Proceedings of Institute of Mathematics of NAS of Ukraine 2002, Vol. 43, Part 2, 712–719

Arnold Diffusion: a Functional Analysis Approach

Massimiliano BERTI

SISSA, via Beirut 2-4, Trieste, Italy
E-mail: *berti@sissa.it*

We present, in the context of nearly integrable Hamiltonian systems, a functional analysis approach to study the "splitting of the whiskers" and the "shadowing problem" developed in collaboration with P. Bolle in the recent papers [1] and [2] . This method is applied to the problem of Arnold diffusion for nearly integrable partially isochronous systems improving known results.

1 Introduction

Topological instability of action variables in multidimensional nearly integrable Hamiltonian systems is known as Arnold diffusion. This phenomenon was pointed out in 1964 by Arnold himself in his famous paper [3]. For autonomous Hamiltonian systems with two degrees of freedom KAM theory generically implies topological stability of the action variables (i.e. the time-evolution of the action variables for the perturbed system stay close to their initial values for all times). On the contrary, for systems with more than two degrees of freedom, outside a wide range of initial conditions (the so-called "Kolmogorov set" provided by KAM theory), the action variables may undergo a drift of order one in a very long, but finite time called the "diffusion time". After thirty years from Arnold's seminal work [3], attention to Arnold diffusion has been renewed by [4], followed by several papers (see e.g. [5, 6] and references therein).

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

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

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

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

(*i*) For $\mu \neq 0$ small enough, the perturbed stable and unstable whiskers $W_\mu^s(\mathcal{T}_I^\mu)$ and $W_\mu^u(\mathcal{T}_I^\mu)$ split and intersect transversally ("splitting of the whiskers");

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

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

The shadowing problem (iii) has been extensively studied in the last years by geometrical (see e.g. [4, 7, 8, 9, 10, 11]) and by variational methods (see e.g. [5, 12]). A rich literature is also available for the splitting problem see e.g. [4, 13, 14, 15, 16, 17, 18] and references therein.

The aim of this note is to summarize the functional analysis approach developed in the recent papers [1, 2] (see also [19]), apt to deal with Arnold diffusion, especially with "**splitting**" (i) and "**shadowing**" (iii) problems. The method is illustrated on a relatively simple class of models, namely harmonic oscillators weakly coupled with a pendulum through purely spatial perturbations. Precisely we consider nearly integrable partially *isochronous* Hamiltonian systems given by

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

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

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

Under assuption $(H1)$ *all* the invariant tori are preserved by the perturbation, being just slightly deformed (in [1] an elementary proof, not based on any KAM technique, is given). As a consequence, the existence of a transition chain (ii) follows immediately once the "splitting of the whiskers" (i) is proved.

As applications of our shadowing theorems and our splitting estimates, we will consider the following two particular cases

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

(b) the frequencies of the harmonic oscillators form a Diophantine vector $\omega_\varepsilon = (1/\sqrt{\varepsilon}, \beta\varepsilon^a)$ with $a \geq 0$, $\mu\varepsilon^{-3/2}$ small and the perturbation $f(\varphi, q) = (1 - \cos q) f(\varphi)$ ("three-time-scales problem" with perturbations preserving all the unperturbed invariant tori). This corresponds, after a time rescaling, to $\omega = (1, \beta\varepsilon^a\sqrt{\varepsilon})$ in (1). Hamiltonian systems with three time scales have been introduced in [4] as a description of the d'Alembert problem in celestial mechanics.

2 The functional analysis approach

We now describe the **functional analysis approach** developed to prove both the results on the shadowing theorem and on the "splitting of the whiskers". It is based on a finite dimensional reduction of Lyapunov–Schmidt type, variational in nature, introduced in [20] and in [21], and later extended in [22, 23] in order to construct shadowing orbits of "multibump" type. For simplicity we describe our approach when the perturbation term $f(\varphi, q) = (1 - \cos q) f(\varphi)$ so that the tori $\mathcal{T}_{I_0}$ are still invariant for $\mu \neq 0$ (we underline however that in [1] the shadowing analysis is carried out also for a general perturbation term $f(\varphi, q)$).

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

$$\dot\varphi = \omega, \qquad \dot I = -\mu(1 - \cos q)\ \nabla f(\varphi), \qquad \dot q = p, \qquad \dot p = \sin q - \mu \sin q\ f(\varphi). \tag{3}$$

The dynamics on the angles φ is given by $\varphi(t) = \omega t + A$ so that (3) are reduced to the quasi-periodically forced pendulum equation

$$-\ddot{q} + \sin q = \mu \ \sin q \ f(\omega t + A), \tag{4}$$

corresponding to the Lagrangian

$$\mathcal{L}_{\mu,A}(q, \dot{q}, t) = \frac{\dot{q}^2}{2} + (1 - \cos q) + \mu(\cos q - 1) f(\omega t + A). \tag{5}$$

For each solution $q(t)$ of (4) one recovers the dynamics of the actions $I(t)$ by quadratures in (3).

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

$$\Phi_{\mu,A}(q) := \int_{\mathbb{R}} \mathcal{L}_{\mu,A}(q(t), \dot{q}(t), t) \ dt. \tag{6}$$

$\Phi_{\mu,A}$ is smooth on $q_0 + H^1(\mathbb{R})$ and critical points q of $\Phi_{\mu,A}$ are homoclinic solutions to 0, mod 2π, of (4). These critical points q are in fact smooth functions of the time t and are exponentially decaying to 0, mod 2π, as $|t| \to +\infty$.

The unperturbed functional $\Phi_0 := \Phi_{0,A}$ does not depend on A and possesses the 1-dimensional manifold of critical points $Z := \{q_\theta \mid \theta \in \mathbb{R}\}$ with tangent space at q_θ spanned by $\dot{q}_\theta$. All the unperturbed critical points q_θ are degenerate since $d^2\Phi_0(q_\theta)[\dot{q}_\theta] = 0$. However q_θ are non-degenerate critical points of the restriction $\Phi_{0|E_\theta}$ for any subspace E_θ supplementary to $\langle \dot{q}_\theta \rangle$. It is then possible to apply a Lyapunov–Schmidt type reduction, based on the Implicit Function Theorem, to find near q_θ, for μ small, critical points $q^\mu_{A,\theta}$ of $\Phi_{\mu,A}$ restricted to E_θ; more precisely $q^\mu_{A,\theta} = q_\theta + w^\mu_{A,\theta}$ with $w^\mu_{A,\theta} \in E_\theta$, $||w^\mu_{A,\theta}|| = O(\mu)$ and $d\Phi_{\mu,A}(q^\mu_{A,\theta})_{|E_\theta} = 0$. We call the functions $q^\mu_{A,\theta}$ "**1-bump pseudo-homoclinic solutions**" of the quasi-periodically forced pendulum (4).

It turns out that the 1-dimensional manifold $Z_\mu = \{q^\mu_{A,\theta} \mid \theta \in \mathbb{R}\}$ is a "natural constraint" for the action functional $\Phi_{\mu,A}$, namely any critical point of $\Phi_{\mu,A|Z_\mu}$ is a critical point of $\Phi_{\mu,A}$, and hence a true solution of equation (4) homoclinic to 0 (mod 2π).

In [1] the above finite dimensional reduction is performed using two different supplementary spaces to $\langle \dot{q}_\theta \rangle$: one is better suited for the shadowing arguments, the other is better suited for studying the splitting problem in presence of "high frequencies".

Shadowing. For dealing with the shadowing problem, we choose as supplementary space

$$E_\theta = \big\{ w : \mathbb{R} \to \mathbb{R} \ \big| \ w(\theta) = 0 \big\}. \tag{7}$$

E_θ and $\langle \dot{q}_\theta \rangle$ are supplementary since $\dot{q}_0(0) \neq 0$. The choice of the supplementary space E_θ is very well suited to perform the shadowing theorem because the corresponding "1-bump pseudo solutions" $q^\mu_{A,\theta}(t)$ are true solutions of (4) except at the instant $t = \theta$ where $\dot{q}^\mu_{A,\theta}(t)$ may have a jump, even though $q^\mu_{A,\theta}(t)$ is continuous at $t = \theta$ and assumes the value $q^\mu_{A,\theta}(\theta) = q_\theta(\theta) + w^\mu_{A,\theta}(\theta) = q_0(0) = \pi$. The corresponding "reduced action functional" turns out to be

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

By the autonomy of the system $F_\mu(A, \theta) = G_\mu(A + \omega\theta)$ where $G_\mu(A) := F_\mu(A, 0)$. The function $G_\mu : \mathbb{T}^n \to \mathbb{R}$, called the **homoclinic function**, has a neat geometric meaning: it is the difference between the generating functions of the exact Lagrangian stable and unstable manifolds $W^{s,u}(\mathcal{T}_{I_0})$ at section $\{q = \pi\}$. Hence, from a geometrical point of view, the choice of the supplementary space E_θ means to study $W^s(\mathcal{T}_{I_0})$ and $W^u(\mathcal{T}_{I_0})$ at the fixed Poincaré section $\{q = \pi\}$.

If $\partial_\theta F_\mu(A,\theta) = 0$ then $q^\mu_{A,\theta}$ is a true homoclinic orbit of the quasi periodically forced pendulum equation (4). Moreover it results that $\partial_A F_\mu(A,\theta) = \int_{-\infty}^{+\infty} \dot{I}_\mu(t)\ dt =$ *"heteroclinic jump"* and hence critical points of $\mathcal{F}_\mu(A,\theta) := F_\mu(A,\theta) - (I'_0 - I_0)\cdot A$ give rise to true heteroclinic orbits between the tori $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I'_0}$. By a Taylor expansion in μ it results that $\mathcal{F}_\mu(A,\theta) = \Phi_0(q_0) + \mu\Gamma(A+\omega\theta) + O(\mu^2) - (I'_0 - I_0)\cdot A$ where $\Gamma(B) = \int_{\mathbb{R}}(\cos q_0(t) - 1) f(\omega t + B)$ is nothing but the Poincaré-Melnikov primitive. Hence, roughly speaking, critical points of Γ give rise, for μ small, $I'_0 - I_0 = O(\mu)$ and $(I'_0 - I_0)\cdot\omega = 0$, to heteroclinic orbits between $\mathcal{T}_{I_0}$ and $\mathcal{T}_{I'_0}$.

Following [22, 23] the above finite dimensional reduction is generalized in [1] in order to find a natural constraint for "k-bump pseudo homoclinic solutions" turning k times near the unperturbed separatrices of the pendulum. The search for "diffusion orbits" is then reduced to find critical points of a finite dimensional functional, which is the natural generalitation of the previous one: the Lagrangian action functional evaluated on the k-bump pseudo homoclinic solutions. In this way under a suitable "splitting condition", satisfied for instance if $G_\mu(A)$ possesses a proper minimum we can prove a general shadowing theorem with explicit estimates on the diffusion time T_d. Denoting by $B_\alpha(A_0)$ the open ball centered at $A_0 \in \mathbb{R}^n$ and of radius α, let assume

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

(*i*) $\inf_{\partial U} G_\mu \geq \inf_U G_\mu + \delta$;

(*ii*) $\sup_{B_\alpha(A_0)} G_\mu \leq \frac{\delta}{4} + \inf_U G_\mu$;

(*iii*) $d(\{A \in U \mid G_\mu(A) \leq \delta/2 + \inf_U G_\mu\}, \{A \in U \mid G_\mu(A) \geq 3\delta/4 + \inf_U G_\mu\}) \geq 2\alpha$.

The following shadowing type theorem holds, where $\rho_U := \mathrm{diam}(\Pi_\omega(U))$ and $\Pi_\omega : \mathbb{R}^n \to \mathbb{R}^n$ denotes the orthogonal projection onto $\omega^\perp$.

Theorem 1. *Assume* $(H1)$ *and the "splitting condition". Then* $\forall\ I_0, I'_0$ *with* $\omega\cdot I_0 = \omega\cdot I'_0$, *there is a heteroclinic orbit connecting the invariant tori* $\mathcal{T}_{I_0}$ *and* $\mathcal{T}_{I'_0}$. *Moreover there exists* $C_3 > 0$ *such that* $\forall\ \eta > 0$ *small enough the "diffusion time"* T_d *needed to go from a* η*-neighbourhood of* $\mathcal{T}_{I_0}$ *to a* η*-neighbourhood of* $\mathcal{T}_{I'_0}$ *is bounded by*

$$T_d \leq C_3 \frac{|I_0 - I'_0|}{\delta}\cdot\rho_U\cdot\max\left(|\log\delta|, \frac{1}{\gamma\alpha^\tau}\right) + C_3|\log(\eta)|. \tag{8}$$

The meaning of the previous estimate (8) is roughly the following: T_d is estimated by the product of the number of heteroclinic transitions k (= number of tori forming the transition chain = heteroclinic jump/splitting) and of the time T_s required for a single transition, namely $T_d = kT_s$. The time for a single transition T_s is bounded by the maximum time between the time needed to "shadow" homoclinic orbits for the quasi-periodically forced pendulum and the "ergodization time" T_e of the torus $\mathbb{T}^n$ run by the linear flow ωt, defined as the time needed for the flow $\{\omega t\}$ to make an α-net of the torus. By a well known result this time can be estimated by $T_e = O(1/\alpha^\tau)$.

The a-priori unstable systems (Case (a)) highlight the improvement of our estimates on diffusion times. In this case it is easy to show that the splitting of the whiskers is $O(\mu)$ using the classical Poincaré–Melnikov function, which for a general perturbation turns out to be $M(A) = \int_{\mathbb{R}}[f(\omega t + A, 0) - f(\omega t + A, q_0(t))]\ dt$ (note that when the perturbation is $f(\varphi, q) = (1 - \cos q) f(\varphi)$ then $M(A)$ reduces to the Poincaré–Melnikov function Γ previously defined). Then our shadowing method yields

Theorem 2. *Assume (H1) and let $M(A)$ possess a proper minimum (or maximum) A_0. Then, for μ small enough, there exist orbits whose action variables undergo a drift of order one, with diffusion time $T_d = O((1/\mu)\log(1/\mu))$.*

Theorem 2 answers a question raised in [24, Section 7] proving that, at least for isochronous systems, it is possible to reach the maximal speed of diffusion $\mu/|\log\mu|$. The estimate on the diffusion time obtained in [4] is $T_d >> O(\exp(1/\mu))$ and that in [8] it is improved to be $T_d = O(\exp(1/\mu))$; recently in [12], by means of Mather theory, the estimate on the diffusion time has been improved to be $T_d = O(1/\mu^{2\tau+1})$; in [10] it is obtained via geometric methods that $T_d = O(1/\mu^{\tau+1})$. It is worth pointing out that the estimates given in [12] and [10], while providing a diffusion time polynomial in the splitting, depend on the diophantine exponent τ and hence on the number of rotators n. Instead our estimate (as well as that discussed in [11]) does not depend upon the number of degrees of freedom.

The main reason for which Theorem 2 improves the polynomial estimates $T_d = O(1/\mu^{2\tau+1})$ and $T_d = O(1/\mu^{\tau+1})$, obtained respectively in [12] and [10], 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 [12] and [10]. This implies that the time spent by our diffusion orbit at each transition is $T_s = O(\log(1/\mu))$. Since the number of tori forming the transition chain is equal to $O(1/\text{splitting}) = O(1/\mu)$ the diffusion time is finally estimated by $T_d = O((1/\mu)\log(1/\mu))$.

As mentioned in the introduction variational methods in the context of Arnold diffusion have been used also in [5] and [12]. One possible advantage of our approach is that it may be used to consider more general critical points of the reduced functional, not only minima. Another advantage is that the same shadowing arguments can be used also when the hyperbolic part is a general Hamiltonian in $\mathbb{R}^{2m}$, $m \geq 1$, possessing one hyperbolic equilibrium and a transversal homoclinic orbit.

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

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

For this reason we would like to extend analytically the "reduced action functional" $F_\mu(A,\theta) = \Phi_{\mu,A}(q^\mu_{A,\theta})$ in a complex strip sufficiently wide in the θ variable. However $F_\mu(A,\theta)$ can not be easily analytically extended. Indeed, for θ complex, the supplementary space $E_\theta = \{w : \mathbb{R} \to \mathbb{C} \mid w(\text{Re }\theta) = 0\}$, appearing naturally when we try to extend the definition of $q^\mu_{A,\theta}$ to $\theta \in \mathbb{C}$, does not depend analytically on θ. This breakdown of analyticity, arising when measuring the "splitting of the whiskers" at the fixed Poincaré section $\{q = \pi\}$, is a well known difficulty and has been compensated in [4, 14, 15] via the introduction of tree techniques which enable to prove cancellations in the power series expansions.

Our method to overcome this "loss of analyticity" is different and relies on the introduction of another supplementary space $\widetilde{E}_\theta$, which depends analytically on θ. This "trick" was yet used in [20] to study the exponentially small splitting in rapidly periodically forced systems. Our supplementary space $\widetilde{E}_\theta$ is defined by

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

where $\psi_0(t) = \cosh^2(t)/(1+\cosh t)^3$. $\widetilde{E}_\theta$ and $\langle \dot{q}_\theta \rangle$ are supplementary spaces since $\int_{\mathbb{R}} \psi_0(t)\dot{q}_0(t)dt \neq 0$ (the above choice of ψ_0 is motivated by the fact that $\psi_0(t)$ decays at zero as $\dot{q}_0(t)$ for $|t| \to$

$\pm\infty$ and its singularities are located at $\pm i\pi$ while those of $\dot{q}_0$ stay at $i\pi/2$). The corresponding reduced functional $\widetilde{F}_\mu(A,\theta) := \Phi_{\mu,A}(Q^\mu_{A,\theta})$, where $Q^\mu_{A,\theta}$ are the corresponding "1-bump pseudo-homoclinics solutions", can be analytically extended in a sufficiently large complex strip. More precisely let $f(\varphi,q) = (1-\cos q)f(\varphi)$ with $f(\varphi)$ analytic in $D := \prod_{j=1}^n(\mathbb{R} + i[-a_j,a_j])$ for some $a_i \geq 0$ and define $||f|| := \sup_{A\in D}|f(A)|$. Since $q_0(t)$ has a analytic extension up to the streep $|\text{Im } t| < \pi/2$ we manage to extend, using the contraction mapping theorem, $\widetilde{F}_\mu(A,\theta)$ in $D \times \{|\text{Im } \theta| < ((\pi/2)-\sigma)\}$, for $\mu\sigma^{-3}||f||$ small enough.

By an estimate of $\widetilde{F}_\mu(A,\theta) - \mu\Gamma(A+\omega\theta)$ over its complex domain and a standard lemma on Fourier coefficients of analytical functions we easily obtain an exponentially small bounds for the Fourier coefficients of the splitting function $\widetilde{G}_\mu := \widetilde{F}_\mu(A,0)$. Setting $\widetilde{G}_\mu(A) = \sum_{k\in\mathbb{Z}^n}\widetilde{G}_k \exp^{ik\cdot A}$, $\Gamma(A) = \sum_{k\in\mathbb{Z}^n}\Gamma_k \exp^{ik\cdot A}$ and $f(A) = \sum_{k\in\mathbb{Z}^n} f_k \exp(ik\cdot A)$, the following theorem holds

Theorem 3. *For $\mu||f||\sigma^{-3}$ small enough, $\forall\, k \neq 0$, $k \in \mathbb{Z}^n$, $\forall\, \sigma \in \left(0, \frac{\pi}{2}\right)$,*

$$|\widetilde{G}_k - \mu\Gamma_k| = O\left(\frac{\mu^2||f||^2}{\sigma^4}\exp\left(-\sum_{i=1}^n a_i|k_i|\right)\exp\left(-|k\cdot\omega|\left(\frac{\pi}{2}-\sigma\right)\right)\right). \tag{9}$$

Γ_k are explicitely given by $\Gamma_k = f_k 2\pi(k\cdot\omega)/\sinh\left(k\cdot\omega\frac{\pi}{2}\right)$.

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

$$F_\mu(A,\theta) = \widetilde{F}_\mu(A, \theta + h_\mu(A,\theta)), \qquad h_\mu = O(\mu). \tag{10}$$

This fact enables to transpose the informations on $\widetilde{F}_\mu$ to F_μ and viceversa. The introduction of $\widetilde{F}_\mu$ (which recover the analiticity) may then be interpreted simply as measuring the splitting with a non fixed Poincaré section. Setting $\theta = 0$ in equation (10) we get

$$G_\mu(A) = \widetilde{G}_\mu(A + g_\mu(A)\omega), \qquad \text{where} \quad g_\mu(A) := h_\mu(A,0), \tag{11}$$

namely the splitting function G_μ and $\widetilde{G}_\mu$ are the same up to the change of variables of the torus $\psi_\mu := Id + g_\mu$ close to identity, i.e. $G_\mu = \widetilde{G}_\mu \circ \psi_\mu$.

3 Systems with three time scales

We consider now Hamiltonians with three time scales

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

namely $\mathcal{H}_\mu$ with $\omega_\varepsilon = \left(\frac{1}{\sqrt{\varepsilon}}, \varepsilon^a\beta\right)$. In this case, as an application of the previous estimate (9), it follows easily lower estimates for the splitting and hence for the diffusion time. Assume that f is analytical w.r.t φ_2. Set $\widetilde{G}_\mu(A) = \sum_{k_1\in\mathbb{Z}}\widetilde{g}_{k_1}(A_2)\exp^{ik_1\cdot A_1}$ and $\Gamma(\varepsilon,A) = \sum_{k_1\in\mathbb{Z}}\Gamma_{k_1}(\varepsilon,A_2)\exp^{ik_1\cdot A_1}$. In [1] it is proved

Theorem 4. *For $\mu||f||\varepsilon^{-3/2}$ small there holds*

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

where $b := \Phi_0(q_0)$

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

This result improves the main Theorem I in [18] which holds for $\mu = \varepsilon^p$, $p > 2 + a$; w.r.t. [14] (which deals with more general systems) we remark that our results hold in any dimension, while the results of [14], based on tree techniques and cancellations, are proved for $n = 2$.

Using Theorem 4, in [1] simple conditions on the perturbation f which imply the "splitting condition", are given. For example, setting $f(A_1, A_2) = \sum_{k_1\in\mathbb{Z}} f_{k_1}(A_2)\exp(ik_1A_1)$, the "splitting condition" is satisfied if (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 $f_1(\overline{A}_2 + i(\pi/2)\beta) \neq 0$.

In systems with three time scales it appears that the splitting is not uniform in all the 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 $I_2 \in \mathbb{R}^{n-1}$, where the splitting is just polynomially small w.r.t. $1/\varepsilon$, in a polynomially long diffusion time $T_d = O(1/\varepsilon^q)$. In [2] we prove that, for $n \geq 3$, this is indeed the case (note that Arnold diffusion can take place in the direction I_2 only for $n \geq 3$ because of the conservation of the energy along the orbits). For example we can prove

Theorem 5. *Let* $f(\varphi) = \sum_{j=1}^{n} \cos\varphi_j$, $n \geq 3$, *and* ω_ε *be a* $(\gamma_\varepsilon, \tau)$*-diophantine vector. Assume* ε, $\mu\varepsilon^{-3/2}$ *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*

$$T_d \leq C\frac{|I_0' - I_0|}{\mu\varepsilon^{a+(1/2)}} \times \max\left\{\frac{1}{\gamma_\varepsilon(\varepsilon^{a+(1/2)})^\tau}, |\ln(\mu)|\right\}. \tag{12}$$

The previous phenomenon can not be seen by the splitting estimates given in [14] and [18] where the size of the splitting is measured by the "determinant of the splitting matrix" which turns out to be exponentially small.

In order to prove Theorem 5 we refine the shadowing Theorem 1. 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 Theorem 1, since the homoclinic orbit decays exponentially fast to 0, the time needed to "shadow" homoclinic orbits for the quasi-periodically forced pendulum (4) 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 to get a polynomial estimate for the "ergodization time" T_e. The crucial improvement of the shadowing Theorem 5 allows the shadowing orbit to approach the homoclinic point only up to a polynomially small distance $\alpha = O(\varepsilon^p)$, $p > 0$, (and not exponentially small as it would be required applying the shadowing Theorem 1). Since the ergodization time is estimated by $T_e = O(1/\alpha^\tau)$, it results that the minimum time after which the homoclinic trajectory can "jump" to another torus is only polynomially long w.r.t. $1/\varepsilon$.

These results are the first steps to prove the existence of this phenomenon also for more general systems (with non isochronous terms and more general perturbations).

Acknowledgments

The author thanks P. Bolle for numerous discussions.

[1] Berti M. and Bolle P., A functional analysis approach to Arnold diffusion, *Annales de l'I. H. P, analyse non lineaire*, to appear.

[2] Berti M. and Bolle P., Fast Arnold's diffusion in systems with three time scales, *Discrete and Continuous Dynamical Systems*, to appear.

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

[4] Chierchia L. and Gallavotti G., Drift and diffusion in phase space, *Annales de l'I.H.P, section Physique Théorique*, 1994, V.60, 1–144; see also Erratum, 1998, V.68, 135.

[5] Bessi U., An approach to Arnold diffusion through the calculus of variations, *Nonlinear Analysis T.M.A.*, 1996, V.26, 1115–1135.

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

[7] Marco J.P., Transitions le long des chaînes de tores invariants pour les systèmes hamiltoniens analytiques, *Annales I.H.P.*, 1995, V.64, 205–252.

[8] Gallavotti G., Arnold's diffusion in isochronous systems, *Mathematical Physics, Analysis and Geometry* 1999, V.1, 295–312.

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

[10] Cresson J., Conjecture de Chirikov et optimalité des exposants de stabilité du théorème de Nekhoroshev, Preprint Univ. Besançon, 2000.

[11] Cresson J. and Guillet C., Periodic orbits and Arnold diffusion, Preprint Univ. Besançon, 2001.

[12] Bessi U., Chierchia L. and Valdinoci E., Upper bounds on Arnold diffusion time via mather theory, *J. Math. Pures Appl*, 2001, V.80, N 1, 105–129.

[13] Delshams A., Gelfreich V., Jorba A. and Seara T., Exponentially small splitting of separatrices under fast quasi-periodic forcing, *Comm. Math. Phys.*, 1997, V.189, 35–71.

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

[15] Gallavotti G., Gentile G. and Mastropietro V., Melnikov approximation dominance. Some examples, *Rev. Math. Phys.*, 1999, V.11, N 4, 451–461.

[16] Sauzin D., A new method for measuring the splitting of invariant manifolds, *Ann. Scient. E.N.S*, to appear.

[17] Lochak P., Marco J.P. and Sauzin D., On the splitting of invariant manifolds in multidimensional Hamiltonian systems, Preprint Université Jussieu, 1999.

[18] Pumarino A. and Valls C., Three time scales systems exhibiting persistent Arnold diffusion, Preprint, 1999.

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

[20] Angenent S., A variational interpretation of Melnikov's function and exponentially small separatrix splitting, Lecture Notes of the London Math. Soc., Symplectic Geometry, Editor Dietmar Salamon, 1993.

[21] Ambrosetti A. and Badiale M., Homoclinics: Poincaré–Melnikov type results via a variational approach, *Annales I.H.P. - Analyse nonlin.*, 1998, V.15, N 2, 233–252.

[22] Berti M. and Bolle P., Homoclinics and chaotic behaviour for perturbed second order systems, *Annali di Mat. Pura e Applicata*, 1999, (IV), V. CLXXVI, 323–378.

[23] Berti M. and Bolle P., Variational construction of homoclinics and chaotic behaviour in presence of a saddle-saddle equilibrium, *Annali della Scuola Normale Superiore di Pisa*, 1998, serie IV, V. XXVII, fasc. 2, 331–377.

[24] Lochak P., Arnold diffusion: a compendium of remarks and questions, in Proceedings "Hamiltonian Systems with Three or More Degrees of Freedom" (1995, Agaro), Editor C. Simó, Kluwer Acad. Publ., Dordrecht, 1999, 168–183.