Communications in Mathematical Physics manuscript No.
(will be inserted by the editor)

A numerical study of Arnold diffusion in a priori unstable systems. ⋆

Massimiliano Guzzo[2], **Elena Lega**[1], **Claude Froeschlé**[1]

[1] UNSA, CNRS UMR 6202, Observatoire de Nice, Bv. de l'Observatoire, B.P. 4229, 06304 Nice cedex 4, France.

[2] Università degli Studi di Padova, Dipartimento di Matematica Pura ed Applicata, via Trieste 63, 35121 Padova, Italy.

Abstract This paper concerns the problem of the numerical detection of Arnold diffusion in a priori unstable systems. Specifically, we introduce a new definition of Arnold diffusion which is adapted to the numerical investigation of the problem, and is based on the numerical computation of the stable and unstable manifolds of the system. Examples of this Arnold diffusion are provided in a model system. In this model, we also find that Arnold diffusion behaves as an approximate Markovian process, thus it becomes possible to compute diffusion coefficients. The values of the diffusion coefficients satisfy the scaling $D(\epsilon) \simeq \epsilon^2$. We also find that this law is correlated to the validity of the Melnikov approximation: in fact, the $D(\epsilon) \simeq \epsilon^2$ law is valid up to the same critical value of ϵ for which the error terms of Melnikov approximations have a sharp increment.

1. Introduction

Diffusion in conservative dynamical systems has been a much studied subject in the last decades. Apart from specific examples, the understanding of the general mechanisms which can produce drift and diffusion in the phase space of such systems is an interesting, and still open, problem. The existence of a slow diffusion of the actions in a specific quasi-integrable system has been proved for the first time by Arnold [1]. The proof of the Arnold diffusion is based on the existence of chains of whiskered tori such that, under the effect of a perturbation, the unstable manifold of one intersects the stable manifold of the next one. The sequence of such invariant tori is called transition chain and the shadowing argument used

⋆ Author–created Preprint version of the manuscript accepted for publication on Communications in Mathematical Physics, Volume 290, Number 2, 557-576, 2009. The original publication is available at www.springerlink.com. Digital Object Identifier (DOI) 10.1007/s00220-009-0846-9.

to prove diffusion through the transition chain is called transition chain mechanism. A non generic feature of Arnold's example is that the hyperbolic invariant manifold along which Arnold proves the existence of diffusion is fibered by invariant tori for all values of the perturbing parameter. That is, the restriction of the dynamical system to the invariant manifold is integrable. Generalizations of Arnold's example consider normally hyperbolic invariant manifolds such that the restriction of the dynamics to them is not integrable. As a consequence, the distribution of the invariant whiskered tori has gaps which correspond to the resonances of the dynamical system restricted to the invariant manifold. In [6] the existence of transition chains in regions of the invariant manifold which do not contain a selected number of main resonances is proved. In [8] transition chains crossing the main resonances are constructed by including also stable and unstable manifolds of invariant sets which are topologically different from invariant whiskered tori. The existence of diffusing motions has been proved also in [2], [3], [4] using different models and techniques, including variational methods based on Mather theory, and in [26], [27] using the so called separatrix map. The most important techniques to prove the existence of transitions chains (used in [6], [8], [26], [27]) are based on the so–called Melnikov theory, which provides first order approximations of the stable and unstable manifolds.

Arnold's paper motivated a great debate about the possibility of numerical detection of Arnold diffusion. Few years after the first numerical detections of chaotic motions [12], problems related to the numerical detection of Arnold diffusion were discussed in [7]. In the following decades, many authors studied numerically the diffusion through resonances, referring to it as Arnold diffusion. For example, explicit reference to possible interpretations of numerical diffusion as Arnold diffusion can be found in [17]. Other papers, such as [9], [29], [20], [16], studied the numerical diffusion of orbits in coupled standard maps by changing the perturbation parameters and the number of coupled maps (for a review see also [19] and references therein). Computations of the stable and unstable manifolds of hyperbolic tori related to an Arnold diffusion problem can be found in [25]. In [18], [11], [10] we studied the diffusion of orbits in quasi–integrable systems for values of the perturbing parameters for which there is numerical evidence of applicability of the KAM and Nekhoroshev theorems.

In this paper we study the problem of the numerical detection of Arnold diffusion for an important class of conservative systems, the so–called *a priori unstable* ones following the terminology introduced in [6]. To do this, we first need to define precisely what is the Arnold diffusion that one can measure with numerical experiments, that can be repeated for a finite number of initial conditions and values of the perturbing parameter. We therefore provide a new definition of Arnold diffusion which is based on the computation of the stable and unstable manifolds of two whiskered tori (or other invariant hyperbolic objects) for specific values of the perturbing parameter ϵ. The perturbing parameter is required to be sufficiently small so that the normally hyperbolic invariant manifold is filled by a large volume of whiskered tori. An ideal verification of Arnold diffusion would correspond to the detection of heteroclinic transitions among the stable and unstable manifolds of different whiskered tori. But, the probability of finding an orbit which passes near a selected number of heteroclinic points is very small (in [6], based on a prescribed selection of the successive passages, this is reflected in superexponential estimates of the time of diffusion, in the sense

used there), and moreover the time needed for an orbit to perform an exact heteroclinic excursion would be infinite. Therefore, we base our definition on the detection of approximate heteroclinic transitions, corresponding to the existence of orbits with initial conditions in a neighbourhood of a whiskered torus which enter a finite neighbourhood of another whiskered torus in some finite time, with suitably 'large' variation of the action variables which are constants of motion for $\epsilon = 0$. On the one hand, the new notion of Arnold diffusion is weaker than the usual one, because it refers to finite values of ϵ and to specific neighbourhoods of the whiskered tori. On the other hand, it provides a numerical verification of the topological mechanism which is behind the diffusion of the actions, inspired by the proofs of Arnold diffusion such as [6] and [8]. We provide numerical examples of this Arnold diffusion, with a control of the numerical errors, including also a transition among whiskered tori with *large gaps* between them.

The detection of Arnold diffusion in the sense stated by the new definition requires the precise detection of one approximate heteroclinic transition. Therefore, we needed to set the numerical precision to the high value of 400 digits in the numerical experiments. Relaxing the numerical precision of the computations (precisely we switched to double precision) we could compute the statistical properties of many of these transitions. We remark that, while the lower precision affects drastically the individual integrated orbits after some Lyapunov times, it affects much less the computation of statistical quantities, such as the Lyapunov exponents and the diffusion coefficients (see, for example, [23]).

We find that, for small values of ϵ, Arnold diffusion behaves as an approximate Markovian process (see Section 4 and 5.2 for the precise meaning of 'approximate') allowing one to compute diffusion coefficients, and the values of these diffusion coefficients satisfy a scaling $D(\epsilon) \simeq \epsilon^2$. For higher values of ϵ, data cannot be fitted by the ϵ^2 law, and we do not try any fit because their statistics is poor (see Section 4 and the technical Section 5.2 for details). It is remarkable that the $D(\epsilon) \simeq \epsilon^2$ law is correlated to the validity of the Melnikov approximation, in the sense that the $D(\epsilon) \simeq \epsilon^2$ law is valid up to the same critical value of ϵ for which the error terms of Melnikov approximations (which increase with ϵ as well) have a sharp increment.

The paper is organized as follows: in Section 2 we introduce a new definition of Arnold diffusion suited for numerical experiments; in Section 3 we provide numerical examples of Arnold diffusion; in Section 4 we numerically show that, for small values of ϵ, Arnold diffusion behaves as an approximate Markovian process and we discuss the relevance of Melnikov approximations. The technical tools used through the paper, that is the tools related to normal hyperbolicity, the computation of the stable and unstable manifolds, the statistical tools and the computation of Melnikov approximations are reported in Section 5. Conclusions are provided in Section 6.

2. A definition of Arnold diffusion suited to numerical experiments

We consider dynamical systems defined by a family of smooth symplectic maps: $(I', \varphi') = \phi_\epsilon(I, \varphi)$, with the action–angle variables (I, φ) defined on a domain $B \times \mathbb{T}^n$, with $B \subseteq \mathbb{R}^n$ open bounded. The family ϕ_ϵ depends smoothly on the parameter ϵ. We assume that:

– some actions $I_j, \ldots, I_n$, with $j > 1$, are constants of motion of the unperturbed map ϕ_0;
– ϕ_0 has an invariant sub–manifold Λ_0 which is normally hyperbolic and symplectic (the definition of normal hyperbolicity is recalled in Section 5.1);
– the restriction of ϕ_0 to Λ_0 is an integrable anisochronous map[1].

We call such maps a priori unstable.

We will consider suitably small $|\epsilon|$ such that the map ϕ_ϵ has an invariant submanifold Λ_ϵ which is normally hyperbolic, symplectic, and canonically smoothly conjugate to Λ_0. Therefore, the KAM theorem for maps applies to the restriction of ϕ_ϵ to Λ_ϵ, and ϵ is the small parameter.

For some $c_0 > 0$, and any small ϵ, we require that any orbit $(I(t), \varphi(t)) = \phi_\epsilon^t(I(0), \varphi(0))$, with $(I(0), \varphi(0)) \in \Lambda_\epsilon$, satisfies:

$$\|I(t) - I(0)\| < c_0 \tag{1}$$

for any $t \in \mathbb{Z}$. That is, the motions of ϕ_ϵ with initial conditions on Λ_ϵ are uniformly bounded in the actions. Therefore, Arnold diffusion concerns the dynamics in neighborhoods of Λ_ϵ.

Definition 1. *The problem of Arnold diffusion for ϕ_ϵ consists in proving that, for any suitably small $\epsilon \neq 0$ and for any neighbourhood V of Λ_ϵ there exist motions such that for some $t \in \mathbb{Z}$ it is: $(I(0), \varphi(0)), (I(t), \varphi(t)) \in V$, and*

$$\|I(t) - I(0)\| > 2c_0 \quad . \tag{2}$$

The above definition of Arnold diffusion is not well suited for the numerical study of the problem, because numerical integrations cannot span any value of the perturbing parameter and any small neighbourhood of Λ_ϵ. Therefore, we give below a definition which is more adapted to the numerical investigation, and it still contains most of the whole complexity of Arnold diffusion.

Definition 2. *The problem of the numerical detection of Arnold diffusion for ϕ_ϵ in the subset $\tilde{\Lambda} \subseteq \Lambda_\epsilon$ consists in the numerical detection of:*

- *two points $x' = (I', \varphi'), x'' = (I'', \varphi'') \in \tilde{\Lambda}$ such that the closures $\mathcal{C}(x'), \mathcal{C}(x'')$ of their orbits have empty intersection;*
- *two vectors $\Delta x' = (\Delta I', \Delta\varphi'), \Delta x'' = (\Delta I'', \Delta\varphi'') \in \mathbb{R}^{2n}$;*
- *a positive $t \in \mathbb{N}$ and an index $k \in \{j, ..., n\}$;*

such that:

- *$x' + \Delta x' \in W_u(x')$, where $W_u(x')$ denotes the unstable manifold of x';*
- *$\phi_\epsilon^t(x' + \Delta x') + \Delta x'' \in W_s(x'')$, where $W_s(x'')$ denotes the stable manifold of x'';*
- *for any $(\tilde{I}', \tilde{\varphi}') \in \mathcal{C}(x')$, $(\tilde{I}'', \tilde{\varphi}'') \in \mathcal{C}(x'')$ it is:*

$$\left|\tilde{I}'_k - \tilde{I}''_k\right| > c_k + |\Delta I'_k| + |\Delta I''_k| \tag{3}$$

where c_k is such that any orbit $(I(h), \varphi(h)) = \phi_\epsilon^h(I(0), \varphi(0))$, with $(I(0), \varphi(0)) \in \tilde{\Lambda}$, satisfies:

$$|I_k(h) - I_k(0)| < c_k \quad \forall \quad h \in \mathbb{Z} \quad ; \tag{4}$$

[1] The complete integrability of $\phi_{0|\Lambda_0}$ is intended with reference to the symplectic form $dI \wedge d\varphi_{|\Lambda}$.

for some values of the perturbing parameter ϵ such that the KAM theorem applies to $\phi_{\epsilon|\Lambda_\epsilon}$ and inequality (1) is satisfied.

We remark that Arnold diffusion in the sense of Definition 2 does not exist for the unperturbed map ϕ^0, because in such a case the actions $I_j, \ldots, I_n$ are constants of motion.

The above definition is clearly inspired by the proofs of existence of Arnold diffusion which show that the stable and unstable manifolds of different invariant tori of Λ_ϵ intersect transversely (for a precise statement we refer to [6]). An ideal verification of Arnold diffusion would correspond to the detection of an heteroclinic intersection among $W_u(x')$ and $W_s(x'')$. But, the probability of finding an orbit passing near a selected number of heteroclinic points is very small (in [6], based on a prescribed selection of the successive passages, this is reflected in superexponential estimates of the time of diffusion, in the sense used there), and moreover the time needed to perform an exact heteroclinic excursion would be infinite. Therefore, Definition 2 is based on the detection of approximate heteroclinic transitions, corresponding to the existence of orbits with initial conditions in a neighbourhood of x' which enter a finite neighbourhood of x'' in some finite time, with variation of the action variable I_k as in (3). On the one hand, Definition 2 is weaker than Definition 1 because it refers to finite values of ϵ and to specific neighbourhoods of the whiskered tori. On the other hand it provides also a numerical verification of the topological mechanism which is behind the diffusion of the actions, inspired by the proofs of Arnold diffusion, such as in [6] and [8].

In Section 3 we provide an example of numerical detection of Arnold diffusion for which equation (3) is verified within the numerical errors, i.e. we find:

$$\left|\tilde{I}'_k - \tilde{I}''_k\right| > c_k + |\Delta I'_k| + |\Delta I''_k| + \rho \quad , \tag{5}$$

where $\rho > 0$ is an estimator of the numerical errors (see technical Section 5.1 for details).

3. Arnold diffusion in a model problem

We detect Arnold diffusion in the a priori unstable system defined by the family of maps:

$$\begin{array}{c} \phi_\epsilon : \mathbb{R}^2 \times \mathbb{T}^2 \longrightarrow \mathbb{R}^2 \times \mathbb{T}^2 \\ (\varphi_1, \varphi_2, I_1, I_2) \longmapsto (\varphi'_1, \varphi'_2, I'_1, I'_2) \end{array} \tag{6}$$

such that:

$$\begin{aligned} \varphi'_1 &= \varphi_1 + I_1 \\ \varphi'_2 &= \varphi_2 + I_2 \\ I'_1 &= I_1 - a \sin \varphi'_1 + \epsilon \frac{\sin \varphi'_1}{(\cos \varphi'_1 + \cos \varphi'_2 + c)^2} \\ I'_2 &= I_2 + \epsilon \frac{\sin \varphi'_2}{(\cos \varphi'_1 + \cos \varphi'_2 + c)^2} \quad , \end{aligned} \tag{7}$$

where $a > 0, \epsilon$ and $c > 2$ are parameters (in all the numerical experiments we set $a = 0.4$, $c = 2.1$). The symplectic structure on $\mathbb{R}^2 \times \mathbb{T}^2$ is the standard one:

$d\varphi_1 \wedge dI_1 + d\varphi_2 \wedge dI_2$. According to the definitions given in Section 2, the family (7) indeed defines an a priori unstable system. In fact:

- the action I_2 is a first integral of the unperturbed map ϕ_0;
- ϕ_0 has an invariant manifold Λ_0 defined by:

$$\Lambda_0 = \{(I_1, \varphi_1, I_2, \varphi_2) : \quad \text{such that} \quad (I_1, \varphi_1) = (0, \pi)\} \quad , \tag{8}$$

  which is normally hyperbolic and symplectic. The stable and unstable manifolds of Λ_0 are the product of the stable and unstable manifolds of the hyperbolic fixed point of the standard map:

$$\varphi_1' = \varphi_1 + I_1 \quad , \quad I_1' = I_1 - a \sin \varphi_1'$$

  with $\mathbb{R} \times \mathbb{T}$, domain of (I_2, φ_2).
- the restriction of ϕ_0 to Λ_0 is represented by the 2–dimensional map:

$$\varphi_2' = \varphi_2 + I_2 \quad , \quad I_2' = I_2 \tag{9}$$

  while $I_1' = I_1$, $\varphi_1' = \varphi_1$. The 2–dimensional map (9) is integrable anisochronous, with first integral I_2.

The manifold Λ_0 is invariant also for the map ϕ_ϵ for any ϵ, and it is also normally hyperbolic if ϵ is suitably small. Therefore, according to the notations and definitions given in Section 2, for all such ϵ it is $\Lambda_\epsilon = \Lambda_0$ and the restriction of ϕ_ϵ to the invariant manifold Λ_ϵ is explicitly represented by the 2–dimensional map:

$$\varphi_2' = \varphi_2 + I_2 \quad , \quad I_2' = I_2 + \epsilon \frac{\sin \varphi_2'}{(\cos \varphi_2' + c - 1)^2} \quad , \tag{10}$$

while $I_1' = I_1$, $\varphi_1' = \varphi_1$. The explicit representation of $\Lambda_\epsilon = \Lambda_0$ and of $\phi_{\epsilon|\Lambda_\epsilon}$ simplifies a lot the technical implementation of the numerical experiments. In the following, to simplify the notations, we denote $\phi = \phi_\epsilon$ and $\Lambda = \Lambda_0 = \Lambda_\epsilon$.

To numerically detect Arnold diffusion for the map ϕ in the sense provided by Definition 2 we need to study with some more detail the dynamics of the restricted map (10). First, we fix an interval $D = [0.26, 0.38]$ of the action I_2 and we determine with a numerical method the value ϵ_c such that the KAM theorem is valid for any $0 = \epsilon \leq \epsilon_c$ in some open domain of (I_2, φ_2) containing $D \times \mathbb{T}$. Precisely, because we detected the presence of KAM curves in numerically computed phase portraits of (10) for $0 \leq \epsilon \leq 0.002$, while we did not detect any KAM curve for $\epsilon = 0.0026$, we inferred that $\epsilon_c \in (0.002, 0.0026)$. Because the KAM curves of (10) are topological barriers for the variation of the action I_2 for any motion with initial condition on Λ, condition (1) is satisfied for some $c_0 > 0$ for any $0 \leq \epsilon < 0.002$.

We numerically find Arnold diffusion for $\epsilon = 10^{-6}, 10^{-4} < \epsilon_c$ (of course, other values can be investigated with the same techniques). The phase portraits of the restricted map $\phi_{|\Lambda}$ are reported in Figure 1 top–left panel ($\epsilon = 10^{-6}$) and bottom–left panel ($\epsilon = 10^{-4}$): in both cases the phase portraits contain several KAM curves, and for $\epsilon = 10^{-4}$ also an evident resonance. We remark that, in studies of Arnold diffusion, this kind of resonances are usually called *large gaps* in the distribution of invariant tori.

ϵ	c_2	$\lvert\Delta I_2''\rvert$	ρ	$\frac{\lvert\tilde{I}_2'-\tilde{I}_2''\rvert}{c_2+\lvert\Delta I_2'\rvert+\lvert\Delta I_2''\rvert+\rho}$
10^{-6}	$4\ 10^{-5}$	$< 10^{-12}$	$< 10^{-173}$	> 6
10^{-4}	$4\ 10^{-3}$	$< 3\,10^{-7}$	$< 10^{-113}$	> 1.6

Table 1. Tolerances for the verification of equation (3) and (5). Because the last column has values bigger than 1, inequalities (3) and (5) are satisfied.

The variations of the action I_2 for initial conditions on Λ are bounded by a constant c_2 (see (4)) which can be computed from the phase portraits (see Table 1). The points x', x'' that we find to satisfy equations (3) and (5) belong to the bold KAM curves marked in the two phase portraits (the bottom ones for x' and the upper ones for x''). In both cases $\epsilon = 10^{-4}, 10^{-6}$ we found that the crucial tolerances for the correction $|\Delta I_2'|, |\Delta I_2''|$ on the action I_2 and the estimator ρ of the numerical error are respected within many orders of magnitude (see Table 1 and Figure 1, right panels). Because inequalities (3) and (5) are satisfied (see Table 1, last column), the two computations are a numerical detection of Arnold diffusion in the sense stated by Definition 2. We remark (see Figure 1, bottom–left panel) the presence of an evident resonance between the invariant tori containing x' and x'', and therefore the unstable manifold of x' has crossed this large gap before arriving near the stable manifold of x''. All details of these numerical computations are reported in Section 5.1.

The above numerical detections of Arnold diffusion are based on the computation of pieces of the stable and unstable manifolds. To give also a global geometric idea of how the stable and unstable manifolds support Arnold diffusion we computed their parametrization with respect to their arc–length. Figure 2 reports the computation of the unstable manifold of a point on the torus containing x' for $\epsilon = 10^{-6}$ (see Section 5.1 for the computational details). On the top right panel it appears clearly that I_2 undergoes relatively large fluctuations. The unstable manifolds, which are contained in a plane of constant I_2 for $\epsilon = 0$, are unrolled along the I_2 direction for $\epsilon > 0$, thus supporting diffusion in the neighborhood of Λ. The (many) returns of $W_u(x')$ near the manifold Λ can be well appreciated in the three–dimensional representation (bottom right panel). To compare the value of the action I_2 of these return points with the variation of I_2 along the torus, we represent in the bottom left panel the KAM curve of x' and the vertical segment which corresponds to the representation on the plane (I_2, φ_2) of the points of the unstable manifold with $|\varphi_1 - \pi| \leq 0.5$ (reducing the tolerance on φ_1 decreases the number of points on the figure, but does not decrease the amplitude of the segment). The amplitude of this segment is larger than c_2, providing indication that these returns of $W_u(x')$ near Λ support Arnold diffusion.

4. Statistical properties of Arnold diffusion and Melnikov approximations

The definitions of Arnold diffusion given in Section 2 characterize it as the possibility of orbits with initial conditions near Λ of returning near Λ with a suitably

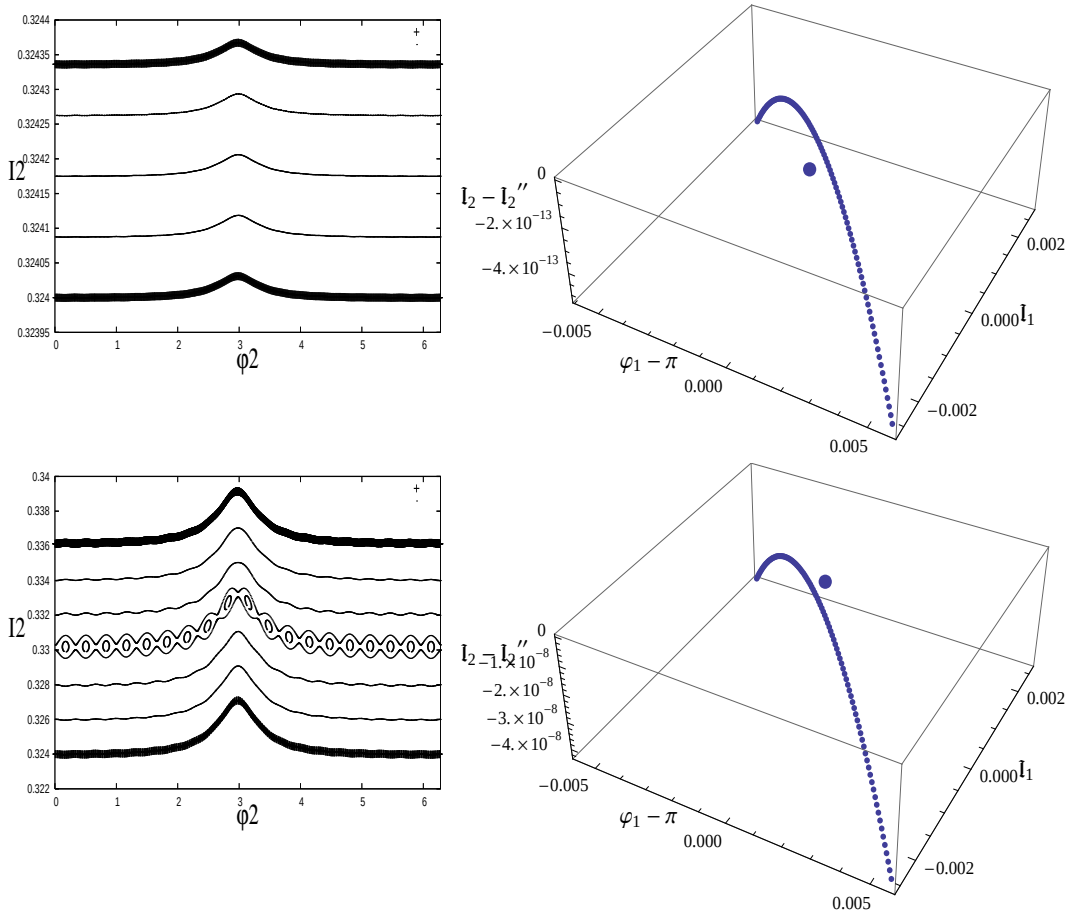

Figure 1. Numerical test of equation (3). **On the top–left.** Phase plane of the map ϕ restricted to Λ for $\epsilon = 10^{-6}$: the invariant tori containing x' and x'' are represented by bold curves. **On the top–right.** Projection on the space of variables $\varphi_1 - \pi, I_1, I_2 - I_2''$ of the point $\phi^T(x' + \Delta x')$ in the orbit of $x' + \Delta x'$ and of the stable manifold of x'', for $\epsilon = 10^{-6}$. We remark that with a correction $\Delta x''$ characterized by $\Delta I_2''$ of order 10^{-13} the point $\phi^T(x' + \Delta x') + \Delta x''$ belongs to the stable manifold of x''. **On the bottom–left.** Phase plane of the map ϕ restricted to Λ for $\epsilon = 10^{-4}$: the invariant tori containing x' and x'' are represented by bold curves. We remark the presence of a large resonance between the two invariant tori. **On the bottom–right.** Projection on the space of variables $\varphi_1 - \pi, I_1, I_2 - I_2''$ of the point $\phi^T(x' + \Delta x')$ in the orbit of $x' + \Delta x'$ and of the stable manifold of x'', for $\epsilon = 10^{-4}$. We remark that with a correction $\Delta x''$ characterized by $\Delta I_2''$ of order 10^{-8} the point $\phi^T(x' + \Delta x') + \Delta x''$ belongs to the stable manifold of x''.

large variation of the actions (compared to c_k, see (3)). In particular, because Definition 2 requires the detection of one orbit returning to Λ, we needed to set the numerical precision to the high number of 400 digits, necessary to detect precisely that return. Relaxing the numerical precision of the computations (precisely we switched to double precision) we computed the statistical properties of many of these returns. We remark that, while the lower precision affects the individual integrated orbits after some Lyapunov times, it affects much less the

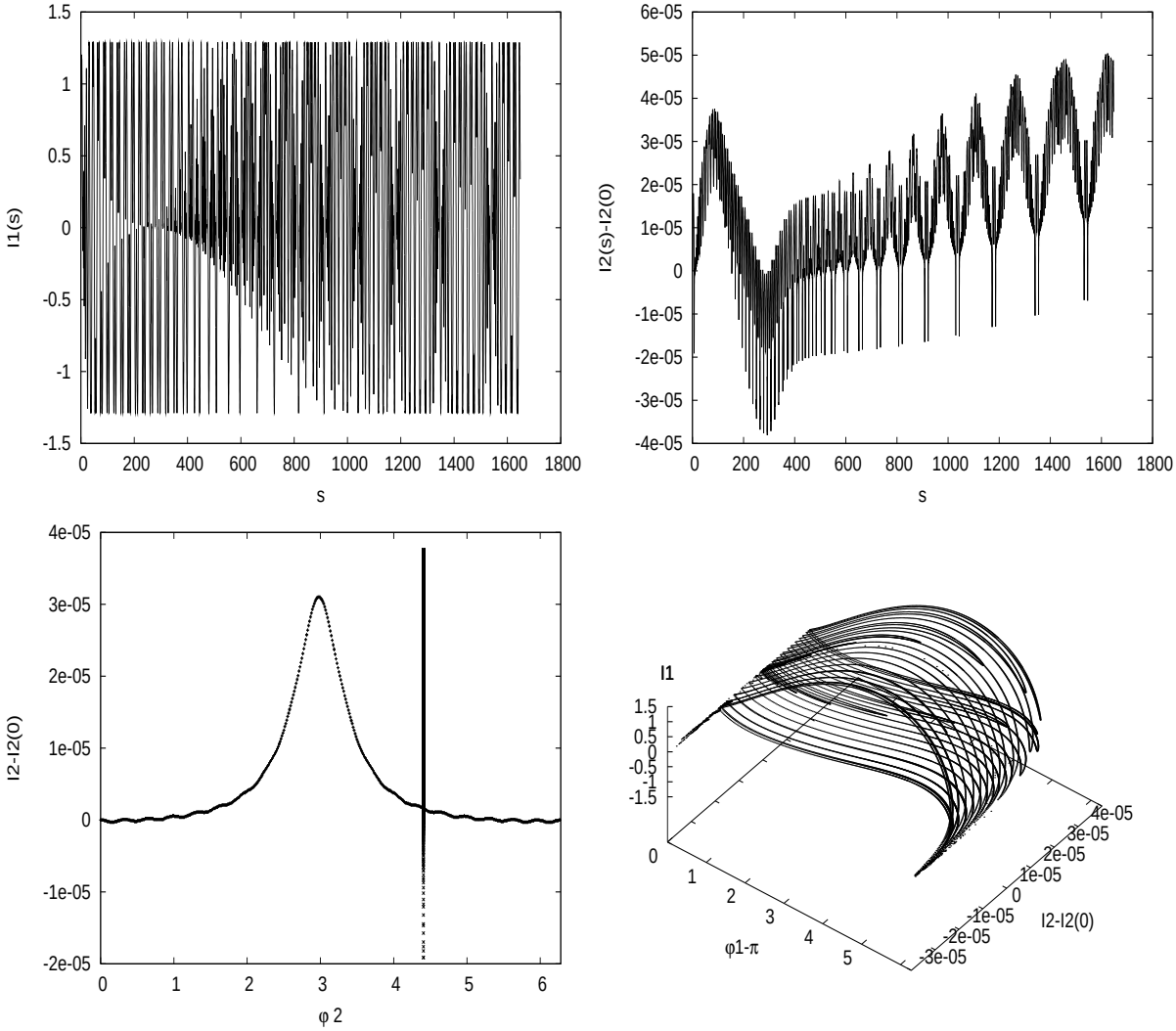

Figure 2. Computation of a parametrization of the unstable manifold of a point $x = (\varphi_1, \varphi_2, I_1, I_2) = \phi^{10^5}(\pi, 0, 0, 0.324)$ on a KAM curve, with respect to its arc–lenght s, for $\epsilon = 10^{-6}$. **On the top:** Representation of $I_1(s)$ (on the left) and $I_2(s)$ (on the right). **On the bottom left:** The orbit of $\phi_{|\Lambda}$ is on a KAM torus. The vertical segment contains the projection on the plane $(I_2 - I_2(0), \varphi_2)$ of the points of $W_u(x)$ with $|\varphi_1 - \pi| \leq 0.5$ (reducing the tolerance on φ_1 decreases the number of points on the figure, but does not decrease the amplitude of the segment). The fluctuations of $W_u(x)$ along I_2 are definitely bigger than the variation of I_2 along the torus. **On the bottom right:** Representation of the unstable manifold of x in the three dimensional space φ_1, $I_2 - I_2(0)$,I_1.

computation of statistical quantities, such as the Lyapunov exponents and the diffusion coefficients (see, for example, [23]).

Our statistical study is based on the following considerations. First, we remark that the map (7) depends periodically on all the actions. This property simplifies the definition of the diffusion process because, in principle, the actions are allowed to diffuse indefinitely on $\mathbb{R}^2$.

Then, we consider the curve $\gamma \subseteq \Lambda$ obtained by fixing all the variables except for the action I_2:

$$\gamma = \{(\varphi_1, \varphi_2, I_1, I_2) \ \text{ with } \ \varphi_1 = \pi, \varphi_2 = 0, I_1 = 0\} \ , \tag{11}$$

we choose a neighbourhood W of γ, and perform a statistical analysis on the variations of I_2 for orbits with initial conditions in W, returning to W after some time. For initial conditions $x = (\varphi_1, \varphi_2, I_1, I_2) \in W \backslash \Lambda$ we denote by $t(x)$ the return time to W and by $\psi(x) = \phi^{t(x)}(x)$ the return map to W. The set W_* on which ψ is defined can be a proper subset of W, but for the Poincaré recurrence theorem (which applies to the present case because the map ϕ is periodic with respect to the actions) it has the same Lebesgue measure as W.

Then, let us denote $I_2(x) = I_2$, and $\Delta I_2^i(x) = I_2(\psi^{i+1}(x)) - I_2(\psi^i(x))$, i.e. the action variation occurred in the i–th return. A statistical approach to the dynamics in W_*, such as the one described in [28], would be justified by the existence of a set $\tilde{W}_* \subseteq W_*$ of points x such that the sequence $\Delta I_2^1, \Delta I_2^2, ...$ is a sequence of independent random variables. This is a very strong requirement that, in our knowledge, can represent only an approximate description of the dynamics of the system. In this spirit, the traditional statistical approaches, such as for example those based on random phases approximations, replace first the true dynamics with an approximate one which behaves as a Markovian process, and then compute statistical quantities that can be defined precisely via the Markovian approximation.

Here, we proceed in a different way: instead of performing statistical approximations on the dynamics, we check that finite sets of initial conditions $x_1, \ldots, x_N$ and the finite sequence $\Delta I_2^1, ..., \Delta I_2^T$ averaged over these initial conditions behave as if the process would be approximately Markovian, i.e. we check that the variable $Y_T = \frac{\Delta I_2^1 + ... + \Delta I_2^T}{T}$ is normally distributed within a tolerance admitted for the central limit theorem convergence (see Section 5.2 for all the technical details). Then, we compute the diffusion coefficient D, of the initial conditions $x_1, \ldots, x_N$, as if the process would be a Markovian one:

$$D = \frac{1}{N} \sum_{j=1}^{N} \frac{1}{T} \sum_{i=1}^{T} \frac{\Delta I_2^i(x_j)^2}{t_i(x_j)}$$

where $t_i(x_j) = t(\psi^{i-1}(x_j))$ denotes the i–th return time of the j–th initial condition.

We remark that positive diffusion coefficients can be measured only for $\epsilon \neq 0$, because, for $\epsilon = 0$, it is $\Delta I_2^i(x) = 0$ for all i, for all $x \in W$, for any choice of W. For $\epsilon \neq 0$, the values of the diffusion coefficients depend also on the choice of W. For example, we expect that the dynamics better approximate a Markovian process by restricting the neighbourhood W.

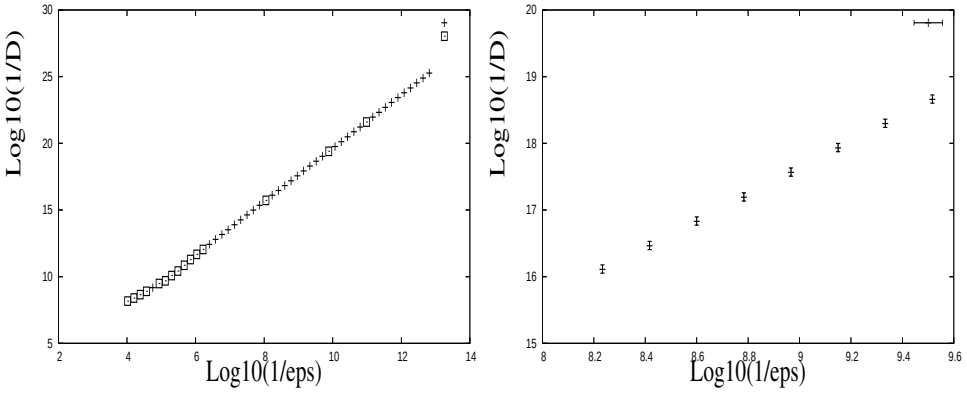

Figure 3. On the left: Computation of the diffusion coefficient for different values of $\epsilon \in (10^{-13}, 10^{-4})$ for a set of $N = 100$ initial conditions in W (the initial conditions are $I_2 = 0.324$, $I_1 \in [-10^{-5}, 10^{-5}]$, $\varphi_1 = \pi$, $\varphi_2 = 0$), using 100 return times to W. Data are very well fitted to a power law $D(\epsilon) \simeq \epsilon^2$ for $10^{-13} \leq \epsilon \leq 10^{-6}$. Data diffusing with regular statistics (precisely, satisfying (s1), (s2), see Section 5.2) are represented with a cross symbol, while the other data are represented by squares. **On the right:** Representation of a zoom of the data of the left panel with their error bars.

In Figure 3 we report the computation of the diffusion coefficient for different values of $\epsilon \in (10^{-13}, 10^{-4})$ for a set of $N = 100$ initial conditions in a set W defined by:

$$W = \{(I_1, \varphi_1, I_2, \varphi_2) : \max\{|I_1|, |\varphi_1 - \pi|, |\varphi_2|\} < 0.01\} \quad . \tag{12}$$

using $T = 100$ returns to W. We find that the values of the diffusion coefficients reported in Figure 3 are well fitted by a power law $D(\epsilon) \simeq \epsilon^2$ for $\epsilon \leq 10^{-6}$. For these small values of ϵ, the sets of integrated initial conditions behave as approximate Markovian processes (that is they satisfy conditions (s1), (s2), see Section 5.2) allowing us to compute diffusion coefficients, and the values of these diffusion coefficients satisfy a nice scaling law $D(\epsilon) \simeq \epsilon^2$.

For higher values of ϵ, that is for $\epsilon \geq 10^{-6}$, data cannot be fitted by the ϵ^2 law, and we do not try any fit because their statistics is poor (i.e. (s1), (s2) are not satisfied, see Section 5.2).

It is remarkable that the critical value $\epsilon = 10^{-6}$ is close to the value for which the error terms of Melnikov approximations (which increase with ϵ) have a sharp increment. Precisely, in Figure 4 we report the computation of a distance (defined in (28)) between the unstable manifold of a point of an invariant KAM curve and the unstable manifold computed using the Melnikov approximation (see Section 5.3 for the technical details). The distance between the two manifolds increases by two orders of magnitude between $\epsilon = 10^{-6}$ and $\epsilon = 10^{-4}$.

5. Technical tools

5.1. Normal hyperbolic invariant manifolds: numerical check of normal hyperbolicity and computation of the stable and unstable manifolds. The notion of

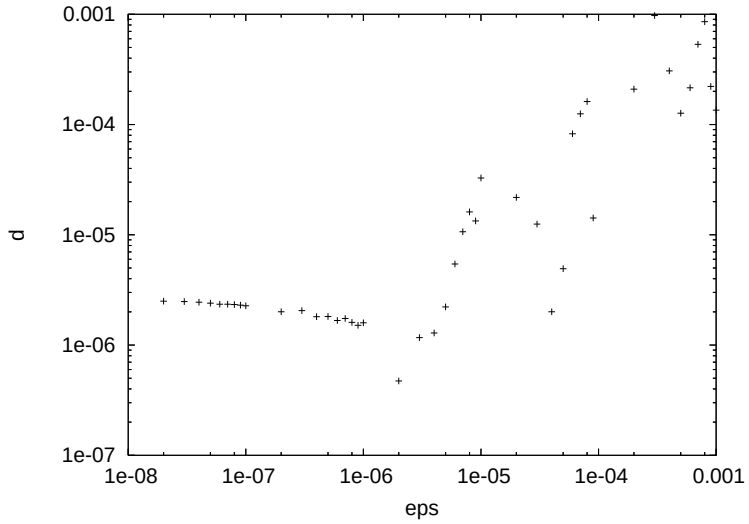

Figure 4. Computation of the d defined in (28) as a function of ϵ. The quantity d represents a distance between the unstable manifold of a point x and its Melnikov approximation, divided by ϵ. From the values reported in the figure we can appreciate that d increases by two orders of magnitude between $\epsilon = 10^{-6}$ and $\epsilon = 10^{-4}$. We also find that d decreases slowly for $\epsilon < 10^{-8}$ (for example, we measured $d \sim 10^{-7}$ for $\epsilon \sim 10^{-20}$).

normally hyperbolic invariant manifolds is extensively studied in [14], and can be stated as follows (see, for example, [14], [13]):

Definition. *Let M be a $\mathcal{C}^q$ ($q \geq 1$) compact connected manifold; let $U \subseteq M$ open and let $\phi : U \to M$ be a $\mathcal{C}^q$ embedding; let Λ be a sub-manifold of M which is invariant by ϕ. The map ϕ is said to be normally hyperbolic on Λ (Λ is also said to be normally hyperbolic invariant manifold) if there exists a Riemannian structure on M such that for any point $x \in \Lambda$ the tangent space T_xM has the following splitting:*

$$T_xM = E^s(x) \oplus T_x\Lambda \oplus E^u(x)$$

which is continuous, invariant, i.e. the linear spaces $E^s(x), E^u(x)$ are invariant by ϕ:

$$D\phi E^s(x) \subseteq E^s(\phi(x)) \quad , \quad D\phi E^u(x) \subseteq E^u(\phi(x)) \quad ,$$

and there exist constants $\lambda_1, \lambda_2, \lambda_3, \mu_1, \mu_2, \mu_3$ satisfying:

$$0 < \lambda_1 \leq \mu_1 < \lambda_2 \leq \mu_2 < \lambda_3 \leq \mu_3 \quad , \quad \mu_1 < 1 < \lambda_3 \quad , \tag{13}$$

such that:

$$\begin{aligned}
\lambda_1 \leq \inf_{\xi \in E^s(x)\setminus 0} \frac{\|D\phi(x)\xi\|}{\|\xi\|} \leq \sup_{\xi \in E^s(x)\setminus 0} \frac{\|D\phi(x)\xi\|}{\|\xi\|} \leq \mu_1 \\
\lambda_2 \leq \inf_{\xi \in T_x\Lambda\setminus 0} \frac{\|D\phi(x)\xi\|}{\|\xi\|} \leq \sup_{\xi \in T_x\Lambda\setminus 0} \frac{\|D\phi(x)\xi\|}{\|\xi\|} \leq \mu_2 \\
\lambda_3 \leq \inf_{\xi \in E^u(x)\setminus 0} \frac{\|D\phi(x)\xi\|}{\|\xi\|} \leq \sup_{\xi \in E^u(x)\setminus 0} \frac{\|D\phi(x)\xi\|}{\|\xi\|} \leq \mu_3 \quad .
\end{aligned} \tag{14}$$

Normally hyperbolic invariant manifolds have stable and unstable manifolds. Precisely, for any $x \in \Lambda$ there exist the smooth manifolds $W_s^{loc}(x), W_u^{loc}(x)$ (see

[14]) such that: $x \in W_s^{loc}(x), W_u^{loc}(x)$, $T_x W_s^{loc}(x) = E^s(x)$, $T_x W_u^{loc}(x) = E^u(x)$ and for any $n \geq 0$:

$$\begin{array}{lll} y \in W_s^{loc}(x) & \Rightarrow & d(\phi^n(x), \phi^n(y)) \leq C(\mu_1 + c)^n d(x, y) \\ y \in W_u^{loc}(x) & \Rightarrow & d(\phi^{-n}(x), \phi^{-n}(y)) \leq C(\lambda_3 - c)^{-n} d(x, y) \end{array} \tag{15}$$

with $C, c > 0$ suitable constants (c suitably small) and where $d(\cdot, \cdot)$ denotes a distance on M. The manifolds $W_s(x), W_u(x)$ are then obtained by iterating the local manifolds $W_s^{loc}(x), W_u^{loc}(x)$ with ϕ^{-1} and ϕ respectively. The local stable and unstable manifolds of Λ are defined by:

$$W_s^{loc} = \cup_{x \in \Lambda} W_s^{loc}(x) \quad , \quad W_u^{loc} = \cup_{x \in \Lambda} W_u^{loc}(x) \quad , \tag{16}$$

while the stable and unstable manifolds of Λ are:

$$W_s = \cup_{x \in \Lambda} W_s(x) \quad , \quad W_u = \cup_{x \in \Lambda} W_u(x) \quad . \tag{17}$$

Below we describe the numerical methods that we use to compute points of the stable and unstable manifolds of the map (10), adapting the method of propagation of sets commonly used for hyperbolic fixed points of two dimensional maps. A sophisticated version of this method providing high precision computations and good visualizations of pieces of the manifold can be found in [24]. Different sophisticated methods can be found in the literature for computing unstable manifolds for the higher dimensional cases (see [15] for a detailed review with applications to the visualization of two dimensional manifolds). The common point of all these methods is that the manifolds are constructed from local linear approximations (see, for example, [5]). A technique specifically adapted to compute stable manifolds of hyperbolic tori is described in [25]. A numerical study of the relation between splittings of stable and unstable manifolds and normal forms is done in [22]. The methods that we used in Section 3 adapt these known techniques (for example of [24]) to the present case, and consist in the following steps.

i) Verification that the manifold Λ is normally hyperbolic. We numerically check that the invariant manifold Λ is normally hyperbolic for $\epsilon = 0.0001$, which is the largest value of the perturbing parameter used in this paper. Precisely, we check that a compact invariant region of Λ, delimited by two invariant KAM curves containing $(I_2, \varphi_2) = (\pm 2, 0)$, is normally hyperbolic with respect to the map ϕ^N for some integer N. For each point x of a grid of initial conditions with $I_2 \in [-2, 2]$, $I_1 = 0$, $\varphi_1 = \pi$, $\varphi_2 = 0$ we first compute the Lyapunov exponents of the map ϕ (up to $N = 10^3$ iterations) for initial tangent vectors in the tangent space $T_x \Lambda^{ort}$ orthogonal to $T_x \Lambda$, i.e. for vectors of the form $\xi = (\xi_{\varphi_1}, 0, \xi_{I_1}, 0)$. We measure a positive Lyapunov exponent bigger than 0.62 for all the points of the grid, and of course a negative Lyapunov exponent smaller than -0.62. This is an indication of the hyperbolic splitting of the space $T_x \Lambda^{ort}$ as a direct sum of a stable space $E^s(x)$ and an unstable space $E^u(x)$. The numerical algorithm for the computation of the Lyapunov characteristic exponents provides also an estimate for $\lambda_1 = \mu_1$ and $\lambda_3 = \mu_3$ related to ϕ^N. It remains to estimate the constants λ_2, μ_2 for the map ϕ^N at the point x. Because in this case the growth of initial tangent vectors $\xi = (0, \xi_{\varphi_2}, 0, \xi_{I_2}) \in$

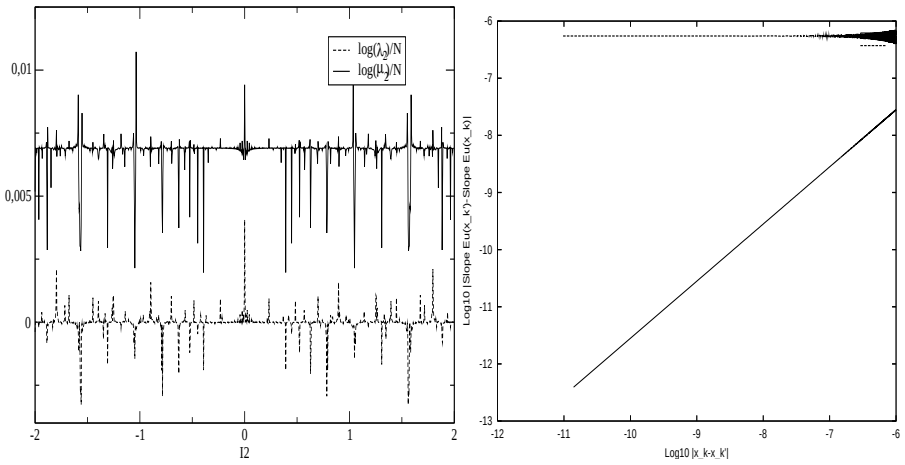

Figure 5. On the right: Numerical estimates of $\log \lambda_2/N$ and $\log \mu_2/N$, $N = 1000$, computed on a grid of 1000 initial conditions with $I_2 \in [-2, 2]$, $I_1 = 0$, $\varphi_1 = \pi$, $\varphi_2 = 0$ and $\epsilon = 10^{-4}$. **On the left:** Test of the numerical precision in the computation of $E^u(x_k)$. The figure reports on logarithmic scale the difference among the slope of $E^u(x_{k'})$ and the slope of $E^u(x_k)$ (on the y axis) versus $|x_k - x_{k'}|$ (on the x axis), for those $k' > k$ such that $|x_k - x_{k'}| \leq 10^{-6}$. The upper curve refers to the case $k = 10$, which provides poor precision of the computation (of order 10^{-6}), the lower curve refers to the case $k = 10^5$, which provides good precision (better than 10^{-12}). The data for $k = 10^5$ can be fitted by a straight line of slope 1. We can therefore infer that, within this precision, $E^u(x)$ is compatible with a Lipschitz condition in a neighbourhood of x.

$T_x\Lambda$ is not always exponential, we do not compute the Lyapunov characteristic exponents, but we computed numerically the two dimensional matrix representing the restriction of $D\phi^N(x)$ to the space $T_x\Lambda$ and the quantities:

$$\lambda_2 \leq \inf_{\xi \in T_x\Lambda \setminus 0} \frac{\|D\phi^N(x)\xi\|}{\|\xi\|} \leq \sup_{\xi \in T_x\Lambda \setminus 0} \frac{\|D\phi^N(x)\xi\|}{\|\xi\|} \leq \mu_2 \quad .$$

Figure 5 (left panel) shows the numerical computation of $\log \lambda_2/N$ and $\log \mu_2/N$ for $N = 1000$. From the comparison of the four computed quantities $\log \lambda_1$, $\log \lambda_2$, $\log \mu_2$, $\log \lambda_3$ we infer that they satisfy (13).

ii) Computation of the linear stable–unstable spaces. To compute numerical approximations of the linear space $E^u(x)$ we can now take advantage of the hyperbolicity of the dynamics. Precisely, we take a generic initial tangent vector $\xi = (\xi_{\varphi_1}, 0, \xi_{I_1}, 0) \in E^s(x) \oplus E^u(x)$ and we define the sequence:

$$\xi_k = D\phi^k(x)\xi = (\xi^k_{\varphi_1}, 0, \xi^k_{I_1}, 0) \in E^s(\phi^k(x)) \oplus E^u(\phi^k(x)) \quad .$$

The components $(\xi^k_{I_1}, \xi^k_{\varphi_1})$ do not necessarily converge to limit values, but we know from hyperbolicity that the component of ξ_k on the space $E^u(\phi^k(x))$ expands exponentially, while the component of ξ_k on the space $E^s(\phi^k(x))$ contracts exponentially. Therefore, if k is a suitably high number (compared to the exponent of the expanding direction), the direction of the unstable

space $E^u(\phi^k(x))$ is determined by ξ_k. For example, for the initial condition $(\varphi_1, \varphi_2, I_1, I_2) = (\pi, 0, 0, 0.324)$, $\epsilon = 10^{-4}$, after $k = 10^5$ iterations we obtain:

$$x_k = \phi^k(x) \sim (\pi, 4.070625, 0, 0.324319) \ , \ \ E^u(x_k) \sim < (0.652, 0, 0.75749, 0) >$$

and $x_j = \phi^j(x), E^u(x_j)$ can be easily computed for any j needed.
A test of the precision reached by these computations is done by computing $E^u(x_{k'})$ for $k' > k$ and by analyzing the variation of the slope of $E^u(x_{k'})$ as $x_{k'}$ approaches x_k. Two computations are reported in Figure 5 (right panel): one for $k = 10^5$ as above, and another one for $k = 10$. The computation for $k = 10^5$ shows that the slope of $E^u(x_{k'})$ converges to the slope of $E^u(x_k)$ as $x_{k'}$ approaches x_k. This confirms that $k = 10^5$ is sufficient to compute the unstable space with an error smaller than 10^{-12}. Moreover, because the data in the figure can be fitted by a straight line of slope 1, we can infer that $E^u(x)$ is compatible with a Lipschitz condition in a neighbourhood of x. In the figure we report for comparison the same computation for $k = 10$: in this case the slope of $E^u(x_{k'})$ does not converge to the slope of $E^u(x_k)$ as $x_{k'}$ approaches x_k, but the difference among the slopes converges to a quantity of order 10^{-6}.

iii) Computation of the stable–unstable manifolds. For any point x_j, denoting by ξ_j the unit vector generating the unstable space $E^u(x_j)$, we use the linear approximation:

$$W_u^{loc}(x_j) \sim \{x_j + s\ \xi_j \ \ , \ \ s \in [0, \rho)\} \ \ , \tag{18}$$

which is good as soon as ρ is very small (we use $\rho = 10^{-10}$ in our computations). Then, we compute finite pieces of the unstable manifold using:

$$\phi^j(W_u^{loc}(x_{-j})) \subseteq W_u(x) \ \ . \tag{19}$$

The small errors done by using the linear approximation for the local manifold do not accumulate at successive iterations, because the hyperbolic dynamics tends to reduce them (see [24]).

iv) Computation of $W_u(x'), W_s(x'')$ of Figure 1. To detect Arnold diffusion in the system (7) we compute points of $W_u(x')$ using equation (19), with $x' = \phi^{10^3}(\pi, 0, 0, 0.324)$, $\epsilon = 10^{-6}, 10^{-4}$, with the high numerical precision of 400 digits. Then, we check if some of the computed points of $W_u(x')$ are good candidates to satisfy condition (3), that is if they have a variation of the action I_2 bigger than c_2, with respect to x'. Then, we choose the correction $\Delta x''$. For both cases $\epsilon = 10^{-6}, 10^{-4}$ we found $|\Delta I_2''|$ of many orders of magnitude smaller than c_2 (see Table 1), and $|\Delta I_2'|$ are even much smaller, because the I_2 component of $E^u(x')$ is 0 and $|\Delta x'| \leq 10^{-10}$.
The error estimator ρ is computed as follows. We considered a set of 10 points in a segment of amplitude 10^{-N} aligned to $E^u(x')$, in a neighbourhood of $x' + \Delta x'$. Then, we computed the orbits of these 10 points for the number of iterations T such that $\phi^T(x' + \Delta x') + \Delta x'' \in W_s(x'')$ with a numerical precision of $2N$. We decide that N is sufficiently large, compared to T, when the map ϕ separates the 10 points by a quantity ρ which is much smaller than the precision required to verify equation (3). For example, for $\epsilon = 10^{-6}$ and $T = 1382$, we found that $\rho < 10^{-93}$ for $N = 120$, $\rho < 10^{-153}$ for $N = 180$, while $\rho = 10^{-173}$ with the actual precision of 400 digits. For $\epsilon = 10^{-4}$ and

$T = 1951$ we found that $\rho < 10^{-33}$ for $N = 120$, $\rho < 10^{-93}$ for $N = 180$, and $\rho < 10^{-113}$ for the actual precision of 400 digits.

v) Computation of the parametrization of the manifold with respect to its arc length.

To compute a parametrization of the manifold with respect to its arc length we proceed in two steps. First, we set K such that:

$$W_K(x) = \cup_{j=1}^{K} \phi^j(W_u^{loc}(x_{-j})) \subseteq W_u(x) \tag{20}$$

can be parametrized by the φ_1 coordinate, so that we can order the points in $W_K(x)$ with respect to φ_1. This allows one to construct a parametrization of $W_K(x)$ with respect to its arc–length, that we denote by:

$$s \longmapsto (\varphi_1(s), \varphi_2(s), I_1(s), I_2(s)) \quad .$$

Then, we reconstruct the unstable manifold for an arc–length much longer than the one obtained at the first step, so that to include many lobes of the manifold. This can be done by mapping with ϕ^K additional points of the linear approximation of the local manifold, but paying attention to obtain a uniform sampling of the manifold with respect to its arc–length. This problem was already discussed in [24] and we use a similar procedure for the choice of the initial conditions on $W_u^{loc}(x_{-K})$. More precisely, let us denote by x^m, x^{m+1} the last two points of $W_u^{loc}(x_{-K})$ used to compute $W_K(x)$, by $\Delta x^m = d(x^m, x^{m+1})$, and by $\Delta s^m = s^{m+1} - s^m$ the difference between the arc–lengths of the points $\phi^K(x^m)$,$\phi^K(x^{m+1})$. The choice of the point x^{m+2} will be done depending on Δs^m as follows:

$$\begin{cases} x^{m+2} = x^{m+1} + \Delta x^m & \text{if } \Delta s_1 < \Delta s^m < \Delta s \\ x^{m+2} = x^{m+1} + \eta \Delta x^m & \text{if } \Delta s^m > \Delta s \\ x^{m+2} = x^{m+1} + \frac{1}{\eta} \Delta x^m & \text{if } \Delta s^m < \Delta s_1 \end{cases} \tag{21}$$

with $\Delta s = 10^{-2}$, $\Delta s_1 = 10^{-3}$ and $\eta = 0.1$. The result of the computation is reported in Figure 2.

5.2. Computation of diffusion coefficients. In this section we describe the method that we use to estimate the diffusion coefficient related to Arnold diffusion for the map (7). We consider the curve $\gamma \subseteq \Lambda$ defined by (11), a neighbourhood W, and we perform a statistical analysis on the variations of I_2 for orbits with initial conditions in W. We define the return map to W as follows: if there exists a minimum integer $t(x) \geq 1$ such that $\phi^{t(x)-1}(x) \notin W$ and $\phi^{t(x)}(x) \in W$, we denote $\psi(x) = \phi^{t(x)}(x)$. Then, let us denote $I_2(x) = I_2$, and by $\Delta I_2^i(x) = I_2(\psi^{i+1}(x)) - I_2(\psi^i(x))$. A statistical approach to the dynamics in W_*, such as the one described in [28], would be justified by the existence of a set $\tilde{W}_* \subseteq W_*$ of points x such that the sequence $\Delta I_2^1, \Delta I_2^2, ...$ is a sequence of independent random variables. This is a very strong requirement that, in our knowledge, can represent only an approximate description of the dynamics of the system. In this spirit, the traditional statistical approaches, such as for example those based on random phases approximations, replace first the true dynamics with an approximate one which behave like a Markovian process, and then

compute statistical quantities that can be defined precisely via the Markovian approximation.

Here, we proceed in a different way: we fix a set W and then, instead of performing statistical approximations on the dynamics, we check that finite sets of initial conditions and the finite sequence $\Delta I_2^1, ..., \Delta I_2^T$ averaged over these initial conditions behave as if the process would be approximately Markovian. Because the variables $\Delta I_2^1, ..., \Delta I_2^T$ have the same mean and variance, but are not necessarily normally distributed, we check that the variable $Y_T = \frac{\Delta I_2^1 + ... + \Delta I_2^T}{T}$ is normally distributed within a tolerance admitted for the central limit theorem convergence. Precisely:

s1) denoting by $E(Y_T) = \frac{1}{N}\sum_{j=1}^{N} Y_T(x_j)$ the average of the variable Y_T over the set of N initial conditions $x_1, \ldots, x_N$, we require:

$$|E(Y_T)| \leq \frac{1}{\sqrt{N}}\sqrt{E(Y_T^2)} \quad ; \tag{22}$$

s2) the cumulative density function Φ_T of $Y_T\frac{\sqrt{T}}{\sigma}$ satisfies the Berry–Essèen inequality (see, form example, [21]):

$$\|\Phi_T(X) - \Phi(X)\| \leq C\frac{\rho}{\sigma^3\sqrt{T}} \quad , \quad \forall X \in \mathbb{R} \quad , \tag{23}$$

with $C = 0.8$, where:

$$\sigma^2 = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{T}\sum_{i=1}^{T}\Delta I_2^i(x_j)^2$$

is the mean variance of $\Delta I_2^1, ..., \Delta I_2^T$ averaged over the N initial conditions,

$$\rho = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{T}\sum_{i=1}^{T}\left|\Delta I_2^i(x_j)\right|^3 \quad ,$$

and $\Phi(x) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{X}{\sqrt{2}}\right)\right)$ is the cumulative normal distribution.

By denoting:

$$T_0 = \min_{i=1,\ldots,N}\Big(\sup_{j=1,\ldots,T} t(\psi^j(x_i))\Big) \quad , \tag{24}$$

we say that a set of N initial conditions has regular statistics in the time interval $[0, T_0]$ if conditions (s1),(s2) are satisfied.

Then, we compute the diffusion coefficient D on the set of N initial conditions $x_1, \ldots, x_N$ as if the process would be a Markovian one, as the following average:

$$D = \frac{1}{N}\sum_{j=1}^{N}\frac{1}{T}\sum_{i=1}^{T}\frac{\Delta I_2^i(x_j)^2}{t_i(x_j)}$$

where $t_i(x_j) = t(\psi^{i-1}(x_j))$ denotes the i–th return time of the j–th initial condition.

Remarks. (i) The quantity D is different from the variance σ^2 because it takes into account the individual return times $t_i(x_j)$.
(ii) In view of the central limit theorem, the diffusion coefficient and the variance of the variable Y_T are computed by averaging over the variables ΔI_2^i, while their errors are estimated as the normal errors of the normal distribution of Y. Therefore, the error on D can be estimated by $D\sqrt{2/N}$.
(iii) The results of this statistical analysis depend on the choice of W: on the one hand, we expect that the dynamics in neighbourhoods of γ better approximates a Markovian process by restricting the neighbourhood W; on the other hand, for $\epsilon = 0$ it is $X_i(x) = 0$ for all i and for all $x \in W$, for any choice of the neighbourhood W.

5.3. Melnikov approximations. The Melnikov approximations of a priori unstable systems are obtained by neglecting the perturbation on the hyperbolic part of the system, as follows:

Definition. *Let us consider the map (7), $x = (\varphi_1, \varphi_2, I_1, I_2) \in \Lambda$ and denote $J = I_2$. The Melnikov approximation of $W_u(x)$ is the unstable manifold of x with respect to the following simplified map $\tilde{\phi}$:*

$$\begin{aligned} \varphi_1' &= \varphi_1 + I_1 \quad & \varphi_2' &= \varphi_2 + J \\ I_1' &= I_1 - a\sin\varphi_1' \quad & I_2' &= I_2 + \epsilon \frac{\sin\varphi_2'}{(\cos\varphi_1' + \cos\varphi_2' + c)^2} \end{aligned} \quad . \tag{25}$$

The numerical computation of the Melnikov approximation of $W_u(x)$ is based on the following representation:

Proposition. *Let us consider $x = (\tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{I}_1, \tilde{I}_2) \in \Lambda$ and denote $J = \tilde{I}_2$. The Melnikov approximation of $W_u(x)$ is represented by all points $z = (\varphi_1, \varphi_2, I_1, I_2)$ such that (φ_1, I_1) is in the unstable manifold W_u^* of the fixed point $(\pi, 0)$ of the standard map:*

$$\varphi_1' = \varphi_1 + I_1 \quad , \quad I_1' = I_1 - a\sin\varphi_1' \quad , \tag{26}$$

while $\varphi_2 = \tilde{\varphi}_2$ and:

$$I_2 = \tilde{I}_2 - \epsilon \sum_{k=-1}^{-\infty} \Big(\frac{\sin(\tilde{\varphi}_2 - kJ)}{(\cos\varphi_1(k) + \cos(\tilde{\varphi}_2 - kJ) + c)^2} - \frac{\sin(\tilde{\varphi}_2 - kJ)}{(\cos(\tilde{\varphi}_2 - kJ) + c - 1)^2} \Big) \tag{27}$$

where $(\varphi_1(j), I_1(j))$ denote the orbit with initial condition $(\varphi_1, I_1) \in W_u^*$ with respect to the map (26).

The proof of this proposition is reported at the end of this section. In Figure 6 we compare two parametrizations $s \mapsto (I_2(s) - I_2(0))$ of the manifold $W_u(x)$: one is obtained with the Melnikov approximation (27), while the other one is obtained using the full map (7). The left panel shows that for $\epsilon = 10^{-6}$ the two parametrizations are indeed very close to one another. The right panel shows that for $\epsilon = 10^{-4}$ the Melnikov approximation is not valid at all. In order to quantify the relevance of the error terms of the Melnikov approximation we have computed for $10^{-8} < \epsilon < 10^{-3}$ the histograms H_f and H_M of $(I_2(s) - I_2(0))/\epsilon$

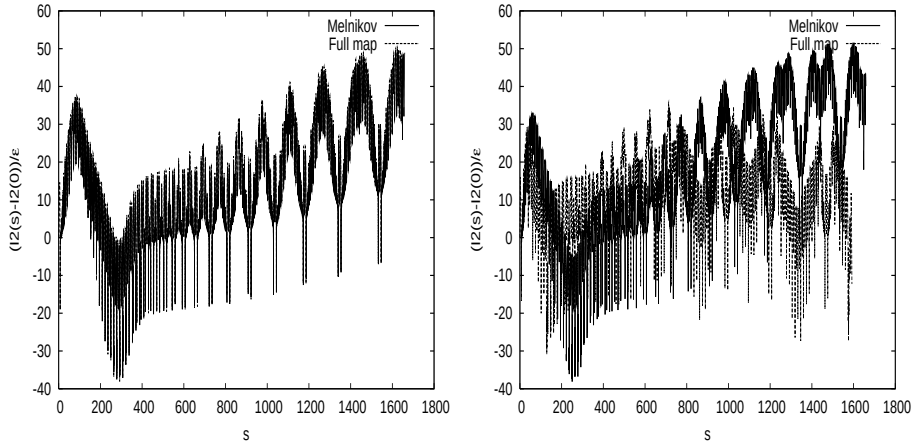

Figure 6. Each panel represents two parametrizations $s \mapsto (I_2(s) - I_2(0))/\epsilon$ of the manifold $W_u(x)$, with $x = \phi^{10^5}(\pi, 0, 0, 0.324)$: one is obtained using the Melnikov approximation, while the other one is obtained using the full map. The left panel is for $\epsilon = 10^{-6}$: the two parametrizations are close one to the other. The right panel is for $\epsilon = 10^{-4}$: the Melnikov approximation is not valid.

for the full map and the Melnikov approximation respectively. We consider as an indicator of the distance between the two distributions the quantity:

$$d = \frac{\sum_{i=1}^{N}(H_f(i) - H_M(i))^2}{N} \tag{28}$$

where $N = 100$ is the number of bins. The quantity d (Figure 4) increases by two orders of magnitude between $\epsilon = 10^{-6}$ and $\epsilon = 10^{-4}$, while it slowly decreases for $\epsilon < 10^{-8}$ (not reported in Figure 4).

Proof of the Proposition. Let us denote by $z(j) = (\varphi_1(j), \varphi_2(j), I_1(j), I_2(j))$ the orbit of $z = z(0) = (\varphi_1, \varphi_2, I_1, I_2)$ and by $x(j) = (\tilde{\varphi}_1(j), \tilde{\varphi}_2(j), \tilde{I}_1(j), \tilde{I}_2(j))$ the orbit of $x = \tilde{x}(0) = (\tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{I}_1, \tilde{I}_2)$ with respect to the map $\tilde{\phi}$. The point z is in the unstable manifold of x if and only if it is: $\lim_{j\to-\infty} \|z(j) - x(j)\| = 0$. Therefore, $(I_1(j), \varphi_1(j))$ tends to $(0, \pi)$ as $j \to -\infty$ if and only if $(I_1(0), \varphi_1(0))$ is in the unstable manifold W_u^* of the fixed point $(\pi, 0)$ of the map (26). Let us now prove (27). For any $j \leq -1$ it holds:

$$\begin{aligned} I_2(j) &= \sum_{k=-1}^{j} (I_2(k) - I_2(k+1)) + I_2(0) \\ &= \epsilon \sum_{k=-1}^{j} \frac{\sin\varphi_2(k+1)}{(\cos\varphi_1(k+1) + \cos\varphi_2(k+1) + c)^2} + I_2 \quad , \end{aligned} \tag{29}$$

as well as:

$$\tilde{I}_2(j) = \sum_{k=-1}^{j} (\tilde{I}_2(k) - \tilde{I}_2(k+1)) + \tilde{I}_2(0) = \epsilon \sum_{k=-1}^{j} \frac{\sin\tilde{\varphi}_2(k+1)}{(\cos\tilde{\varphi}_2(k+1) + c - 1)^2} + \tilde{I}_2 \quad .$$

Therefore, $\lim_{j\to-\infty} \|I_2(j) - \tilde{I}_2(j)\| = 0$ if and only if (27) holds.

6. Conclusions

We have studied the Arnold diffusion along a normally hyperbolic manifold in a model of a priori unstable dynamical systems. We have introduced a definition of Arnold diffusion which is adapted to the numerical investigation of the problem, and is based on the numerical computation of the stable and unstable manifolds of the system. We have shown that the numerically computed stable and unstable manifolds indeed support this kind of Arnold diffusion. We also performed a numerical statistical study of Arnold diffusion, and we found that, for small values of ϵ, Arnold diffusion behaves as an approximate Markovian process, allowing one to compute diffusion coefficients. The dependence of the diffusion coefficient D on the perturbing parameter satisfies the scaling $D(\epsilon) \simeq \epsilon^2$ for small values of ϵ. We also find that this law is correlated to the validity of the Melnikov approximation, in the sense that it is valid up to the same critical value of ϵ for which the error terms of Melnikov approximations have a sharp increment. This suggests that the Melnikov approximation is not only a technical tool which allows one to compute accurate approximations of the manifolds at small values of the perturbing parameters, but is related to a dynamical regime, and possibly it could be used to explain the statistical properties of Arnold diffusion.

Acknowledgments. M. Guzzo acknowledges the project CPDA063945/06 of the University of Padova.

References

1. Arnold V.I.: Instability of dynamical systems with several degrees of freedom. Sov. Math. Dokl., 6, 581–585, 1964.
2. Berti M., Biasco L. and Bolle P.: Drift in phase space: a new variational mechanism with optimal diffusion time. J. Math. Pures Appl., Vol. 9, 82, no. 6, 613–664, 2003.
3. Berti M. and Bolle P.: A functional analysis approach to Arnold diffusion. Annales de lInstitut Henri Poincare (C) Non Linear Analysis, Vol. 19, 4, 395-450, 2002.
4. Bessi U., Chierchia L. and Valdinoci E.: Upper bounds on Arnold diffusion times via Mather theory. J. Math. Pures Appl., Vol. 80, 105–129, 2001.
5. Broer H.W., Osinga H.M. and Vegter G.: Algorithms for computing normally hyperbolic invariant manifolds. ZAMP, 48, 480-524, 1997.
6. Chierchia L. and Gallavotti G.: Drift and diffusion in phase space. Ann. Inst. H.Poincaré, Vol. 60, 1–144, 1994.
7. Chirikov B.V.: Research concerning the theory of nonlinear resonance and stochasticity", Preprint N 267, Institute of Nuclear Physics, Novosibirsk (1969) Engl. Trans., CERN Trans. 71-40 (1971).
8. Delshams A., de la Llave R. and Seara T.M.: A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model. Mem. Amer. Math. Soc. 179, no. 844, 2006.
9. Efthymiopoulos, C.; Voglis and Contopoulos, G.: Diffusion and Transient Spectra in 4-dimensional symplectic mapping. *In Analysis and Modelling of discrete dynamical systems*, D. Benest and C. Froeschlé eds. Gordon and Breach Science Publishers, 1998.
10. Froeschlé C., Guzzo M. and Lega E.: "Local and global diffusion along resonant lines in discrete quasi–integrable dynamical systems", Celestial Mechanics and Dynamical Astronomy, vol. 92, n. 1-3, 243-255, 2005.

11. Guzzo M., Lega E. and Froeschlé C.: "First Numerical Evidence of Arnold diffusion in quasi–integrable systems". DCDS B, vol. 5, n. 3, 2005.
12. Hénon M. and Heiles C.: The Applicability of the Third Integral of Motion: Some Numerical Experiments. The Astronomical Journal, 69, p. 73–79, (1964).
13. Hasselblatt B. and Pesin Y.: Partially hyperbolic dynamical systems. Handbook of dynamical systems. Vol. 1B, 1–55, Elsevier B. V., Amsterdam, 2006.
14. Hirsch M.W., Pugh C.C. and Shub M.: Invariant Manifolds. Lecture Notes in Mathematics, Vol. 583. Springer-Verlag, Berlin-New York, 1977.
15. Krauskopf, B., Osinga, H.M., Doedel, E.J., Henderson, enheimer, J., Vladimirsky, A., Dellnitz, M., Junge, O.: A survey of methods for computing (un)stable manifolds of vector fields. Int. J. Bif. and Chaos, 15, 763-791, 2005.
16. Konishi T. and Kaneko K.: Diffusion in Hamiltonian chaos and its size dependence. J. Phys. A: Math. Gen., 23, 15, L715-L720.
17. Laskar J.: Frequency analysis for multi-dimensional systems. Global dynamics and diffusion. Physica D, 67, pp 257–281, (1993).
18. Lega E., Guzzo M. and Froeschlé C.: "Detection of Arnold diffusion in Hamiltonian systems". Physica D, vol. 182, p. 179–187, 2003.
19. Lega E., Froeschlé C. and Guzzo M.: "Diffusion in Hamiltonian quasi–integrable systems." In Lecture Notes in Physics 729, "Topics in gravitational dynamics", Benest, Froeschlé, Lega eds., Spinger, 2007.
20. Lichtemberg A. and Aswani M.A.: Arnold diffusion in many weakly coupled mappings. Phys. Rev. E, 57, 5, 5325–5321, 1998.
21. Manoukian, E.: Modern Concepts and Theorems of Mathematical Statistics. Springer, 1986.
22. Morbidelli A. and Giorgilli A.: On the role of high order resonances in normal forms and in separatrix splitting. Physica D, vol. 102, 195–207, 1997.
23. Ralston, A. and Rabinowitz, P. A First Course in Numerical Analysis (2nd ed.) McGraw-Hill, 1978.
24. Simó C.: On the analytical and numerical approximation of invariant manifolds, in Modern Methods in Celestial Mechanics, D. Benest, Cl. Froeschlé eds, Editions Frontières, 285-329, 1989.
25. Simó C. and Valls C.: A formal approximation of the splitting of separatrices in the classical Arnold's example of diffusion with two equal parameters. Nonlinearity 14, no. 6, 1707–1760, 2001.
26. Treschev D.: Trajectories in a neighbourhood of asymptotic surfaces of a priori unstable Hamiltonian systems. Nonlinearity 15 2033-2052, 2002.
27. Treschev D.: Evolution of slow variables in a priori unstable Hamiltonian systems. Nonlinearity 17 1803-1841, 2004.
28. Varvoglis H.: "Chaos, random walks and diffusion in Hamiltonian systems." In Hamiltonian systems and Fourier Analysis, 247–287. Benest, Froeschlé and Lega editors. Cambridge Scientific Publishers, 2005.
29. Wood B.P., Lichtenberg A. and Lieberman M.A.: Arnold diffusion in weakly coupled standard map. Phys. Rev. A, 42, 5885-5893, (1990).