Arnold's Diffusion in nearly integrable isochronous Hamiltonian systems

Massimiliano Berti and Philippe Bolle

Abstract: We consider the problem of Arnold's diffusion for nearly integrable isochronous Hamiltonian systems. We prove a shadowing theorem which improves the known estimates for the diffusion time. We also develop a new method for measuring the splitting of the separatrices. As an application we justify, for three time scales systems, that the splitting is correctly predicted by the Poincaré-Melnikov function. [1]

Keywords: Arnold's diffusion, shadowing theorem, splitting of separatrices, heteroclinic orbits, variational methods.

1 Introduction

Through this paper we consider nearly integrable isochronous Hamiltonian systems as

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

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. Hamiltonian $\mathcal{H}_\mu$ describes a system of n isochronous harmonic oscillators of frequencies ω weakly coupled with a pendulum.

When $\mu = 0$ the energy $\omega_i I_i$ of each oscillator is a constant of the motion. The problem of *Arnold's diffusion* is whether, for $\mu \neq 0$, there exist motions whose net effect is to transfer energy from one oscillator to others. This problem has been broadly investigated by many authors also for non-isochronous systems, see for example [3], [9], [10], [11] and [25]. In this paper we focus on isochronous systems for which, 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

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

The existence of Arnold's diffusion is usually proved following the mechanism proposed in [3]. First one remarks that, for $\mu = 0$, Hamiltonian $\mathcal{H}_\mu$ admits a continuous family of n-dimensional partially hyperbolic invariant tori $\mathcal{T}_{I_0} = \{(\varphi, I, q, p) \in \mathbf{T}^n \times \mathbf{R}^n \times \mathbf{T}^1 \times \mathbf{R}^1 \mid I = I_0,\ q = p = 0\}$ possessing stable and unstable manifolds $W^s(\mathcal{T}_{I_0}) = W^u(\mathcal{T}_{I_0}) = \{(\varphi, I, q, p) \in \mathbf{T}^n \times \mathbf{R}^n \times \mathbf{T}^1 \times \mathbf{R}^1 \mid I = I_0,\ p^2/2 + (\cos q - 1) = 0\}$. Then Arnold's mechanism is based on the following three main steps.

Step (i) To prove that, for μ small enough, the perturbed stable and unstable manifolds $W^s_\mu(\mathcal{T}^\mu_{I_0})$ and $W^u_\mu(\mathcal{T}^\mu_{I_0})$ split and intersect transversally ("splitting of the separatrices");

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

Step (iii) To prove, by a shadowing type argument, the existence of an orbit such that the action variables I undergo a variation of $O(1)$ in a certain time T_d called the *diffusion time*.

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

We point out that for isochronous systems assumption $(H1)$ implies that *all* the unperturbed tori $\mathcal{T}_{I_0}$, with their stable and unstable manifolds, persist, for μ small enough, being just sligthly deformed. For this reason the construction of the "transition chain" of step (ii) is a straightforward consequence of step (i). This also happens, for the peculiar choice of the perturbation, in the non-isochronous system considered in [3]. On the other hand, 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 difficult matter, see [11]. We quote paper [25] for a somewhat different mechanism of diffusion where step (ii) is bypassed using Mather's theory.

In the present paper we address, for isochronous systems, the following two main questions

1) Shadowing theorems and estimates of the diffusion time;

2) Splitting of separatrices.

Problem 1) has been intensively studied in the last years, see for example [10],[11],[12],[14] and [21] (we underline that [10],[11],[12] and [21] deal also with non-isochronous systems). Our general shadowing theorem (*thm.2.3-thm.3.2*) improves -for isochronous systems- the estimates on the diffusion time obtained in the forementioned papers.

The estimate on the diffusion time that we obtain (see expression (2.23)), once it is verified that the stable and the unstable manifolds split, is roughly the following: the diffusion time 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.

In order to highlight the improvement of our estimate of the diffusion time let us consider the particular case of "a-priori unstable" systems, i.e. when the frequency vector ω is considered as a constant independent of any parameter. In such a case it is easy to evaluate, using the classical Poincaré-Melnikov theory, that the splitting of the separatrices is $O(\mu)$. Then our shadowing theorem yields the estimate for the diffusion time $T_d = O((1/\mu)\log(1/\mu))$, see *thm. 2.4-thm.3.3*.

Such estimate answers to a question raised in [19] (sec.7) proving that, at least for isochronous systems, it is possible to reach the maximal speed of diffusion $\mu/|\log\mu|$. On the contrary the estimate on the diffusion time obtained in [11] is $T_d >> O(\exp(1/\mu))$ and is improved in [14] to be $T_d = O(\exp(1/\mu))$. Recently in [10] by means of Mather's theory the estimate on the diffusion time has been improved to be $T_d = O(1/\mu^{2\tau+1})$. In [12] it is obtained via geometric methods that $T_d = O(1/\mu^{\tau+1})$. It is worth pointing out that the estimates given in [10] and [12], which yet provide a diffusion time polinomial in the splitting, depend on the diophantine exponent τ and hence on the number of rotators n. On the contrary our estimate is independent of n.

The main reason for which we are able to improve also the estimates of [10] and [12] 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 [10] and [12]. 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 $T_d = O((1/\mu)\log(1/\mu))$.

Regarding the method of proof, we use a variational technique inspired by [5] and [6]. One advantage of this approach is that the same arguments can be also used 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 in the case that the hyperbolic part is the standard one-dimensional pendulum because it is the model equation to study Arnold's diffusion near a simple-resonance.

Furthermore we also remark that our proof of theorem 3.2 is completely self-contained in the sense that, unlike the known approaches (excepted [25]), we do not make use of any KAM-type result for proving, under assumption $(H1)$, the persistence of invariant tori, see *thm. 3.1*.

In sections 4-5 we study problem 2). Detecting the splitting of the separatrices becomes a very 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 oscillations 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 ε. Much literature in the last years has been devoted to overcome this problem, see for example [13],[15], [16], [20] and [23]. In the present paper, in order to justify the dominance of the Poincaré-Melnikov function when $\mu = O(\varepsilon^p)$, we extend the approach originally used in [2] for dealing with rapidly periodic forced systems. Up to a change of variables close to the identity, we prove (see *thm. 4.2*) an exponentially small upper bound for the Fourier coefficients of the splitting. As an application we provide some results (*thm. 5.1, thm. 5.4)* on the splitting of the separatrices and the diffusion time (*thm. 5.2*) for three time scales systems

$$\mathcal{H}_\varepsilon = \frac{1}{\sqrt{\varepsilon}} I_1 + \varepsilon^a \beta \cdot I_2 + \frac{p^2}{2} + (\cos q - 1) + \mu(\cos q - 1) f(\varphi),\ I_1 \in \mathbf{R}^1, I_2, \beta \in \mathbf{R}^{n-1},\ n \geq 2,$$

for $\mu\varepsilon^{-3/2}$ small. This improves the main theorem I in [22] which holds for $\mu = \varepsilon^p$, $p > 2 + a$. With respect to [15], which deals for more general systems, we remark that our results hold in any dimension, while the results of [15] are proved for 2 rotators only.

Theorem 4.2 is also the starting point for proving the splitting of the separatrices in presence of two high frequencies, assuming as in [13],[16], [20] and [23] suitable hypotheses on the perturbation term. We do not address this problem in this paper.

On the other hand theorem 5.1 is the starting point to prove, for $n \geq 3$, the existence of diffusion solutions such that the action variables I_2 undergo a variation $O(1)$ in polynomial time (while the I_1 action variable does not change considerably). This phenomenon can not be deduced by the estimates, given in [15] and [22], on the "determinant of the splitting" which does not distinguish among "slow" and "fast" directions and would give rise to exponentially large diffusion times. This type of results are contained in the forthcoming paper [8].

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 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.

The results of this paper have been announced in [7].

After this paper was completed we learned by prof. Bolotin about the recent preprint [24] which deals with a-priori unstable Hamiltonian systems time periodically forced. Among many results, in theorem 2 of [24] a shadowing theorem for a symplectic separatrix map is proved providing an estimate on the diffusion speed $\mu/|\log \mu|$, like our. However theorem 2 applies just to an approximation of the symplectic map describing the true dynamics of the system and, moreover, requires the hyperbolic part to be just two dimensional.

Acknowledgments: The first author wishes to thank Prof. G. Gallavotti for stimulating discussions.

2 The shadowing theorem

We first develop 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

$$\dot{\varphi} = \omega, \qquad \dot{I} = -\mu(1 - \cos q)\ \partial_\varphi f(\varphi), \qquad \dot{q} = p, \qquad \dot{p} = \sin q - \mu \sin q\ f(\varphi). \tag{2.1}$$

The dynamics on the angles φ is given by $\varphi(t) = \omega t + A$ so that (2.1) are reduced to the quasi-periodically forced pendulum

$$-\ddot{q} + \sin q = \mu\ \sin q\ f(\omega t + A) \tag{2.2}$$

corresponding to the Lagrangian

$$\mathcal{L}_\mu(q, \dot{q}, t) = \frac{\dot{q}^2}{2} + (1 - \cos q) + \mu(\cos q - 1) f(\omega t + A). \tag{2.3}$$

For each solution $q(t)$ of (2.2) one recovers the dynamics of the actions $I(t)$ by quadratures in (2.1).

2.1 1-bump homoclinic and heteroclinic solutions

For $\mu = 0$ equation (2.2) is autonomous and possesses the homoclinic (mod. 2π) solutions $q_\theta(t) = 4\ \mathrm{arctg}(\exp(t-\theta))$, $\theta \in \mathbf{R}$. Using the Contraction Mapping Theorem we now prove that, near the unperturbed homoclinic solutions $q_\theta(t)$, there exist, for μ small enough, *"pseudo-homoclinic solutions"* $q^\mu_{A,\theta}(t)$ of equation (2.2). $q^\mu_{A,\theta}(t)$ are true solutions of (2.2) 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.

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

- *(i) $q^\mu_{A,\theta}(t)$ is a solution of (2.2) in each interval $(-\infty,\theta)$ and $(\theta,+\infty)$ and $q^\mu_{A,\theta}(\theta)=\pi$;*
- *(ii) $\max\Big(|q^\mu_{A,\theta}(t)-q_\theta(t)|, |\dot q^\mu_{A,\theta}(t)-\dot q_\theta(t)|\Big)\le C_0\mu\exp(-\frac{|t-\theta|}{2})$, $\forall t\in\mathbf{R}$;*
- *(iii) $q^\mu_{A,\theta}(t)=q^\mu_{A+k2\pi,\theta}(t)$, $\forall k\in\mathbf{Z}^n$;*
- *(iv) $q^\mu_{A,\theta+\eta}(t+\eta)=q^\mu_{A+\omega\eta,\theta}(t)$, $\forall\theta,\eta\in\mathbf{R}$.*
- *(v) $\max\Big(|\partial_A q^\mu_{A,\theta}(t)|, |\partial_A\dot q^\mu_{A,\theta}(t)|, |\omega\cdot\partial_A q^\mu_{A,\theta}(t)|, |\omega\cdot\partial_A\dot q^\mu_{A,\theta}(t)|\Big)=O\Big(\mu\exp(-\frac{|t-\theta|}{2})\Big)$.*

Proof. Proof in the appendix. Note that in (v) the bound of $|\omega\cdot\partial_A q^\mu_{A,\theta}(t)|, |\omega\cdot\partial_A\dot q^\mu_{A,\theta}(t)|$ is uniform in ω. ■

We can then define the function $F_\mu:\mathbf{T}^n\times\mathbf{R}\to\mathbf{R}$ as the action functional of Lagrangian (2.3) evaluated on the "1-bump pseudo-homoclinic solutions" $q^\mu_{A,\theta}(t)$, namely

$$F_\mu(A,\theta)=\int_{-\infty}^{\theta}\mathcal{L}_\mu(q^\mu_{A,\theta}(t),\dot q^\mu_{A,\theta}(t),t)\ dt+\int_{\theta}^{+\infty}\mathcal{L}_\mu(q^\mu_{A,\theta}(t),\dot q^\mu_{A,\theta}(t),t)\ dt \tag{2.4}$$

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

$$G_\mu(A)=F_\mu(A,0). \tag{2.5}$$

Since $q^\mu_{A,\theta}(t)$ converges exponentially fast to 0, mod 2π, the integrals in (2.4) are convergent. Note that the homoclinic function G_μ is independent of I_0. By property (v) of lemma 2.1 the following invariance property holds

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

and in particular

$$F_\mu(A,\theta)=G_\mu(A+\omega\theta),\quad\forall\theta\in\mathbf{R}. \tag{2.6}$$

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 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 (2.2) 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 (2.2) near $q_0(t)$ with $q^\mu_{A,q_0}(0)=q_0$ and $\lim_{t\to-\infty}q^\mu_{A,q_0}(t)=0$.

Lemma 2.2 *The derivative of $\theta \to F_\mu(A,\theta)$ satisfies*

$$\partial_\theta F_\mu(A,\theta) = \frac{(\dot q^\mu_{A,\theta})^2(\theta^+)}{2} - \frac{(\dot q^\mu_{A,\theta})^2(\theta^-)}{2}. \tag{2.7}$$

Proof. There holds

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

Integrating by parts and using also that $q^\mu_{A,\theta}(t)$ solves (2.2), we obtain

$$\partial_\theta F_\mu(A,\theta) = \frac{1}{2}(\dot q^\mu_{A,\theta})^2(\theta^-) - \frac{1}{2}(\dot q^\mu_{A,\theta})^2(\theta^+) + \Big[\partial_\theta q^\mu_{A,\theta}(t)\dot q^\mu_{A,\theta}(t)\Big]_{-\infty}^{\theta^-} + \Big[\partial_\theta q^\mu_{A,\theta}(t)\dot q^\mu_{A,\theta}(t)\Big]_{\theta^+}^{+\infty}. \tag{2.8}$$

Since $\forall \theta \in \mathbf{R}$ $q^\mu_{A,\theta}(\theta) = \pi$, deriving in θ we obtain $\partial_\theta q^\mu_{A,\theta}(\theta) + \dot q^\mu_{A,\theta}(\theta) = 0$; hence from (2.8) and using that $\lim_{t\to\pm\infty} \dot q^\mu_{A,\theta}(t) = 0$ we deduce lemma 2.2. ■

By lemma 2.2 if $\partial_\theta F_\mu(A,\theta) = 0$ then $q^\mu_{A,\theta}(t)$ is a true homoclinic (mod. 2π) solution of (2.2). Then, for each $I_0 \in \mathbf{R}^n$,

$$\Big(\omega t + A, I_\mu(t), q^\mu_{A,\theta}(t), \dot q^\mu_{A,\theta}(t)\Big) \tag{2.9}$$

where

$$I_\mu(t) = I_0 - \mu \int_{-\infty}^{t} (1 - \cos q^\mu_{A,\theta}(s))\partial_\varphi f(\omega s + A) ds \tag{2.10}$$

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 from $\mathcal{T}_{I_0}$ to $\mathcal{T}_{I_\mu(+\infty)}$ when the jump is not zero. Moreover the next lemma says that such 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 (2.10) satisfies*

$$\partial_A F_\mu(A,\theta) = \int_{-\infty}^{+\infty} \dot I_\mu(t)\ dt = I_\mu(+\infty) - I_0 < +\infty. \tag{2.11}$$

In particular if (A,θ) is a critical point of $F_\mu(A,\theta)$ then (2.9) in a homoclinic orbit to torus $\mathcal{T}_{I_0}$.

Proof. There holds

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

Integrating by parts, since $q^\mu_{A,\theta}(t)$ solves (2.2), and using that $\lim_{t\to\pm\infty} \dot q^\mu_{A,\theta}(t) = 0$, we deduce

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

We deduce from (2.12) equality (2.11). ■

By the invariance property (2.6) if B is a critical point of the homoclinic function G_μ, then, for all (A,θ) such that $A + \omega\theta = B$, (2.9) are homoclinic solutions to each torus $\mathcal{T}_{I_0}$. These homoclinics 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 Lusternik-Schirelman 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 [20])

Theorem 2.1 *Let $0 < \mu \leq \mu_0$. $\forall I_0 \in \mathbf{R}^n$ there exist at least $n+1$ homoclinic orbits geometrically dinstict 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

$$\omega \cdot I_0 = \omega \cdot I_0'. \tag{2.13}$$

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 $\mathcal{T}_{I_0}$ to $\mathcal{T}_{I_0'}$. If the energy condition (2.13) holds then the function $F_{\mu,I_0,I_0'}(A,\theta)$ satisfies the invariance property

$$F_{\mu,I_0,I_0'}(A,\theta) = G_\mu(A+\omega\theta) - (I_0' - I_0)\cdot(A+\omega\theta) = G_{\mu,I_0,I_0'}(A+\omega\theta). \tag{2.14}$$

where

$$G_{\mu,I_0,I_0'}(B) := G_\mu(B) - (I_0' - I_0)\cdot B. \tag{2.15}$$

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

2.2 The k-bump pseudo-homoclinic solutions

We prove in the next lemma the existence of pseudo-homoclinic solutions $q^L_{A,\theta}(t)$ of the quasi-periodically forced pendulum (2.2) 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 $\forall\omega \in \mathbf{R}^n$, $\forall 0 < \mu \leq \mu_0$, $\forall k \in \mathbf{N}$, $\forall L > L_1$, $\forall \theta_1 < \ldots < \theta_k$ with $\min_i |\theta_{i+1} - \theta_i| > L$, 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 (2.2) 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) $q^L_{A,\theta+\eta}(t+\eta) = q^L_{A+\omega\eta,\theta}(t)$, $\forall\theta,\eta \in \mathbf{R}$.*

Proof. In the appendix. ■

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

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

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

By lemma 2.4-v the following invariance property holds

$$F_\mu^k(A,\theta+\eta) = F_\mu^k(A+\eta\omega,\theta), \qquad \forall \theta,\eta \in \mathbf{R}. \tag{2.16}$$

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

$$\mathcal{F}_\mu^k(A,\theta) := F_\mu^k(A,\theta) - (I_0' - I_0)\cdot A. \tag{2.17}$$

Arguing as in lemma 2.3 we have

Lemma 2.5 *$\forall I_0, I_0' \in \mathbf{R}^n$, if (A,θ) is a critical point of $\mathcal{F}_\mu^k(A,\theta)$, then $(\omega t + A, I_\mu(t),\ q_{A,\theta}^L(t), \dot q_{A,\theta}^L(t))$ where $I_\mu(t) = I_0 - \mu\int_{-\infty}^t (1-\cos q_{A,\theta}^L(s))\partial_\varphi f(\omega s + A)ds$ 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 $\mathcal{F}_\mu^k(A,\theta)$. When $\min_i(\theta_{i+1}-\theta_i) \to +\infty$ the "k-bump homoclinic function" $F_\mu^k(A,\theta)$ turns out to be 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 $\min_i(\theta_{i+1}-\theta_i) > L$*

$$F_\mu^k(A,\theta_1,\cdots,\theta_k) = \sum_{i=1}^k F_\mu(A,\theta_i) + \sum_{i=1}^k R_i(\mu,A,\theta_{i-1},\theta_i,\theta_{i+1}), \tag{2.18}$$

with

$$|R_i(\mu,A,\theta_{i-1},\theta_i,\theta_{i+1})| \le C_2\exp(-C_2L).$$

Proof. We can write

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

We define

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

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

where q_{A,θ_i}^μ is the 1-bump pseudo-homoclinic solution obtained in lemma 2.1. Recalling the definition 2.4 of $F_\mu(A,\theta)$ we have

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

Setting $R_i = R_i^- + R_i^+$ we derive the expression (2.18). In order to complete the proof, it is enough 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 $i = 1,\ldots,k$

$$|R_i^\pm(\mu,A,\theta_i,\theta_{i+1})| \le C_2\exp(-C_2L). \tag{2.19}$$

We write the proof for R_i^+. We have

$$\begin{aligned} R_i^+(\mu, A, \theta_i, \theta_{i+1}) &= \int_{\theta_i}^{(\theta_i+\theta_{i+1})/2} (\mathcal{L}_\mu(q^L_{A,\theta}(t), \dot{q}^L_{A,\theta}(t), t) - \mathcal{L}_\mu(q^\mu_{A,\theta_i}(t), \dot{q}^\mu_{A,\theta_i}(t), t)\ dt \\ &- \int_{(\theta_i+\theta_{i+1})/2}^{+\infty} \mathcal{L}_\mu(q^\mu_{A,\theta_i}(t), \dot{q}^\mu_{A,\theta_i}(t), t)\ dt. \end{aligned} \tag{2.20}$$

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

$$\Big| \int_{(\theta_i+\theta_{i+1})/2}^{+\infty} \mathcal{L}_\mu(q^\mu_{A,\theta_i}(t), \dot{q}^\mu_{A,\theta_i}(t), t)\ dt \Big| = O(e^{-L/2}). \tag{2.21}$$

From lemma 2.4-(ii) we also deduce that

$$\Big(\int_{\theta_i}^{(\theta_i+\theta_{i+1})/2} \mathcal{L}_\mu(q^L_{A,\theta}(t), \dot{q}^L_{A,\theta}(t), t) - \mathcal{L}_\mu(q^\mu_{A,\theta_i}(t), \dot{q}^\mu_{A,\theta_i}(t), t)\ dt \Big) = O(e^{-CL}). \tag{2.22}$$

From (2.20), (2.21) and (2.22) we deduce (2.19) and hence the lemma. ■

2.3 The diffusion orbit

We are now able to consider the existence of the shadowing orbit. We give an example of condition on G_μ which implies the existence of diffusion orbits.

Condition 2.1 *("Splitting condition") There exist $A_0 \in \mathbf{T}^n$, $\delta > 0$, $0 < \alpha < \rho$ such that*

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

Remark 2.2 *If G_μ possesses a non-degenerate minimum in A_0 the "splitting condition" above is satisfied, for ρ sufficiently small, choosing $\delta = (\min \lambda_i)\rho^2/4$ and $\alpha = (\rho/8)\sqrt{(\min_i \lambda_i)/(\max_i \lambda_i)}$ where λ_i are the positive eigenvalues of $D^2G^\mu(A_0)$.*

Remark 2.3 *$B_\rho(A_0)$, open ball of radius ρ in $\mathbf{R}^n$ (the covering space of $\mathbf{T}^n$), could be replaced by a bounded open subset U of $\mathbf{R}^n$.*

The following shadowing type theorem holds

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*

$$T_d \leq C_3 \frac{|I_0 - I_0'|}{\delta} \rho \max\Big(|\ln \delta|, \frac{1}{\gamma\alpha^\tau} \Big) + C_3 |\ln(\eta)|. \tag{2.23}$$

Remark 2.4 *The meaning of (2.23) 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 $|\ln\delta|$ needed to "shadow" homoclinic orbits for the forced pendulum equation. We use here that these homoclinic orbits are exponentially asymptotic to the equilibrium.*

Remark 2.5 *The following 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 [10] and [25], 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. For example it holds assuming as in [14] the condition $D^2G_\mu(A_0)\omega\cdot\omega \neq 0$.*

Proof. Assume with no loss of generality that $A_0 = 0$ and $\inf_{B_\rho(0)} G_\mu(A) = 0$. Let us choose the number of bumps k as

$$k = \Big[\frac{24\cdot\rho\cdot|I_0' - I_0|}{\delta}\Big] + 1. \tag{2.24}$$

By lemma 2.4-(i) and lemma 2.1-(ii), the trajectory 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 $\mathcal{F}_\mu^k$, defined in (2.17), such that for some positive constant K_1

$$|\overline{\theta}_k - \overline{\theta}_1| \leq K_1 \frac{|I_0 - I_0'|}{\delta}\ \rho \max\Big(|\ln\delta|, \frac{1}{\gamma\alpha^\tau}\Big). \tag{2.25}$$

More precisely we shall enforce

$$K_2|\ln\delta| < |\overline{\theta}_{i+1} - \overline{\theta}_i| < K_3 \max\Big(|\ln\delta|, \frac{1}{\gamma\alpha^\tau}\Big) \qquad \forall i = 1,\dots,k, \tag{2.26}$$

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 $\mathcal{F}_\mu^k$ we introduce suitable coordinates $(a_1,\dots,a_n,s_1,\dots,s_k) \in \mathbf{R}^n \times (-\rho,\rho)^k$ defined by

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

where η_i are constants to be chosen later. In these new coordinates the heteroclinic function defined in (2.17) is given by

$$\widetilde{\mathcal{F}}_\mu^k(a_1,a_2,\dots,a_n,s_1,\dots,s_k) = F_\mu\Big(\sum_{j=1}^n a_j\Omega_j, \frac{\eta_1 + s_1 - a_1}{|\omega|},\dots,\frac{\eta_k + s_k - a_1}{|\omega|}\Big) - |I_0' - I_0|a_2. \tag{2.27}$$

Using the invariance property (2.16) we see that $\widetilde{\mathcal{F}}_\mu^k$ does not depend on the new variable a_1 :

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

For simplicity of notation we will still denote $\widetilde{\mathcal{F}}^k_\mu(a_2,\dots,a_n,s_1,\dots,s_k) := \widetilde{\mathcal{F}}^k_\mu(0,a_2,\dots,a_n,s_1,\dots,s_k)$. We now choose the contants η_i. Let

$$D = \frac{|\omega|}{C_2}\Big|\ln\Big(\frac{24C_2}{\delta}\Big)\Big| + 2\rho, \tag{2.28}$$

where C_2 is the constant appearing in lemma 2.6. We shall use the following fact (see [4]): 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

$$d(\theta\omega, 2\pi\mathbf{Z}^n) < \alpha. \tag{2.29}$$

By (2.29) there is $(\eta_1,\dots,\eta_k) \in \mathbf{R}^k$ such that

$$\eta_i\Omega_1 \equiv \chi_i,\ \mathrm{mod} 2\pi\mathbf{Z}^n, \quad |\chi_i| < \alpha \quad \text{and} \quad \chi_i\cdot\Omega_1 = 0,\ i.e.\ \chi_i = \sum_{j=2}^n \chi_{i,j}\Omega_j. \tag{2.30}$$

$$\eta_1 = 0, \quad D \le \eta_{i+1} - \eta_i \le \Big(D + \frac{\overline{C}|\omega|}{\gamma\alpha^\tau}\Big). \tag{2.31}$$

By (2.28), (2.31), since $s_i \in (-\rho,\rho)$ we have that $\theta_{i+1} - \theta_i \ge \frac{1}{C_2}|\ln(\frac{24C_2}{\delta})|$; hence, by lemma 2.6, setting

$$\begin{aligned}\widetilde{R}_i &= \widetilde{R}_i(a_2,\dots,a_n,s_{i-1},s_i,s_{i+1})\\ &= R_i\Big(\sum_{j=2}^n a_j\Omega_j, \frac{s_{i-1}+\eta_{i-1}-a_1}{|\omega|}, \frac{s_i+\eta_i-a_1}{|\omega|}, \frac{s_{i+1}+\eta_{i+1}-a_1}{|\omega|}\Big)\end{aligned}$$

we get

$$|\widetilde{R}_i(a_2,\dots,a_n,s_{i-1},s_i,s_{i+1})| \le \frac{\delta}{24}. \tag{2.32}$$

By lemma 2.6, the invariance property (2.16), (2.30) and since G_μ is $2\pi\mathbf{Z}^n$-periodic, we have

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

where $\widetilde{G}_\mu(a_2,\dots,a_n,s) = G_\mu(\sum_{j=2}^n a_j\Omega_j + s\Omega_1)$. Since the basis $(\Omega_1,\dots,\Omega_n)$ is orthonormal the function $\widetilde{G}_\mu$ satisfies the same properties as G_μ, *i.e.*

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

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

$$W = \Big\{(a_2,\dots,a_n,s) \in \mathbf{R}^{n-1}\times\mathbf{R}^k \ \Big|\ (a_2,\dots,a_n,s_i) \in B_\rho(0), \quad \forall i = 1,\dots,k\Big\}.$$

$\widetilde{\mathcal{F}}_\mu^k$ attains its minimum over $\overline{W}$ at some point $(\overline{a},\overline{s})$. Notice that by (2.32)

$$\inf_{\overline{W}} \widetilde{\mathcal{F}}_\mu^k \leq \widetilde{\mathcal{F}}_\mu^k(0,0) = \sum_{i=1}^k \widetilde{G}_\mu(0,\chi_i) + k\frac{\delta}{24}.$$

Since $|\chi_i| < \alpha$ for all $i = 1,\dots,k$ and $\sup_{B_\alpha(0)} \widetilde{G}_\mu \leq \delta/4$, we have

$$\inf_{\overline{W}} \widetilde{\mathcal{F}}_\mu^k \leq k\frac{\delta}{4} + k\frac{\delta}{24} = k\frac{7\delta}{24}. \tag{2.33}$$

The theorem is proved if we show that $(\overline{a},\overline{s}) \in W$. Arguing by contradiction assume that $(\overline{a},\overline{s}) \in \partial W$. Then there is some $l \in \{1,\cdots,k\}$ such that $(\overline{a}+\chi_l,\overline{s}_l) \in \partial B_\rho(0)$, so that $\widetilde{G}_\mu(\overline{a}+\chi_l,\overline{s}_l) \geq \delta$. We now prove that $(\overline{a}+\chi_l,t); t \in (-\rho,\rho)\} \cap B_\rho(0) \subset Z := \{x \in B_\rho(0) \mid \widetilde{G}_\mu(x) \geq 3\delta/4\}$. Indeed, if not, by (2.32), for some $t \in (-\rho,\rho)$ such that $(\overline{a}+\chi_l,t) \in B_\rho(0)$,

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

which is wrong since $(\overline{a},\overline{s})$ is the minimum of $\widetilde{\mathcal{F}}_\mu^k$ over $\overline{W}$. We deduce in particular that, for all i, $(\overline{a}+\chi_l,\overline{s_i}) \in Z \cup B_\rho(0)^c$. Now, as $\widetilde{G}_\mu \geq 3\delta/4$ in a neighbourhood of $\partial B_\rho(0)$, our splitting condition implies that

$$d(\{x \in B_\rho(0) \mid \widetilde{G}_\mu(x) \geq 3\delta/4\} \cup B_\rho(0)^c, \{x \in B_\rho(0) \mid \widetilde{G}_\mu(x) \leq \delta/2\}) \geq 2\alpha.$$

We derive by (2.30) that for all i,

$$\widetilde{G}_\mu(\overline{a}+\chi_i,\overline{s_i}) \geq \delta/2. \tag{2.34}$$

As a consequence, noting that, from (2.24), $|I_0' - I_0|\ \rho \leq (k\delta)/24$, we deduce that

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

contradicting (2.33). The proof of the theorem is complete. ■

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

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

where $\Gamma : \mathbf{T}^n \to \mathbf{R}$ is nothing but the Poincaré-Melnikov primitive

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

Hence, if Γ possesses a proper minimum (resp. maximum) in $A_0 \in \mathbf{R}^n$, i.e $\exists r > 0$ 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 $\delta = O(\mu)$, $\rho = O(1)$ and $\alpha = O(1)$. We remark that the previous $B_r(A_0)$ 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. $\exists r > 0$ such that $\inf_{\partial B_r(A_0)} \Gamma > \Gamma(A_0)$. Then, for μ small enough, the same statement of 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 when dealing 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

$$\dot{\varphi} = \omega, \qquad \dot{I} = -\mu\partial_\varphi f(\varphi, q), \qquad \dot{q} = p, \qquad \dot{p} = \sin q - \mu\partial_q f(\varphi, q), \tag{3.1}$$

corresponding to the quasi-periodically forced pendulum

$$-\ddot{q} + \sin q = \mu\ \partial_q f(\omega t + A, q). \tag{3.2}$$

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 [18] for a 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*

$$\mathcal{T}^\mu_{I_0} = \Big\{ I = I_0 + a^\mu(\psi),\ \varphi = \psi,\ q = Q^\mu(\psi),\ p = P^\mu(\psi), \quad \psi \in \mathbf{T}^n \Big\}, \tag{3.3}$$

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

We first determine the functions $Q^\mu(\cdot), P^\mu(\cdot)$ in (3.3). Using the standard Implicit Function Theorem we prove that there exists a unique quasi-periodic solution $q^\mu_A(t)$ for the quasi-periodically forced pendulum (3.2) which bifurcates from the hyperbolic equilibrium 0.

Lemma 3.1 *Let $f \in C^l(\mathbf{T}^n \times \mathbf{T})$. For μ small enough there exists a unique quasi-periodic solution $(q^\mu_A(t), p^\mu_A(t))$ of (3.2) with $(q^\mu_A(t), p^\mu_A(t)) = O(\mu)$, 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 $(q^\mu_A(t), p^\mu_A(t)) = (Q^\mu(\omega t + A), P^\mu(\omega t + A))$.*

Proof. Let L be the Green operator of the differential operator $h \to -D^2h + h$ 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 $\mathbf{R}$, 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^μ_A 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 $q^\mu_A = O(\mu)$. By (3.2) $q^\mu_A \in C^{l+1}(\mathbf{R})$; 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

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

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

$$q^\mu_A(s) = q^\mu_{A+\omega s}(0) := Q^\mu(A + \omega s) \quad \text{and} \quad p^\mu_A(s) = p^\mu_{A+\omega s}(0) := P^\mu(A + \omega s),\ \forall s \in \mathbf{R}$$

proving the lemma.■

We now define the functions $a^\mu(\psi)$ of (3.3). 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 (3.1); hence the functions $a^\mu(\psi)$ must satisfy the following system of equations

$$(\omega \cdot \nabla) a^\mu(\psi) = \mu g^\mu(\psi), \quad \text{where} \quad g^\mu(\psi) := -(\nabla_\psi f)(\psi, Q^\mu(\psi)). \tag{3.4}$$

In order to solve (3.4) 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

$$i(k\cdot\omega)a_k = \mu g_k, \qquad \forall k\in\mathbf{Z}^n. \tag{3.5}$$

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*

$$\int_{\mathbf{T}^n} (\nabla_\psi f)(\psi, Q^\mu(\psi))\ d\psi = 0. \tag{3.6}$$

Proof. For all $i = 1,\dots,n$

$$\partial_{\psi_i} f(\psi, Q^\mu(\psi)) = \frac{d}{d\psi_i} f(\psi, Q^\mu(\psi)) - \partial_q f(\psi, Q^\mu(\psi))\partial_{\psi_i} Q^\mu(\psi). \tag{3.7}$$

Since $(q_A^\mu(t), p_A^\mu(t))$ satisfies the pendulum equation $\sum_{j=1}^n \omega_j\partial_{\psi_j} Q^\mu(\psi) = P^\mu(\psi)$, $\sum_{j=1}^n \omega_j\partial_{\psi_j} P^\mu(\psi) = \sin Q^\mu\psi) - \mu\partial_q f(\psi, Q^\mu(\psi))$ and we deduce that

$$-\partial_q f(\psi, Q^\mu(\psi))\partial_{\psi_i} Q^\mu(\psi) = \frac{1}{\mu}\Big(\frac{d}{d\psi_i}\cos Q^\mu(\psi) + \sum_{j=1}^n \omega_j\partial_{\psi_j} P^\mu \partial_{\psi_i} Q^\mu(\psi)\Big). \tag{3.8}$$

We now prove that

$$\sum_{j=1}^n \omega_j\partial_{\psi_j} P^\mu\partial_{\psi_i} Q^\mu = \frac{d}{d\psi_i}\frac{(P^\mu(\psi))^2}{2} + \sum_{j\neq i}\omega_j\Big(\partial_{\psi_i}(Q^\mu\partial_{\psi_j}P^\mu) - \partial_{\psi_j}(Q^\mu\partial_{\psi_i}P^\mu)\Big). \tag{3.9}$$

Indeed

$$\begin{aligned}
\sum_{j=1}^n \omega_j\partial_{\psi_j} P^\mu\partial_{\psi_i} Q^\mu &= \omega_i\partial_{\psi_i} P^\mu\partial_{\psi_i} Q^\mu + \sum_{j\neq i}\omega_j\partial_{\psi_j} P^\mu\partial_{\psi_i} Q^\mu \\
&= \Big(\sum_{j=1}^n \omega_j\partial_{\psi_j} Q^\mu\Big)\partial_{\psi_i} P^\mu + \sum_{j\neq i}\omega_j\Big(\partial_{\psi_j} P^\mu\partial_{\psi_i} Q^\mu - \partial_{\psi_j} Q^\mu\partial_{\psi_i} P^\mu\Big) \\
&= \frac{d}{d\psi_i}\frac{(P^\mu(\psi))^2}{2} + \sum_{j\neq i}\omega_j\Big(\partial_{\psi_i}(Q^\mu\partial_{\psi_j}P^\mu) - \partial_{\psi_j}(Q^\mu\partial_{\psi_i}P^\mu)\Big).
\end{aligned}$$

From (3.7), (3.8), (3.9) we finally obtain that

$$\partial_{\psi_i} f(\psi, Q^\mu) = \frac{d}{d\psi_i}\Big(f(\psi, Q^\mu) + \frac{1}{\mu}\cos Q^\mu + \frac{1}{\mu}\frac{P^{\mu 2}}{2}\Big) + \frac{1}{\mu}\sum_{j\neq i}\omega_j\Big(\partial_{\psi_i}(Q^\mu\partial_{\psi_j}P^\mu) - \partial_{\psi_j}(Q^\mu\partial_{\psi_i}P^\mu)\Big)$$

from which property (3.6) follows. ■

Since ω satisfies $(H1)$ and f is sufficiently smooth the function a^μ defined by

$$a^\mu(\psi) = \sum_{k\in\mathbf{Z}^n, k\neq 0}\frac{g_k}{i(k\cdot\omega)}e^{ik\cdot\psi}, \tag{3.10}$$

which formally solves equation (3.5), is well defined and smooth. Indeed since $f\in C^l$ the function g^μ defined in (3.4) is C^{l-1} and there exists $M>0$ such that $|g_k|\le M/|k|^{l-1}$, $\forall k\in\mathbf{Z}^n$, $k\neq 0$. By $(H1)$ it follows that $|a_k|\le M/|k|^{l-1}|\omega\cdot k| \le M|k|^\tau/(\gamma|k|^{l-1})$. 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 we can prove the following lemma

Lemma 3.3 *The transformation of coordinates* $(J,\psi,u,v)\to(I,\varphi,q,p)$ *defined on the covering space* $\mathbf{R}^{2(n+1)}$ *of* $\mathbf{T}^n\times\mathbf{R}^n\times\mathbf{T}\times\mathbf{R}$ *by*

$$I=a^\mu(\psi)+u\partial_\psi P^\mu(\psi)-v\partial_\psi Q^\mu(\psi)+J,\quad \varphi=\psi,\quad q=Q^\mu(\psi)+u,\quad p=P^\mu(\psi)+v \tag{3.11}$$

is symplectic.

Proof. Set $dI\wedge d\varphi=\sum_{i=1}^n dI_i\wedge d\varphi_i$ and $dJ\wedge d\psi=\sum_{i=1}^n dJ_i\wedge d\psi_i$. We have

$$\begin{aligned} dI\wedge d\varphi+dp\wedge dq &= \sum_{i=1}^n da_i^\mu(\psi)\wedge d\psi_i+d(u\partial_{\psi_i}P^\mu(\psi))\wedge d\psi_i-d(v\partial_{\psi_i}Q^\mu(\psi))\wedge d\psi_i\\ &+ dJ\wedge d\psi+dv\wedge du+dP^\mu(\psi)\wedge dQ^\mu(\psi)+dP^\mu(\psi)\wedge du+dv\wedge dQ^\mu(\psi).\end{aligned}$$

Using that the tori $\mathcal{T}^\mu_{I_0}$ are isotropic, that is $\sum_{i=1}^n da_i^\mu(\psi)\wedge d\psi_i+dP^\mu(\psi)\wedge dQ^\mu(\psi)=0$, and noticing that $\sum_{i,j}u\partial^2_{\psi_i,\psi_j}P^\mu(\psi)d\psi_j\wedge d\psi_i=0=\sum_{i,j}v\partial^2_{\psi_i,\psi_j}Q^\mu(\psi)d\psi_j\wedge d\psi_i$ we deduce

$$\begin{aligned} dI\wedge d\varphi+dp\wedge dq &= \sum_{i=1}^n d(u\partial_{\psi_i}P^\mu(\psi))\wedge d\psi_i-d(v\partial_{\psi_i}Q^\mu(\psi))\wedge d\psi_i\\ &+ dJ\wedge d\psi+dv\wedge du+dP^\mu(\psi)\wedge du+dv\wedge dQ^\mu(\psi)\\ &= dJ\wedge d\psi+dv\wedge du+\sum_{i=1}^n\partial_{\psi_i}P^\mu(\psi)du\wedge d\psi_i-\partial_{\psi_i}Q^\mu(\psi)dv\wedge d\psi_i\\ &+ dP^\mu(\psi)\wedge du+dv\wedge dQ^\mu(\psi)\\ &= dJ\wedge d\psi+dv\wedge du,\end{aligned}$$

and the transformation (3.11) 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 $\mathcal{T}^\mu_0=\{(a^\mu(\psi),\psi,Q^\mu(\psi),P^\mu(\psi));\psi\in\mathbf{T}^n\}$. Hamiltonian $(\mathcal{K}_\mu)$ corresponds to the quasi-periodically forced pendulum equation

$$-\ddot u+\sin u=\partial_u P_0(\mu,u,\omega t+A). \tag{3.12}$$

of Lagrangian

$$L_\mu=\frac{\dot u^2}{2}+(1-\cos u)-P_0(\mu,u,\omega t+A). \tag{3.13}$$

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 (3.12) 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, $\forall i \in \mathbf{Z}$, 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. The term $2\pi \dot q^\mu_A(\theta)$ takes into account that the stable and the unstable manifolds of the tori $\mathcal{T}^\mu_{I_0}$ are not exact Lagrangian manifolds, see [20]. We define the *"homoclinic function"* $\mathcal{G}_\mu : \mathbf{T}^n \to \mathbf{R}$ as

$$\mathcal{G}_\mu(A) = \mathcal{F}_\mu(A, 0). \tag{3.14}$$

It holds also $\mathcal{F}_\mu(A,\theta) = \mathcal{G}_\mu(A + \omega\theta), \quad \forall \theta \in \mathbf{R}$. Arguing as in lemma 2.4 we can prove the existence of k-bump pseudo-homoclinic solutions $u^L_{A,\theta}$, which is a true solution of (3.12) in each interval $(-\infty,\theta_1)$, (θ_i,θ_{i+1}) $(i = 1,\dots,k-1)$, $(\theta_k,+\infty)$, and satisfying all the properties of lemma 2.4. Then we define the "k-bump heteroclinic function"

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

If $\partial_{\theta_i}\mathcal{F}^k_\mu(A,\theta_1,\dots,\theta_k) = (\dot u^L_{A,\theta})^2(\theta_i^-)/2 - (\dot u^L_{A,\theta})^2(\theta_i^+)/2 = 0$ then $u^L_{A,\theta}$ is a true solution of the quasi-periodically forced pendulum (3.12). As in the previous section the variation in the action variables is given by the partial derivative with respect to A, that is

$$\partial_A \mathcal{F}^k_\mu(A,\theta) = \int_{-\infty}^{+\infty} -\mu\Big(\partial_\varphi f(\omega t + A, q^\mu_A(t) + u^L_{A,\theta}(t)) - \partial_\varphi f(\omega t + A, q^\mu_A(t))\Big)\ dt - (I'_0 - I_0). \tag{3.15}$$

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 (3.15) it is easy to verify that the solutions of (3.1) $(I_\mu(t), \omega t + A, q^\mu_A + u^L_{A,\theta}, \dot q^\mu_A + \dot u^L_{A,\theta})$, 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 given in theorem 2.3 holds.*

A Taylor expansion in μ gives

Lemma 3.5 *For μ small enough*

$$\mathcal{G}_\mu(A) = const + \mu M(A) + O(\mu^2), \qquad \forall A \in \mathbf{T}^n \tag{3.16}$$

where $M(A)$ is the Poincaré-Melnikov primitive $M(A) = \int_{-\infty}^{+\infty} \Big[f(\omega t + A, q_0(t)) - f(\omega t + A, 0) \Big]\, dt$.

Proof. We develop with a Taylor expansion in μ the Lagrangian L_μ defined in (3.13)

$$L_\mu = \frac{\dot u^2}{2} + (1-\cos u) + \mu\Big((u - \sin u)\gamma + f(\omega t + A, u) - f(\omega t + A, 0) - \partial_q f(\omega t + A, 0) u \Big) + \mathcal{R}(\mu, u, t) \tag{3.17}$$

where $\gamma(t) := \partial_{\mu_{|\mu=0}} q_A^\mu(t)$, $|\mathcal{R}(\mu, u, t)| = (\mu^2)$, $\mathcal{R}(\mu, 0, t) = 0$ and $\partial_u \mathcal{R}(\mu, 0, t) = 0$. The Melnikov function corresponding to Lagrangian (3.17) is

$$M^*(A) = \int_{\mathbf{R}} (q_0(t) - \sin q_0(t))\gamma(t) + f(\omega t + A, q_0(t)) - f(\omega t + A, 0) - \partial_q f(\omega t + A, 0) q_0(t)\, dt. \tag{3.18}$$

Integrating by parts, since $-\ddot\gamma + \gamma = \partial_q f(\omega t + A, 0)$, we have

$$\int_{\mathbf{R}} (q_0(t) - \sin q_0(t))\gamma(t)\, dt = \int_{\mathbf{R}} (q_0(t) - \ddot q_0(t))\gamma(t)\, dt = \int_{\mathbf{R}} (-\ddot\gamma(t) + \gamma(t)) q_0(t) dt = \int_{\mathbf{R}} \partial_q f(\omega t + A, 0) q_0(t) dt,$$

and we deduce from (3.18) that $M^*(A) = M(A) = \int_{\mathbf{R}} [f(\omega t + A, q_0(t)) - f(\omega t + A, 0)]\, dt$. ■

Theorem 3.3 *Assume $(H1)$ and let M possess a proper minimum (or maximum) A_0, i.e. $\exists r > 0$ such that $\inf_{\partial B_r(A_0)} \Gamma > \Gamma(A_0)$. Then, for μ small enough, the same statement of 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 [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 [10], 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 is a suitable positive constant.

4 Splitting of separatrices

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 analytical, the oscillations of the Melnikov function along some directions turn out to be exponentially small with respect to ε. Hence the development (3.16) 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. The classical way to overcome this difficulty would be to extend analytically the function $F_\mu(A, \theta)$ for complex values of the variables, see [2]-[13] and [23]. 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 $q_{A,\theta}^\mu(Re\ \theta) = \pi$ appearing naturally when we try to extend the definition of $q_{A,\theta}^\mu$ to $\theta \in \mathbf{C}$ breaks analyticity). We bypass this problem considering the action functional evaluated on different "1-bump pseudo-homoclinic solutions" $Q_{A,\theta}^\mu$. This new "reduced action functional" $\widetilde F_\mu(A, \theta) = \widetilde G_\mu(A + \omega\theta)$ has the advantage to have an analytical 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

$$\forall s \in \mathbf{N},\ \exists C_s > 0 \quad \text{such that} \quad |f_k| \leq \frac{C_s}{|k|^s} \exp\Big(- \sum_{i=1}^n r_i |k_i| \Big),\ \forall k \in \mathbf{Z}^n. \tag{4.1}$$

Condition (4.1) 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

$$||f|| := \sup_{A \in D} |f(A)|. \tag{4.2}$$

It will be used from subsection 4.2.

4.1 The change of coordinates

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 $\int_{\mathbf{R}} \psi_0(t) \dot{q}_0(t)\, dt = \gamma \neq 0$. Arguing as in lemma 2.1 we can prove

Lemma 4.1 *For μ small enough (independently of ω), $\forall \theta \in \mathbf{R}$, 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)* $\max\Big(|Q^\mu_{A,\theta}(t) - q_\theta(t)|, |\dot{Q}^\mu_{A,\theta}(t) - \dot{q}_\theta(t)| \Big) = O\Big(\mu \exp(-\frac{|t-\theta|}{2}) \Big)$;
- *(iv)* $\max\Big(|\partial_A Q^\mu_{A,\theta}(t)|, |\partial_A \dot{Q}^\mu_{A,\theta}(t)|, |\omega.\partial_A Q^\mu_{A,\theta}(t)|, |\omega.\partial_A \dot{Q}^\mu_{A,\theta}(t)| \Big) = O\Big(\mu \exp(-\frac{|t-\theta|}{2}) \Big)$.

Moreover $Q^\mu_{A,\theta}(t) = Q^\mu_{A+k2\pi,\theta}(t)$, $\forall k \in \mathbf{Z}^n$ and $Q^\mu_{A,\theta+\eta}(t+\eta) = Q^\mu_{A+\omega\eta,\theta}(t)$, $\forall \theta, \eta \in \mathbf{R}$.

We define the function $\widetilde{F}_\mu(A, \theta) : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$ as the action functional of Lagrangian 2.3 evaluated on the "1-bump pseudo-homoclinic solutions" $Q^\mu_{A,\theta}(t)$ obtained in lemma 4.1, namely

$$\widetilde{F}_\mu(A, \theta) := \int_{\mathbf{R}} \mathcal{L}_\mu(Q^\mu_{A,\theta}(t), \dot{Q}^\mu_{A,\theta}(t), t)\ dt \tag{4.3}$$

and $\widetilde{G}_\mu(A) : \mathbf{T}^n \to \mathbf{R}$ as $\widetilde{G}_\mu(A) = \widetilde{F}_\mu(A, 0)$. 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 $\widetilde{F}_\mu(A, \theta) = \widetilde{G}_\mu(A + \omega\theta)$, $\forall \theta \in \mathbf{R}$.

Remark 4.1 *By lemma 4.1-(i)-(ii), if $\partial_\theta \widetilde{F}_\mu(A, \theta) = 0$ then $Q^\mu_{A,\theta}$ is a true solution of (2.2). More precisely we have $|\alpha^\mu_{A,\theta}| \leq C |\partial_\theta \widetilde{F}_\mu(A, \theta)|$, for a suitable positive contant $C > 0$. In addition we could easily prove using lemma 4.1 that*

$$|\nabla^s \widetilde{G}_\mu(A)| = O(\mu), \qquad |\nabla^s \widetilde{F}_\mu(A, \theta)| = O(\mu), \qquad s = 1, 2. \tag{4.4}$$

The relation between the two functions $F_\mu(A, \theta)$ and $\widetilde{F}_\mu(A, \theta)$ is given below. The next theorem is formulated to handle with also non-analytical perturbations f. For the analytical case see remark 4.3.

Theorem 4.1 *For μ small enough (independently of ω) there exists a Lipschitz continuous function $\overline{h}_\mu : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$, with $\overline{h}_\mu(A,\theta) = O(\mu)$, $|\overline{h}_\mu(A',\theta') - \overline{h}_\mu(A,\theta)| = O(\sqrt{\mu}(|A'-A| + |\theta'-\theta|))$, $\overline{h}_\mu(A,\theta+\eta) = \overline{h}_\mu(A+\eta\omega,\theta)$, such that $F_\mu(A,\theta) = \widetilde{F}_\mu(A,\theta+\overline{h}_\mu(A,\theta))$. In particular, setting $\overline{g}_\mu(A) = \overline{h}_\mu(A,0)$, the homeomorphism $\psi_\mu : \mathbf{T}^n \to \mathbf{T}^n$ given by $\psi_\mu(A) = A + \overline{g}_\mu(A)\omega$ satisfies $G_\mu = \widetilde{G}_\mu \circ \psi_\mu$.*

In order to prove theorem 4.1 we need the next two lemmas, proved in the appendix.

Lemma 4.2 *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 $V_\mu(A,\theta) = F_\mu(A,\theta + l_\mu(A,\theta))$.

Lemma 4.3 *There exists a positive constant C_4 such that, for all $(A,\theta) \in \mathbf{T}^n \times \mathbf{R}$, there holds*

$$|\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta)| \le C_4 |\partial_\theta \widetilde{F}_\mu(A,\theta)|^2.$$

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

Proof of Theorem 4.1. By lemma 4.2, there is a smooth function $\overline{l}_\mu$ such that $\overline{l}_\mu(A,\theta) = O(\mu)$, $\nabla \overline{l}_\mu(A,\theta) = O(\mu)$ $\overline{l}_\mu(A,\theta+\eta) = \overline{l}_\mu(A+\eta\omega,\theta)$ and $F_\mu(A,\theta) = V_\mu(A,\theta+\overline{l}_\mu(A,\theta))$. So it is enough to find $h = h_\mu(A,\theta)$ such that

$$V_\mu(A,\theta) = \widetilde{F}_\mu(A,\theta+h). \tag{4.5}$$

$\overline{h}_\mu(A,\theta)$ will be then defined by

$$\overline{h}_\mu(A,\theta) = \overline{l}_\mu(A,\theta) + h_\mu(A,\theta+\overline{l}_\mu(A,\theta)). \tag{4.6}$$

Note that if $\partial_\theta \widetilde{F}_\mu(A,\theta) = 0$ then, by lemma 4.3, equation (4.5) is solved by $h = 0$. In general we look for h of the form $h = \partial_\theta \widetilde{F}_\mu(A,\theta) g$. Then we can write

$$\begin{aligned} \widetilde{F}_\mu(A,\theta+h) &= \widetilde{F}_\mu(A,\theta) + \partial_\theta \widetilde{F}_\mu(A,\theta) h + R_\mu(A,\theta,h) h^2 \\ &= \widetilde{F}_\mu(A,\theta) + (\partial_\theta \widetilde{F}_\mu(A,\theta))^2 g + R_\mu(A,\theta,\partial_\theta \widetilde{F}_\mu(A,\theta) g)(\partial_\theta \widetilde{F}_\mu(A,\theta))^2 g^2 \end{aligned} \tag{4.7}$$

where

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

is smooth and, by the estimates (4.4) on the derivatives of $\widetilde{F}_\mu$, it satisfies $R_\mu(A,\theta,h) = O(\mu)$, $\partial_h R_\mu(A,\theta,h) = O(\mu/|h|)$. By (4.7) equation (4.5) is then equivalent to

$$\frac{V_\mu(A,\theta) - \widetilde{F}_\mu(A,\theta)}{(\partial_\theta \widetilde{F}_\mu(A,\theta))^2} = g + R_\mu(A,\theta,\partial_\theta \widetilde{F}_\mu(A,\theta) g) g^2$$

We have $R_\mu(A,\theta,\partial_\theta \widetilde{F}_\mu(A,\theta) g) g^2 = O(\mu g^2)$ and $\partial_g\Big(R_\mu(A,\theta,\partial_\theta \widetilde{F}_\mu(A,\theta) g) g^2\Big) = O(\mu g)$. By the contraction mapping theorem, for μ small enough, for all $y \in \mathbf{R}$ such that $|y| < 2C_4$, there exists a unique solution $g = \varphi(\mu,A,\theta,y)$ of the equation

$$y = g + R_\mu(A,\theta,\partial_\theta \widetilde{F}_\mu(A,\theta) g) g^2, \tag{4.8}$$

such that $|g| < 3C_4$. Moreover, the function φ defined in this way is smooth. Setting

$$h_\mu(A,\theta) := \varphi\Big(\mu, A, \theta, \frac{V_\mu(A,\theta) - \widetilde{F}_\mu(A,\theta)}{(\partial_\theta \widetilde{F}_\mu(A,\theta))^2}\Big) \partial_\theta \widetilde{F}_\mu(A,\theta) \tag{4.9}$$

if $\partial_\theta \widetilde{F}_\mu(A,\theta) \neq 0$ and $h_\mu(A,\theta) = 0$ if $\partial_\theta \widetilde{F}_\mu(A,\theta) = 0$, we get a continuous function h_μ which satisfies (4.5) and $|h_\mu(A,\theta)| \le 3C_4 |\partial_\theta \widetilde{F}_\mu(A,\theta)|$, which implies $|h_\mu| = O(\mu)$. Moreover h_μ is the unique function that enjoys this properties.

By (4.9) the restriction of h_μ to

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

is smooth. Deriving the identity $V_\mu(A,\theta) = \widetilde{F}_\mu(A,\theta + h_\mu(A,\theta))$ we obtain

$$\partial_\theta \widetilde{F}_\mu(A,\theta + h_\mu(A,\theta))\nabla h_\mu(A,\theta)) = \nabla V_\mu(A,\theta) - \nabla \widetilde{F}_\mu(A,\theta + h_\mu(A,\theta))$$

for $(A,\theta) \in U_\mu$. Now, from

$$|\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta)| \leq C_4 |\partial_\theta \widetilde{F}_\mu(A,\theta)|^2, \quad \partial^2_{\theta\theta} V_\mu(A,\theta) = O(\mu), \quad \partial^2_{\theta\theta} \widetilde{F}_\mu(A,\theta) = O(\mu),$$

we can derive by an elementary argument whose main ingredient is Taylor formula that

$$|\nabla \widetilde{F}_\mu(A,\theta) - \nabla V_\mu(A,\theta)| = O(\sqrt{\mu}|\partial_\theta \widetilde{F}_\mu(A,\theta)|). \tag{4.10}$$

Moreover, by the estimates of $|\partial^2_{\theta\theta}\widetilde{F}_\mu|$ and $|h_\mu|$, $(\partial_\theta \widetilde{F}_\mu)(A,\theta + h_\mu(A,\theta)) = (1 + O(\mu))\partial_\theta \widetilde{F}_\mu(A,\theta)$. Hence by (4.10) $|\nabla h_\mu| = O(\sqrt{\mu})$ uniformly in U_μ. Since h_μ is continuous and $h_\mu(A,\theta) = 0$ if $(A,\theta) \notin U_\mu$, the Lipschitz continuity of h_μ follows.

To complete the proof, we observe that $h_\mu(A,\theta+\eta) = h_\mu(A+\eta\omega,\theta)$, which is a consequence of uniqueness. Hence, by (4.6) and the properties of $\overline{l}_\mu$, $\overline{h}_\mu$ satisfies the same. ■

Remark 4.2 *Assume that $\widetilde{G}_\mu$ satisfies the "splitting condition" 2.1 (or its generalisation introduced in remark 2.3), with bounds δ and α. Then, by theorem 4.1, G_μ too satisfies this condition, with constants δ', α'. We can take $\delta' = \delta$; moreover, at least if $\sqrt{\mu}|\omega|$ is small, we can take $\alpha' = \alpha/2$. As a consequence, the results that we shall obtain in the next section proving a "splitting condition" for $\widetilde{G}_\mu$, may be used to apply the shadowing theorem 2.3.*

Remark 4.3 *Assume that $r_i > 0$ for all i (i.e. that the perturbation f is analytical). 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 analytical 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 analytical. One could check that l_μ and V_μ, defined in lemma 4.2 have analytical extensions too, and that the inequality of lemma 4.3 still holds in the new set of definition. Moreover in the next lemma 4.4 it is proved that $\widetilde{F}_\mu$ is analytical 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 analytical, 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 analytical diffeomorphism.*

4.2 Analytical extension

The unpertubed homoclinic $q_0(t) = 4\text{arctg } e^t$ can be extended to a holomorphic function over the strip $S := \mathbf{R} + i(-\pi/2, \pi/2)$. Moreover equation (2.2) may be considered also for complex values of q and, for $\mu = 0$, q_θ is a solution of (2.2) for all $\theta \in S$. We shall use the notation $S_\delta = \mathbf{R} + i(-(\frac{\pi}{2} - \delta), \frac{\pi}{2} - \delta)$, for $\delta \in (0, \pi/2)$. 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 $\theta \in S_\delta, Re(\theta) = 0$. The following estimates hold, where $t \in \mathbf{R}$.

$$|\dot{q}_\theta(t)| \leq \frac{C}{\min\{(|t|+\delta), 1\}} \exp(-|t|); \tag{4.11}$$

$$|\sin q_\theta(t)| \leq \frac{C}{\min\{(|t|+\delta)^2, 1\}} \exp(-|t|); \tag{4.12}$$

$$|\cos q_\theta(t)| \leq \frac{C}{\min\{(|t|+\delta)^2, 1\}}; \tag{4.13}$$

$$\frac{1}{|\dot{q}_\theta(t)|} \leq C \exp(|t|) \min\{(|t|+\delta), 1\}. \tag{4.14}$$

In what follows we consider the Banach spaces

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

and

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

endowed with norm

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

The next lemma extends lemma 4.1 for complex values of the variables. First note that the function $\psi_0(t)$ can be extended to a holomorphic function on $\mathbf{R} + i(-\pi, \pi)$. Recalling the definition for $||f||$ given in (4.2), we have

Lemma 4.4 *There exist positive constants η, C_5 such that for all $\delta \in (0, \pi/2)$, $\forall 0 < \mu \leq (\eta\delta^3)/||f||$, for all ω, for all $A \in D$, for all $\theta \in S_\delta$ 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 $||w^\mu_{A,\theta}||_{2,\delta} + |\nu^\mu_{A,\theta}| + |\alpha^\mu_{A,\theta}| \leq C_5\mu||f||/\delta^2$;*
- *$-\ddot{Q}^\mu_{A,\theta}(t) + \sin Q^\mu_{A,\theta}(t) = \mu \ \sin Q^\mu_{A,\theta}(t) \ f(\omega t + A) + \alpha^\mu_{A,\theta}\psi_\theta(t)$;*
- *$\int_{\mathbf{R}} (Q^\mu_{A,\theta}(t) - q_\theta(t))\psi_\theta(t) \ dt = 0$.*

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

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

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

For $\theta' \in S_{\delta/2}$ such that $|\theta' - \theta| \leq \delta/2$ 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 $-\ddot{y} + \cos q_{\theta'} y = 0$ we can compute the inverse of $L_{\theta'}$. It is given by $L^{-1}_{\theta'}(g) = (w, \alpha)$ with

$$\alpha = -\frac{\int_{\mathbf{R}} g(t)\dot{q}_{\theta'}(t) \ dt}{\int_{\mathbf{R}} \psi_\theta(t)\dot{q}_{\theta'}(t) \ dt}, \tag{4.15}$$

$$w(t) = \dot{q}_{\theta'}(t)\Big[\int_0^t -\frac{1}{\dot{q}^2_{\theta'}(s)} \Big(\int_{-\infty}^s (g(\sigma) + \alpha\psi_\theta(\sigma))\dot{q}_{\theta'}(\sigma) \ d\sigma \Big) \ ds \Big] \tag{4.16}$$

$$= \dot{q}_{\theta'}(t)\Big[\int_0^t \frac{1}{\dot{q}^2_{\theta'}(s)} \Big(\int_s^{+\infty} (g(\sigma) + \alpha\psi_\theta(\sigma))\dot{q}_{\theta'}(\sigma) \ d\sigma \Big) \ ds \Big]. \tag{4.17}$$

Note that since $|\theta - \theta'| \leq \delta/2$, $Re(\theta') \leq \delta/2$. Therefore estimates (4.11)-(4.14) hold as well (with perhaps different constants) when θ is replaced by θ'. We derive from (4.15)-(4.17) that

$$|\alpha| + ||w||_{2,\delta} \leq \frac{C}{\delta}||g||_{-1,\delta}. \tag{4.18}$$

2nd Step. We shall search Q as $Q = q_{\theta+\nu} + w$ with $|\nu| < \delta/2$, $w \in \overline{X}$. Let B denote the open ball of radius $\delta/2$ 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 $||(\nu, w, \alpha)||_2 = |\nu| + ||w||_{2,\delta} + |\alpha|$ on $B \times \overline{X} \times \mathbf{C}$ and $||(g, \beta)||_{-1} = ||g||_{-1,\delta} + |\beta|$ 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 $||(\nu, w, \alpha)||_2/\delta$ and $\mu||f||/\delta^3$ 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, for a suitable positive constant C

$$||T_\nu^{-1}(g, \beta)||_2 \le \frac{C}{\delta}||(g, \beta)||_{-1}. \tag{4.19}$$

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

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

As a consequence, recalling (4.19), if $\mu||f||/\delta^3 \le K_0$ and $||w||_{2,\delta}/\delta \le K_0$, for K_0 small enough, then $DJ_\mu(\mu, w, \alpha)$ is invertible and

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

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 $B(K_2\delta)$, ball centered at the origin and of radius $K_2\delta$ 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, there holds

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

hence there is a constant C' such that

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

So, if $J_0(\nu, w, \alpha) = 0$ then, by (4.19)

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

Let $K_2 < 1/(CC')$. By the latter inequality, if $J_0(\nu, w, \alpha) = 0$ and $||(\nu, w, \alpha)||_2 \leq K_2\delta$, 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 $||(\nu, w, \alpha)||_2 \leq K_2\delta$) $\iff$ $\nu = w = \alpha = 0$;
- (ii) If $|\nu| < \delta/2$, $||w||_{2,\delta} \leq K_0\delta$, $\mu||f|| \leq K_0\delta^3$ then $DJ_\mu(\nu, w, \alpha)$ is invertible and $||(DJ_\mu(\nu, w, \alpha))^{-1}|| \leq K_1/\delta$.

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

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

We say that $(i), (ii), (iii)$ imply that there is η such that, for all $0 < \mu < \eta\delta^3/||f||$, the equation $J_\mu(\nu, w, \alpha) = 0$ has a unique solution such that $||(\nu, w, \alpha)||_2 < K_2\delta/2$. In addition $||(\nu, w, \alpha)||_2 = O(\mu||f||/\delta^2)$. To prove existence, we can proceed as follows. Let $\mathcal{S}$ denote the set of all $\mu \in [0, K_0\delta^3/||f||]$ such that there exists a C^1 function $S_\mu : [0, \mu] \to \{(\nu, w, \alpha) \ : \ ||(\nu, w, \alpha)||_2 < K_2\delta/2\}$ 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 $S : [0, \overline{\mu}) \to \{(\nu, w, \alpha) \ : \ ||(\nu, w, \alpha)||_2 < K_2\delta/2\}$ such that $S(t) = S_\mu(t)$ for all $\mu \in (0, \overline{\mu})$. By (ii) and (iii), we can write, for all $t \in (0, \overline{\mu})$,

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

Hence

$$||S(t)||_2 \leq \frac{K_1K_3}{\delta^2}||f||t|. \tag{4.20}$$

Now, since $S'(t)$ is bounded, $S(t)$ converges to some $\overline{S}$ as $t \to \overline{\mu}$. Either $\overline{\mu} = K_0\delta^3/||f||$ or $||\overline{S}||_2 = K_2\delta/2$ (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 (4.20),

$$\overline{S} = \frac{K_2\delta}{2} \leq \frac{K_1K_3}{\delta^2}\overline{\mu}||f||.$$

So the existence assertion holds for $0 < \mu < \eta\delta^3/||f||$, 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 $||b_i||_2 < K_2\delta/2$, $J_\mu(b_i) = 0$. Then, by the same argument as previously, we can prove the existence of two functions of class C^1 $S_1, S_2 \ : \ [0, \mu] \to \{b \ : \ ||b||_2 < K_2\delta\}$ 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 $||w^\mu_{A,\theta}||_{2,\delta} + |\nu^\mu_{A,\theta}| + |\alpha^\mu_{A,\theta}|$ given in the statement is a direct consequence of (4.20).

To complete the proof, we point out that J_μ is analytical on (A, θ). 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 analytical extension of the function $\widetilde{F}_\mu(A,\theta)$ for $(A,\theta)\in D\times S_\delta$

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

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

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

We have $\Gamma(A+\omega\theta)=M(A,\theta)$. We now prove

Lemma 4.5 *For $\mu||f||\delta^{-3}$ small enough, for all $(A,\theta)\in D\times S_\delta$, we have*

$$\widetilde{F}_\mu(A,\theta)=const+\mu M(A,\theta)+O\Big(\frac{\mu^2||f||^2}{\delta^4}\Big). \tag{4.21}$$

Proof. We have $Q^\mu_{A,\theta}=q_{\theta+\nu^\mu_{A,\theta}}+w^\mu_{A,\theta}$ and we set for brevity $Q^\mu_{A,\theta}=q_{\theta+\nu}+w$.

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

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

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

For example we can get that $\int_{\mathbf{R}}\cos q_{\theta+\nu}-\cos(q_{\theta+\nu}+w)-(\sin q_{\theta+\nu})w=O(\mu^2||f||^2/\delta^4)$ 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 (4.12)-(4.13) togheter with $||w||_{2,\delta}\le\mu||f||/\delta^2$. Moreover

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

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

$$\Gamma_k=f_k\frac{2\pi(k\cdot\omega)}{\sinh((k\cdot\omega)\frac{\pi}{2})}. \tag{4.22}$$

By estimate (4.21), since $\widetilde{F}_\mu(A,\theta)=\widetilde{G}_\mu(A+\omega\theta)$ and $M(A,\theta)=\Gamma(A+\omega\theta)$, via a standard lemma on Fourier coefficients of analytical functions (lemma 3 in [13]), we obtain the following result (compare with theorem 3.4.5 in [20]).

Theorem 4.2 *There exists a positive constant C_6 such that, for $\mu||f||\delta^{-3}$ small enough, then $\forall k\neq 0, k\in\mathbf{Z}^n$, for all $\delta\in(0,\frac{\pi}{2})$, for all ω,*

$$|\widetilde{G}_k-\mu\Gamma_k|\le\frac{C_6\mu^2||f||^2}{\delta^4}\exp\Big(-\sum_{i=1}^n r_i|k_i|\Big)\exp\Big(-|k\cdot\omega|\Big(\frac{\pi}{2}-\delta\Big)\Big). \tag{4.23}$$

5 Three time scales

We consider in this section three time scales systems as (see [15] and [22])

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

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 $\beta = (\beta_2, \ldots, \beta_n) \in \mathbf{R}^{n-1}$ is given.

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

Given $\kappa_2 = (k_2, \ldots, k_n) \in \mathbf{Z}^{n-1}$, we shall use the notation $\kappa_2^+ := (|k_2|, \ldots, |k_n|)$. Moreover we shall use the abbreviation $\rho_2 := (r_2, \ldots, r_n)$, so that $\kappa_2^+ \cdot \rho_2 := \sum_{i=2}^n r_i |k_i|$. We recall that $r_1, \ldots, r_n$ are defined in formula (4.1).

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 analytical w.r.t φ_2. More precisely, $r_1 = 0$ and for $i \geq 2$, $r_i > 0$. If $a = 0$, we impose in addition that $r_i > |\beta_i| \pi/2$ for $i \geq 2$.

We shall use (4.23) in order to give an expansion for the "homoclinic function"

$$\widetilde{G}_\mu(A) = \sum_{(k_1, \kappa_2) \in \mathbf{Z} \times \mathbf{Z}^{n-1}} \widetilde{G}_{k_1, \kappa_2} \exp(i(k_1 A_1 + \kappa_2 \cdot A_2)) = \sum_{k_1 \in \mathbf{Z}} \widetilde{g}_{k_1}(A_2) \exp(i k_1 A_1).$$

We start with

Lemma 5.1 *There exists a positive constant C_7 such that, for $\mu ||f|| \varepsilon^{-3/2}$ small enough,*

$$\sum_{\kappa_2 \in \mathbf{Z}^{n-1}, |k_1| \geq 2} |\widetilde{G}_{k_1, \kappa_2}| \leq C_7 \frac{\mu ||f||}{\sqrt{\varepsilon}} \exp(-\frac{\pi}{\sqrt{\varepsilon}}). \tag{5.1}$$

Proof. Choosing $\delta = \sqrt{\varepsilon}$, we get from (4.22) and (4.23)

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

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 (5.1). ■

Let

$$\Gamma(\varepsilon, A) = \sum_{(k_1, \kappa_2) \in \mathbf{Z} \times \mathbf{Z}^{n-1}} \Gamma_{k_1, \kappa_2} \exp(i(k_1 A_1 + \kappa_2 \cdot A_2)) = \sum_{k_1 \in \mathbf{Z}} \Gamma_{k_1}(\varepsilon, A_2) \exp(i k_1 A_1).$$

Lemma 5.2 *We have*

$$\widetilde{g}_0(A_2) = \mu\Gamma_0(\varepsilon, A_2) + O(\mu^2||f||^2).$$

PROOF. A summation over κ_2 in estimate (4.23) (where we chose $\delta = \pi/2$ and $k_1 = 0$) yields immediately the estimate. ■

Lemma 5.3 *We have*

$$\widetilde{g}_{\pm 1}(A_2) = \mu\Gamma_{\pm 1}(\varepsilon, A_2) + O\Big(\frac{\mu^2||f||^2}{\varepsilon^2}\exp(-\frac{\pi}{2\sqrt{\varepsilon}})\Big).$$

PROOF. By (4.23) (where we chose $\delta = \sqrt{\varepsilon}$ and $k_1 = \pm 1$), we can obtain as in the proof of lemma 5.1

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

■

Since $\Gamma(A)$ and $G_\mu(A)$ are real 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 three lemmas the following result.

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

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

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 [22] which require $\mu = \varepsilon^p$ with $p > 2 + a$.*

(ii) Theorem 5.1 certainly holds in any dimension, while the results of [15], which hold for more general systems, are proved for 2 rotators only.

(iii) Theorem 5.1 is not in contradiction with the counterexample given in [17].

(iv) In order to prove a splitting condition using theorem 5.1 it is necessary, according with [15] and [22], that $\exists m, l \in \mathbf{Z}^{n-1}$ such that $f_{0,l}, f_{1,m} \neq 0$. Otherwise, recalling (4.22), it 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 5.4 *Assume that there are $\overline{A}_2 \in \mathbf{R}^{n-1}$ and $d_0, c_0 > 0$ such that, for all small $\varepsilon > 0$,*

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

Then there is $c_1 > 0$ such that, for $\mu||f||\varepsilon^{-3/2}$ small enough, condition 2.1 is satisfied by $\widetilde{G}_\mu$, with $\alpha = c_1 e^{-\pi/(2\sqrt{\varepsilon})}/\sqrt{\varepsilon}$ and $\delta = c_0\mu/(2\sqrt{\varepsilon})e^{-\pi/(2\sqrt{\varepsilon})}$.

Proof. First we can derive from (4.22) and (4.23) in the same way as in the proof of lemmas 5.1 and 5.3 that

$$\begin{aligned} |\widetilde{g}_1(A_2)| + |\nabla \widetilde{g}_1(A_2)| &\leq \sum_{\kappa_2 \in \mathbf{Z}^{n-1}} (1+|\kappa_2|)|\widetilde{G}_{1,\kappa_2}| \\ &\leq C\frac{\mu||f||}{\sqrt{\varepsilon}} e^{-\pi/(2\sqrt{\varepsilon})}. \end{aligned} \tag{5.2}$$

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

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

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 function defined in B_{d_0} ((5.2) and the previous lower bound of $|\widetilde{g}_1(A_2)|$ provide a bound of $\nabla\phi(A_2)$).

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

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

Let

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

If $A_2 \in \partial B_{d_0}$ then

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

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

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

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

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

Using that $|\nabla \widetilde{G}_\mu| = O(\mu)$, we can easily derive that condition 2.1 (not with a ball B_ρ but the bounded open set U, according to remark 2.3) is satisfied with $\delta = (c_0/2)\mu e^{-\pi/2\sqrt{\varepsilon}}/\sqrt{\varepsilon}$, $\alpha = c_1 e^{-\pi/2\sqrt{\varepsilon}}/\sqrt{\varepsilon}$ for some positive constant c_1. ■

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

$$f(A_1, A_2) = \sum_{(k_1, \kappa_2) \in \mathbf{Z} \times \mathbf{Z}^{n-1}} f_{k_1, \kappa_2} \exp(i(k_1 A_1 + \kappa_2 \cdot A_2)) = \sum_{k_1 \in \mathbf{Z}} f_{k_1}(A_2) \exp(ik_1 A_1).$$

If f is analytical in some domain then also $f_{k_1}(\cdot)$ can be analitically extended is the same domain, as $f_{k_1}(s) = (1/2\pi)\int_0^{2\pi} f(\sigma, s)e^{-ik_1\sigma}\, d\sigma$.

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 $f_1(\overline{A}_2) \neq 0$
(ii) $a = 0$, $f_0(A_2)$ admits a strict local minimum at the point $\overline{A}_2$ and $f_1(\overline{A}_2 + i(\pi/2)\beta) \neq 0$.

Then, for all small ε such that $\omega_\varepsilon = (1/\sqrt{\varepsilon}, \beta\varepsilon^a)$ satisfies

$$\omega_\varepsilon \cdot \mathbf{k} \geq \frac{\gamma_\varepsilon}{|\mathbf{k}|^\tau}, \ \forall k \in \mathbf{Z}^n, k \neq 0$$

for all I_0, I_0' with $\omega_\varepsilon \cdot I_0 = \omega_\varepsilon \cdot I_0'$, there is a heteroclinic orbit 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 $O(|I_0 - I_0'|(\sqrt{\varepsilon}/\mu)e^{\pi/(2\sqrt{\varepsilon})}[(\gamma_\varepsilon)^{-1}(\sqrt{\varepsilon}e^{\pi/(2\sqrt{\varepsilon})})^\tau + |\ln \mu|] + |\ln(\eta)|)$.

Proof. It is enough to prove that, if (i) or (ii) is satisfied, then the condition given in lemma 5.4 holds. The statement is then a direct consequence of theorem 2.3.

We first assume that condition (i) is satisfied. In what follows, the notation $u = O(v)$ means that $|u| \leq C|v|$, 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

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

provided ε is small enough. It is then clear that condition (i) implies that assumption of lemma 5.4 holds.

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

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

We observe also that, if $a = 0$, then $\Gamma_0(\varepsilon, A_2)$ is independent of ε. It follows easily that condition (ii) implies that the assumption of lemma 5.4 holds true. ■

Remark 5.2 *In many examples, condition (i) or condition (ii) is satisfied. However we need that $f_0(A_2)$ and $f_1(A_2)$ do not vanish everywhere, see remark 5.1-(iv).*

6 Appendix

In the proof of the following lemmas we will closely follow the arguments developed in the papers [5]-[6] 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 (2.2) 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 $t \in [0, +\infty)$,

$$u(t) = (\mathcal{L}_0 h)(t) = \frac{1}{2} \int_0^{+\infty} \Big(e^{-|t-s|} - e^{-(t+s)} \Big) h(s) ds. \tag{6.1}$$

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

$$H(\mu, A, w) := w - \mathcal{L}_0\Big(-\Big(\sin(q_0 + w) - \sin q_0 - w\Big) + \mu \sin(q_0 + w) f(\omega t + A)\Big). \tag{6.2}$$

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 then 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

$$-\ddot{W} + \cos q_0 W = 0. \tag{6.3}$$

Multiplying by $\dot{q}_0$ in (6.3) 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 then $W = 0$. 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 unicity of ${w'}^{\mu}_{A} : (-\infty, 0] \to \mathbf{R}$ which satisfies analogous properties over the interval $(-\infty, 0]$. We can define $q^{\mu}_{A,0}$ by $q^{\mu}_{A,0}(t) = q_0(t) + w^{\mu}_{A}(t)$ if $t \geq 0$, $q^{\mu}_{A,0} = q_0(t) + {w'}^{\mu}_{A}(t)$ if $t < 0$. This is the unique function for which (i), (ii) (with $\theta = 0$) and (iii) 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 $q^{\mu}_{A,\theta}$ which satisfies (i), (ii), defined by $q^{\mu}_{A,\theta} = T_\theta q^{\mu}_{A+\omega\theta,0}$, i.e. $q^{\mu}_{A,\theta}(t) = q^{\mu}_{A+\omega\theta,0}(t-\theta)$; (iii) and (iv) clearly hold. The regularity of $q^{\mu}_{A,\theta}$ w.r.t. A, θ, μ is a consequence of the regularity of w^{μ}_{A} and ${w'}^{\mu}_{A}$ w.r.t. A and μ. (v) follows from

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

provided we can justify that $||\partial_A H(\mu, A, w^{\mu}_{A})||_1 = O(\mu)$, $\quad ||\omega \cdot \partial_A H(\mu, A, w^{\mu}_{A})||_1 = O(\mu)$. The second bound (uniform in ω) is not so obvious. We just point out that

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

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 (2.2) over $(0, \theta)$ with $w(0) = w(\theta) = 0$, where $q^*_{0,\theta}$ is the following smooth "approximate solution"

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

where

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

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:

$$\begin{cases} -\ddot{u} + u = h \\ u(0) = 0 \ , \ u(\theta) = 0. \end{cases} \tag{6.4}$$

An explicit computation shows that for $t \in [0, \theta]$ the solution u of (6.4) 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 $||\mathcal{L}_{0,\theta} W||_{1,\infty} \leq K ||W||_\infty$, 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

$$||\partial^2_{ww} H^\theta(\mu, A, w)[W, W]|| = O(||W||^2_\infty). \tag{6.5}$$

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

$$||H^\theta(\mu,A,0)||_{1,\infty}=O(\exp(-\theta/2)). \tag{6.6}$$

$q^*_{0,\theta}+w$ is a solution of (2.2) 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

$$\Big|\Big|\left(\partial_w H^\theta(\mu,A,0)\right)^{-1}\Big|\Big|\le\overline{C}. \tag{6.7}$$

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 [5]). 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 $W_n\in W^{1,\infty}([0,\theta_n]), ||W_n||_{1,\infty}=1$,

$$||\partial_w H^{\theta_n}(\mu_n,A_n,0)W_n||_{1,\infty}\to 0. \tag{6.8}$$

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 (6.8) 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 (6.8) 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 (6.8), 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 (6.5) and (6.6) in a standard way that, for all $||w||_{1,\infty},||w'||_{1,\infty}\le\rho$, $|\mu|<\overline{\mu}$, $\theta>\overline{L}$ there holds

$$\|F^\theta_{\mu,A}(w)\|_{1,\infty}=O(\exp(-\theta/2)+\rho^2)\quad;\quad\|F^\theta_{\mu,A}(w)-F^\theta_{\mu,A}(w')\|=O(\rho||w'-w||). \tag{6.9}$$

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 (6.9) 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 like in [5]. ■

Proof of lemma 4.2. 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

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

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,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 $\partial_A Q^\mu_{A,\theta}$, $\omega \cdot \partial_A Q^\mu_{A,\theta}$, there holds $|\nabla l_\mu(A,\theta)| = O(\mu)$. ■

Proof of lemma 4.3. The first step is to prove that

$$\max\Big(|q^\mu_{A,\theta+l_\mu(A,\theta)}(t) - Q^\mu_{A,\theta}(t)|, |\dot{q}^\mu_{A,\theta+l_\mu(A,\theta)}(t) - \dot{Q}^\mu_{A,\theta}(t)|\Big) \leq K_0 |\partial_\theta \widetilde{F}_\mu(A,\theta)| \exp(-\frac{|t-\theta|}{2}), \forall t \in \mathbf{R}. \tag{6.10}$$

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 $|\alpha^\mu_{A',\theta'}| = O(|\partial_\theta \widetilde{F}_\mu(A,\theta)|)$ we deduce estimate (6.10).

We can now estimate $\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta)$. For $q \in \mathbf{X}$ let

$$\mathcal{G}^\mu_{A'}(q) = \int_{\mathbf{R}} \mathcal{L}_{\mu,A'}(q, \dot{q}, t)\ dt = \int_{\mathbf{R}} \frac{1}{2}(\dot{q}(t))^2 + (1 - \cos(q(t))) + \mu(\cos(q(t)) - 1) f(A' + \omega t)\ dt.$$

By standard arguments, $\mathcal{G}^\mu_{A'} : \mathbf{X} \to \mathbf{R}$ is smooth. Moreover for $|\mu| < \mu_0$,

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

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

$$\mathcal{G}^\mu_{A'}(q^\mu_{A',0} + w) = \mathcal{G}^\mu_{A'}(q^\mu_{A',0}) + O(||w||_1^2)$$

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

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

We obtain by (6.10) that

$$|\widetilde{F}_\mu(A,\theta) - V_\mu(A,\theta)| \leq C_4 \Big(\partial_\theta \widetilde{F}_\mu(A,\theta)\Big)^2$$

for some positive constant C_4. ■

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, p. 233-252.

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

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

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

[5] 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.

[6] 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.

[7] M. Berti, P. Bolle: *Diffusion time and splitting of separatrices for nearly integrable isochronous Hamiltonian systems*, to appear in Rend. Mat. Acc. Naz. Lincei.

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

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

[10] U. Bessi, L. Chierchia, E. Valdinoci: *Upper Bounds on Arnold Diffusion Time via Mather theory*, preprint.

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

[12] J. Cresson: *Conjecture de Chirikov et Optimalité des exposants de stabilité du théorèm de Nekhoroshev*, preprint univ. Besancon.

[13] 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, 35-71, 1997.

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

[15] G. Gallavotti, G. Gentile, V. Mastropietro: *Separatrix splitting for systems with three times scale*, Commun. Math. Phys. 202, 197-236, 1999.

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

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

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

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

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

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

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

[23] D. Sauzin: *A new method for measuring the splitting of invariant manifolds*, to appear in Ann. scient. Ec. Norm. Sup.

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

[25] Z. Xia: *Arnold diffusion: a variational construction*, Documenta Matematica, extra vol. ICM, 198, 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