# Fast Arnold Diffusion in systems with three time scales

Massimiliano Berti and Philippe Bolle

**Abstract:** We consider the problem of Arnold Diffusion for nearly integrable partially isochronous Hamiltonian systems with three time scales. By means of a careful shadowing analysis, based on a variational technique, we prove that, along special directions, Arnold diffusion takes place with fast (polynomial) speed, even though the "splitting determinant" is exponentially small. [1]

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

## 1 Introduction

In a previous paper [6] (see also [7]) we introduced, in the context of nearly integrable Hamiltonian systems, a functional analysis approach to the "splitting of separatrices" and to the "shadowing problem". We applied our method to the problem of Arnold Diffusion, i.e. topological instability of action variables, for nearly integrable partially isochronous systems. The aim of this paper is to improve the shadowing theorem of [6] and to apply this new theorem to the system with three time scales (1.1) below , in order to prove that along special directions Arnold diffusion takes place with "very fast speed", namely a speed polynomial in $\varepsilon$. To that effect, we use the results on the splitting provided in [6].

Hamiltonian systems with three time scales have been introduced in [11] as a description of the D'Alembert problem in Celestial Mechanics. Later on systems with three time scales have been reconsidered for example in [16], [17], [22], [10], [6], [19].

In this paper we focus on isochronous systems with three time scales as

$$\mathcal{H}_\mu = \frac{1}{\sqrt{\varepsilon}} I_1 + \varepsilon^a \beta \cdot I_2 + \frac{p^2}{2} + (\cos q - 1)(1 + \mu f(\varphi)), \tag{1.1}$$

where $(\varphi_1, \varphi_2, q) \in \mathbf{T}^1 \times \mathbf{T}^{n-1} \times \mathbf{T}^1$ are the angle variables, $(I_1, I_2, p) \in \mathbf{R}^1 \times \mathbf{R}^{n-1} \times \mathbf{R}^1$ are the action variables, $\beta = (\beta_2, \ldots, \beta_n) \in \mathbf{R}^{n-1}$, $n \geq 3$, $a > 0$ and $\varepsilon > 0$, $\mu \geq 0$ are small real parameters. We will assume that $\mu = O(\min\{\varepsilon^{3/2}, \varepsilon^{2a+1}\})$. Hamiltonian $\mathcal{H}_\mu$ describes a system of $n$ isochronous harmonic oscillators with a Diophantine frequency vector $\omega_\varepsilon = (1/\sqrt{\varepsilon}, \varepsilon^a \beta)$ , with one fast frequency $\omega_{\varepsilon,1} = 1/\sqrt{\varepsilon}$ and $(n-1)$ slow frequencies $\omega_{\varepsilon,2} = \varepsilon^a \beta$, weakly coupled with a pendulum.

When $\mu = 0$ the energy $\omega_{\varepsilon,i} I_i$ of each oscillator is a constant of the motion. The problem of *Arnold diffusion* in this context is whether, for $\mu \neq 0$, there exist motions whose net effect is to transfer $O(1)$-energy from one oscillator to others in a certain time $T_d$ called the diffusion time.

The existence of Arnold diffusion is usually proved following the mechanism proposed in [3]. 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 $W^s(\mathcal{T}_{I_0}) = W^u(\mathcal{T}_{I_0}) = \{(\varphi, I, q, p) \in \mathbf{T}^n \times \mathbf{R}^n \times \mathbf{T}^1 \times \mathbf{R}^1 \mid I = I_0,\ p^2/2 + (\cos q - 1) = 0\}$ called "whiskers" by Arnold. For $\mu$ small enough the perturbed stable and unstable manifolds $W^s_\mu(\mathcal{T}^\mu_{I_0})$ and $W^u_\mu(\mathcal{T}^\mu_{I_0})$ may split and intersect transversally, giving rise to a chain of tori connected by heteroclinic orbits. By a shadowing type argument one can then prove 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*. In order to prove the existence

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

of diffusion orbits following the previous mechanism one encounters two different problems: 1) *Splitting of the whiskers*; 2) *Shadowing problem.*

The "splitting of the whiskers" for Hamiltonian $\mathcal{H}_\mu$, when $\mu = O(\varepsilon^p)$, $p > 0$ and $\varepsilon \to 0$, has been studied in [16], [19], [22] and [6]. In [16]-[19] and [22] the size of the splitting is measured by the "determinant of the splitting matrix" which turns out to be exponentially small, precisely $O(\exp(-(\pi/2)\varepsilon^{-1/2}))$. We underline that papers [16]-[19] deal also with non-isochronous systems and more general perturbation terms (but two rotators only).

In [6], the splitting of stable and unstable manifolds is related to the variations of the "homoclinic function" $G_\mu : \mathbf{T}^n \to \mathbf{R}$ (defined in (2.6)), which is the difference between the generating functions of stable and unstable manifolds at section $\{q = \pi\}$. $\nabla G_\mu(A)$ provides a measure of the distance between stable and unstable manifolds, so that a critical point $\overline{A}$ of $G_\mu$ gives rise to a homoclinic intersection. Usually det $D^2 G_\mu(\overline{A})$ is called the "splitting determinant". The use of the "homoclinic function" $G_\mu$ for measuring the splitting has two advantages. Firstly, it is very well suited to deal with the shadowing problem by means of variational techniques because $G_\mu$ is nothing but the difference of the values of the Lagrangian action functional associated to the quasi-periodically forced pendulum (2.3) at two true solutions, lying respectively on the stable and unstable manifolds $W_\mu^{s,u}(\mathcal{T}_{I_0})$, see (2.5). Secondly it may shed light on a "non uniform" splitting which would not be given by the splitting determinant, when the variations of $G_\mu$ in different directions are of different orders. With this regard we quote paper [20] where, for a general nearly integrable Hamiltonian system, detailed estimates for the "eigenvalues and the eigenspaces of the splitting", rather than for the determinant, are given.

For the system with three time scales associated to Hamiltonian $\mathcal{H}_\mu$, "non uniform" splitting is suggested by the behaviour of the first order expansion of $G_\mu$ in $\mu$, called the Poincaré-Melnikov approximation. In fact the first order term, which is given by the Poincaré-Melnikov primitive defined in (2.9), has exponentially small oscillations in the fast angle $A_1$, and polynomially small ones in the slow angles $A_2$. Naively this hints the splitting to be exponentially small in the direction $I_1$ and just polynomially small in the directions $I_2$.

However, in general, for $\mu = O(\varepsilon^p)$ and $\varepsilon \to 0$ the homoclinic function $G_\mu$ is not well approximated by the Poincaré-Melnikov primitive. In [16]-[19] the asymptotic validity of Melnikov's integrals for computing the exponentially small "splitting determinant" is proved to hold only after exhibiting many cancellations.

In [6] the naive Poincaré-Melnikov approximation for Hamiltonian $\mathcal{H}_\mu$ has been rigourously justified for $\mu\varepsilon^{-3/2}$ sufficiently small, in a different way. We define another "splitting function" $\widetilde{G}_\mu$, see (2.8), whose critical points as well give rise to homoclinic intersections. $\widetilde{G}_\mu$ is well approximated, for $\mu = O(\varepsilon^p)$ and $\varepsilon \to 0$, by the Poincaré-Melnikov primitive and has exponentially small oscillations in $A_1$, see theorem 2.2. The crucial observation is that $G_\mu$ and $\widetilde{G}_\mu$ are the same function up to a diffeomorphism $\psi_\mu$ of the torus close to identity, namely $\widetilde{G}_\mu = G_\mu \circ \psi_\mu$, see theorem 2.1.

After the works [8], [9], [21], [12], [10], [6], [13], [14] and references therein, it is a well established fact that the diffusion time is estimated by a polynomial inverse power of the splitting. For instance, using the estimate on the size of the splitting of [16] and [19] an exponentially long diffusion time has been obtained in [10], namely $T_d = O(\exp(C/\varepsilon^b))$ for some $b > 0$ (see also theorem 5.2 of [6]).

However the properties of $G_\mu$ (oscillations of different amplitude orders according to the direction) suggest that Arnold diffusion can take place with different speed along different directions; since, for larger splitting one would expect a faster speed of diffusion, one could guess the existence of diffusion orbits that drift along the "fast" directions $I_2 \in \mathbf{R}^{n-1}$, where the splitting is just polynomially small w.r.t. $1/\varepsilon$, in a polynomially long diffusion time $T_d = O(1/\varepsilon^q)$; see also the discussion in chapter 2 of [20]. The aim of this paper is to prove that this is indeed the case and to provide explicit and careful estimates on the diffusion time. In order to prove this phenomenon (see *theorem 4.1* for the general case and *theorem 4.2* for an application) we refine the shadowing theorem 2.3 of [6] for dealing with the present "non-uniform" splitting. Note that, because of the preservation of the energy along the orbits, Arnold diffusion can take place in the direction $I_2$ for $n \geq 3$ only.

In order to justify heuristically our result we recall how the diffusion time $T_d$ is estimated in [6], once it is verified that stable and unstable manifolds split. $T_d$ is, roughly, 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" $T_e$ of the torus $\mathbf{T}^n$ run by the linear flow $\omega_\varepsilon t$, and the time needed to "shadow" homoclinic orbits for the corresponding quasi-periodically forced pendulum equation 2.3.

The reasons for which we are able to move in polynomial time w.r.t $1/\varepsilon$ along the fast $I_2$ directions are the following three ones. $(i)$ As in [6], since the homoclinic orbit decays exponentially fast to 0, the time needed to "shadow" homoclinic orbits for the quasi-periodically forced pendulum (2.3) is only polynomial. $(ii)$ Since the splitting is polynomially small in the directions $I_2$, we can choose just a polynomially large number of tori forming the transition chain $k = O(1/\varepsilon^p)$ to get a $O(1)$-drift of $I_2$. $(iii)$ Finally, the most difficult task is getting a polynomial estimate for the "ergodization time" $T_e$ -defined as the time needed for the flow $\{\omega_\varepsilon t\}$ to make an $\alpha$-net of the torus- with $\alpha$ appropriately small. By a result of [4] this time satisfies $T_e = O(1/\alpha^\tau)$. Let us explain how this estimate enters into play. In order to apply our "gluing" variational technique, the projection of our shadowing orbit on the torus $\mathbf{T}^n$, namely $\{\omega_\varepsilon t + A_0\}$, must approach, at each transition, sufficiently close to the homoclinic point $\overline{A}$ to be capable to "see" the homoclinic critical point $\overline{A}$ of $G_\mu$. The crucial improvement of the shadowing theorem 4.1 allows the shadowing orbit to approach $\overline{A}$ only up to a polynomially small distance $\alpha = O(\varepsilon^p)$, $p > 0$, (and not exponentially small as it would be required when applying the shadowing theorem of [6]). By the forementioned estimate on the ergodization time $T_e = O(1/\alpha^\tau)$ it results that the minimum time after which the homoclinic trajectory can "jump" to another torus is only polynomially long w.r.t $1/\varepsilon$. Actually this allows to improve as well the exponential estimate on the diffusion time required to move also in the $I_1$ direction, see remark 4.3.

Theorems 4.1 and 4.2 are the first steps to prove the existence of this phenomenon also for more general systems (with non isochronous terms and more general perturbations). We refer to [20], where the splitting problem is studied in a quite general framework.

The paper is organized as follows: in section 2 we recall some preliminary results taken from [6]. In section 3 we introduce the general "splitting condition" which will be used in section 4 to prove the shadowing theorems.

Through the paper $C_i$ and $\delta_i$ will denote positive constants which are independent of $\varepsilon$ and $\mu$.

## 2 Preliminaries

In this section we recall the results of [6] that will be used in the sequel. We refer to [6] for complete details and for the description of the general functional analysis approach based on a Lyapunov Schmidt type reduction. With respect to the notations of [6] we remark that we have changed the sign of the perturbation $f$ in Hamiltonian $\mathcal{H}_\mu$.

The equations of motion derived by Hamiltonian $\mathcal{H}_\mu$ are

$$\dot\varphi = \omega_\varepsilon, \qquad \dot I = -\mu(\cos q - 1)\, \partial_\varphi f(\varphi), \qquad \dot q = p, \qquad \dot p = \sin q\, (1 + \mu\, f(\varphi)). \tag{2.2}$$

The angles $\varphi$ evolve as $\varphi(t) = \omega_\varepsilon t + A$; therefore equations (2.2) can be reduced to the quasi-periodically forced pendulum equation

$$-\ddot q + \sin q\, (1 + \mu f(\omega_\varepsilon t + A)) = 0, \tag{2.3}$$

corresponding to the Lagrangian

$$\mathcal{L}_{\mu,A}(q, \dot q, t) = \frac{\dot q^2}{2} + (1 - \cos q)(1 + \mu f(\omega_\varepsilon t + A)). \tag{2.4}$$

For each solution $q(t)$ of (2.3) one recovers the dynamics of the actions $I(t)$ by quadratures in (2.2).

For $\mu = 0$ equation (2.3) possesses the one parameter family of homoclinic solutions to 0, mod $2\pi$, $q_\theta(t) = 4\arctan(\exp(t - \theta))$, $\theta \in \mathbf{R}$. Using the Implicit Function Theorem one can prove (lemma 2.1 of [6]) that there exist, near the unperturbed homoclinic solutions $q_\theta(t)$, for $0 < \mu < \mu_0$ small enough independently of $\omega_\varepsilon$, *"pseudo-homoclinic solutions"* $q^\mu_{A,\theta}(t)$ of equation (2.3). These are true solutions of

(2.3) in each interval $(-\infty, \theta)$ 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 $2\pi$. We can then define the function $F_\mu : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$ as the action functional of Lagrangian (2.4) evaluated on the "1-bump pseudo-homoclinic solutions" $q^{\mu}_{A,\theta}(t)$, namely

$$F_\mu(A,\theta) = \int_{-\infty}^{\theta} \mathcal{L}_{\mu,A}(q^{\mu}_{A,\theta}(t), \dot{q}^{\mu}_{A,\theta}(t), t)\ dt + \int_{\theta}^{+\infty} \mathcal{L}_{\mu,A}(q^{\mu}_{A,\theta}(t), \dot{q}^{\mu}_{A,\theta}(t), t)\ dt, \tag{2.5}$$

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

$$G_\mu(A) = F_\mu(A, 0). \tag{2.6}$$

There holds

$$F_\mu(A,\theta) = G_\mu(A + \omega_\varepsilon \theta), \forall \theta \in \mathbf{R}. \tag{2.7}$$

**Remark 2.1** *The homoclinic function $G_\mu$ is the difference between the generating functions $\mathcal{S}^{\pm}_{\mu,I_0}(A,q)$ of the stable and the unstable manifolds $W^{s,u}_{\mu}(\mathcal{T}_{I_0})$ (which in this case are exact Lagrangian manifolds) at the fixed section $\{q = \pi\}$, namely $G_\mu(A) = \mathcal{S}^{-}_{\mu,I_0}(A,\pi) - \mathcal{S}^{+}_{\mu,I_0}(A,\pi)$. A critical point of $G_\mu$ gives rise to a homoclinic orbit to torus $\mathcal{T}_{I_0}$, see lemma 2.3 of [6].*

In order to justify the dominance of the Poincaré-Melnikov function when $\mu = O(\varepsilon^p)$ one would need to extend analytically the function $F_\mu(A,\theta)$ for complex values of the variables. Since the condition $q^{\mu}_{A,\theta}(Re\ \theta) = \pi$, appearing naturally when trying to extend the definition of $q^{\mu}_{A,\theta}$ to $\theta \in \mathbf{C}$, breaks analyticity, the function $F_\mu(A,\theta)$ can not be easily analytically extended in a sufficiently wide complex strip. To overcome this problem, in [6] the Lagrangian action functional is evaluated on different "1-bump pseudo-homoclinic solutions" $Q^{\mu}_{A,\theta}$. Define $\psi_0 : \mathbf{R} \to \mathbf{R}$ by $\psi_0(t) = \cosh^2(t)/(1 + \cosh t)^3$ and set $\psi_\theta(t) = \psi_0(t - \theta)$. Two important properties of the function $\psi_0(t)$ are that $\int_{\mathbf{R}} \psi_0(t)\dot{q}_0(t)\ dt \neq 0$ and that it can be extended to a holomorphic function on $\mathbf{R} + i(-\pi,\pi)$ (while the homoclinic solution $q_0(t)$ can be extended to a holomorphic function only up to $\mathbf{R} + i(-\pi/2, \pi/2)$). By the Contraction Mapping Theorem there exist (lemma 4.1 of [6]) near $q_\theta$, for $\mu$ small enough, *pseudo-homoclinic solutions* $Q^{\mu}_{A,\theta}(t)$ and a constant $\alpha^{\mu}_{A,\theta}$ defined by

$$-\ddot{Q}^{\mu}_{A,\theta} + \sin Q^{\mu}_{A,\theta}(1 + \mu\ f(\omega_\varepsilon t + A)) = \alpha^{\mu}_{A,\theta}\psi_\theta(t) \quad \text{and} \quad \int_{\mathbf{R}} \Big(Q^{\mu}_{A,\theta}(t) - q_\theta(t)\Big)\psi_\theta(t)\ dt = 0.$$

We define the function $\widetilde{F}_\mu : \mathbf{T}^n \times \mathbf{R} \to \mathbf{R}$ as the action functional of Lagrangian (2.4) evaluated on the "1-bump pseudo-homoclinic solutions" $Q^{\mu}_{A,\theta}(t)$, namely

$$\widetilde{F}_\mu(A,\theta) = \int_{\mathbf{R}} \mathcal{L}_{\mu,A}(Q^{\mu}_{A,\theta}(t), \dot{Q}^{\mu}_{A,\theta}(t), t)\ dt \tag{2.8}$$

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

**Remark 2.2** *Also critical points of $\widetilde{G}_\mu$ give rise to homoclinic solutions to torus $\mathcal{T}_{I_0}$, see lemma 4.2 of [6]. By theorem 2.1 below, from a geometrical point of view the introduction of the "homoclinic function" $\widetilde{G}_\mu$ may be interpreted simply as measuring the splitting with a non constant Poincaré section, see the introduction of [6].*

The crucial point is now to observe that the homoclinic functions $G_\mu$ and $\widetilde{G}_\mu$ are the same up to a change of variables close to the identity, as stated by the following theorem (see theorem 4.1 of [6])

**Theorem 2.1** *For $\mu$ small enough (independently of $\omega_\varepsilon$) there exists a Lipschitz homeomorphism (a real analytic diffeomorphism if $f$ is analytic) $\psi_\mu : \mathbf{T}^n \to \mathbf{T}^n$ of the form $\psi_\mu(A) = A + k_\mu(A)\omega_\varepsilon$ with $k_\mu : \mathbf{T}^n \to \mathbf{R}$ satisfying $k_\mu(A) = O(\mu)$, $|k_\mu(A) - k_\mu(A')| = O(\mu)|A - A'|$ such that $\widetilde{G}_\mu = G_\mu \circ \psi_\mu$.*

Let $\Gamma(\varepsilon, A)$ denote the Poincaré-Melnikov primitive

$$\Gamma(\varepsilon, A) = \int_{\mathbf{R}} (1 - \cos q_0(t)) f(\omega_\varepsilon t + A) \, dt. \tag{2.9}$$

Develop in Fourier series w.r.t. the first variable the homoclinic function $\widetilde{G}_\mu(A) = \sum_{k_1 \in \mathbf{Z}} \widetilde{g}_{k_1}(A_2) e^{ik_1 \cdot A_1}$ and the Poincaré-Melnikov primitive $\Gamma(\varepsilon, A) = \sum_{k_1 \in \mathbf{Z}} \Gamma_{k_1}(\varepsilon, A_2) e^{ik_1 \cdot A_1}$. Assume that the perturbation $f$ is analytic w.r.t $(\varphi_2, \ldots, \varphi_n)$. More precisely assume that there exist $r_i > 0$ for $i = 2, \ldots, n$, such that $f$ has a $C^\infty$ extension in $D := \mathbf{R} \times (\mathbf{R} + i[-r_2, r_2]) \times \ldots \times (\mathbf{R} + i[-r_n, r_n])$, holomorphic w.r.t. $(\varphi_2, \ldots, \varphi_n)$. Denote the supremum of $|f|$ over $D$ as $||f|| := \sup_{\varphi \in D} |f(\varphi)|$. The following theorem about the splitting of stable and unstable manifolds in systems with three time scales, holds (see theorem 5.1 of [6])

**Theorem 2.2** *For $\mu ||f|| \varepsilon^{-3/2}$ small there holds, for all $A_2 \in \mathbf{T}^{n-1}$*

$$\begin{aligned} \widetilde{G}_\mu(A_1, A_2) &= \widetilde{g}_0(A_2) + 2\mathrm{Re}\Big[\widetilde{g}_1(A_2) e^{iA_1}\Big] + \widetilde{R}(A_1, A_2) \\ &= Const + \Big(\mu \Gamma_0(\varepsilon, A_2) + R_0(\varepsilon, \mu, A_2)\Big) + 2 \ \mathrm{Re} \ \Big[\mu \Gamma_1(\varepsilon, A_2) + R_1(\varepsilon, \mu, A_2)\Big] e^{iA_1} \\ &+ \widetilde{R}(A_1, A_2), \end{aligned}$$

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

$$\widetilde{R}(A_1, A_2) = O\Big(\mu \varepsilon^{-1/2} ||f|| \exp\Big(-\frac{\pi}{\sqrt{\varepsilon}}\Big)\Big).$$

In order to prove our shadowing theorem we need also to recall the definition of the **k-bump pseudo-homoclinic solutions** $q^L_{A,\theta}(t)$ for the quasi-periodically forced pendulum (2.3). Such pseudo solutions turn $k$ times along the separatrices and are asymptotic to the equilibrium 0, mod $2\pi$, for $t \to \pm\infty$. More precisely in lemma 2.4 of [6] it is proved that for all $k \in \mathbf{N}$, for all $\theta_1 < \ldots < \theta_k$ with $\min_i(\theta_{i+1} - \theta_i) > L$, with $L$ sufficiently large, independently of $\omega_\varepsilon$ and $\mu$, there exists a unique pseudo-homoclinic solution $q^L_{A,\theta}(t) : \mathbf{R} \to \mathbf{R}$ which is a true solution of (2.3) in each interval $(-\infty, \theta_1)$, $(\theta_i, \theta_{i+1})$ $(i = 1, \ldots, k-1)$, $(\theta_k, +\infty)$ and $q^L_{A,\theta}(\theta_i) = \pi(2i - 1)$, $q^L_{A,\theta}(t) = q^\mu_{A,\theta_1}(t)$ in $(-\infty, \theta_1)$ and $q^L_{A,\theta}(t) = 2\pi k + q^\mu_{A,\theta_k}(t)$ in $(\theta_k, +\infty)$. Such pseudo-homoclinic orbits are found via the Contraction Mapping Theorem, as small perturbations of a chain of "1-bump homoclinic solutions" $q^\mu_{A,\theta_i}$.

Then we consider the Lagrangian action functional evaluated on these pseudo-homoclinic orbits $q^L_{A,\theta}$ depending on $n + k$ variables

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

Setting $e_k = (1, \ldots, 1) \in \mathbf{R}^k$, the following invariance property, inherited from the autonomy of $\mathcal{H}_\mu$, holds

$$F^k_\mu(A, \theta + \eta e_k) = F^k_\mu(A + \eta \omega_\varepsilon, \theta), \qquad \forall \theta \in \mathbf{R}^k, \eta \in \mathbf{R}. \tag{2.10}$$

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

$$\mathcal{F}^k_\mu(A, \theta) := F^k_\mu(A, \theta) - (I'_0 - I_0) \cdot A. \tag{2.11}$$

**Lemma 2.1** *$\forall I_0, I'_0 \in \mathbf{R}^n$, if $(A, \theta)$ is a critical point of the "$k$-bump heteroclinic function" $\mathcal{F}^k_\mu(A, \theta)$, then $(I_\mu(t), \omega_\varepsilon t + A, \ q^L_{A,\theta}(t), \dot{q}^L_{A,\theta}(t))$ where $I_\mu(t) = I_0 - \mu \int_{-\infty}^t (\cos q^L_{A,\theta}(s) - 1) \partial_\varphi f(\omega_\varepsilon s + A) ds$ is an heteroclinic solution connecting $\mathcal{T}_{I_0}$ to $\mathcal{T}_{I'_0}$.*

By lemma 2.1, in order to get heteroclinic solutions connecting $\mathcal{T}_{I_0}$ to $\mathcal{T}_{I'_0}$, we need to find critical points of $\mathcal{F}^k_\mu(A, \theta)$. When $\min_i(\theta_{i+1} - \theta_i) \to +\infty$ the "$k$-bump homoclinic function" $F^k_\mu(A, \theta)$ turns out to be well approximated simply by the sum of the functions $F_\mu(A, \theta_i)$ according to the following lemma. We set $\theta_0 = -\infty$ and $\theta_{k+1} = +\infty$.

**Lemma 2.2** *There exist positive constants* $C_1, L_1 > 0$ *and functions* $R_i(\mu, A, \theta_{i-1}, \theta_i, \theta_{i+1})$ *such that* $\forall \varepsilon > 0$, $\forall 0 < \mu < \mu_0$, $\forall L > L_1$, $\forall \theta_1 < \ldots < \theta_k$ *with* $\min_i(\theta_{i+1} - \theta_i) > L$

$$F_\mu^k(A, \theta_1, \cdots, \theta_k) = \sum_{i=1}^k F_\mu(A, \theta_i) + \sum_{i=1}^k R_i(\mu, A, \theta_{i-1}, \theta_i, \theta_{i+1}), \tag{2.12}$$

*with*

$$|R_i(\mu, A, \theta_{i-1}, \theta_i, \theta_{i+1})| \leq C_1 \exp(-C_1 L). \tag{2.13}$$

# 3 The splitting condition

We now give a general "splitting condition" on the homoclinic function $G_\mu$ well suited to describe the non-uniform splitting of stable and unstable manifolds which takes place in systems with three time scales. Roughly, the "splitting condition" 3.1 below states that $G_\mu$ possesses a maximum and provides explicit estimates of the non-uniform splitting. It will be used, in the next section, to prove the shadowing theorem 4.1. As a paradigmatic example, we will verify, in lemma 3.2, that, when the perturbation $f(\varphi) = \sum_{j=1}^n \cos \varphi_j$, the "splitting condition" is satisfied, see also remark 3.1.

**Condition 3.1 "Splitting Condition".** *There exist* $\overline{A} \in \mathbf{R}^n$ *and a basis* $\{\Omega_1, \ldots, \Omega_n\}$ *of* $\mathbf{R}^n$, $n \geq 3$, *such that* $\omega_\varepsilon \in \mathbf{R}_+ \Omega_1$, $1/2 \leq |\Omega_i| \leq 2$, $\det\{\Omega_1, \ldots, \Omega_n\} \geq 1/2$, $\{\Omega_3, \ldots, \Omega_n\}$ *is an orthonormal basis of* $\{\Omega_1, \Omega_2\}^\perp$, *and which enjoy the following properties : let us define* $H_\mu(a_1, \ldots, a_n)$ *as the homoclinic function* $G_\mu(A)$ *in the new basis, namely*

$$H_\mu(a_1, \ldots, a_n) = G_\mu(\overline{A} + a_1 \Omega_1 + \ldots + a_n \Omega_n). \tag{3.14}$$

*Then there exist positive constants* $\rho, \sigma, \delta_1, \delta_2, \delta_3 > 0$, *with* $3\sigma < \rho$, $\delta_2 < \delta_3$, *and two continuous functions* $l_1, l_2 : [-\rho, \rho] \times \overline{B}_\rho^{n-2} \to \mathbf{R}$ *with* $l_1(x) < l_2(x)$ *for all* $x \in [-\rho, \rho] \times \overline{B}_\rho^{n-2}$, *such that:*

- *(i) for* $x = (a_2, \ldots, a_n) \in [-\rho, \rho] \times \overline{B}_\rho^{n-2}$

$$\mathcal{J}(x) := \sup_{a_1 \in [l_1(x), l_2(x)]} H_\mu(a_1, x) \geq \max\Big\{H_\mu(l_1(x), x), H_\mu(l_2(x), x)\Big\} + \delta_1;$$

- *(ii) for all* $y = (a_3, \ldots, a_n) \in \overline{B}_\rho^{n-2}$,

$$\forall a_2 \in [-\sigma, \sigma], \quad \mathcal{J}(a_2, y) \geq \mathcal{J}(0, y) - \frac{\delta_2}{2},$$

$$\forall a_2 \in [-\rho, -\rho + 2\sigma] \cup [\rho - 2\sigma, \rho], \quad \mathcal{J}(a_2, y) \leq \mathcal{J}(0, y) - \delta_2;$$

- *(iii)*

$$\forall a_2 \in [-\sigma, \sigma], \quad \forall y \in \overline{B}_\sigma^{n-2}, \quad \mathcal{J}(a_2, y) \geq \mathcal{J}(0, 0) - \frac{\delta_3}{2},$$

$$\forall a_2 \in [-\rho, \rho], \quad \forall y \in \overline{B}_\rho^{n-2} \backslash B_{\rho - 2\sigma}^{n-2}, \quad \mathcal{J}(a_2, y) \leq \mathcal{J}(0, 0) - \delta_3.$$

The next lemma states that the former "splitting condition" is satisfied by the homoclinic function $G_\mu$ if (and only if) it holds for the homoclinic function $\widetilde{G}_\mu$.

**Lemma 3.1** *Assume that* $\widetilde{G}_\mu$ *satisfies the splitting condition 3.1 with maps* $\widetilde{l}_1, \widetilde{l}_2$ *and parameters* $\rho$, $\sigma$, $\delta_1$, $\delta_2$, $\delta_3$. *Then* $G_\mu$ *satisfies the splitting condition 3.1 as well, for some maps* $l_{1,2} = \widetilde{l}_{1,2} + O(\mu/\sqrt{\varepsilon})$ *and with the same parameters. The converse is also true.*

PROOF. By theorem 2.1, $\widetilde{G}_\mu = G_\mu \circ \psi_\mu$, where $\psi_\mu(A) = A + k_\mu(A)\omega_\varepsilon$ and $\psi_\mu$ is a homeomorphism. Set $\widetilde{H}_\mu(a_1,\dots,a_n) = \widetilde{G}_\mu(\overline{A} + a_1\Omega_1 + \dots + a_n\Omega_n)$. We have

$$\widetilde{H}_\mu(a_1, a_2, \dots, a_n) = H_\mu\Big(a_1 + \overline{k}_\mu(a_1,\dots,a_n)\frac{|\omega_\varepsilon|}{|\Omega_1|}, a_2, \dots, a_n\Big),$$

where $\overline{k}_\mu(a_1,\dots,a_n) := k_\mu(\overline{A} + a_1\Omega_1 + \dots + a_n\Omega_n)$.

Assume that $\widetilde{G}_\mu$ satisfies condition 3.1 with maps $\widetilde{l}_1, \widetilde{l}_2$. For all $x = (a_2,\dots,a_n) \in [-\rho,\rho] \times \overline{B}_\rho^{n-2}$, the map $a_1 \mapsto a_1 + \overline{k}_\mu(a_1,x)|\omega_\varepsilon|/|\Omega_1|$ is a homeomorphism from the interval $(\widetilde{l}_1(x), \widetilde{l}_2(x))$ to the interval $(l_1(x), l_2(x))$, where $l_j(x) := \widetilde{l}_j(x) + \overline{k}_\mu(\widetilde{l}_j(x),x)|\omega_\varepsilon|/|\Omega_1|$ $(j = 1,2)$. There results that, for all $x = (a_2,\dots,a_n) \in [-\rho,\rho] \times \overline{B}_\rho^{n-2}$

$$\widetilde{\mathcal{J}}(x) := \sup_{a_1 \in [\widetilde{l}_1(x), \widetilde{l}_2(x)]} \widetilde{H}_\mu(a_1,x) = \sup_{a_1 \in [l_1(x), l_2(x)]} H_\mu(a_1,x) = \mathcal{J}(x).$$

Therefore $G_\mu$ satisfies the splitting condition 3.1, with maps $\widetilde{l}_j$ replaced by $l_j$, and the same positive parameters. Since $\overline{k}_\mu = O(\mu)$ and $|\omega_\varepsilon| = O(1/\sqrt{\varepsilon})$ we have $|l_j(x) - \widetilde{l}_j(x)| = O(\mu/\sqrt{\varepsilon})$. ∎

We now give a paradigmatic example where the former "splitting condition" is satisfied. Assume that the perturbation $f$ is given by $f(\varphi_1,\dots,\varphi_n) = \sum_{j=1}^n \cos\varphi_j$. In the next lemma we show that the corresponding homoclinic function $\widetilde{G}_\mu$ satisfies the "splitting condition" 3.1 and hence, by lemma 3.1, $G_\mu$ as well satisfies the "splitting condition" 3.1.

**Lemma 3.2** *Assume that $f(\varphi) = \sum_{j=1}^n \cos\varphi_j$. There exist a basis $\{\Omega_1,\dots,\Omega_n\}$ and a positive constant $\delta_0$ such that, if $\varepsilon$ is small, $0 < \mu\varepsilon^{-3/2} < \delta_0$ and $0 < \mu\varepsilon^{-2a-1} < \delta_0$, then $\widetilde{G}_\mu$ satisfies the "splitting condition" 3.1, with $\overline{A} = 0$, $\rho = \pi\varepsilon^{a+1/2}$, $\sigma = \rho/6$, $\delta_1 = \delta_3 = \mu\rho^2/2$, $\delta_2 = 3\pi\mu\varepsilon^{-1/2}\exp(-\pi/(2\sqrt{\varepsilon}))$, $\widetilde{l}_1(x) = -2\pi, \widetilde{l}_2(x) = 2\pi$.*

PROOF. In order to simplify the notations we give the proof for $n = 3$ and we assume that $|\beta| = 1$. We will prove that $\widetilde{G}_\mu$ satisfies the "splitting condition" 3.1 with $\overline{A} = 0$ and w.r.t the basis

$$\Omega_1 = (1, \varepsilon^{a+1/2}\beta), \qquad \Omega_2 = (0,\beta), \qquad \Omega_3 = (0,\beta'),$$

where $|\beta'| = 1$ and $\beta\cdot\beta' = 0$. We set $\rho = \pi\varepsilon^{a+1/2}$ and we assume that $0 < \mu \le \delta\rho^2$, $0 < \mu \le \delta\varepsilon^{3/2}$, where $\delta$ is a small constant (independent of $\varepsilon$) to be specified later. Let $\overline{\delta} > 0$ be such that theorem 2.2 holds for $0 < \mu \le \overline{\delta}\varepsilon^{3/2}$. We shall always choose $0 < \delta \le \overline{\delta}$.

¿From now on, notation $K_i$ will be used for positive universal constants, whereas notation $c_i(\delta)$ will be used for positive constants depending only on $\delta$. Notation $u = O(v)$ will mean that there exists a universal constant $K$ such that $|u| \le K|v|$.

Our first aim is to prove expression (3.27) below. It easily results that, if $f(\varphi) = \sum_{j=1}^3 \cos\varphi_j$,

$$\Gamma_0(\varepsilon, A_2) = \sum_{j=2}^{3} \frac{2\pi\beta_j\varepsilon^a}{\sinh(\beta_j\varepsilon^a\frac{\pi}{2})}\cos A_j \qquad \text{and} \qquad \Gamma_1(\varepsilon, A_2) = \frac{\pi}{\sqrt{\varepsilon}\sinh(\frac{\pi}{2\sqrt{\varepsilon}})}. \tag{3.15}$$

By thereom 2.2 we have

$$\widetilde{G}_\mu(A_1,A_2,A_3) = \widetilde{g}_0(A_2,A_3) + 2\mathrm{Re}\,\Big[\widetilde{g}_1(A_2,A_3)e^{iA_1}\Big] + O\Big(\mu\varepsilon^{-1/2}e^{-\pi/\sqrt{\varepsilon}}\Big) \tag{3.16}$$

and, by (3.15), up to a constant that we shall omit,

$$\widetilde{g}_0(A_2,A_3) = \frac{\mu 2\pi\beta_2\varepsilon^a}{\sinh(\beta_2\varepsilon^a\frac{\pi}{2})}\cos A_2 + \frac{\mu 2\pi\beta_3\varepsilon^a}{\sinh(\beta_3\varepsilon^a\frac{\pi}{2})}\cos A_3 + O(\mu^2), \tag{3.17}$$

$$\widetilde{g}_1(A_2,A_3)=\frac{\mu\pi}{\sqrt{\varepsilon}\sinh(\frac{\pi}{2\sqrt{\varepsilon}})}+O\Big(\frac{\mu^2}{\varepsilon^2}e^{-\pi/2\sqrt{\varepsilon}}\Big). \tag{3.18}$$

In this proof we shall use the abbreviations

$$C_\varepsilon=\frac{2\pi\beta_2\varepsilon^a}{\sinh(\beta_2\varepsilon^a\frac{\pi}{2})}+\frac{2\pi\beta_3\varepsilon^a}{\sinh(\beta_3\varepsilon^a\frac{\pi}{2})},\qquad D_\varepsilon=\frac{2\pi}{\sqrt{\varepsilon}\sinh(\frac{\pi}{2\sqrt{\varepsilon}})}.$$

Note that, as $\varepsilon\to 0$, we have

$$\frac{2\pi\beta_j\varepsilon^a}{\sinh(\beta_j\varepsilon^a\frac{\pi}{2})}=4+O(\varepsilon^a),\qquad D_\varepsilon=\frac{4\pi}{\sqrt{\varepsilon}}e^{-\pi/(2\sqrt{\varepsilon})}\Big(1+O(e^{-\pi/\sqrt{\varepsilon}})\Big). \tag{3.19}$$

We shall consider $\varepsilon$ small so that

$$\frac{3\pi}{\sqrt{\varepsilon}}e^{-\pi/2\sqrt{\varepsilon}}\le D_\varepsilon\le\frac{5\pi}{\sqrt{\varepsilon}}e^{-\pi/2\sqrt{\varepsilon}} \tag{3.20}$$

By (3.16) and (3.18), since $0<\mu\le\delta\varepsilon^{3/2}$,

$$\widetilde{G}_\mu(A_1,A_2,A_3)=\widetilde{g}_0(A_2,A_3)+\mu D_\varepsilon\cos A_1+O\Big(\mu\varepsilon^{-1/2}e^{-\pi/\sqrt{\varepsilon}}+\mu\delta\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}\Big). \tag{3.21}$$

Since $(A_1,A_2,A_3)=a_1\Omega_1+a_2\Omega_2+a_3\Omega_3=(a_1,(a_1\varepsilon^{a+1/2}+a_2)\beta+a_3\beta')$, the homoclinic function $\widetilde{G}_\mu$ in the new basis $\{\Omega_1,\Omega_2,\Omega_3\}$ writes

$$\widetilde{H}_\mu(a_1,a_2,a_3)=\widetilde{G}_\mu\Big(a_1\Omega_1+a_2\Omega_2+a_3\Omega_3\Big)=\widetilde{G}_\mu\Big(a_1,(a_1\varepsilon^{a+1/2}+a_2)\beta+a_3\beta'\Big). \tag{3.22}$$

Define $\widetilde{h}_0(b_2,a_3)=\widetilde{g}_0(b_2\beta+a_3\beta')$. By (3.22) and (3.21), there exists $c_0(\delta)>0$ such that, for all $0<\varepsilon\le c_0(\delta)$,

$$\widetilde{H}_\mu(a_1,a_2,a_3)=\widetilde{h}_0\Big(a_1\varepsilon^{a+1/2}+a_2,a_3\Big)+\mu D_\varepsilon\cos a_1+O\Big(\mu\delta\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}\Big). \tag{3.23}$$

We derive from this latter expression and (3.20) that

$$\widetilde{H}_\mu(a_1,a_2,a_3)=\widetilde{h}_0\Big(a_1\varepsilon^{a+1/2}+a_2,a_3\Big)+O\Big(\mu\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}\Big). \tag{3.24}$$

By (3.17) and (3.19)

$$\widetilde{g}_0(A_2,A_3)=\mu C_\varepsilon-2\mu(A_2^2+A_3^2)+O\Big(\mu\varepsilon^a(A_2^2+A_3^2)\Big)+O\Big(\mu(A_2^4+A_3^4)\Big)+O(\mu^2). \tag{3.25}$$

We shall assume in the sequel of the proof that $a_2,a_3\in[-\rho,\rho]$, $a_1\in[-2\pi,2\pi]$, so that, since $\rho=\pi\varepsilon^{a+1/2}$, there results $a_1\varepsilon^{a+1/2}\in[-2\rho,2\rho]$, $b_2=a_1\varepsilon^{a+1/2}+a_2\in[-3\rho,3\rho]$ and $b_2^4+a_3^4=O(\rho^4)$. Moreover we have that $\mu^2\le\mu\delta\rho^2$ and there exists $c_1(\delta)\in(0,c_0(\delta))$ such that, if $0<\varepsilon\le c_1(\delta)$, then $\varepsilon^a\le\delta$ and $\rho^4\le\delta\rho^2$. Note also that, since $\beta,\beta'$ are orthonormal vectors, we have $A_2^2+A_3^2=b_2^2+a_3^2$. Finally we derive from (3.25) that, for $0<\varepsilon\le c_1(\delta)$,

$$\widetilde{h}_0(b_2,a_3)=\mu C_\varepsilon-2\mu(b_2^2+a_3^2)+O(\mu\delta\rho^2). \tag{3.26}$$

Since $\rho=\pi\varepsilon^{a+1/2}$ we have $\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}=o(\rho^2)$ as $\varepsilon\to 0$; therefore, by (3.24) and (3.26), there exist $K_0>0$, $c_2(\delta)\in(0,c_1(\delta))$ such that, for all $0<\varepsilon\le c_2(\delta)$,

$$\widetilde{H}_\mu(a_1,a_2,a_3)=\mu C_\varepsilon-2\mu(b_2^2+a_3^2)+r_0(a_1,a_2,a_3),\quad |r_0(a_1,a_2,a_3)|\le K_0\mu\delta\rho^2, \tag{3.27}$$

where $b_2=a_1\varepsilon^{a+1/2}+a_2$.

We now prove that point $(i)$ of the "splitting condition" 3.1 is satisfied by $\widetilde{G}_\mu$ with $\delta_1=\mu\rho^2/2$, $\widetilde{l}_1(x)=$

$-2\pi$ and $\widetilde{l}_2(x) = 2\pi$ where $x := (a_2, a_3)$. Let us consider $\mathcal{J}(a_2, a_3) := \sup_{a_1 \in [-2\pi, 2\pi]} \widetilde{H}_\mu(a_1, a_2, a_3)$. Since $a_2 \in [-\rho, \rho]$, $-a_2\varepsilon^{-(a+1/2)} \in [-\pi, \pi]$ and we can derive from (3.27) that

$$\mathcal{J}(a_2, a_3) \geq \widetilde{H}_\mu\Big( -a_2\varepsilon^{-(a+1/2)}, a_2, a_3\Big) \geq \mu C_\varepsilon - 2\mu a_3^2 - K_0\mu\delta\rho^2. \tag{3.28}$$

If $a_1 = \pm 2\pi$ then $b_2 = a_1\varepsilon^{a+1/2} + a_2 = a_2 \pm 2\rho$ and then, since $a_2 \in [-\rho, \rho]$, we get $|b_2| \geq \rho$. As a consequence, by (3.27) and (3.28),

$$\widetilde{H}_\mu\Big( \pm 2\pi, a_2, a_3\Big) \leq \mu C_\varepsilon - 2\mu(\rho^2 + a_3^2) + K_0\mu\delta\rho^2 \leq \mathcal{J}(a_2, a_3) - 2\mu\rho^2 + 2K_0\mu\delta\rho^2. \tag{3.29}$$

Choosing $\delta < 1/2K_0$, we get in (3.29) that $\widetilde{H}_\mu(\pm 2\pi, a_2, a_3) \leq \mathcal{J}(a_2, a_3) - \mu\rho^2/2$. It results that condition 3.1-$(i)$ is satisfied with $\delta_1 = \mu\rho^2/2$, $l_1(x) = -2\pi$ and $l_2(x) = 2\pi$ where $x = (a_2, a_3)$.

We now turn to the proof of $(ii)$ and $(iii)$. If $a_2, a_3 \in [-\rho, \rho]$, $a_1 \in [-2\pi, 2\pi]$ and $|b_2| = |a_1\varepsilon^{a+1/2} + a_2| \geq \sqrt{2K_0\delta}\rho$, then, by (3.27) and (3.28),

$$\widetilde{H}_\mu(a_1, a_2, a_3) \leq \mu C_\varepsilon - 2\mu(a_3^2 + 2K_0\delta\rho^2) + K_0\mu\delta\rho^2 \leq \mu C_\varepsilon - 2\mu a_3^2 - 3K_0\mu\delta\rho^2 < \mathcal{J}(a_2, a_3).$$

Hence

$$\mathcal{J}(a_2, a_3) = \sup\Big\{\widetilde{H}_\mu(a_1, a_2, a_3)\ ;\ a_1\varepsilon^{a+1/2} \in \Big[ -a_2 - \sqrt{2K_0\delta}\rho, -a_2 + \sqrt{2K_0\delta}\rho\Big]\Big\}. \tag{3.30}$$

We use here that, since $2K_0\delta < 1$, $[-a_2 - \sqrt{2K_0\delta}\rho, -a_2 + \sqrt{2K_0\delta}\rho] \subset [-2\rho, 2\rho] = [-2\pi\varepsilon^{a+1/2}, 2\pi\varepsilon^{a+1/2}]$. Writing $a_1 = (b_2 - a_2)\varepsilon^{-(a+1/2)}$, we derive from (3.30) and (3.23) that

$$\begin{aligned}\mathcal{J}(a_2, a_3) &= \sup_{b_2 \in [-\sqrt{2K_0\delta}\rho, \sqrt{2K_0\delta}\rho]} \widetilde{H}_\mu\Big((b_2 - a_2)\varepsilon^{-a-1/2}, a_2, a_3\Big) \\ &= \sup_{b_2 \in [-\sqrt{2K_0\delta}\rho, \sqrt{2K_0\delta}\rho]} \Big(\widetilde{h}_0(b_2, a_3) + \mu D_\varepsilon \cos\Big(\frac{b_2 - a_2}{\varepsilon^{a+1/2}}\Big)\Big) + O\Big(\delta\mu\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}\Big). \end{aligned}\tag{3.31}$$

Now, if $b_2 \in [-\sqrt{2K_0\delta}\rho, \sqrt{2K_0\delta}\rho]$ then $b_2\varepsilon^{-a-1/2} \in [-\pi\sqrt{2K_0\delta}, \pi\sqrt{2K_0\delta}]$, so we can write that

$$\cos\Big(\frac{b_2 - a_2}{\varepsilon^{a+1/2}}\Big) = \cos\Big(\frac{-a_2}{\varepsilon^{a+1/2}}\Big) + O(\sqrt{\delta}). \tag{3.32}$$

As a consequence, by (3.31) and (3.20) there holds

$$\begin{aligned}\mathcal{J}(a_2, a_3) &= \sup_{b_2 \in [-\sqrt{2K_0\delta}\rho, \sqrt{2K_0\delta}\rho]} \Big(\widetilde{h}_0(b_2, a_3) + \mu D_\varepsilon \cos\Big(\frac{-a_2}{\varepsilon^{a+1/2}}\Big)\Big) + O\Big(\sqrt{\delta}\mu\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}\Big) \\ &= \widetilde{m}(a_3) + \mu D_\varepsilon \cos\Big(\frac{a_2}{\varepsilon^{a+1/2}}\Big) + O\Big(\sqrt{\delta}\mu\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}\Big), \end{aligned}\tag{3.33}$$

where we have set

$$\widetilde{m}(a_3) := \sup_{b_2 \in [-\sqrt{2K_0\delta}\rho, \sqrt{2K_0\delta}\rho]} \widetilde{h}_0(b_2, a_3). \tag{3.34}$$

Finally, there exists $K_1 > 0$ such that, by (3.33),

$$\mathcal{J}(a_2, a_3) = \mathcal{J}(0, a_3) + \mu D_\varepsilon\Big(\cos\Big(\frac{a_2}{\varepsilon^{a+1/2}}\Big) - 1\Big) + r_1(a_2, a_3), \quad |r_1(a_2, a_3)| \leq K_1\sqrt{\delta}\mu\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}. \tag{3.35}$$

We are now in position to prove condition 3.1-$(ii)$. Assume $0 < \delta \leq \pi^2/4K_1^2$ and choose $\sigma = \rho/6 = \varepsilon^{a+1/2}\pi/6$ and $\delta_2 = 3\pi\mu\varepsilon^{-1/2}e^{-\pi/2\sqrt{\varepsilon}}$. If $a_2 \in [-\sigma, \sigma]$ then $\cos(a_2/\varepsilon^{a+1/2}) - 1 \geq -1 + \sqrt{3}/2 \geq -1/6$. This readily implies, by (3.35) and (3.20), that $\mathcal{J}(a_2, a_3) \geq \mathcal{J}(0, a_3) - \delta_2/2$. If $a_2 \in [-\rho, -\rho + 2\sigma] \cup [\rho - 2\sigma, \rho]$ then $a_2/\varepsilon^{a+1/2} \in [-\pi, -2\pi/3] \cup [2\pi/3, \pi]$, so that $\cos(a_2/\varepsilon^{a+1/2}) - 1 \leq -3/2$. It follows, still by (3.35) and (3.20), that $\mathcal{J}(a_2, a_3) \leq \mathcal{J}(0, a_3) - \delta_2$. This proves condition 3.1-$(ii)$.

In order to prove condition 3.1-($iii$), we notice that, by (3.26) and the definition of $\widetilde{m}$ given in (3.34),

$$\widetilde{m}(a_3) = \mu C_\varepsilon - 2\mu a_3^2 + O(\delta\mu\rho^2). \tag{3.36}$$

Hence there exist $K_2 > 0$ and $c_3(\delta) \in (0, c_2(\delta))$ such that, for all $0 < \varepsilon \le c_3(\delta)$, by (3.36), (3.33) and (3.20)

$$\mathcal{J}(a_2, a_3) = \mathcal{J}(0,0) - 2\mu a_3^2 + r_2(a_2, a_3), \quad \text{with} \quad |r_2(a_2,a_3)| \le K_2\delta\mu\rho^2. \tag{3.37}$$

Let us assume $\delta \le 1/6K_2$ and let $\delta_3 = \mu\rho^2/2$. By (3.37) and (3.20), if $a_3 \in \overline{B}_\sigma^{n-2}$, then $\mathcal{J}(a_2,a_3) \ge \mathcal{J}(0,0) - \delta_3/2$; if $a_3 \in \overline{B}_\rho^{n-2} \backslash B_{\rho-2\sigma}^{n-2}$ then $\mathcal{J}(a_2,a_3) \le \mathcal{J}(0,0) - \delta_3$.

As a conclusion, lemma 3.2 holds with $\delta_0 = \min\{\overline{\delta}, 1/2K_0, \pi^2/4K_1^2, 1/6K_2\}$. ■

**Remark 3.1** *The former splitting condition holds also also for more general perturbations $f(\varphi_1, \ldots, \varphi_n)$ for which $f_0(\varphi_2, \ldots, \varphi_n)$ possesses a nondegenerate maximum at $(\overline{\varphi}_2, \ldots, \overline{\varphi}_n)$ and $f_1(\overline{\varphi}_2, \ldots, \overline{\varphi}_n) \neq 0$ where $f_{k_1}(\varphi_2, \ldots, \varphi_n) = (1/2\pi)\int_0^{2\pi} f(\sigma, \varphi_2, \ldots, \varphi_n)e^{-ik_1\sigma}\, d\sigma$. This kind of condition is considered in theorem 5.2 of [6].*

# 4 The shadowing theorem

In this section we shall prove, under the "splitting condition" 3.1, our general shadowing theorem.

**Theorem 4.1** *Let $n \ge 3$ and assume that the homoclinic function $G_\mu$ satisfies the splitting condition 3.1. Let $\omega_\varepsilon$ be a $(\gamma_\varepsilon, \tau)$-diophantine vector, i.e. $|\omega_\varepsilon \cdot k| \ge \gamma_\varepsilon/|k|^\tau$ $\forall k \in \mathbf{Z}^n \backslash \{0\}$. Then, for all $I_0, I_0' \in \mathbf{R}^n$ such that $(I_0' - I_0) \in \mathrm{Span}\{\Omega_3, \ldots, \Omega_n\}$, there exists an heteroclinic trajectory from $\mathcal{T}_{I_0}$ to $\mathcal{T}_{I_0'}$ which connects a $\eta$-neighbourhood of torus $\mathcal{T}_{I_0}$ to a $\eta$-neighbourhood of torus $\mathcal{T}_{I_0'}$ in the "diffusion time"*

$$T_d \le C\frac{\rho|I_0' - I_0|}{\delta_3} \max\Big\{\frac{1}{\gamma_\varepsilon\sigma^\tau}, |\ln\delta_1|, |\ln\delta_2|, \frac{\Delta}{|\omega_\varepsilon|}\Big\} + |\ln\eta|, \tag{4.38}$$

*where* $\Delta := \Big\{\max_{x\in[-\rho,\rho]\times\overline{B}_\rho^{n-2}} l_2(x) - \min_{x\in[-\rho,\rho]\times\overline{B}_\rho^{n-2}} l_1(x)\Big\}$.

**Remark 4.1** *The diophantine condition on the frequency vector $\omega_\varepsilon$ restricts the values of $\varepsilon$ and $\beta$ that we consider. In any case, if for instance $\beta$ is $(\gamma, n-2)$-diophantine then for $\tau \ge n-1$ there exist $c_0 > 0$ and a sequence $\varepsilon_j \to 0$ such that $\omega_\varepsilon$ is $(\gamma_\varepsilon, \tau)$-diophantine with $\gamma_\varepsilon = c_0\varepsilon^a$, see for example [16].*

**Remark 4.2** *The meaning of (4.38) is the following: the diffusion time $T_d$ is estimated by the product of the number of heteroclinic transitions $k =$ ( heteroclinic jump / splitting ) $= |I_0' - I_0|/\delta_3$, and of the time $T_s$ required for a single transition, that is $T_d \approx k \cdot T_s$. The time for a single transition $T_s$ is bounded by the maximum time between the "ergodization time" $(1/\gamma_\varepsilon\sigma^\tau)$, i.e. the time needed for the flow $\omega t$ to make an $\sigma$-net of the torus, and the time $\max\{|\ln\delta_1|, |\ln\delta_2|, \Delta/|\omega_\varepsilon|\}$ needed to "shadow" homoclinic orbits for the forced pendulum equation. We use here that these homoclinic orbits are exponentially asymptotic to the equilibrium.*

*We could prove also the existence of connecting orbits for all $I_0' - I_0 \in$ Span $\{\Omega_2, \ldots, \Omega_n\}$. In this case the number $k$ of heteroclinic transitions would depend also on $\delta_2$, see remark 4.3.*

Proof. Still for simplicity of notation we write the proof for $n = 3$. Then $I_0' - I_0 = \pm|I_0' - I_0|\Omega_3$; we assume for definitiveness that $I_0' - I_0 = |I_0' - I_0|\Omega_3$, so that $(I_0' - I_0) \cdot (\sum_{j=1}^3 a_j\Omega_j) = |I_0' - I_0|a_3$.

We choose the number of heteroclinic transitions as

$$k = \Big[\frac{8|I_0' - I_0|\rho}{\delta_3}\Big] + 1. \tag{4.39}$$

By lemma 2.1, in order to prove the theorem, it is sufficient to find a critical point of the $k$-bump heteroclinic function $\mathcal{F}^k_\mu : \mathbf{T}^3 \times \mathbf{R}^k \to \mathbf{R}$ such that

$$\theta_k - \theta_1 = O\Big(\frac{\rho|I_0' - I_0|}{\delta_3}\max\Big\{\frac{1}{\gamma_\varepsilon\sigma^\tau}, |\ln\delta_1|, |\ln\delta_2|, \frac{\Delta}{|\omega_\varepsilon|}\Big\}\Big). \tag{4.40}$$

We introduce suitable coordinates $(a_1, a_2, a_3, s_1, \ldots, s_k) \in \mathbf{R}^3 \times (-2\pi, 2\pi)^k$ defined by

$$A = \overline{A} + \sum_{j=1}^{3} a_j\Omega_j \quad \text{and} \quad \forall i = 1, \ldots, k, \quad \theta_i = \frac{(\eta_i + s_i - a_1)|\Omega_1|}{|\omega_\varepsilon|}, \tag{4.41}$$

where $\eta_i$ are constants to be chosen later. Let $H^k_\mu(a,s) = F^k_\mu(A,\theta)$ be the "$k$-bump homoclinic function" and $\mathcal{H}^k_\mu(a,s) = \mathcal{F}^k_\mu(A,\theta)$ be the "$k$-bump heteroclinic function" expressed in the new variables $(a,s)$.

The function $\mathcal{H}^k_\mu$ does not depend on $a_1$, since, by the invariance property (2.10) (we recall that $\Omega_1 = |\Omega_1|\omega_\varepsilon/|\omega_\varepsilon|$),

$$\begin{aligned} \mathcal{H}^k_\mu(a,s) &= F^k_\mu\Big(\overline{A} + \sum_{j=1}^{3} a_j\Omega_j, \frac{(\eta_1 + s_1 - a_1)|\Omega_1|}{|\omega_\varepsilon|}, \ldots, \frac{(\eta_k + s_k - a_1)|\Omega_1|}{|\omega_\varepsilon|}\Big) - (I_0' - I_0)\cdot\Big(\sum_{j=1}^{3} a_j\Omega_j\Big) \\ &= F^k_\mu\Big(\overline{A} + \sum_{j=2}^{3} a_j\Omega_j, \frac{(\eta_1 + s_1)|\Omega_1|}{|\omega_\varepsilon|}, \ldots, \frac{(\eta_k + s_k)|\Omega_1|}{|\omega_\varepsilon|}\Big) - |I_0' - I_0|a_3. \end{aligned} \tag{4.42}$$

In the sequel of the proof we shall use the abbreviation $\mathcal{H}^k_\mu = \mathcal{H}^k_\mu(0, a_2, a_3, s)$.

We now choose the constants $(\eta_1, \ldots, \eta_k) \in \mathbf{R}^k$. Note that, since $\omega_\varepsilon$ is $(\gamma_\varepsilon, \tau)$-diophantine, $\Omega_1$ satisfies the diophantine condition

$$|\Omega_1 \cdot k| \geq \frac{\gamma_\varepsilon|\Omega_1|}{|\omega_\varepsilon||k|^\tau}, \quad \forall k \in \mathbf{Z}^n\backslash\{0\}.$$

Hence, by the results of [4], there exists $\overline{C} > 0$ such that the "ergodization time" $T_e$ of the torus $\mathbf{T}^3$ run by the linear flow $\Omega_1 t$, i.e the smallest time for which $\{\Omega_1 t \; ; \; 0 \leq t \leq T_e\}$ is a $\sigma-$ net of the torus, can be bounded from above by $\overline{C}|\omega_\varepsilon|/(\gamma_\varepsilon\sigma^\tau)$. Hence for each interval $J$ of length greater or equal to $\overline{C}|\omega_\varepsilon|/(\gamma_\varepsilon\sigma^\tau)$ there exists $\eta \in J$ such that

$$d(\eta\Omega_1, 2\pi\mathbf{Z}^3) < \sigma. \tag{4.43}$$

In particular there exists a constant $C_2$ and there exist $\eta_i$ such that

$$\frac{|\omega_\varepsilon|}{C_1|\Omega_1|}\ln\Big(8C_1\frac{1}{\min\{\delta_1, \delta_2\}}\Big) + \Delta \leq \eta_{i+1} - \eta_i \leq \frac{|\omega_\varepsilon|}{C_1|\Omega_1|}\ln\Big(8C_1\frac{1}{\min\{\delta_1, \delta_2\}}\Big) + \frac{C_2|\omega_\varepsilon|}{\gamma_\varepsilon\sigma^\tau} + \Delta, \tag{4.44}$$

$$\eta_i\Omega_1 \equiv \chi_i, \; \text{mod} 2\pi\mathbf{Z}^3, \quad \chi_i = y_i\Omega_2 + z_i\Omega_3 \quad \text{with } |y_i| < \sigma, \; |z_i| < \sigma. \tag{4.45}$$

In order to prove the theorem we just need to prove the existence of a critical point of $\mathcal{H}^k_\mu$ in $\mathbf{R}^2 \times (\min l_1, \max l_2)^k$. The upperbound of the diffusion time given in (4.38) will then be a consequence of (4.44) and (4.39). Indeed, by (4.41) and (4.44) we get that

$$\theta_{i+1} - \theta_i = \frac{(\eta_{i+1} - \eta_i)|\Omega_1|}{|\omega_\varepsilon|} + \frac{(s_{i+1} - s_i)|\Omega_1|}{|\omega_\varepsilon|} \leq \frac{1}{C_1}\ln\Big(8C_1\frac{1}{\min\{\delta_1, \delta_2\}}\Big) + \frac{C_2|\Omega_1|}{\gamma_\varepsilon\sigma^\tau} + \frac{2\Delta|\Omega_1|}{|\omega_\varepsilon|}. \tag{4.46}$$

By (4.46) there exists $C > 0$ such that the time $\theta_{i+1} - \theta_i$ "spent for a single transition" is bounded by

$$T_s := \max_i(\theta_{i+1} - \theta_i) \leq C\max\Big\{\frac{1}{\gamma_\varepsilon\sigma^\tau}, |\ln\delta_1|, |\ln\delta_2|, \frac{\Delta}{|\omega_\varepsilon|}\Big\}. \tag{4.47}$$

¿From (4.47) and (4.39) we derive immediately (4.40) and then (4.38).

We now provide, using lemma 2.2, a suitable expression of the $k$-bump heteroclinic function $\mathcal{H}_\mu^k$. By lemma 2.2, the invariance property (2.7), (4.45) and since $G_\mu : \mathbf{T}^3 \to \mathbf{R}$, we get

$$\begin{aligned}
\mathcal{H}_\mu^k(a_2,a_3,s) &= \sum_{i=1}^k \Big[F_\mu\Big(\overline{A}+\sum_{j=2}^3 a_j\Omega_j, \frac{(\eta_i+s_i)|\Omega_1|}{|\omega_\varepsilon|}\Big) + S_i(a_2,a_3,s_{i-1},s_i,s_{i+1}) - \frac{|I_0'-I_0|}{k}a_3\Big] \\
&= \sum_{i=1}^k \Big[G_\mu\Big(\overline{A}+\sum_{j=2}^3 a_j\Omega_j + \eta_i\Omega_1 + s_i\Omega_1\Big) + S_i(a_2,a_3,s_{i-1},s_i,s_{i+1}) - \frac{|I_0'-I_0|}{k}a_3\Big] \\
&= \sum_{i=1}^k \Big[H_\mu\Big(s_i, a_2+y_i, a_3+z_i\Big) - \frac{|I_0'-I_0|}{k}a_3 + S_i\Big],
\end{aligned} \tag{4.48}$$

where $S_i := S_i(a_2,a_3,s_{i-1},s_i,s_{i+1}) = R_i(A,\theta_{i-1},\theta_i,\theta_{i+1})$ after the change of variables (4.41). The left hand side inequality in (4.44) implies that

$$\theta_{i+1}-\theta_i \geq \frac{(\eta_{i+1}-\eta_i-\Delta)|\Omega_1|}{|\omega_\varepsilon|} \geq \frac{1}{C_1}\ln\Big(8C_1\frac{1}{\min\{\delta_1,\delta_2\}}\Big);$$

hence, by (2.13),

$$|S_i| \leq \frac{\min\{\delta_1,\delta_2\}}{8}. \tag{4.49}$$

We will maximize $\mathcal{H}_\mu^k$ in the open set

$$U = \Big\{(a_2,a_3,s)\in\mathbf{R}^{k+2} \;\Big|\; \forall i\ \ a_2+y_i\in(-\rho,\rho), a_3+z_i\in(-\rho,\rho),\ s_i\in\Big(l_1(a_2+y_i,a_3+z_i), l_2(a_2+y_i,a_3+z_i)\Big)\Big\}.$$

$U\neq\emptyset$ since $\{0\}\times\{0\}\times\Pi_{i=1}^k(l_1(y_i,z_i),l_2(y_i,z_i))\subset U$. Since $U$ is bounded, $\mathcal{H}_\mu^k$ attains its maximum over $\overline{U}$ at some point $(\overline{a},\overline{s}) = (\overline{a}_2,\overline{a}_3,\overline{s})$. It is enough to prove that $(\overline{a},\overline{s})\in U$.

- We first prove that for all $i$, $\overline{s}_i\in(l_1(\overline{a}_2+y_i,\overline{a}_3+z_i), l_2(\overline{a}_2+y_i,\overline{a}_3+z_i))$. Since $(\overline{a},\overline{s})$ is a maximum point of $\mathcal{H}_\mu^k$ in $\overline{U}$, for any $t\in[l_1(\overline{a}_2+y_i,\overline{a}_3+z_i), l_2(\overline{a}_2+y_i,\overline{a}_3+z_i)]$, replacing $\overline{s}_i$ with $t$ does not increase $\mathcal{H}_\mu^k$. Since such a substitution alters at most three terms among $S_1,\ldots,S_k$ in (4.48), we obtain, using (4.49), that for any $i$, for any $t\in[l_1(\overline{a}_2+y_i,\overline{a}_3+z_i), l_2(\overline{a}_2+y_i,\overline{a}_3+z_i)]$,

$$H_\mu(\overline{s}_i,\overline{a}_2+y_i,\overline{a}_3+z_i) \geq H_\mu(t,\overline{a}_2+y_i,\overline{a}_3+z_i) - \frac{3}{4}\min\{\delta_1,\delta_2\}.$$

  Hence

$$H_\mu(\overline{s}_i,\overline{a}_2+y_i,\overline{a}_3+z_i) \geq \mathcal{J}(\overline{a}_2+y_i,\overline{a}_3+z_i) - \frac{3\delta_1}{4},$$

  and, by condition 3.1-$(i)$, this implies that $\overline{s}_i\in(l_1(\overline{a}+\overline{\chi}_i), l_2(\overline{a}+\overline{\chi}_i))$, where we have set $\overline{\chi}_i = (y_i,z_i)$.

- We now prove that for all $i$, $\overline{a}_2+y_i\in(-\rho,\rho)$. Indeed we have by (4.48) and (4.49)

$$\begin{aligned}
\mathcal{H}_\mu^k(\overline{a}_2,\overline{a}_3,\overline{s}) &\leq \sum_{i=1}^k \Big[H_\mu(\overline{s}_i,\overline{a}_2+y_i,\overline{a}_3+z_i) + \frac{\min\{\delta_1,\delta_2\}}{8} - \frac{|I_0'-I_0|}{k}\overline{a}_3\Big] \\
&\leq \sum_{i=1}^k \Big[\mathcal{J}(\overline{a}_2+y_i,\overline{a}_3+z_i) + \frac{\delta_2}{8} - \frac{|I_0'-I_0|}{k}\overline{a}_3\Big].
\end{aligned} \tag{4.50}$$

  On the other hand, still by (4.48) and (4.49), choosing $s=\widetilde{s}=(\widetilde{s}_1,\ldots,\widetilde{s}_k)$ with $\widetilde{s}_i\in(l_1(y_i,\overline{a}_3+z_i), l_2(y_i,\overline{a}_3+z_i))$ so that $H_\mu(\widetilde{s}_i,y_i,\overline{a}_3+z_i) = \mathcal{J}(y_i,\overline{a}_3+z_i)$, we get

$$\mathcal{H}_\mu^k(0,\overline{a}_3,\widetilde{s}) = \sum_{i=1}^k \Big[H_\mu\Big(\widetilde{s}_i,y_i,\overline{a}_3+z_i\Big) - \frac{|I_0'-I_0|}{k}\overline{a}_3 + S_i\Big]$$

$$\geq \sum_{i=1}^{k}\Big[\mathcal{J}(y_i,\overline{a}_3+z_i)-\frac{\delta_2}{8}-\frac{|I_0'-I_0|}{k}\overline{a}_3\Big]$$
$$\geq \sum_{i=1}^{k}\Big[\mathcal{J}(0,\overline{a}_3+z_i)-\frac{5\delta_2}{8}-\frac{|I_0'-I_0|}{k}\overline{a}_3\Big], \tag{4.51}$$

since $|y_i|<\sigma$ and by condition 3.1-$(ii)$. Since $\mathcal{H}_\mu^k(\overline{a},\overline{s})\geq\mathcal{H}_\mu^k(0,\overline{a}_3,\widetilde{s})$, we can derive from (4.50) and (4.51) that there exists $i_0\in\{1,\ldots,k\}$ such that $\mathcal{J}(\overline{a}_2+y_{i_0},\overline{a}_3+z_{i_0})\geq\mathcal{J}(0,\overline{a}_3+z_{i_0})-3\delta_2/4$. Still by condition 3.1-$(ii)$, $\overline{a}_2+y_{i_0}\in(-\rho+2\sigma,\rho-2\sigma)$. As a result, since (by (4.45)) $|y_i-y_{i_0}|\leq 2\sigma$ for all $i$, we get that $\overline{a}_2+y_i\in(-\rho,\rho)$.

- At last we prove that, for all $i$, $\overline{a}_3+z_i\in(-\rho,\rho)$. By (4.48) and (4.49), choosing $s=\widetilde{s}=(\widetilde{s}_1,\ldots,\widetilde{s}_k)$ such that $\widetilde{s}_i\in(l_1(y_i,z_i),l_2(y_i,z_i))$ so that $H_\mu(\widetilde{s}_i,y_i,z_i)=\mathcal{J}(y_i,z_i)$,

$$\mathcal{H}_\mu^k(0,0,\widetilde{s})=\sum_{i=1}^{k}\Big[H_\mu\Big(\widetilde{s}_i,y_i,z_i\Big)+S_i\Big]\geq\sum_{i=1}^{k}\Big[\mathcal{J}(y_i,z_i)-\frac{\min\{\delta_1,\delta_2\}}{8}\Big]\geq\sum_{i=1}^{k}\Big[\mathcal{J}(0,0)-\frac{5\delta_3}{8}\Big], \tag{4.52}$$

since $\delta_2<\delta_3$, $|y_i|,|z_i|<\sigma$ and by condition 3.1-$(iii)$. Hence, since $\mathcal{H}_\mu^k(\overline{a},\overline{s})\geq\mathcal{H}_\mu^k(0,0,\widetilde{s})$, by (4.50) there exists $i_0\in\{1,\ldots,k\}$ such that

$$\mathcal{J}(\overline{a}_2+y_{i_0},\overline{a}_3+z_{i_0})+\frac{\delta_3}{8}-\frac{|I_0'-I_0|}{k}\overline{a}_3\geq\mathcal{J}(0,0)-\frac{5\delta_3}{8}. \tag{4.53}$$

Since $|a_3|\leq\rho$ and by (4.39) we get $|I_0'-I_0||\overline{a}_3|/k\leq\delta_3/8$. Hence by (4.53), $\mathcal{J}(\overline{a}_2+y_{i_0},\overline{a}_3+z_{i_0})\geq\mathcal{J}(0,0)-7\delta_3/8$. By condition 3.1-$(iii)$, $\overline{a}_3+z_{i_0}\in(-\rho+2\sigma,\rho-2\sigma)$ and as a consequence, since for all $i$ $|z_i|<\sigma$, we deduce that $\overline{a}_3+z_i\in(-\rho,\rho)$.

We have proved that the maximum point $(\overline{a},\overline{s})\in U$, which completes the proof of the theorem. ■

As a consequence of the general shadowing theorem 4.1 and of lemma 3.2 we get the following theorem

**Theorem 4.2** *Let $f(\varphi)=\sum_{j=1}^{n}\cos\varphi_j$, $n\geq 3$, and $\omega_\varepsilon$ be a $(\gamma_\varepsilon,\tau)$-diophantine vector. Assume $\epsilon$, $\mu\varepsilon^{-3/2}$ and $\mu\varepsilon^{-2a-1}$ to be sufficiently small. Then, for all $I_0,I_0'$ with $\omega_\varepsilon\cdot I_0=\omega_\varepsilon\cdot I_0'$ and $(I_0)_1=(I_0')_1$ there exists a heteroclinic orbit connecting the invariant tori $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I_0'}$ with a diffusion time*

$$T_d\leq C\frac{|I_0'-I_0|}{\mu\varepsilon^{a+(1/2)}}\times\max\Big\{\frac{1}{\gamma_\varepsilon(\varepsilon^{a+(1/2)})^\tau},|\ln(\mu)|\Big\} \tag{4.54}$$

**Remark 4.3** *The number $k$ of heteroclinic transitions used in the proof of theorem 4.2 is polynomial w.r.t $1/\varepsilon$ since our shadowing orbit moves along the directions $(I_2,\ldots,I_n)\in\mathbf{R}^{n-1}$ only ("directions of large splitting"). On the contrary the shadowing orbit connecting tori $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I_0'}$ with $(I_0')_1\neq(I_0)_1$ would shadow an exponentially large number of heteroclinic transitions and the diffusion time would depend also on the contant $\delta_2=3\pi\mu\varepsilon^{-1/2}\exp(-\pi/(2\sqrt{\varepsilon}))$ which describes the exponentially small splitting. In any case, at each transition, the shadowing orbit approaches the homoclinic point only up $\rho=O(\varepsilon^{a+1/2})$ and therefore the time $T_s$ spent for each single transition is polynomial w.r.t $1/\varepsilon$. In this way we deduce that the diffusion time $T_d$ is estimated, up to inverse powers of $1/\varepsilon$, by an exponential $T_d=O(\exp(\pi/(2\sqrt{\varepsilon})))$. Since the (determinant of the) splitting $\Delta=O(\exp(-\pi/(2\sqrt{\varepsilon})))$ we get that $T_d\approx 1/\Delta$, while in [10] and [6] the diffusion time is estimated by $T_d\approx 1/\Delta^p$ for some positive constant $p$.*

**Acknowledgments:** Much of this paper was written when the second author was visiting S.I.S.S.A. in Trieste. He thanks Prof. A. Ambrosetti and the staff of S.I.S.S.A. for their kind hospitality.

# References

[1] A. Ambrosetti, M. Badiale, *Homoclinics: Poincaré-Melnikov type results via a variational approach*, C. R. Acad. Sci. Paris, t. 323, Série I, 1996, 753-758, and 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, 1993, 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: *A functional analysis approach to Arnold Diffusion*, to appear in Annales de l'I.H.P, analyse non linéaire.

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

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

[9] U. Bessi: *Arnold's example with three rotators*, Nonlinearity, 10, 763-781, 1997.

[10] U. Bessi, L. Chierchia, E. Valdinoci: *Upper Bounds on Arnold Diffusion Time via Mather theory*, J. Math. Pures Appl. vol. 80, 1, 105–129, 2001.

[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: *Temps d' instabilité des systèmes hamiltoniens initialement hyperboliques*, C. R. Acad. Sci, t. 332, Serie I, 2001.

[13] J. Cresson: *Conjecture de Chirikov et Optimalité des exposants de stabilité du théorème de Nekhoroshev*, preprint univ. Besançon.

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

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

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

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

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

[19] G. Galllavotti, G. Gentile, V. Mastropietro: *Lindstedt series and Hamilton–Jacobi equation for hyperbolic tori in three time scales problems*, to appear in Journal of Mathematical Physics.

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

[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's Diffusion*, preprint.

*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*