A functional analysis approach to Arnold Diffusion

Massimiliano Berti and Philippe Bolle

Abstract: We discuss in the context of nearly integrable Hamiltonian systems a functional analysis approach to the "splitting of separatrices" and to the "shadowing problem". As an application we apply our method to the problem of Arnold Diffusion for nearly integrable partially isochronous systems improving known results. [1]

Keywords: Arnold diffusion, shadowing theorem, splitting of separatrices, heteroclinic orbits, variational methods, nonlinear functional analysis

AMS subject classification: 34J40, 37J45.

Une approche de la diffusion d'Arnold par l'analyse fonctionnelle

Résumé: Nous discutons dans le contexte des systèmes hamiltoniens presque intégrables une approche basée sur l'analyse fonctionnelle des problèmes de "splitting" des séparatrices et de "shadowing". Nous appliquons notre méthode au problème de la diffusion d'Arnold pour des systèmes presque intégrables partiellement isochrones, améliorant des résultats connus.

1 Introduction

Topological instability of action variables in multidimensional nearly integrable Hamiltonian systems is known as Arnold Diffusion. This terminology was introduced by Chirikov in [17], years after Arnold discovered this phenomenon in his famous paper [3]. For autonomous Hamiltonian systems with two degrees of freedom KAM theory generically implies topological stability of the action variables (i.e. the time-evolution of the action variables for the perturbed system stay close to their initial values for all times). On the contrary, for systems with more than two degrees of freedom, outside a wide range of initial conditions (the so-called "Kolmogorov set" provided by KAM theory), the action variables may undergo a drift of order one in a very long, but finite time called the "diffusion time".

After thirty years from Arnold's seminal work [3], attention to Arnold diffusion has been renewed by [15], followed by several papers (see e.g. [14], [8], [24], [45], [28] and references therein).

The Hamiltonian models which are usually studied (as suggested by normal form theory near simple resonances) have the form

$$H(I,\varphi,p,q) = \frac{1}{2}I_1^2 + \omega \cdot I_2 + \frac{1}{2}p^2 + \varepsilon(\cos q - 1) + \varepsilon\mu f(I,\varphi,p,q) \tag{1.1}$$

where ε and μ are small parameters (the "natural" order for μ being ε^d for some positive d); (I_1, I_2, p) and (φ, q) are standard symplectic action-angle variables ($I_i \in \mathbf{R}^{n_i}$, $n_1 + n_2 = n$, $\varphi = (\varphi_1, \varphi_2) \in \mathbf{T}^n$, $(p,q) \in \mathbf{R} \times \mathbf{T}$, $\mathbf{T}$ being the standard torus $\mathbf{R}/2\pi\mathbf{Z}$). In Arnold's model $I_1, I_2 \in \mathbf{R}$, $\omega = 1$, $f(I,\varphi,p,q) = (\cos q - 1)(\sin\varphi_1 + \cos\varphi_2)$ and in [3] diffusion is proved for μ exponentially small w.r.t. $\sqrt{\varepsilon}$. Physically (1.1) describes a system of n_1 "rotators" and n_2 harmonic oscillators weakly coupled with a pendulum through a perturbation term.

The existence of Arnold diffusion is usually proved following the mechanism proposed in [3]. For $\mu = 0$, Hamiltonian H admits a continuous family of n-dimensional partially hyperbolic invariant tori $\mathcal{T}_I$ possessing stable and unstable manifolds $W_0^s(\mathcal{T}_I) = W_0^u(\mathcal{T}_I)$, called "whiskers" by Arnold. Arnold's mechanism is then based on the following three main steps.

[1]Supported by M.U.R.S.T. Variational Methods and Nonlinear Differential Equations.

(i) For $\mu \neq 0$ small enough, the perturbed stable and unstable whiskers $W^s_\mu(\mathcal{T}^\mu_I)$ and $W^u_\mu(\mathcal{T}^\mu_I)$ split and intersect transversally ("splitting of the whiskers");

(ii) Prove the existence of a chain of "transition" tori connected by heteroclinic orbits ("transition chain");

(iii) Prove the existence of an orbit, "shadowing" the transition chain, for which the action variables I undergo a variation of $O(1)$ in a certain time T_d called the *diffusion time*.

The shadowing problem (iii) has been extensively studied in the last years by geometrical (see e.g. [15], [37], [24], [18], [19], [20]) and by variational methods (see e.g. [8], [11], [12], [9]). More rich and older literature is available on the splitting problem see e.g. [34], [33], [23], [15], [29], [21], [25], [26], [30], [43], [41], [42], [31], [36], [38], [39] and references therein.

The aim of this paper is to provide a functional analysis approach apt to deal with Arnold diffusion, especially with "splitting" (i) and "shadowing" (iii) problems.

Rather than formulating our results in an abstract setting we shall illustrate the method (in a complete and self-contained way) on a relatively simple class of models, namely harmonic oscillators weakly coupled with a pendulum through purely spatial perturbations.

In this context defining a transition chain (ii) is a straighforward consequence of the "splitting of the whiskers" (i) since *all* the invariant tori are preserved by the perturbation, being just slightly deformed. This also happens, for the peculiar choice of the perturbation, in the non-isochronous system considered in [3]. Obviously this is not the case for general non-isochronous systems where the surviving perturbed tori are separated by the gaps appearing in KAM constructions, making the existence of chains of tori a more difficult matter, see [15]. We also refer to [45] for a somewhat different mechanism of diffusion where step (ii) is bypassed for systems with three degrees of freedom using Mather theory.

We now give a rough description of the **main results** of this paper. We consider nearly integrable partially *isochronous* Hamiltonian systems given by

$$\mathcal{H}_\mu = \omega \cdot I + \frac{p^2}{2} + (\cos q - 1) + \mu f(\varphi, q), \tag{1.2}$$

where $(\varphi, q) \in \mathbf{T}^n \times \mathbf{T}^1 := (\mathbf{R}^n / 2\pi \mathbf{Z}^n) \times (\mathbf{R}/2\pi\mathbf{Z})$ are the angle variables, $(I, p) \in \mathbf{R}^n \times \mathbf{R}^1$ are the action variables and $\mu \geq 0$ is a small real parameter. When $\mu = 0$ the energy $\omega_i I_i$ of each oscillator is a constant of the motion. The unperturbed Hamitonian possesses n-dimensional invariant tori $\mathcal{T}_{I_0} = \{(\varphi, I, q, p) \in \mathbf{T}^n \times \mathbf{R}^n \times \mathbf{T}^1 \times \mathbf{R}^1 \mid I = I_0, \ q = p = 0\}$ with stable and unstable manifolds $W^s(\mathcal{T}_{I_0}) = W^u(\mathcal{T}_{I_0}) = \{(\varphi, I, q, p) \in \mathbf{T}^n \times \mathbf{R}^n \times \mathbf{T}^1 \times \mathbf{R}^1 \mid I = I_0, \ p^2/2 + (\cos q - 1) = 0\}$. The problem of *Arnold diffusion* in this context is whether, for $\mu \neq 0$, there exist motions whose net effect is to transfer energy from one oscillator to the others. In order to exclude trivial drifts of the actions due to resonance phenomena, it is standard to assume a diophantine condition for the frequency vector ω. Precisely we will always suppose that ω is (γ, τ)-diophantine, i.e.

- ($H1$) $\exists \ \gamma > 0$, $\tau \geq n - 1$ such that $|\omega \cdot k| \geq \gamma / |k|^\tau$, $\forall k \in \mathbf{Z}^n, k \neq 0$.

In the present paper we first prove general shadowing theorems (*theorems 2.3 and 3.2*) which improve -for isochronous systems- known estimates on the diffusion time. Secondly we discuss a new method for the splitting of the whiskers, providing general estimates on the Fourier coefficients of some "splitting function" (see *theorem 4.2*). As applications we consider the following two cases

(a) the frequencies of the harmonic oscillators form a diophantine vector ω of order 1 ("a priori-unstable case");

(b) the frequencies of the harmonic oscillators form a diophantine vector $\omega_\varepsilon = (1/\sqrt{\varepsilon}, \beta \varepsilon^a)$ with $a \geq 0$, $\mu \varepsilon^{-3/2}$ small and the perturbation $f(\varphi, q) = (1 - \cos q) f(\varphi)$ ("three-time-scales problem" with perturbations preserving all the unperturbed invariant tori). This corresponds, after a time rescaling, to $\omega = (1, \beta \varepsilon^a \sqrt{\varepsilon})$ in (1.1).

Case (a) highlights the improvement of our estimates on diffusion times. In this case it is easy to show, using the classical Poincaré-Melnikov function

$$M(A) = \int_{\mathbf{R}} \Big[f(\omega t + A, 0) - f(\omega t + A, q_0(t)) \Big] \, dt,$$

$q_0(t) = 4 \arctan(\exp t)$ being the unperturbed separatrix of the pendulum, that the splitting of the whiskers is $O(\mu)$. Then our shadowing method yields (see for a more precise statement *theorems 2.4 and 3.3*)

Theorem 1.1 *Assume $(H1)$ and let M possess a proper minimum (or maximum) A_0. Then, for μ small enough, there exist orbits whose action variables undergo a drift of order one, with diffusion time $T_d = O((1/\mu)\log(1/\mu))$.*

Theorem 1.1 answers a question raised in [35] (sec.7) proving that, at least for isochronous systems, it is possible to reach the maximal speed of diffusion $\mu/|\log \mu|$; for related results compare also with [20] and [44]. We recall that the estimate on the diffusion time obtained in [15] is $T_d >> O(\exp(1/\mu))$ and that in [24] it is improved to be $T_d = O(\exp(1/\mu))$; recently in [9], by means of Mather theory, the estimate on the diffusion time has been improved to be $T_d = O(1/\mu^{2\tau+1})$; in [19] it is obtained via geometric methods that $T_d = O(1/\mu^{\tau+1})$. It is worth pointing out that the estimates given in [9] and [19], while providing a diffusion time polynomial in the splitting, depend on the diophantine exponent τ and hence on the number of rotators n. Instead our estimate (as well as that discussed in [20]) does not depend upon the number of degrees of freedom.

The three-time-scales system illustrates an application of our estimate of the splitting (*theorem 5.1*), which, together with our general shadowing theorem 2.3, yields Arnold diffusion in case (b). Roughly (see *theorem 5.2* for a precise statement) we get

Theorem 1.2 *Under a suitable non degeneracy condition on f, for all ε small such that ω_ε is $(\gamma_\varepsilon, \tau)$ diophantine, for $\mu\varepsilon^{-3/2}$ small enough, stable and unstable manifolds split and there exist orbits whose action variables undergo a drift of order one with diffusion time $T_d = O((\sqrt{\varepsilon}/\mu)e^{\pi/(2\sqrt{\varepsilon})}[(\gamma_\varepsilon)^{-1}(\sqrt{\varepsilon}e^{\pi/(2\sqrt{\varepsilon})})^\tau + |\log \mu|])$.*

The choice of three time scales frequencies enables to derive lower estimates for the splitting relatively easily. Incidentally we mention that theorem 5.1 improves the main theorem I in [38] which holds for $\mu = \varepsilon^p$, $p > 2 + a$; w.r.t. [25] (which deals with more general systems) we remark that our results hold in any dimension, while the results of [25], based on tree techniques and cancellations, are proved for $n = 2$.

We also mention that, for $n \geq 3$, with a more careful shadowing analysis (exploiting the anisotropy of this splitting) we can prove that, along special directions, Arnold diffusion takes place in polynomial time w.r.t $1/\varepsilon$ (see the forecoming paper [7] and remark 5.2).

Furthermore we also remark that our proof of theorem 3.2 is completely self-contained in the sense that, unlike the known approaches (excepted [45]), we do not make use of any KAM-type result for proving, under assumption $(H1)$, the persistence of invariant tori, see *theorem 3.1.*

The results of this paper have been announced in [6].

We now describe the **functional analysis approach** developed in this paper to prove both the results on the shadowing theorem and on the "splitting of the whiskers". It is based on a finite dimensional reduction of Lyapunov-Schmidt type, variational in nature, introduced in [2] and in [1], and later extended in [4]-[5] in order to construct shadowing orbits. For simplicity we describe our approach when the perturbation term $f(\varphi, q) = (1 - \cos q) f(\varphi)$ so that the tori $\mathcal{T}_{I_0}$ are still invariant for $\mu \neq 0$. The equations of motion derived by Hamiltonian $\mathcal{H}_\mu$ are

$$\dot{\varphi} = \omega, \qquad \dot{I} = -\mu(1 - \cos q) \, \nabla f(\varphi), \qquad \dot{q} = p, \qquad \dot{p} = \sin q - \mu \sin q \; f(\varphi). \tag{1.3}$$

The dynamics on the angles φ is given by $\varphi(t) = \omega t + A$ so that (1.3) are reduced to the quasi-periodically forced pendulum equation

$$-\ddot{q} + \sin q = \mu \; \sin q \; f(\omega t + A), \tag{1.4}$$

corresponding to the Lagrangian

$$\mathcal{L}_{\mu,A}(q,\dot{q},t)=\frac{\dot{q}^2}{2}+(1-\cos q)+\mu(\cos q-1)f(\omega t+A). \tag{1.5}$$

For each solution $q(t)$ of (1.4) one recovers the dynamics of the actions $I(t)$ by quadratures in (1.3).

For $\mu=0$ equation (1.4) is autonomous and possesses the one parameter family of homoclinic solutions (mod. 2π) $q_\theta(t)=4\arctan(\exp(t-\theta))$, $\theta\in\mathbf{R}$. Consider the Lagrangian action functional $\Phi_{\mu,A}$: $q_0+H^1(\mathbf{R})\to\mathbf{R}$ associated to the quasi-periodically forced pendulum (1.4)

$$\Phi_{\mu,A}(q):=\int_{\mathbf{R}}\mathcal{L}_{\mu,A}(q(t),\dot{q}(t),t)\,dt. \tag{1.6}$$

$\Phi_{\mu,A}$ is smooth on $q_0+H^1(\mathbf{R})$ and critical points q of $\Phi_{\mu,A}$ are homoclinic solutions to 0, mod. 2π, of (1.4). These critical points q are in fact smooth functions of the time t and are exponentially decaying to 0, mod. 2π, as $|t|\to+\infty$.

The unperturbed functional $\Phi_0:=\Phi_{0,A}$ does not depend on A and possesses the 1-dimensional manifold of critical points $Z:=\{q_\theta\mid\theta\in\mathbf{R}\}$ with tangent space at q_θ spanned by $\dot{q}_\theta$. All the unperturbed critical points q_θ are degenerate since $d^2\Phi_0(q_\theta)[\dot{q}_\theta]=0$. However q_θ are non-degenerate critical points of the restriction $\Phi_{0|E_\theta}$ for any subspace E_θ supplementary to $\langle\dot{q}_\theta\rangle$. It is then possible to apply a Lyapunov-Schmidt type reduction, based on the Implicit Function Theorem, to find near q_θ, for μ small, critical points $q^\mu_{A,\theta}$ of $\Phi_{\mu,A}$ restricted to E_θ; more precisely $q^\mu_{A,\theta}=q_\theta+w^\mu_{A,\theta}$ with $w^\mu_{A,\theta}\in E_\theta$, $||w^\mu_{A,\theta}||=O(\mu)$ and $d\Phi_{\mu,A}(q^\mu_{A,\theta})_{|E_\theta}=0$. We will call the functions $q^\mu_{A,\theta}$ "1-bump pseudo-homoclinic solutions" of the quasi-periodically forced pendulum (1.4).

It turns out that the 1-dimensional manifold $Z_\mu=\{q^\mu_{A,\theta}\mid\theta\in\mathbf{R}\}$ is a "natural constraint" for the action functional $\Phi_{\mu,A}$, namely any critical point of $\Phi_{\mu,A|Z_\mu}$ is a critical point of $\Phi_{\mu,A}$, and hence a true solution of (1.4) homoclinic to 0 (mod. 2π).

Through this paper we will perform the above finite dimensional reduction using two different supplementary spaces to $\langle\dot{q}_\theta\rangle$: one is better suited for the shadowing arguments, the other is better suited for studying the splitting problem in presence of "high frequencies".

Shadowing. In sections 2 and 3, dealing with the shadowing theorem, we choose

$$E_\theta=\Big\{w:\mathbf{R}\to\mathbf{R}\;\Big|\;w(\theta)=0\Big\}. \tag{1.7}$$

E_θ and $\langle\dot{q}_\theta\rangle$ are supplementary spaces since $\dot{q}_0(0)\neq0$. The corresponding "reduced action functional" $F_\mu(A,\theta):=\Phi_{\mu,A}(q^\mu_{A,\theta})$, see (2.8), turns out to have a neat geometric meaning, see remark 2.1, related to the generating functions of the exact Lagrangian stable and unstable manifolds $W^{s,u}(\mathcal{T}_{I_0})$.

¿From a geometrical point of view, the choice of the supplementary space defined by (1.7) means considering $W^s(\mathcal{T}_{I_0})$ and $W^u(\mathcal{T}_{I_0})$ at the fixed Poincaré section $\{q=\pi\}$, since $q^\mu_{A,\theta}(\theta)=q_\theta(\theta)+w^\mu_{A,\theta}(\theta)=q_0(0)=\pi$.

The supplementary space E_θ is very well suited to perform the shadowing theorem because the "1-bump pseudo solutions" $q^\mu_{A,\theta}(t)$ are true solutions of (1.4) except at the instant $t=\theta$ where $\dot{q}^\mu_{A,\theta}(t)$ may have a jump, see lemma 2.1. Following [4]-[5] we will generalize in subsection 2.2 the above finite dimensional reduction, finding a natural constraint for "k-bump pseudo homoclinic solutions" turning k times near the unperturbed separatrices of the pendulum, see lemma 2.4.

In this way under a suitable "splitting condition", satisfied for instance if $G_\mu(A):=F_\mu(A,0)$ possesses a proper minimum (see condition 2.1), we can prove a general shadowing theorem with explicit estimates on the diffusion time T_d, see theorem 2.3. This estimate (see expression (2.28)), is roughly the following: T_d is estimated by the product of the number of heteroclinic transitions k (= number of tori forming the transition chain = heteroclinic jump/splitting) and of the time T_s required for a single transition, namely $T_d=kT_s$. The time for a single transition T_s is bounded by the maximum time between the "ergodization time" of the torus $\mathbf{T}^n$ run by the linear flow ωt, and the time needed to "shadow" homoclinic orbits for the quasi-periodically forced pendulum.

Theorem 1.1 (for a priori unstable systems) is a straightforward consequence of the general shadowing theorem, noting that, by the classical Poincaré-Melnikov theory, the splitting of the whiskers is $O(\mu)$. The main reason for which theorem 1.1 improves the polynomial estimates $T_d = O(1/\mu^{2\tau+1})$ and $T_d = O(1/\mu^{\tau+1})$, obtained respectively in [9] and [19], is that our shadowing orbit can be chosen, at each transition, to approach the homoclinic point only up to a distance $O(1)$ and not $O(\mu)$ like in [9] and [19]. This implies that the time spent by our diffusion orbit at each transition is $T_s = O(\log(1/\mu))$. Since the number of tori forming the transition chain is equal to $O(1/\text{splitting}) = O(1/\mu)$ the diffusion time is finally estimated by $T_d = O((1/\mu)\log(1/\mu))$.

As well, theorem 1.2 (and theorem 5.2) for three-time-scales systems, is a consequence of the general shadowing theorem 2.3, but in this situation the splitting is by far less easy to measure, because it is exponentially small.

As mentioned above variational methods in the context of Arnold diffusion have been already used in [8], see remark 2.2 for comments on that. One possible advantage of our approach is that it may be used to consider more general critical points of the reduced functional, not only minima. Another advantage is that the same shadowing arguments can be used also when the hyperbolic part is a general Hamiltonian in $\mathbf{R}^{2m}$, $m \geq 1$, possessing one hyperbolic equilibrium and a transversal homoclinic orbit. Nevertheless we have developed all the details when the hyperbolic part is the standard one-dimensional pendulum because it is the model equation to study Arnold diffusion near a simple-resonance.

Splitting. Detecting and measuring the splitting of the whiskers is a difficult problem when the frequency vector $\omega = \omega_\varepsilon$ depends on some small parameter ε and contains some "fast frequencies" $\omega_i = O(1/\varepsilon^b)$, $b > 0$. Indeed, in this case, the variations of the Melnikov function along some directions turn out to be exponentially small with respect to ε and then the naive Poincaré-Melnikov expansion provides a valid measure of the splitting only for μ exponentially small with respect to some power of ε.

The typical argument to estimate exponentially small splittings, used virtually in all papers dealing with this problem, goes back to Arnold himself and is based on Fourier analysis on complex domains.

For this reason we would like to extend analytically the "reduced action functional" $F_\mu(A,\theta) = \Phi_{\mu,A}(q^\mu_{A,\theta})$ in a complex strip sufficiently wide in the θ variable. However $F_\mu(A,\theta)$ can not be easily analytically extended. Indeed, for θ complex, the supplementary space $E_\theta = \{w : \mathbf{R} \to \mathbf{C} \mid w(Re\,\theta) = 0\}$, appearing naturally when we try to extend the definition of $q^\mu_{A,\theta}$ to $\theta \in \mathbf{C}$, does not depend analytically on θ. This breakdown of analyticity, arising when measuring the "splitting of the whiskers" at the fixed Poincaré section $\{q = \pi\}$, is a well known difficulty and has been compensated in [15], [25], [26] via the introduction of tree techniques which enable to prove cancellations in the power series expansions.

Our method to overcome this "loss of analyticity" is different and relies on the introduction of another supplementary space $\widetilde{E}_\theta$, which depends analytically on θ. It is defined by

$$\widetilde{E}_\theta = \Big\{ w : \mathbf{R} \to \mathbf{R} \;\Big|\; \int_{\mathbf{R}} \psi_\theta(t) w(t)\, dt = 0 \Big\}$$

where $\psi_0(t) = \cosh^2(t)/(1 + \cosh t)^3$. $\widetilde{E}_\theta$ and $\langle \dot{q}_\theta \rangle$ are supplementary spaces since $\int_{\mathbf{R}} \psi_0(t) \dot{q}_0 dt \neq 0$ (see remark 4.1 for the choice of ψ_0). The corresponding reduced functional $\widetilde{F}_\mu(A,\theta) := \Phi_{\mu,A}(Q^\mu_{A,\theta})$, defined in (4.63), where $Q^\mu_{A,\theta}$ are the corresponding "1-bump pseudo-homoclinics solutions", can be analytically extended in a sufficiently large complex strip. This enables to find easily the exponentially small bounds on the Fourier coefficients for the splitting, thanks to the important invariance property $\widetilde{F}_\mu(A,\theta+\eta) = \widetilde{F}_\mu(A+\omega\eta,\theta)$, which is a consequence of the autonomy of Hamiltonian $\mathcal{H}_\mu$. This type of Lyapunov-Schmidt reduction was used in [2] to study the exponentially small splitting in rapidly periodically forced systems.

The crucial point is now to observe that "reduced action functionals" corresponding to different choices of the supplementary space are equivalent: it turns out (*theorem 4.1*) that the reduced functionals F_μ and $\widetilde{F}_\mu$ are simply the same up to a change of variables close to the identity,

$$F_\mu(A,\theta) = \widetilde{F}_\mu(A, \theta + h_\mu(A,\theta)), \qquad h_\mu = O(\mu).$$

This fact enables to transpose the informations on $\widetilde{F}_\mu$ to F_μ and viceversa. We point out that the proof of theorem 4.1 relies only on the standard Implicit Function Theorem. The introduction of $\widetilde{F}_\mu$ may be interpreted simply as measuring the splitting with a non constant Poincaré section.

A different approach for establishing exponentially small splittings for a rapidly forced pendulum has been used by [23] and recently extended to the setting considered here in [21] and [38]-[39] by the mean of normal form theory. For the more difficult case of the standard map, where the Melnikov method fails to give the correct prediction for the splitting, we refer to the pioneering paper [34] and to [30].

We also mention that the recent papers [43] and [36], even though quite different in spirit from ours, have also several technical similarities with our method (the use of generating functions, non fixed Poincaré section, ...). We add that these papers deal also with non-isochronous systems.

The paper is organized as follows: in section 2 we prove the shadowing theorem when the perturbation term is $f(\varphi, q) = (1 - \cos q) f(\varphi)$. In section 3 we show how to prove the shadowing theorem for general perturbation terms $f(\varphi, q)$. In section 4 we provide the theorem on the Fourier coefficients of the splitting and in section 5 we consider three-time-scales systems.

Acknowledgments: The authors thank Prof. L. Chierchia for numerous comments. The first author thanks Prof. G. Gallavotti for stimulating discussions. Much of this paper was written when the second author was visiting S.I.S.S.A. in Trieste. He thanks Prof. A. Ambrosetti and the staff of S.I.S.S.A. for their kind hospitality.

2 The shadowing theorem

2.1 1-bump homoclinic and heteroclinic solutions

Using the Implicit Function Theorem we now prove that, near the unperturbed homoclinic solutions $q_\theta(t)$, there exist, for μ small enough, *"1-bump pseudo-homoclinic solutions"* $q^\mu_{A,\theta}(t)$ of equation (1.4). $q^\mu_{A,\theta}(t)$ are true solutions of (1.4) in $(-\infty, \theta)$ and $(\theta, +\infty)$; at time $t = \theta$ such pseudo-solutions are glued with continuity at value $q^\mu_{A,\theta}(\theta) = \pi$ and for $t \to \pm\infty$ are asymptotic to the equilibrium 0 mod 2π. The $q^\mu_{A,\theta}$ are critical points for $\Phi_{\mu,A}$ restricted to $q_\theta + E_\theta$ (E_θ is defined in (1.7)).

Lemma 2.1 *There exist $\mu_0, C_0 > 0$ such that $\forall 0 < \mu \leq \mu_0$, $\forall \omega \in \mathbf{R}^n$, $\forall \theta \in \mathbf{R}$, there exists a unique function $q^\mu_{A,\theta}(t) : \mathbf{R} \to \mathbf{R}$, smooth in (A, μ), such that*

- *(i) $q^\mu_{A,\theta}(t)$ is a solution of (1.4) in each interval $(-\infty, \theta)$ and $(\theta, +\infty)$ and $q^\mu_{A,\theta}(\theta) = \pi$;*
- *(ii) $\max\Big(|q^\mu_{A,\theta}(t) - q_\theta(t)|, |\dot{q}^\mu_{A,\theta}(t) - \dot{q}_\theta(t)|\Big) \leq C_0 \mu \exp(-\frac{|t-\theta|}{2})$, $\forall t \in \mathbf{R}$;*
- *(iii) $q^\mu_{A,\theta}(t) = q^\mu_{A+k2\pi,\theta}(t)$, $\forall k \in \mathbf{Z}^n$;*
- *(iv) $q^\mu_{A,\theta+\eta}(t) = q^\mu_{A+\omega\eta,\theta}(t - \eta)$, $\forall \theta, \eta \in \mathbf{R}$.*
- *(v) $\max\Big(|\partial_A q^\mu_{A,\theta}(t)|, |\partial_A \dot{q}^\mu_{A,\theta}(t)|, |\omega \cdot \partial_A q^\mu_{A,\theta}(t)|, |\omega \cdot \partial_A \dot{q}^\mu_{A,\theta}(t)|\Big) \leq C_0 \mu \exp(-\frac{|t-\theta|}{2})$.*

PROOF. In the appendix. ■

We can then define the function $F_\mu : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$ as the action functional of Lagrangian (1.5) evaluated on the 1-dimensional manifold $Z_\mu := \{q^\mu_{A,\theta} \mid \theta \in \mathbf{R}\}$ of "1-bump pseudo-homoclinic solutions", namely

$$F_\mu(A, \theta) = \Phi_{\mu,A}(q^\mu_{A,\theta}) = \int_{-\infty}^{\theta} \mathcal{L}_{\mu,A}(q^\mu_{A,\theta}(t), \dot{q}^\mu_{A,\theta}(t), t)\, dt + \int_{\theta}^{+\infty} \mathcal{L}_{\mu,A}(q^\mu_{A,\theta}(t), \dot{q}^\mu_{A,\theta}(t), t)\, dt \qquad (2.8)$$

and the *"homoclinic function"* $G_\mu : \mathbf{T}^n \to \mathbf{R}$ as

$$G_\mu(A) = F_\mu(A, 0). \qquad (2.9)$$

Since $q^\mu_{A,\theta}(t)$ converges exponentially fast to 0, mod 2π, the integrals in (2.8) are convergent.

By property (iv) of lemma 2.1 the following invariance property holds

$$F_\mu(A,\theta+\eta) = F_\mu(A+\omega\eta,\theta), \quad \forall \theta,\eta \in \mathbf{R}, \tag{2.10}$$

and in particular

$$F_\mu(A,\theta) = G_\mu(A+\omega\theta), \quad \forall \theta \in \mathbf{R}. \tag{2.11}$$

Remark 2.1 *The homoclinic function G_μ is the difference between the generating functions $\mathcal{S}^{\pm}_{\mu,I_0}(A,q_0)$ of the stable and the unstable manifolds $W^{s,u}_\mu(\mathcal{T}_{I_0})$ (which in this case are exact Lagrangian manifolds) at section $q_0=\pi$, namely $G_\mu(A) = \mathcal{S}^-_{\mu,I_0}(A,\pi) - \mathcal{S}^+_{\mu,I_0}(A,\pi)$. Indeed it can be easily verified that*

$$\mathcal{S}^+_{\mu,I_0}(A,q_0) = I_0\cdot A - \int_0^{+\infty} \frac{(\dot q^\mu_{A,q_0}(t))^2}{2} + (1-\cos q^\mu_{A,q_0}(t)) + \mu(\cos q^\mu_{A,q_0}(t)-1)f(\omega t + A)\ dt,$$

where $q^\mu_{A,q_0}(t)$ is the unique solution of (1.4) near $q_0(t)$ with $q^\mu_{A,q_0}(0)=q_0$ and $\lim_{t\to+\infty} q^\mu_{A,q_0}(t) = 2\pi$. Analogously

$$\mathcal{S}^-_{\mu,I_0}(A,q_0) := I_0\cdot A + \int_{-\infty}^0 \frac{(\dot q^\mu_{A,q_0}(t))^2}{2} + (1-\cos q^\mu_{A,q_0}(t)) + \mu(\cos q^\mu_{A,q_0}(t)-1)f(\omega t + A)\ dt,$$

where $q^\mu_{A,q_0}(t)$ is the unique solution of (1.4) near $q_0(t)$ with $q^\mu_{A,q_0}(0)=q_0$ and $\lim_{t\to-\infty} q^\mu_{A,q_0}(t) = 0$.

There holds

Lemma 2.2 *$F_\mu : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$ is smooth and*

$$\partial_A F_\mu(A,\theta) = \int_{\mathbf{R}} \mu(\cos q^\mu_{A,\theta}(t) - 1)\partial_\varphi f(\omega t + A)\ dt, \tag{2.12}$$

$$\partial_\theta F_\mu(A,\theta) = \mu \int_{\mathbf{R}} f(\omega t + A)\sin q^\mu_{A,\theta}(t)\dot q^\mu_{A,\theta}(t)\ dt = \frac{(\dot q^\mu_{A,\theta})^2(\theta^+)}{2} - \frac{(\dot q^\mu_{A,\theta})^2(\theta^-)}{2}. \tag{2.13}$$

PROOF. Since $q^\mu_{A,\theta}$ depends smoothly on A so does $F_\mu(A,\theta)$ and there holds

$$\begin{aligned} \partial_A F_\mu(A,\theta) &= \int_{-\infty}^{+\infty} \dot q^\mu_{A,\theta}(t)\partial_A \dot q^\mu_{A,\theta}(t) + \sin q^\mu_{A,\theta}(t)\partial_A q^\mu_{A,\theta}(t) \\ &- \mu \sin q^\mu_{A,\theta}(t) f(\omega t + A)\partial_A q^\mu_{A,\theta}(t) + \mu(\cos q^\mu_{A,\theta}(t) - 1)\partial_\varphi f(\omega t + A)\ dt. \end{aligned}$$

Integrating by parts and using that $\lim_{t\to\pm\infty} \dot q^\mu_{A,\theta}(t) = 0$ and $\partial_A q^\mu_{A,\theta}(\theta) = 0$, we get

$$\begin{aligned} \int_{-\infty}^{+\infty} \dot q^\mu_{A,\theta}(t)\partial_A \dot q^\mu_{A,\theta}(t)\ dt &= \int_{-\infty}^{\theta} \ldots + \int_\theta^{+\infty} \dot q^\mu_{A,\theta}(t)\partial_A \dot q^\mu_{A,\theta}(t)\ dt \\ &= -\int_{-\infty}^{\theta} \ddot q^\mu_{A,\theta}(t)\partial_A q^\mu_{A,\theta}(t)\ dt - \int_\theta^{+\infty} \ddot q^\mu_{A,\theta}(t)\partial_A q^\mu_{A,\theta}(t)\ dt. \end{aligned}$$

Hence, since $q^\mu_{A,\theta}(t)$ solves (1.4) on each interval $(-\infty,\theta)$ and $(\theta,+\infty)$,

$$\partial_A F_\mu(A,\theta) = \int_{-\infty}^{+\infty} \mu(\cos q^\mu_{A,\theta}(t) - 1)\partial_\varphi f(\omega t + A)\ dt \tag{2.14}$$

and (2.12) is proved.

$\partial_\theta q^\mu_{A,\theta}$ is not defined in $W^{1,\infty}(\mathbf{R})$ because, $\dot q^\mu_{A,\theta}$ may have a jump at $t=\theta$. However, by the invariance property (2.10), $F_\mu(A,\theta)$ is smooth also in θ and we have that

$$\begin{aligned}\partial_\theta F_\mu(A,\theta) &= \omega\cdot\partial_A F_\mu(A,\theta)\\ &= \mu\int_{\mathbf{R}}(\cos q^\mu_{A,\theta}(t)-1)\frac{d}{dt}f(\omega t+A)\ dt\\ &= \mu\int_{\mathbf{R}}\sin q^\mu_{A,\theta}(t)\dot q^\mu_{A,\theta}(t)f(\omega t+A)\ dt,\end{aligned}$$

by an integration by parts. Now since $q^\mu_{A,\theta}(t)$ solves (1.4) in $(\theta,+\infty)$

$$\begin{aligned}\int_\theta^{+\infty}\mu\sin q^\mu_{A,\theta}(t)\dot q^\mu_{A,\theta}(t)f(\omega t+A)\ dt &= \int_\theta^{+\infty}-\ddot q^\mu_{A,\theta}\dot q^\mu_{A,\theta}+\sin q^\mu_{A,\theta}(t)\dot q^\mu_{A,\theta}(t)\ dt\\ &= \frac{(\dot q^\mu_{A,\theta})^2(\theta^+)}{2}+\cos q^\mu_{A,\theta}(\theta).\end{aligned}$$

In the same way

$$\int_\infty^\theta\mu\sin q^\mu_{A,\theta}(t)\dot q^\mu_{A,\theta}(t)f(\omega t+A)\ dt=-\frac{(\dot q^\mu_{A,\theta})^2(\theta^-)}{2}-\cos q^\mu_{A,\theta}(\theta).$$

This proves the lemma. ■

It follows that $Z_\mu:=\{q^\mu_{A,\theta}\mid\theta\in\mathbf{R}\}$ is a "natural constraint" for $\Phi_{\mu,A}$. Indeed, by (2.13) and lemma 2.1-(ii), if $\partial_\theta F_\mu(A,\theta)=0$ then $q^\mu_{A,\theta}(t)$ is a true homoclinic (mod. 2π) solution of (1.4). Then, for each $I_0\in\mathbf{R}^n$,

$$\Big(\omega t+A,I_\mu(t),q^\mu_{A,\theta}(t),\dot q^\mu_{A,\theta}(t)\Big)\tag{2.15}$$

where

$$I_\mu(t)=I_0-\mu\int_{-\infty}^t(1-\cos q^\mu_{A,\theta}(s))\partial_\varphi f(\omega s+A)ds\tag{2.16}$$

is a solution of $\mathcal{H}_\mu$ emanating at $t=-\infty$ from torus $\mathcal{T}_{I_0}$. Since $q^\mu_{A,\theta}$ converges exponentially fast to the equilibrium, the "jump" in the action variables $I_\mu(+\infty)-I_0$ is finite. We shall speak of homoclinic orbit to the torus $\mathcal{T}_{I_0}$ when the jump is zero, and of heteroclinic orbit from $\mathcal{T}_{I_0}$ to $\mathcal{T}_{I_\mu(+\infty)}$ when the jump is not zero. As a direct consequence of (2.12) we have that such a jump is given by $\partial_A F_\mu(A,\theta)$:

Lemma 2.3 *Let $\partial_\theta F_\mu(A,\theta)=0$ then $I_\mu(t)$ given in (2.16) satisfies*

$$\partial_A F_\mu(A,\theta)=\int_{-\infty}^{+\infty}\dot I_\mu(t)\ dt=I_\mu(+\infty)-I_0\in\mathbf{R}^n.\tag{2.17}$$

In particular if (A,θ) is a critical point of $F_\mu(A,\theta)$ then (2.15) defines a homoclinic orbit to torus $\mathcal{T}_{I_0}$.

By the invariance property (2.11) if B is a critical point of the homoclinic function G_μ, then, for all (A,θ) such that $A+\omega\theta=B$, (2.15) provides, for each I_0, a homoclinic solution to the torus $\mathcal{T}_{I_0}$. These homoclinic orbits are not geometrically distinct since, by the autonomy of $\mathcal{H}_\mu$, they are all obtained by time translation of the same homoclinic orbit. By the Ljusternik-Schnirelman. category theory, since cat $\mathbf{T}^n=n+1$, the function $G_\mu:\mathbf{T}^n\to\mathbf{R}$ has at least $n+1$ distinct critical points. This proves (see also [36])

Theorem 2.1 *Let $0<\mu\le\mu_0$. $\forall I_0\in\mathbf{R}^n$ there exist at least $n+1$ geometrically distinct homoclinic orbits to $\mathcal{T}_{I_0}$.*

¿From the conservation of energy a heteroclinic orbit between $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I_0'}$, if any, must satisfy the energy relation

$$\omega \cdot I_0 = \omega \cdot I_0'. \tag{2.18}$$

By lemma 2.3 a critical point of $F_{\mu,I_0,I_0'}(A,\theta)$, defined by $F_{\mu,I_0,I_0'}(A,\theta) = F_\mu(A,\theta) - (I_0' - I_0) \cdot A = G_\mu(A+\omega\theta) - (I_0' - I_0) \cdot A$, gives rise to a heteroclinic solution joining the tori $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I_0'}$. If the energy condition (2.18) holds then the function $F_{\mu,I_0,I_0'}(A,\theta)$ satisfies the invariance property

$$F_{\mu,I_0,I_0'}(A,\theta) = G_\mu(A+\omega\theta) - (I_0' - I_0) \cdot (A+\omega\theta) = G_{\mu,I_0,I_0'}(A+\omega\theta). \tag{2.19}$$

where

$$G_{\mu,I_0,I_0'}(B) := G_\mu(B) - (I_0' - I_0) \cdot B. \tag{2.20}$$

Note that $G_{\mu,I_0,I_0'}$ is not $2\pi\mathbf{Z}^n$-periodic, and it might possess no critical point even for $|I_0' - I_0|$ small. However near a homoclinic orbit to $\mathcal{T}_{I_0}$ satisfying some "transversality condition" there exist heteroclinic solutions connecting nearby tori $\mathcal{T}_{I_0'}$. As an example, the following theorem holds, where $B_\rho(A_0)$ denotes the open ball of radius ρ centered at A_0 in $\mathbf{R}^n$ (covering space of $\mathbf{T}^n$).

Theorem 2.2 *Assume that there exist* $A_0 \in \mathbf{T}^n$, $\delta > 0$ *and* $\rho > 0$ *such that* $\inf_{\partial B_\rho(A_0)} G_\mu \geq \inf_{B_\rho(A_0)} G_\mu + \delta$. *Then for all* $I_0, I_0' \in \mathbf{R}^n$ *satisfying* $(I_0 - I_0') \cdot \omega = 0$ *and* $|I_0 - I_0'| < \delta/(2\rho)$ *there exists a heteroclinic solution of* $\mathcal{H}_\mu$ *connecting* $\mathcal{T}_{I_0}$ *to* $\mathcal{T}_{I_0'}$.

Proof. Let us consider $\mathcal{G}_\mu(A) := G_\mu(A) - (I_0' - I_0) \cdot (A - A_0)$ which differs from $G_{\mu,I_0,I_0'}$, defined in (2.20), only by a constant. $\mathcal{G}$ attains an absolute minimum in $\overline{B_\rho(A_0)}$ at some point $\overline{A}$. The theorem is proved if we show that $\overline{A} \in B_\rho(A_0)$. Arguing by contradiction we assume that $\overline{A} \in \partial B_\rho(A_0)$. Denote by $\widetilde{A} \in B_\rho(A_0)$ a minimum point for G_μ. Then $\mathcal{G}_\mu(\overline{A}) \geq G_\mu(\overline{A}) - |I_0' - I_0|\rho > \min_{B_\rho(A_0)} G_\mu + \delta/2 = G_\mu(\widetilde{A}) + \delta/2$. However $\mathcal{G}_\mu(\widetilde{A}) \leq G_\mu(\widetilde{A}) + \delta/2$ and we get a contradiction. ∎

2.2 The k-bump pseudo-homoclinic solutions

We prove in the next lemma the existence of "k-bump pseudo-homoclinic solutions" $q^L_{A,\theta}(t)$ of the quasi-periodically forced pendulum (1.4) which turn k times along the separatrices and are asymptotic to the equilibrium for $t \to \pm\infty$. Such pseudo-homoclinics $q^L_{A,\theta}(t)$ are found, via the Contraction Mapping Theorem, as small perturbations of a chain of "1-bump pseudo-homoclinic solutions" obtained in lemma 2.1.

Lemma 2.4 *There exist* $C_1, L_1 > 0$ *such that* $\forall \omega \in \mathbf{R}^n$, $\forall 0 < \mu \leq \mu_0$, $\forall k \in \mathbf{N}$, $\forall L > L_1$, $\forall \theta = (\theta_1, \ldots, \theta_k) \in \mathbf{R}^k$ *such that* $\min_i(\theta_{i+1} - \theta_i) > L$, *there exists a unique pseudo-homoclinic solution* $q^L_{A,\theta}(t) : \mathbf{R} \to \mathbf{R}$, *smooth in* (A,θ,μ) *which is a true solution of (1.4) in each interval* $(-\infty, \theta_1)$, (θ_i, θ_{i+1}) $(i = 1, \ldots, k-1)$, $(\theta_k, +\infty)$ *and*

- *(i)* $q^L_{A,\theta}(\theta_i) = \pi(2i-1)$, $q^L_{A,\theta}(t) = q^\mu_{A,\theta_1}(t)$ *in* $(-\infty, \theta_1)$ *and* $q^L_{A,\theta}(t) = 2\pi(k-1) + q^\mu_{A,\theta_k}(t)$ *in* $(\theta_k, +\infty)$;
- *(ii)* $||q^L_{A,\theta} - q^\mu_{A,\theta_i}||_{W^{1,\infty}(J_i)} \leq C_1 \exp(-C_1 L)$ where $J_i = (\theta_i, (\theta_i + \theta_{i+1})/2)$, $\forall\, i = 1, \ldots, k-1$;
- *(iii)* $||q^L_{A,\theta} - q^\mu_{A,\theta_{i+1}}||_{W^{1,\infty}(J_i')} \leq C_1 \exp(-C_1 L)$ where $J_i' = ((\theta_i + \theta_{i+1})/2, \theta_{i+1})$, $\forall\, i = 1, \ldots, k-1$;
- *(iv)* $q^L_{A,\theta}(t) = q^L_{A+k2\pi,\theta}(t)$, $\forall k \in \mathbf{Z}^n$;
- *(v)* $q^L_{A,\theta+\eta e_k}(t) = q^L_{A+\omega\eta,\theta}(t-\eta)$, $\forall \eta \in \mathbf{R}, \theta \in \mathbf{R}^k$, *where* $e_k = (1, \ldots, 1) \in \mathbf{R}^k$.

Proof. In the appendix. ∎

Remark 2.2 *The "k-bump pseudo-homoclinic solutions" $q^L_{A,\theta}$ of lemma 2.4 could also be obtained minimizing the Lagrangian action functional with fixed end-points as in [8] using the directs method of the calculus of variations. For general systems, where the homoclinic is not a minimum for the action functional, as for the Duffing equation, this approach can not be applied. Moreover we prefer to use the Contraction Mapping Theorem since it also provides immediately the approximation estimates of lemma 2.4-$(ii)-(iii)$.*

We consider the Lagrangian action functional evaluated on the pseudo-homoclinic solutions $q^L_{A,\theta}$ given by lemma 2.4 depending on $n+k$ variables

$$F^k_\mu(A_1,\dots,A_n,\theta_1,\dots,\theta_k) = \int_{-\infty}^{\theta_1} \mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t)\ dt +$$

$$\sum_{i=1}^{k-1}\int_{\theta_i}^{\theta_{i+1}} \mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t)\ dt + \int_{\theta_k}^{+\infty} \mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t)\ dt.$$

By lemma 2.4-(v) the following invariance property holds

$$F^k_\mu(A,\theta+\eta e_k) = F^k_\mu(A+\eta\omega,\theta), \qquad \forall\eta\in\mathbf{R},\theta\in\mathbf{R}^k. \tag{2.21}$$

Let $\mathcal{F}^k_\mu : \mathbf{T}^n\times\mathbf{R}^k\to\mathbf{R}$ be the "k-bump heteroclinic function" defined by

$$\mathcal{F}^k_\mu(A,\theta) := F^k_\mu(A,\theta) - (I'_0 - I_0)\cdot A. \tag{2.22}$$

It turns out that the k-dimensional manifold $Z_k := \{q^L_{A,\theta} \mid \theta\in\mathbf{R}^k\}$ of "k-bump pseudo-homoclinic solutions" of (1.4) is a natural constraint for $\Phi_{\mu,A}$, see [4]. Arguing as in lemma 2.2 we have

Lemma 2.5 *$\forall I_0, I'_0\in\mathbf{R}^n$, if (A,θ) is a critical point of $\mathcal{F}^k_\mu(A,\theta)$, then $(\omega t + A, I_\mu(t),\ q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t))$ where $I_\mu(t) = I_0 - \mu\int_{-\infty}^t (1-\cos q^L_{A,\theta}(s))\partial_\varphi f(\omega s + A)ds$ is a heteroclinic solution connecting $\mathcal{T}_{I_0}$ to $\mathcal{T}_{I'_0}$.*

By lemma 2.5 we need to find critical points of $\mathcal{F}^k_\mu(A,\theta)$. When $\min_i(\theta_{i+1}-\theta_i)\to+\infty$ the "k-bump homoclinic function" $F^k_\mu(A,\theta)$ is well approximated simply by the sum of $F_\mu(A,\theta_i)$ according to the following lemma. We set $\theta_0=-\infty$ and $\theta_{k+1}=+\infty$.

Lemma 2.6 *There exist $C_2, L_2 > 0$ such that $\forall\omega\in\mathbf{R}^n$, $\forall 0<\mu\le\mu_0$, $\forall L>L_2$, $\forall\theta_1<\ldots<\theta_k$ with $\min_i(\theta_{i+1}-\theta_i) > L$*

$$F^k_\mu(A,\theta_1,\cdots,\theta_k) = \sum_{i=1}^k F_\mu(A,\theta_i) + \sum_{i=1}^k R_i(\mu,A,\theta_{i-1},\theta_i,\theta_{i+1}), \tag{2.23}$$

with

$$|R_i(\mu,A,\theta_{i-1},\theta_i,\theta_{i+1})| \le C_2\exp(-C_2 L).$$

PROOF. We can write

$$\begin{aligned} F^k_\mu(A,\theta_1,\dots,\theta_k) \;&=\; \Big(\int_{-\infty}^{\theta_1}\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t) + \int_{\theta_1}^{(\theta_1+\theta_2)/2}\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t)\Big)\\ &+\; \sum_{i=2}^{k-1}\Big(\int_{(\theta_{i-1}+\theta_i)/2}^{\theta_i}\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t) + \int_{\theta_i}^{(\theta_i+\theta_{i+1})/2}\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t)\Big)\\ &+\; \Big(\int_{(\theta_{k-1}+\theta_k)/2}^{\theta_k}\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t) + \int_{\theta_k}^{+\infty}\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t)\Big). \end{aligned}$$

We define

$$R_i^-(\mu,A,\theta_{i-1},\theta_i) = \int_{(\theta_{i-1}+\theta_i)/2}^{\theta_i}\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t),\dot q^L_{A,\theta}(t),t)\ dt - \int_{-\infty}^{\theta_i}\mathcal{L}_{\mu,A}(q^\mu_{A,\theta_i}(t),\dot q^\mu_{A,\theta_i}(t),t)\ dt,$$

$$R_i^+(\mu, A, \theta_i, \theta_{i+1}) = \int_{\theta_i}^{(\theta_i+\theta_{i+1})/2} \mathcal{L}_{\mu,A}(q^L_{A,\theta}(t), \dot q^L_{A,\theta}(t), t)\ dt - \int_{\theta_i}^{+\infty} \mathcal{L}_{\mu,A}(q^\mu_{A,\theta_i}(t), \dot q^\mu_{A,\theta_i}(t), t)\ dt$$

where q^μ_{A,θ_i} is the 1-bump pseudo-homoclinic solution obtained in lemma 2.1. Recalling the definition 2.8 of $F_\mu(A, \theta)$ we have

$$\begin{aligned} F^k_\mu(A, \theta_1, \dots, \theta_k) &= F_\mu(A, \theta_1) + R_1^+(\mu, A, \theta_1, \theta_2) \\ &+ \sum_{i=2}^{k-1} F_\mu(A, \theta_i) + \Big(R_i^-(\mu, A, \theta_{i-1}, \theta_i) + R_i^+(\mu, A, \theta_i, \theta_{i+1})\Big) \\ &+ F_\mu(A, \theta_k) + R_k^-(\mu, A, \theta_{k-1}, \theta_k). \end{aligned}$$

Setting $R_1 = R_1^+$, $R_i = R_i^- + R_i^+$ $(i = 2, \dots, k-1)$ and $R_k = R_k^-$ we derive the expression (2.23). In order to complete the proof, we have to show the existence of $C_2, L_2 > 0$ such that $\forall \omega \in \mathbf{R}^n$, for all $0 < \mu \le \mu_0$, $\forall L > L_2$, $\forall \theta_1 < \ldots < \theta_k$ with $\min_i(\theta_{i+1} - \theta_i) > L$, for all $i = 1, \dots, k$

$$|R_i^\pm(\mu, A, \theta_i, \theta_{i+1})| \le C_2 \exp(-C_2 L). \tag{2.24}$$

We write the proof for R_i^+. We have

$$\begin{aligned} R_i^+(\mu, A, \theta_i, \theta_{i+1}) &= \int_{\theta_i}^{(\theta_i+\theta_{i+1})/2} (\mathcal{L}_{\mu,A}(q^L_{A,\theta}(t), \dot q^L_{A,\theta}(t), t) - \mathcal{L}_{\mu,A}(q^\mu_{A,\theta_i}(t), \dot q^\mu_{A,\theta_i}(t), t)\ dt \\ &- \int_{(\theta_i+\theta_{i+1})/2}^{+\infty} \mathcal{L}_{\mu,A}(q^\mu_{A,\theta_i}(t), \dot q^\mu_{A,\theta_i}(t), t)\ dt. \end{aligned} \tag{2.25}$$

By lemma 2.1-(ii) the homoclinic orbit satisfies $\max(|q^\mu_{A,\theta_i}(t)|, |\dot q^\mu_{A,\theta_i}(t)|) \le C\exp(-|t - \theta_i|/2)$. Hence $\mathcal{L}_{\mu,A}(q^\mu_{A,\theta_i}(t), \dot q^\mu_{A,\theta_i}(t), t) = O(\exp(-|t - \theta_i|))$ and for all $\theta_1 < \ldots < \theta_k$ with $\min_i(\theta_{i+1} - \theta_i) > L$,

$$\Big| \int_{(\theta_i+\theta_{i+1})/2}^{+\infty} \mathcal{L}_{\mu,A}(q^\mu_{A,\theta_i}(t), \dot q^\mu_{A,\theta_i}(t), t)\ dt \Big| = O(e^{-L/2}). \tag{2.26}$$

¿From lemma 2.4-(ii) we also deduce that

$$\Big(\int_{\theta_i}^{(\theta_i+\theta_{i+1})/2} \mathcal{L}_{\mu,A}(q^L_{A,\theta}(t), \dot q^L_{A,\theta}(t), t) - \mathcal{L}_{\mu,A}(q^\mu_{A,\theta_i}(t), \dot q^\mu_{A,\theta_i}(t), t)\ dt \Big) = O(e^{-CL}). \tag{2.27}$$

for some constant $C > 0$. ¿From (2.25), (2.26) and (2.27) we deduce (2.24). ■

2.3 The diffusion orbit

We now give an example of condition on G_μ which implies the existence of diffusion orbits. $B_\alpha(A_0) \subset \mathbf{R}^n$ denotes the open ball centered at $A_0 \in \mathbf{R}^n$ and of radius α.

Condition 2.1 *("Splitting condition") There exist $A_0 \in \mathbf{T}^n$, $\alpha > 0$, a bounded open set $U \subset \mathbf{R}^n$ (the covering space of $\mathbf{T}^n$) such that $B_\alpha(A_0) \subset U$ and a positive constant $\delta > 0$ such that*

- *(i) $\inf_{\partial U} G_\mu \ge \inf_U G_\mu + \delta$;*
- *(ii) $\sup_{B_\alpha(A_0)} G_\mu \le \frac{\delta}{4} + \inf_U G_\mu$;*
- *(iii) $d(\{A \in U \mid G_\mu(A) \le \delta/2 + \inf_U G_\mu\}, \{A \in U \mid G_\mu(A) \ge 3\delta/4 + \inf_U G_\mu\}) \ge 2\alpha$.*

Remark 2.3 *If G_μ possesses a non-degenerate minimum in A_0 the "splitting condition" 2.1 is satisfied, for example, choosing $U = B_\rho(A_0)$ with $\rho = \min \lambda_i/(3 \sup_{A\in\mathbf{T}^n} |D^3 G_\mu|)$, $\delta = 4(\min \lambda_i)\rho^2/9$ and $\alpha = \rho((\sqrt{6} - \sqrt{5})/2\sqrt{10})\sqrt{(\min_i \lambda_i)/(\max_i \lambda_i)}$ where λ_i are the (positive) eigenvalues of $D^2 G^\mu(A_0)$. Details are given in the appendix.*

The following shadowing type theorem holds, where

$$\rho_U := \mathrm{diam}\Big(\Pi_\omega(U)\Big)$$

and $\Pi_\omega : \mathbf{R}^n \to \mathbf{R}^n$ denotes the orthogonal projection onto $\omega^\perp$.

Theorem 2.3 *Assume $(H1)$ and the "splitting condition" 2.1. Then $\forall I_0, I_0'$ with $\omega \cdot I_0 = \omega \cdot I_0'$, there is a heteroclinic orbit connecting the invariant tori $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I_0'}$. Moreover there exists $C_3 > 0$ such that $\forall \eta > 0$ small enough the "diffusion time" T_d needed to go from a η-neighbourhood of $\mathcal{T}_{I_0}$ to a η-neighbourhood of $\mathcal{T}_{I_0'}$ is bounded by*

$$T_d \leq C_3 \frac{|I_0 - I_0'|}{\delta} \cdot \rho_U \cdot \max\Big(|\log\delta|, \frac{1}{\gamma\alpha^\tau}\Big) + C_3|\log(\eta)|. \tag{2.28}$$

Remark 2.4 *The meaning of (2.28) is the following: the diffusion time T_d is estimated by the product of the number of heteroclinic transitions $k =$ (heteroclinic jump / splitting) $= |I_0' - I_0|/\delta$, and of the time T_s required for a single transition, that is $T_d = k \cdot T_s$. The time for a single transition T_s is bounded by the maximum time between the "ergodization time" $(1/\gamma\alpha^\tau)$, i.e. the time needed for the flow ωt to make an α-net of the torus, and the time $|\log\delta|$ needed to "shadow" homoclinic orbits for the forced pendulum equation. We use here that these homoclinic orbits are exponentially asymptotic to the equilibrium.*

PROOF. Assume with no loss of generality that $A_0 = 0$ and $\inf_U G_\mu(A) = 0$. Let us choose the number of bumps k as

$$k = \Big[\frac{24 \cdot \rho_U \cdot |I_0' - I_0|}{\delta}\Big] + 1. \tag{2.29}$$

By lemma 2.4-(i) and lemma 2.1-(ii), the trajectory $q^L_{A,\theta}$ converges exponentially fast to $\mathcal{T}_{I_0}$ (resp. $\mathcal{T}_{I_0'}$) as $t \to -\infty$ (resp. $+\infty$) from θ_1 (resp. θ_k). Therefore it is enough to prove the existence of a critical point $(\overline{A}, \overline{\theta}) \in \mathbf{T}^n \times \mathbf{R}^k$ of the k-bump heteroclinic function $\mathcal{F}^k_\mu$, defined in (2.22), such that for some positive constant K_1

$$|\overline{\theta}_k - \overline{\theta}_1| \leq K_1 \frac{|I_0 - I_0'|}{\delta} \cdot \rho_U \cdot \max\Big(|\log\delta|, \frac{1}{\gamma\alpha^\tau}\Big). \tag{2.30}$$

More precisely we shall enforce

$$K_2|\log\delta| < |\overline{\theta}_{i+1} - \overline{\theta}_i| < K_3 \max\Big(|\log\delta|, \frac{1}{\gamma\alpha^\tau}\Big) \qquad \forall i = 1, \ldots, k-1, \tag{2.31}$$

for some positive constants K_2,K_3. Let $(\Omega_1, \cdots, \Omega_n)$ be an orthonormal basis of $\mathbf{R}^n$ where

$$\Omega_1 = \frac{\omega}{|\omega|} \quad \text{and} \quad \Omega_2 = \frac{I_0' - I_0}{|I_0' - I_0|};$$

We recall that $\omega \cdot (I_0' - I_0) = 0$. In order to find a critical point of $\mathcal{F}^k_\mu$ we introduce suitable coordinates $(a_1, \ldots, a_n, s_1, \ldots, s_k) \in \mathbf{R}^n \times (-\mathrm{diam}U, \mathrm{diam}U)^k$ defined by

$$A = \sum_{j=1}^n a_j \Omega_j, \qquad \theta_i = \frac{\eta_i + s_i - a_1}{|\omega|} \quad \forall i = 1, \ldots, k$$

where η_i are constants to be chosen later. In these new coordinates the heteroclinic function defined in (2.22) is given by

$$\overline{\mathcal{F}}^k_\mu(a_1, a_2, \ldots, a_n, s_1, \ldots, s_k) = F^k_\mu\Big(\sum_{j=1}^n a_j\Omega_j, \frac{\eta_1 + s_1 - a_1}{|\omega|}, \ldots, \frac{\eta_k + s_k - a_1}{|\omega|}\Big) - |I_0' - I_0| a_2. \tag{2.32}$$

Using the invariance property (2.21) we see that $\overline{\mathcal{F}}_\mu^k$ does not depend on the new variable a_1 :

$$\begin{aligned}\overline{\mathcal{F}}_\mu^k(a_1,a_2,\dots,a_n,s_1,\dots,s_k) &= F_\mu^k\Big(\sum_{j=2}^n a_j\Omega_j,\frac{\eta_1+s_1}{|\omega|},\dots,\frac{\eta_k+s_k}{|\omega|}\Big)-|I_0'-I_0|a_2\\ &= \overline{\mathcal{F}}_\mu^k(0,a_2,\dots,a_n,s_1,\dots,s_k).\end{aligned}$$

For simplicity we will still use the notation $\overline{\mathcal{F}}_\mu^k(a_2,\dots,a_n,s_1,\dots,s_k)$ for $\overline{\mathcal{F}}_\mu^k(0,a_2,\dots,a_n,s_1,\dots,s_k)$. We now choose the constants η_i. Let

$$D=\frac{|\omega|}{C_2}\Big|\log\Big(\frac{24C_2}{\delta}\Big)\Big|+2\mathrm{diam}U\ , \tag{2.33}$$

where C_2 is the constant appearing in lemma 2.6. We shall use the following fact (see [13]): there is $\overline{C}>0$ such that, for all intervals $J\subset\mathbf{R}$ of length greater or equal to $\overline{C}/(\gamma\alpha^\tau)$, there is $\theta\in J$ such that

$$d(\theta\omega,2\pi\mathbf{Z}^n)<\alpha. \tag{2.34}$$

By (2.34) there is $(\eta_1,\dots,\eta_k)\in\mathbf{R}^k$ such that

$$\eta_i\Omega_1\equiv\chi_i,\ \mathrm{mod}2\pi\mathbf{Z}^n,\quad |\chi_i|<\alpha\quad\text{and}\quad \chi_i\cdot\Omega_1=0,\ i.e.\ \chi_i=\sum_{j=2}^n\chi_{i,j}\Omega_j, \tag{2.35}$$

$$\eta_1=0,\quad D<\eta_{i+1}-\eta_i<\Big(D+\frac{\overline{C}|\omega|}{\gamma\alpha^\tau}\Big). \tag{2.36}$$

By (2.33), (2.36), since $s_i\in(-\mathrm{diam}U,\mathrm{diam}U)$ we have that $\theta_{i+1}-\theta_i\geq\frac{1}{C_2}|\log(\frac{24C_2}{\delta})|$; hence, by lemma 2.6, setting

$$\begin{aligned}\overline{R}_i &= \overline{R}_i(a_2,\dots,a_n,s_{i-1},s_i,s_{i+1})\\ &= R_i\Big(\mu,\sum_{j=2}^n a_j\Omega_j,\frac{s_{i-1}+\eta_{i-1}}{|\omega|},\frac{s_i+\eta_i}{|\omega|},\frac{s_{i+1}+\eta_{i+1}}{|\omega|}\Big)\end{aligned}$$

we get

$$|\overline{R}_i(a_2,\dots,a_n,s_{i-1},s_i,s_{i+1})|<\frac{\delta}{24}. \tag{2.37}$$

By lemma 2.6, the invariance property (2.21), (2.35) and since G_μ is $2\pi\mathbf{Z}^n$-periodic, we have

$$\begin{aligned}\overline{\mathcal{F}}_\mu^k(a_2,\dots,a_n,s_1,\dots,s_k) &= \sum_{i=1}^k F_\mu\Big(\sum_{j=2}^n a_j\Omega_j,\frac{\eta_i+s_i}{|\omega|}\Big)+\overline{R}_i-|I_0'-I_0|a_2\\ &= \sum_{i=1}^k F_\mu\Big(\sum_{j=2}^n a_j\Omega_j+\chi_i+s_i\Omega_1,0\Big)+\overline{R}_i-|I_0'-I_0|a_2\\ &= \sum_{i=1}^k G_\mu\Big(\sum_{j=2}^n(a_j+\chi_{i,j})\Omega_j+s_i\Omega_1\Big)+\overline{R}_i-|I_0'-I_0|a_2\\ &= \sum_{i=1}^k \overline{G}_\mu(a_2+\chi_{i,2},\dots,a_n+\chi_{i,n},s_i)+\overline{R}_i-|I_0'-I_0|a_2\end{aligned} \tag{2.38}$$

where $\overline{G}_\mu(a_2,\dots,a_n,s)=G_\mu(\sum_{j=2}^n a_j\Omega_j+s\Omega_1)$. By (2.38), setting $\overline{\chi}_i:=(\chi_{i,2},\dots,\chi_{i,n})\in\mathbf{R}^{n-1}$ and $a=(a_2,\dots,a_n)$, we get

$$\overline{\mathcal{F}}_\mu^k(a_2,\dots,a_n,s_1,\dots,s_k)=\sum_{i=1}^k\overline{G}_\mu(a+\overline{\chi}_i,s_i)+\overline{R}_i-|I_0'-I_0|a_2. \tag{2.39}$$

Since the basis $(\Omega_1,\ldots,\Omega_n)$ is orthonormal the function $\overline{G}_\mu$ satisfies the same properties as G_μ, *i.e.*

$$\sup_{B_\alpha(0)} \overline{G}_\mu \le \frac{\delta}{4},\ \inf_{\partial\widetilde{U}} \overline{G}_\mu \ge \delta \text{ and } d\Big(\{x\in\widetilde{U}\mid \overline{G}_\mu(x)\le \delta/2\},\{x\in\widetilde{U}\mid \overline{G}_\mu(x)\ge 3\delta/4\}\Big)\ge 2\alpha,$$

where $\widetilde{U}=\{(a_2,\ldots,a_n,s)\in\mathbf{R}^n \mid \sum_{j=2}^n a_j\Omega_j+s\Omega_1\in U\}$. Note that $\mathrm{diam}U=\mathrm{diam}\widetilde{U}$. In particular, since $(0,0)\in\widetilde{U}$, for all $(a,s)\in\widetilde{U}$, $s\in(-\mathrm{diam}U,\mathrm{diam U})$. Moreover, for all $(a,s)\in\widetilde{U}$, $|a|=|\Pi_\omega(\sum_{j=2}^n a_j\Omega_j+s\Omega_1)|\le\rho_U$.

We shall find a critical point of $\overline{\mathcal{F}}^k_\mu$ in

$$W=\Big\{(a,s)\in\mathbf{R}^{n-1}\times\mathbf{R}^k\ \Big|\ (a+\overline{\chi}_i,s_i)\in\widetilde{U},\quad \forall i=1,\ldots,k\Big\}.$$

As W is bounded, $\overline{\mathcal{F}}^k_\mu$ attains its minimum over $\overline{W}$ at some point $(\overline{a},\overline{s})$. Notice that by (2.39) and (2.37)

$$\inf_{\overline{W}}\overline{\mathcal{F}}^k_\mu \le \overline{\mathcal{F}}^k_\mu(0,0)=\sum_{i=1}^k \overline{G}_\mu(\overline{\chi}_i,0)+k\frac{\delta}{24}.$$

Hence, since $|\overline{\chi}_i|<\alpha$ for all $i=1,\ldots,k$ and $\sup_{B_\alpha(0)}\overline{G}_\mu\le\delta/4$, we have

$$\inf_{\overline{W}}\overline{\mathcal{F}}^k_\mu \le k\frac{\delta}{4}+k\frac{\delta}{24}=k\frac{7\delta}{24}. \tag{2.40}$$

The theorem is proved if we show that $(\overline{a},\overline{s})\in W$. Arguing by contradiction we assume that $(\overline{a},\overline{s})\in\partial W$. Then there is some $l\in\{1,\cdots,k\}$ such that $(\overline{a}+\overline{\chi}_l,\overline{s}_l)\in\partial\widetilde{U}$, so that $\overline{G}_\mu(\overline{a}+\overline{\chi}_l,\overline{s}_l)\ge\delta$. We now prove that $\{(\overline{a}+\overline{\chi}_l,t);t\in(-\mathrm{diam}U,\mathrm{diam}U\ )\}\cap\widetilde{U}\subset Z:=\{x\in\widetilde{U}\mid \overline{G}_\mu(x)\ge 3\delta/4\}$. Indeed, if not, by (2.37), for some $t\in(-\mathrm{diam}U,\mathrm{diam}U)$ such that $(\overline{a}+\overline{\chi}_l,t)\in\widetilde{U}$,

$$\begin{aligned}
\overline{\mathcal{F}}^k_\mu(\overline{a},\overline{s}_1,\cdots,\overline{s}_{l-1},t,\overline{s}_{l+1},\cdots,\overline{s}_k) \ &\le\ \overline{\mathcal{F}}^k_\mu(\overline{a},\overline{s})+(\overline{G}_\mu(\overline{a}+\overline{\chi}_l,t)-\overline{G}_\mu(\overline{a}+\overline{\chi}_l,\overline{s}_l)))\\
&+\ |\overline{R}_{l-1}(\overline{a},\overline{s}_{l-2},\overline{s}_{l-1},s_l)-\overline{R}_{l-1}(\overline{a},\overline{s}_{l-2},\overline{s}_{l-1},t)|\\
&+\ |\overline{R}_l(\overline{a},s_{l-1},t,\overline{s}_l)-\overline{R}_l(\overline{a},\overline{s}_{l-1},\overline{s}_l,\overline{s}_{l+1})|\\
&+\ |\overline{R}_{l+1}(\overline{a},\overline{s}_l,\overline{s}_{l+1},t)-\overline{R}_{l+1}(\overline{a},\overline{s}_l,\overline{s}_{l+1},\overline{s}_{l+2})|\\
&<\ \overline{\mathcal{F}}^k_\mu(\overline{a},\overline{s})-\frac{\delta}{4}+\frac{6\delta}{24}=\overline{\mathcal{F}}^k_\mu(\overline{a},\overline{s}),
\end{aligned}$$

which is wrong since $(\overline{a},\overline{s})$ is the minimum of $\overline{\mathcal{F}}^k_\mu$ over $\overline{W}$. We deduce in particular that, for all i,

$$(\overline{a}+\overline{\chi}_l,\overline{s_i})\in Z\cup\widetilde{U}^c. \tag{2.41}$$

We now prove that for all i,

$$\overline{G}_\mu(\overline{a}+\overline{\chi}_i,\overline{s_i})>\frac{\delta}{2}. \tag{2.42}$$

Indeed, assume that $\overline{G}_\mu(\overline{a}+\overline{\chi}_i,\overline{s_i})\le\delta/2$. Since $\overline{G}_\mu\ge 3\delta/4$ in a neighbourhood of $\partial\widetilde{U}$, $d(x,\widetilde{U}^c)\ge2\alpha$ for all $x\in\widetilde{U}$ such that $\overline{G}_\mu(x)\le\delta/2$, so our splitting condition implies that

$$d\Big(Z\cup\widetilde{U}^c,\{x\in\widetilde{U}\mid\overline{G}_\mu(x)\le\delta/2\}\Big)\ge 2\alpha.$$

Therefore we get, recalling (2.41), that $|(\overline{a}+\overline{\chi}_l,\overline{s_i})-(\overline{a}+\overline{\chi}_i,\overline{s_i})|=|\overline{\chi}_l-\overline{\chi}_i|\ge2\alpha$. This contradicts $|\overline{\chi}_i|,|\overline{\chi}_l|<\alpha$ in (2.35).

By (2.39), (2.42), (2.37), and noting that, by (2.29), $|I_0'-I_0||a_2|\le|I_0'-I_0|\ \rho_U\ \le(k\delta)/24$, we deduce that

$$\overline{\mathcal{F}}^k_\mu(\overline{a},\overline{s})\ge k\frac{\delta}{2}-k\frac{\delta}{24}-k\frac{\delta}{24}=k\frac{10\delta}{24}>k\frac{7\delta}{24},$$

contradicting (2.40). The proof of the theorem is complete. ∎

Remark 2.5 *The latter proof works if G_μ possesses a local maximum which satisfies a non-degeneracy type condition like the "splitting condition" 2.1, while in the approaches developed in [9] and [45], based on Mather's theory, diffusion orbits are always built from local minima of G_μ. The proof of the shadowing theorem when the homoclimic point A_0 is a saddle point requires slightly different arguments involving the topological degree. For example it holds assuming as in [24] the condition $D^2 G_\mu(A_0)\omega \cdot \omega \neq 0$. This condition enables to prove that $deg(\nabla \overline{\mathcal{F}}^k_\mu(a,s), W, 0) \neq 0$ for a suitable W.*

When the frequency vector ω is considered as a constant, independent of any parameter ("a priori-unstable case") it is easy to justify the splitting condition 2.1 using the first-order approximation given by the Poincaré-Melnikov primitive. With a Taylor expansion in μ we can easily prove that for μ small enough

$$G_\mu(B) = b + \mu\Gamma(B) + O(\mu^2), \ \forall B \in \mathbf{T}^n,$$

where $b = \Phi_0(q_0)$ and $\Gamma : \mathbf{T}^n \to \mathbf{R}$ is nothing but the Poincaré-Melnikov primitive

$$\Gamma(B) = \int_{\mathbf{R}} (\cos q_0(t) - 1) f(\omega t + B) \ dt. \tag{2.43}$$

Hence, if Γ possesses a proper minimum (resp. maximum) in $A_0 \in \mathbf{R}^n$, i.e if $\exists r > 0$ such that $\inf_{\partial B_r(A_0)} \Gamma > \Gamma(A_0)$ (resp. $\sup_{\partial B_r(A_0)} \Gamma < \Gamma(A_0)$) then, for μ small enough, the "splitting" condition 2.1 holds with $U = B_\rho(A_0)$, $\rho_U = O(1)$, $\delta = c\mu$ and $\alpha = c'$ for some positive constants c and c' depending on Γ. We remark that $B_r(A_0)$ could be replaced by a bounded open subset U of $\mathbf{R}^n$. Applying theorem 2.3 we deduce

Theorem 2.4 *Assume $(H1)$ and let Γ possess a proper minimum (or maximum) A_0, i.e. suppose that $\exists r > 0$ such that $\inf_{\partial B_r(A_0)} \Gamma > \Gamma(A_0)$. Then, for μ small enough, the same statement as in theorem 2.3 holds with a diffusion time $T_d = O((1/\mu)\log(1/\mu))$.*

3 More general perturbation terms

In this section we show how to adapt the arguments of the previous section in order to deal with a more general perturbation term $f(\varphi, q)$. Regarding regularity it is sufficient to have finite large enough smoothness for f. The equation of motion derived by Hamiltonian $\mathcal{H}_\mu$ are

$$\dot\varphi = \omega, \qquad \dot I = -\mu\partial_\varphi f(\varphi, q), \qquad \dot q = p, \qquad \dot p = \sin q - \mu\partial_q f(\varphi, q), \tag{3.44}$$

corresponding to the quasi-periodically forced pendulum

$$-\ddot q + \sin q = \mu \ \partial_q f(\omega t + A, q). \tag{3.45}$$

3.1 Invariant tori in the perturbed system

The first step is to prove the persistence of invariant tori for $\mu \neq 0$ small enough. It appears that no more than the standard Implicit Function Theorem is required to prove the following well known result (see for example [32] for a different proof)

Theorem 3.1 *Let ω satisfy $(H1)$. For μ small enough and $\forall I_0 \in \mathbf{R}^n$ system $\mathcal{H}_\mu$ possesses n-dimensional invariant tori $\mathcal{T}^\mu_{I_0} \approx \mathcal{T}_{I_0}$ of the form*

$$\mathcal{T}^\mu_{I_0} = \Big\{ I = I_0 + a^\mu(\psi), \ p = P^\mu(\psi), \ \varphi = \psi, \ q = Q^\mu(\psi), \quad \psi \in \mathbf{T}^n \Big\}, \tag{3.46}$$

with $Q^\mu(\psi), P^\mu(\psi) = O(\mu)$, $a^\mu(\psi) = O(\mu)$. Moreover the dynamics on $\mathcal{T}^\mu_{I_0}$ is conjugated to the rotation of speed ω for ψ.

We first determine the functions Q^μ, P^μ in (3.46). Using the standard Implicit Function Theorem we prove that there exists a unique quasi-periodic solution $q_A^\mu(t)$ for the quasi-periodically forced pendulum equation (3.45) which bifurcates from the hyperbolic equilibrium 0.

Lemma 3.1 *Let $f \in C^l(\mathbf{T}^n \times \mathbf{T})$, $l \geq 2$. For μ small enough there exists a unique quasi-periodic solution $q_A^\mu(t)$ of (3.45) with $(q_A^\mu(t), \dot{q}_A^\mu(t)) = (q_A^\mu(t), p_A^\mu(t)) = O(\mu)$, C^{l-1}-smooth in A. More precisely there exist functions $Q^\mu, P^\mu : \mathbf{T}^n \to \mathbf{R}$ of class C^{l-1}, such that $(q_A^\mu(t), p_A^\mu(t)) = (Q^\mu(\omega t + A), P^\mu(\omega t + A))$.*

PROOF. Let L be the Green operator of the differential operator $h \to -D^2 h + h$ with Dirichlet boundary conditions at $\pm\infty$. L is explicitely given by $L(f) = \int_{\mathbf{R}} e^{-|t-s|} f(s)\, ds/2$. It results that L is a continuous linear operator in the Banach space of the continuous bounded functions from $\mathbf{R}$ to $\mathbf{R}$ endowed with the sup-norm, which we shall denote by E. We consider the non-linear operator $S : \mathbf{R} \times \mathbf{T}^n \times E \to E$

$$S(\mu, A, q) := q - L(q - \sin q) - \mu L(\partial_q f(\omega t + A, q)).$$

S is of class C^{l-1}. We are looking for a solution q_A^μ of $S(\mu, A, q) = 0$. Since $S(0, A, 0) = 0$ and $\partial_q S(0, A, 0) = Id$, by the Implicit Function Theorem there exists, for μ small enough, a unique solution $q_A^\mu = O(\mu)$. By (3.45) $q_A^\mu \in C^{l+1}(\mathbf{R})$; moreover it is C^{l-1}-smooth in A. We define the C^{l-1}-maps $Q^\mu(\cdot), P^\mu(\cdot) : \mathbf{T}^n \to \mathbf{R}$ by

$$Q^\mu(A) := q_A^\mu(0), \qquad P^\mu(A) := \dot{q}_A^\mu(0).$$

By uniqueness we deduce that $q_A^\mu(s + t) = q_{A+\omega s}^\mu(t)$, $\forall s, t \in \mathbf{R}$. For $t = 0$ this yields

$$q_A^\mu(s) = q_{A+\omega s}^\mu(0) := Q^\mu(A + \omega s) \quad \text{and} \quad p_A^\mu(s) = p_{A+\omega s}^\mu(0) := P^\mu(A + \omega s), \ \forall s \in \mathbf{R},$$

proving the lemma.■

We now define the functions $a^\mu(\psi)$ of (3.46). We impose that $(\omega t + A, I_0 + a^\mu(\omega t + A), Q^\mu(\omega t + A), P^\mu(\omega t + A))$ satisfy the equations of motions (3.44); hence the functions $a^\mu(\psi)$ must satisfy the following system of equations

$$(\omega \cdot \nabla) a^\mu(\psi) = \mu g^\mu(\psi), \quad \text{where} \quad g^\mu(\psi) := -(\nabla_\psi f)(\psi, Q^\mu(\psi)). \tag{3.47}$$

In order to solve (3.47) we expand in Fourier series the functions $a^\mu(\psi) = \sum_{k \in \mathbf{Z}^n} a_k e^{ik\cdot\psi}$, $g^\mu(\psi) = \sum_{k \in \mathbf{Z}^n} g_k e^{ik\cdot\psi}$. Each Fourier coefficient a_k must then satisfy

$$i(k \cdot \omega) a_k = \mu g_k, \qquad \forall k \in \mathbf{Z}^n. \tag{3.48}$$

It is necessary for the existence of a solution that $g_0 = \int_{\mathbf{T}^n} g^\mu(\psi) d\psi = 0$. This property can be checked directly, that is

Lemma 3.2 *We have*

$$\int_{\mathbf{T}^n} (\nabla_\psi f)(\psi, Q^\mu(\psi))\, d\psi = 0. \tag{3.49}$$

PROOF. For a proof see [16] or appendix A12 of [15] . ■

Since ω satisfies $(H1)$, if f is smooth enough, then the function a^μ defined by

$$a^\mu(\psi) = \sum_{k \in \mathbf{Z}^n, k \neq 0} \frac{g_k}{i(k \cdot \omega)} e^{ik\cdot\psi}, \tag{3.50}$$

which formally solves equation (3.48), is well defined and smooth. Indeed since $f \in C^l$ the function g^μ defined in (3.47) is C^{l-1} and there exists $M > 0$ such that $|g_k| \leq M/|k|^{l-1}$, $\forall k \in \mathbf{Z}^n$, $k \neq 0$. By $(H1)$ it follows that $|a_k| \leq M/|k|^{l-1}|\omega \cdot k| \leq M|k|^\tau/(\gamma|k|^{l-1})$. The proof of theorem 3.1 is complete.

3.2 The new symplectic coordinates

In order to reduce to the previous case we want to put the tori $\mathcal{T}^\mu_{I_0}$ at the origin by a symplectic change of variables. Recalling that the tori $\mathcal{T}^\mu_{I_0}$ are *isotropic* submanifolds, i.e. that the symplectic 2 form $dI \wedge d\psi + dp \wedge dq$ vanishes on each tangent space to $\mathcal{T}^\mu_{I_0}$ (or equivalently $da^\mu(\psi) \wedge d\psi + dP^\mu(\psi) \wedge dQ^\mu(\psi) = 0$), we can prove the following lemma

Lemma 3.3 *The transformation of coordinates $(J, v, \psi, u) \to (I, p, \varphi, q)$ defined on the space $\mathbf{R}^n \times \mathbf{R} \times \mathbf{T}^n \times \mathbf{R}$ by*

$$I = a^\mu(\psi) + u \partial_\psi P^\mu(\psi) - v \partial_\psi Q^\mu(\psi) + J, \quad p = P^\mu(\psi) + v, \quad \varphi = \psi, \quad q = Q^\mu(\psi) + u \tag{3.51}$$

is symplectic.

PROOF. Since the tori $\mathcal{T}^\mu_{I_0}$ are isotropic the 1-form on $\mathbf{T}^n$ $\lambda = \sum_{i=1}^n a_i^\mu d\varphi_i + P^\mu dQ^\mu$ is closed. Therefore there exists a unique $c \in \mathbf{R}^n$ and there exists $b^\mu : \mathbf{T}^n \to \mathbf{R}$ such that $\lambda = \sum_{i=1}^n c_i d\varphi_i + db^\mu$. Then transformation (3.51) has the generating function

$$\Phi(J, v, \varphi, q) = (J + c) \cdot \varphi + vq + b^\mu(\varphi) - P^\mu(\varphi) Q^\mu(\varphi) + q P^\mu(\varphi) - v Q^\mu(\varphi)$$

and hence is symplectic. ■

In the new coordinates each invariant torus $\mathcal{T}^\mu_{I_0}$ is simply described by $\{J = I_0,\ \psi \in \mathbf{T}^n,\ u = v = 0\}$ and the new Hamiltonian writes

$$(\mathcal{K}_\mu) \qquad \mathcal{K}_\mu = E_\mu + \omega \cdot J + \frac{v^2}{2} + (\cos u - 1) + P_0(\mu, u, \psi)$$

where

$$P_0(\mu, u, \psi) = \Big(\cos(Q^\mu + u) - \cos Q^\mu + (\sin Q^\mu) u + 1 - \cos u \Big) + \mu \Big(f(\psi, Q^\mu + u) - f(\psi, Q^\mu) - \partial_q f(\psi, Q^\mu) u \Big)$$

and E_μ is the energy of the perturbed invariant torus $\mathcal{T}^\mu_0 = \{(a^\mu(\psi), \psi, Q^\mu(\psi), P^\mu(\psi)); \psi \in \mathbf{T}^n\}$. To Hamiltonian $(\mathcal{K}_\mu)$ is associated the quasi-periodically forced pendulum equation

$$-\ddot{u} + \sin u = \partial_u P_0(\mu, u, \omega t + A). \tag{3.52}$$

of Lagrangian

$$L_\mu = \frac{\dot{u}^2}{2} + (1 - \cos u) - P_0(\mu, u, \omega t + A). \tag{3.53}$$

Since the Hamiltonian $\mathcal{K}_\mu$ is no more periodic in the variable u we can not directly apply theorem 2.3 and the arguments of the previous sections require some modifications. Arguing as in lemma 2.1 we deduce that, there exists, for μ small enough, a unique 1-bump pseudo-homoclinic solution $u^\mu_{A,\theta}(t)$, true solution of (3.52) in $(-\infty, \theta), (\theta, +\infty)$, satisfying all the properties of lemma 2.1. Then we define the function $\mathcal{F}_\mu : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$ as

$$\begin{aligned} \mathcal{F}_\mu(A, \theta) &= \int_{-\infty}^\theta \frac{(\dot{u}^\mu_{A,\theta})^2}{2} + (1 - \cos u^\mu_{A,\theta}) - P_0(\mu, u^\mu_{A,\theta}, \omega t + A)\, dt \\ &+ \int_\theta^{+\infty} \frac{(\dot{u}^\mu_{A,\theta})^2}{2} + (1 - \cos u^\mu_{A,\theta}) - P_1(\mu, u^\mu_{A,\theta}, \omega t + A)\, dt + 2\pi \dot{q}^\mu_A(\theta), \end{aligned}$$

where, $q^\mu_A(t) = Q^\mu(\omega t + A)$ and $\forall i \in \mathbf{Z}$, we have set

$$\begin{aligned} P_i(\mu, u, \omega t + A) &= \Big(\cos(q^\mu_A(t) + u) - \cos q^\mu_A(t) + \sin q^\mu_A(t)\ (u - 2\pi i) + 1 - \cos u \Big) \\ &+ \mu \Big(f(\omega t + A, q^\mu_A(t) + u) - f(\omega t + A, q^\mu_A(t)) - (\partial_q f)(\omega t + A, q^\mu_A(t))\ (u - 2\pi i) \Big). \end{aligned}$$

Since $u^{\mu}_{A,\theta}$ converges exponentially fast to 0 for $t \to -\infty$ and to 2π for $t \to +\infty$ the above integrals are convergent. We define the *"homoclinic function"* $\mathcal{G}_{\mu} : \mathbf{T}^n \to \mathbf{R}$ as

$$\mathcal{G}_{\mu}(A) = \mathcal{F}_{\mu}(A, 0). \tag{3.54}$$

There holds also $\mathcal{F}_{\mu}(A,\theta) = \mathcal{G}_{\mu}(A + \omega\theta)$, $\forall \theta \in \mathbf{R}$. Arguing as in lemma 2.4 we can prove the existence of k-bump pseudo-homoclinic solutions $u^{L}_{A,\theta}$, which are true solutions of (3.52) in each interval $(-\infty, \theta_1)$, (θ_i, θ_{i+1}) $(i = 1, \ldots, k-1)$, $(\theta_k, +\infty)$, and satisfy all the properties of lemma 2.4. Then we define the "k-bump heteroclinic function"

$$\begin{aligned}
\mathcal{F}^k_{\mu}(A, \theta_1, \ldots, \theta_k) \;&=\; \int_{-\infty}^{\theta_1} \frac{(\dot u^L_{A,\theta})^2}{2} + (1 - \cos u^L_{A,\theta}) - P_0(\mu, u^L_{A,\theta}, \omega t + A)\; dt + 2\pi \dot q^{\mu}_A(\theta_1) \\
&+ \sum_{i=1}^{k-1} \int_{\theta_i}^{\theta_{i+1}} \frac{(\dot u^L_{A,\theta})^2}{2} + (1 - \cos u^L_{A,\theta}) - P_i(\mu, u^L_{A,\theta}, \omega t + A)\; dt + 2\pi \dot q^{\mu}_A(\theta_{i+1}) \\
&+ \int_{\theta_k}^{+\infty} \frac{(\dot u^L_{A,\theta})^2}{2} + (1 - \cos u^L_{A,\theta}) - P_k(\mu, u^{\mu}_{A,\theta}, \omega t + A)\; dt - (I'_0 - I_0) \cdot A
\end{aligned}$$

If $\partial_{\theta_i} \mathcal{F}^k_{\mu}(A, \theta_1, \ldots, \theta_k) = (\dot u^L_{A,\theta})^2(\theta_i^-)/2 - (\dot u^L_{A,\theta})^2(\theta_i^+)/2 = 0$ then $u^L_{A,\theta}$ is a true solution of the quasi-periodically forced pendulum equation (3.52). As in the previous section the variation in the action variables is given by the partial derivative with respect to A, that is

$$\partial_A \mathcal{F}^k_{\mu}(A,\theta) = \int_{-\infty}^{+\infty} -\mu\Big(\partial_{\varphi} f(\omega t + A, q^{\mu}_A(t) + u^L_{A,\theta}(t)) - \partial_{\varphi} f(\omega t + A, q^{\mu}_A(t))\Big)\; dt - (I'_0 - I_0). \tag{3.55}$$

Lemma 3.4 *Let (A, θ) be a critical point of $\mathcal{F}^k_{\mu}$. Then there exists a heteroclinic orbit connecting the tori $\mathcal{T}^{\mu}_{I_0}$ and $\mathcal{T}^{\mu}_{I'_0}$.*

PROOF. By (3.55) it is easy to verify that the solution of (3.44) $(I_{\mu}(t), \omega t + A, q^{\mu}_A + u^L_{A,\theta}, \dot q^{\mu}_A + \dot u^L_{A,\theta})$, with $I_{\mu}(t) = C - \mu \int_0^t \partial_{\varphi} f(\omega s + A, q^{\mu}_A(s) + u^L_{A,\theta}(s))\, ds$ and $C = I'_0 + a_{\mu}(A) + \mu \int_0^{+\infty} \partial_{\varphi} f(\omega t + A, q^{\mu}_A(t) + u^L_{A,\theta}(t)) - \partial_{\varphi} f(\omega t + A, q^{\mu}_A(t))\, dt$, is a heteroclinic solution connecting $\mathcal{T}^{\mu}_{I_0}$ and $\mathcal{T}^{\mu}_{I'_0}$. ■

Finally, arguing as in the proof of theorem 2.3, we obtain

Theorem 3.2 *Assume $(H1)$ and let $\mathcal{G}_{\mu}$ satisfy the "splitting condition" 2.1. Then $\forall I_0, I'_0$ with $\omega \cdot I_0 = \omega \cdot I'_0$, there is a heteroclinic orbit connecting the invariant tori $\mathcal{T}^{\mu}_{I_0}$ and $\mathcal{T}^{\mu}_{I'_0}$. The same estimate on the diffusion time as in theorem 2.3 holds.*

A Taylor expansion in μ gives

Lemma 3.5 *For μ small enough*

$$\mathcal{G}_{\mu}(A) = b + \mu M(A) + O(\mu^2), \qquad \forall A \in \mathbf{T}^n \tag{3.56}$$

where $b = \Phi_0(q_0)$ and $M(A)$ is the Poincaré-Melnikov primitive $M(A) = \int_{\mathbf{R}} \Big[f(\omega t + A, 0) - f(\omega t + A, q_0(t))\Big]\; dt$.

PROOF. We develop in μ the Lagrangian $L_{\mu,i} = \dot u^2/2 + (1 - \cos u) - P_i(\mu, u, \omega t + A)$ for $i = 0, 1$.

$$L_{\mu,i} = \frac{\dot u^2}{2} + (1 - \cos u) - \mu\Big((u - 2\pi i - \sin u)\gamma + f(\omega t + A, u) - f(\omega t + A, 0) - \partial_q f(\omega t + A, 0)(u - 2\pi i)\Big) + \mathcal{R}_i(\mu, u, t)$$

where $\gamma(t) := \partial_{\mu_{|\mu=0}} q^\mu_A(t)$ and $|\mathcal{R}(\mu,u,t)| = O(\mu^2(u-2\pi i)^2)$. Hence $\mathcal{G}_\mu(A) = b + \mu M^*(A) + O(\mu^2)$ where

$$\begin{aligned} M^*(A) &= -\int_{-\infty}^{0} (q_0(t) - \sin q_0(t))\gamma(t) + f(\omega t + A, q_0(t)) - f(\omega t + A, 0) - \partial_q f(\omega t + A, 0) q_0(t)\ dt \\ &- \int_{0}^{+\infty} (q_0(t) - 2\pi - \sin q_0(t))\gamma(t) + f(\omega t + A, q_0(t)) - f(\omega t + A, 0)\ dt \\ &+ \int_{0}^{+\infty} \partial_q f(\omega t + A, 0)(q_0(t) - 2\pi)\ dt + 2\pi\dot{\gamma}(0). \end{aligned} \tag{3.57}$$

The lemma is proved if we show that $M^*(A) = M(A)$. Integrating by parts, since $\sin q_0(t) = \ddot{q}_0(t)$ and $-\ddot{\gamma} + \gamma = \partial_q f(\omega t + A, 0)$, we have

$$\int_{-\infty}^{0} (q_0(t) - \sin q_0(t))\gamma(t) = \int_{-\infty}^{0} q_0(t)\partial_q f(\omega t + A, 0) - \dot{q}_0(0)\gamma(0)\ dt + \pi\dot{\gamma}(0) \tag{3.58}$$

and

$$\int_{0}^{+\infty} (q_0(t) - 2\pi - \sin q_0(t))\gamma(t) = \int_{0}^{+\infty} (q_0(t) - 2\pi)\partial_q f(\omega t + A, 0)\ dt + \dot{q}_0(0)\gamma(0) + \pi\dot{\gamma}(0). \tag{3.59}$$

Finally we deduce from (3.57), (3.58) and (3.59) that $M^*(A) = M(A) = \int_{\mathbf{R}} [f(\omega t + A, 0) - f(\omega t + A, q_0(t))]\ dt$. ■

Theorem 3.3 *Assume $(H1)$ and let M possess a proper minimum (or maximum) A_0, i.e. $\exists r > 0$ such that $\inf_{\partial B_r(A_0)} M > M(A_0)$. Then, for μ small enough, the same statement as in theorem 3.2 holds where the diffusion time is $T_d = O((1/\mu)\log(1/\mu))$.*

Remark 3.1 *By theorems 3.1-3.2 we obtain that, for a priori-stable, isochronous, degenerate systems considered in [9] (see also [10])*

$$\mathcal{H}_\varepsilon = \varepsilon\omega \cdot I + \frac{p^2}{2} + \varepsilon^d(\cos q - 1) + \mu f(\varphi, q) \quad \text{with} \quad 1 < d < 2,$$

for $\mu = \delta\varepsilon^d$, δ being a small constant, the diffusion time is bounded by $T_d = O(C(\delta)/\varepsilon^d)$. This improves the result of [9], which holds for $\mu = O(\varepsilon^{d'})$, $d' > d/2 + 3$, and provides the upper bound on the diffusion time $T_d = O(1/\varepsilon^{C+2(\tau+1)(2d'-1-d/2)})$, C being a suitable positive constant.

4 Splitting of whiskers

If the frequency vector $\omega = \omega_\varepsilon$ contains some "fast frequencies" $\omega_i = O(1/\varepsilon^b)$, $b > 0$, ε being a small parameter, and if the perturbation is analytic, the oscillations of the Melnikov function along some directions turn out to be exponentially small with respect to ε. Hence the development (3.56) will provide a valid measure of the splitting only for μ exponentially small with respect to ε. In order to justify the dominance of the Poincaré-Melnikov function when $\mu = O(\varepsilon^p)$ we need more refined estimates for the error. However it turns out that the function $F_\mu(A,\theta)$ can not be easily analytically extended in a sufficiently wide complex strip (roughly speaking, the condition $q^\mu_{A,\theta}(Re\ \theta) = \pi$ appearing naturally when we try to extend the definition of $q^\mu_{A,\theta}$ to $\theta \in \mathbf{C}$ breaks analyticity). We bypass this problem considering the action functional evaluated on different "1-bump pseudo-homoclinic solutions" $Q^\mu_{A,\theta}$. This new "reduced action functional" $\widetilde{F}_\mu(A,\theta) = \widetilde{G}_\mu(A + \omega\theta)$ has the advantage to have an analytic extension in (A,θ) in a wide complex strip. Moreover we will show that the homoclinic functions G_μ, $\widetilde{G}_\mu$ corresponding to both reductions are the same up to a change of variables of the torus close to the identity. This enables to recover enough information on the homoclinic function G_μ to construct diffusion orbits.

We assume that $f(\varphi,q)=(1-\cos q)f(\varphi)$, $f(\varphi)=\sum_{k\in\mathbf{Z}^n} f_k \exp(ik\cdot\varphi)$ and that there are $r_i\geq 0$ such that

$$\forall s\in\mathbf{N},\ \exists D_s>0 \quad \text{such that} \quad |f_k|\leq \frac{D_s}{|k|^s}\exp\Big(-\sum_{i=1}^n r_i|k_i|\Big),\ \forall k\in\mathbf{Z}^n. \tag{4.60}$$

Condition (4.60) means that f has a C^∞ extension defined in

$$D:=(\mathbf{R}+i[-r_1,r_1])\times\ldots\times(\mathbf{R}+i[-r_n,r_n])$$

which is holomorphic w.r.t. the variables for which $r_i>0$ in $(\mathbf{R}+iI_1)\times\ldots\times(\mathbf{R}+iI_n)$, where $I_i=\{0\}$ if $r_i=0$, $I_i=(-r_i,r_i)$ if $r_i>0$. We denote the supremum of $|f|$ over D as

$$||f||:=\sup_{A\in D}|f(A)|. \tag{4.61}$$

Note that we can take in (4.60) $D_0=\sup_{\varphi\in\mathbf{T}^n}|f(\varphi)|\leq||f||$. It will be used starting from subsection 4.2.

4.1 The change of coordinates

In order to obtain an analytic "reduced functional" we perform a Lyapunov-Schmidt reduction with a supplementary space $\widetilde{E}_\theta$ to $\langle\dot{q}_\theta\rangle$ which depends analytically on θ. We define $\psi_0:\mathbf{R}\to\mathbf{R}$ by $\psi_0(t)=\cosh^2 t/(1+\cosh t)^3$ and set $\psi_\theta(t)=\psi(t-\theta)$. Note that, since,

$$\int_{\mathbf{R}}\psi_0(t)\dot{q}_0(t)\ dt=\overline{K}\neq 0. \tag{4.62}$$

$$\widetilde{E}_\theta:=\Big\{w\in H^1(\mathbf{R})\ \Big|\ \int_{\mathbf{R}}\psi_\theta(t)w(t)\ dt=0\Big\}$$

is a supplementary space to $\langle\dot{q}_\theta\rangle$.

Remark 4.1 *We could use any function ψ decaying at infinity, analytic in the complex strip $S:=\{\theta\in\mathbf{C}\mid|\mathrm{Im}\theta|<\pi/2\}$ such that (4.62) holds. We use the function $\psi_0(t)$ instead of the more natural $\dot{q}_0(t)$ because the closest singularities to the real line of $\psi_0(z)$ are located in the complex plain at $\pm i\pi$, while the closest singularities of $\dot{q}_0(t)$ are located at $\pm i(\pi/2)$. Hence near the boundary of the complex strip S the analytic function ψ_0 is regular and this improves the bounds of lemmas 4.6 and 4.7.*

Arguing as in lemma 2.1 we can prove, for μ small, the existence of "1 bump pseudo-homoclinics" $Q^\mu_{A,\theta}\in q_\theta+\widetilde{E}_\theta$ which is near q_θ.

Lemma 4.1 *There exist $\mu_1>0$ (independent of ω) and $C_4>0$ such that for all $0<\mu<\mu_1$, for all $\theta\in\mathbf{R}$, for all $A\in\mathbf{R}^n$ there exists a unique function $Q^\mu_{A,\theta}(t):\mathbf{R}\to\mathbf{R}$, and a constant $\alpha^\mu_{A,\theta}$ smooth in (A,θ,μ), such that*

- *(i)* $-\ddot{Q}^\mu_{A,\theta}(t)+\sin Q^\mu_{A,\theta}(t)=\mu\ \sin Q^\mu_{A,\theta}(t)f(\omega t+A)+\alpha^\mu_{A,\theta}\psi_\theta(t)$;
- *(ii)* $\int_{\mathbf{R}}(Q^\mu_{A,\theta}(t)-q_\theta(t))\psi_\theta(t)\ dt=0$;
- *(iii)* $\max\Big(|Q^\mu_{A,\theta}(t)-q_\theta(t)|,|\dot{Q}^\mu_{A,\theta}(t)-\dot{q}_\theta(t)|\Big)\leq C_4\mu\exp(-\frac{|t-\theta|}{2})$;
- *(iv)* $\max\Big(|\partial_A Q^\mu_{A,\theta}(t)|,|\partial_A\dot{Q}^\mu_{A,\theta}(t)|,|\omega\cdot\partial_A Q^\mu_{A,\theta}(t)|,|\omega\cdot\partial_A\dot{Q}^\mu_{A,\theta}(t)|\Big)\leq C_4\mu\exp(-\frac{|t-\theta|}{2})$.
- *(v)* $Q^\mu_{A,\theta}(t)=Q^\mu_{A+k2\pi,\theta}(t),\ \forall k\in\mathbf{Z}^n$ *and* $Q^\mu_{A,\theta+\eta}(t)=Q^\mu_{A+\omega\eta,\theta}(t-\eta),\ \forall\theta,\eta\in\mathbf{R}$.

We define the function $\widetilde{F}_\mu(A,\theta) : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$ as the action functional of Lagrangian 1.5 evaluated on the "1-bump pseudo-homoclinic solutions" $Q^\mu_{A,\theta}(t)$ obtained in lemma 4.1, namely

$$\widetilde{F}_\mu(A,\theta) := \Phi_{\mu,A}(Q^\mu_{A,\theta}) = \int_{\mathbf{R}} \mathcal{L}_{\mu,A}(Q^\mu_{A,\theta}(t), \dot{Q}^\mu_{A,\theta}(t), t)\ dt \tag{4.63}$$

and $\widetilde{G}_\mu(A) : \mathbf{T}^n \to \mathbf{R}$ as $\widetilde{G}_\mu(A) = \widetilde{F}_\mu(A,0)$. By lemma (4.1)-($v$) the following invariance property holds $\widetilde{F}_\mu(A,\theta+\eta) = \widetilde{F}_\mu(A+\omega\eta,\theta)$, $\forall\theta,\eta \in \mathbf{R}$; in particular $\widetilde{F}_\mu(A,\theta) = \widetilde{G}_\mu(A+\omega\theta)$, $\forall\theta \in \mathbf{R}$.

Lemma 4.1-(i)-(ii) suggests that $\{Q^\mu_{A,\theta} \mid \theta \in \mathbf{R}\}$ is "natural constraint" for $\Phi_{\mu,A}$. For completeness we prove it in detail (it is the analogue of (2.13)), together with useful bounds of the derivatives of $\widetilde{F}_\mu$ (and $\widetilde{G}_\mu$) up to order 2.

Lemma 4.2 *Let $0 < \mu < \min\{\mu_1, (\overline{K}/8C_4)\}$, where $\overline{K}$ is defined in (4.62). If $\partial_\theta \widetilde{F}_\mu(A,\theta) = 0$ then $Q^\mu_{A,\theta}$ is a true solution of (1.4). More precisely*

$$|\alpha^\mu_{A,\theta}| \le \frac{2}{|\overline{K}|} |\partial_\theta \widetilde{F}_\mu(A,\theta)|. \tag{4.64}$$

Moreover we also have

$$|\nabla^s \widetilde{G}_\mu(A)| = O(\mu), \qquad |\nabla^s \widetilde{F}_\mu(A,\theta)| = O(\mu), \qquad s = 1,2. \tag{4.65}$$

Proof. We have

$$\partial_\theta \widetilde{F}_\mu(A,\theta) = \int_{\mathbf{R}} \dot{Q}^\mu_{A,\theta}(t)\partial_\theta \dot{Q}^\mu_{A,\theta}(t) + \Big(\sin Q^\mu_{A,\theta}(t) - \mu \sin Q^\mu_{A,\theta}(t) f(\omega t + A)\Big)\partial_\theta Q^\mu_{A,\theta}(t)\ dt.$$

Integrating by parts and using also that $Q^\mu_{A,\theta}(t)$ solves (i) in lemma 4.1 we obtain

$$\partial_\theta \widetilde{F}_\mu(A,\theta) = \Big[\partial_\theta Q^\mu_{A,\theta}(t)\dot{Q}^\mu_{A,\theta}(t)\Big]_{-\infty}^{+\infty} + \int_{\mathbf{R}} \partial_\theta Q^\mu_{A,\theta}(t)\alpha^\mu_{A,\theta}\psi_\theta(t)\ dt. \tag{4.66}$$

By lemma 4.1-(iii)-(iv) the boundary term in (4.66) vanishes. Moreover, deriving w.r.t. η the second equality in lemma 4.1-(v), we get

$$\partial_\theta Q^\mu_{A,\theta}(t) = \omega \cdot \partial_A Q^\mu_{A,\theta}(t) - \dot{Q}^\mu_{A,\theta}(t). \tag{4.67}$$

Since, by lemma 4.1-(ii), $\int_{\mathbf{R}} Q^\mu_{A,\theta}\psi_\theta\ dt = \int_{\mathbf{R}} q_\theta\psi_\theta\ dt$ is independent of A, $\int_{\mathbf{R}} \partial_A Q^\mu_{A,\theta}\psi_\theta(t) = 0$. Hence

$$\begin{aligned}
\partial_\theta \widetilde{F}_\mu(A,\theta) &= \alpha^\mu_{A,\theta}\int_{\mathbf{R}} \Big(\omega \cdot \partial_A Q^\mu_{A,\theta}(t) - \dot{Q}^\mu_{A,\theta}(t)\Big)\psi_\theta(t)\ dt = -\alpha^\mu_{A,\theta}\int_{\mathbf{R}} \dot{Q}^\mu_{A,\theta}(t)\psi_\theta(t)\ dt \\
&= -\alpha^\mu_{A,\theta}\Big(\overline{K} + \int_{\mathbf{R}} (\dot{Q}^\mu_{A,\theta}(t) - \dot{q}_\theta(t))\psi_\theta\ dt\Big).
\end{aligned} \tag{4.68}$$

By lemma 4.1-(iii), noting that $|\psi_\theta| \le 1$, we have

$$\Big|\int_{\mathbf{R}} (\dot{Q}^\mu_{A,\theta}(t) - \dot{q}_\theta(t))\psi_\theta(t)\ dt\Big| \le 4C_4\mu. \tag{4.69}$$

Finally by (4.69) and (4.68), if $0 < \mu 4C_4 \le |\overline{K}|/2$, then

$$|\partial_\theta \widetilde{F}_\mu(A,\theta)| \ge |\alpha^\mu_{A,\theta}|(|\overline{K}| - 4\mu C_4) \ge |\alpha^\mu_{A,\theta}||\overline{K}|/2,$$

namely (4.64).

The proof of (4.65) can be performed with similar computations. Let us for instance justify that $\partial^2_\theta \widetilde{F}_\mu(A,\theta) = O(\mu)$. We have

$$\begin{aligned}
\partial_\theta \widetilde{F}_\mu(A,\theta) &= -\alpha^\mu_{A,\theta} \int_{\mathbf{R}} \dot{Q}^\mu_{A,\theta}(t)\psi_\theta(t)\, dt \\
&= \int_{\mathbf{R}} \ddot{Q}^\mu_{A,\theta}\dot{Q}^\mu_{A,\theta} - \sin Q^\mu_{A,\theta}\dot{Q}^\mu_{A,\theta} + \mu f(\omega t + A)\sin Q^\mu_{A,\theta}\dot{Q}^\mu_{A,\theta}\, dt \\
&= \mu \int_{\mathbf{R}} f(\omega t + A)\sin Q^\mu_{A,\theta}\dot{Q}^\mu_{A,\theta}\, dt. \qquad (4.70)
\end{aligned}$$

Hence

$$\partial^2_\theta \widetilde{F}_\mu(A,\theta) = \mu \int_{\mathbf{R}} (\cos Q^\mu_{A,\theta}\partial_\theta Q^\mu_{A,\theta}\dot{Q}^\mu_{A,\theta} + \sin Q^\mu_{A,\theta}\partial_\theta \dot{Q}^\mu_{A,\theta}) f(\omega t + A)\, dt.$$

The bound of $\partial^2_\theta \widetilde{F}_\mu(A,\theta)$ (independent of ω) is then a straightforward consequence of (4.67) and lemma 4.1-$(i)-(iii)-(iv)$. ■

The relation between the two reduced action functionals $F_\mu(A,\theta) = \Phi_{\mu,A}(q^\mu_{A,\theta})$ and $\widetilde{F}_\mu(A,\theta) = \Phi_{\mu,A}(Q^\mu_{A,\theta})$ is given below: they are the same "up to a change of variables near the identity". The next theorem is formulated to handle also non-analytic perturbations f.

Theorem 4.1 *Assume that f is smooth. There exist $\mu_2 > 0$ and $C_5 > 0$ (independent of ω) such that, for $0 < \mu < \mu_2$, there exist a Lipschitz continuous function $h_\mu : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$, with $|h_\mu(A,\theta)| \le C_5\mu$, $|h_\mu(A',\theta') - h_\mu(A,\theta)| \le C_5\mu(|A'-A| + |\theta'-\theta|)$, $h_\mu(A,\theta+\eta) = h_\mu(A+\eta\omega,\theta)$, such that*

$$F_\mu(A,\theta) = \widetilde{F}_\mu(A,\theta + h_\mu(A,\theta)). \qquad (4.71)$$

In particular, setting $g_\mu(A) = h_\mu(A,0)$, $\psi_\mu : \mathbf{T}^n \to \mathbf{T}^n$ defined by $\psi_\mu(A) = A + g_\mu(A)\omega$ is a homeomorphism provided that $C_5\mu \le 1/2$ and satisfies

$$G_\mu = \widetilde{G}_\mu \circ \psi_\mu.$$

The inverse of ψ_μ writes $\psi_\mu^{-1}(A) = A + k_\mu(A)\omega$ where k_μ is $2C_5\mu$-Lipschitz continuous. Moreover, if f is analytic and satisfies (4.60) with $r_i > 0$ for all $i = 1,\ldots,n$, then the homeomorphism ψ_μ is a real analytic diffeomorphism.

In order to prove theorem 4.1 we need the next two lemmas, proved in the appendix. The next lemma states that $Q^\mu_{A,\theta}$ belongs to the supplementary space $E_{\theta + l_\mu(A,\theta)}$ for some $l_\mu(A,\theta) = O(\mu)$.

Lemma 4.3 *For μ small enough (independently of ω) there exists a smooth function $l_\mu(A,\theta)$ with $l_\mu(A,\theta) = O(\mu), \nabla l_\mu(A,\theta) = O(\mu)$, $l_\mu(A,\theta+\eta) = l_\mu(A+\eta\omega,\theta)$ such that $Q^\mu_{A,\theta}(\theta + l_\mu(A,\theta)) = \pi$.*

Define $V_\mu(A,\theta) := F_\mu(A,\theta + l_\mu(A,\theta)) = \Phi_{\mu,A}(q^\mu_{A,\theta+l_\mu(A,\theta)})$.

Lemma 4.4 *There exists a positive constant C_6 such that, for all $(A,\theta) \in \mathbf{T}^n \times \mathbf{R}$, there holds*

$$|\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta)| \le C_6|\partial_\theta \widetilde{F}_\mu(A,\theta)|^2. \qquad (4.72)$$

In particular if $\partial_\theta \widetilde{F}_\mu(A,\theta) = 0$ then $\widetilde{F}_\mu(A,\theta) = V_\mu(A,\theta)$. In addition

$$|\nabla \widetilde{F}_\mu(A,\theta) - \nabla V_\mu(A,\theta)| \le C_6\mu|\partial_\theta \widetilde{F}_\mu(A,\theta)|. \qquad (4.73)$$

PROOF OF THEOREM 4.1. By lemma 4.3, there is a smooth function $\bar{l}_\mu$ such that $\bar{l}_\mu(A,\theta) = O(\mu)$, $\nabla \bar{l}_\mu(A,\theta) = O(\mu)$, $\bar{l}_\mu(A,\theta+\eta) = \bar{l}_\mu(A+\eta\omega,\theta)$ and $F_\mu(A,\theta) = V_\mu(A,\theta + \bar{l}_\mu(A,\theta))$. $\bar{l}_\mu(A,\theta)$ is the unique solution of the equation $x = -l_\mu(A,\theta + x)$. It is enough to find, for all θ, $y = \bar{h}_\mu(A,\theta)$ such that

$$V_\mu(A,\theta) = \widetilde{F}_\mu(A,\theta + y). \qquad (4.74)$$

Indeed, then

$$F_\mu(A,\theta) = V_\mu(A,\theta+\bar l_\mu(A,\theta)) = \widetilde F_\mu(A,(\theta+\bar l_\mu(A,\theta)) + \bar h_\mu(A,\theta+\bar l_\mu(A,\theta)))$$

and $h_\mu(A,\theta)$ will be defined by

$$h_\mu(A,\theta) = \bar l_\mu(A,\theta) + \bar h_\mu(A,\theta+\bar l_\mu(A,\theta)). \tag{4.75}$$

Note that if $\partial_\theta \widetilde F_\mu(A,\theta)=0$ then, by lemma 4.4, equation (4.74) is solved by $y=0$. In general we look for y of the form $y = \partial_\theta \widetilde F_\mu(A,\theta) z$. Then we can write

$$\begin{aligned} \widetilde F_\mu(A,\theta+y) &= \widetilde F_\mu(A,\theta) + \partial_\theta \widetilde F_\mu(A,\theta) y + R_\mu(A,\theta,y) y^2 \\ &= \widetilde F_\mu(A,\theta) + (\partial_\theta \widetilde F_\mu(A,\theta))^2 z + R_\mu(A,\theta,\partial_\theta \widetilde F_\mu(A,\theta)z)(\partial_\theta \widetilde F_\mu(A,\theta))^2 z^2, \end{aligned} \tag{4.76}$$

where

$$R_\mu(A,\theta,y) = \frac{1}{y^2}\Big[\widetilde F_\mu(A,\theta+y) - \widetilde F_\mu(A,\theta) - \partial_\theta \widetilde F_\mu(A,\theta) y\Big]$$

is smooth and, by the estimates (4.65) on the derivatives of $\widetilde F_\mu$, satisfies $R_\mu(A,\theta,y) = O(\mu)$, $\partial_y R_\mu(A,\theta,y) = O(\mu/|y|)$. By (4.76) equation (4.74) is then equivalent to

$$\frac{V_\mu(A,\theta) - \widetilde F_\mu(A,\theta)}{(\partial_\theta \widetilde F_\mu(A,\theta))^2} = z + R_\mu(A,\theta,\partial_\theta \widetilde F_\mu(A,\theta)z) z^2.$$

We have $R_\mu(A,\theta,\partial_\theta \widetilde F_\mu(A,\theta)z)z^2 = O(\mu z^2)$ and $\partial_z\Big(R_\mu(A,\theta,\partial_\theta \widetilde F_\mu(A,\theta)z)z^2\Big) = O(\mu z)$. By the contraction mapping theorem, for μ small enough, for all $u \in \mathbf{R}$ such that $|u| < 2C_6$, there exists a unique solution $z = \varphi(\mu,A,\theta,u)$ of the equation

$$u = z + R_\mu(A,\theta,\partial_\theta \widetilde F_\mu(A,\theta)z) z^2, \tag{4.77}$$

such that $|z| < 3C_6$. Moreover, the function φ defined in this way is smooth and is real analytic if $\widetilde F_\mu$ is real analytic. Setting

$$\bar h_\mu(A,\theta) := \varphi\Big(\mu, A, \theta, \frac{V_\mu(A,\theta) - \widetilde F_\mu(A,\theta)}{(\partial_\theta \widetilde F_\mu(A,\theta))^2}\Big) \partial_\theta \widetilde F_\mu(A,\theta) \tag{4.78}$$

if $\partial_\theta \widetilde F_\mu(A,\theta) \neq 0$ and $\bar h_\mu(A,\theta) = 0$ if $\partial_\theta \widetilde F_\mu(A,\theta) = 0$, we get a continuous function $\bar h_\mu$ which satisfies (4.74) and $|\bar h_\mu(A,\theta)| \le 3C_6|\partial_\theta \widetilde F_\mu(A,\theta)|$, which implies $|\bar h_\mu| = O(\mu)$. Moreover $\bar h_\mu$ is the unique function that enjoys these properties. By (4.78) the restriction of $\bar h_\mu$ to

$$U_\mu := \Big\{(A,\theta) \in \mathbf{T}^n \times \mathbf{R} \ : \ \partial_\theta \widetilde F_\mu(A,\theta) \neq 0\Big\}$$

is smooth. Deriving the identity $V_\mu(A,\theta) = \widetilde F_\mu(A,\theta+\bar h_\mu(A,\theta))$ we obtain

$$(\partial_\theta \widetilde F_\mu)(A,\theta+\bar h_\mu(A,\theta)) \nabla \bar h_\mu(A,\theta) = \nabla V_\mu(A,\theta) - \nabla \widetilde F_\mu(A,\theta+\bar h_\mu(A,\theta))$$

for $(A,\theta) \in U_\mu$. By (4.65)

$$\nabla \widetilde F_\mu(A,\theta+\bar h_\mu(A,\theta)) = \nabla \widetilde F_\mu(A,\theta) + O(\mu|\bar h_\mu(A,\theta)|) = \nabla \widetilde F_\mu(A,\theta) + O(\mu|\partial_\theta \widetilde F_\mu(A,\theta)|).$$

Hence, if $\partial_\theta \widetilde F_\mu(A,\theta) \neq 0$ then, by (4.73),

$$\nabla \bar h_\mu = \frac{\nabla V_\mu(A,\theta) - \nabla \widetilde F_\mu(A,\theta) + O(\mu|\partial_\theta \widetilde F_\mu(A,\theta)|)}{\partial_\theta \widetilde F_\mu(A,\theta)(1+O(\mu))} = O(\mu).$$

uniformly in U_μ. Since $\overline{h}_\mu$ is continuous and $\overline{h}_\mu(A,\theta)=0$, if $(A,\theta)\notin U_\mu$, the Lipschitz continuity of $\overline{h}_\mu$ follows.

Regarding the properties of h_μ, we observe that $\overline{h}_\mu(A,\theta+\eta)=\overline{h}_\mu(A+\eta\omega,\theta)$, which is a consequence of uniqueness.

Hence, by (4.75) and since $\bar{l}_\mu(A,\theta+\eta)=\bar{l}_\mu(A+\eta\omega,\theta)$, $h_\mu(A,\theta+\eta)=h_\mu(A+\eta\omega,\theta)$. Moreover, since $|\nabla\bar{l}_\mu(A,\theta)|=O(\mu)$, by the Lipschitz continuity of $\overline{h}_\mu$, there is a constant C_5 such that $|h_\mu(A,\theta)|\le C_5\mu$ and h_μ is Lipschitz continuous with ratio $C_5\mu$.

To complete the proof, we remark that for a given $A\in\mathbf{T}^n$, the equation in x, $\psi_\mu(A+x\omega)=A$ is equivalent to

$$x=-g_\mu(A+\omega x)=-h_\mu(A,x), \tag{4.79}$$

where h_μ is $C_5\mu$-Lipschitz continuous. By the contracting mapping theorem, if $C_5\mu\le 1/2$ then (4.79) has a unique solution $k_\mu(A)\in\mathbf{R}$. Moreover, one can prove without much effort that $|k_\mu(A)|\le C_5\mu$ and k_μ is Lipschitz continuous of ratio $2C_5\mu$. As a consequence ψ_μ is a homeomorphism and its inverse is defined by $\psi_\mu^{-1}(A)=A+k_\mu(A)\omega$.

The proof of the analyticity of ψ_μ when f is analytic, is given in the next remark 4.2 and in the Appendix. ■

By theorem 4.1 we deduce the following lemma which enables us to apply the shadowing theorem 2.3 once a "splitting condition" for $\widetilde{G}_\mu$ is verified.

Lemma 4.5 *Assume that $\widetilde{G}_\mu$ satisfies the "splitting condition" 2.1 for some bounded open set U, $A_0\in U$, with constants δ and α. Then, for $\mu|\omega|C_5<1/2$, G_μ too satisfies the "splitting condition" with $A_0'=\psi_\mu^{-1}(A_0)$, $U'=\psi_\mu^{-1}(U)$ and the constants $\delta'=\delta$ and $\alpha'=\alpha/2$.*

PROOF. In the appendix. ■

Remark 4.2 *Assume that $r_i>0$ for all i (i.e. that the perturbation f is analytic). Then we can prove, using the arguments of the next subsection, that the homoclinic function $G_\mu(\cdot)=F_\mu(\cdot,0)$ can be extended to a complex analytic function over the interior of D. Hence $F_\mu(A,\theta)=G_\mu(A+\omega\theta)$ can be defined in an open neighbourhood of $\mathbf{T}^n\times\mathbf{R}$ in $(\mathbf{T}^n+i\mathbf{R}^n)\times\mathbf{C}$, so that the extension is analytic. One could check that l_μ and V_μ, defined in lemma 4.3 have analytic extensions too, and that the inequality of lemma 4.4 still holds in the new set of definition. Moreover in the next lemma 4.6 it is proved that $\widetilde{F}_\mu$ is analytic w.r.t. (A,θ). As a consequence, $(V_\mu(A,\theta)-\widetilde{F}_\mu(A,\theta))/(\partial_\theta\widetilde{F}_\mu(A,\theta))^2$ is real analytic, and so is the function h_μ defined in the proof of theorem 4.1. Therefore if $r_i>0$ for all i, then the homeomorphism ψ_μ defined in theorem 4.1 is a real analytic diffeomorphism. For completeness, the proof of the claims included in this remark is spelled out in the appendix.*

4.2 Analytic extension

The unpertubed homoclinic $q_0(t)=4\arctan e^t$ can be extended to a holomorphic function over the strip $S:=\mathbf{R}+i(-\pi/2,\pi/2)$. Moreover equation (1.4) may be considered also for complex values of q and, for $\mu=0$, q_θ is a solution of (1.4) for all $\theta\in S$. The unperturbed manifold of critical points $Z=\{q_\theta\mid\theta\in S\}$ can be seen as a holomorphic curve in the complex affine space $q_\theta+H^1(\mathbf{R},\mathbf{C})$.

We want to perform the Lyapunov-Schmidt reduction on the complex strip

$$S_\sigma:=\mathbf{R}+i\Big(-(\frac{\pi}{2}-\sigma),\frac{\pi}{2}-\sigma\Big),$$

for $\sigma\in(0,\pi/2)$. We have

$$\dot q_\theta(z)=\frac{2}{\cosh(z-\theta)},\quad \ddot q_\theta(z)=\sin q_\theta(z)=-2\frac{\sinh(z-\theta)}{\cosh^2(z-\theta)},\quad (1-\cos q_\theta(z))=\frac{2}{\cosh^2(z-\theta)}.$$

Assume that $\theta \in S_\sigma, Re(\theta) = 0$. The following estimates hold, where $t \in \mathbf{R}$

$$|\dot{q}_\theta(t)| \leq \frac{C}{\min\{(|t| + \sigma), 1\}} \exp(-|t|); \tag{4.80}$$

$$|\sin q_\theta(t)| \leq \frac{C}{\min\{(|t| + \sigma)^2, 1\}} \exp(-|t|); \tag{4.81}$$

$$|\cos q_\theta(t)| \leq \frac{C}{\min\{(|t| + \sigma)^2, 1\}}; \tag{4.82}$$

$$\frac{1}{|\dot{q}_\theta(t)|} \leq C \exp(|t|) \min\{(|t| + \sigma), 1\}. \tag{4.83}$$

In what follows we consider the Banach spaces

$$X = \left\{ w \in C^2(\mathbf{R}, \mathbf{C}) \;\Big|\; \sup_t \exp(|t|/2)(|w(t)| + |\dot{w}(t)| + |\ddot{w}(t)|) < +\infty \right\}$$

and

$$\overline{X} = \left\{ w \in X \;\Big|\; w(0) = 0 \right\},$$

endowed with norm

$$||w||_{2,\sigma} = \sup_{|t|>1} \Big(|w(t)| + |\dot{w}(t)| + |\ddot{w}(t)| \Big) \exp(\frac{|t|}{2}) + \sup_{|t|<1} \Big(\frac{|w(t)|}{(|t| + \sigma)^2} + \frac{|\dot{w}(t)|}{(|t| + \sigma)} + |\ddot{w}(t)| \Big).$$

Note that the function $\psi_0(t)$ can be extended to a holomorphic function on $\mathbf{R} + i(-\pi, \pi)$. With the definition for $||f||$ given in (4.61), we have

Lemma 4.6 *There exist positive constants η, C_7 such that for all $\sigma \in (0, \pi/2)$, $\forall 0 < \mu \leq (\eta\sigma^3)/||f||$, for all ω, for all $A \in D$, for all $\theta \in S_\sigma$ there exist a unique $Q^\mu_{A,\theta} : \mathbf{R} \to \mathbf{C}$ and a unique $\alpha^\mu_{A,\theta} \in \mathbf{C}$ such that*

- *$Q^\mu_{A,\theta} = q_{\theta + \nu^\mu_{A,\theta}} + w^\mu_{A,\theta}$, where $\nu^\mu_{A,\theta} \in \mathbf{C}$, $w \in \overline{X}$ and $||w^\mu_{A,\theta}||_{2,\sigma} + |\nu^\mu_{A,\theta}| + |\alpha^\mu_{A,\theta}| \leq C_7 \mu ||f||/\sigma^2$;*
- *$-\ddot{Q}^\mu_{A,\theta}(t) + \sin Q^\mu_{A,\theta}(t) = \mu \ \sin Q^\mu_{A,\theta}(t) \ f(\omega t + A) + \alpha^\mu_{A,\theta} \psi_\theta(t)$;*
- *$\int_{\mathbf{R}} (Q^\mu_{A,\theta}(t) - q_\theta(t)) \psi_\theta(t) \ dt = 0$.*

Moreover $Q^\mu_{A,\theta}$ and $\alpha^\mu_{A,\theta}$ depend analytically on θ and on the A_i for which $r_i > 0$.

Proof. **1st step**. Let us consider the Banach space

$$Y = \left\{ v \in C(\mathbf{R}, \mathbf{C}) \;\Big|\; \sup_t |v(t)| \exp\left(\frac{|t|}{2}\right) < +\infty \right\}$$

endowed with norm $||v||_{-1,\sigma} = \sup_{|t|>1} |v(t)| \exp(\frac{|t|}{2}) + \sup_{|t|<1} (|t| + \sigma)|v(t)|$. Let $\theta \in S_\sigma$ be given once for all. We may assume without loss of generality that $Re(\theta) = 0$.

For $\theta' \in S_{\sigma/2}$ such that $|\theta' - \theta| \leq \sigma/2$ we introduce the linear operator $L_{\theta'} : \overline{X} \times \mathbf{C} \to Y$ defined by

$$L_{\theta'}(w, \alpha) = -\ddot{w} + (\cos q_{\theta'}) w - \alpha \psi_\theta.$$

Using that $\dot{q}_{\theta'}$ is a solution of $-\ddot{y} + \cos q_{\theta'} y = 0$ and that $\dot{q}_{\theta'}$ does not vanish anywhere, we can compute the inverse of $L_{\theta'}$. It is given by $L^{-1}_{\theta'}(g) = (w, \alpha)$ with

$$\alpha = -\frac{\int_{\mathbf{R}} g(t) \dot{q}_{\theta'}(t) \ dt}{\int_{\mathbf{R}} \psi_\theta(t) \dot{q}_{\theta'}(t) \ dt}, \tag{4.84}$$

$$w(t) = \dot q_{\theta'}(t)\Big[\int_0^t -\frac{1}{\dot q_{\theta'}^2(s)}\Big(\int_{-\infty}^s (g(\sigma)+\alpha\psi_\theta(\sigma))\dot q_{\theta'}(\sigma)\ d\sigma\Big)\ ds\Big] \tag{4.85}$$

$$= \dot q_{\theta'}(t)\Big[\int_0^t \frac{1}{\dot q_{\theta'}^2(s)}\Big(\int_s^{+\infty} (g(\sigma)+\alpha\psi_\theta(\sigma))\dot q_{\theta'}(\sigma)\ d\sigma\Big)\ ds\Big]. \tag{4.86}$$

Note that since $|\theta-\theta'|\le\sigma/2$, $Re(\theta')\le\sigma/2$. Therefore estimates (4.80)-(4.83) hold as well (with perhaps different constants) when θ is replaced by θ'. We derive from (4.84)-(4.86) that

$$|\alpha| + ||w||_{2,\sigma} \le \frac{C}{\sigma}||g||_{-1,\sigma}. \tag{4.87}$$

2nd Step. We shall search Q as $Q = q_{\theta+\nu} + w$ with $|\nu| < \sigma/2$, $w\in\overline{X}$. Let B denote the open ball of radius $\sigma/2$ in $\mathbf{C}$ centered at 0. Let $J_\mu : B\times\overline{X}\times\mathbf{C}\to Y\times\mathbf{C}$ be defined by

$$J_\mu(\nu,w,\alpha) = \Big(-\ddot q_{\theta+\nu} - \ddot w + \sin(q_{\theta+\nu}+w) - \mu\sin(q_{\theta+\nu}+w)\ f(\varphi) - \alpha\psi_\theta, \int_{\mathbf{R}}(q_{\theta+\nu}+w-q_\theta)\psi_\theta(t)\Big).$$

¿From now we shall use the norms $||(\nu,w,\alpha)||_2 = |\nu| + ||w||_{2,\sigma} + |\alpha|$ on $B\times\overline{X}\times\mathbf{C}$ and $||(g,\beta)||_{-1} = ||g||_{-1,\sigma} + |\beta|$ on $Y\times\mathbf{C}$. J_μ is of class C^1 and

$$\begin{aligned} DJ_\mu(\nu,w,\alpha)[z,W,a] &= \Big(-z\ \dddot q_{\theta+\nu} - \ddot W + \cos(q_{\theta+\nu}+w)(z\dot q_{\theta+\nu}+W)\\ &- \mu\cos(q_{\theta+\nu}+w)(z\dot q_{\theta+\nu}+W)\ f(\varphi) - a\psi_\theta, \int_{\mathbf{R}}(z\dot q_{\theta+\nu}+W)\psi_\theta(t)\Big). \end{aligned}$$

We shall prove that, provided $||(\nu,w,\alpha)||_2/\sigma$ and $\mu||f||/\sigma^3$ are small enough $DJ_\mu(\nu,w,\alpha)$ is invertible. We first consider the case when $w=0$ and $\mu=0$. Let $T_\nu = DJ_0(\nu,0,\alpha)$ (independent of α). Observing that $-\dddot q_{\theta+\nu} + \cos(q_{\theta+\nu})\dot q_{\theta+\nu} = 0$, we obtain

$$T_\nu[z,W,a] = \Big(-\ddot W + \cos q_{\theta+\nu}W - a\psi_\theta, \int_{\mathbf{R}}(z\dot q_{\theta+\nu}+W)\psi_\theta(t)\Big).$$

Using the first step we derive that T_ν is invertible and that, for a suitable positive constant C,

$$||T_\nu^{-1}(g,\beta)||_2 \le \frac{C}{\sigma}||(g,\beta)||_{-1}. \tag{4.88}$$

Now we estimate $||(DJ_\mu(\nu,w,\alpha)-T_\nu)[z,W,a]||_{-1}$. We have

$$\begin{aligned} (DJ_\mu(\nu,w,\alpha)-T_\nu)[z,W,a] &= \Big((\cos(q_{\theta+\nu}+w) - (\cos q_{\theta+\nu}))(z\dot q_{\theta+\nu}+W)\\ &- \mu\cos(q_{\theta+\nu}+w)(z\dot q_{\theta+\nu}+W)f(\varphi), 0\Big). \end{aligned}$$

We easily get

$$\begin{aligned} ||(DJ_\mu(\nu,w,\alpha)-T_\nu)[z,W,a]||_{-1} &\le C||w||_{2,\sigma}\Big(||W||_{2,\sigma}+|z|\Big) + \frac{\mu||f||}{\sigma^2}|z| + |\mu|||f||||W||_{2,\sigma}\\ &\le C\Big(||w||_{2,\sigma} + \frac{\mu||f||}{\sigma^2}\Big)||(z,W,a)||_2. \end{aligned}$$

As a consequence, by (4.88), if $\mu||f||/\sigma^3 \le K_0$ and $||w||_{2,\sigma}/\sigma \le K_0$, for K_0 small enough, then $DJ_\mu(\mu,w,\alpha)$ is invertible and

$$||(DJ_\mu(\nu,w,\alpha))^{-1}|| \le \frac{K_1}{\sigma}$$

for a suitable positive constant K_1.

3rd Step. We now prove the existence of a constant K_2 such that $(0,0,0)$ is the unique solution of the equation $J_0(\nu,w,\alpha)=0$ in $B(K_2\sigma)$, ball centered at the origin and of radius $K_2\sigma$ for the norm $||\ ||_2$. First we observe that, since $\ddot{q}_{\theta+\nu}=\sin(q_{\theta+\nu})$, there holds

$$J_0(\nu,w,\alpha)=T_\nu[\nu,w,\alpha]+\Big(\sin(q_{\theta+\nu}+w)-\sin(q_{\theta+\nu})-\cos(q_{\theta+\nu})w,\int_{\mathbf{R}}(q_{\theta+\nu}-q_\theta-\nu\dot{q}_{\theta+\nu})\psi_\theta\Big).$$

Moreover, by the analyticity of $q_0,\dot{q}_0,\psi_0$ over S,

$$\int_{\mathbf{R}}(q_{\theta+\nu}(t)-q_\theta(t)-\nu\dot{q}_{\theta+\nu}(t))\psi_\theta(t)\ dt=\int_{\mathbf{R}}(q_\nu(t)-q_0(t)-\nu\dot{q}_\nu(t))\psi_0(t)\ dt.$$

Hence there is a constant C' such that

$$\Big|\Big|\Big(\sin(q_{\theta+\nu}+w)-\sin(q_{\theta+\nu})-\cos(q_{\theta+\nu})w,\int_{\mathbf{R}}(q_{\theta+\nu}-q_\theta-\nu\dot{q}_{\theta+\nu})\psi_\theta\Big)\Big|\Big|_{-1}\le C'(||w||_{2,\sigma}^2+|\nu|^2).$$

So, if $J_0(\nu,w,\alpha)=0$ then, by (4.88)

$$\begin{aligned}||(\nu,w,\alpha)||_2 &= \Big|\Big|-T_\nu^{-1}\Big(\sin(q_{\theta+\nu}+w)-\sin(q_{\theta+\nu})-\cos(q_{\theta+\nu})w,\int_{\mathbf{R}}(q_{\theta+\nu}-q_\theta-\nu\dot{q}_{\theta+\nu})\psi_\theta\Big)\Big|\Big|_2\\ &\le \frac{CC'}{\sigma}||(\nu,w,\alpha)||_2^2.\end{aligned}$$

Let $K_2<1/(CC')$. By the latter inequality, if $J_0(\nu,w,\alpha)=0$ and $||(\nu,w,\alpha)||_2\le K_2\sigma$, then $\nu=0$, $w=0$, $\alpha=0$.

4th step. By the previous steps we know that there exist positive constants K_0, K_1 and K_2 such that

- (i) ($J_0(\nu,w,\alpha)=0$ and $||(\nu,w,\alpha)||_2\le K_2\sigma$) $\iff$ $\nu=w=\alpha=0$;
- (ii) If $|\nu|<\sigma/2$, $||w||_{2,\sigma}\le K_0\sigma$, $\mu||f||\le K_0\sigma^3$ then $DJ_\mu(\nu,w,\alpha)$ is invertible and $||(DJ_\mu(\nu,w,\alpha))^{-1}||\le K_1/\sigma$.

Moreover there exists a constant $K_3>0$ such that

- (iii) $||\partial_\mu J_\mu(\nu,w,\alpha)||_{-1}=\Big|\Big|(\sin(q_{\theta+\nu}+w)f(\varphi),0)\Big|\Big|_{-1}\le ||f||K_3/\sigma$.

We say that $(i),(ii),(iii)$ imply that there is η such that, for all $0<\mu<\eta\sigma^3/||f||$, the equation $J_\mu(\nu,w,\alpha)=0$ has a unique solution such that $||(\nu,w,\alpha)||_2<K_2\sigma/2$. In addition $||(\nu,w,\alpha)||_2=O(\mu||f||/\sigma^2)$. To prove existence, we can proceed as follows. Let $\mathcal{S}$ denote the set of all $\mu\in[0,K_0\sigma^3/||f||]$ such that there exists a C^1 function $S_\mu:[0,\mu]\to\{(\nu,w,\alpha)\ :\ ||(\nu,w,\alpha)||_2<K_2\sigma/2\}$ such that $S_\mu(0)=0$, $J_t(S_\mu(t))=0$ for all $t\in[0,\mu]$. $\mathcal{S}$ is a bounded interval. Let us call $\overline{\mu}$ its supremum. By (ii) and the Implicit Function Theorem, $\overline{\mu}>0$. In addition, for $\mu\in\mathcal{S}$, there is a unique function S_μ with the required properties. As a consequence, for $0<\mu<\mu'$, $S_\mu=S_{\mu'|[0,\mu]}$ and we can define a C^1 function $S:[0,\overline{\mu})\to\{(\nu,w,\alpha)\ :\ ||(\nu,w,\alpha)||_2<K_2\sigma/2\}$ such that $S(t)=S_\mu(t)$ for all $\mu\in(0,\overline{\mu})$. By (ii) and (iii), we can write, for all $t\in(0,\overline{\mu})$,

$$||S'(t)||_2=\Big|\Big|\Big[DJ_t(S(t))\Big]^{-1}\cdot\Big(\frac{\partial J_t}{\partial t}(S(t))\Big)\Big|\Big|_2\le\frac{K_1K_3||f||}{\sigma^2}.$$

Hence

$$||S(t)||_2\le\frac{K_1K_3}{\sigma^2}||f||\,|t|. \tag{4.89}$$

Now, since $S'(t)$ is bounded, $S(t)$ converges to some $\overline{S}$ as $t\to\overline{\mu}$. Either $\overline{\mu}=K_0\sigma^3/||f||$ or $||\overline{S}||_2=K_2\sigma/2$ (If not, by the Implicit Function Theorem, we could extend the solution S to an interval $[0,\overline{\mu}+\xi)$, $\xi>0$, contradicting the definition of $\overline{\mu}$). In the latter case, by (4.89),

$$||\overline{S}||_2=\frac{K_2\sigma}{2}\le\frac{K_1K_3}{\sigma^2}\overline{\mu}||f||.$$

So the existence assertion holds for $0<\mu<\eta\sigma^3/||f||$, where $\eta=\min(K_0,K_2/(2K_1K_3))$.

In order to prove uniqueness, we assume that there are b_1,b_2 such that $||b_i||_2<K_2\sigma/2$, $J_\mu(b_i)=0$. Then, by the same argument as previously, we can prove the existence of two functions of class C^1 $S_1,S_2\;:\;[0,\mu]\to\{b\;:\;||b||_2<K_2\sigma\}$ such that $S_i(\mu)=b_i$, $J_t(S_i(t))=0$. Moreover, by (ii) and the Implicit Function Theorem, $S_1(\mu)\neq S_2(\mu)$ implies that $S_1(t)\neq S_2(t)$ for all $t\in[0,\mu]$, which contradicts (i), proving uniqueness.

The bound of $||w^\mu_{A,\theta}||_{2,\sigma}+|\nu^\mu_{A,\theta}|+|\alpha^\mu_{A,\theta}|$ given in the statement is a direct consequence of (4.89).

To complete the proof, we point out that J_μ is complex differentiable w.r.t all its variables. Therefore, as a consequence of the Implicit Function Theorem (see for example [2]), $Q^\mu_{A,\theta}=q_{\theta+\nu^\mu(A,\theta)}+w^\mu(A,\theta)$ depends analytically on θ and on A_i if $r_i>0$. ■

We now consider the analytic extension of the function $\widetilde{F}_\mu$ defined for $(A,\theta)\in D\times S_\sigma$ by

$$\widetilde{F}_\mu(A,\theta)=\Phi_{\mu,A}(Q^\mu_{A,\theta})=\int_{\mathbf{R}}\frac{(\dot{Q}^\mu_{A,\theta})^2(t)}{2}+(1-\cos Q^\mu_{A,\theta}(t))+\mu(\cos Q^\mu_{A,\theta}(t)-1)f(\omega t+A)\;dt.$$

Let us consider also the analytic extension for $(A,\theta)\in D\times S_\sigma$ of the Melnikov function

$$M(A,\theta)=\int_{\mathbf{R}}(\cos q_\theta(t)-1)f(\omega t+A)\;dt.$$

We have $M(A,\theta)=\Gamma(A+\omega\theta)$, where Γ is defined in (2.43). We now prove

Lemma 4.7 *For $\mu||f||\sigma^{-3}$ small enough, we have*

$$\widetilde{F}_\mu(A,\theta)=b+\mu M(A,\theta)+O\Big(\frac{\mu^2||f||^2}{\sigma^4}\Big). \tag{4.90}$$

uniformly for $(A,\theta)\in D\times S_\sigma$, where $b=\Phi_0(q_0)$.

PROOF. We have $Q^\mu_{A,\theta}=q_{\theta+\nu^\mu_{A,\theta}}+w^\mu_{A,\theta}$ which we denote for brevity $Q^\mu_{A,\theta}=q_{\theta+\nu}+w$.

$$\begin{aligned}
\widetilde{F}_\mu(A,\theta)\;&=\;\int_{\mathbf{R}}\frac{(\dot{q}_{\theta+\nu}+\dot{w})^2}{2}+(1-\cos(q_{\theta+\nu}+w))+\mu(\cos(q_{\theta+\nu}+w)-1)f(\omega t+A)\;dt\\
&=\;b+\int_{\mathbf{R}}\Big[-\ddot{q}_{\theta+\nu}w+\frac{1}{2}\dot{w}^2+(\cos q_{\theta+\nu}-\cos(q_{\theta+\nu}+w))\\
&+\;\mu(\cos(q_{\theta+\nu})-1)f(\omega t+A)+\mu(\cos(q_{\theta+\nu}+w)-\cos q_{\theta+\nu})f(\omega t+A)\Big]\;dt\\
&=\;b+\mu M(\theta+\nu,A)+\int_{\mathbf{R}}\frac{1}{2}\dot{w}^2+\Big(-\cos(q_{\theta+\nu}+w)+\cos q_{\theta+\nu}-\sin q_{\theta+\nu}w\Big)\;dt\\
&+\;\mu\int_{\mathbf{R}}\Big(\cos(q_{\theta+\nu}+w)-\cos q_{\theta+\nu}\Big)f(\omega t+A)\;dt.
\end{aligned}$$

By the estimate $||w||_{2,\sigma}\le C\mu||f||/\sigma^2$, it follows easily

$$\widetilde{F}_\mu(A,\theta)=b+\mu M(A,\theta+\nu)+O\Big(\frac{\mu^2||f||^2}{\sigma^4}\Big).$$

For example we can get that $\int_{\mathbf{R}}\cos q_{\theta+\nu}-\cos(q_{\theta+\nu}+w)-(\sin q_{\theta+\nu})w=O(\mu^2||f||^2/\sigma^4)$ by writing $\cos q_{\theta+\nu}-\cos(q_{\theta+\nu}+w)-(\sin q_{\theta+\nu})w=w^2\int_0^1(1-s)\cos(q_{\theta+\nu}+sw)\;ds$ and using (4.81)-(4.82) together with $||w||_{2,\sigma}\le\mu||f||/\sigma^2$. Moreover

$$|M(A,\theta+\nu)-M(A,\theta)|=O\Big(\frac{|\nu|}{\sigma^2}\Big)=O\Big(\frac{\mu||f||}{\sigma^4}\Big),$$

which completes the proof of the lemma. ■

The Fourier coefficients of the Melnikov function $\Gamma(A)=\sum_k \Gamma_k \exp(ikA)$ are explicitely given by

$$\Gamma_k = f_k \frac{2\pi(k\cdot\omega)}{\sinh((k\cdot\omega)\frac{\pi}{2})}\ (k\neq 0), \quad \Gamma_0 = 4f_0 \tag{4.91}$$

Since $\widetilde{F}_\mu(A,\theta)=\widetilde{G}_\mu(A+\omega\theta)$ and $M(A,\theta)=\Gamma(A+\omega\theta)$, we obtain, via a standard argument on Fourier coefficients of analytic functions, the following result

Theorem 4.2 *There exists a positive constant C_8 such that, if $\mu||f||\sigma^{-3}$ is small enough, then $\forall k \in \mathbf{Z}^n\backslash\{0\}$, for all $\sigma\in(0,\frac{\pi}{2})$, for all ω,*

$$|\widetilde{G}_k-\mu\Gamma_k| \leq \frac{C_8\mu^2||f||^2}{\sigma^4}\exp\Big(-\sum_{i=1}^n r_i|k_i|\Big)\exp\Big(-|k\cdot\omega|\Big(\frac{\pi}{2}-\sigma\Big)\Big). \tag{4.92}$$

Proof. $\widetilde{F}_\mu(A,\theta)-b-\mu M(A,\theta)=\widetilde{G}_\mu(A+\omega\theta)-b-\mu\Gamma(A+\omega\theta)$ is analytic in $D\times S_\sigma$ and, by (4.90), its modulus is bounded by $C_8\mu^2/(||f||^2\sigma^4)$ where C_8 is a positive contant. By lemma 3 in [21], (4.92) follows. ■

Remark 4.3 *Compare with theorem 3.4.5 in [36] which gives the exponentially small estimate of the Fourier coefficients.*

5 Three time scales

We consider in this section three-time-scales systems, introduced and investigated in connection with the problem of Arnold Diffusion in [15] (see also [25] and [38]), as

$$\mathcal{H}=\frac{I_1}{\sqrt{\varepsilon}}+\varepsilon^a\beta\cdot I_2+\frac{p^2}{2}+(\cos q-1)+\mu(\cos q-1)f(\varphi_1,\varphi_2), \quad \varepsilon>0,\ a\geq 0$$

with $n\geq 2$, $\varphi_1\in\mathbf{T}^1$, $\varphi_2\in\mathbf{T}^{n-1}$, $I_1\in\mathbf{R}^1$, $I_2\in\mathbf{R}^{n-1}$, $\beta\in\mathbf{R}^{n-1}$ and ε is a positive small parameter. The frequency vector is $\omega=(1/\sqrt{\varepsilon},\varepsilon^a\beta)$, where $\beta=(\beta_2,\dots,\beta_n)\in\mathbf{R}^{n-1}$ is fixed.

We assume through this section that $\mu||f||\varepsilon^{-3/2}$ and ε are small.

Given $\kappa_2=(k_2,\dots,k_n)\in\mathbf{Z}^{n-1}$, we shall use the notation $\kappa_2^+:=(|k_2|,\dots,|k_n|)$. Moreover we shall use the abbreviation $\rho_2:=(r_2,\dots,r_n)$, so that $\kappa_2^+\cdot\rho_2:=\sum_{i=2}^n r_i|k_i|$. We recall that $r_1,\dots,r_n$ are defined in formula (4.60). Writing

$$f(\varphi_1,\varphi_2)=\sum_{(k_1,\kappa_2)\in\mathbf{Z}\times\mathbf{Z}^{n-1}} f_{k_1,\kappa_2}\exp(i(k_1\varphi_1+\kappa_2\cdot\varphi_2)),$$

we assume that f is analytic w.r.t φ_2. More precisely for $i\geq 2$, $r_i>0$, whereas r_1 may be zero in (4.60). If $a=0$, we impose in addition that $r_i>|\beta_i|\pi/2$ for $i\geq 2$.

We shall use (4.92) in order to give an expansion for the "homoclinic function"

$$\widetilde{G}_\mu(A)=\sum_{(k_1,\kappa_2)\in\mathbf{Z}\times\mathbf{Z}^{n-1}}\widetilde{G}_{k_1,\kappa_2}\exp(i(k_1A_1+\kappa_2\cdot A_2))=\sum_{k_1\in\mathbf{Z}}\widetilde{g}_{k_1}(A_2)\exp(ik_1A_1).$$

where

$$\widetilde{g}_{k_1}(A_2)=\frac{1}{2\pi}\int_0^{2\pi}\widetilde{G}_\mu(A_1,A_2)\exp(-ik_1A_1)dA_1=\sum_{\kappa_2\in\mathbf{Z}^{n-1}}\widetilde{G}_{k_1,\kappa_2}\exp(i\kappa_2\cdot A_2).$$

We start with

Lemma 5.1 *There exists a positive constant C_9 such that, for $\mu||f||\varepsilon^{-3/2}$ small enough,*

$$\Big|\sum_{|k_1|\geq 2}\widetilde{g}_{k_1}(A_2)\exp(ik_1A_1)\Big| \leq \sum_{\kappa_2\in\mathbf{Z}^{n-1},|k_1|\geq 2}|\widetilde{G}_{k_1,\kappa_2}| \leq C_9\frac{\mu||f||}{\sqrt{\varepsilon}}\exp(-\frac{\pi}{\sqrt{\varepsilon}}). \tag{5.93}$$

PROOF. Choosing $\sigma=\sqrt{\varepsilon}$ and recalling that by (4.60) $|f_k|\leq||f||e^{-\kappa_2^+\cdot\rho_2}$, we get from (4.91) and (4.92) that

$$\begin{aligned}
|\widetilde{G}_{k_1,\kappa_2}| &\leq \mu|\Gamma_{k_1,\kappa_2}|+|\widetilde{G}_{k_1,\kappa_2}-\mu\Gamma_{k_1,\kappa_2}|\\
&\leq C\mu||f||e^{-\kappa_2^+\cdot\rho_2}\Big(\Big|\frac{k_1}{\sqrt{\varepsilon}}+\kappa_2\cdot\beta\varepsilon^a\Big|+1\Big)e^{-|\frac{k_1}{\sqrt{\varepsilon}}+\kappa_2\cdot\beta\varepsilon^a|\pi/2}\\
&+ C\frac{\mu^2}{\varepsilon^2}||f||^2e^{-\kappa_2^+\cdot\rho_2}e^{-|\frac{k_1}{\sqrt{\varepsilon}}+\kappa_2\cdot\beta\varepsilon^a|(\pi/2-\sqrt{\varepsilon})}
\end{aligned}$$

Since $|\frac{k_1}{\varepsilon}+\kappa_2\cdot\beta\varepsilon^a|\geq|\frac{k_1}{\varepsilon}|-|\kappa_2\cdot\beta\varepsilon^a|$ we get,

$$\begin{aligned}
|\widetilde{G}_{k_1,\kappa_2}| &\leq C\frac{\mu||f||}{\sqrt{\varepsilon}}(|k_1|+|\kappa_2|)e^{-\kappa_2^+\cdot\rho_2+|\kappa_2\cdot\beta|\varepsilon^a\pi/2}\ e^{-\frac{|k_1|}{\sqrt{\varepsilon}}\pi/2}\\
&+ C\frac{\mu^2}{\varepsilon^2}||f||^2e^{-\kappa_2^+\cdot\rho_2+|\kappa_2\cdot\beta|\varepsilon^a(\pi/2-\sqrt{\varepsilon})}\ e^{-|k_1|(\frac{\pi}{2\sqrt{\varepsilon}}-1)}\\
&\leq C\frac{\mu||f||}{\sqrt{\varepsilon}}(|k_1|+|\kappa_2|)\exp(-\sum_{j=2}^{n}|k_j|(r_j-|\beta_j|\varepsilon^a\pi/2))\exp(-|k_1|(\frac{\pi}{2\sqrt{\varepsilon}}-1)).
\end{aligned}$$

We have used in the last line that $\mu||f||/\varepsilon^{3/2}=O(1)$. Now $r_j-|\beta_j|\varepsilon^a\pi/2>0$ for ε small enough both if $a=0$ or if $a>0$. Summing in $|k_1|>2$ and in $\kappa_2\in\mathbf{Z}^{n-1}$ we obtain (5.93). ■

The Poincaré-Melnikov primitive defined in (2.43) can be written as

$$\Gamma(\varepsilon,A)=\sum_{(k_1,\kappa_2)\in\mathbf{Z}\times\mathbf{Z}^{n-1}}\Gamma_{k_1,\kappa_2}\exp(i(k_1A_1+\kappa_2\cdot A_2))=\sum_{k_1\in\mathbf{Z}}\Gamma_{k_1}(\varepsilon,A_2)\exp(ik_1A_1).$$

Lemma 5.2 *Define*

$$R_0(\varepsilon,\mu,A_2)=\widetilde{g}_0(A_2)-b-\mu\Gamma_0(\varepsilon,A_2)$$

and

$$R_1(\varepsilon,\mu,A_2)=\widetilde{g}_{\pm1}(A_2)-\mu\Gamma_{\pm1}(\varepsilon,A_2).$$

The following estimates hold

$$R_0(\varepsilon,\mu,A_2)=O\Big(\mu^2||f||^2\Big)\quad\text{and}\quad R_1(\varepsilon,\mu,A_2)=O\Big(\frac{\mu^2||f||^2}{\varepsilon^2}\exp(-\frac{\pi}{2\sqrt{\varepsilon}})\Big).$$

PROOF. We get immediately the estimate for $R_0(\varepsilon,\mu,A_2)$ by taking the mean value w.r.t. $A_1\in\mathbf{T}$ in formula (4.90) where $\theta=0$, $A_2\in\mathbf{T}^{n-1}$ and $\sigma=\pi/4$.

We now prove the estimate for $R_1(\varepsilon,\mu,A_2)$. By (4.92) (where we choose $\sigma=\sqrt{\varepsilon}$ and $k_1=\pm1$), we can obtain as in the proof of lemma 5.1

$$\begin{aligned}
|\widetilde{g}_{\pm1}(A_2)-\mu\Gamma_{\pm1}(\varepsilon,A_2)| &\leq C\frac{\mu^2}{\varepsilon^2}||f||^2\sum_{\kappa_2\in\mathbf{Z}^{n-1}}\exp(-\sum_{j=2}^{n}|k_j|(r_j-|\beta_j|\varepsilon^a\pi/2))\exp(-(\frac{\pi}{2\sqrt{\varepsilon}}-1).\\
&\leq C\frac{\mu^2}{\varepsilon^2}||f||^2e^{-\frac{\pi}{2\sqrt{\varepsilon}}}.
\end{aligned}$$

■

Since $\Gamma(A)$ and $G_\mu(A)$ are real valued functions we have that $\widetilde{g}_{-1}(A_2)=\overline{\widetilde{g}}_1(A_2)$ and $\Gamma_{-1}(A_2)=\overline{\Gamma}_1(A_2)$, where $\overline{z}$ denotes the complex conjugate of the complex number z. We deduce from the previous two lemmas the following result.

Theorem 5.1 *For $\mu||f||\varepsilon^{-3/2}$ small there holds*

$$\begin{aligned}\widetilde{G}_\mu(A_1,A_2) &= \widetilde{g}_0(A_2)+2\mathrm{Re}\left[\widetilde{g}_1(A_2)e^{iA_1}\right]+\sum_{|k_1|\geq 2}\widetilde{g}_{k_1}(A_2)\exp(ik_1A_1)\\ &= b+\Big(\mu\Gamma_0(\varepsilon,A_2)+R_0(\varepsilon,\mu,A_2)\Big)+2\mathrm{Re}\left[\Big(\mu\Gamma_1(\varepsilon,A_2)+R_1(\varepsilon,\mu,A_2)\Big)e^{iA_1}\right]\\ &+ O\Big(\mu\varepsilon^{-1/2}||f||\exp(-\frac{\pi}{\sqrt{\varepsilon}})\Big)\end{aligned}$$

where

$$R_0(\varepsilon,\mu,A_2)=O\Big(\mu^2||f||^2\Big)\quad\text{and}\quad R_1(\varepsilon,\mu,A_2)=O\Big(\frac{\mu^2||f||^2}{\varepsilon^2}\exp(-\frac{\pi}{2\sqrt{\varepsilon}})\Big).$$

Remark 5.1 *(i) This improves the results in [38] which require $\mu=\varepsilon^p$ with $p>2+a$.*

(ii) Theorem 5.1 certainly holds in any dimension, while the results of [25], which hold for more general systems, are proved for 2 rotators only.

(iii) In order to prove a splitting condition using theorem 5.1 it is necessary, accordingly with [25] and [38], that $\exists m,l\in\mathbf{Z}^{n-1}$ such that $f_{0,l},f_{1,m}\neq 0$. If not, recalling (4.91), there results that $\Gamma_0(\varepsilon,A_2)=\sum_{\kappa_2\in\mathbf{Z}^{n-1}}\Gamma_{0,\kappa_2}\exp^{i\kappa_2\cdot A_2}=0$ and also $\Gamma_1(\varepsilon,A_2)=\sum_{\kappa_2\in\mathbf{Z}^{n-1}}\Gamma_{1,\kappa_2}\exp^{i\kappa_2\cdot A_2}=0$.

Theorem 5.1 enables us to provide conditions implying the existence of diffusion orbits. For instance we obtain the following result.

Lemma 5.3 *Assume that there are $\overline{A}_2\in\mathbf{R}^{n-1}$ and $d_0,c_0>0$ such that, for all small $\varepsilon>0$,*

$$\begin{array}{ll}(i) & |\Gamma_1(\varepsilon,A_2)|>(c_0/\sqrt{\varepsilon})e^{-\pi/(2\sqrt{\varepsilon})},\quad\forall A_2\in\mathbf{R}^{n-1}\ \text{ such that }\ |A_2-\overline{A}_2|\leq d_0;\\ (ii) & \Gamma_0(\varepsilon,A_2)>\Gamma_0(\varepsilon,\overline{A}_2)+c_0,\quad\forall A_2\in\mathbf{R}^{n-1},\ \text{ such that }\ |A_2-\overline{A}_2|=d_0.\end{array}$$

Then there is $\nu(c_0,d_0)>0$ and a constant $K>0$ such that, for all ε,μ with $\varepsilon,\mu||f||\varepsilon^{-3/2}<\nu(c_0,d_0)$, condition 2.1 is satisfied by $\widetilde{G}_\mu$ for some bounded open set $U\in\mathbf{R}^n$ and $\delta=(c_0\mu/2\sqrt{\varepsilon})e^{-\pi/(2\sqrt{\varepsilon})}$, $\alpha=Kc_0e^{-\pi/(2\sqrt{\varepsilon})}/2\sqrt{\varepsilon}$, $\rho_U\leq 3d_0$.

Proof. First we can derive from (4.91) and (4.92) in the same way as in the proof of lemmas 5.1 and 5.2 that

$$\begin{aligned}|\widetilde{g}_1(A_2)|+|\nabla\widetilde{g}_1(A_2)| &\leq \sum_{\kappa_2\in\mathbf{Z}^{n-1}}(1+|\kappa_2|)|\widetilde{G}_{1,\kappa_2}|\\ &\leq K_0\frac{\mu||f||}{\sqrt{\varepsilon}}e^{-\pi/(2\sqrt{\varepsilon})}\end{aligned}\tag{5.94}$$

for some constant K_0. By the bounds of R_0 and R_1 in theorem 5.1, for ε and $\mu||f||\varepsilon^{-3/2}$ small enough, we have

$$\begin{array}{ll}(a) & |\widetilde{g}_1(A_2)|=|(\mu\Gamma_1+R_1)(A_2)|>(\mu c_0/(2\sqrt{\varepsilon}))e^{-\pi/(2\sqrt{\varepsilon})}\quad\forall A_2\in B_{d_0}\\ (b) & \widetilde{g}_0(A_2)=(\mu\Gamma_0+R_0)(A_2)>(\mu\Gamma_0+R_0)(\overline{A}_2)+(\mu c_0)/2\quad\forall A_2\in\partial B_{d_0},\end{array}$$

where B_{d_0} is the open ball centered at $\overline{A}_2$ of radius d_0.

So we can write $\widetilde{g}_1(A_2)=|\widetilde{g}_1(A_2)|e^{i\phi(A_2)}$, where ϕ is a smooth real function defined in B_{d_0} : (5.94) and the previous lower bound of $|\widetilde{g}_1(A_2)|$ provide a bound of $\nabla\phi(A_2)$. Precisely,

$$|\nabla\phi(A_2)|\leq\frac{|\nabla\widetilde{g}_1(A_2)|}{|\widetilde{g}_1(A_2)|}\leq 2||f||\frac{K_0}{c_0}.\tag{5.95}$$

For $A_2\in B_{d_0}$, by theorem 5.1 we have

$$\widetilde{G}_\mu(A_1,A_2)=b+(\mu\Gamma_0+R_0)(\varepsilon,\mu,A_2)+2|(\mu\Gamma_1+R_1)(\varepsilon,\mu,A_2)|\cos(A_1+\phi(A_2))+O(\mu\varepsilon^{-1/2}||f||\exp(-\frac{\pi}{\sqrt{\varepsilon}})).$$

Let

$$U=\Big\{(A_1,A_2)\in\mathbf{R}\times\mathbf{R}^{n-1}\;\Big|\;A_2\in B_{d_0},\;|A_1+\phi(A_2)-\pi|<\frac{\pi}{2}\Big\}.$$

We now prove that $\widetilde{G}_\mu$ satisfies point (i) of condition 2.1, with $\delta=c_0\mu e^{-\pi/2\sqrt{\varepsilon}}/2\sqrt{\varepsilon}$ and $A_0=(\overline{A}_1,\overline{A}_2)=(\pi-\phi(\overline{A}_2),\overline{A}_2)$. Assume that $(A_1,A_2)\in\partial U$. Either $A_2\in\partial B_{d_0}$ and then

$$\widetilde{G}_\mu(A_1,A_2)-\widetilde{G}_\mu(\pi-\phi(\overline{A}_2),\overline{A}_2)\geq\frac{\mu c_0}{2}+O\Big(\mu\varepsilon^{-1/2}||f||\exp(-\frac{\pi}{2\sqrt{\varepsilon}})\Big),\tag{5.96}$$

or $|A_1+\phi(A_2)-\pi|=\pi/2$ and then

$$\begin{aligned}\widetilde{G}_\mu(A_1,A_2)-\widetilde{G}_\mu(\pi-\phi(A_2),A_2)&=2|(\mu\Gamma_1+R_1)(A_2)|+O\Big(\mu\varepsilon^{-1/2}||f||\exp(-\frac{\pi}{\sqrt{\varepsilon}})\Big)\\&\geq\frac{\mu c_0}{\sqrt{\varepsilon}}e^{-\pi/(2\sqrt{\varepsilon})}+O\Big(\mu\varepsilon^{-1/2}||f||\exp(-\frac{\pi}{\sqrt{\varepsilon}})\Big).\end{aligned}\tag{5.97}$$

Hence, by (5.96) and (5.97), for ε and $\mu||f||\varepsilon^{-3/2}$ small enough,

$$\inf_{\partial U}\widetilde{G}_\mu\geq\inf_U\widetilde{G}_\mu+\frac{c_0\mu e^{-\pi/2\sqrt{\varepsilon}}}{2\sqrt{\varepsilon}}.\tag{5.98}$$

Thus $(\widetilde{G}_\mu,U)$ satisfies point (i) of condition 2.1, with $\delta=c_0\mu e^{-\pi/2\sqrt{\varepsilon}}/2\sqrt{\varepsilon}$. Now $|\nabla\widetilde{G}_\mu|=O(\mu)$, hence there exists a positive constant K_1 such that, for all $A,B\in\mathbf{R}^n$, $|\widetilde{G}_\mu(A)-\widetilde{G}_\mu(B)|\leq K_1\mu|A-B|$. Let $\alpha=K\delta=Kc_0e^{-\pi/2\sqrt{\varepsilon}}/2\sqrt{\varepsilon}$, where $K=1/8K_1$. Then

$$\forall(A,B)\in\mathbf{R}^n\times\mathbf{R}^n\quad|A-B|\leq2\alpha\Longrightarrow|\widetilde{G}_\mu(A)-\widetilde{G}_\mu(B)|\leq\frac{\delta}{4}.\tag{5.99}$$

Let $A^0=(A_1^0,A_2^0)\in U$ be such that $\inf_U\widetilde{G}_\mu=\widetilde{G}_\mu(A^0)$. Then by (5.99) and (5.98), $\sup_{B_\alpha(A^0)}\widetilde{G}_\mu\leq\inf_U\widetilde{G}_\mu+\delta/4$, and the ball $B_\alpha(A^0)$ is included in U. So (ii) is satisfied with the above choice of α.

That (iii) is satisfied is a straightforward consequence of (5.99).

To complete the proof, we observe that, by (5.95), $|\phi(A_2)-\phi(\overline{A}_2)|\leq d_0K_0||f||/c_0$ for all $A_2\in B_{d_0}$. Hence any $A\in U$ can be written as

$$A=(\pi-\phi(\overline{A}_2),\overline{A}_2)+(l_1,l_2),\qquad|l_1|\leq\pi+\frac{d_0K_0}{c_0}||f||,\quad|l_2|\leq d_0.$$

Now, since $(l_1,l_2)=\sqrt{\varepsilon}l_1\omega+(0,l_2-\varepsilon^{a+1/2}l_1\beta)$,

$$|\Pi_\omega(l_1,l_2)|=|\Pi_\omega(0,l_2-\varepsilon^{a+1/2}l_1\beta)|\leq|l_2-\varepsilon^{a+1/2}l_1\beta|\leq d_0+O(\sqrt{\varepsilon}).$$

Hence $\rho_U\leq2d_0+O(\sqrt{\varepsilon})$, and for ε small enough, $\rho_U\leq3d_0$. ■

The condition given in the previous lemma is not easily handable. We now want to provide simpler conditions, involving properties of the perturbation f. For $A=(A_1,A_2)\in\mathbf{T}^1\times\mathbf{T}^{n-1}$, let

$$f(A_1,A_2)=\sum_{(k_1,\kappa_2)\in\mathbf{Z}\times\mathbf{Z}^{n-1}}f_{k_1,\kappa_2}\exp(i(k_1A_1+\kappa_2\cdot A_2))=\sum_{k_1\in\mathbf{Z}}f_{k_1}(A_2)\exp(ik_1A_1).$$

Since f is analytic w.r.t A_2, $f_{k_1}(A_2)=(1/2\pi)\int_0^{2\pi}f(\sigma,A_2)e^{-ik_1\sigma}\,d\sigma$ depends analytically of A_2.

Theorem 5.2 *Assume that f satisfies one of the following conditions:*
(i) $a>0$, $f_0(A_2)$ admits a strict local minimum at the point $\overline{A}_2$ and $f_1(\overline{A}_2)\neq0$
(ii) $a=0$, $f_0(A_2)$ admits a strict local minimum at the point $\overline{A}_2$ and $f_1(\overline{A}_2+i(\pi/2)\beta)\neq0$.
Then, for all small ε such that $\omega_\varepsilon=(1/\sqrt{\varepsilon},\beta\varepsilon^a)$ satisfies

$$|\omega_\varepsilon\cdot k|\geq\frac{\gamma_\varepsilon}{|k|^\tau},\;\forall k\in\mathbf{Z}^n,k\neq0$$

for all I_0, I_0' with $\omega_\varepsilon \cdot I_0 = \omega_\varepsilon \cdot I_0'$, for all $\mu||f||\varepsilon^{-3/2}$ small enough, there is a heteroclinic orbit of the perturbed three-time-scales system, connecting the invariant tori $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I_0'}$. In addition, for all $\eta > 0$ small enough the "diffusion time" T_d needed to go from a η-neighbourhood of $\mathcal{T}_{I_0}$ to a η-neighbourhood of $\mathcal{T}_{I_0'}$ is $O(|I_0 - I_0'|(\sqrt{\varepsilon}/\mu)e^{\pi/(2\sqrt{\varepsilon})}[(\gamma_\varepsilon)^{-1}(\sqrt{\varepsilon}e^{\pi/(2\sqrt{\varepsilon})})^\tau + |\log\mu|] + |\log(\eta)|)$.

PROOF. It is enough to prove that, if (i) or (ii) is satisfied, then the condition given in lemma 5.3 holds for some $c_0, d_0 > 0$. The statement is then a direct consequence of theorem 2.3 and lemma 4.5.

We first assume that condition (i) is satisfied. In what follows, the notation $u = O(v)$ means that $|u| \le C|v|$, where C is a universal constant. We have

$$\begin{aligned}
\Gamma_0(\varepsilon, A_2) &= \sum_{\kappa_2 \in \mathbf{Z}^{n-1}} \frac{2\pi\kappa_2 \cdot \beta\varepsilon^a}{\sinh(\pi\kappa_2 \cdot \beta\varepsilon^a/2)} f_{0,\kappa_2} e^{i\kappa_2 \cdot A_2} \\
&= \sum_{\kappa_2 \in \mathbf{Z}^{n-1}} (4 + O(\varepsilon^{2a}|\kappa_2|^2)) f_{0,\kappa_2} e^{i\kappa_2 \cdot A_2} \\
&= 4f_0(A_2) + O\Big(\sum_{\kappa_2 \in \mathbf{Z}^{n-1}} \varepsilon^{2a}|\kappa_2|^2 e^{-\kappa_2^+ \cdot \rho_2} \Big) \\
&= 4f_0(A_2) + O(\varepsilon^{2a}).
\end{aligned}$$

Moreover

$$\begin{aligned}
\Gamma_1(\varepsilon, A_2) &= \sum_{\kappa_2 \in \mathbf{Z}^{n-1}} \frac{2\pi(\varepsilon^{-1/2} + \kappa_2 \cdot \beta\varepsilon^a)}{\sinh((\pi/2)(\varepsilon^{-1/2} + \kappa_2 \cdot \beta\varepsilon^a))} f_{1,\kappa_2} e^{i\kappa_2 \cdot A_2} \\
&= \sum_{\kappa_2 \in \mathbf{Z}^{n-1}} \frac{4\pi}{\sqrt{\varepsilon}} e^{-(\pi/2)(\varepsilon^{-1/2} + \kappa_2 \cdot \beta\varepsilon^a)} \Big(1 + O\Big((|\kappa_2| + 1)\varepsilon^{a+(1/2)}\Big)\Big) f_{1,\kappa_2} e^{i\kappa_2 \cdot A_2}.
\end{aligned}$$

Since, by (4.60), $|f_{1,\kappa_2}| = O(e^{-\kappa_2^+ \cdot \rho_2})$, we have

$$\begin{aligned}
\Gamma_1(\varepsilon, A_2) &= \frac{4\pi}{\sqrt{\varepsilon}} e^{-(\pi/2)\varepsilon^{-1/2}} \Big[f_1(A_2) + O\Big(\sum_{\kappa_2 \in \mathbf{Z}^{n-1}} e^{-\kappa_2^+ \cdot \rho_2} \Big(|e^{-(\pi/2)\kappa_2 \cdot \beta\varepsilon^a} - 1| + (|\kappa_2| + 1)\varepsilon^{a+(1/2)} e^{|\kappa_2 \cdot \beta|\varepsilon^a} \Big) \Big) \Big] \\
&= \frac{4\pi}{\sqrt{\varepsilon}} e^{-(\pi/2)\varepsilon^{-1/2}} \Big[f_1(A_2) + O\Big(\sum_{\kappa_2 \in \mathbf{Z}^{n-1}} (|\kappa_2| + 1)\varepsilon^a \exp\Big(- \sum_{j=2}^{n} |k_j|(r_j - (\pi/2)|\beta_j|\varepsilon^a) \Big) \Big) \Big] \\
&= \frac{4\pi}{\sqrt{\varepsilon}} e^{-(\pi/2)\varepsilon^{-1/2}} \Big[f_1(A_2) + O(\varepsilon^a) \Big],
\end{aligned}$$

provided that ε is small enough. Now, since $f_1(\overline{A}_2) \neq 0$ and f_0 admits a strict local minimum at $\overline{A}_2$, by the continuity of f_1, there is $d_0 > 0$ such that $|f_1(A_2)| > |f_1(\overline{A}_2)|/2$ for all $|A_2 - \overline{A}_2| \le d_0$ and $\Delta(d_0) := \inf_{|A_2 - \overline{A}_2| = d_0} f_0(A_2) - f_0(\overline{A}_2) > 0$. Hence, for ϵ small enough, the assumptions of lemma 5.3 hold with $c_0 = \min(\pi|f_1(\overline{A}_2)|, 2\Delta(d_0))$.

We now assume that condition (ii) is satisfied. As previously, we have

$$\begin{aligned}
\Gamma_1(\varepsilon, A_2) &= \sum_{\kappa_2 \in \mathbf{Z}^{n-1}} \frac{4\pi}{\sqrt{\varepsilon}} e^{-(\pi/2)(\varepsilon^{-1/2} + \kappa_2 \cdot \beta)} \Big(1 + O((|\kappa_2| + 1)\varepsilon^{1/2})\Big) f_{1,\kappa_2} e^{i\kappa_2 \cdot A_2} \\
&= \frac{4\pi}{\sqrt{\varepsilon}} e^{-(\pi/2)\varepsilon^{-1/2}} \Big[f_1(A_2 + i(\pi/2)\beta) + O\Big(\sum_{\kappa_2 \in \mathbf{Z}^{n-1}} (|\kappa_2| + 1)\varepsilon^{1/2} \exp(-\sum_{j=2}^{n} |k_j|(r_j - (\pi/2)|\beta_j|)) \Big) \Big] \\
&= \frac{4\pi}{\sqrt{\varepsilon}} e^{-(\pi/2)\varepsilon^{-1/2}} \Big[f_1(A_2 + i(\pi/2)\beta) + O(\sqrt{\varepsilon}) \Big].
\end{aligned}$$

We observe also that, if $a = 0$, then $\Gamma_0(\varepsilon, A_2) = 4f_0(A_2)$ is independent of ε. It follows easily that condition (ii) as well implies that the assumption of lemma 5.3 holds true. ■

Remark 5.2 *In the time estimate obtained in theorem 5.2, τ is the exponent of an exponetially large number, since, in order to apply the shadowing theorem 2.3, which relies on condition 2.1, the diffusion orbit must approach to the homoclinic point at an exponentially small distance (namely $\alpha = O(\exp -(\pi/2\sqrt{\varepsilon}))$. This result can be improved via a shadowing theorem suited to the case when the splitting is exponentially small in one direction only. This allows to find diffusion orbits which approach to the homoclinic point just at a polynomially small distance in the orthogonal directions to ω. This is proved in the forthcoming paper [7]. Moreover, for this reason, we are able to obtain diffusion orbits which drift in polynomial time w.r.t $1/\varepsilon$ along the fast directions $(I_2, \ldots, I_n)$ (no diffusion in I_1), see [7].*

6 Appendix

In the proof of the following lemmas we will closely follow the arguments developed in the papers [4]-[5] to which we refer for further details. In the sequel the notation $u = O(v)$ (resp. $u = o(v)$) will mean that there is a constant C (resp. a function $\varepsilon(v)$) independent of anything except f such that $|u| \leq C|v|$ (resp. $|u| \leq \varepsilon(v)|v|$ and $\lim_{v\to 0} \varepsilon(v) = 0$).

PROOF OF LEMMA 2.1. We first assume that $\theta = 0$ and give the existence proof in $[0, +\infty)$. We are looking for a solution of (1.4) in the form of $q = q_0 + w$ with $w(0) = 0$ and $\lim_{t\to+\infty} w(t) = 0$. The function w must satisfy the equation

$$-\ddot{w} + w = -\Big(\sin(q_0 + w) - \sin q_0 - w\Big) + \mu \sin(q_0 + w) f(\omega t + A).$$

Let

$$\mathbf{X} = \Big\{ w(\cdot) \in W^{1,\infty}([0, +\infty)) \;\Big|\; ||w||_1 := \sup_{t\in\mathbf{R}} \max(|w(t)|, |\dot{w}(t)|) \exp(\frac{|t|}{2}) < +\infty \Big\}$$

and

$$\mathbf{X}' = \Big\{ w(\cdot) \in L^{\infty}([0, +\infty)) \;\Big|\; ||w||_0 := \sup_{t\in\mathbf{R}} |w(t)| \exp(\frac{|t|}{2}) < +\infty \Big\}.$$

$\mathbf{X}$ and $\mathbf{X}'$, endowed respectively with norms $||\ ||_1$ and $||\ ||_0$, are Banach spaces. Let $\mathcal{L}_0$ be the linear operator which assigns to $h \in \mathbf{X}'$ the unique solution $u = \mathcal{L}_0 h$ of the problem:

$$\begin{cases} -\ddot{u} + u = h \\ u(0) = 0\ ,\ \lim_{t\to+\infty} u(t) = 0. \end{cases}$$

An explicit computation shows that, for $t \in [0, +\infty)$,

$$u(t) = (\mathcal{L}_0 h)(t) = \frac{1}{2} \int_0^{+\infty} \Big(e^{-|t-s|} - e^{-(t+s)}\Big) h(s) ds. \tag{6.100}$$

As an easy consequence $\mathcal{L}_0$ sends $\mathbf{X}'$ into $\mathbf{X}$ continuously.

We define the non-linear operator $H : \mathbf{R} \times \mathbf{R}^n \times \mathbf{X} \to \mathbf{X}$ by

$$H(\mu, A, w) := w - \mathcal{L}_0\Big(-\Big(\sin(q_0 + w) - \sin q_0 - w\Big) + \mu \sin(q_0 + w) f(\omega t + A)\Big). \tag{6.101}$$

H is smooth, $2\pi\mathbf{Z}^n$-periodic w.r.t. A and we have $H(0, A, 0) = 0$. The unknown w must solve the equation $H(\mu, A, w) = 0$. We can apply the Implicit Function Theorem. In fact, let us check that

$$\partial_w H(0, A, 0) :\ W \to W - \mathcal{L}_0\Big[(1 - \cos q_0)W\Big]$$

is invertible. Since $\lim_{t\to\infty}(1 - \cos q_0(t)) = 0$, $\partial_w H(0, A, 0)$ is of the type"Identity + Compact" and therefore it is sufficient to show that it is injective. W is in the kernel of $\partial_w H(0, A, 0)$ iff $W(0) = 0$ and W satisfies in $(0, +\infty)$ the equation

$$-\ddot{W} + \cos q_0 W = 0. \tag{6.102}$$

Multiplying by $\dot{q}_0$ in (6.102) and integrating over $[0,+\infty)$ by parts twice we obtain that $\dot{W}(0)\dot{q}_0(0)=0$. Since $\dot{q}_0(0)\neq 0$ we get also $\dot{W}(0)=0$ and as a consequence $W=0$. Thus the kernel of $\partial_w H(0,A,0)$ is reduced to 0, and this operator is invertible. We derive by the Implicit Function Theorem that there are $\rho_0>0$ and $\mu_0>0$ such that, for all $|\mu|<\mu_0$, for all $A\in\mathbf{R}^n$, the equation $H(\mu,A,w)=0$ has a unique solution w_A^μ in $\mathbf{X}$ such that $||w_A^\mu||<\rho_0$.

Note that μ_0 and ρ_0 may be chosen independent of A (and of ω too) because $\partial_w H(0,A,0)$ is independent of A and ω, $\partial_\mu H(0,A,0)$ is uniformly bounded, and $\partial_w H(\mu,A,w)$ (resp. $\partial_\mu H(\mu,A,w)$) tend to $\partial_w H(0,A,0)$ (resp. $\partial_\mu H(0,A,0)$) as $(\mu,w)\to(0,0)$ uniformly in (A,ω).

Since H is smooth w_A^μ depends smoothly on μ and A and $w_{A+2\pi k}^\mu=w_A^\mu$ by the $2\pi\mathbf{Z}^n$-periodicity of H w.r.t. A. By the properties of $\partial_\mu H$ mentioned above, $||w_A^\mu||_1=O(\mu)$.

In a similar way we can prove the existence and uniqueness of $w'^\mu_A:(-\infty,0]\to\mathbf{R}$ which satisfies analogous properties over the interval $(-\infty,0]$. We can define $q_{A,0}^\mu$ by $q_{A,0}^\mu(t)=q_0(t)+w_A^\mu(t)$ if $t\geq 0$, $q_{A,0}^\mu=q_0(t)+w'^\mu_A(t)$ if $t<0$. This is the unique function for which (i), (ii) (with $\theta=0$) hold.

If $\theta\neq 0$, we observe that q satisfies (i) iff

$$\begin{cases} -(T_{-\theta}q)''+\sin(T_{-\theta}q)=\mu\sin(T_{-\theta}q)f(\omega t+A+\omega\theta)\\ (T_{-\theta}q)(0)=\pi,\end{cases}$$

where $T_{-\theta}q(t)=q(t+\theta)$. Hence there is a unique $q_{A,\theta}^\mu$ which satisfies (i), (ii), defined by $q_{A,\theta}^\mu=T_\theta q_{A+\omega\theta,0}^\mu$, i.e. $q_{A,\theta}^\mu(t)=q_{A+\omega\theta,0}^\mu(t-\theta)$; (iii) and (iv) clearly hold. The regularity of $q_{A,\theta}^\mu$ w.r.t. A,μ is a consequence of the regularity of w_A^μ and w'^μ_A w.r.t. A and μ. (v) follows from

$$\partial_A w_A^\mu=-\Big[\partial_w H(\mu,A,w_A^\mu)\Big]^{-1}(\partial_A H)(\mu,A,w_\mu^A)$$

provided we can justify that $||(\partial_A H)(\mu,A,w_A^\mu)||_1=O(\mu)$, $\quad ||\omega\cdot\partial_A H(\mu,A,w_A^\mu)||_1=O(\mu)$. The second bound (uniform in ω) is less obvious. We just point out that

$$\omega\cdot\partial_A H(\mu,A,w_A^\mu)=-\mathcal{L}_0(\mu\sin(q_0+w_A^\mu)\frac{d}{dt}f(\omega t+A))$$

and that we can use the "regularizing" properties of $\mathcal{L}_0$. ■

Proof of lemma 2.4. We give the proof in the interval $[\theta_1,\theta_2]$. We may assume without loss of generality that $\theta_1=0$ since, by the remark at the end of the proof of lemma 2.1, a translation of the time by $-\theta_1$ amounts to adding $\omega\theta_1$ to A. For simplicity of notations, we shall write $\theta_2=\theta$.

We are looking for a solution $q=q_{0,\theta}^*+w$ of (1.4) over $[0,\theta]$ with $w(0)=w(\theta)=0$, where $q_{0,\theta}^*$ is the following smooth "approximate solution"

$$q_{0,\theta}^*(t)=\begin{cases} q_{A,0}^\mu(t) \text{ if } t\in(0,\theta/2-1),\\ r_\theta^*(t) \text{ if } t\in[\theta/2-1,\theta/2+1],\\ 2\pi+q_{A,\theta}^\mu(t) \text{ if } t\in(\theta/2+1,\theta),\end{cases}$$

where

$$r_\theta^*(t)=(1-R(t-\theta/2))q_{A,0}^\mu(t)+R(t-\theta/2)(q_{A,\theta}^\mu(t)+2\pi),$$

and $R:\mathbf{R}\to[0,1]$ is a C^∞ function such that $R(s)=0$ if $s\leq -1$, $R(s)=1$ if $s\geq 1$. Let $\mathcal{L}_{0,\theta}$ be the linear operator which assigns to $h\in L^\infty([0,\theta])$ the unique solution $u=\mathcal{L}_{0,\theta}h$ of the problem:

$$\begin{cases} -\ddot{u}+u=h\\ u(0)=0\ ,\ u(\theta)=0.\end{cases} \tag{6.103}$$

An explicit computation shows that for $t\in[0,\theta]$ the solution u of (6.103) is given by

$$u(t)=\frac{1}{\sinh(\theta)}\Big[\int_0^t h(s)\sinh(s)\sinh(\theta-t)\,ds\ +\ \int_t^\theta h(s)\sinh(\theta-s)\sinh(t)\,ds\Big].$$

Note that $\mathcal{L}_{0,\theta}$ sends $L^\infty([0,\theta])$ into $W^{1,\infty}([0,\theta])$ ($W^{2,\infty}([0,\theta])$ in fact) and that there is a constant K independent of θ such that $||\mathcal{L}_{0,\theta}W||_{1,\infty} \le K||W||_\infty$, where $||\ ||_\infty$ denotes the infty norm in $[0,\theta]$ and $||W||_{1,\infty} := ||W||_\infty + ||\dot{W}||_\infty$.

We define the smooth non-linear operator $H^\theta : \mathbf{R} \times \mathbf{R}^n \times W^{1,\infty}([0,\theta]) \to W^{1,\infty}([0,\theta])$ by

$$H^\theta(\mu, A, w) := w - \mathcal{L}_{0,\theta}\Big(- \Big(\sin(q^*_{0,\theta} + w) - \ddot{q}^*_{0,\theta} - w \Big) + \mu \sin(q^*_{0,\theta} + w) f(\omega t + A) \Big).$$

We immediately remark for further purpose that

$$||\partial^2_w H^\theta(\mu, A, w)[W, W]|| = O(||W||^2_\infty). \tag{6.104}$$

Moreover, by lemma 2.1-(i) and the definition of $q^*_{0,\theta}$, $|| - \sin q^*_{0,\theta} + \ddot{q}^*_{0,\theta} + \mu \sin(q^*_{0,\theta}) f(\omega t + A)||_\infty = O(\exp(-\theta/2))$ hence

$$||H^\theta(\mu, A, 0)||_{1,\infty} = O(\exp(-\theta/2)). \tag{6.105}$$

$q^*_{0,\theta} + w$ is a solution of (1.4) with the appropriate boundary conditions iff $H^\theta(\mu, A, w) = 0$. We shall show that there exist $\overline{C}, \overline{L}, \overline{\mu} > 0$ such that $\forall \theta > \overline{L}$, for all $|\mu| < \overline{\mu}$, for all A and ω, $\partial_w H^\theta(\mu, A, 0)$ is invertible and

$$\Big|\Big| \left(\partial_w H^\theta(\mu, A, 0)\right)^{-1} \Big|\Big| \le \overline{C}. \tag{6.106}$$

Since $\partial_w H^\theta(\mu, A, 0)$ is of the type "Id + Compact", it is enough to prove that

$$\forall W \in W^{1,\infty}([0,\theta]) \quad ||\partial_w H^\theta(\mu, A, 0)W||_{1,\infty} \ge \frac{1}{\overline{C}} ||W||_{1,\infty}.$$

We shall just sketch the proof of this assertion (see also lemma 2 of [4]). Arguing by contradiction, we assume that there are sequences $(\mu_n) \to 0$, $(\theta_n) \to \infty$, (A_n), (ω_n), (W_n) such that $W_n \in W^{1,\infty}([0,\theta_n])$, $||W_n||_{1,\infty} = 1$,

$$||\partial_w H^{\theta_n}(\mu_n, A_n, 0)W_n||_{1,\infty} \to 0. \tag{6.107}$$

Let $\xi_n \in [0, \theta_n]$ be such that $m_n := \max_{t \in [0,\theta_n]} |W_n(t)| = W_n(\xi_n)$. By (6.107) and the properties of $\mathcal{L}_{0,\theta}$, $||W_n||_{1,\infty} = O(m_n)$. Hence $\liminf(m_n) > 0$. Taking a subsequence, we may assume that (ξ_n) is bounded or $(\theta_n - \xi_n)$ is bounded or ($(\xi_n) \to \infty$ and $(\theta_n - \xi_n) \to \infty$).

In the first case, still up to a subsequence $W_n \to W \neq 0$ uniformly in compact subsets of $[0, \infty)$. Taking limits in (6.107) we obtain that $W(0) = 0$, $-\ddot{W} + \cos q_0 W = 0$, which contradicts $W \neq 0$. The second case can be dealt with similarly. In the third case, up to a subsequence, $W_n(\cdot + \xi_n) \to W \neq 0$ uniformly in compact subsets of $\mathbf{R}$, with $|W(t)| \le |W(0)|$ for all $t \in \mathbf{R}$. Taking limits in (6.107), we obtain that $-\ddot{W} + W = 0$ over $\mathbf{R}$, which contradicts $W \neq 0$ bounded.

¿From now we shall assume that $|\mu| < \mu_0 \le \overline{\mu}$, $\theta > \overline{L}$. Let

$$R^\theta(\mu, A, w) = H^\theta(\mu, A, w) - H^\theta(\mu, A, 0) - \partial_w H^\theta(\mu, A, 0)w.$$

By the previous assertion,

$$H^\theta(\mu, A, w) = 0 \quad \Leftrightarrow \quad w = -\left(\partial_w H^\theta(\mu, A, , 0)\right)^{-1} H^\theta(\mu, A, 0) - \left(\partial_w H^\theta(\mu, A, 0)\right)^{-1} R^\theta(\mu, A, w) := F^\theta_{\mu,A}(w).$$

We just have to show that $F^\theta_{\mu,A}$ is a contraction in some ball $B(0,\rho) \subset W^{1,\infty}([0,\theta])$. For this, we derive from (6.104) and (6.105) in a standard way that, for all $||w||_{1,\infty}, ||w'||_{1,\infty} \le \rho$, $|\mu| < \overline{\mu}$, $\theta > \overline{L}$ there holds

$$||F^\theta_{\mu,A}(w)||_{1,\infty} = O(\exp(-\theta/2) + \rho^2) \quad ; \quad ||F^\theta_{\mu,A}(w) - F^\theta_{\mu,A}(w')|| = O(\rho||w' - w||). \tag{6.108}$$

We can deduce that $F^\theta_{\mu,A}$ is a contraction in $\overline{B}(0,\rho)$, with $\rho = C\exp(-\theta/2)$, for some constant C, provided that $\theta > \overline{L}$, $\overline{L}$ large enough. Applying the Contraction Mapping Theorem we conclude that there is a unique solution $||w^L_\mu(A,\theta)||_{1,\infty} \le C\exp(-\theta/2)$ of the equation $H^\theta_{\mu,A}(w) = 0$. Note that by (6.108) uniqueness holds in $B(0,\rho_0)$ for some $\rho_0 > 0$ independent of θ. The regularity of the solutions in (A, θ, μ) follows as in [4]. ■

PROOF OF REMARK 2.3. Assume with no loss of generality that $A_0 = 0$ and that $D^2G_\mu(0) = \mathrm{diag}\,\{\lambda_1, \ldots, \lambda_n\}$ (this can be always obtained by an orthogonal change of variables). Assume also that $\lambda_1 = \min \lambda_i$ and that $\lambda_n = \max \lambda_i$. Let $Q(A) = \frac{1}{2}(\lambda_1 A_1^2 + \ldots + \lambda_n A_n^2)$. By the Taylor expansion

$$G_\mu(A) = Q(A) + \frac{1}{6}D^3G_\mu(\overline{A})A^3, \tag{6.109}$$

for a suitable $\overline{A}$ with $|\overline{A}| < |A|$. For $A \in B_\rho(0)$, with $\rho = \lambda_1/(3\sup_{A\in\mathbf{T}^n}|D^3G_\mu|)$, there holds $|(1/6)D^3G_\mu(\overline{A})A^3| \le (1/9)Q(A)$, hence, by (6.109)

$$\frac{8}{9}Q(A) \le G_\mu(A) \le \frac{10}{9}Q(A). \tag{6.110}$$

If we choose $U = B_\rho(0)$ condition (i) is then satisfied with $\delta = 4\lambda_1\rho^2/9$ (we have $\inf_{B_\rho(0)} G_\mu = 0$). Moreover, by (6.110) we have that

$$\Big\{G_\mu \le \frac{\delta}{2}\Big\} \subset \Big\{Q(A) \le \frac{9\delta}{16}\Big\}$$

and

$$\Big\{G_\mu \ge \frac{3\delta}{4}\Big\} \subset \Big\{Q(A) \ge \frac{27\delta}{40}\Big\}.$$

One can easily compute that $\mathrm{dist}\,(\{Q(A) \le \frac{9\delta}{16}\}, \{Q(A) \ge \frac{27\delta}{40}\}) = ((\sqrt{6}-\sqrt{5})/\sqrt{10})\rho\sqrt{\lambda_1/\lambda_n} = 2\alpha$, hence $\mathrm{dist}\,(\{G_\mu \le \frac{\delta}{2}\}, \{G_\mu \ge \frac{3\delta}{4}\}) \ge 2\alpha$.

Condition (ii) holds as well since, by (6.110), $\{G_\mu(A) \le \delta/4\}$ contains $\{Q(A) \le \delta/5\}$, and $B_\alpha(0) \subset \{Q(A) \le \delta/5\}$. ■

PROOF OF LEMMA 4.3. Let us consider the function $H : \mathbf{R} \times \mathbf{T}^n \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ defined by

$$H(\mu, A, \theta, l) = Q^\mu_{A,\theta}(\theta + l) - \pi.$$

The unknown $l_\mu(A,\theta)$ can be implicitely defined by the equation $H(\mu, A, \theta, l) = 0$. We have $H(0, A, \theta, 0) = 0$ and

$$\partial_l H(0, A, \theta, 0) = \dot{q}_\theta(\theta) = \dot{q}_0(0) \neq 0.$$

Hence by the Implicit function theorem, for μ small enough (independently of A, θ, ω because $\partial_l H$ and $\partial_\mu H$ are continuous uniformly in A, θ with a modulus of continuity independent of ω), there exists a unique smooth solution $l_\mu(A,\theta) = O(\mu)$ of $H(\mu, A, \theta, l) = 0$. Moreover, by the uniform estimates in A and ω that we can obtain for $\partial_A Q^\mu_{A,\theta}$, $\omega \cdot \partial_A Q^\mu_{A,\theta}$ and by (4.67), $|\partial_A H| + |\partial_\theta H| = O(\mu)$. Hence there holds $|\nabla l_\mu(A,\theta)| = O(\mu)$. ■

PROOF OF LEMMA 4.4. The first step is to prove that

$$\max\Big(|q^\mu_{A,\theta+l_\mu(A,\theta)}(t) - Q^\mu_{A,\theta}(t)|, |\dot{q}^\mu_{A,\theta+l_\mu(A,\theta)}(t) - \dot{Q}^\mu_{A,\theta}(t)|\Big) \le K_0|\partial_\theta \widetilde{F}_\mu(A,\theta)|\exp(-\frac{|t-\theta|}{2}), \forall t \in \mathbf{R}. \tag{6.111}$$

By lemma 2.1-(iv) we have $q^\mu_{A,\theta+l_\mu(A,\theta)} = T_{\theta+l_\mu(A,\theta)} q^\mu_{A',0}$; $Q^\mu_{A,\theta} = T_{\theta+l_\mu(A,\theta)} Q^\mu_{A',\theta'}$, where $A' = A + \omega(\theta + l_\mu(A,\theta))$ and $\theta' = -l_\mu(A,\theta)$. So it is enough to prove the estimate for $w := Q^\mu_{A',\theta'} - q^\mu_{A',0}$.

Note that $Q^\mu_{A',\theta'}(0) = q^\mu_{A',0}(0) = \pi$. So $Q^\mu_{A',\theta'} - q_0$, $q^\mu_{A',0} - q_0$ belong to $\mathbf{X}$ and satisfy

$$H(\mu, A', q^\mu_{A',0} - q_0) = 0 \quad ; \quad H(\mu, A', Q^\mu_{A',\theta'} - q_0) = \alpha^\mu_{A',\theta'}\mathcal{L}_0(\psi_{\theta'}),$$

where $\mathbf{X}, H$ and $\mathcal{L}_0$ are defined in the proof of lemma 2.1. Therefore

$$\alpha^\mu_{A',\theta'}\mathcal{L}_0(\psi_{\theta'}) = \partial_w H(\mu, A', q^\mu_{A',0} - q_0)(Q^\mu_{A',\theta'} - q^\mu_{A',0}) + o(||Q^\mu_{A',\theta'} - q^\mu_{A',0}||_1).$$

Moreover $||Q^\mu_{A',\theta'} - q_0||_1 + ||q^\mu_{A',0} - q_0||_1 = O(\mu)$. Hence, by the properties of H mentioned in the proof of lemma 2.1 (in particular the fact that $\partial_w H(0, A', 0)$ is invertible) we obtain, for μ small enough the following bound:

$$||Q^\mu_{A',\theta'} - q^\mu_{A',0}||_1 = O(|\alpha^\mu_{A',\theta'}|).$$

Since, by lemma 4.2, $|\alpha^{\mu}_{A',\theta'}| = O(|\partial_\theta \widetilde{F}_\mu(A,\theta)|)$ we deduce estimate (6.111).

We can now estimate $\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta)$. Consider the action functional $\Phi_{\mu,A'}$ defined in (1.6). For $|\mu| < \mu_0$,

$$D^2\Phi_{\mu,A'}(q)[w,w] = \int_{\mathbf{R}} \dot{w}^2 + \cos(q)w^2 - \mu\cos(q)w^2 f(A'+\omega t)\, dt = O(||w||_1^2).$$

By the definition of $q^{\mu}_{A',0}$ ((i) in lemma 2.1), we easily obtain with an integration by parts that $D\Phi_{\mu,A'}(q^{\mu}_{A',0})w = 0$ for all $w \in \mathbf{X}$ such that $w(0) = 0$. Therefore

$$\Phi_{\mu,A'}(q^{\mu}_{A',0} + w) = \Phi_{\mu,A'}(q^{\mu}_{A',0}) + O(||w||_1^2)$$

for all $w \in \mathbf{X}$ such that $w(0) = 0$. Hence since $(Q^{\mu}_{A',\theta'} - q^{\mu}_{A',0})(0) = 0$,

$$\begin{aligned}\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta) &= \widetilde{F}_\mu(A,\theta + l_\mu(A,\theta) + \theta') - F_\mu(A,\theta + l_\mu(A,\theta)) = \widetilde{F}_\mu(A',\theta') - F_\mu(A',0)\\ &= \Phi_{\mu,A'}(Q^{\mu}_{A',\theta'}) - \Phi_{\mu,A'}(q^{\mu}_{A',0}) = O(||Q^{\mu}_{A',\theta'} - q^{\mu}_{A',0}||_1^2).\end{aligned}$$

We obtain by (6.111) that

$$|\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta)| = O\Big(\partial_\theta \widetilde{F}_\mu(A,\theta)^2\Big),$$

namely (4.72). We now prove (4.73). As a consequence of (4.70), (2.13) and (6.111), we have

$$\Big|(\partial_\theta F_\mu)(A,\theta + l_\mu(A,\theta)) - \partial_\theta \widetilde{F}_\mu(A,\theta)\Big| = O(\mu|\partial_\theta \widetilde{F}_\mu(A,\theta)|). \tag{6.112}$$

In the same way, using that

$$\nabla_A F_\mu(A,\theta) = \mu\int_{\mathbf{R}} \partial_\varphi f(\omega t + A)(\cos(q^{\mu}_{A,\theta}(t)) - 1)\, dt,$$

and an analogous expression for $\nabla_A \widetilde{F}_\mu(A,\theta)$, we get

$$\Big|(\nabla_A F_\mu)(A,\theta + l_\mu(A,\theta)) - \nabla_A \widetilde{F}_\mu(A,\theta)\Big| = O(\mu|\partial_\theta \widetilde{F}_\mu(A,\theta)|). \tag{6.113}$$

We compute

$$\begin{aligned}\nabla_A V_\mu(A,\theta) - \nabla_A \widetilde{F}_\mu(A,\theta) &= (\nabla_A F_\mu)(A,\theta + l_\mu(A,\theta)) - \nabla_A \widetilde{F}_\mu(A,\theta) + (\partial_\theta F_\mu)(A,\theta + l_\mu(A,\theta))\nabla_A l_\mu(A,\theta)\\ &= O(\mu|\partial_\theta \widetilde{F}_\mu(A,\theta)|) + O(\mu|(\partial_\theta F)(A,\theta + l_\mu(A,\theta)) - \partial_\theta \widetilde{F}_\mu(A,\theta)|),\end{aligned}$$

by (6.113) and since $|\nabla_A l_\mu(A,\theta)| = O(\mu)$. By (6.112) we get

$$\Big|\nabla_A V_\mu(A,\theta) - \nabla_A \widetilde{F}_\mu(A,\theta)\Big| = O\Big(\mu|\partial_\theta \widetilde{F}_\mu(A,\theta)|\Big).$$

The estimate

$$\Big|\partial_\theta V_\mu(A,\theta) - \partial_\theta \widetilde{F}_\mu(A,\theta)\Big| = O\Big(\mu|\partial_\theta \widetilde{F}_\mu(A,\theta)|\Big)$$

is obtained in a similar way. ■

Proof of lemma 4.5. By theorem 4.1,for $C_5\mu|\omega| \le 1/2$, for all $A, A' \in \mathbf{R}^n$ there holds

$$\begin{aligned}|(\psi_\mu)(A) - (\psi_\mu)(A')| &= |A - A' + (g_\mu(A) - g_\mu(A'))\omega|\\ &\le |A - A'| + |g_\mu(A) - g_\mu(A')||\omega| \le |A - A'| + C_5\mu|\omega||A - A'|\\ &\le 3|A - A'|/2.\end{aligned}$$

Hence, for all $A, A' \in \mathbf{R}^n$,

$$|\psi_\mu^{-1}(A) - \psi_\mu^{-1}(A')| \ge \frac{2}{3}|A - A'| \ge \frac{1}{2}|A - A'|. \tag{6.114}$$

We now verify that the image of $B_\alpha(A_0)$ under the homeomorphism ψ_μ^{-1} contains the ball $B_{\alpha/2}(\psi_\mu^{-1}(A_0))$ of radius $\alpha' = \alpha/2$ centered at $\psi_\mu^{-1}(A_0)$. Since ψ_μ^{-1} is a homeomorphism, $\partial\psi_\mu^{-1}B_\alpha(A_0) = \psi_\mu^{-1}(\partial B_\alpha(A_0))$. Let $A \in \partial B_\alpha(A_0)$, namely $|A - A_0| = \alpha$. Then by (6.114) $|\psi_\mu^{-1}(A) - \psi_\mu^{-1}(A_0)| \geq \alpha/2$ and then $\mathrm{dist}\{\partial\psi_\mu^{-1}B_\alpha(A_0), \psi_\mu^{-1}(A_0)\} > \alpha/2$. This means that $B_{\alpha/2}(\psi_\mu^{-1}(A_0)) \subset \psi_\mu^{-1}(B_\alpha)$. Recalling that $\widetilde{G}_\mu = G_\mu \circ \psi_\mu^{-1}$ we have that $\inf_{\psi_\mu^{-1}(U)} G_\mu = \inf_U \widetilde{G}_\mu$, $\inf_{\partial\psi_\mu^{-1}(U)} G_\mu = \inf_{\partial U} \widetilde{G}_\mu$, $\sup_{\psi_\mu^{-1}(B_\alpha)} G_\mu = \sup_{B_\alpha} \widetilde{G}_\mu$ and then we easily deduce (i)-(ii). Properties (iii) follows as well since by (6.114) the homeomorphism ψ_μ^{-1} reduces the distance between sets at most by a factor $1/2$. ■

Proof of remark 4.2. We shall use in this proof the results of subsection 4.2.

(i) We first prove that $F_\mu(A,\theta)$ has a complex analytic extension. Let

$$Y = \Big\{ w \in W^{1,\infty}(\mathbf{R},\mathbf{C}) \;\Big|\; ||w|| := \sup_{t\in\mathbf{R}} (|w(t)| + |\dot w(t)|) e^{|t|/2} < \infty \Big\}$$

and $\overline{Y} = \{w \in Y \mid w(0) = 0\}$. We assume that f has an analytic extension defined in an open neighbourhood of

$$D_r := \mathbf{T}^n + i[-r,r]^n$$

for some $r > 0$ and we introduce $|f|_r := \sup_{A\in D_r} |f(A)| < \infty$. Let $\Phi_\mu : D_r \times Y \to \mathbf{C}$ be defined by

$$\Phi_\mu(A,w) = \int_{-\infty}^{\infty} \frac{1}{2}(\dot q_0(t) + \dot w(t))^2 + (1 - \cos(q_0(t) + w(t))) - \mu(1 - \cos(q_0(t) + w(t)))f(A + \omega t)\, dt.$$

Φ_μ is a smooth functional. Furthermore, since $\cos : \mathbf{C} \to \mathbf{C}$ is analytic, Φ_μ is complex differentiable, which means that its differential $D\Phi_\mu$ is $\mathbf{C}$-linear everywhere. We shall denote by $\overline{\Phi}_\mu$ the restriction of Φ_μ to $D_r \times \overline{Y}$; $w \in \overline{Y}$ is a critical point of $\overline{\Phi}_\mu(A,\cdot)$ iff the restriction of $q := q_0 + w$ to each interval $(-\infty,0)$ and $(0,\infty)$ is smooth and satisfies $-\ddot q(t) + \sin q(t) = \mu \sin q(t) f(A + \omega t)$.

Note that, in analogy with the real case, if $w \in \overline{Y}$ is a solution of the linear equation $-\ddot w + (\cos q_0)w = 0$ in each interval $(-\infty,0)$ and $(0,\infty)$, then $w = 0$.

As a consequence we can apply the Implicit Function Theorem and obtain in the same way as in the proof of lemma 2.1 that there is $\rho > 0$ such that, for $\mu|f|_r$ small enough, for all $A \in D_r$ there is a unique $q^\mu_{A,0} = q_0 + w^\mu_A$ in $q_0 + \overline{Y}$ such that $||w^\mu_A|| < \rho$ and

$$-\ddot q^\mu_{A,0}(t) + \sin q^\mu_{A,0}(t) = \mu \sin q^\mu_{A,0} f(A + \omega t) \quad \text{in} \quad (-\infty,0) \quad \text{and} \quad (0,\infty), \quad q^\mu_A(0) = \pi.$$

Moreover $q^\mu_{A,0}$ depends smoothly on A.

For later purpose, we remark that, in connection with the Implicit Function Theorem, there are two constants $\rho', C > 0$ such that, if $Q \in q_0 + \overline{Y}$ is a solution of $-\ddot q + \sin q = \mu \sin q f(A + \omega t) = g(t)$ in each interval $(-\infty,0)$ and $(0,+\infty)$ with $g \in Y$, $||g|| < \rho'$ then

$$||Q - q^\mu_{A,0}|| \leq C||g||. \tag{6.115}$$

¿From (6.115) we can derive as in the proof of lemma 4.4 that

$$\begin{aligned} \overline{\Phi}_\mu(A, Q - q_0) - \overline{\Phi}_\mu(A, q^\mu_{A,0} - q_0) &= D_w\overline{\Phi}_\mu(A, q^\mu_{A,0} - q_0)\cdot(Q - q^\mu_{A,0}) + O(||Q - q^\mu_{A,0}||^2) \\ &= O(||Q - q^\mu_{A,0}||^2). \end{aligned} \tag{6.116}$$

By the complex differentiability of Φ_μ, the differential of the map $(A \mapsto q^\mu_{A,0})$ is $\mathbf{C}$-linear at all point, hence $q^\mu_{A,0}$ depends analytically on $A \in D_r$. As a result, $G_\mu : \mathbf{T}^n \to \mathbf{R}$ has an analytic extension which we shall still denote by G_μ, defined by $G_\mu(A) = \Phi_\mu(w^\mu_A)$. Now let

$$U_{r,\omega} = \{\theta \in \mathbf{C} \;:\; |\omega||Im\theta| < r/2\}.$$

Since $F_\mu(A,\theta) = G_\mu(A + \omega\theta)$ for all $A \in \mathbf{T}^n, \theta \in \mathbf{R}$, F_μ too has an analytic extension, defined in $D_{r/2} \times U_{r,\omega}$ by $F_\mu(A,\theta) = G_\mu(A + \theta\omega)$.

(ii) Now we justify that l_μ too has an analytic extension and that the estimate $|V_\mu - \widetilde{F}_\mu| = O(|\partial_\theta \widetilde{F}_\mu|^2)$ may be extended to complex values of (A, θ). Let

$$S' = \Big\{\theta \in \mathbf{C} \ : \ \mathrm{Im}\theta \in (-\frac{\pi}{4}, \frac{\pi}{4})\Big\}.$$

As a straightforward consequence of lemma 4.6 (where σ is not small but for instance equal to $\pi/6$) for $\mu|f|_r$ small enough, for all $A \in D_r$, $\theta \in S'$, there is a unique $Q^\mu_{A,\theta} \in q_0 + Y = q_\theta + Y$, there is a unique $\alpha^\mu_{A,\theta} \in \mathbf{C}$ such that $||Q^\mu_{A,\theta} - q_\theta|| + |\alpha^\mu_{A,\theta}| \le \rho$ and

$$\left\{ \begin{array}{l} -\ddot{Q}^\mu_{A,\theta} + \sin Q^\mu_{A,\theta} = \mu \sin Q^\mu_{A,\theta} f(A + \omega t) + \alpha^\mu_{A,\theta} \psi_\theta \\ \displaystyle\int_{-\infty}^{\infty} (Q^\mu_{A,\theta} - q_\theta) \psi_\theta \ dt = 0. \end{array} \right.$$

Moreover $||Q^\mu_{A,\theta} - q_\theta|| + |\alpha^\mu_{A,\theta}| = O(\mu|f|_r)$,

$$||\partial_A Q^\mu_{A,\theta}|| = O(\mu|f|_r) \tag{6.117}$$

and $Q^\mu_{A,\theta}$, $\alpha^\mu_{A,\theta}$ depend analytically on $(A, \theta) \in D_r \times S'$.

We have $\widetilde{F}_\mu(A, \theta) = \Phi_\mu(A, Q^\mu_{A,\theta} - q_\theta)$. In a similar way as in lemma 4.2 we obtain

$$|\alpha^\mu_{A,\theta}| = O(|\partial_\theta \widetilde{F}_\mu(A, \theta)|). \tag{6.118}$$

Now by lemma 4.1,

$$Q^\mu_{A,0}(t + \eta) = Q^\mu_{A+\omega\eta,-\eta}(t) \tag{6.119}$$

for all $A \in \mathbf{T}^n$, $\eta \in \mathbf{R}$. Hence, by analyticity, (6.119) holds for all $A \in D_r$, $\eta \in \mathbf{R}$. In particular $Q^\mu_{A,0}(\eta) = Q^\mu_{A+\omega\eta,-\eta}(0)$. This implies that for $A \in D_{r/2}$, $Q^\mu_{A,0}$ has an analytic extension defined in $U_{r,\omega}$ (provided $r/(2|\omega|) \le \pi/4$). We can write for $A \in D_{r/2}$, $\eta \in U_{r,\omega}$,

$$\frac{d}{d\eta} Q^\mu_{A+\omega\eta,-\eta}(0) = \dot{Q}^\mu{}_{A+\omega\eta,-\eta}(0) = \dot{q}_0(\eta) + O(\mu) = \dot{q}_0(0) + O(|\eta|) + O(\mu). \tag{6.120}$$

We already know (lemma 4.3) that, for $A \in \mathbf{T}^n$, there is $\widetilde{l}(A) = l_\mu(A, 0) = O(\mu) \in \mathbf{R}$ such that $Q^\mu_{A,0}(\widetilde{l}(A)) = \pi$, which writes also $Q^\mu_{A+\omega\widetilde{l}(A),-\widetilde{l}(A)}(0) = \pi$. Hence by (6.117), for $A' \in [-r, r]^n$ and $B = A + iA' + \omega\widetilde{l}(A)$,

$$Q^\mu_{B,-\widetilde{l}(A)}(0) = \pi + O(\mu|A'|).$$

Hence, since $\dot{q}_0(0) \neq 0$, by (6.120) and by the Implicit Function Theorem, if $\mu|\omega||A'|$ is small enough, there is $\widetilde{l}(A + iA') = \widetilde{l}(A) + O(\mu|A'|)$ such that

$$Q^\mu_{A+\omega\widetilde{l}(A+iA'),-\widetilde{l}(A+iB)}(0) = \pi.$$

Moreover $\widetilde{l}$ thus defined is analytic in $D_{r'}$ for r' small enough (how small depending on $|\omega|$). For a given $A \in D_{r'}$, let $q = Q^\mu_{A+\omega\widetilde{l}(A),-\widetilde{l}(A)}$. Then, with the abbreviation $\alpha = \alpha^\mu_{A+\omega\widetilde{l}(A),-\widetilde{l}(A)}$,

$$\left\{ \begin{array}{l} -\ddot{q} + \sin q = \mu \sin q f(A + \omega\widetilde{l}(A) + \omega t) + \alpha\psi_{-\widetilde{l}(A)} \\ q(0) = \pi, \end{array} \right. .$$

Therefore, by (6.116),

$$\begin{aligned} G_\mu(A + \widetilde{l}(A)\omega) - \widetilde{G}_\mu(A) &= G_\mu(A + \widetilde{l}(A)\omega) - \widetilde{F}_\mu(A + \omega\widetilde{l}(A), -\widetilde{l}(A)) \\ &= \overline{\Phi}_\mu(q^\mu_{A+\widetilde{l}(A)\omega} - q_0) - \overline{\Phi}_\mu(Q^\mu_{A+\widetilde{l}(A)\omega,-\widetilde{l}(A)} - q_0) \\ &= O(||\alpha\psi_{-\widetilde{l}(A)}||^2) \\ &= O(|\partial_\theta \widetilde{F}_\mu(A + \omega\widetilde{l}(A), -\widetilde{l}(A))|^2) = O(|\partial_\theta \widetilde{F}_\mu(A, 0)|^2). \end{aligned}$$

Let l_μ be the analytic map defined in $D_{r'/2} \times U_{r'/2,\omega}$ by $l_\mu(A,\theta) = \widetilde{l}(A+\omega\theta)$. Let

$$V_\mu(A,\theta) = F_\mu(A, \theta + l_\mu(A,\theta))$$

This is an analytic extension of the map defined in lemma 4.3. By the previous estimates,

$$|V_\mu - \widetilde{F}_\mu| = O(|\partial_\theta \widetilde{F}_\mu|^2), \tag{6.121}$$

uniformly for $(A,\theta) \in D_{r'/2} \times U_{r'/2,\omega}$.

(iii) To complete the proof, we remark that, if $f, g : U \to \mathbf{C}$ are analytic in U, open subset of $\mathbf{C}^m$ and $g \not\equiv 0$, $f = O(g)$ locally in U, then f/g has an analytic extension defined in the whole set U (see for instance [40], p. 32, thm 3.4). Hence $(V_\mu - \widetilde{F}_\mu)/(\partial_\theta \widetilde{F}_\mu)^2$ is real analytic in $\mathbf{T}^n \times \mathbf{R}$. As a consequence, $\overline{h}_\mu(A,\theta)$ given by (4.78) is real analytic w.r.t. (A,θ). Moreover, as l_μ is an analytic function, so is $\overline{l}_\mu$. Finally $h_\mu(A,\theta) = \overline{l}_\mu(A,\theta) + \overline{h}_\mu(A, \theta + \overline{l}_\mu(A,\theta))$ is real analytic. ■

References

[1] A. Ambrosetti, M. Badiale: *Homoclinics: Poincaré-Melnikov type results via a variational approach*, Annales I. H. P. - Analyse nonlin., vol. 15, n.2, 1998, pp. 233-252.

[2] S. Angenent: *A variational interpretation of Melnikov's function and exponentially small separatrix splitting*, Lecture notes of the London Math. Soc, 1993, Symplectic geometry, ed. Dietmar Salamon.

[3] V. I. Arnold: *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Dokl. 6, 1964, pp. 581-585.

[4] M. Berti, P. Bolle: *Homoclinics and Chaotic Behaviour for Perturbed Second order Systems*, Annali di Mat. Pura e Applicata, (IV), vol. CLXXVI, 1999, pp. 323-378.

[5] M. Berti, P. Bolle: *Variational construction of Homoclinics and Chaotic Behaviour in presence of a saddle-saddle equilibrium*, Annali della Scuola Normale Superiore di Pisa, serie IV, vol. XXVII, fasc. 2, 1998, pp. 331-377.

[6] M. Berti, P. Bolle: *Diffusion time and splitting of separatrices for nearly integrable isochronous Hamiltonian systems*, Rend. Mat. Acc. Naz. Lincei, s. 9, vol. 11, fasc. 4, 2000, pp. 235-243.

[7] M. Berti, P. Bolle: *Fast Arnold's diffusion in three time scales systems*, preprint SISSA.

[8] U. Bessi: *An approach to Arnold diffusion through the calculus of variations*, Nonlinear Analysis T. M. A., 26, 1996, pp. 1115-1135.

[9] U. Bessi, L. Chierchia, E. Valdinoci: *Upper Bounds on Arnold Diffusion Time via Mather theory*, to appear in Journal de Mathématiques Pures et Appliquées.

[10] L. Biasco: *Stime analitiche sui tempi di instabilitá per perturbazioni di sistemi hamiltoniani degeneri*, Tesi di Laurea, 1999, Univ. Roma 3.

[11] S. V. Bolotin: *Homoclinic orbits to invariant tori of Hamiltonian systems*, Amer. Math. Soc. Transl., ser 2, 168, 1995, pp. 21-90.

[12] S. V. Bolotin: *Infinite number of homoclinic orbits to hyperbolic invariant tori of Hamiltonian systems*, Regular and Chaotic Dynamics, v. 5, n.2, 2000.

[13] J. Bourgain, F. Golse, B. Wennberg: *On the distribution of free path lengths for periodic Lorentz gas*, Comm. Math. Phys., vol. 190, 1998, pp. 491-508.

[14] L. Chierchia: *Arnold instability for nearly-integrable analytic hamiltonian systems*, Proceedings of the workshop "Variational and local methods in the study of Hamiltonian systems", A. Ambrosetti, G. F. Dell'Antonio eds., Trieste, 1994.

[15] L. Chierchia, G. Gallavotti: *Drift and diffusion in phase space*, Annales de l'IHP, section Physique Théorique, 60, 1994, pp. 1-144; see also Erratum in Vol. 68, 135, 1998.

[16] L. Chierchia, E. Zehnder: *Asymptotic expansions of quasi-periodic motions*, Annali della Scuola Normale Superiore di Pisa, serie IV, vol. XVI, 1989, pp. 245-258.

[17] B. V. Chirikov: *A universal instability of many dimensional oscillator systems*, Physics Reports 52, 1979, pp. 263-379.

[18] J. Cresson: *A λ-lemma for partially hyperbolic tori and the obstruction property*, Lett. Math. Phys., 42, 1997, pp. 363-377.

[19] J. Cresson: *Conjecture de Chirikov et optimalité des exposants de stabilité du théorème de Nekhoroshev*, preprint univ. Besancon.

[20] J. Cresson, C. Guillet: *Periodic orbits and Arnold diffusion*, preprint univ. Besancon.

[21] A. Delshams, V. G. Gelfreich, V. G. Jorba, T. M. Seara: *Exponentially small splitting of separatrices under fast quasi-periodic forcing*, Comm. Math Ph. 189, 1997, pp. 35-71.

[22] A. Delshams, V. Gutiérrez: *Homoclinic orbits to invariant tori in Hamiltonian systems*, In Multiple-Time-Dynamical systems, C. Jones, S. Wiggins, A Khibnik, F. Dumortier and D. Terman eds., IMA Vol. in Math and its Applications, Springer-Verlag, to appear.

[23] A. Delshams, T. M. Seara: *An asymptotic expression for the splitting of separatrices of the rapidly forced pendulum*, Comm. Math. Phys. 150, 1992, pp. 433-463.

[24] G. Gallavotti: *Arnold's Diffusion in Isochronous Systems*, Mathematical Physics, Analysis and Geometry 1, 1999, pp. 295-312.

[25] G. Gallavotti, G. Gentile, V. Mastropietro: *Separatrix splitting for systems with three time scales*, Commun. Math. Phys. 202, 1999, pp. 197-236.

[26] G. Gallavotti, G. Gentile, V. Mastropietro: *Melnikov approximation dominance. Some examples*, to appear in Rev. Math. Phys.

[27] G. Gallavotti, G. Gentile, V. Mastropietro: *On homoclinic splitting problems*, Physica D, 137, 2000, pp. 202-204.

[28] G. Gallavotti, G. Gentile, V. Mastropietro: *Hamilton-Jacobi equation and existence of heteroclinic chains in three time scales systems*, Nonlinearity, 13, 2000, pp. 323-340.

[29] V. Gelfreich: *Melnikov method and exponentially small splitting of separatrices*, Physica D, 101, 1997, pp. 227-248.

[30] V. Gelfreich: *A proof of the exponentially small transversality of the separatrices for the standard map*, Comm. Math. Phys., 201, 1999, pp. 155-216.

[31] V. Gelfreich: *Separatrix splitting for a high-frequency perturbation of the pendulum*, Russian Journal of Mathematical Physics, vol. 7, n. 1, 2000, pp. 48-71.

[32] G. Gentile: *A proof of existence of whiskered tori with quasi flat homoclinic intersection in a class of almost integrable systems*, Forum Mathematicum, 7, 1995, pp. 709-753.

[33] P. Holmes, J. Marsden, J. Scheurle: *Exponentially small splittings of separatrices with applications to KAM theory and degenerate bifurcations*, Contemp. Math., 81, 1988, pp. 213-244.

[34] V. F. Lazutkin: Splitting of separatrices for the Chirikov's standard map, preprint VINITI 6372/84, 1984.

[35] P. Lochak: *Arnold diffusion: a compendium of remarks and questions*, Proceedings of 3DHAM's Agaro, 1995.

[36] P. Lochak, J. P. Marco, D. Sauzin: *On the splitting of invariant manifolds in multidimensional Hamiltonian systems*, preprint Université Jussieu.

[37] J. P. Marco: *Transitions le long des chaînes de tores invariants pour les systèmes hamiltoniens analytiques*, Annales I. H. P., vol. 64, 1995, pp. 205-252.

[38] A. Pumarino, C. Valls: *Three time scales systems exhibiting persistent Arnold Diffusion*, preprint.

[39] A. Pumarino, C. Valls: *Exponentially small splitting of separatrices in open sets of frequencies*, preprint.

[40] R. M. Range: *Holomorphic functions and integral representations in several complex variables*, Springer Verlag, 1986.

[41] M. Rudnev, S. Wiggins: *Existence of exponentially small separatrix splittings and homoclinic connections between whiskered tori in weakly hyperbolic near integrable Hamiltonian systems*, Physica D, 114, 1998, 3-80. See also Erratum in Physica D, 145, 2000, pp. 349-354.

[42] M. Rudnev, S. Wiggins: *On a homoclinic splitting problem*, Regular and Chaotic Dynamics, vol. 5, n. 2, 2000, pp. 227-242.

[43] D. Sauzin: *A new method for measuring the splitting of invariant manifolds*, to appear in Ann. scient. E.N.S.

[44] D. Treschev: *Multidimensional symplectic separatrix maps*, preprint.

[45] Z. Xia: *Arnold diffusion: a variational construction*, Documenta Matematica, extra vol. ICM, 1998, II, pp. 867-877.

Massimiliano Berti, S.I.S.S.A., Via Beirut 2-4, 34014, Trieste, Italy, berti@sissa.it.

Philippe Bolle, Département de mathématiques, Université d'Avignon, 33, rue Louis Pasteur, 84000 Avignon, France, philippe.bolle@univ-avignon.fr