 + \|W\|_{\mathcal{C}^{r-1}})^{r-1}, \\ \|\hat{P}(\cdot, W(\cdot)\|_{\mathcal{C}^{r-1}} &\leqslant C\|\hat{P}\|_{\mathcal{C}^{r-1}}(1 + \|W\|_{\mathcal{C}^{r-1}})^{r-1}. \end{aligned}$$

Hence, under smallness assumptions on $\|\hat{N}\|_{\mathcal{C}^{r-1}}$ and δ, and assuming

$$\sigma(\delta, \|\hat{N}\|_{\mathcal{C}^{r-1}}) \leqslant \delta, \tag{175}$$

we obtain that R from (170) is a $\mathcal{C}^{r-1}$ perturbation of the identity, hence, R^{-1} will be defined for all functions W in $\mathcal{B}$. Moreover, $\|R_W^{-1} - \mathrm{Id}\,\|_{\mathcal{C}^{r-1}}$ will be uniformly small for all functions in W.

We denote by

$$\gamma = \gamma(\delta, \|\hat{N}\|_{\mathcal{C}^{r-1}}) = \sup_{W \in \mathcal{B}} \max(\|R_W - \mathrm{Id}\,\|_{\mathcal{C}^{r-1}}, \|R_W^{-1} - \mathrm{Id}\,\|_{\mathcal{C}^{r-1}})$$

and we note that $\gamma = \gamma(\delta, \|\hat{N}\|_{\mathcal{C}^{r-1}})$ is small if ε, $\|N\|_{\mathcal{C}^{r-1}}$ are small.

The above considerations show that the operator $\mathcal{T}$ is well defined for all the functions $W \in \mathcal{B}$.

B.5.2. The range of the operator $\mathcal{T}$

The fact that $\mathcal{T}$ maps the space $\mathcal{B}$ into itself follows because, by applying (172), (174) and the triangle inequality to the definition of $\mathcal{T}$, we obtain

$$\begin{aligned}
\|\mathcal{T}_s[W]\|_{\mathcal{C}^{r-1}} &\leqslant \|e^{A_s}\| C\sigma(1+\gamma)^{r-1} \\
&\quad + C(\|\hat{N}\|_{\mathcal{C}^{r-1}} + \delta\|\hat{P}\|_{\mathcal{C}^{r-1}}) \cdot (1 + C\sigma(1+\gamma)^{r-1})^{r-1}, \\
\|\mathcal{T}_u[W]\|_{\mathcal{C}^{r-1}} &\leqslant \|e^{-A_u}\| [C\sigma(1+\gamma)^{r-1} \\
&\quad + C(\|\hat{N}\|_{\mathcal{C}^{r-1}} + \delta\|\hat{P}\|_{\mathcal{C}^{r-1}}) \cdot (1+\sigma)^{r-1})]
\end{aligned} \tag{176}$$

We see that (176) have the structure

$$\|\mathcal{T}\|_{\mathcal{C}^{r-1}} \leqslant \lambda\sigma\left(\delta, \left\|\hat{N}\right\|_{\mathcal{C}^{r-1}}\right) + \mu,$$

where $\lambda < 1$ and $\mu = \mu(\delta, \left\|\hat{N}\right\|_{\mathcal{C}^{r-1}})$ is a polynomial expression which goes to zero if $\delta, \left\|\hat{N}\right\|_{\mathcal{C}^{r-1}}$ converge to 0.

It, therefore, suffices to take

$$\sigma < \frac{\mu}{1-\lambda} \tag{177}$$

to ensure that the set $\mathcal{B}$ is mapped into itself.

B.5.3. Contraction properties of $\mathcal{T}$

The estimates for the contraction in $\mathcal{C}^0$ are not very difficult, even if a bit repetitive.

The only tools needed are elementary techniques such as adding and subtracting, the triangle inequality and the following estimate, which is an easy consequence of the mean value theorem:

$$\|f \circ W - f \circ \widetilde{W}\|_{\mathcal{C}^0} \leqslant \|f\|_{\mathcal{C}^1} \|W - \widetilde{W}\|_{\mathcal{C}^0}. \tag{178}$$

A good heuristic guide is to consider that all the terms involving $\hat{N}$, $\hat{P}$ are ignorable since they contribute to quantities that have Lipschitz constants that are arbitrarily small. As we will see, the estimates needed to justify this heuristics are straightforward—albeit long—applications of the above elementary techniques.

We start by estimating R_W in detail. As in the previous proof, we write μ to denote terms which are arbitrarily small by choosing $\delta \left\|\hat{N}\right\|_{\mathcal{C}^{r-1}}$ small. We observe that

$$\begin{aligned}
\|\Gamma(\cdot, W(\cdot)) - \Gamma(\cdot, \widetilde{W}(\cdot))\|_{\mathcal{C}^0} &\leqslant \mu\|W - \widetilde{W}\|_{\mathcal{C}^0}, \\
\delta\|\hat{P}(\cdot, W) - \hat{P}(\cdot, \widetilde{W}(\cdot))\|_{\mathcal{C}^0} &\leqslant \mu\|W - \widetilde{W}\|_{\mathcal{C}^0}.
\end{aligned}$$

Therefore

$$\|R_W - R_{\widetilde{W}}\|_{\mathcal{C}^0} \leqslant \mu \|W - \widetilde{W}\|_{\mathcal{C}^0}. \tag{179}$$

Since we have that

$$\|R - \mathrm{Id}\|_{\mathcal{C}^1}, \left\|R^{-1} - \mathrm{Id}\right\|_{\mathcal{C}^1} \leqslant 1/2,$$

we obtain from (179),

$$\|R_W - R_{\widetilde{W}}\|_{\mathcal{C}^0} \leqslant \mu \|W - \widetilde{W}\|_{\mathcal{C}^0} \tag{180}$$

and

$$\begin{aligned}\|W \circ R_W - \widetilde{W} \circ R_{\widetilde{W}}\|_{\mathcal{C}^0} &\leqslant \|W \circ R_W - \widetilde{W} \circ R_W\|_{\mathcal{C}^0} + \|\widetilde{W} \circ R_W - \widetilde{W} \circ R_{\widetilde{W}}\|_{\mathcal{C}^0} \\ &\leqslant \|W - \widetilde{W}\|_{\mathcal{C}^0} + \|\widetilde{W}\|_{\mathcal{C}^1} \mu \|W - \widetilde{W}\|_{\mathcal{C}^0} \\ &\leqslant (1+\mu)\|W - \widetilde{W}\|_{\mathcal{C}^0}. \end{aligned} \tag{181}$$

Similarly, we obtain

$$\|W \circ R_W^{-1} - \widetilde{W} \circ R_{\widetilde{W}}^{-1}\|_{\mathcal{C}^0} \leqslant (1+\mu)\|W - \widetilde{W}\|_{\mathcal{C}^0}. \tag{182}$$

From (182) we obtain

$$\begin{aligned}&\|\hat{N}_s(R_W^{-1}, W \circ R_W^{-1}) - \hat{N}_s(R_{\widetilde{W}}^{-1}, \widetilde{W} \circ R_W^{-1})\|_{\mathcal{C}^0} \\ &\quad\leqslant \|\hat{N}\|_{\mathcal{C}^1} \left(\|R_W^{-1} - R_{\widetilde{W}}^{-1}\|_{\mathcal{C}^0} + \|W \circ R_W^{-1} - \widetilde{W} \circ R_{\widetilde{W}}^{-1}\|_{\mathcal{C}^0}\right) \\ &\quad\leqslant \|\hat{N}\|_{\mathcal{C}^1}(1+2\mu)\|W - \widetilde{W}\|_{\mathcal{C}^0} \\ &\quad\leqslant \mu\|W - \widetilde{W}\|_{\mathcal{C}^0}.\end{aligned}$$

Identical estimates show

$$\|\delta\hat{P}_s(R_W^{-1}, W \circ R_W^{-1}) - \delta\hat{P}_s(R_{\widetilde{W}}^{-1}, \widetilde{W} \circ R_{\widetilde{W}}^{-1})\|_{\mathcal{C}^0} \leqslant \mu\|W - \widetilde{W}\|_{\mathcal{C}^0},$$

$$\|\hat{N}_u(\cdot, W) - \hat{N}_u(\cdot, \widetilde{W})\|_{\mathcal{C}^0} \leqslant \mu\|W - \widetilde{W}\|_{\mathcal{C}^0},$$

$$\|\delta\hat{P}_u(\cdot, W) - \delta\hat{P}_u(\cdot, \widetilde{W})\|_{\mathcal{C}^0} \leqslant \mu\|W - \widetilde{W}\|_{\mathcal{C}^0}.$$

The previous estimates show that

$$\|\mathcal{T}_s[W] - \mathcal{T}_s[\widetilde{W}]\|_{\mathcal{C}^0} \leqslant (\|e^{A_s}\| + \mu)\|W - \widetilde{W}\|_{\mathcal{C}^0},$$
$$\|\mathcal{T}_u[W] - \mathcal{T}[\widetilde{W}]\|_{\mathcal{C}^0} \leqslant (\|e^{-A_u}\| + \mu)\|W - \widetilde{W}\|_{\mathcal{C}^0}.$$

Therefore if δ, $\|\hat{N}\|_{\mathcal{C}^1}$ is small enough, we obtain that, as claimed, $\mathcal{T}$ is a contraction in $\mathcal{C}^0$ norm for the functions in $\mathcal{B}$.

Applying the contraction mapping theorem, we obtain that there is a fixed point in the C^0 closure of $\mathcal{B}$. By Ascoli-Arzelà theorem, this closure consists of functions which are $C^{r-2+\mathrm{Lip}}$.

This finishes the proof of the theorem with the slightly smaller regularity $\mathcal{C}^{r-2+\mathrm{Lip}}$ rather that with the claimed regularity $\mathcal{C}^{r-1}$.

B.5.4. Sharp regularity

To obtain the $\mathcal{C}^{r-1}$ regularity claimed, we note that we know by Rademacher's theorem that the $r-1$ derivative exists almost everywhere.

If we take derivatives of order $r-1$ of Eq. (171), we obtain, writing explicitly the terms which involve derivatives of W of order $r-1$,

$$\begin{aligned} D^{r-1}W_s(x) &= e^{A_s}D^{r-1}W_s \circ R^{-1}(x)(DR^{-1}(x))^{\otimes(r-1)} \\ &\quad + D_2\hat{N}_s(R^{-1}x, W \circ R^{-1}(x))D^{r-1}W(x)(DR^{-1}(x))^{\otimes(r-1)} \\ &\quad + \delta D_2\hat{P}_s(R^{-1}(x), W \circ R^{-1}(x))D^{r-1}W(x)(DR^{-1}(x))^{\otimes(r-1)} \\ &\quad + M_s(x), \\ D^{r-1}W_u(x) &= e^{-A_u}D^{r-1}W_u \circ R(x)(DR(x))^{\otimes r-1} \\ &\quad - D_2\hat{N}_u(x, W(x))D^{r-1}W(x) \\ &\quad - \delta D_2\hat{P}_u(x, W(x))D^{r-1}W_u(x) + M_u(x), \end{aligned} \tag{183}$$

where M_s, M_u are polynomial expressions involving derivatives of W, W_s, W_u up to order $r-2$ evaluated at x or at $R(x)$ or at $R^{-1}(x)$ and derivatives of $\hat{N}$, $\hat{P}$ of order up to $r-1$ evaluated at appropriate places.

The main observation is that M_s, M_u are continuous and are uniformly bounded.

We consider (183) as a fixed point equation for $D^{r-1}W$. We note that, because of the assumptions on the rates of contraction, we have that the right-hand side is a contraction in L^∞ provided that $\left\|\hat{N}\right\|_{\mathcal{C}^{r-1}}$ and δ are small enough. We also have that, because of the continuity of $M = (M_s, M_u)$, the operator given by the right-hand side maps $\mathcal{C}^0$ into $\mathcal{C}^0$. In this circumstances, we obtain the the L^∞ fixed point has to be in $\mathcal{C}^0$.

This establishes that the fixed point of (171) is $\mathcal{C}^{r-1}$. This is, the regularity that we had claimed for the manifold in Theorem 4.1.

By noting that $\hat{N}$, $\hat{P}$ have uniformly continuous derivatives of order $r-1$, we obtain that the right-hand side of (183) maps uniformly continuous functions into uniformly continuous functions. Since the uniform limit of uniformly continuous functions is uniformly continuous, we obtain that the $r-1$ derivatives of W are uniformly continuous.

B.6. The stable and unstable manifolds of $\tilde{\Lambda}$

In this section we establish the existence of the the stable and unstable manifolds to the invariant manifold $\tilde{\Lambda}$, we just constructed.

We consider first the unstable manifold, since its existence can be done by the graph transform.

We consider a function S that given θ, I, δ, u produces the coordinate s. We denote $y = (\theta, I, \delta, u)$. We will denote a point in the full space as (y, s). This is slightly inconsistent with the order we have used for the variables, but we hope it will not lead to confusion.

We consider a point $(y, S(y))$ in the graph of S. Its image under map (167) is in the graph of a map $\mathcal{G}(S)$ given by

$$\mathcal{G}[S](y) = e^{A_s} S(T^{-1}(y)) + N_s(T^{-1}(y), S(T^{-1}(y))) + \varepsilon^2 P_s(T^{-1}(y), S(T^{-1}(y))), \tag{184}$$

where T is the map given by (again, we temporarily suppress from the notation that T depends on S)

$$\begin{aligned} T(\theta, I, \delta, u) = \; & \theta + v/\tau\Gamma(y, S(y)) + \delta\hat{P}_\theta(y, S(y)), \\ & I + \hat{N}_I(y, S(y)) + \delta\hat{P}_I(y, S(y)), \\ & \delta, \\ & e^{A_u} u + \hat{N}_u(y, S(y)) + \delta\hat{P}_u(y, S(y))). \end{aligned} \tag{185}$$

The formulas above for the map are easily obtained noting that T is just θ', I', δ', u' obtained applying map (167) to the point $(y, S(y))$. Then, we just express the coordinate s' as function of those. Noting that $\hat{N}$, $\delta\hat{P}$ are $\mathcal{C}^1$ small, we can apply the implicit function theorem and obtain that T is indeed invertible and that $\mathcal{G}(S)$ is indeed a well defined map.

By following arguments very similar to—indeed somewhat simpler than—those in the previous section, it follows that, under strong enough smallness assumptions on the mapping, $\mathcal{G}$ maps a $\mathcal{C}^{r-1}$ ball to itself and it is a $\mathcal{C}^0$ contraction there. This establishes that the manifold is $\mathcal{C}^{r-2-\mathrm{Lip}}$. As before, we establish that it is $\mathcal{C}^{r-1}$ by studying the equation which is satisfied by the $r-1$ derivative. We leave details to the reader since they are identical to the argument presented in Section B.5.4.

Hence, we can construct a $\mathcal{C}^{r-1}$ invariant manifold which is a $\mathcal{C}^{r-1}$ perturbation of the coordinate manifold corresponding to (θ, I, δ). We will show that this manifold is $W^{\mathrm{u}}_{\tilde{\Lambda}}$.

We also note that the fact that $\mathcal{G}$ is a contraction in $\mathcal{C}^0$ tells us that the iterates of the coordinate manifold are converging exponentially fast in $\mathcal{C}^0$ to the invariant manifold.

Applying the results above to the inverse mapping we can obtain a $\mathcal{C}^{r-1}$ invariant manifold which is modelled on the coordinate manifold corresponding to (θ, I, δ, s). We will denote this manifold by $W^{\mathrm{s}}_{\tilde{\Lambda}}$.

We note that the manifold $W^{\mathrm{s}}_{\tilde{\Lambda}} \cap W^{\mathrm{u}}_{\tilde{\Lambda}}$ is invariant and is a small perturbation of the coordinate manifold along (θ, I, δ).

Since the manifold $\tilde{\Lambda}$ produced in the previous section is also invariant and is also $\mathcal{C}^0$ close to the coordinate manifold, and the solutions of (171) are unique—we are using the contraction mapping principle—, we obtain that the manifold $\tilde{\Lambda}$ produced in the previous section agrees with the intersection. That is

$$W^{\mathrm{u}}_{\tilde{\Lambda}} \cap W^{\mathrm{s}}_{\tilde{\Lambda}} = \tilde{\Lambda}.$$

Furthermore, by the fact that $\mathcal{G}$ preserves a $\mathcal{C}^1$ neighborhood of the coordinate map and that it is a contraction in the $\mathcal{C}^0$ distance there, we obtain that

$$\mathrm{dist}_{\mathcal{C}^0}(\mathcal{G}^n(0) \cap W^{\mathrm{s}}_{\tilde{\Lambda}}, \tilde{\Lambda}) \leqslant C\tilde{\lambda}^n$$

for $n > 0$, where $\tilde{\lambda}$ is a number which is arbitrarily close to λ if $\|\hat{N}\|_{\mathcal{C}^1}, \delta$ are small enough.

Recalling the definition of $\mathcal{G}$, we conclude that, given any point $x \in W^{\mathrm{s}}_{\tilde{\Lambda}}$, we have

$$\mathrm{dist}(F^n(x), \tilde{\Lambda}) \leqslant C\tilde{\lambda}^n \tag{186}$$

for $n \geqslant 0$.

Applying the argument for F^{-1}, we obtain that given any point $x \in W^{\mathrm{s}}_{\tilde{\Lambda}}$, we have

$$\mathrm{dist}(F^n(x), \tilde{\Lambda}) \leqslant C\tilde{\lambda}^n \tag{187}$$

for $n \leqslant 0$.

The following is a strengthening of a converse to (186) and (187). The arguments will play an important role in the discussion of uniqueness.

Proposition B.3. *Under the standard assumptions that* $\|\hat{N}\|_{\mathcal{C}^1}, \delta$ *are small enough, then all the points that satisfy* $\mathrm{dist}(F^n(x), \tilde{\Lambda}) \leqslant C$ *for all* $n \geqslant 0$, *are points in* $W^{\mathrm{s}}_{\tilde{\Lambda}}$.

Conversely, if $\operatorname{dist}(F^n(x), \tilde{\Lambda}) \leqslant C$ *for all* $n \leqslant 0$, *then* x *in* $W^{\mathrm{u}}_{\tilde{\Lambda}}$.

If $\operatorname{dist}(F^n, \tilde{\Lambda}) \leqslant C$ *for all* $n \in \mathbb{Z}$, *then* $x \in \tilde{\Lambda}$.

The proof of Proposition B.3 consists in showing that there is a field of cones that is preserved by the dynamics. In this cones, the unstable directions stretch. From this, it follows that given (θ, I, δ, s), there can only be a value of u for which the iterates remain in a neighborhood in the future. Since the points in $W^{\mathrm{s}}_{\tilde{\Lambda}}$ satisfy this, and $W^{\mathrm{s}}_{\tilde{\Lambda}}$ is a graph, it follows that the only points which remain bounded in the future are the points in the stable manifold. Similarly for the unstable manifold.

The proof of the persistence of cone fields is quite standard. We refer to [KH95, p. 245].

Remark B.4. Many expositions of the theory of normally hyperbolic manifolds prove the persistence of the invariant manifold by proving the persistence of the stable and the unstable manifold as in Section B.6 and then, constructing the invariant manifold as the intersection. Calculations of invariant manifolds using the stable and unstable approach have been undertaken in [BOV97].

From the point of view of numerical implementations, the proof presented here has the advantage that the algorithms for fixed δ have only to work with functions of two variables. The proof based on stable and unstable manifolds has to deal with functions of three variables.

B.6.1. The stable and unstable bundles of $\tilde{\Lambda}$

The following result about stability of bundles is quite standard. See for example [HP70].

The invariant manifold $\tilde{\Lambda}$ is normally hyperbolic in the sense that

Proposition B.5. *There is a splitting of the tangent space of the coordinate space* $\mathcal{D}$ *in* (165)

$$T_x\mathcal{D} = T_x\tilde{\Lambda} \oplus E^{\mathrm{s}}_x \oplus E^{\mathrm{u}}_x,$$

where there exist constants $C > 0$, $0 < \tilde{\lambda} < 1$ *such that*

$$v \in E^{\mathrm{s}}_x \iff |DF^n(x)v| \leqslant C|v|\tilde{\lambda}^n, \qquad n \geqslant 0,$$

$$v \in E^{\mathrm{u}}_x \iff |DF^n(x)v| \leqslant C|v|\tilde{\lambda}^{|n|}, \qquad n \leqslant 0.$$

Moreover, the mappings that to x *assign* $E^{\mathrm{s,u}}_x$ *are* $\mathcal{C}^{r-2}$.

The Proposition B.5 is a very standard perturbation result. Full details can be found in [HP70] in more general situations. We will just indicate the most important ideas of the proof.

Note that the result is true when $\hat{N} \equiv 0$ and $\delta = 0$.

We can use the coordinate spaces as coordinates in $T_x\mathcal{D}$. Denoting by C_x^σ the different coordinate spaces of the tangent bundle, we will write

$$
\begin{aligned}
T_x\mathcal{D} &= C_x^\theta \oplus C_x^I \oplus C_x^\delta \oplus C_x^s \oplus C_x^u \\
&= C_x^s \oplus C_x^c,
\end{aligned} \tag{188}
$$

where C_x^c denotes all the variables which are not s.

We obtain the stable bundle E_x^{s} as the graph of a linear map L_x from C_x^s to C_x^c.

Corresponding to splitting (188), we can express the matrix of the derivative of the map is expressed as

$$
DF(x) = \begin{pmatrix} (e^{A_s} + \mu_{ss}(x) & \mu_{sc}(x) \\ \mu_{cs}(x) & B + \mu_{cc}(x) \end{pmatrix},
$$

where all the μ's above indicate matrices that are $\mathcal{C}^{r-2}$ small if $\|N\|_{\mathcal{C}^{r-2}}$, ε are small.

The matrix B is readily available. It has diagonal elements which are 1 in the θ and I directions, e^{A_u} in the unstable direction. There are no non-diagonal elements. We note that the matrix B is invertible.

We can see $(\delta, L_x\delta) \in T_x\mathcal{D}$ gets mapped under $DF(x)$ into

$$
((e^{A_s} + \mu_{ss}(x) + \mu_{sc}(x)L_x)\delta, (\mu_{cs}(x) + BL_x)\delta.
$$

Therefore, the condition of invariance of the graph is

$$
A_{F(x)}(e^{A_s} + \mu_{ss}(x) + \mu_{sc}(x)L_x) = \mu_{cs}(x) + BL_x.
$$

Equivalently,

$$
L_x = B^{-1}(L_{F(x)}(e^{A_s} + \mu_{ss}(x) + \mu_{sc}(x)L_x) - \mu_{cs}(x)). \tag{189}
$$

We consider (189) as a fixed point equation. We will argue that the operator defined by the right-hand side on the space of C^{r-2} functions is a contraction in C^{r-2}.

We note that, taking derivatives of (189) we obtain that

$$
D_x^i B^{-1}(L_{f(x)}(e^{A_s} + \mu_{ss}(x) + \mu_{sc}(x)L_x) - \mu_{cs}(x)) = D_x^i L_{f(x)}(D^F(x))^{\otimes i}(e^{A_s}) + S_i,
$$

where we denote by f the map induced in the coordinates given the invariant manifold by the map F. where S_i is a polynomial involving derivatives of L up to order i and derivatives of F and of μ also to order i. All the terms in S contain at least a derivative of μ.

We note that the derivative of *Df* has norm as close sa desired to 1 since it is the restriction of F to the invariant manifold.

Hence, we can conclude that the RHS of (189) is a contraction operator in C^{r-2} provided that we make strong enough assumptions on $\|N\|_{C^{r-1}}$, ε.

B.6.2. The stable/unstable manifolds for a point

From the exponential convergence of the points in $W^s_{\tilde{\Lambda}}$ to $\tilde{\Lambda}$, it follows that the orbit of each point in $W^{\mathrm{s}}_{\tilde{\Lambda}}$ is asymptotic in the future to the orbit of a point in $\tilde{\Lambda}$. More precisely, $z \in W^{\mathrm{s}}_{\tilde{\Lambda}}$ implies that there exists x such that

$$\operatorname{dist}(F^n(x), F^n(z)) \leqslant C\tilde{\lambda}^n \quad n \geqslant 0. \tag{190}$$

We note that the point x is defined uniquely because

$$\operatorname{dist}(F^n(x), F^n(\tilde{x})) \leqslant C\tilde{\lambda}^n \quad n \geqslant 0, x, \tilde{x} \in \tilde{\Lambda} \Longrightarrow x = \tilde{x}. \tag{191}$$

We denote the set of points satisfying (190) by W^{s}_x. Clearly, the W^{s}_x constitutes a foliation of $W^{\mathrm{s}}_{\tilde{\Lambda}}$, since, (191) is an equivalence relation.

The following result is a particular case of the results in [Fen74,Fen77].

Proposition B.6. *With the notations above, the manifolds W^{s}_x are C^{r-1} manifolds. We have*

$$T_x W^{\mathrm{s}}_x = E^{\mathrm{s}}_x. \tag{192}$$

Moreover, there is a $\mathcal{C}^{r-1}$ mapping $\mathcal{S} : \tilde{\Lambda} \times B_\sigma$ in such a way that $\mathcal{S}(\{x\}, B_\sigma)$ is a local diffeomorphism into a neighborhood of x in W^{s}_x.

Remark B.7. In the general theory of normally hyperbolic manifolds, the manifolds W^{s}_x are as regular as the mapping, but the regularity of $\mathcal{S}$ is limited not only by the regularity of the map but also by ratios of exponents along the normal directions and along the manifold.

In our case, however, since the exponents along the manifold are as close to 1 as desired, we obtain that the only limit to the regularity of $\mathcal{S}$ is the regularity of the mapping F.

Besides the proofs in [Fen74,Fen77], a good exposition of the existence and regularity of W^{s}_x and their characterization by exponential rates, is in [KH95, p. 244 ff.].

To prove the existence of the manifold W^{s}_x, we construct it as the graph of a mapping s_x from a ball in E^{s}_x to its complementary space E^c_x. The fact that $x \in W^{\mathrm{s}}_x$ is equivalent to $s_x(0) = 0$. Eq. (192) is equivalent to $Ds_x(0) = 0$.

We introduce the notation that $x + v$ where $v \in T_x\mathcal{D}$ to mean the regular addition of the coordinates (of course in the angle coordinates, they are taken mod 1). We introduce the mapping $F_x : T_x\mathcal{D} \to T_{F(x)}\mathcal{D}$ by $F_x(v) \equiv F(x+v) - F(x)$. This construction is standard. See e.g. [HP70] and, even if now we have used the Euclidean structure of the coordinate space, analogous constructions can be carried out in any manifold using the exponential mapping from Riemannian geometry.

We furthermore introduce the notation $F_x = DF(x) + N_x$ (we suppress the dependence on ε from the notation for the sake of brevity). We also use $DF^{s,s}(x)$ and similar notations to denote the components of the matrix for $DF(x)$ along the splitting $T_x\mathcal{D} = E_x^s \oplus E_x^c$. Because the splitting is invariant we have $DF(x)^s c(x) = 0$, $DF^{cs}(x) = 0$.

We use N_x^s to denote the projections of N along the same splitting. All the mappings N_x, N_x^σ are $\mathcal{C}^{r-1}$ and that the $r-1$ derivatives of all of them have a uniform modulus of continuity in B_σ, which is also independent of x. $N_x(z)$, $N_x^\sigma(z)$ is $\mathcal{C}^{r-2}$ jointly in x, z with uniformly continuous $r-2$ derivatives. Similarly, $DF(x)$, $DF^{\sigma,\sigma'}(x)$ are $\mathcal{C}^2$ with uniformly continuous $r-2$ derivatives.

Under the assumption that $\|N\|_{\mathcal{C}^{r-1}}, \varepsilon$ are small enough, we obtain that $\sup_x \|N_x\|_{\mathcal{C}^{r-1}}$, $\sup_x \left\|N_x^\sigma\right\|_{\mathcal{C}^{r-1}}$, $\|N.(\cdot)\|_{\mathcal{C}^{r-1}}$, $\left\|N_\cdot^\sigma(\cdot)\right\|_{\mathcal{C}^{r-1}}$ are as small as desired. We also have that DF_x^{ss} is close to e^{A_s}, in particular is a contraction.

Proceeding in a way similar to the derivation of (189) in Section B.6.1, we obtain that a point $(t, s_x(t))$ in the graph of s_x is mapped by F_x to

$$(DF^{ss})_x t + N_x^s(y, s_x(t)), DF^{cc})_x w_s(t)t + N_x^c(y, s_x(t))) \tag{193}$$

Note that, under our assumptions that $DF^{ss}(x)$ is a contraction and that N^s is small we obtain that $DF^{ss})_x t + N_x^s(y, s_x(t)) \subset B_\sigma$, therefore, point (193) is in the graph of $s_{f(x)}$ if and only if we have

$$s_{f(x)}(DF^{ss})_x t + N_x^s(y, s_x(t))) = DF^{cc})_x t + N_x^c(y, s_x(t)). \tag{194}$$

Equivalently

$$s_x(t) = DF^c c(x)^{-1}[s_{f(x)}(DF^{ss})_x t + N_x^s(y, s_x(t))) - N_x^c(y, s_x(t))]. \tag{195}$$

Exactly the same analysis that we have performed before shows that the RHS of (195) defines an operator that sends a ball in the space of mappings which are C^{r-1} in t and C^{r-2} in t, x into itself. Moreover, it is a contraction in the $\mathcal{C}^0$ distance.

This shows that it has a fixed point which is $\mathcal{C}^{r-2+\mathrm{Lip}}$ in t and jointly $\mathcal{C}^{r-3+\mathrm{Lip}}$ in x, t.

The existence of derivatives of order $r-1$ in t and $r-2$ jointly in x, t as well as their uniform continuity are obtained, as before by examining the equation satisfied by the derivatives of highest order by differentiating (195).

B.7. Uniqueness of the manifold $\tilde{\Lambda}$. Invariance

Since we are applying the contraction mapping theorem, we obtain that the solutions of Eqs. (171) are unique among all the bounded solutions. Similarly, for the stable and unstable manifolds.

This gives that the invariant manifolds of the extended equations produced in Section B.2 are unique in a $\mathcal{C}^0$ neighborhood.

Since the extended equations agree with the original ones domain $E \geqslant E_0, s^2+u^2 \leqslant \delta$, we obtain that invariant manifold W invariant for the extended equations is locally invariant for the original equations.

Unfortunately, the extension process involves arbitrary choices and it is, in principle possible that the manifolds produced by two such extension processes are different. Indeed, in the general theory of normally hyperbolic manifolds it is easy to construct examples of systems with an infinitude of locally invariant manifolds.

When $\tilde{\Lambda}$ is only locally invariant, the notion of the stable manifolds to $\tilde{\Lambda}$ and that of stable manifolds to a point in $\tilde{\Lambda}$ become somewhat delicate. Note that, the characterization of the stable manifold of a point x by the asymptotic behavior after a large number of iterates becomes quite problematic if the orbit of a point steps out of $\tilde{\Lambda}$ after a finite number of iterates. Note that, if we wander out of $\tilde{\Lambda}$, then the orbit of x is determined by the extensions that we have chosen and, of course, the orbits that approach this extended orbit also depend on the choice of extensions.

In [Fen77] one can find a discussion of these issues in the general case.

B.7.1. Invariance in our case

However, in our case, one can do better than in the general theory. The key observation is that since the manifold is $2+d$ dimensional—the d corresponds to the phases of the quasi-periodic perturbation—the KAM tori, produced in Section 4.3.6 separate the space since they are $1+d$ dimensional. (Note that the proof of the KAM theorem only requires the local invariance since it only involves transformations in a very small neighborhood.)

Hence, all the locally invariant manifolds produced so far are, actually invariant, after perhaps, reducing them slightly so that the boundaries are KAM tori.

Since the locally manifolds are invariant, the cone argument presented in Proposition B.3 shows that they have to agree.

This makes the invariant manifolds independent of the extension. As a consequence, the characterization of stable manifolds by the asymptotic convergence, which was established for the extended system, becomes valid in the original system.

Note that, in contrast with the arguments of existence of the locally invariant manifold which only require that the map is $\mathcal{C}^1$, the arguments on uniqueness of the manifold require enough differentiability so that we can apply the KAM theorem A.26.

References

[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) (1963) 9–36.

[Arn64] V.I. Arnold, Instability of dynamical systems with several degrees of freedom, Sov. Math. Dokl. 5 (1964) 581–585.

[AA67] V.I. Arnold, A. Avez, Ergodic Problems of Classical Mechanics, Benjamin, New York, 1967.

[AKN88] V.I. Arnold, V.V. Kozlov, A.I. Neishtadt, Dynamical Systems III, Encyclopaedia Mathematical Science, vol. 3, Springer, Berlin, 1988.

[BLW96] A. Banyaga, R. de la Llave, C.E. Wayne, Cohomology equations near hyperbolic points and geometric versions of Sternberg linearization theorem, J. Geom. Anal. 6 (4) (1996) 613–649.

[Ban88] V. Bangert, Mather sets for twist maps and geodesics on tori, in: U. Kirchgraber, H.-O. Walter (Eds.), Dynamics Reported, vol. 1, Teubner, Stuttgart, 1988, pp. 1–56.

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

[BB02] M. Berti, P. Bolle, A functional analysis approach to Arnold diffusion, Ann. Inst. H. Poincaré Anal. Non Linéaire 19 (4) (2002) 395–450.

[Bes96] U. Bessi, An approach to Arnold's diffusion through the calculus of variations, Nonlinear Anal. 26 (6) (1996) 1115–1135.

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

[BT99] S. Bolotin, D. Treschev, Unbounded growth of energy in nonautonomous Hamiltonian systems, Nonlinearity 12 (2) (1999) 365–388.

[BHTB90] H.W. Broer, G.B. Huitema, F. Takens, B.L.J. Braaksma, Unfoldings and bifurcations of quasi-periodic tori, Mem. Amer. Math. Soc. 83 (421) (1990) viii+175.

[BOV97] H.W. Broer, H.M. Osinga, G. Vegter, Algorithms for computing normally hyperbolic invariant manifolds, Z. Angew. Math. Phys. 48 (3) (1997) 480–524.

[BW02] K. Burns, H. Weiss, Spheres with positive curvature and nearly dense orbits for the geodesic flow, Ergodic Theory Dyn. Systems 22 (2) (2002) 329–348.

[CG82] L. Chierchia, G. Gallavotti, Smooth prime integrals for quasi-integrable Hamiltonian systems, Nuovo Cimento B (11) 67 (2) (1982) 277–295.

[CG94] L. Chierchia, G. Gallavotti, Drift and diffusion in phase space, Ann. Inst. H. Poincaré Phys. Théor. 60 (1) (1994) 1–144.

[CG98] L. Chierchia, G. Gallavotti, Erratum: drift and diffusion in phase space, Ann. Inst. H. Poincaré Phys. Théor 68 (1998) 135.

[CQC03] J.Y. Chong-Qing Cheng, Existence of diffusion orbits in a priori unstable hamiltonian systems, MPARC # 03-360, 2003.

[CL55] E.A. Coddington, N. Levinson, Theory of Ordinary Differential Equations, McGraw-Hill, New York, Toronto, London, 1955.

[CBP02] G. Contreras-Barandiarán, G.P. Paternain, Genericity of geodesic flows with posi tive topological entropy on $\{S^2\}$, J. Differential Geom. 61 (1) (2002) 1–49.

[Cre97] J. Cresson, A λ-lemma for partially hyperbolic tori and the obstruction property, Lett. Math. Phys. 42 (4) (1997) 363–377.

[CG03] J. Cresson, C. Guillet, Periodic orbits and Arnold diffusion, Discrete Continuous Dyn. Systems 9 (2) (2003) 451–470.

[dlL01] R. de la Llave, A tutorial on KAM theory, in: A. Katok, R. de la Llave, Y. Pesin, H. Weiss (Eds.), Smooth Ergodic Theory and its Applications (Seattle, WA, 1999), Amer. Math. Soc., Providence, RI, 2001, pp. 175–292.

[dlL04] R. de la Llave, Improved estimates on measure occupied and the regularity of invariant tori in hamiltonian systems, preprint, 2004.

[dlLO99] R. de la Llave, R. Obaya, Regularity of the composition operator in spaces of Hólder functions, Discrete Continuous Dyn. Systems 5 (1) (1999) 157–184.

[LV01] R. de la Llave, J. Vano. A Whitney–Zehnder implicit function theorem, manuscript, 2001.

[dlLW04] R. de la Llave, C.E. Wayne, Whiskered and low dimensional tori in nearly integrable hamiltonian systems, Math. Phys. Electron. J. 10 (Paper 5) (2004) 45pp (electronic).

[DLS00] A. Delshams, R. de la Llave, 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) (2000) 353–392.

[DdlLMS03] A. Delshams, R. de la Llave, T.M.-Seara, A geometric mechanism for diffusion in hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model, MPARC # 03-182, Mem. Amer. Math. Soc. 2003, to appear.

[DdlLS03] A. Delshams, R. de la Llave, 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 (2003) 125–134 (electronic).

[DdlLMS04] A. Delshams, R. de la Llave, T.M. -Seara, Geometric properties of the scattering map to a normally hyperbolic manifold, preprint, 2004.

[DG00] A. Delshams, P. Gutiérrez, Splitting potential and Poincare–Melnikov method for whiskered tori in hamiltonian systems, J. Nonlinear Sci. 10 (4) (2000) 433–476.

[DR97] A. Delshams, R. Ramirez-Ros, Melnikov potential for exact symplectic maps, Comm. Math. Phys. 190 (1997) 213–245.

[Din71] E.I. Dinaburg, A connection between various entropy characterizations of dynamical systems, Izv. Akad. Nauk SSSR Ser. Mat. 35 (1971) 324–366.

[Don88] V.J. Donnay, Geodesic flow on the two-sphere. I. Positive measure entropy, Ergodic Theory Dyn. Systems 8 (4) (1988) 531–553.

[Don95] V.J. Donnay, Transverse homoclinic connections for geodesic flows, in: H.S. Dumas, K.R. Meyer, D.S. Schmidt (Eds.), Hamiltonian Dynamical Systems, Springer, New York, 1995, pp. 115–125.

[Eas78] R.W. Easton, Homoclinic phenomena in Hamiltonian systems with several degrees of freedom, J. Differential Equations 29 (2) (1978) 241–252.

[Eas89] R. Easton, Isolating blocks and epsilon chains for maps, Physica D 39 (1) (1989) 95–110.

[EMR01] R.W. Easton, J.D. Meiss, G. Roberts, Drift by coupling to an anti-integrable limit, Physica D 156 (3–4) (2001) 201–218.

[FL92] C. Falcolini, R. de la Llave, A rigorous partial justification of Greene's criterion, J. Statist. Phys. 67 (3-4) (1992) 609–643.

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

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

[Fen77] N. Fenichel, Asymptotic stability with rate conditions. II, Indiana Univ. Math. J. 26 (1) (1977) 81–93.

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

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

[Fur61] H. Furstenberg, Strict ergodicity and transformation of the torus, Amer. J. Math. 83 (1961) 573–601.

[Gal94] G. Gallavotti, Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review, Rev. Math. Phys. 6 (3) (1994) 343–411.

[Gal97] G. Gallavotti, Hamilton–Jacobi's equation and Arnold's diffusion near invariant tori in a priori unstable isochronous systems, Rend. Sem. Mat. Univ. Politec. Torino 55 (4) (1999) 291–302 (1997).

[Gal99] G. Gallavotti, Arnold's diffusion in isochronous systems, Math. Phys. Anal. Geom. 1 (4) (1998/99) 295–312.

[GGM99] G. Gallavotti, G. Gentile, V. Mastropietro, A field theory approach to Lindstedt series for hyperbolic tori in three time scales problems, J. Math. Phys. 40 (12) (1999) 6430–6472.

[GGM00] G. Gallavotti, G. Gentile, V. Mastropietro, Hamilton–Jacobi equation, heteroclinic chains and Arnol'd diffusion in three time scale systems, Nonlinearity 13 (2) (2000) 323–340.

[GR03] M. Gidea, C. Robinson, Topologically crossing heteroclinic connections to invariant tori, J. Differential Equations 193 (1) (2003) 49–74.

[GR04] M. Gidea, C. Robinson, Symbolic dynamics for transition tori-II, in: New Advances in Celestial Mechanics and Hamiltonian Systems, Kluwer/Plenum Press, New York, 2004, pp. 95–108.

[GJSM01a] G. Gómez, Á. Jorba, C. Simó, J. Masdemont, Advanced methods for collinear points, Dynamics and Mission Design Near Libration Points, vol. III, World Scientific Monograph Series in Mathematics, vol. 4, World Scientific Publishing, River Edge, NJ, 2001.

[GJSM01b] G. Gómez, Á. Jorba, C. Simó, J. Masdemont, Advanced methods for triangular points, Dynamics and Mission Design Near Libration Points, vol. IV, World Scientific Monograph Series in Mathematics, vol. 5, World Scientific Publishing, River Edge, NJ, 2001.

[GLMS01] G. Gómez, J. Llibre, R. Martínez, C. Simó, Fundamentals: the case of collinear libration points, Dynamics and Mission Design Near Libration Points, vol. I, World Scientific Monograph Series in Mathematics, vol. 2, World Scientific Publishing, River Edge, NJ, 2001.

[GSLM01] G. Gómez, C. Simó, J. Llibre, R. Martínez, Dynamics and mission design near libration poits, vol. I, Fundamentals: the case of triangular libration points, Dynamics and Mission Design Near Libration Points. vol. II, World Scientific Monograph Series in Mathematics, vol. 3, World Scientific Publishing, River Edge, NJ, 2001.

[Hed32] G.A. Hedlund, Geodesics on a two-dimensional Riemannian manifold with periodic coefficients, Ann. of Math. 33 (1932) 719–739.

[HP70] M.W. Hirsch, C.C. Pugh, Stable manifolds and hyperbolic sets, in: S. Chern, S. Smale (Eds.), Global Analysis; Proceedings of the Symposium on Pure Mathematics, vol. XIV, Berkeley, CA, 1968, Amer. Math. Soc., Providence, R.I., 1970, pp. 133–163.

[HPS77] M.W. Hirsch, C.C. Pugh, M. Shub, Invariant Manifolds, vol. 583, Lecture Notes in Mathematics, Springer, Berlin, 1977.

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

[IdlLNV02] A.P. Itin, R. de la Llave, A.I. Neishtadt, A.A. Vasiliev, Transport in a slowly perturbed convective cell flow, Chaos 12 (4) (2002) 1043–1053.

[Itu96] R. Iturriaga, Minimizing measures for time-dependent Lagrangians, Proc. London Math. Soc. (3) 73 (1) (1996) 216–240.

[ISM00] R. Iturriaga, H. Sánchez-Morgado, Finsler metrics and action potentials, Proc. Amer. Math. Soc. 128 (11) (2000) 3311–3316 (electronic).

[Kat80] A. Katok, Lyapunov exponents, entropy and periodic orbits for diffeomorphisms, Inst. Hautes Etudes Sci. Publ. Math. 51 (1980) 137–173.

[Kat82] A. Katok, Entropy and closed geodesics, Ergodic Theory Dyn. Systems 2 (3–4) (1983) 339–365 (1982).

[KH95] A. Katok, B. Hasselblatt, Introduction to the Modern Theory of Dynamical Systems, Cambridge University Press, Cambridge, 1995.

[Kli78] W. Klingenberg, Lectures on Closed Geodesics, Springer, Berlin, 1978; Grundlehren der Mathematischen Wissenschaften, vol. 230.

[Kli83] W. Klingenberg, Closed Geodesics on Riemannian Manifolds, CBMS Regional Conference Series in Mathematics, vol. 53, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1983.

[KW02] G. Knieper, H. Weiss, C^∞ genericity of positive topological entropy for geodesic flows on S^2, J. Differential Geom. 62 (1) (2002) 127–141.

[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 (1954) 527–530 (English translation in Stochastic Behavior in Classical and Quantum Hamiltonian Systems (Volta Memorial Conference Como, 1977), Lecture Notes in Physics, vol. 93, Springer, Berlin, 1979, pp 51–56.).

[LI73] O.E. Lanford III, Bifurcation of periodic solutions into invariant tori: the work of Ruelle and Takens, in: I. Stakgold, D.D. Joseph, D.H. Sattinger (Eds.), Nonlinear Problems in the Physical Sciences and Biology: Proceedings of a Battelle Summer Institute, Lecture Notes in Mathematics, vol. 322, Springer, Berlin, 1973, pp. 159–192.

[Lev97] M. Levi, Shadowing property of geodesics in Hedlund's metric, Ergodic Theory Dyn. Systems 17 (1) (1997) 187–203.

[Lla02] R. Llave, Orbits of unbounded energy in perturbations of geodesic flows by periodic potentials. A simple construction, preprint, 2002.

[LM88] P. Lochak, C. Meunier, Multiphase Averaging for Classical Systems, Applied Mathematical Science, vol. 72, Springer, New York, 1988.

[MS02] J.-P. Marco, D. Sauzin, Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems, Publ. Math. Inst. Hautes Études Sci. 96 (2003) 199–275 (2002).

[Mat2] J. Mather, Arnold diffusion I: announcement of results, 2–2, preprint.

[Mat96] J.N. Mather, Graduate Course at Princeton, 95–96, and Lectures at Penn State, Spring 96, Paris, Summer 96, Austin, Fall 96.

[Moe96] R. Moeckel, Transition tori in the five-body problem, J. Differential Equations 129 (2) (1996) 290–314.

[Moe02] R. Moeckel, Generic drift on Cantor sets of annuli, in: Celestial Mechanics (Evanston, IL, 1999), Amer. Math. Soc., Providence, RI, 2002, pp. 163–171.

[Mor24] M. Morse, A fundamental class of geodesics on any closed surface of genus greater than one, Trans. Amer. Math. Soc. 26 (1924) 26–60.

[Mos66a] J. Moser, A rapidly convergent iteration method and non-linear differential equations. II, Ann. Scuola Norm. Sup. Pisa (3) 20 (1966) 499–535.

[Mos66b] J. Moser, A rapidly convergent iteration method and non-linear partial differential equations I, Ann. Scuola Norm. Sup. Pisa (3) 20 (1966) 265–315.

[Nei81] A.I. Neishtadt, Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions, Prikl. Mat. Mekh. 45 (6) (1981) 1016–1025.

[New77] S.E. Newhouse, Quasi-elliptic periodic points in conservative dynamical systems, Amer. J. Math. 99 (5) (1977) 1061–1087.

[Pat99] G.P. Paternain, Geodesic Flows, Birkhäuser, Boston, MA, 1999.

[Pet96] D. Petroll, Existenz und Transversalitaet homoklinen und heteroklinen Orbits beim geodatischen Fluss, Technical Report, Univ. Freiburg, Math. Fak., 1996.

[Pol93] M. Pollicott, Lectures on Ergodic Theory and Pesin Theory on Compact Manifolds, London Mathematical Society Lecture Note Series, vol. 180, Cambridge University Press, Cambridge, 1993.

[Pos82] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, Comm. Pure Appl. Math. 35 (5) (1982) 653–696.

[Pos01] J. Pöschel, A lecture on the classical KAM theorem, in: Smooth Ergodic Theory and its Applications (Seattle, WA, 1999), Proceedings of the Symposium Pure Mathematics, vol. 69, Amer. Math. Soc., Providence, RI, 2001. pp. 707–732.

[RS02] P.H. Rabinowitz, E.W. Stredulinsky, A variational shadowing method, in: Celestial Mechanics (Evanston, IL, 1999), Contemporary Mathematics, vol. 292, Amer. Math. Soc., Providence, RI, 2002, pp. 185–197.

[Rob70a] R.C. Robinson, Generic properties of conservative systems, Amer. J. Math. 92 (1970) 562–603.

[Rob70b] R.C. Robinson, Generic properties of conservative systems. II, Amer. J. Math. 92 (1970) 897–906.

[Rob88] C. Robinson, Horseshoes for autonomous Hamiltonian systems using the Melnikov integral, Ergodic Theory Dyn. Systems 8 (1988) 395–409.

[Rue78] D. Ruelle, An inequality for the entropy of differentiable maps, Bol. Soc. Brasil. Mat. 9 (1) (1978) 83–87.

[Rus75] H. Rüssmann, On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus, in: Dynamical Systems, Theory and Applications (Battelle Recontres, Seattle, Washington, 1974), Lecture Notes in Physics, vol. 38, Springer, Berlin, 1975, pp. 598–624.

[Rus76a] H. Russmann, Note on sums containing small divisors, Comm. Pure Appl. Math. 29 (6) (1976) 755–758.

[Rus76b] H. Russmann, On optimal estimates for the solutions of linear difference equations on the circle, Celestial Mech. 14 (1) (1976) 33–37.

[Sal86] D. Salamon, The Kolmogorov-Arnold-Moser theorem, Math. Phys. Electron. J. 10 (2004) (Paper 3, 37pp).

[SGJM95] C. Simó, G. Gómez, A. Jorba, J. Masdemont, The bicircular model near the triangular libration points of the RTBP, in: From Newton to Chaos (Cortina d'Ampezzo, 1993), Plenum, New York, 1995, pp. 343–370.

[Sor02] A. Sorrentino, Sulle soluzioni quasi-periodiche di sistemi Hamiltoniani differenziabili, Univ. di Roma Tre, Tesi di Laurea, 2002.

[Ste70] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton Mathematical Series, vol. no. 30, Princeton University Press, Princeton, NJ, 1970.

[Sva80] N.V. Svanidze, Small perturbations of an integrable dynamical system with an integral invariant, Trudy Mat. Inst. Steklov. 147 (1980) 124–146, 204. (English translation: Proc. Steklov Inst. Math. (2) (1981)).

[Tak70] F. Takens, Hamiltonian systems: generic propeties of closed orbits and local perturbations, Math. Ann. 188 (1970) 304–312.

[Tre94] D.V. Treschev, Hyperbolic tori and asymptotic surfaces in Hamiltonian systems, Russian J. Math. Phys. 2 (1) (1994) 93–110.

[Tre02a] D. Treschev, Evolution of slow variables in near-integrable Hamiltonian systems, in: Progress in Nonlinear Science, vol. 1 (Nizhny Novgorod, 2001), RAS, Inst. Appl. Phys., Nizhniĭ Novgorod, 2002, pp. 166–169.

[Tre02b] D. Treschev, Multidimensional symplectic separatrix maps, J. Nonlinear Sci. 12 (1) (2002) 27–58.

[Tre02c] D. Treschev, Trajectories in a neighbourhood of asymptotic surfaces of a priori unstable Hamil tonian systems, Nonlinearity 15 (6) (2002) 2033–2052.

[Tre04] D. Treschev, Evolution of slow variables in a priori unstable Hamiltonian systems, Nonlinearity 17 (5) (2004) 1803–1841.

[Wei77] A. Weinstein, Lectures on Symplectic Manifolds, CBMS Regional Conference Series in Mathematics, vol. 29, Amer. Math. Soc., Providence, RI, 1977.

[Wig90] S. Wiggins, Global Bifurcations and Chaos: Analytical Methods, Applied Mathematical Science, vol. 73, Springer, New York, 1990.

[Xia98] Z. Xia, Arnold diffusion: a variational construction, in: Proceedings of the International Congress of Mathematicians, vol. II (Berlin, 1998), number extra vol. II, 1998, pp. 867–877 (electronic).

[Zeh73] E. Zehnder, Homoclinic points near elliptic fixed points, Comm. Pure Appl. Math. 26 (1973) 131–182.

[Zeh75] E. Zehnder, Generalized implicit function theorems with applications to some small divisor problems/I, Comm. Pure Appl. Math. 28 (1975) 91–140.

[Zeh76a] E. Zehnder, Generalized implicit function theorems with applications to some small divisor problems/II, Comm. Pure Appl. Math. 29 (1976) 49–111.

[Zeh76b] E. Zehnder, Moser's implicit function theorem in the framework of analytic smoothing, Math. Ann. 219 (2) (1976) 105–121.

[Zha00] V. Zharnitsky, Invariant curve theorem for quasiperiodic twist mappings and stability of motion in the Fermi–Ulam problem, Nonlinearity 13 (4) (2000) 1123–1136.