Semi-analytic computations of the speed of Arnold diffusion along single resonances in a priori stable Hamiltonian systems

Massimiliano Guzzo
Università degli Studi di Padova
Dipartimento di Matematica 'Tullio Levi-Civita'
Via Trieste 63 - 35121 Padova, Italy

Christos Efthymiopoulos
Research Center for Astronomy
and Applied Mathematics
Academy of Athens
Soranou Efessiou 4, 11527 Athens, Greece

Rocío I. Paez
Università degli Studi di Padova
Dipartimento di Matematica 'Tullio Levi-Civita'
Via Trieste 63 - 35121 Padova, Italy

Abstract

Cornerstone models of Physics, from the semi-classical mechanics in atomic and molecular physics to planetary systems, are represented by quasi-integrable Hamiltonian systems. Since Arnold's example, the long-term diffusion in Hamiltonian systems with more than two degrees of freedom has been represented as a slow diffusion within the 'Arnold web', an intricate web formed by chaotic trajectories. With modern computers it became possible to perform numerical integrations which reveal this phenomenon for moderately small perturbations. Here we provide a semi-analytic model which predicts the extremely slow-time evolution of the action variables along the resonances of multiplicity one. We base our model on two concepts: (i) By considering a (quasi-)stationary phase approach to the analysis of the Nekhoroshev normal form, we demonstrate that only a small fraction of the terms of the associated optimal remainder provide meaningful contributions to the evolution of the action variables. (ii) We provide rigorous analytical approximations to the Melnikov integrals of terms with stationary or quasi-stationary phase. Applying our model to an example of three degrees of freedom steep Hamiltonian provides the speed of Arnold diffusion, as well as a precise representation of the evolution of the action variables, in very good agreement (over several orders of magnitude) with the numerically computed one.

1 Introduction

Fundamental problems of Physics are often modeled with small Hamiltonian perturbations of integrable systems. For example, the problem of the stability and long–term evolution of the Solar System can be modeled in terms of perturbations to Kepler's motion of each planet under the gravity of the Sun.

This is a post-peer-review, pre-copyedit version of an article published in Journal of Nonlinear Science. The final authenticated version is available online at: http://dx.doi.org/10.1007/s00332-019-09594-9.

Similar perturbative approaches are employed in some of the most classical problems of Mechanics appearing from the microscopic scale (e.g. the semi-classical treatment of atomic and molecular dynamics) up to the astronomical one (e.g. solar systems and galaxies). The above and other important applications, as e.g. in plasma and accelerator physics, or statistical mechanics, have rendered Hamiltonian near-integrable systems a fundamental topic in physics (see [56] for a collection of basic papers and reviews in this field).

One of the most interesting questions in near-integrable systems is the long–term fate of trajectories which belong to the so-called 'Arnold web'. Following the pioneering work of V.I. Arnold [1], the Arnold web is understood as an intricate in shape and connected set in phase–space which contains chaotic trajectories. The Arnold web is tightly related to the existence, in phase–space, of a corresponding 'web of resonances', i.e., domains where the trajectories undergo near-oscillatory motions with a commensurable set of frequencies. 'Arnold diffusion' is a theoretically predicted phenomenon, according to which a trajectory with initial conditions within the Arnold web undergoes slow chaotic diffusion. When the number n of the degrees of freedom is equal to 3 or larger, such diffusion renders possible, in principle, to connect every part of the Arnold web within sufficiently long times. Let us therefore consider a n-degree of freedom Hamiltonian of the form:

$$H_\varepsilon(I,\varphi) = H_0(I) + \varepsilon f(I,\varphi) \tag{1}$$

where $(I,\varphi) \in \mathcal{G} \times \mathbb{T}^n$ are action-angle variables, $\mathcal{G} \subseteq \mathbb{R}^n$ is open bounded, the integrable approximation H_0 and the perturbation f are real analytic, ε is a small parameter. The problem we address in this paper is the following:

PROBLEM 1: For given H_0, f, small $\varepsilon > 0$, and $I_* \in \mathcal{G}$ such that $\ell \cdot \nabla H_0(I_*) = 0$ for a unique $\ell \in \mathbb{Z}^n \backslash 0$ (with its multiples), provide a formula which gives the maximum speed of the drift along the resonance $\ell \cdot \nabla H_0(I) = 0$ (averaged on time intervals T longer than $1/\varepsilon$) among all the solutions of Hamilton's equations with initial conditions $I(0), \varphi(0)$ with $\varphi(0) \in \mathbb{T}^n$ and $I(0)$ in the ball $B(I_*, C\sqrt{\varepsilon}) \subseteq \mathcal{G}$ of center I_* and radius $C\sqrt{\varepsilon}$, with some $C > 0$.

Remarks:

(i) For special choices of H_0, f, and of the resonant vector $\ell \in \mathbb{Z}^n \backslash 0$, the previous problem has a simple solution. In fact, Nekhoroshev provided a class of quasi-integrable Hamiltonian systems with variations of the actions of order 1 already on times of order $1/\varepsilon$ which can be explicitly computed with a simple quadrature.

(ii) For $n = 2$, instability on times of order $1/\varepsilon$ has been proved for a class of non-convex Hamiltonians H_0 [11]. If instead H_0 is iso–energetically non–degenerate, the KAM theorem provides a topological obstruction to the drift along the resonances of the system, so Problem 1 is not interesting in this case.

(iii) For H_0 satisfying a transversality condition, called by Nekhoroshev "steepness", the stability time of the action variables improves dramatically to an exponential order in $1/\varepsilon$ [57, 58]: precisely, there exist positive constants a, b and ε_0 such that for any $0 \le \varepsilon < \varepsilon_0$ the solutions $(I(t), \varphi(t))$ of the Hamilton equations of $H_\varepsilon(I,\varphi)$ satisfy

$$|I(t) - I(0)| \le \varepsilon^b \;\; \text{for} \;\; |t| \;\le\; T_N := \frac{1}{\varepsilon}\, \exp\Big(\frac{1}{\varepsilon^a}\Big). \tag{2}$$

According to Nekhoroshev's theorem, any large drift of the action variables needs time intervals longer than the exponentially long time T_N. For systems with $n \geq 3$ satisfying the hypotheses of Nekhoroshev's theorem, proving the existence of orbits with variations of the actions of order 1 in some suitable long time for any small value of $|\varepsilon|$, is highly non trivial, and these are the conditions under which Problem 1 is interesting and, up to now, unsolved.

(iv) We do not provide here a rigorous solution to Problem 1, but we provide formulas which match the very slow drifts observed in numerical experiments. These formulas are obtained by combining: a semi-analytic argument including the computer assisted computation of normal forms, whose coefficients are provided in floating point arithmetics; a rigorous approximation of the Melnikov integrals using methods of asymptotic analysis based on the so called stationary-phase approximation; a random–phase assumption used in the Melnikov approximation. We call our approach 'semi-analytical', since it combines rigorous results (in the stationary-phase method, see Section 3) with ones based on the numerical (computer-assisted) computations of the Nekhoroshev normal forms. Whether these ideas can be transformed into a fully rigorous argument is a question beyond the purpose of this paper.

(v) There is a range of values for the constant C appearing in Problem 1 on which our resuls apply. In fact, with mild hypotheses on H_0 and f, by neglecting the exponentially small remainder in the Nekhoroshev normal form adapted to the resonance $\ell \cdot \nabla H_0(I) = 0$, one remains with an integrable Hamiltonian with separatrix loops of amplitude ΔI in the action variables proportional to $\sqrt{\varepsilon}$. The constant C must be chosen so that $C\sqrt{\varepsilon} > \Delta I$ (see Section 2 for all the details, and Eq. (23) for the explicit choice of C).

(vi) The estimated speed of drift along a resonance expected from the solution of Problem 1 should depend on ε, I_* and ℓ. Of course, close to I_* the resonance $\ell \cdot \nabla H_0(I) = 0$ may intersect an infinite number of other resonances $\tilde{\ell} \cdot \nabla H_0(I) = 0$ with $\tilde{\ell} \in \mathbb{Z}^n \backslash 0$ independent on ℓ, and solutions with initial conditions close to I_* may leave the resonance $\ell \cdot \nabla H_0(I) = 0$ and drift along different resonances, possibly of different multiplicities. Problem 1 concerns only the orbits which drift along the fixed resonance $\ell \cdot \nabla H_0(I) = 0$.

(vii) For systems of $n = 3$ degrees of freedom a solution of Problem 1 gives the opportunity to compare the time needed to diffuse along the resonances of multiplicity 1 with the stability time T_N of the Nekhoroshev theorem. In fact for $n = 3$ distant points of the action–space on the same energy level are connected through paths of the Arnold web which are mostly contained in resonances of multiplicity one, where Problem 1 is applicable. The resonances of multiplicity two are just at the points of intersection of the resonances of multiplicity one. The transit of the orbits through resonances of multiplicity two, the so–called 'large gap problem', is one of the hardest theoretical difficulties in rigorously proving the existence of Arnold diffusion. Numerical studies, instead, provide overwhelming evidence for the existence of such transits, see [45, 46], while the key question regarding the quantification of Arnold diffusion is Problem 1.

(viii) Throughout this paper, and mostly in Section 5, we compare our semi-analytic solution of Problem 1 with numerical experiments. The long-term behaviour of Hamiltonian systems, including Arnold diffusion, can be numerically investigated with symplectic integrators (see [5, 59, 34]). In fact, depending on the order of the integration scheme, every step ϕ^τ of the integrator is exponentially close, with exponential factor $-1/\tau$, to the exact Hamiltonian flow of a modified Hamiltonian

$$K_\varepsilon(I,\varphi) = H_\varepsilon(I,\varphi) + \tau^\nu W(I,\varphi;\varepsilon) \ ,$$

where the integer ν and W both depend on the integration scheme. Therefore, for suitably small τ the spurious term $\tau^\nu W(I,\varphi;\varepsilon)$ is just a perturbation of the original Hamiltonian, and the exponential factor becomes negligible with respect to any observed diffusion.

(ix) While the KAM and Nekhoroshev theorems, as well as the examples of Arnold diffusion (and also Problem 1), are usually formulated for quasi-integrable Hamiltonians (1), many quasi-integrable systems of interest for Physics and Celestial Mechanics are characterized by degeneracies and singularities of the action variables which introduce additional complications. Proofs of the KAM and Nekhoroshev theorems, as well as of the existence of hyperbolic tori, have been provided also for many cases of interest for Celestial Mechanics, see for example, [16, 3, 37, 39, 17, 60, 24].

Problem 1 is an applicative spin-off of the problem of Arnold diffusion, which started with the fundamental paper published by Arnold in 1964 [1], first providing a quasi–integrable Hamiltonian system with non trivial long–term instability. Since Arnold's pioneering paper, a rich literature has appeared on attempts to prove of existence of Arnold diffusion for more general quasi–integrable Hamiltonian systems, called, in the context of Arnold diffusion, *a priori stable* systems. A simpler, albeit still highly non trivial, case if the one of *a priori unstable* systems. In the latter case, the existence of diffusing motions has been proved using different models and techniques, including Mather's variational methods, geometrical methods and the so-called separatrix and scattering maps (among the rich literature see [16, 7, 21, 64, 19, 6, 35, 49, 48, 22, 12] and references therein).

Due to the long timescales involved, also numerical or experimental observations of Arnold diffusion are hard to achieve. Already few years after the first numerical detection of chaotic motions [47], the long-term instability in Hamiltonian systems was discussed from both an analytical and numerical point of view in [18]. However, only modern computers rendered possible to simulate the phenomenon in simple physical models. In the last decades, diffusion through the resonances has been clearly identified [51, 25, 50, 36, 52, 45, 32, 42, 36, 27, 28, 63, 43, 44]. Then, in the series of papers [52, 45, 32, 42, 44], diffusion of orbits was detected for values of the perturbation parameters so small that the set of resonant motions has the structure of the Arnold web embedded in a large volume of invariant tori (the distributions of resonances and tori being computed numerically with chaos indicators [31, 52, 43]). In these experiments, the instability was characterized by diffusion coefficients decreasing faster than power laws in ε, compatibly with the exponential stability result of Nekhoroshev's theorem. This was confirmed by a direct comparison of the numerical diffusion coefficient with the size of the optimal remainder of the Nekhoroshev normal form in [27].

In this paper we propose a semi-analytic solution to Problem 1 which is obtained through the following steps:

(a) Given ε and I_* we construct a computer assisted normal form adapted to the local resonance properties at I_* up to an optimal normalization order, by following the construction of normal forms which appears within the proof of the Nekhoroshev theorem. The computer

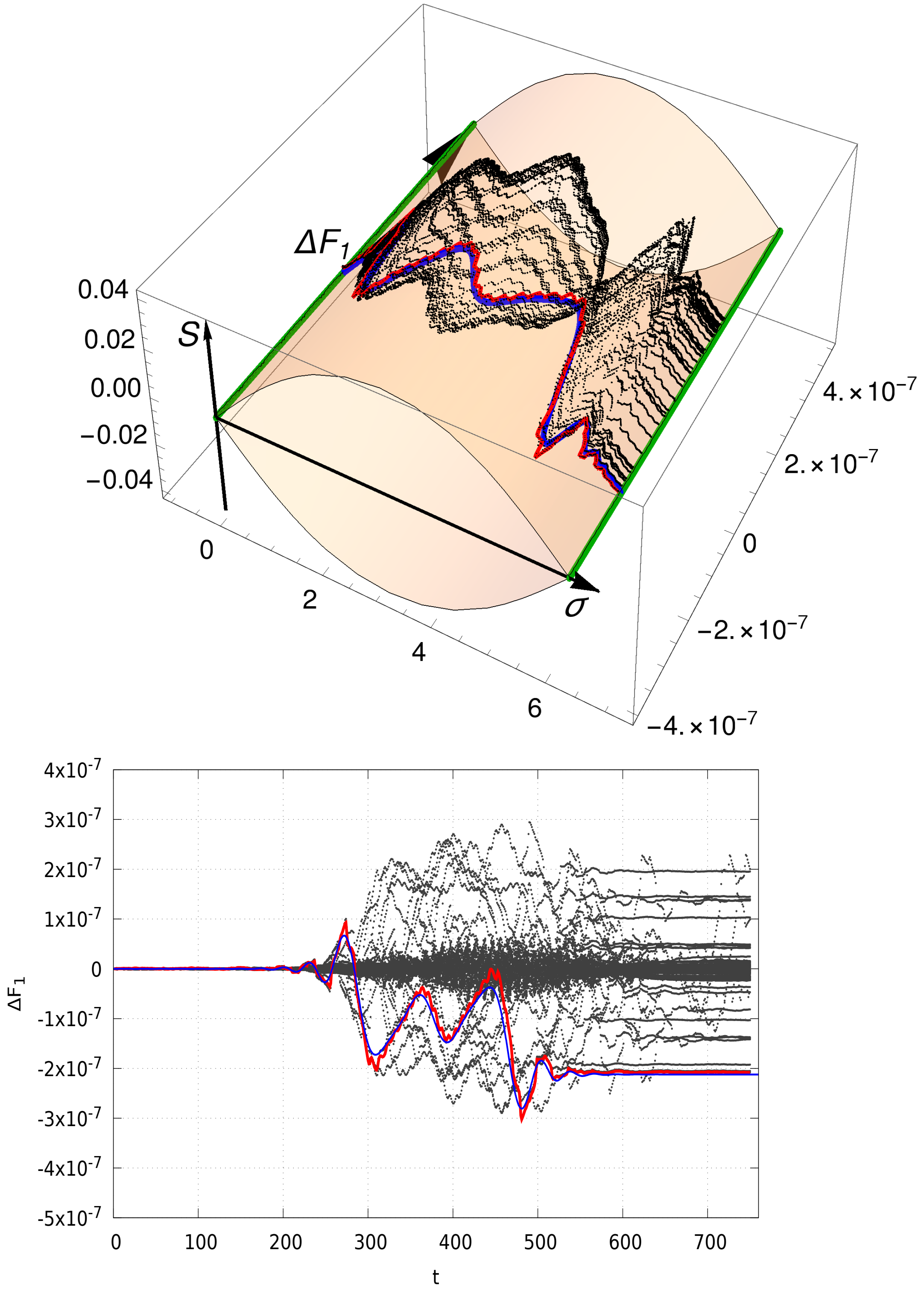

Figure 1: In the top panel we report the projection of a swarm of 100 orbits in the resonance defined by $\ell = (1, 1, 0)$ on the space of the normal form variables $\sigma, \hat{F}_1, \hat{S}$, during a full circulation of the resonant angle σ, for $\varepsilon = 0.01$ in the numerical experiments conducted with the Hamiltonian (3) (see text). The 100 initial conditions have been chosen in a small neighbourhood of the stable/unstable manifolds (represented as shaded surfaces in the picture) of a family of 2-dimensional normally hyperbolic tori of an approximated normal form Hamiltonian (which projects on the green lines in the picture), parameterized by the angles ϕ_1, ϕ_2 (not included in the picture). The 100 orbits are represented in black; the red curve represents the orbit of the swarm leading to the maximum variation $\Delta\hat{F}_1$ after a full circulation of σ; the blue curve represent the analytic computation of the same orbit with the theory presented in Sections 2, 3 and 4. Apart from small oscillations, we have a very good agreement between the numerically computed orbit and the prediction of the analytic computation. In the bottom panel we represent the time variation of the action $\hat{F}_1$ for the same orbits.

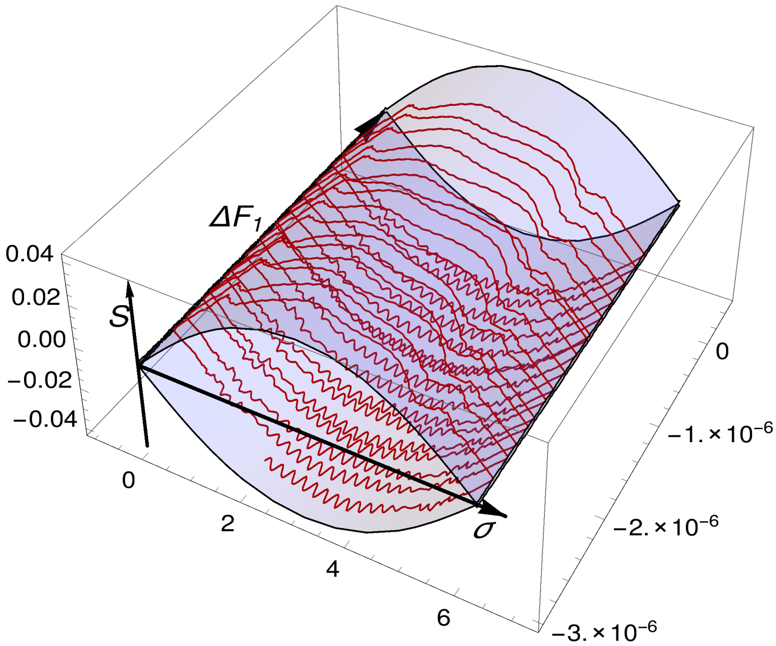

Figure 2: Projection on the space of the normal form variables $\sigma, \hat{F}_1, \hat{S}$ of an orbit leading to a maximum variation $\Delta\hat{F}_1$ comparable with the values predicted by the semi-analytic theory presented in this paper (and reported in Figure 3 below) through a sequence of several circulations/librations of the resonant angle σ. The initial condition is in the resonance defined by $\ell = (1, 1, 0)$ and the value of the perturbation parameter is $\varepsilon = 0.01$.

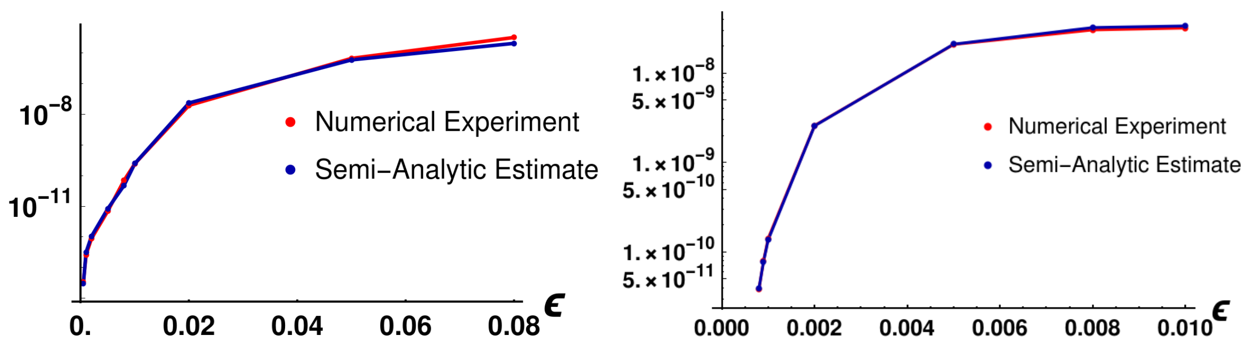

Figure 3: Comparison of the maximum speed of Arnold diffusion along the resonances $\ell = (1, 1, 0)$ (left panel) and $\ell = (1, 3, 0)$ (right panel) measured from the numerical integrations of different initial conditions (red dots) with the values obtained using the semi-analytic theory developed in this paper (the blue dots are correspond to the ratios of $|\Delta F_1|_P$ or $|\Delta F_1|_{NP}$ and T_α reported in Tables 1 and 2) computed for several values of ε.

assisted construction of a normal form is mandatory, since our purpose is to compare the predicted values of the drifts with the numerically observed ones. As it is well known (see [13, 14] for the KAM theorem, and [15, 39] for the Nekhoroshev theorem) purely analytic estimates which do not use computer assisted methods are highly unrealistic.

(b) We represent the variation of the actions along the resonance with Melnikov integrals defined from the normal form constructed as indicated in (a). For $n = 3$ degrees of freedom Hamiltonians the remainders of the normal forms are represented as expansions of already millions of very small terms. Hence, the variation of the actions is represented as a sum of millions of Melnikov integrals, which have to be computed in order to solve Problem 1. The problem becomes prohibitive if the goal is to maximize the result with respect to some variables in order to compute the orbits with largest instability. To overcome this difficulty, we require an analytic method that allows to distinguish between the terms of the remainder associated with large contribution to the variations of the actions and those of negligible contribution, therefore reducing the total amount of terms to consider.

(c) We represent the Melnikov integrals with a method from asymptotic analysis, the so-called *method of the stationary-phase* (see [10]), recently used also in the related context of the computation of the splitting of separatrices [29]. In fact, for quasi-integrable systems, the Melnikov integrals can be reformulated as integrals with a rapidly oscillating phase, and the computation of the critical points of this phase provides an estimate of the integral. We find that only the Melnikov integrals whose phase either 1) has critical points, or 2) has a suitably small derivative with respect to the slow angle variable of the resonance is suitably small, provide major contributions to the Arnold diffusion. We call the corresponding terms in the remainder stationary or quasi-stationary, respectively. The Melnikov integrals whose phase is neither stationary nor quasi-stationary represent the large majority of terms, and their cumulative contribution to the Arnold diffusion is negligible with respect to the cumulative contribution of the stationary or quasi-stationary terms. Therefore, we provide a rigorous criterion to select, from the millions of harmonics of the remainder of the Nekhoroshev normal form, a few thousand ones. All the relevant integrals of Melnikov theory can be explicitly represented with an asymptotic formula or directly computed by quadratures. The asymptotic formula, providing the variation of the actions during a resonant libration, depends on the initial phases $\varphi(0)$. For all possible values of these phases the formula represents closely the spread of the actions which is observed with numerical integrations.

(d) Finally, by maximizing with respect to $\varphi(0)$ the variations of the actions obtained from the Melnikov integrals we obtain the orbits with largest variation of the actions at each homoclinic loop, as well as the rare initial conditions whose orbit, in a sequence of homoclinic loops, have a systematic variation of the action variables. Thus we predict which orbits undergo the 'fastest' Arnold diffusion which we can observe.

For simplicity, we develop our theory for the resonances defined by integer vectors[1] $\ell \in \mathbb{Z}^N \backslash 0$ such that, by denoting with

$$f(I,\varphi) = \sum_{K\in\mathbb{Z}^n} f_K(I) e^{iK\cdot\varphi}$$

the complex Fourier expansion of f with respect to the angles $\varphi_1, \ldots, \varphi_n$, the periodic func-

[1] The vector ℓ must be chosen so that the integer lattice Λ generated by ℓ is not properly contained in any lattice of the same dimension.

tion

$$v_\ell(\sigma) = \sum_{\nu \in \mathbb{Z}} f_{\nu\ell}(I_*) e^{i\nu\sigma}$$

for the considered I_* has a unique non-degenerate local maximum and a unique non-degenerate local minimum with respect to the slow angle of the resonance $\sigma = \ell \cdot \varphi$. If this hypothesis is not satisfied, for example because $v_\ell(\sigma)$ has more than one local maximum, then one has to perform a straightforward reformulation of the theory presented in this paper; if instead $v'_\ell(\sigma) = 0$ for all σ, then our method needs major modifications.

From steps (a), (b), (c), (d) above we have a qualitative and quantitative description of the drift along the resonances of multiplicity one. The qualitative picture of the diffusion is in agreement with the idea having its roots in Chirikov's fundamental paper [18] and recently recovered e.g. in [20], namely that the diffusion along a resonance is not uniform in time, but it is produced by impulsive 'kicks' or 'jumps' at every homoclinic loop, see [2, 23, 62]. The new quantitative analysis allows us to determine the frequency of occurrence and amplitude of these jumps as the resonant angle becomes critical for some Melnikov integrals; also, we are able to select the initial conditions whose orbits have the fastest Arnold diffusion. Therefore, for given values of ε, we are able to predict the minimum timescales needed to observe long–term diffusion along any single resonance.

We finally provide numerical demonstrations of this theory. To this end, we compute the diffusion in the resonances of multiplicity 1 for the 3-degrees of freedom Hamiltonian (introduced following [42, 63])

$$H_\varepsilon = \frac{I_1^2}{2} - \frac{I_2^2}{2} + \frac{I_2^3}{3\pi} + 2\pi I_3 + \frac{\varepsilon}{\cos\varphi_1 + \cos\varphi_2 + \cos\varphi_3 + 4} \ , \tag{3}$$

satisfying the hypotheses of the Nekhoroshev theorem (H_0 is steep and the perturbation is analytic). Precisely, after selecting a vector ℓ, for example $\ell = (1, 1, 0)$, and a value of the actions I_* on the resonance $\ell \cdot (I_1, -I_2 + I_2^2/\pi, 2\pi) = 0$, we compare the time evolution of the action variables obtained from the output of numerical simulations with the theory presented in this paper. Precisely,

- For fixed values of ε, we perform step (a) above by constructing, via a comptuter program, an explicit canonical transformation in a neighbourhood of I_*

  $$(\hat{S}, \hat{F}_1, \hat{F}_2, \sigma, \phi_1, \phi_2) = \mathcal{C}(I_1, I_2, I_3, \varphi_1, \varphi_2, \varphi_3; \varepsilon) \tag{4}$$

  conjugating the Hamiltonian H_ε to:

  $$H^N = \omega\hat{F}_1 + 2\pi\hat{F}_2 - \hat{F}_1\hat{S} + \frac{A}{2}\hat{S}^2 + \frac{1}{2}\hat{F}_1^2 + \frac{\hat{S}^3}{3\pi} + \varepsilon f_\ell(\hat{S}, \hat{F}, \sigma; \varepsilon) + r_\ell(\hat{S}, \hat{F}, \sigma, \phi; \varepsilon) \tag{5}$$

  defined in a neighbourhood of $(\hat{S}, \hat{F}_1, \hat{F}_2) = (0, 0, 0)$, where r_ℓ represents the remainder of the normal form after an optimal number of normalization steps. Precisely, the norm of r_ℓ decreases exponentially with ε (see Tables 1 and 2 for the precise values of the norm of r_ℓ). If we neglect, as a first approximation, the exponentially small remainder r_ℓ in the normal form (5), the actions $\hat{F}_1, \hat{F}_2$ are first integrals, while the motion of $(\hat{S}, \sigma)$ is integrable, admiting an unstable equilibrium solution along with its associated separatrices (provided $f_\ell(0, 0, \sigma; \varepsilon)$ is a non constant function of σ); these equilibria for the reduced system $(\hat{S}, \sigma)$ define hyperbolic 2-dimensional hyperbolic tori parameterized by the angles ϕ_1, ϕ_2 in the complete system, with stable and unstable manifolds contained in the sets $\hat{F}_1, \hat{F}_2 = constant$. This approximate description of the motions does not include Arnold diffusion, since for $r_\ell = 0$ the actions $\hat{F}_1, \hat{F}_2$ are first integrals. Nevertheless it allows us to compute a neighbourhood of the stable/unstable manifolds of the hyperbolic tori where we expect to find Arnold diffusion. Such an example is shown in Fig. 1. We select a swarm of 100 initial conditions in the neighborhood of the hyperbolic torus of the $(1, 1, 0)$ resonance for the parameters discussed in Section 5, within an energy interval comparable to the norm of the remainder. We then back–transform these initial conditions to the original variables (I, φ) through the inverse of the transformation (4), and we numerically compute the solutions of Hamilton's equations of the complete Hamiltonian H_ε for this swarm of trajectories. Transforming again to the optimal variables $(\hat{S}, \hat{F}_1, \hat{F}_2, \sigma, \phi_1, \phi_2)$, Fiq. 1 shows the projection of these orbits in the space $\sigma, \hat{F}_1, \hat{S}$, for a complete circulation of σ. Since for the numerical integration we use the complete Hamiltonian, the actions $\hat{F}_1, \hat{F}_2$ are not first integrals, but undergo a slow variation determining a drift of the orbits along the resonance. Due to the different initial conditions $\phi(0)$ we observe a spread of the action $\hat{F}_1$ during this circulation.

- Now, the aim is to predict the variation of the normal form action $\hat{F}_1$ during a circulation without performing the numerical integrations. The red curve in Fig. 1 represents the orbit of the swarm yielding the maximum negative jump for $\hat{F}_1$. The blue curve represents the evolution computed analytically using the stationary phase approximations that we will describe in the paper (see (b) and (c) above). We have a very good agreement between the numerical integrations and the predictions of our model.

- The fastest drift along the resonance is produced by a systematic variation of the action $\hat{F}_1$ caused by a sequence of resonant circulations (or librations), each one producing a variation $\Delta\hat{F}_1$ of the same sign. The amplitude and sign of $\Delta\hat{F}_1$ are determined by the values of the phases ϕ_1, ϕ_2 at the beginning of each circulation from one circulation to the other. We here assume a random variation of these phases. This is a heuristic assumption, justified by the fact that the dynamics is chaotic, and the phases $\phi_1(t), \phi_2(t)$ are fast with respect to the periods of the circulations. Random variation of the phases yields a random walk of $\hat{F}_1$ and, by selecting an initial condition such that the values of the phases at each circulation/libration produce the maximum $\Delta\hat{F}_1$ allowed by the analytic formulas, we obtain a monotonic ballistic motion along the resonance, such as the one represented in Fig. 1. Our model provides a formula to compute the speed $\mathcal{D}$ of these motions along the resonance (see Eq. (83)). In Fig. 3 we compare the speed $\mathcal{D}$ of Arnold diffusion along the resonance $\ell = (1, 1, 0)$ measured from numerical experiments (red dots) with the values obtained using the semi-analytic theory developed in this paper (blue points) (i.e. the values of $\mathcal{D}$ computed using Eq. (83)) computed for several values of $\varepsilon \in [0.0005, 0.08]$; we find a very good agreement between the two quantities.

The paper is organized as follows. In Section 2 we define the Melnikov integrals from the normal forms of Nekhoroshev theory. In Section 3 we provide the asymptotic representations of the Melnikov integrals using stationary phase approximations. In Section 4 we present a semi-analytic solution to Problem 1. Section 5 is devoted to a numerical demonstration of the theory presented in Sections 2, 3, 4.

2 Nekhoroshev normal forms and Melnikov integrals

The long–term dynamics of the quasi–integrable Hamiltonian (1) is traditionally studied using the averaging method. In the refined version of the method defined within the proof of Nekhoroshev's theorem, for a d–dimensional lattice $\Lambda \subseteq \mathbb{Z}^n$ defining the resonance

$$\mathcal{R}_\Lambda = \{I \in \mathcal{G} : \quad \ell \cdot \nabla H_0(I) = 0, \quad \forall \ell \in \Lambda\},$$

one constructs a canonical transformation

$$(S, F, \sigma, \phi) = \mathcal{C}(\Gamma^{-T} I, \Gamma\varphi; \varepsilon)$$

defined for $(S, F, \sigma, \phi) \in \mathcal{G}_\Lambda \times \mathbb{T}^n$, where

- Γ is a matrix with $\Gamma_{ij} \in \mathbb{Z}$ and $\det \Gamma = 1$, that defines a linear canonical map (see [4])

$$(\tilde{S}, \tilde{F}, \tilde{\sigma}, \tilde{\phi}) = (\Gamma^{-T} I, \Gamma\varphi) \tag{6}$$

  and conjugates $H_0(I)$ to $h(\tilde{S}, \tilde{F})$ such that the resonance $\mathcal{R}_\Lambda$ is transformed into

$$\tilde{\mathcal{R}} = \{(\tilde{S}, \tilde{F}) \in \Gamma^{-T}\mathcal{G} : \quad \frac{\partial}{\partial \tilde{S}_j} h(\tilde{S}, \tilde{F}) = 0, \quad \forall j = 1, \ldots, d\} \ ,$$

  since in the new variables the resonant lattice Λ is transformed into the lattice $\tilde{\Lambda}$ generated by $e_1, \ldots, e_d$, ($e_1, \ldots, e_n$ denotes the canonical basis of $\mathbb{R}^n$).

- $\sigma \in \mathbb{T}^d, \phi \in \mathbb{T}^{n-d}$ are angles conjugate to the actions $S \in \mathbb{R}^d, F \in \mathbb{R}^{n-d}$.

- The action domain $\mathcal{G}_\Lambda \subseteq \Gamma^{-T}\mathcal{G}$ is an open set whose definition depends both on the resonant lattice Λ and on ε. In particular, $\mathcal{G}_\Lambda$ is a neighbourhood of the resonance $\tilde{\mathcal{R}}$ except for gaps at the crossing with different resonances. In Nekhoroshev theory the gaps are defined with reference to a cut-off order $N := N_\varepsilon$ depending on ε, so that for any $(S_*, F_*) \in \mathcal{G}_\Lambda \cap \tilde{\mathcal{R}}$ and for any $(\nu_1, \ldots, \nu_d, k_1, \ldots, k_{n-d}) \in \mathbb{Z}^n \backslash \tilde{\Lambda}$ with $0 < \sum_j |\nu_j| + \sum_j |k_j| \leq N$ we have

$$|k \cdot \omega_*| \geq \frac{c}{N^q} \tag{7}$$

  where $k = (k_1, \ldots, k_{n-d})$,

$$\omega_* = \left(\frac{\partial h}{\partial F_1}(S_*, F_*), \ldots, \frac{\partial h}{\partial F_{n-d}}(S_*, F_*) \right) , \tag{8}$$

  c is a constant independent on ε, $q = n - d - 1$ if H_0 is quasi-convex.

- $\mathcal{C}$ is a near to the identity transformation which, when composed with (6), conjugates H_ε to the Nekhoroshev normal form Hamiltonian

$$H_{\varepsilon,\Lambda} = h(S, F) + \varepsilon f_\Lambda(S, F, \sigma; \varepsilon) + r_\Lambda(S, F, \sigma, \phi; \varepsilon) \ , \tag{9}$$

  where f_Λ does not depend on the angles ϕ, and the remainder r_Λ has norm bounded by a factor exponentially small with respect to $-1/\varepsilon^a$ (see [57, 61, 54, 55, 41] for precise definitions and statements).

The integer $d \in 1, \ldots, n-1$ is called the multiplicity of the resonance.

Although the proof of Nekhoroshev's theorem grants the existence of normal forms (9) for suitably small values $0 \le \varepsilon \le \varepsilon_0$ (see Eq. (2)), the precise value of the threshold parameter ε_0 is believed to be largely underestimated by the general proofs, while the Fourier coefficients of f_Λ, r_Λ are estimated in norm, but not explicitly provided. Both problems can be overcome by constructing the normal forms (9) with computer assisted methods [37, 15, 38, 27]. We call Hamiltonian normalizing algorithm (HNA) a computer-algebraic implementation which provides the coefficients of the Nekhoroshev normal form (9). We use the HNA introduced in [27], which normalizes quasi-integrable Hamiltonians H_ε. The HNA is constructed by composing N elementary transformations; the input of the HNA is the Hamilton function, a resonance lattice Λ, a domain $G \times \mathbb{T}^n$ where the transformation is defined; the output of the HNA is a canonical transformation $(S, F, \sigma, \phi) = \mathcal{C}^N(I, \varphi)$ and a normal form Hamiltonian

$$H^N = h(S,F) + \varepsilon f^N(S,F,\sigma) + r^N(S,F,\sigma,\phi) \ , \tag{10}$$

conjugate to H_ε by $\mathcal{C}^N$. Both f^N and the remainder r^N are provided as a Taylor-Fourier series expanded at[2] $(S_*, F_*) \in G \times \tilde{\mathcal{R}}$,

$$f^N(S,F,\sigma) = \sum_{m \in \mathbb{N}^d} \sum_{p \in \mathbb{N}^{n-d}} \sum_{\nu \in \mathbb{Z}^d} f_\nu^{m,p} \ (S - S_*)^m (F - F_*)^p e^{i\nu\cdot\sigma}, \tag{11}$$

$$r^N(S,F,\sigma,\phi) = \sum_{m \in \mathbb{N}^d} \sum_{p \in \mathbb{N}^{n-d}} \sum_{\nu \in \mathbb{Z}^d} \sum_{k \in \mathbb{Z}^{n-d}} r_{\nu,k}^{m,p} \ (S - S_*)^m (F - F_*)^p e^{i\nu\cdot\sigma + ik\cdot\phi} \tag{12}$$

with computer-evaluated truncations involving a large number of terms (for example, we need $\sim 10^7, 10^8$ for $n = 3$ and $d = 1$). We remark that, since f^N, r^N are produced for any needed value of ε by a numerical algorithm, their dependence on ε is hidden in the numerical values of the coefficients $f_\nu^{m,p}, r_{\nu,k}^{m,p}$. The HNA computes also the value of the norms:

$$\left\| r^N \right\|_{\rho_S, \rho_F} = \sum_{m \in \mathbb{N}^d} \sum_{p \in \mathbb{N}^{n-d}} \sum_{\nu \in \mathbb{Z}^d} \sum_{k \in \mathbb{Z}^{n-d}} \left| r_{\nu,k}^{m,p} \right| \ \rho_S^{|m|} \rho_F^{|p|}, \tag{13}$$

for convenient weights ρ_S, ρ_F, which have to be chosen larger than the variations $\Delta S_j, \Delta F_j$ during the Arnold diffusion. Since the variations ΔS_j satisfy $|\Delta S_j| \le C\sqrt{\varepsilon}$ and ΔF_j is extremely small, we set $\rho_S := C\sqrt{\varepsilon}$ and we compute the approximate value for the norm of $r^N(S, F_*, \sigma, \phi)$:

$$\left\| r^N \right\| := \left\| r^N \right\|_{\rho_S, 0} = \sum_{m \in \mathbb{N}^d} \sum_{\nu \in \mathbb{Z}^d} \sum_{k \in \mathbb{Z}^{n-d}} \left| r_{\nu,k}^{m} \right| \ \rho_S^{|m|}, \quad r_{\nu,k}^{m} := r_{\nu,k}^{m,0} \quad . \tag{14}$$

Finally, the normalization order N is chosen so that $\left\| r^1 \right\| > \ldots > \left\| r^N \right\|$ and $\left\| r^{N+1} \right\| > \left\| r^N \right\|$, thus the normal form is called optimal, r^N the optimal remainder and N the optimal normalization order.

If we artificially suppress the remainder r_Λ in Eq. (9), or correspondingly r^N in (10), we obtain a small perturbation of the original Hamiltonian which possibly exhibits chaotic motions due to homoclinic and heteroclinic phenomena (for $d > 1$), but in which the actions F_j, which we call 'adiabatic', remain constant in time. Therefore, in the flow of the complete

[2] For any $p \in \mathbb{N}^{n-d}, m \in \mathbb{N}^d$, we use the multi-index notation for $(F - F_*)^p$ and $(S - S_*)^m$ which denote, respectively, $(F_1 - F_{*,1})^{p_1} \cdots (F_{n-d} - F_{*,n-d})^{p_{n-d}}$ and $(S_1 - S_{*,1})^{m_1} \cdots (S_d - S_{*,d})^{m_d}$; we also denote by $|p|, |m|$ the lengths $\sum_{j=1}^{n-d} |p_j|$, $\sum_{j=1}^{d} |m_j|$.

Hamiltonian, any long–term evolution of the adiabatic actions is due to the accumulation of the effects of the very small remainder on very long times. In particular, the adiabatic actions F_j have a long-term variation, representing the drift along the resonance, bounded for an exponentially long time by

$$|F_j(t) - F_j(0)| \leq |t| \left\| \frac{\partial r_\Lambda}{\partial \phi_j} \right\|_{\rho_S,\rho_F} . \tag{15}$$

The *a priori* estimate (15) obtained from the Nekhoroshev normal form (9) provides an upper bound to the average variation of the adiabatic actions; establishing a *lower bound* to $|F_j(t) - F_j(0)|$ for a subset of orbits is the fundamental brick in the theory of Arnold diffusion.

From now on we focus our discussion on resonances of multiplicity $d = 1$. Denoting by (S, σ) the resonant action-angle pair, we first consider the dynamics of the approximated normal form which is obtained from (10) just by dropping the remainder $r^N(S, F, \sigma, \phi)$:

$$\overline{H}^N = h(S, F) + \varepsilon f^N(S, F, \sigma). \tag{16}$$

Since the corresponding Hamiltonian $\overline{H}^N$ depends only on one angle, it is integrable, and we represent its motions as follows. Following [4], we first expand $\overline{H}^N$ at $(S_*, F_*) = \Gamma^{-T} I_*$ identifying the center of the resonance, precisely such that

$$\frac{\partial h}{\partial S}(S_*, F_*) = 0, \tag{17}$$

and then we represent $\overline{H}^N_*(\hat{S}, \hat{F}, \sigma) = \overline{H}^N(S_* + \hat{S}, F_* + \hat{F}, \sigma)$ by

$$\overline{H}^N_* = \overline{H}_0 + \overline{H}_1 \quad , \quad \overline{H}_0 = \omega_* \cdot \hat{F} + \frac{A}{2}\hat{S}^2 + \hat{S}B \cdot \hat{F} + \frac{1}{2}C\hat{F} \cdot \hat{F} + \varepsilon v(\sigma) \ , \tag{18}$$

where $\hat{F} = F - F_*$, $\hat{S} = S - S_*$; $A \in \mathbb{R}$, $\omega_*, B \in \mathbb{R}^{n-1}$, the square matrix C and the function $v(\sigma)$ depend parametrically on S_*, F_*. Since the actions $\hat{F}$ are constants of motion for the Hamiltonian flow of $\overline{H}^N_*$, the dynamics of the variables $\hat{S}, \sigma$ is determined by the reduced one degree of freedom Hamiltonian $\overline{H}^N_*(\hat{S}, \hat{F}, \sigma)$ where $\hat{F}$ are treated as parameters, and we consider the case $\hat{F} = 0$ (which corresponds to $F(0) = F_*$). Following again [4], we rescale the action $\hat{S}$ and the time t:

$$\hat{S} = \sqrt{\varepsilon}\mathcal{S} \quad , \quad t = \tau/\sqrt{\varepsilon},$$

so that $\mathcal{S}(\tau), \sigma(\tau)$ are solutions of the Hamilton equations of the effective Hamiltonian:

$$\mathcal{H}(\mathcal{S}, \sigma; \varepsilon) := \mathcal{H}_0(\mathcal{S}, \sigma) + \sqrt{\varepsilon}\mathcal{H}_1(\mathcal{S}, \sigma; \varepsilon) \tag{19}$$

where

$$\mathcal{H}_0(\mathcal{S}, \sigma) = \frac{A}{2}\mathcal{S}^2 + v(\sigma) \tag{20}$$

is independent of ε and $\sqrt{\varepsilon}\mathcal{H}_1$, defined from the Taylor expansion of $\overline{H}_1$ (see [4] for details), is of order $\sqrt{\varepsilon}$. Since the Hamiltonian $\mathcal{H}_0$ represents better the normal form dynamics as soon as ε goes to zero, we are going to use it as the reference dynamics to conveniently approximate the Melnikov integrals; we need to state the following hypothesis:

(i) For simplicity we assume that $v(\sigma)$ has a unique non-degenerate local maximum and a unique non-degenerate local minimum. If this hypothesis is not satisfied, for example because $v(\sigma)$ has more than one local maximum, then one has to reformulate the theory presented in this paper accordingly; if instead $v'(\sigma) = 0$ for all σ, then our method needs major modifications. We denote by

$$M = \max_{\sigma\in[0,2\pi]} v(\sigma), \quad \bar{M} = \min_{\sigma\in[0,2\pi]} v(\sigma) \ ,$$

and (without loss of generality) we assume that the maximum is at $\sigma = 0$, $M > 0 > \bar{M}$, and $A > 0$, $\varepsilon > 0$.

(ii) For any suitably small value of $\varepsilon \geq 0$, we have that Hamilton's equations of (19) admits an unstable equilibrium $\mathcal{S}^*_\varepsilon, \sigma^*_\varepsilon$ close to the equilibrium $(\mathcal{S}^*_0, \sigma^*_0) = (0, 0)$ of the Hamilton's equations of (20).

(iii) For suitably small values of $\varepsilon \geq 0$ and $\alpha > 0$, the equation

$$\mathcal{H}_0(\mathcal{S}, \sigma) + \sqrt{\varepsilon}\mathcal{H}^1(\mathcal{S}, \sigma; \varepsilon) = \mathcal{H}(\mathcal{S}^*_\varepsilon, \sigma^*_\varepsilon; \varepsilon)(1 + \alpha) \tag{21}$$

is solvable with respect to any $\sigma \in [0, 2\pi]$, by providing the function:

$$\mathcal{S} := \mathcal{S}_\alpha(\sigma; \varepsilon).$$

We remark that for $\varepsilon = 0$ the equation is solvable for any $\sigma \in [0, 2\pi]$ providing:

$$s_\alpha(\sigma) := \mathcal{S}_\alpha(\sigma; 0) = \pm\sqrt{\frac{2}{|A|}(M(1 + \alpha) - v(\sigma))}. \tag{22}$$

For any $\alpha \neq 0$, we denote by T^ε_α the period of the corresponding solutions of Hamilton's equations under the flow of $\overline{H}^N_*$ (for $\hat{F} = 0$), and we set $T_\alpha := T^0_\alpha$.

Remark:

(x) One is tempted to invoke an implicit function theorem to prove the existence of the equilibria $(\mathcal{S}^*_\varepsilon, \sigma^*_\varepsilon)$ and of the function $\mathcal{S}_\alpha(\sigma; \varepsilon)$ for all ε in a neighbourhood of $\varepsilon = 0$. Unfortunately, since the construction of the Nekhoroshev normal form is not continuous with respect to ε (the number $N := N_\varepsilon$ of normalizations is a discontinuous function of ε), applications of classical formulations of the implicit function theorem are prevented. To gain continuity with respect to ε one may consider a modification of the normal form Hamiltonian coinciding with the original normal form Hamiltonian except for ε in a sequence of suitably small intervals A_j (which will be excluded by the analysis) centered at the discontinuity values ε_j of N_ε, with $N_\varepsilon = j$ for all $\varepsilon \in [\varepsilon_j, \varepsilon_{j-1})$.[3]

[3]For example, consider the intervals $B_j = [\varepsilon_j, \varepsilon_{j-1})\backslash(A_j \cup A_{j-1})$ and a set of smooth functions χ^j_ε satisfying

$$\chi^j_\varepsilon = \begin{cases} 1 & \text{if } \ \varepsilon \in B_j \\ 0 & \text{if } \ \varepsilon \notin A_j \cup B_j \cup A_{j-1} \end{cases} .$$

The modified Hamiltonian is obtained by replacing the function $\mathcal{H}_1$ with the function $\chi_\varepsilon\mathcal{H}_1$, where $\chi_\varepsilon = \sum_{j\geq 1} \chi^j_\varepsilon$, so that we have: $\chi_\varepsilon\mathcal{H}_1 = \mathcal{H}_1$ for all $\varepsilon \notin \cup_{j\geq 1} A_j$, as well as:

$$\lim_{\varepsilon\to 0} \sqrt{\varepsilon}\chi_\varepsilon\mathcal{H}_1(\mathcal{S}, \sigma; \varepsilon) = 0,$$

thus gaining continuity.

(xi) The value of C in Problem 1 can be set as follows:

$$C = 2\sqrt{\frac{2}{|A|}(M - \overline{M})}. \tag{23}$$

When considering the solutions $(S(t), F(t), \sigma(t), \phi(t))$ of Hamilton's equations of the complete Hamiltonian (10), the adiabatic actions F_j can have a slow evolution forced by the remainder r^N, whose variation in a time interval $[0, T]$ is given by

$$\Delta F_j(T) = F_j(T) - F_j(0) = -\sum_{m,p,\nu,k} \int_0^T ik_j r_{\nu,k}^{m,p} \hat{F}(t)^p \hat{S}(t)^m e^{i\nu\sigma(t) + ik\cdot\phi(t)} dt =: \sum_{m,p,\nu,k} \Delta F_{j,T}^{m,p,\nu,k} \ . \tag{24}$$

According to the well known Melnikov approach (see [18] for a review) we approximate $\Delta F_j(T)$ with

$$-\sum_{m,\nu,k} \int_0^T ik_j r_{\nu,k}^m \hat{S}^N(t)^m e^{i\nu\sigma^N(t) + ik\cdot\phi^N(t)} dt \quad , \tag{25}$$

obtained by replacing the solution $(S(t), F(t), \sigma(t), \phi(t))$ in the integrals with

$$(S^\varepsilon(t), F_*, \sigma^\varepsilon(t), \phi^\varepsilon(t)) = (S_*, F_*, 0, 0) + (\hat{S}^N(t), 0, \sigma^N(t), \phi^N(t))$$

where $(\hat{S}^N(t), 0, \sigma^N(t), \phi^N(t))$ is solution of Hamilton's equations of $\overline{H}_*^N$.

For simplicity we only consider the rotational solutions of the normal form dynamics, i.e. initial conditions characterized by a value of the parameter $\alpha > 0$ (see Eq. (21)); then, we change the integration variable in (25) to $\sigma = \sigma^N(t)$, obtaining

$$\Delta F_{j,T}^{m,\nu,k} := \Delta F_{j,T}^{m,0,\nu,k} \simeq \Delta^0 F_{j,T}^{m,\nu,k} := -ik_j r_{\nu,k}^m \varepsilon^{\frac{m-1}{2}} e^{ik\cdot\phi(0)} \int_0^{\sigma^\varepsilon(T)} \frac{\mathcal{S}_\alpha(\sigma;\varepsilon)^m}{\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon), \sigma;\varepsilon)} e^{i\Theta_{\nu,k}(\sigma;\varepsilon)} d\sigma \tag{26}$$

where the phase $\Theta_{\nu,k}(\sigma;\varepsilon)$ is defined by:

$$\Theta_{\nu,k}(\sigma;\varepsilon) = \nu\sigma + \int_0^\sigma \frac{k\cdot\frac{\partial\bar{H}_*^N}{\partial F}(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon), 0, \sigma)}{\sqrt{\varepsilon}\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon), \sigma;\varepsilon)} d\sigma.$$

The change of the adiabatic action F_j over a full cycle of the resonant angle σ is therefore approximated by:

$$\Delta^0 F_j = \sum_{m,\nu,k} -ik_j r_{\nu,k}^m \varepsilon^{\frac{m-1}{2}} e^{ik\cdot\phi(0)} \ \mathcal{I}_{m,\nu,k} \tag{27}$$

where $\mathcal{I}_{m,\nu,k}$ denote the integrals:

$$\mathcal{I}_{m,\nu,k} = \int_0^{2\pi} \frac{\mathcal{S}_\alpha(\sigma;\varepsilon)^m}{\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon), \sigma;\varepsilon)} e^{i\Theta_{\nu,k}(\sigma;\varepsilon)} d\sigma. \tag{28}$$

Since for typical computations (see Section 5 for explicit examples) we have to consider millions of remainder terms $r_{\nu,k}^m$, the direct numerical computation of millions of integrals (28) is hardly tractable in practice. As a matter of fact, we find that the direct numerical computation of all the integrals (28) is not necessary, since most of the terms in (25) (including some with the largest $\left|r_{\nu,k}^m\right|$) contribute very little to the sum (25). This fact can be explained by invoking methods of asymptotic analysis inspired by the so–called principle of the stationary phase (PSP hereafter).

Remark (xii). Usually, Melnikov approximations are introduced to compute the splittings of stable-unstable manifolds, so that integrals like (24) are approximated by choosing $(\hat{S}^0(t), \sigma^0(t), \phi^0(t))$ to be the solution of the approximate normal form $\overline{H}_0$ corresponding to a separatrix homoclinic loop ($\alpha = 0$ in our notation). Our method exposed below differs from the usual Melnikov approach since, in order to find the orbits which diffuse in shorter time along the resonance, we evaluate the integrals for a solution $(\hat{S}^0(t), \sigma^0(t), \phi^0(t))$ which is suitably close to, but *not exactly on* the separatrix, precisely $\alpha \sim \left\| r^N \right\|$, with finite period T_α.

3 Stationary phase approximation of Melnikov integrals

3.1 The principle of stationary phase

In its classical formulation (e.g., see [10]) the principle concerns the asymptotic behaviour of the parametric integrals

$$I_\lambda = \int_a^b \eta(\sigma) e^{i\lambda\Phi(\sigma)} d\sigma \ , \tag{29}$$

when the parameter λ is large. With mild conditions on the amplitude function $\eta(\sigma)$, we have the following cases:

(A) The phase $\Phi(\sigma)$ has no stationary points, i.e. $\Phi(\sigma) \neq 0$ for all $\sigma \in [a, b]$, for large λ we have

$$I_\lambda \sim \frac{\eta(\sigma)}{\lambda\Phi'(\sigma)} e^{i(\lambda\Phi(\sigma))} \Big|_a^b , \tag{30}$$

where the neglected contributions are of order smaller than $1/\lambda$.

(B) The phase $\Phi(\sigma)$ has a non–degenerate stationary point $\sigma_c \in (a, b)$, i.e. $\Phi'(\sigma_c) = 0$ and $\Phi''(\sigma_c) \neq 0$. Then, if $\eta(\sigma_c) \neq 0$, for large λ we have

$$I_\lambda \sim \eta(\sigma_c) e^{i\left(\lambda\Phi(\sigma_c) \pm \frac{\pi}{4}\right)} \sqrt{\frac{2\pi}{\lambda \left|\Phi''(\sigma_c)\right|}} \ , \tag{31}$$

where the $\pm$ is chosen according to the sign of $\Phi''(\sigma_c)$. If there are more stationary points in (a, b), we must sum all the corresponding terms.

(C) The phase $\Phi(\sigma)$ has a degenerate stationary point $\sigma_c \in (a, b)$, i.e. $\Phi'(\sigma_c) = 0, \Phi''(\sigma_c) = 0$ and we assume $\Phi'''(\sigma_c) \neq 0$. Then, if $\eta(\sigma_c) \neq 0$, for large λ we have

$$I_\lambda \sim \eta(\sigma_c) e^{i\lambda\Phi(\sigma_c)} \sqrt{3}\Gamma(4/3) \left(\frac{6}{\lambda \left|\Phi'''(\sigma_c)\right|}\right)^{\frac{1}{3}} . \tag{32}$$

3.2 Heuristic discussion of the PSP for Melnikov integrals

In this Subsection we discuss the heuristic arguments that lead to produce asymptotic expansions for (27); rigorous results will be presented in Subsections 3.3 and 3.4.

We first notice that all the terms appearing in (27) are proportional to the coefficients $r^m_{\nu,k}$, which converge exponentially fast to zero for increasing values of $|\nu|+|k|$; therefore we limit to consider terms of the expansion such that $|\nu|+|k|\le \hat{N}$, with suitable $\hat{N}\ge N$. Then, in each integral (28)

$$\mathcal{I}_{m,\nu,k}=\int_0^{2\pi}\frac{\mathcal{S}_\alpha(\sigma;\varepsilon)^m}{\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon),\sigma;\varepsilon)}e^{i\Theta_{\nu,k}(\sigma;\varepsilon)}d\sigma \tag{33}$$

we identify the phase function $\Theta_{\nu,k}(\sigma;\varepsilon):=\lambda\Phi(\sigma)$ to apply the method of the stationary phase. In fact, by expressing the derivative of $\Theta_{\nu,k}(\sigma;\varepsilon)$ profitting of the representations of the normal form Hamiltonian considered in Section 2 we obtain:

$$\begin{aligned}\frac{d\Theta_{\nu,k}}{d\sigma}&=\nu+\frac{1}{\sqrt{\varepsilon}}\frac{k\cdot\frac{\partial\bar{H}^N_*}{\partial F}(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon),0,\sigma)}{\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon),\sigma;\varepsilon)}\\
=\left(\nu+\frac{k\cdot B}{A}\right)&+\frac{1}{\sqrt{\varepsilon}\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon),\sigma;\varepsilon)}k\cdot\Bigg[\omega_*-B\frac{\varepsilon}{A}\frac{\partial\mathcal{H}_1}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon),\sigma;\varepsilon)\\
&+\mathcal{O}_2(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon))+\varepsilon\frac{\partial f^N}{\partial F}(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon),F_*,\sigma)\Bigg]\end{aligned}\tag{34}$$

where $\mathcal{O}_2(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon))$ denotes terms at least quadratic in $\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon)$. We notice that $\frac{d\Theta_{\nu,k}}{d\sigma}$ contains:

- the term:

$$\frac{k\cdot\omega_*}{\sqrt{\varepsilon}\,\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon),\sigma;\varepsilon)} \tag{35}$$

  which is large for $\varepsilon<\varepsilon_0$, with ε_0 suitably chosen positive threshold, if one assumes:

$$|k\cdot\omega_*|\ge\Gamma_0\sqrt{\varepsilon},\quad \forall\varepsilon<\varepsilon_0 \tag{36}$$

  with Γ_0 suitably large parameter independent of ε. By Nekhoroshev theory equation (36) is in particular satisfied by all $k\in\mathbb{Z}^n\backslash 0$ with $|k|\le N$ (see Subsection 3.4); for a discussion about the case $N<|k|\le\hat{N}$ with suitable $\hat{N}>N$ see Subsection 3.4; the terms with $|k|\ge\hat{N}$ will be neglected.

- the term:

$$\frac{1}{\sqrt{\varepsilon}\frac{\partial\mathcal{H}}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon),\sigma;\varepsilon)}k\cdot\left[-B\frac{\varepsilon}{A}\frac{\partial\mathcal{H}_1}{\partial\mathcal{S}}(\mathcal{S}_\alpha(\sigma;\varepsilon),\sigma;\varepsilon)+\mathcal{O}_2(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon))+\varepsilon\frac{\partial f^N}{\partial F}(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\varepsilon),F_*,\sigma)\right]$$

  which, for $|k|\le N$, converges pointwise to zero as ε tends to zero.

- the constant term

$$\mathcal{N}=\nu+\frac{k\cdot B}{A} \tag{37}$$

  which, for $|\nu|+|k|\le N$, is divergent at most as $|\mathcal{N}|\le\max(1,\|B\|/|A|)N$.

Therefore, the derivative of the phase contains terms which are divergent for ε going to zero, and we will justify in the next sub-section the use of the method of the stationary phase by identifying the large parameter λ with the parameter $|\mathcal{W}|$, where

$$\mathcal{W}=\pm\frac{k\cdot\omega_*}{\sqrt{2\,|A|\,\varepsilon(M-\bar{M})}} \tag{38}$$

with the sign $\pm$ chosen according to the signs of $\mathcal{S}_\alpha(\sigma;0)$ and A. For the terms with k satisfying (36) the parameter $|\mathcal{W}|$ takes values in the large interval:

$$\frac{\Gamma_0}{\sqrt{2\,|A|\,(M-\bar{M})}} \leq |\mathcal{W}| \leq \frac{|k|}{\sqrt{\varepsilon}}\frac{\|\omega_*\|}{\sqrt{2\,|A|\,(M-\bar{M})}}.$$

Therefore, also for fixed ε,α, it is meaningful to estimate the asymptotic value of the Melnikov integrals in the limit of large values of $|\mathcal{W}|$.

According to this idea, the Melnikov integrals whose oscillating phase $\Theta_{\nu,k}(\sigma;\varepsilon)$ has critical points are expected to be dominant over those whose oscillating phase has no critical points. However, we find that the asymptotic behaviour of the Melnikov integrals is more complicated than the behaviour of the integrals (29) thus rendering necessary to use a refinement of the stationary phase method: specifically, we need to consider also a case which is intermediate between (A) and (B), called hereafter the quasi-stationary case, produced by the disappearance of couples of non-degenerate critical points σ_c^1,σ_c^2 after they merge into a degenerate critical point. The quasi-stationary case represents a transition between the stationary and the non-stationary case, which is not considered in the usual formulations of the PSP method. To be more precise, depending on the values of ν,k, we will consider three cases

(I) *the phase $\Theta_{\nu,k}(\sigma;\varepsilon)$ is 'fast' for all $\sigma\in[0,2\pi]$, i.e. we have $\left|\Theta'_{\nu,k}(\sigma;\varepsilon)\right| > |\mathcal{W}|^{1/3}$ (this is the case, for example, when $\mathcal{N}$ and $\mathcal{W}$ have the same sign).*

In this case the integral in (26) is estimated smaller than order $|\mathcal{W}|^{-1/3}$, see Eq. (30), and we will assume that the contribution of $\Delta F_{j,T}^{m,\nu,k}(T)$ to the series expansion in Eq. (24) can be neglected. In fact, we find that for the vast majority of remainder terms one has $\left|\Theta'_{\nu,k}(\sigma;\varepsilon)\right| > |\mathcal{W}|$, in which case the estimate (30) reduces to contributions of order $1/|\mathcal{W}|$, which is much smaller than $1/|\mathcal{W}|^{1/3}$.

(II) *the phase $\Theta_{\nu,k}(\sigma;\varepsilon)$ is 'slow' only close to non-degenerate critical points σ_c, provided they are distant enough with respect to $1/|\mathcal{W}|^{\frac{1}{3}}$.*

In this case, by invoking the PSP (see (31)), the integral in (26) will be approximated by the sum of a contribution of order $1/|\mathcal{W}|^{\frac{1}{2}}$ for each stationary point σ_c (see Lemma 1).

(III) *the phase $\Theta_{\nu,k}(\sigma;\varepsilon)$ is quasi-stationary, i.e. $\left|\Theta'_{\nu,k}(\sigma;\varepsilon)\right| \leq |\mathcal{W}|^{1/3}$ in an interval of size of order $1/|\mathcal{W}|^{\frac{1}{3}}$ (notice that this condition can occur in absence of critical points, or in presence of two very close non-degenerate critical points or of one degenerate critical point).*

While in this case we cannot directly apply (A), (B), (C), we will obtain an asymptotic formula for the integrals stemming from formula (C) (see Lemma 2).

Remarks:

(xiii) From estimates (A) and (B), any individual integral estimated using (31) is of order $1/\sqrt{|\mathcal{W}|}$, larger with respect to the integrals estimated using (30) which are of order $1/|\mathcal{W}|$ (when $|\Theta'|\geq|\mathcal{W}|$). The transition between the case (I) where $|\Theta'|\geq|\mathcal{W}|$ and the case (III) will be considered in Section 3.3. Moreover, in the non–stationary case

(I) the integrals in (30) are estimated only by differences of functions computed at the border values a, b. Since Arnold diffusion is produced by the variations of F_j through a sequence of circulations or librations, the border values of the sequence may cancel (as a matter of fact they may cancel only partially, since from a libration/circulation to the next one there can be small variations of α and a change of the fast phases $\phi(0)$ in the factor multiplying the integral in (26)).

(xiv) The practical classification of the integrals in one of the categories (I), (II), or (III) will be done by a fast algorithmic criterion (see below), based only on each term's integer labels m, ν, k. Since for the large majority of ν, k the phase satisfies (I) (see Section 5, Tables 1 and 2, for numerical examples on a $n = 3$ degrees of freedom Hamiltonian), we have a criterion to select the few harmonics ($\sim 1\text{‰}$ in the numerical examples of Section 5) belonging to (II) and (III), hence, producing the dominant terms in the time evolution of F_j.

(xv) Despite being more complicated than (II), the inclusion in the computation of the quasi–stationary terms (III) is essential, since these individual contributions can be as large as those of (II) and quite often we find algebraic near-cancellations between terms of the groups (II) and (III) leaving residuals of order only few percent of the absolute values of the corresponding Melnikov integrals.

Finally, we discuss a further approximation valid for small values of ε. In fact, let us introduce an auxiliary parameter ξ and the integrals

$$\hat{\mathcal{I}}_{m,\nu,k}(\varepsilon, \alpha, \xi) = \int_0^{2\pi} \frac{\mathcal{S}_\alpha(\sigma;\xi)^m}{\frac{\partial \mathcal{H}_0}{\partial \mathcal{S}}(\mathcal{S}_\alpha(\sigma;\xi), \sigma) + \sqrt{\xi}\frac{\partial \mathcal{H}_1}{\partial \mathcal{S}}(\mathcal{S}_\alpha(\sigma;\xi), \sigma; \xi)} e^{i\hat{\Theta}_{\nu,k}(\sigma;\varepsilon,\alpha,\xi)} d\sigma$$

with

$$\hat{\Theta}_{\nu,k}(\sigma;\varepsilon,\alpha,\xi) = \nu\sigma + \int_0^\sigma \frac{k \cdot \left(\frac{\partial \bar{H}_0}{\partial F}(\sqrt{\varepsilon}\mathcal{S}_\alpha(\sigma;\xi), 0, \sigma) + \frac{\partial \bar{H}_1}{\partial F}(\sqrt{\xi}\mathcal{S}_\alpha(\sigma;\xi), 0, \sigma; \xi)\right)}{\sqrt{\varepsilon}\left(\frac{\partial \mathcal{H}_0}{\partial \mathcal{S}}(\mathcal{S}_\alpha(\sigma;\xi), \sigma) + \sqrt{\xi}\frac{\partial \mathcal{H}_1}{\partial \mathcal{S}}(\mathcal{S}_\alpha(\sigma;\xi), \sigma; \xi)\right)} d\sigma.$$

By assuming the continuity of $\mathcal{S}_\alpha(\sigma;\xi)$ in $\xi = 0$ (see Remark (x)), by the Dominated Convergence Theorem, for every $\varepsilon > 0$ and $\alpha > 0$ we have:

$$\lim_{\xi \to 0} \hat{\mathcal{I}}_{m,\nu,k}(\varepsilon, \alpha, \xi) = \hat{\mathcal{I}}_{m,\nu,k}(\varepsilon, \alpha, 0) = \int_0^{2\pi} [s_\alpha(\sigma)]^{m-1} e^{i\theta_{\nu,k}(\sigma;\varepsilon)} d\sigma \tag{39}$$

where:

$$\theta_{\nu,k}(\sigma;\varepsilon) = \mathcal{N}\sigma + \frac{k \cdot \omega_*}{A\sqrt{\varepsilon}} \int_0^\sigma \frac{dx}{s_\alpha(x)} \tag{40}$$

with $\mathcal{N} = \nu + k \cdot B/A$ (from Eq. (37)).

Remark.

(xvi) We are not allowed to compute the limit of the integral $\mathcal{I}_{m,\nu,k}$ for ε going to zero using the Dominated Convergence Theorem. In fact, since the phase contains terms proportional to $1/\sqrt{\varepsilon}$, the integrand has not pointwise limit.

Since $\mathcal{I}_{m,\nu,k} = \hat{\mathcal{I}}_{m,\nu,k}(\varepsilon, \alpha, \sqrt{\varepsilon})$, we will identify $\hat{\mathcal{I}}_{m,\nu,k}(\varepsilon, \alpha, 0)$ with the main approximation of $\mathcal{I}_{m,\nu,k}$ in the sense given by the limit (39). Even if such approximation is not strictly necessary, we find remarkable that it gives an excellent agreement with the numerical results. Furthermore, the computation of the critical points of the phase of $\hat{\mathcal{I}}_{m,\nu,k}(\varepsilon, \alpha, 0)$ can be done explicitly using $\mathcal{H}_0$ which, in turn, can be retrieved directly from the original Hamiltonian. Therefore, the only information needed from the Hamiltonian normalizing algorithm to compute the approximation of $\Delta\hat{F}_j$ are the values of the coefficients $r^m_{\nu,k}$ multiplying the integrals $\mathcal{I}_{m,\nu,k}$ in equation (26).

3.3 Asymptotic formulas for Melnikov integrals

In this Subsection we state the results which allow us to assign all the terms in Eq. (26), labeled by the integers $m \in \mathbb{N}, \nu \in \mathbb{Z}$ and $k \in \mathbb{Z}^{n-1}$, into the categories (I), (II) or (III). We then provide asymptotic representations for the integrals of the terms in (II) and (III). To simplify the discussion, we assume $\alpha > 0$; the modifications needed to represent the case $\alpha < 0$ are straightforward.

We first analyze the conditions on m, ν, k which imply that the corresponding term has a phase with stationary points. To simplify the analysis we assume mild conditions on the potential $v(\sigma)$ in the normal form Hamiltonian (18). As discussed before, we assume that $v(\sigma)$ has only one relative maximum and one relative minimum. Up to a translation of the angle σ we assume that this maximum is at $\sigma = 0$, so that with the notations introduced in Section 2, we have $v(0) = M$. For simplicity, we hereafter denote the phase $\theta_{\nu,k}(\sigma;\varepsilon)$ of Eq. (40) as $\theta(\sigma)$. Then, Eq. (40) takes the form:

$$\theta(\sigma) = \mathcal{N}\sigma + \mathcal{W}\sqrt{1 - \frac{\bar{M}}{M}} \int_0^\sigma \frac{dx}{\sqrt{1 + \alpha - \frac{v(x)}{M}}}, \tag{41}$$

where $\mathcal{N}, \mathcal{W}$ are defined in (37), (38). In particular, the dependence on ε is included in the parameter $\mathcal{W}$ and we study the asymptotic development of the integrals $\int_0^{2\pi} [s_\alpha(\sigma)]^{m-1} e^{i\theta(\sigma)} d\sigma$ with respect to $\mathcal{W}$. We have the following:

Lemma 1. *Consider a phase $\theta(\sigma)$ defined by the labels ν, k. Then:*

- *If $\mathcal{N} \cdot \mathcal{W} > 0$ the phase $\theta(\sigma)$ has no stationary points;*
- *If $\mathcal{N} \cdot \mathcal{W} < 0$, the phase $\theta(\sigma)$ has stationary points if and only if*

$$\frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\sqrt{1 + \alpha - \frac{\bar{M}}{M}}} \le \frac{|\mathcal{N}|}{|\mathcal{W}|} \le \frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\sqrt{\alpha}}. \tag{42}$$

Furthermore, suppose that $v(\sigma)$ has only one non-degenerate local maximum at $\sigma = 0$, one non-degenerate local minimum at $\sigma = \bar{\sigma}$, and $v'(\sigma) \neq 0$ elsewhere. For any given $\Delta_{Max} > \sqrt{1 - \frac{\bar{M}}{M}}$, consider ε suitably small and

$$\alpha \in \left(0, \min\left(\left\|r^N\right\|, \frac{1 - \frac{\bar{M}}{M}}{2^8 \Delta^2_{Max}}\right)\right). \tag{43}$$

Then

(i) If the inequality (42) is strictly satisfied, the phase $\theta(\sigma)$ has two non–degenerate critical points σ_c^1, σ_c^2 and for all $\mathcal{N} = -\mathcal{W}\Delta$ with $\Delta \in \left(\frac{\sqrt{1-\frac{\bar{M}}{M}}}{\sqrt{1+\alpha-\frac{\bar{M}}{M}}}, \Delta_{Max} \right)$, we have

$$\int_0^{2\pi} [s_\alpha(\sigma)]^{m-1} e^{i\theta(\sigma)} d\sigma = \mathcal{I}(\sigma_c^1) a_1(\mathcal{W}, \mathcal{N}) + \mathcal{I}(\sigma_c^2) a_2(\mathcal{W}, \mathcal{N}) \ , \tag{44}$$

where

$$\mathcal{I}(\sigma_c^j) = \frac{1}{|\mathcal{W}|^{\frac{1}{2}}} \frac{\sqrt{2\pi}\, |A|^{\frac{3}{4}}}{[2(M-\overline{M})]^{\frac{1}{4}}} \sqrt{\frac{\left|s_\alpha(\sigma_c^j)\right|^3}{\left|v'(\sigma_c^j)\right|}} [s_\alpha(\sigma_c^j)]^{m-1} e^{i\theta(\sigma_c^j) \pm i\frac{\pi}{4}} \ , \tag{45}$$

with the $\pm$ depending on the sign of $v'(\sigma_c)$, and the functions a_1, a_2 satisfy

$$\lim_{|\mathcal{W}| \to +\infty} a_1(\mathcal{W}, -\mathcal{W}\Delta) = 1 \ , \ \lim_{|\mathcal{W}| \to +\infty} a_2(\mathcal{W}, -\mathcal{W}\Delta) = 1 \ . \tag{46}$$

(ii) If $\frac{|\mathcal{N}|}{|\mathcal{W}|}$ is equal to the lowermost bound of Eq. (42), the two non-degenerate critical points merge into one degenerate critical point $\sigma_c = \bar{\sigma}$.

In the statement of Lemma 1 (as well as of Lemma 2 below) the parameters N, $\left\| r^N \right\|$, α, $|\mathcal{W}|$, $|\mathcal{N}|$ depend on ε, or are constrained to intervals depending on the value of ε. In the practical applications of the Lemmas for specific values of ε, the numerical values of these parameters are obtained from the output of the Hamiltonian Normalizing Algorithm, so we directly check if the hypotheses of Lemmas 1 and 2 are satisfied. In Subsection 3.4 we address the problem of the expected asymptotic dependence of the parameters on ε by assuming the system in the domain of application of Nekhoroshev's theorem, and afterwards we discuss the solvability of condition (42) in the limit of small ε for a set of ν, k.

Proof of Lemma 1. Since we have

$$\theta'(\sigma) = \mathcal{N} + \mathcal{W}\sqrt{1 - \frac{\bar{M}}{M}} \frac{1}{\sqrt{1 + \alpha - \frac{v(\sigma)}{M}}},$$

if $\mathcal{N} \cdot \mathcal{W} > 0$ there are no stationary points. If $\mathcal{N} \cdot \mathcal{W} < 0$ the stationary points are the solutions of the equation

$$\frac{|\mathcal{N}|}{|\mathcal{W}|} = \frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\sqrt{1 + \alpha - \frac{v(\sigma)}{M}}} \ ,$$

which exist if and only if $|\mathcal{N}|/|\mathcal{W}|$ satisfies (42). By assumption, $v(\sigma)$ has only one local maximum $\sigma = 0$ and one local minimum $\bar{\sigma}$. If (42) is strictly satisfied, the function

$$\tilde{\theta}' = |\mathcal{N}| - |\mathcal{W}| \sqrt{1 - \frac{\bar{M}}{M}} \frac{1}{\sqrt{1 + \alpha - \frac{v(\sigma)}{M}}},$$

has a strict maximum at $\sigma = \bar{\sigma}$ with $\tilde{\theta}'(\bar{\sigma}) > 0$ and converges to strictly negative values for σ tending to 0 or 2π. Therefore there are two values $\sigma_c^1, \sigma_c^2 \in (0, 2\pi) \backslash \{\bar{\sigma}\}$ such that $\tilde{\theta}'(\sigma_c^i) = 0$. Then, from

$$\theta''(\sigma) = \frac{\mathcal{W}}{2M} \sqrt{1 - \frac{\bar{M}}{M}} \frac{v'(\sigma)}{\left(1 + \alpha - \frac{v(\sigma)}{M}\right)^{\frac{3}{2}}} \ ,$$

we have $\theta''(\sigma_c^i) \neq 0$, and consequently the two critical points are non-degenerate. Instead, when the inequality (42) is satisfied at its lower extremum, we have only one critical point $\sigma_c = \bar{\sigma}$, which is degenerate.

It remains to prove Eq. (44). With no loss of generality, we consider the case $\mathcal{N} > 0, \mathcal{W} < 0$. Setting $\mathcal{N}$ within the phase $\theta(\sigma)$ with $\mathcal{N} = -\mathcal{W}\Delta = \Delta\,|\mathcal{W}|$, we obtain $\theta(\sigma) = |\mathcal{W}|\,\Phi(\sigma)$ with

$$\Phi(\sigma) = \left(\Delta\sigma - \sqrt{1 - \frac{\bar{M}}{M}} \int_0^\sigma \frac{dx}{\sqrt{1 + \alpha - \frac{v(\sigma)}{M}}}\right) .$$

Since, depending on the values of m, the integral

$$\int_0^{2\pi} [s_0(\sigma)]^{m-1} e^{i\theta(\sigma)} d\sigma$$

may not be smooth at $\sigma = 0, 2\pi$, we use the technique called *neutralization* of the extremals. Precisely, choose a small $\mu > 0$, depending possibly on the given Δ_{Max}, but independent of $|\mathcal{W}|$ and Δ. In the following we denote by $k_1, k_2, \ldots$ suitable constants which do not depend on $m, \mathcal{W}, \Delta, \varepsilon$, while they may depend on μ, Δ_{Max}.

We first prove that in the hypothesis of the Lemma there exists a small μ such that both critical points σ_c^j are in $(\mu, 2\pi - \mu)$, and for any $\sigma \in [0, \mu]$, we have

$$\left|\Phi'(\sigma)\right| \geq k_1 \quad , \left|\Phi'(\sigma) s_\alpha(\sigma)\right| \geq k_1 \ . \tag{47}$$

In fact, for the given $\Delta, \mathcal{W}$, the critical points σ_c^i satisfy

$$\sqrt{1 + \alpha - \frac{v(\sigma_c^i)}{M}} = \frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\Delta} \geq \frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\Delta_{Max}}$$

or equivalently

$$\frac{v(\sigma_c^i)}{M} \leq 1 + \alpha - \frac{1 - \frac{\bar{M}}{M}}{\Delta_{Max}^2} \ .$$

Choosing a small μ satisfying

$$\frac{v(\mu)}{M} > 1 - \frac{7}{8}\,\frac{1 - \frac{\bar{M}}{M}}{\Delta_{Max}^2} \tag{48}$$

and using (43), we obtain

$$\frac{v(\sigma_c^i)}{M} \leq 1 + \alpha - \frac{1 - \frac{\bar{M}}{M}}{\Delta_{Max}^2} \leq 1 - \frac{7}{8}\,\frac{1 - \frac{\bar{M}}{M}}{\Delta_{Max}^2} < \frac{v(\mu)}{M}$$

and therefore we have $\sigma_c^i \in (\mu, 2\pi - \mu)$. Then, for all $\sigma \in [0, \mu]$, we have

$$\left|\Phi'(\sigma)\right| \geq \frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\sqrt{1 + \alpha - \frac{v(\sigma)}{M}}} - \Delta \geq \frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\sqrt{1 + \alpha - \frac{v(\mu)}{M}}} - \Delta_{Max} \geq \frac{\sqrt{1 - \frac{\bar{M}}{M}}}{16\sqrt{1 + \alpha - \frac{v(\mu)}{M}}}$$

as soon as

$$\Delta_{Max} \leq \frac{15}{16}\,\frac{\sqrt{1 - \frac{\bar{M}}{M}}}{\sqrt{1 + \alpha - \frac{v(\mu)}{M}}} \ , \tag{49}$$

which is satisfied if

$$\frac{v(\mu)}{M} \geq 1+\alpha-\left(\frac{15}{16}\right)^2 \frac{1-\frac{\bar{M}}{M}}{\Delta^2_{Max}}.$$

From (43) and (48) we obtain

$$1+\alpha-\left(\frac{15}{16}\right)^2 \frac{1-\frac{\bar{M}}{M}}{\Delta^2_{Max}} \leq 1-\frac{7}{8}\frac{1-\frac{\bar{M}}{M}}{\Delta^2_{Max}} \leq \frac{v(\mu)}{M} \ .$$

Therefore, for μ satisfying (48), for all $\sigma \in [0, \mu]$ we have

$$\left|\Phi'(\sigma)\right| \geq \frac{\Delta_{Max}}{15}.$$

Analogously, we have

$$\left|\Phi'(\sigma) s_\alpha(\sigma)\right| \geq \sqrt{\frac{2M}{|A|}}\sqrt{1-\frac{\bar{M}}{M}}\left(1-\Delta\frac{\sqrt{1+\alpha-\frac{v(\sigma)}{M}}}{\sqrt{1-\frac{\bar{M}}{M}}}\right) \geq$$

$$\geq \sqrt{\frac{2M}{|A|}}\sqrt{1-\frac{\bar{M}}{M}}\left(1-\Delta_{Max}\frac{\sqrt{1+\alpha-\frac{v(\mu)}{M}}}{\sqrt{1-\frac{\bar{M}}{M}}}\right) \geq \frac{1}{16}\sqrt{\frac{M}{|A|}}\sqrt{1-\frac{\bar{M}}{M}}.$$

Let us now consider an infinitely differentiable window function $\rho(x, \mu)$ such that $\rho(x, \mu) = 0$ for $x \leq \mu/2$ and $x \geq 2\pi - \mu/2$; $\rho(x, \mu) = 1$ for $x \in [\mu, 2\pi - \mu]$. Then, we define

$$\eta(\sigma) = \rho(\sigma; \mu)[s_\alpha(\sigma)]^{m-1}$$

and

$$\hat{I}(|\mathcal{W}|, \Delta) = \int_0^{2\pi} \eta(\sigma) e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma.$$

The integrand of $\hat{I}(|\mathcal{W}|, \Delta)$ is smooth and bounded for all $m \geq 0$, and vanishes at the extrema together with all its derivatives. Therefore it has the form suitable for the application of the rigorous version of PSP (see, for example, [10]). As a consequence, taking into account that the phase $\Phi(\sigma)$ has two non-degenerate critical points σ_c^1, σ_c^2, $\hat{I}(|\mathcal{W}|, \Delta)$ is represented by (44) with a_1, a_2 satisfying the limits (46). It remains to estimate the integral

$$\int_0^{2\pi} s_\alpha(\sigma)^{m-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma - \hat{I}(|\mathcal{W}|, \Delta) = \int_0^{2\pi} (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma$$

$$= \int_0^{\mu} (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma + \int_{2\pi-\mu}^{2\pi} (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma.$$

We prove that there exists a constant κ independent on $|\mathcal{W}|$, Δ and m, such that

$$\left|\int_0^{2\pi} s_\alpha(\sigma)^{m-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma - \hat{I}(|\mathcal{W}|, \Delta)\right| \leq C_m \frac{\kappa}{|\mathcal{W}|} \quad , \quad C_m = \left\{ \begin{array}{cc} \|s_\alpha\|^{m-1} & if m \geq 1 \\ 1 & if m = 0 \end{array} \right. \tag{50}$$

so that we have (44) with a_1, a_2 satisfying the limits (46).

Since the phase $\Phi(\sigma)$ has no stationary points in $[0, \mu]$, and if $m \geq 1$, integrating by parts we obtain

$$\int_0^{\mu} (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma = -\frac{[s_\alpha(0)]^{m-1} e^{i\theta(0)}}{i\,|\mathcal{W}|\,\Phi'(0)} - \int_0^{\mu} \left(\frac{(1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1}}{i\,|\mathcal{W}|\,\Phi'(\sigma)}\right)' e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma. \tag{51}$$

We consider the following cases:

- *If* $m \geq 3$ *or* $m = 1$, using (47) we obtain

$$\left| \frac{[s_\alpha(0)]^{m-1} e^{i\theta(0)}}{i\,|\mathcal{W}|\,\Phi'(0)} \right| \leq |[s_\alpha(0)]|^{m-1} \frac{k_2}{|\mathcal{W}|}.$$

Then we estimate

$$\left| \int_0^\mu \left(\frac{(1-\rho(\sigma;\mu)) s_\alpha^{m-1}(\sigma)}{i\,|\mathcal{W}|\,\Phi'(\sigma)} \right)' e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma \right| \leq \int_0^\mu \left| \frac{\rho'(\sigma;\mu) s_\alpha^{m-1}(\sigma)}{i\,|\mathcal{W}|\,\Phi'(\sigma)} \right| d\sigma$$

$$+ \int_0^\mu \left| \frac{(1-\rho(\sigma;\mu))(m-1) s_\alpha^{m-3}(\sigma) v'(\sigma)}{iA\,|\mathcal{W}|\,\Phi'(\sigma)} \right| d\sigma + \int_0^\mu \left| (1-\rho(\sigma;\mu)) s_\alpha^{m-1}(\sigma) \right| \left| \left(\frac{1}{|\mathcal{W}|\,\Phi'(\sigma))} \right)' \right| d\sigma \,. \tag{52}$$

Using again (47) we obtain

$$\int_0^\mu \left| \frac{\rho'(\sigma;\mu) s_\alpha^{m-1}(\sigma)}{i\,|\mathcal{W}|\,\Phi'(\sigma)} \right| d\sigma + \int_0^\mu \left| \frac{(1-\rho(\sigma;\mu))(m-1) s_\alpha^{m-3}(\sigma)}{iA\,|\mathcal{W}|\,\Phi'(\sigma)} \right| d\sigma \leq \|s_\alpha\|^{m-1} \frac{k_3}{|\mathcal{W}|} \,.$$

Since $1/\Phi'(\sigma)$ is strictly monotone in $[0,\mu]$, its derivative has the same sign in $[0,\mu]$, and therefore we have

$$\int_0^\mu \left| (1-\rho(\sigma;\mu)) s_\alpha^{m-1}(\sigma) \right| \left| \left(\frac{1}{|\mathcal{W}|\,\Phi'(\sigma))} \right)' \right| d\sigma \leq \|s_\alpha\|^{m-1} \int_0^\mu \left| \left(\frac{1}{|\mathcal{W}|\,\Phi'(\sigma))} \right)' \right| d\sigma =$$

$$= \|s_\alpha\|^{m-1} \left| \int_0^\mu \left(\frac{1}{|\mathcal{W}|\,\Phi'(\sigma))} \right)' d\sigma \right| = \|s_\alpha\|^{m-1} \left| \frac{1}{|\mathcal{W}|\,\Phi'(\mu)} - \frac{1}{|\mathcal{W}|\,\Phi'(0)} \right| \leq \|s_\alpha\|^{m-1} \frac{k_4}{|\mathcal{W}|} \,.$$

By collecting all these estimates, we obtain

$$\left| \int_0^\mu (1-\rho(\sigma;\mu)) [s_\alpha(\sigma)]^{m-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma \right| \leq \|s_\alpha\|^{m-1} \frac{k_5}{|\mathcal{W}|}$$

with $k_5 > 0$ independent on $|\mathcal{W}|$, Δ and m.

- *If* $m = 2$, we estimate the border contribution in (51), and the first and the third integrals in (52) as in the case $m \geq 3$. It remains to estimate the second integral, whose denominator is only apparently divergent, since because of (47) we have $|\Phi'(\sigma) s_\alpha(\sigma)| \geq k_1$. Therefore we have

$$\int_0^\mu \left| \frac{(1-\rho(\sigma;\mu))}{i\,|A|\,|\mathcal{W}|\,\Phi'(\sigma) s_\alpha(\sigma)} \right| d\sigma \leq \frac{k_6}{|\mathcal{W}|} \,.$$

- *If* $m = 0$, for any arbitrary small $\xi \in (0,\mu)$, we have

$$\int_\xi^\mu (1-\rho(\sigma;\mu)) [s_\alpha(\sigma)]^{-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma = -\frac{1-\rho(\xi;\mu)}{i\,|\mathcal{W}|\,\Phi'(\xi) s_\alpha(\xi)} e^{i|\mathcal{W}|\Phi(\xi)} +$$

$$+ \int_\xi^\mu \frac{\rho'(\sigma;\mu)}{i\,|\mathcal{W}|\,\Phi'(\sigma) s_\alpha(\sigma)} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma - \int_\xi^\mu (1-\rho(\sigma;\mu)) \left(\frac{1}{i\,|\mathcal{W}|\,\Phi'(\sigma) s_\alpha(\sigma)} \right)' e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma.$$

Using (47) we obtain, uniformly on ξ,

$$\left| \frac{1-\rho(\xi;\mu)}{i\,|\mathcal{W}|\,\Phi'(\xi) s_\alpha(\xi)} e^{i|\mathcal{W}|\Phi(\xi)} + \int_\xi^\mu \frac{\rho'(\sigma;\mu)}{i\,|\mathcal{W}|\,\Phi'(\sigma) s_\alpha(\sigma)} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma \right| \leq \frac{k_7}{|\mathcal{W}|} \,.$$

Since $1/(\Phi'(\sigma)s_\alpha(\sigma))$, is strictly monotone in $[0,\mu]$, by proceeding as in the cases $m \geq 1$, we obtain

$$\left| \int_\xi^\mu (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma \right| \leq \frac{k_8}{|\mathcal{W}|}$$

uniformly in ξ, so that

$$\left| \int_0^\mu (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{-1} e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma \right| \leq \frac{k_8}{|\mathcal{W}|}.$$

By repeating the argument to estimate the integral on $[2\pi-\mu, 2\pi]$ we obtain (50). $\square$

Lemma 1 provides an explicit criterion allowing to classify Melnikov integrals as belonging to the categories (I) or (II) depending on the values of of $\mathcal{N}, \mathcal{W}$. Precisely, if $\mathcal{N} \cdot \mathcal{W} > 0$ the phase is considered in the category (I), and the contribution of the corresponding Melnikov integral to the Arnold diffusion will be considered negligible. Instead, if $\mathcal{N} \cdot \mathcal{W} < 0$, and $|\mathcal{N}|/|\mathcal{W}|$ satisfies strictly (42), the phase has two non-degenerate critical points. Then we distinguish two subcases:

- $|\mathcal{N}|/|\mathcal{W}|$ is not too close to its lower extremum, according to to a criterion specified by Lemma 2 below. Then, the phase is considered in the category (II) and the contribution of the corresponding Melnikov integral to the Arnold diffusion can be estimated analytically using (44).

- $\mathcal{N} \cdot \mathcal{W} < 0$ and $\frac{|\mathcal{N}|}{|\mathcal{W}|}$ suitably close to its lower extremum. We find that such a term, while formally 'stationary', contributes to the Melnikov integral similarly as 'quasi-stationary' terms satisfying

$$\frac{|\mathcal{N}|}{|\mathcal{W}|} < \frac{\sqrt{1-\frac{\bar{M}}{M}}}{\sqrt{1+\alpha-\frac{\bar{M}}{M}}} = Q_\alpha. \tag{53}$$

In fact, a careful investigation of the transition of $\frac{|\mathcal{N}|}{|\mathcal{W}|}$ from values higher than Q_α to smaller ones, reveals that the transition corresponds to a degenerate critical point for the phase. The corresponding integral blows to values of order $1/|\mathcal{W}|^{\frac{1}{3}}$, at the transition, and depending linearly on the distance

$$\delta = Q_\alpha - \frac{|\mathcal{N}|}{|\mathcal{W}|}\ , \tag{54}$$

for small $|\delta|$. Thus, these intermediate cases will be considered in the quasi-stationary category (III).

The values of the Melnikov integrals for terms in category (III) are estimated according to the following:

Lemma 2. *Let the potential $v(\sigma)$ have only one local non–degenerate maximum at $\sigma = 0$ and one local non–degenerate minimum at $\sigma = \bar{\sigma}$. Let us consider ε suitably small and $\alpha \in (0, \left\|r^N\right\|)$. For any phase $\theta(\sigma)$ defined by the labels ν, k such that $\mathcal{N} > 0$, $\mathcal{W} < 0$ and*

$$|\mathcal{N}| = |\mathcal{W}| \left(\frac{\sqrt{1-\frac{\bar{M}}{M}}}{\sqrt{1+\alpha-\frac{\bar{M}}{M}}} - \delta \right) \tag{55}$$

with some $\delta > 0$, and defining

$$I(|\mathcal{W}|, \delta) = \int_0^{2\pi} [s_\alpha(\sigma)]^{m-1} e^{i\theta(\sigma)} d\sigma, \tag{56}$$

we have

$$I(|\mathcal{W}|, \delta) = \hat{I}(|\mathcal{W}|, \delta) + \frac{b(|\mathcal{W}|, \delta)}{|\mathcal{W}|} \tag{57}$$

where

$$\left(\frac{\partial^j}{\partial\delta^j}\hat{I}(|\mathcal{W}|, \delta)\right)_{\Big|\delta=0} = e^{i\theta_*} c_j\, |\mathcal{W}|^{\frac{2j-1}{3}} a_j(|\mathcal{W}|) \quad , \quad j \geq 0 \tag{58}$$

$$\lim_{|\mathcal{W}|\to+\infty} a_j(|\mathcal{W}|) = 1, \quad j \geq 0 \tag{59}$$

$$|b(|\mathcal{W}|, \delta)| \leq \kappa \tag{60}$$

with constants κ and c_j independent on $|\mathcal{W}|$ and δ, and $\theta_* = \theta(\bar\sigma)_{|\delta=0}$. In particular, we have

$$c_0 = \frac{\sqrt{3}\Gamma(4/3)\left(\frac{2M}{|A|}(1+\alpha-\frac{\bar M}{M})\right)^{\frac{m-1}{2}}}{\left(\frac{\sqrt{1-\frac{\bar M}{M}}}{12M\left(1+\alpha-\frac{\bar M}{M}\right)^{\frac{3}{2}}}v''(\bar\sigma)\right)^{\frac{1}{3}}} \quad , \quad c_1 = -\frac{\frac{\Gamma(2/3)}{\sqrt{3}}\left(\frac{2M}{|A|}(1+\alpha-\frac{\bar M}{M})\right)^{\frac{m-1}{2}}}{\left(\frac{\sqrt{1-\frac{\bar M}{M}}}{12M\left(1+\alpha-\frac{\bar M}{M}\right)^{\frac{3}{2}}}v''(\bar\sigma)\right)^{\frac{2}{3}}}\ , \tag{61}$$

and, for all $j \geq 2$,

$$|c_j| = \frac{\frac{\Gamma((j+1)/3)}{3}\left(\frac{2M}{|A|}(1+\alpha-\frac{\bar M}{M})\right)^{\frac{m-1}{2}}}{\left(\frac{\sqrt{1-\frac{\bar M}{M}}}{12M\left(1+\alpha-\frac{\bar M}{M}\right)^{\frac{3}{2}}}v''(\bar\sigma)\right)^{\frac{j+1}{3}}}\alpha_j \tag{62}$$

with $\alpha_j \in \{0, 1, \sqrt{3}, 2\}$ depending on j.

Proof of Lemma 2. Let us choose $\mu > 0$ small enough, but independent on $|\mathcal{W}|$, δ and α; let us define an infinitely differentiable window function $\rho(x,\mu)$ such that $\rho(x,\mu) = 0$ for $x \leq \mu/2$ and $x \geq 2\pi - \mu/2$; $\rho(x,\mu) = 1$ for $x \in [\mu, 2\pi-\mu]$. Then, we define

$$\eta(\sigma) = \rho(\sigma;\mu)[s_\alpha(\sigma)]^{m-1}$$

and

$$\hat{I}(|\mathcal{W}|, \delta) = \int_0^{2\pi} \eta(\sigma) e^{i\theta(\sigma)} d\sigma.$$

Let us preliminarly write the phase $\theta(\sigma)$ and its derivative by replacing $\mathcal{N}$ using (55):

$$\theta = \theta(\bar\sigma) + |\mathcal{W}|\left[\left(\frac{\sqrt{1-\frac{\bar M}{M}}}{\sqrt{1+\alpha-\frac{\bar M}{M}}} - \delta\right)(\sigma-\bar\sigma) - \sqrt{1-\frac{\bar M}{M}}\int_{\bar\sigma}^{\sigma}\frac{dx}{\sqrt{1+\alpha-\frac{v(x)}{M}}}\right] \tag{63}$$

$$\theta' = -|\mathcal{W}|\left[\delta + \sqrt{1-\frac{\bar M}{M}}\left(\frac{1}{\sqrt{1+\alpha-\frac{v(\sigma)}{M}}} - \frac{1}{\sqrt{1+\alpha-\frac{\bar M}{M}}}\right)\right] \tag{64}$$

as well as the expansions at $\sigma=\bar\sigma$

$$\theta=\theta(\bar\sigma)-|\mathcal{W}|\left[\delta(\sigma-\bar\sigma)+\frac{\sqrt{1-\frac{\bar M}{M}}}{12M}\frac{v''(\bar\sigma)}{\left(1+\alpha-\frac{\bar M}{M}\right)^{\frac{3}{2}}}(\sigma-\bar\sigma)^3+\mathcal{O}(\sigma-\bar\sigma)^4\right] \tag{65}$$

In the following we denote by $k_1,k_2,\ldots$ suitable constants which do not depend on $m,\mathcal{W},\delta,\varepsilon,\alpha$, while they may depend on μ. Since (for small μ) $\bar\sigma\in(\mu,2\pi-\mu)$, for all $\sigma\in[0,\mu]$ we have (see (64))

$$\left|\theta'(\sigma)\right|\ge|\mathcal{W}|\sqrt{1-\frac{\bar M}{M}}\left(\frac{1}{\sqrt{1+\alpha-\frac{v(\sigma)}{M}}}-\frac{1}{\sqrt{1+\alpha-\frac{\bar M}{M}}}\right)\ge|\mathcal{W}|\,k_1\ , \tag{66}$$

as well as

$$\left|\theta'(\sigma)s_\alpha(\sigma)\right|\ge|\mathcal{W}|\sqrt{1-\frac{\bar M}{M}}\sqrt{\frac{2M}{|A|}}\left(1-\frac{\sqrt{1+\alpha-\frac{v(\sigma)}{M}}}{\sqrt{1+\alpha-\frac{\bar M}{M}}}\right)\ge|\mathcal{W}|\,k_1\ . \tag{67}$$

To estimate

$$\frac{b(|\mathcal{W}|\,,\delta)}{|\mathcal{W}|}=I(|\mathcal{W}|\,,\delta)-\hat I(|\mathcal{W}|\,,\delta)=\int_0^{2\pi}(1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1}e^{i\theta(\sigma)}d\sigma$$

$$=\int_0^{\mu}(1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1}e^{i\theta(\sigma)}d\sigma+\int_{2\pi-\mu}^{2\pi}(1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1}e^{i\theta(\sigma)}d\sigma\ ,$$

we first notice that the phase $\theta(\sigma)$ has no stationary points in $[0,\mu]$ and therefore integrating by parts we obtain

$$\int_0^{\mu}(1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1}e^{i\theta(\sigma)}d\sigma=-\frac{[s_\alpha(0)]^{m-1}e^{i\theta(0)}}{i\theta'(0)}-\int_0^{\mu}\left(\frac{(1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1}}{i\theta'(\sigma)}\right)'e^{i\theta(\sigma)}d\sigma\ . \tag{68}$$

We consider the following cases:

- *If $m\ge3$ or $m=1$,* using (66) we obtain

$$\left|\frac{[s_\alpha(0)]^{m-1}e^{i\theta(0)}}{i\theta'(0)}\right|\le|[s_\alpha(0)]|^{m-1}\,\frac{k_2}{|\mathcal{W}|}.$$

  Then we can estimate

$$\left|\int_0^{\mu}\left(\frac{(1-\rho(\sigma;\mu))s_\alpha^{m-1}(\sigma)}{i\theta'(\sigma)}\right)'e^{i\theta(\sigma)}d\sigma\right|\le\int_0^{\mu}\left|\frac{\rho'(\sigma;\mu)s_\alpha^{m-1}(\sigma)}{i\theta'(\sigma)}\right|d\sigma+$$

$$+\int_0^{\mu}\left|\frac{(1-\rho(\sigma;\mu))(m-1)s_\alpha^{m-3}(\sigma)v'(\sigma)}{iA\theta'(\sigma)}\right|d\sigma+\int_0^{\mu}\left|(1-\rho(\sigma;\mu))s_\alpha^{m-1}(\sigma)\right|\left|\left(\frac{1}{\theta'(\sigma)}\right)'\right|d\sigma\ . \tag{69}$$

  Using (66) we obtain

$$\int_0^{\mu}\left(\left|\frac{\rho'(\sigma;\mu)s_\alpha^{m-1}(\sigma)}{i\theta'(\sigma)}\right|+\left|\frac{(1-\rho(\sigma;\mu))(m-1)s_\alpha^{m-3}(\sigma)}{iA\theta'(\sigma)}\right|\right)d\sigma\le\|s_\alpha\|^{m-1}\,\frac{k_3}{|\mathcal{W}|}\ .$$

Since $1/\theta'(\sigma)$ is strictly monotone in $[0,\mu]$, its derivative has the same sign in $[0,\mu]$, and therefore we have

$$\int_0^\mu \left|(1-\rho(\sigma;\mu))s_\alpha^{m-1}(\sigma)\right| \left|\left(\frac{1}{\theta'(\sigma)}\right)'\right| d\sigma \le \|s_\alpha\|^{m-1} \int_0^\mu \left|\left(\frac{1}{\theta'(\sigma)}\right)'\right| d\sigma =$$

$$= \|s_\alpha\|^{m-1} \left|\int_0^\mu \left(\frac{1}{\theta'(\sigma)}\right)' d\sigma\right| = \|s_\alpha\|^{m-1} \left|\frac{1}{\theta'(\mu)} - \frac{1}{\theta'(0)}\right| \le \|s_\alpha\|^{m-1} \frac{k_4}{|\mathcal{W}|} \ .$$

By collecting all these estimates, we obtain

$$\left|\int_0^\mu (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{m-1} e^{i\theta(\sigma)} d\sigma\right| \le \|s_\alpha\|^{m-1} \frac{k_5}{|\mathcal{W}|} \ .$$

- *If* $m = 2$, we estimate the border contribution in (68), and first and the third integral in (69) as in the case $m \ge 3$. It remains to estimate the second integral, whose denominator is only apparently divergent, since because of (67) we have $|\theta'(\sigma)s_\alpha(\sigma)| \ge |\mathcal{W}|\, k_1$. Therefore we have

$$\int_0^\mu \left|\frac{(1-\rho(\sigma;\mu))}{i\,|A|\,\theta'(\sigma)s_\alpha(\sigma)}\right| d\sigma \le \frac{k_6}{|\mathcal{W}|}.$$

- *If* $m = 0$, for any arbitrary small $\xi \in (0,\mu)$, we have

$$\int_\xi^\mu (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{-1} e^{i\theta(\sigma)} d\sigma = -\frac{1-\rho(\xi;\mu)}{i\theta'(\xi)s_\alpha(\xi)} e^{i\theta(\xi)} +$$

$$+ \int_\xi^\mu \frac{\rho'(\sigma;\mu)}{i\theta'(\sigma)s_\alpha(\sigma)} e^{i\theta(\sigma)} d\sigma - \int_\xi^\mu (1-\rho(\sigma;\mu)) \left(\frac{1}{i\theta'(\sigma)s_\alpha(\sigma)}\right)' e^{i\theta(\sigma)} d\sigma.$$

Using (67) we obtain uniformly on ξ

$$\left|\frac{1-\rho(\xi;\mu)}{i\theta'(\xi)s_\alpha(\xi)} e^{i\theta(\xi)} + \int_\xi^\mu \frac{\rho'(\sigma;\mu)}{i\theta'(\sigma)s_\alpha(\sigma)} e^{i\theta(\sigma)} d\sigma\right| \le \frac{k_7}{|\mathcal{W}|}.$$

Since $1/(\Phi'(\sigma)s_\alpha(\sigma))$, is strictly monotone in $[0,\mu]$, by proceeding as in the cases $m \ge 1$, we obtain

$$\left|\int_\xi^\mu (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{-1} e^{i\theta(\sigma)} d\sigma\right| \le \frac{k_8}{|\mathcal{W}|} \ ,$$

uniformly in ξ, so that

$$\left|\int_0^\mu (1-\rho(\sigma;\mu))[s_\alpha(\sigma)]^{-1} e^{i\theta(\sigma)} d\sigma\right| \le \frac{k_8}{|\mathcal{W}|} \ .$$

By repeating the argument to estimate the integral on $[2\pi-\mu, 2\pi]$ we obtain that there exists a constant κ independent on $|\mathcal{W}|$ and δ such that $|b(|\mathcal{W}|\,,\delta)| \le \kappa$.

Then, since the integral $\hat{I}(|\mathcal{W}|\,,\delta)$ is smooth with respect to δ at $\delta = 0$, we have

$$\left(\frac{\partial^j}{\partial\delta^j}\hat{I}(|\mathcal{W}|\,,\delta)\right)\Bigg|_{\delta=0} = (-i)^j\, |\mathcal{W}|^j\, e^{i\theta_*} \int_0^{2\pi} \eta(\sigma)(\sigma-\bar{\sigma})^j e^{i|\mathcal{W}|\Phi(\sigma)} d\sigma$$

where

$$\Phi(\sigma) = \left[\frac{\sqrt{1-\frac{\bar{M}}{M}}}{\sqrt{1+\alpha-\frac{\bar{M}}{M}}}(\sigma-\bar{\sigma}) - \sqrt{1-\frac{\bar{M}}{M}} \int_{\bar{\sigma}}^{\sigma} \frac{dx}{\sqrt{1+\alpha-\frac{v(\sigma)}{M}}} \right].$$

Therefore, by using the degenerate version of the principle of stationary-phase (see [30]), by identifying in $|\mathcal{W}|$ the large parameter, and by considering that $\eta(\sigma)(\sigma-\bar{\sigma})^j$ vanishes with all its derivatives at $\sigma = 0, 2\pi$, we obtain (58), (59), (61), (62). □

Lemma 2 allows us to study the transition in the representations of the Melnikov integrals from the regime of stationary phases to the regime of non stationary phases. Hence:

- *In the case* $\delta \geq 0$, a non-stationary phase is considered quasi-stationary, and the corresponding Melnikov integral is computed by approximating $\int_0^{2\pi}[s_\alpha(\sigma)]^{m-1}e^{i\theta(\sigma)}d\sigma$ with

$$I(|\mathcal{W}|, \delta) = c_0 \frac{e^{i\theta_*}}{|\mathcal{W}|^{\frac{1}{3}}} - |c_1|\, \delta e^{i\theta_*} |\mathcal{W}|^{\frac{1}{3}} + \dots \tag{70}$$

  if

$$0 \leq \delta \leq \delta_c = \frac{|c_1|}{c_0} \frac{1}{|\mathcal{W}|^{2/3}}, \tag{71}$$

  otherwise the phase is considered non-stationary and the corresponding Melnikov integral is neglected.

- *In the case* $\delta < 0$, since $I(|\mathcal{W}|, \delta)$ is smooth in δ, at $\delta = 0$, we compare two estimates: one coming from Lemma 1 (stationary phase approximation) and another coming from the extension of the linear law (70) (quasi–stationary phase approximation) to small negative δ. We find that, for negative δ suitably close to 0 the quasi-stationary phase approximation provides a better estimate with respect to the stationary phase approximation. To determine a threshold to decide which one to use, we compared the numerical computation of the integrals with the estimates provided by both Lemmas. Let us, for example, consider $v(\sigma) = \cos\sigma$; for $\alpha = 0$, for all $\sigma \in (0, 2\pi)$ we have:

$$\theta(\sigma) = \theta(\pi) + \mathcal{N}(\sigma-\pi) + 2\mathcal{W} \ln\tan(\sigma/4) \ , \tag{72}$$

  and the integrals (for $\theta(\pi) = 0$):

$$\Delta\mathcal{I} = \int_0^{2\pi} \cos\left(\mathcal{N}(\sigma-\pi) + 2\mathcal{W}\ln\tan(\sigma/4)\right) d\sigma \ . \tag{73}$$

Figure 4 shows the values of the integrals $\Delta\mathcal{I}$, computed numerically, for several values of $\mathcal{W} < 0$, and fixed δ (left panel), or for fixed $\mathcal{W}$ and several values of δ (right panel). The left panel shows that the value of the integrals as computed numerically by solving Eq. (73) (blue dots) is well approximated by the corresponding asymptotic law $1/|\mathcal{W}|^{\frac{1}{3}}$ for the values of $|\mathcal{W}|$ considered (blue line). In the right-panel we compare the numerical computations of (73) (blue dots), with the corresponding estimate provided by the stationary-phase approximation (green curve) and with the linear law (70) (red line), for the sample value $|\mathcal{W}| = 15$ (very similar pictures are obtained for different values). We see that the stationary phase estimates reproduce well the values of the integrals for $\delta \leq -\delta_c/2$. For $\delta \in [-\delta_c/2, 0]$ we have a divergence of the stationary phase approximation formula, indicating

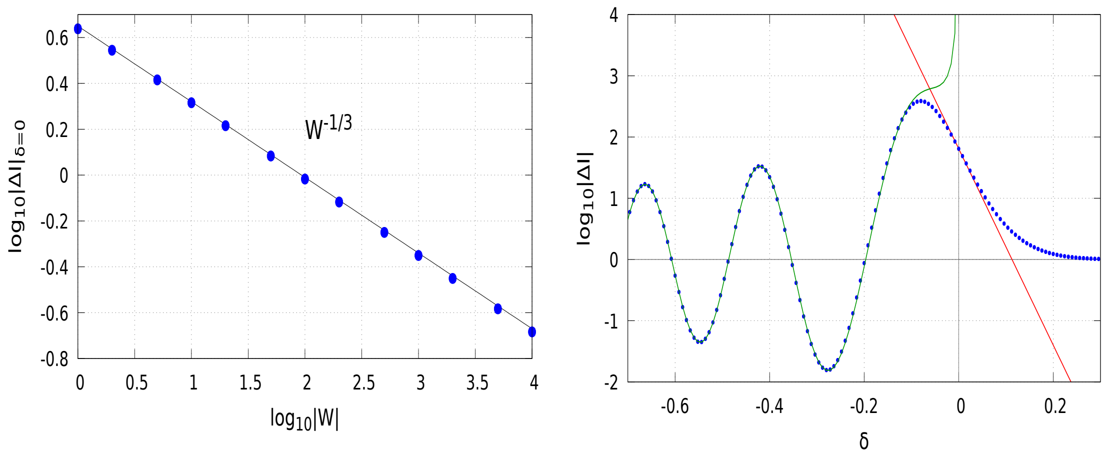

Figure 4: In the left panel we computed numerically the values of $\Delta\mathcal{I}$ defined in (73) for $\delta = 0$ and several values of $|\mathcal{W}|$ (blue points). The blue line corresponds to the asymptotic law $1/|\mathcal{W}|^{\frac{1}{3}}$, that well approximate the values of the integrals for the whole interval of $|\mathcal{W}|$ considered. In the right panel we compare the numerical values of the integrals $\Delta\mathcal{I}$ (the blue dots), with the corresponding estimate provided by the stationary-phase approximation (green curve) and by the linear law (70) (red line), for the sample value $|\mathcal{W}| = 15$ and several values of δ.

that the approximation is no more valid since we are entering the regime of quasi-stationary phase. In fact, we observe that the linear law (70) represents much better the value of the integral for both positive and negative δ in the interval $-\delta_c/2 < \delta < \delta_c$, and therefore, we use Eq. (70) with c_0, c_1 given by Lemma 2 also to estimate those integrals. By using the formula down to $\delta = \delta_c$ we introduce some errors, which could be reduced by considering the non linear corrections (see Remark (xvii) below). On the other hand, δ_c as computed from (70) (represented by the point at which the linear law crosses the x–axis) is clearly underestimated, since the non-linear contributions determine that $\Delta\mathcal{I}$ has a tail extending only asymptotically to zero (see remark (xvii)).

Remarks:

(xvii) The non-linear terms of the expansion (70) provide corrections to the critical value δ_c necessary to discriminate between the quasi-stationary and the non stationary phases.

(xviii) Lemma 1 and 2 are derived by considering the upper branch of the separatrix solution $\theta(\sigma)$ and $\alpha > 0$. Equivalent results are found for the lower branch, and for $\alpha < 0$, after some obvious modifications in the formulas.

3.4 Dependence of all the parameters on ε

In the statements of Lemmas 1 and 2 the parameters N, $\left\|r^N\right\|$, α, $|\mathcal{W}|$, $|\mathcal{N}|$ depend on ε, or are constrained to intervals depending on the value of ε. In this Subsection we specify this dependence by assuming the system in the domain of application of Nekhoroshev's theorem. For simplicity, we discuss the problem in the hypothesis of quasi-convex $h(S, F)$.

The dependence of N on ε. Following [61], Theorem 3, we define a cut-off:

$$N_\varepsilon = \left(\frac{\varepsilon_0}{\varepsilon}\right)^{\frac{1}{2(n-1)}}, \tag{74}$$

and we cover a neighbourhood of radius $\rho = \rho_0\sqrt{\varepsilon}$ of the resonance:

$$\tilde{R} = \left\{(S,F): \ \frac{\partial h}{\partial S} = 0\right\}$$

with open sets $\mathcal{G}_{\mathcal{M}}$ where one constructs normal form Hamiltonians $H_{\varepsilon,\mathcal{M}}$ which are non-resonant with respect to some N_ε–lattice $\mathcal{M} \subseteq \mathbb{Z}^n$ containing the lattice Λ generated by $(1, 0, \dots, 0)$. In particular, for $\mathcal{M} = \Lambda$, there is an open set $\mathcal{G}^*_\Lambda$ containing the point (S_*, F_*) along the resonance, such that for $(S, F, \sigma, \varphi) \in \mathcal{G}^*_\Lambda \times \mathbb{T}^n$ we have the normal form Hamiltonian H^N (see (10)) with $N = N_\varepsilon$. Therefore, in this subsection we identify the optimal normalization order N with the cut-off N_ε.

The dependence of $\|r^N\|$ on ε. According to the proof of the Nekhoroshev theorem, for $N = N_\varepsilon$ the norm of the remainder within the domain $\mathcal{G}^*_\Lambda \times \mathbb{T}^n$ is bounded by:

$$\left\|r^N\right\| \le e^{-N_\varepsilon \sigma_0}\mathcal{F},$$

where $\sigma_0, \mathcal{F}$ are positive parameters independent of ε.

The dependence of α on ε. Any solution of Hamilton's equations of the complete normal form Hamiltonian:

$$H = \overline{H}^N(S, F, \sigma) + r^N(S, F, \sigma, \phi)$$

has the Hamiltonian H as first integral. Therefore, after any complete circulation of the resonant variables S, σ, provided all variables remain in the domain $\mathcal{G}^*_\Lambda \times \mathbb{T}^n$, we have a variation of $\overline{H}^N(S, F, \sigma)$ satisfying

$$\left|\Delta\overline{H}^N\right| = \left|\Delta r^N\right| \le 2\left\|r^N\right\|.$$

Therefore, even for an initial condition very close on the separatrix of $\overline{H}^N(S, F, \sigma)$ (and so with $\alpha \sim 0$), for the next loop we have to consider a level value of $\overline{H}^N$ not larger than $2\left\|r^N\right\|$; as a consequence, we limit our considerations to a small interval of α proportional to $\left\|r^N\right\|$. Precisely, in both lemmas we assume $\alpha \in (0, \left\|r^N\right\|)$.

The dependence of $\mathcal{W}$ on ε. As a consequence of the Geometric Lemma of [61], since by hypothesis $(S_*, F_*) \in \tilde{\mathcal{R}} \cap \mathcal{G}_\Lambda$, from (7) and (74) there is a constant $\gamma > 0$ such that for any $k \in \mathbb{Z}^{n-1}\backslash 0$ with $|k| \le N_\varepsilon$ we have

$$|k \cdot \omega_*| \ge \gamma N_\varepsilon \sqrt{\varepsilon}$$

implying

$$|\mathcal{W}| \ge \frac{\gamma}{\sqrt{2\,|A|\,(M - \overline{M})}} N_\varepsilon. \tag{75}$$

Correspondingly, $|\mathcal{W}|$ is large for all small ε.

Solvability of inequalities (42). We consider the solvability of inequalities (42) for some vectors ν, k such that $|\nu| + |k| \le \hat{N}_\varepsilon$, at any small values of ε. We first notice that, since α is exponentially small with respect to ε, the left-hand side is satisfied by all the vectors ν, k such that $|\nu| + |k| \le N_\varepsilon$, and the factor $\sqrt{1 - \frac{\overline{M}}{M}}/\sqrt{1 + \alpha - \frac{\overline{M}}{M}}$ is close to 1. Since we aim to find solutions to the inequality with the shortest possible values of $|\nu| + |k|$ (corresponding to the larges values of $\left|r^m_{\nu,k}\right|$), the most favorable case is represented by the vectors k with $\mathcal{W} \sim \frac{c}{\sqrt{2|A|(M-\overline{M})}} N_\varepsilon$ and the ν, k such that:

$$\left|\nu + \frac{k \cdot B}{A}\right| \sim \frac{\gamma}{\sqrt{2\,|A|\,(M-\overline{M})}} N_\varepsilon. \tag{76}$$

Therefore, we have the opportunity of finding terms satisfying (42) for any small values of ε within the terms with $|\nu| + |k| \leq N_\varepsilon$, provided that $|k \cdot \omega_*| \sim \gamma N_\varepsilon \sqrt{\varepsilon}$ and and $|\nu| \sim N_\varepsilon$. Moreover, since $|\nu| + |k| \sim N_\varepsilon$ we expect to find a small number of these terms.

Remark:

(xix) We also discuss terms ν, k with $N_\varepsilon < |k| \leq \hat{N}_\varepsilon$, where $\hat{N}_\varepsilon$ is a second cut off larger than N_ε (in practice N_ε is defined by the maximum truncation order of the remainder r^N, which must be such as to ensure the practical convergence of the remainder, see [27] and Section 5 below). By fixing a large parameter W_0, within all these terms, if

$$|k \cdot \omega_*| \geq W_0 \sqrt{2\,|A|\,(M-\overline{M})}\sqrt{\varepsilon},$$

we have $|\mathcal{W}| \geq W_0$ and we have the opportunity to satisfy (42) with $|\nu| \geq W_0$. Therefore we remain with the terms with ν, k such that $N_\varepsilon < |k| \leq \hat{N}_\varepsilon$ and such that the orthogonal projection $P_{\omega_*} k$ of the vector k on the frequency vector ω_* satisfies:

$$\|P_{\omega_*} k\| = \frac{|k \cdot \omega_*|}{\|\omega_*\|} < \frac{W_0}{\|\omega_*\|}\sqrt{2\,|A|\,(M-\overline{M})}\sqrt{\varepsilon}.$$

Again, we have a small subset of all the terms of the remainder r^{N_ε} in this condition, and the few corresponding Melnikov integrals can be numerically evaluated in short CPU time.

4 A semi-analytic solution to Problem 1

Following the analytical results of Section 3, the semi-analytic solution to Problem 1 that we provide in this paper goes through the following steps.

Semi-analytic representation of ΔF_j during a resonant circulation. On the basis of Lemma 1 and Lemma 2 we first define the algorithm which approximates the variation of the adiabatic actions ΔF_j in the time interval $[0, T_\alpha]$ ($t = 0$ is chosen so that $\sigma(0) = 0$ and T_α is the circulation period of the dynamics of the resonant normal forms, see Section 2), with the function (see Eq. (24))

$$\begin{aligned} \Delta F_j : \mathbb{T}^{n-1} &\to \mathbb{R} \\ \phi(0) &\longmapsto \sum_{m,\nu,k} f_{j,m,\nu,k} e^{ik\cdot\phi(0)} \end{aligned} \tag{77}$$

where the coefficients $f_{j,m,\nu,k}$ are provided as floating point numbers obtained by replacing the integral in

$$f_{j,m,\nu,k} = -ik_j \frac{r^m_{\nu,k}\varepsilon^{\frac{m-1}{2}}}{A} \int_0^{2\pi} [s_\alpha(\sigma)]^{m-1} e^{i\theta_{\nu,k}(\sigma;\varepsilon)} d\sigma \tag{78}$$

according to the following fast algorithm:

- For any m, ν, k such that $r^m_{\nu,k} \neq 0$ compute $k \cdot \omega_*$, $\mathcal{N}$ and $\mathcal{W}$ (see (37), (38)).
- If $\mathcal{N} \cdot \mathcal{W} > 0$, or if $|k \cdot \omega_*| > 1$, set $f_{j,m,\nu,k} = 0$. If $\mathcal{N} \cdot \mathcal{W} < 0$ and if $|k \cdot \omega_*| \leq 1$, check if condition (42) and $\delta < -\delta_c/2$ are satisfied (δ defined in Eq. (54), δ_c in Eq. (71)). In such a case, compute $f_{j,m,\nu,k}$ by replacing the integral with its asymptotic expression as indicated in Section 3, Lemma 1.
- If $-\delta_c/2 \leq \delta < \delta_c$ compute $f_{j,m,\nu,k}$ by replacing the integral with its asymptotic expression as indicated in Eq. (70), with the coefficients c_0 and c_1 given in Section 3, Lemma 2. Otherwise, if $\delta \geq \delta_c$ set $f_{j,m,\nu,k} = 0$.

Randomization of the phases, a refinement of equation (78). The above representation is obtained by first approximating the integrals in (24) with Melnikov integrals, and then by computing the Melnikov integrals using Lemmas 1 and 2. Then, we describe the long–term diffusion of the actions caused by a sequence of resonant circulations by applying iteratively formula (78) and by updating the values of the phases $\phi(0)$ at the beginning of each circulation, assuming a random variation. In fact, during any resonant circulation, the angles $\phi(t)$ deviate from the approximation considered in the Melnikov integrals. Since the dynamics is chaotic and the phases $\phi(t)$ are fast, we expect that this deviation is random. The small errors introduced by the randomization of the phases during a circulation period is reduced if we split the period $[0, T_\alpha]$ into two time intervals:

$$[0, \bar{T}_\alpha] \quad , \quad [\bar{T}_\alpha, T_\alpha]$$

where $\sigma(\bar{T}_\alpha) = \bar{\sigma}$, and we compute two semi-analytic formulas:

$$\sum_{m,\nu,k} f^1_{j,m,\nu,k} e^{ik\cdot\phi(0)} \quad , \quad \sum_{m,\nu,k} f^2_{j,m,\nu,k} e^{ik\cdot\phi(\bar{T}_\alpha)} \tag{79}$$

representing the change of the adiabatic actions in the first part of the resonant circulation ($\sigma(t) \in [0, \bar{\sigma}]$) and in the second part ($\sigma(t) \in [\bar{\sigma}, 2\pi]$) respectively. The value of the phases ϕ are then updated also when $\sigma = \bar{\sigma}$.

The reason for this improvement is the symmetry of the distribution of the critical points with respect to the minimum $\bar{\sigma}$, so that the change of the actions is really split into two well differentiated parts, the first one taking place before $\bar{\sigma}$ and the second one taking place after $\bar{\sigma}$.

Largest ΔF_j during a resonant libration. From the numerical values of the coefficients $f_{j,m,\nu,k}, f^i_{j,m,\nu,k}$, one can estimate the maximum variation ΔF_j during a resonant libration by computing the maximum of the function

$$\left| \sum_{m,\nu,k} f_{j,m,\nu,k} e^{ik\cdot\phi(0)} \right| \tag{80}$$

with respect to all the possible initial values of the phases $\phi(0)$.

We remark that to obtain results which compare to the numerical experiments we are not allowed to replace the maximum of the series (80) with the value of the majorant series $\sum_{m,\nu,k} |f_{j,m,\nu,k}|$. In fact, there are examples (see Section 5) where the majorant series is one order of magnitude larger than (80).

Orbits with the fastest long-term instability, ballistic diffusion. The long-term instability of an orbit may arise from a sequence of circulations/librations of S, σ, which produce very small jumps of F_j and α, while the phases ϕ are treated as random variables. Since $\Delta F_j, \Delta\alpha$ are very small at each step, their variations along several circulations/librations are mainly determined by the values of the phases ϕ at the beginning of each circulation/libration. Random variation of the phases yields a random walk of F_j and, by selecting an initial condition such that the values of the phases at each circulation/libration produce the maximum ΔF_j, we obtain a monotonic ballistic motion along the resonance. The conditions to observe these ballistic motions from swarms of $\mathcal{K}$ diffusive orbits are determined as follows: by assuming a randomization of the phases occurring at each resonant libration (random phase approximation), half of the orbits will have $\Delta F_j \geq 0$ and the other half $\Delta F_j < 0$, within the range determined by $|\Delta F_j(T_\alpha)|$ computed as indicated in the previous step. Therefore, we observe orbits with ΔF_j of the same sign for a number of $\mathcal{M}$ randomizations as soon as $\mathcal{K}/2^{\mathcal{M}} \geq 1$. Correspondingly, given $\mathcal{K}$, we observe orbits with F_j which increases (or decreases) almost monotonically in time for a time interval $(\log \mathcal{K}/\log 2)T_\alpha$. By denoting with 10^{-p} the precision of the numerical integration, this time interval is bounded by $(p/\log 2)T_\alpha$.

The speed of the ballistic diffusion in the sequence of resonant librations is represented by

$$\frac{\sum_{i=1}^{\mathcal{M}} \left|\Delta F_j^{(i)}\right|}{\sum_{i=1}^{\mathcal{M}} T_{\alpha^{(i)}}},$$

where we estimate the variation $\left|\Delta F_j^{(i)}\right|$ occurring at the i–th step by the maximum value computed using the semi-analytic theory previously indicated. The sum of the libration periods $T_{\alpha^{(i)}}$ is instead estimated by assuming an average period needed by any libration

$$T_\alpha = \int_0^{2\pi} \frac{d\sigma}{\sqrt{\varepsilon}\sqrt{2|A|(M(1+\alpha) - v(\sigma))}} \tag{81}$$

with $\alpha = \left\|r^N\right\|$. In the case $v(\sigma) = \varepsilon M \cos\sigma$, we find

$$T_\alpha = \frac{1}{\sqrt{A\varepsilon M}} \ln \frac{32 A\varepsilon M}{\|r^N\|} \tag{82}$$

obtained when the energy of the libration differs from the separatrix value by the norm of the remainder. Therefore, we obtain the formula for the average speed of the ballistic diffusion as

$$\mathcal{D} = \max_j \frac{\max_{\phi(0)} |\Delta F_j|}{\frac{1}{\sqrt{A\varepsilon M}} \ln \frac{32 A\varepsilon M}{\|r^N\|}} \quad . \tag{83}$$

Numerical computation of $\Delta F_j(T)$ during a resonant libration. The analytic formulas provided above allow us to compute the maximum speed of diffusion along the resonance. If we are also interested in following, for any given value of the phase $\phi(0)$ at the beginning of the resonant libration, the individual variation $\Delta F_j(T)$ for all $T \in [0, T_\alpha]$, it is possible to compute numerically the function

$$\begin{aligned} \Delta F_j^N : [0, T_\alpha] \times \mathbb{T}^{n-1} &\to \mathbb{R} \\ (T, \phi(0)) &\longmapsto \sum_{m,\nu,k} f_{j,m,\nu,k}(T) e^{ik\cdot\phi(0)} \end{aligned} \tag{84}$$

where the coefficients

$$f_{j,m,\nu,k}(T) = -ik_j \frac{r_{\nu,k}^m \varepsilon^{\frac{m-1}{2}}}{A} \int_0^{\sigma_0(T)} [s_\alpha(\sigma)]^{m-1} e^{i\theta_{\nu,k}(\sigma;\varepsilon)} d\sigma \tag{85}$$

are obtained, only for the terms in the category (II) or (III), by evaluating numerically the integrals in (85). For example, in the $n = 3$ degrees of freedom system considered in Section 5, the terms in the category (II) or (III) are just 1/1000 of terms of the remainder, and the numerical computation of all these integrals is well within the possibility of modern computers.

Remark: (xx) Formula (83) has been obtained from the analysis of the optimal normal form constructed as indicated in the Nekhoroshev theorem. Therefore it represents an improvement of the a priori estimate obtained from the same normal form, for the diffusion along the resonances of multiplicity 1. In the examples of Section 5 the improvement is of some orders of magnitude.

5 Numerical demonstrations on a three degrees of freedom steep Hamiltonian model

We illustrate our theory for the 3-degrees of freedom Hamiltonian (3)

$$H_\varepsilon = \frac{I_1^2}{2} - \frac{I_2^2}{2} + \frac{I_2^3}{3\pi} + 2\pi I_3 + \frac{\varepsilon}{\cos\varphi_1 + \cos\varphi_2 + \cos\varphi_3 + 4} \ ,$$

satisfying the hypotheses of the Nekhoroshev theorem (H_0 is steep and the perturbation is analytic) close to $I_* = (0.664887, 0.955495, 1)$ in the $\ell = (1,1,0)$ resonance and to $I_* = (1.510988, 0.630, 1)$ in the $\ell = (1,3,0)$ resonance.

The upper value of ε. Using the method of the Fast Lyapunov Indicator (see [31, 43], FLI hereafter) we preliminary checked that for the largest value of ε that we considered in our experiments the resonance $\mathcal{R}_\ell$ close to I_* is embedded in a domain dominated by regular motions, with the other resonances forming a web, a circumstance ensuring that the diffusion occurs mainly along the resonance $\mathcal{R}_\ell$. Evidently, the condition persists for smaller values of ε.

Computation of the normal form using a HNA. For a sample of values of ε we computed the normal form of Hamiltonian (3) by implementing the HNA described in [27]. For example, by following the notations of Lemma 2 for the resonance determined by $\ell = (1,1,0)$ (similar analysis can be done for the resonance $\ell = (1,3,0)$), we preliminary define the canonical transformation

$$(\tilde{S}, \tilde{F}_1, \tilde{F}_2) = \Gamma^{-T}(I_1, I_2, I_3) = (I_2, I_2 - I_1, I_3)$$

$$(\tilde{\sigma}_1, \tilde{\phi}_1, \tilde{\phi}_2) = \Gamma(\varphi_1, \varphi_2, \varphi_3) = (\varphi_1 + \varphi_2, -\varphi_1, \varphi_3)$$

with

$$\Gamma = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and then, implementing the HNA, we obtain a canonical transformation

$$(S, F, \sigma, \phi) = \tilde{\mathcal{C}}(\tilde{S}, \tilde{F}, \tilde{\sigma}, \tilde{\phi}) \tag{86}$$

conjugating the Hamiltonian to the normal form (10)

$$H^N = h(S, F) + \varepsilon f^N(S, F, \sigma) + r^N(S, F, \sigma, \phi)$$

with optimal normalization order N depending on the specific value of ε. For all the details about the HNA we refer the reader to [27]. Nevertheless we provide below some details about the output of the algorithm for the case treated in this paper.

- *Truncation order, optimal normalization order, optimal reminder.* Since the HNA is implemented on a computer algebra system, any function $Z(S, F, \sigma, \phi)$ is stored in the memory of the computer as a Taylor–Fourier expansion defined by its series of terms[4]

  $$(S - S_*)^m (F_1 - F_{*,1})^p e^{i(\nu\sigma + k\cdot\phi)},$$

  truncated to some suitably large truncation order. To define the truncation order, as well as other orders within the algorithm, the series is modified by multiplying each term by

  $$\xi^{\left(m+p+\frac{2\mu(|\nu|+|k_1|+|k_2|)}{\ln(1/\varepsilon)}\right)}$$

  where ξ is a formal parameter (which at the end of the computation will be set equal to 1), and μ is defined so that the perturbation is analytic in the complex domain $\{\varphi : \ |\Im\varphi_j| \le \mu\}$. Then we represent the modified series $\mathcal{Z}(S, F, \sigma, \phi, \xi)$ obtained in this way as a Taylor expansion with respect to the parameter ξ

  $$\mathcal{Z} = \sum_{j=1}^{\mathcal{J}} \xi^j \mathcal{Z}_j(S, F, \sigma, \phi)$$

  truncated at some suitable order $\mathcal{J}$. The truncation order of Z is decided as the truncation order of the Taylor expansion of $\mathcal{Z}$ with respect to the parameter ξ.

  The expansions in the formal parameter ξ are used also to define the optimal normalization order. In fact, if we consider all the intermediate Hamiltonians which are constructed within the algorithm

  $$H^i = h(S, F) + \varepsilon f^i(S, F, \sigma) + r^i(S, F, \sigma, \phi) \quad , \quad i = 1, \ldots, N,$$

  any remainder r^i has a truncated Taylor expansion in the formal parameter

  $$\mathcal{R}^i = \sum_{j=\mathcal{J}_i}^{\mathcal{J}} \xi^j \mathcal{R}^i_j(S, F, \sigma, \phi)$$

  starting from a minimum order $\mathcal{J}_i$ such that $\mathcal{J}_{i+1} = \mathcal{J}_i + 1$. The optimal value of N is then chosen so that

  $$\|r^1\| > \|r^2\| > \ldots > \|r^{N-1}\| \ge \|r^N\| \quad , \quad \|r^N\| < \|r^{N+1}\| .$$

[4] Recall that for the specific Hamiltonian (3) the action $F_2 = I_3$ appears only as an isolated linear term.

A necessary condition for the correct execution of the algorithm is that the truncation order $\mathcal{J}$ is larger than $\mathcal{J}_N$. Therefore, the practical limitation for its implementation is due to the limited memory of the computer to store all the series expansions required by the HNA to work within the truncation order. Since the optimal normalization order increases as ε decreases, for any given computer memory we have a lower bound on the value of ε such that we are able to construct the normal form Hamiltonian. For the practical purpose of this work, we considered a lower bound of $\varepsilon = 0.0005$.

- *Domains of the normal forms.* In order to solve Problem 1 we need to provide an estimate of the norms of the normal form remainders (computed as the series of the absolute values of the Taylor-Fourier coefficients) in a domain of the actions S, F which is bounded, in principle, by order $\sqrt{\varepsilon}$. The numerical bounds of the action variables are chosen, for each value of ε, according to the amplitude of the separatrices of the resonant motions.

- *Estimates on the canonical transformation.* The canonical transformation $\tilde{\mathcal{C}}$ (see (86)) is near to the identity, and in particular the difference $\left|F_j - \tilde{F}_j\right|$ can be uniformly bounded by ε^b (with some $b > 0$ defined as in (2)) which is a quantity much larger than the norm of the optimal remainder r^N. As a consequence, even if we suppress from the normal form the remainder r^N, so that the normalized actions F_j are constants of motion, the non-normalized actions have a variation of order ε^b which cannot be ascribed to the Arnold diffusion which is instead produced by a variation of the normalized actions $\tilde{F}_j$ (the so–called 'deformation' in Nekhoroshev theory).

In Tables 1 and 2 we summarize the values of the orders of truncation and of optimal normalization, as well as the norm of the optimal remainders, computed in two resonances ($\ell = (1,1,0)$ for Table 1 and $\ell = (1,3,0)$ for Table 2) for a sample of values of ε between $\varepsilon = 0.08$ and $\varepsilon = 0.0005$. The computations were performed with double floating point precision for the largest values of ε, and with quadruple floating point precision for the smaller ones. The CPU time required by the execution of the HNA on a modern fast multi-processor workstation ranges from few minutes for $\varepsilon = 0.08$ to some hours for $\varepsilon = 0.0005$ for the $\ell = (1,1,0)$ case; the strongest limitations are due to the RAM memory, since for smaller ε we have a largest number of terms to consider within the order N. We notice that the norm of the optimal remainder spans 9 orders of magnitude in this range of variation of ε for $\ell = (1,1,0)$.

To provide an idea of the efficiency of the normalizing transformations, in Fig. 5 we compare the time evolution of F_1 with the time evolution of $\tilde{F}_1$ for a swarm of solutions with initial conditions in a small neighborhood of the separatrix of the resonant normal form, for $\varepsilon = 0.01$. The solutions $(\tilde{S}(t), \tilde{F}(t), \tilde{\sigma}(t), \tilde{\phi}(t))$ have been obtained from a numerical integration of Hamilton's equations of the original Hamiltonian (3); the evolution of the adiabatic action $F_1(t)$ has been obtained by transforming the numerical solution with the canonical transformation $\tilde{\mathcal{C}}$: $(S(t), F(t), \sigma(t), \phi(t)) = \tilde{\mathcal{C}}(\tilde{S}(t), \tilde{F}(t), \tilde{\sigma}(t), \tilde{\phi}(t))$. In the left panel we see that the variation of $\tilde{F}_1$ produces a swarm of points rapidly oscillating in a band of width 6×10^{-3}, which is due to the terms of order ε^b which bound $\left|F_1 - \tilde{F}_1\right|$. A totally different picture appears in the right panel, where the variation of the normalized action F_1 is represented: in this case the slow time evolution is well defined, characterized by jumps of order 10^{-7} (typical values of long–term diffusion of the action variables for this

ε	ΔS	$\mathcal{J}$	$\mathcal{J}_N$	$\left\lVert r^N \right\rVert$	T_α	$\lvert\Delta F_1\rvert_{Nekh}$
0.08	0.114	9	6	$1.179{\times}10^{-4}$	165.0	$1.95{\times}10^{-2}$
0.05	0.090	9	6	$3.01{\times}10^{-5}$	265.9	$7.99{\times}10^{-3}$
0.02	0.057	10	7	$2.13{\times}10^{-6}$	519.6	$1.11{\times}10^{-3}$
0.01	0.040	12	9	$2.07{\times}10^{-7}$	864.2	$1.78{\times}10^{-4}$
0.008	0.036	13	10	$8.43{\times}10^{-8}$	1028.4	$8.67{\times}10^{-5}$
0.005	0.029	13	10	$1.24{\times}10^{-8}$	1468.5	$1.82{\times}10^{-5}$
0.002	0.018	13	10	$2.85{\times}10^{-10}$	2705.0	$7.70{\times}10^{-7}$
0.001	0.013	13	10	$3.05{\times}10^{-11}$	4216.4	$1.29{\times}10^{-7}$
0.0005	0.009	13	10	$4.01{\times}10^{-12}$	6442.7	$2.58{\times}10^{-8}$

ε	$\lvert\Delta F_1\rvert_{Max}$	(II)+(III)$_1$	$\lvert\Delta F_1\rvert_P$	$\lvert\Delta F_1\rvert_{NP}$	(II)+(III)$_2$	$\lvert\Delta F_1\rvert_{sa}$
0.08	$5.22\ {\times}10^{-4}$	1334	$3.45{\times}10^{-4}$	$3.35{\times}10^{-4}$	648(83)	$5.42{\times}10^{-4}$
0.05	$1.75{\times}10^{-4}$	1124	$1.56{\times}10^{-4}$	$1.32{\times}10^{-4}$	513(48)	$2.28{\times}10^{-4}$
0.02	$1.0{\times}10^{-5}$	1696	$1.2{\times}10^{-5}$	$5.2{\times}10^{-6}$	601(50)	$8.34{\times}10^{-5}$
0.01	$2.08{\times}10^{-7}$	3838	$3.6{\times}10^{-7}$	$2.1{\times}10^{-7}$	1300(202)	$2.91{\times}10^{-6}$
0.008	$7.0{\times}10^{-8}$	5470	$4.8{\times}10^{-8}$	$2.5{\times}10^{-8}$	1703(180)	$4.36{\times}10^{-7}$
0.005	$1.0{\times}10^{-8}$	6476	$1.2{\times}10^{-8}$	$6.7{\times}10^{-9}$	634(201)	$6.79{\times}10^{-9}$
0.002	$2.36{\times}10^{-9}$	7346	$4.38{\times}10^{-9}$	$2.79{\times}10^{-9}$	916(7)	$6.66{\times}10^{-9}$
0.001	$1.08{\times}10^{-9}$	8350	$1.32{\times}10^{-9}$	$3.87{\times}10^{-10}$	885(12)	$7.68{\times}10^{-10}$
0.0005	$2.12{\times}10^{-10}$	9364	$1.87{\times}10^{-10}$	$1.86{\times}10^{-10}$	836(26)	$2.04{\times}10^{-10}$

Table 1: Summary of the numerical experiments on the resonance $\ell = (1, 1, 0)$ of Hamiltonian (3), The upper table reports the parameters of the Hamiltonian normalizing algorithm and some of the informations that we can extract from its output: ΔS denotes the amplitude of the domain in the resonant action S, $\mathcal{J}$ the truncation order, $\mathcal{J}_N$ the optimal normalization order, $\left\lVert r^N \right\rVert$ the norm of the remainder expansion (12) close to $I_* = (0.664887, 0.955495, 1)$, T_α the period of the resonant variables computed using (82); $|\Delta F_1|_{Nekh}$ represents the a *priori* upper bound of the maximum variation of F_1 over a period T_α forced by the remainder r^N. The lower table concerns the numerical computation and the analytic estimates about the variations of the normalized adiabatic action F_1 during a resonant period: $|\Delta F_1|_{Max}$ denotes the maximum variation of F_1 after a full resonant period for a swarm of 100 orbits with initial actions close to I_* obtained from numerical integrations of the Hamilton equations; $|\Delta F_1|_{NP}$ denotes the semi-analytic estimate of the maximum variation obtained by computing numerically the Melnikov integral whose phase is stationary or quasi-stationary (since the numerical computation of the Melnikov integrals is more precise than the linear approximation, we include for safety a larger number of terms in the category (III), by checking directly the value δ_c for which the integrals are negligible with respect to $\delta = 0$; the number of terms, reported in the column (II)+(III)$_1$, is still in a ratio of $1 \sim 1000$ of the total number); $|\Delta F_1|_P$ is analogous to $|\Delta F_1|_N$, but obtained with the 'patched' formula (87); $|\Delta F_1|_{sa}$ is the value obtained using the asymptotic expansions of Lemmas 1 or 2; (II)+(III)$_2$ represents the number of terms included in this computation, in parenthesis the number of stationary terms).

ε	ΔS	$\mathcal{J}$	$\mathcal{J}_N$	$\lVert r^N \rVert$	T_α	$\lvert\Delta F_1\rvert_{Nekh}$
0.01	0.0036	10	7	1.387×10^{-7}	1204.7	5.66×10^{-4}
0.008	0.0032	10	7	7.986×10^{-8}	1394.2	2.42×10^{-4}
0.005	0.0025	11	8	2.582×10^{-8}	1882.9	1.18×10^{-4}
0.002	0.0016	12	9	1.793×10^{-9}	3474.5	1.55×10^{-5}
0.001	0.0011	15	12	1.532×10^{-10}	5620.6	2.15×10^{-5}
0.0009	0.00108	15	12	9.861×10^{-11}	6065.9	1.49×10^{-6}
0.0008	0.00102	16	13	5.581×10^{-11}	6635.6	9.26×10^{-7}

ε	$\lvert\Delta F_1\rvert_{Max}$	(II)+(III)$_1$	$\lvert\Delta F_1\rvert_P$	$\lvert\Delta F_1\rvert_{NP}$	(II)+(III)$_2$	$\lvert\Delta F_1\rvert_{sa}$
0.01	3.80×10^{-5}	2384	1.63×10^{-4}	4.01×10^{-5}	386(308)	2.98×10^{-4}
0.008	4.19×10^{-5}	2388	1.13×10^{-4}	4.45×10^{-5}	384(298)	2.30×10^{-4}
0.005	3.89×10^{-5}	3888	6.44×10^{-5}	3.92×10^{-5}	532(364)	1.42×10^{-4}
0.002	8.90×10^{-6}	6048	8.83×10^{-6}	8.77×10^{-6}	566(402)	2.07×10^{-5}
0.001	7.89×10^{-7}	17110	7.60×10^{-7}	7.49×10^{-7}	1212(934)	4.93×10^{-7}
0.0009	4.77×10^{-7}	17110	4.70×10^{-7}	4.53×10^{-7}	1212(708)	1.66×10^{-7}
0.0008	2.52×10^{-7}	22886	2.56×10^{-7}	2.57×10^{-7}	1300(1000)	2.30×10^{-7}

Table 2: Summary of the numerical experiments in the resonance $\ell = (1, 3, 0)$ of Hamiltonian (3) close to $I_* = (1.510988, 0.630, 1)$ (see the caption of Table 1 for the explanation of the column titles). We have a very good match between the values $|\Delta F_1|_{Max}$ provided by the numerical experiments and the semi-analytic patched or the non-patched estimates $|\Delta F_1|_P$, $|\Delta F_1|_{NP}$; for the larger values of ε the purely analytic estimate $|\Delta F_1|_{sa}$ is affected by the nearby crossing of the resonance $\ell = (1, 3, 0)$ by the higher order resonance $\ell = (25, -2, 0)$.

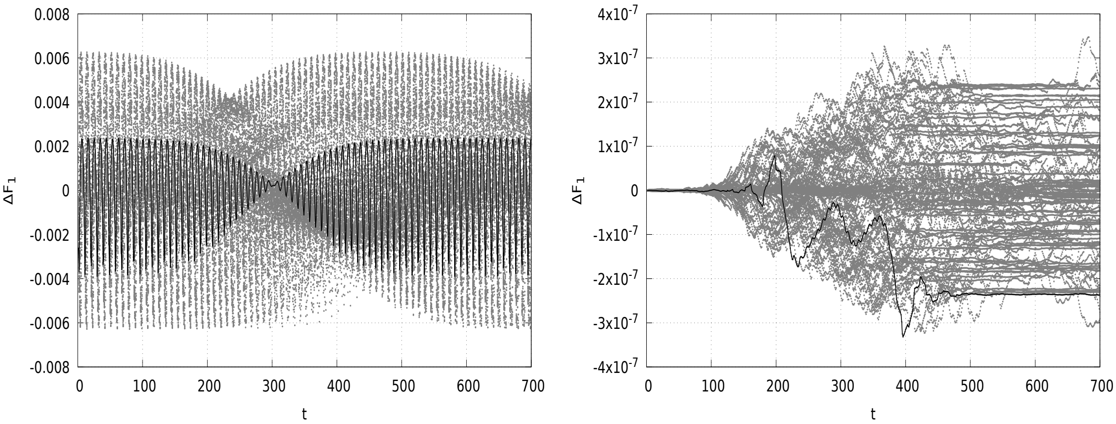

Figure 5: Time evolution of the action $\tilde{F}_1$ (left panel) and of the normalized adiabatic action F_1 (right panel) for the same swarm of 100 solutions with initial conditions in a small neighborhood of the separatrix of the resonance $\ell = (1, 1, 0)$, for $\varepsilon = 0.01$. The bold curves in both panels represent the same sample solution.

value of ε), that are detectable on such time intervals only thanks to the implementation of the normalizing transformation.

Before applying the theory developed in Sections 2, 3, 4, we provide, as in the proof of Nekhoroshev theorem, an upper bound to the variation of the action variables by computing the right-hand side of inequality (15). The upper bound computed for a period (82) is reported in the column $|\Delta F_1|_{Nekh}$ and is larger up to two order of magnitudes with the numerically computed variations $|\Delta F_1|_{Max}$. We therefore proceed by estimating these variations with the Melnikov integrals.

Estimate of ΔF_1 during a resonant libration. Let us analyze more in detail the variation of the adiabatic action F_1. In Fig. 6, as before, we represent the time evolution of $F_1(t)$ during a circulation of the variables σ, S obtained from a numerical integration of Hamilton's equations of (3) for a swarm of 100 orbits, for two sample values of ε. The spread of $F_1(t)$ after the circulation is due to the different values of $\phi(0)$. We are now able to predict the time evolution of *all* these orbits by using the semi-analytic theory developed in Section 3.

Since, due to the discrimination between phases, the number of Melnikov integrals to take into account is now small, we have the opportunity to compute these integrals also numerically for all the intermediate times $t \in [0, T_\alpha]$. For these computations, we can safely extend the value of δ_c computed from the linear approximation as soon as the phases with $\delta > 0$ provide non negligible contributions, and still have a small number of terms (see Table 1, column (II)+(III)$_1$).

The red curves of Fig. 6 represent the orbits yielding the maximum negative jump obtained for the numerical integration of the Hamilton equations, while the black and blue curves represent the Melnikov approximations (without and with the patched formula (87), respectively). One sees that for both values of ε all the curves are sticked up to a time corresponding approximately to half a period of a complete homoclinic loop. In the middle of the homoclinic loop, we distinguish two cases. In the first case the jump is due mostly to remainder terms which become locally stationary at angles σ_c sufficiently far from $\bar{\sigma} = \pi$, while the slope $d\theta/d\sigma$ is substantially larger than unity at $\sigma = \pi$. In such cases, the jumps are localized around the two stationary values symmetric with respect to the middle of the loop, while the associated remainder terms yield a rapid oscillatory evolution of the actions F_1 in between the two jumps. Since the motion is in reality chaotic, the orbits during the rapid oscillations undergo also a randomization of the phases, implying that the predictions obtained by computing (26) may introduce an error. This can be remedied using both representations (79): precisely, the blue curve represents the 'patched' evolution given by:

$$\begin{aligned} \Delta F_1(t) &:= \Delta F_1^N(t) \ \text{ if } \ t < T_\alpha/2 \ , \\ \Delta F_1(t) &:= 2\Delta F_1^N(T_\alpha/2) - \Delta F_1^N(T_\alpha - t) \ \text{ if } \ t \geq T_\alpha/2 \ . \end{aligned} \tag{87}$$

On the other hand, in cases where important quasi-stationary terms enter into play, $\theta_{\nu,k}(\sigma;\varepsilon)$ remains at small values over a large interval around $\sigma = \pi$. Then no rapid oscillations of the fast variables are observed, and the variations become predictable along the whole homoclinic loop using the original estimate (77). In fact, these are cases where the method illustrates its full power, as it is able to capture large cancellations taking place between stationary terms (II), which, however, exhibit near-stationarity in the whole interval between the two (symmetric with respect to π) critical values σ_c, and true quasi-stationary terms (III). An example is provided in Fig 7: the terms of groups (II) and (III) independently produce jumps of order 10^{-5}, that nearly cancel, leaving a residual of order 10^{-7} which fits exactly the numerical evolution of the action F_1. Since no rapid oscillations are observed in the

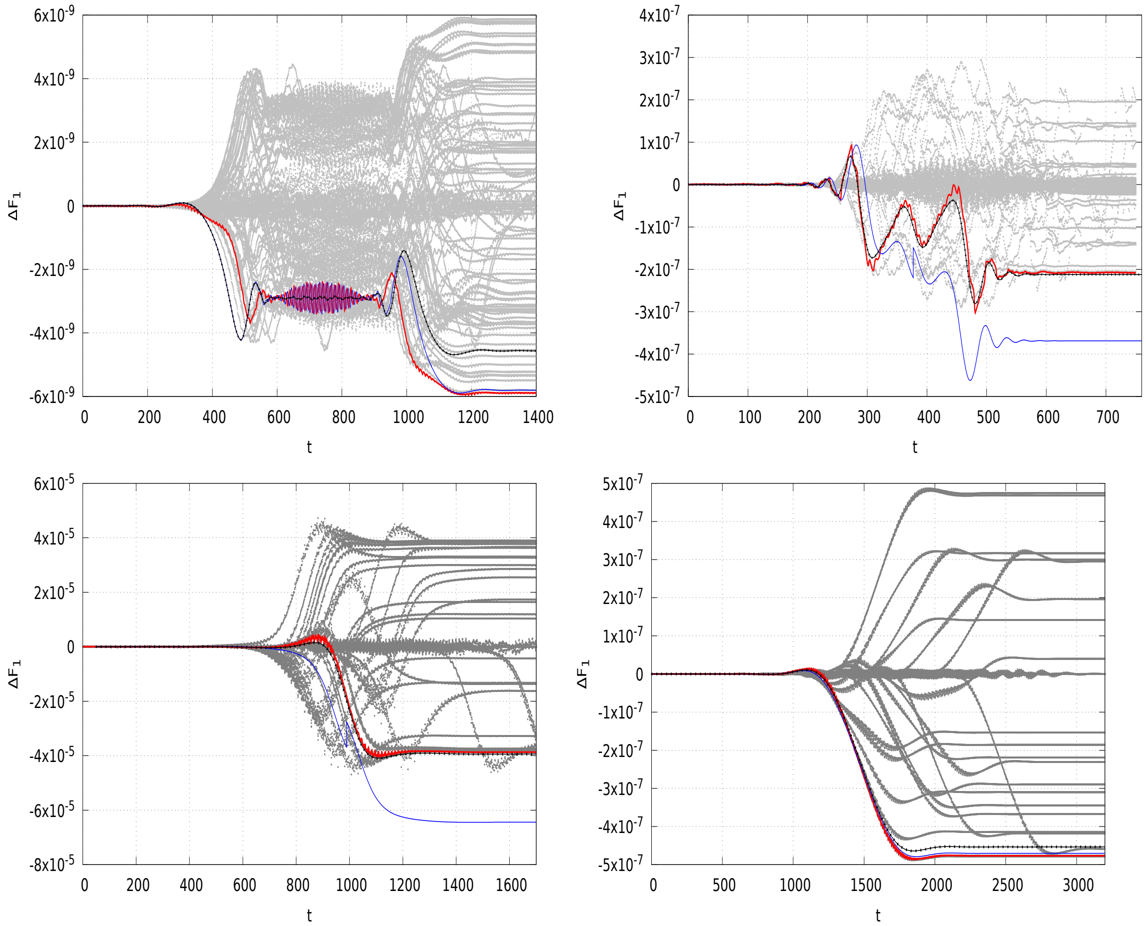

Figure 6: Evolution of the normalized action $F_1(t)$ numerically computed for Hamiltonian (3) for a swarm initial conditions in the resonance $\ell = (1, 1, 0)$ and $\varepsilon = 0.003,\ 0.01$ (top-left and top-right panels resp.) and in the resonance $\ell = (1, 3, 0)$ for $\varepsilon = 0.008, 0.0009$ (bottom-left and bottom-right panels resp.). The initial conditions have been randomly chosen in a two-dimensional square neighbourhood of $(S, \sigma, F, \phi) = (0, 0, F_*, 0, 0)$ (parameterized by ϕ_1, S, and with values of FLI larger than 3 over a time interval of $T = 1000$) and performing a circulation in the S, σ variables. Hamilton's equations have been numerically integrated in the original variables (I, φ); $F_1(t)$ has been then computed from the numerical solution using the canonical transformation defined by the HNA. The red line highlights the evolution with the largest ΔF_1 over a circulation. The (dotted) black and (thin) blue lines show the evolutions obtained by numerically integrating the Melnikov integrals (85) whose phase satisfies (II) or (III), without and with the patched correction (87), respectively.

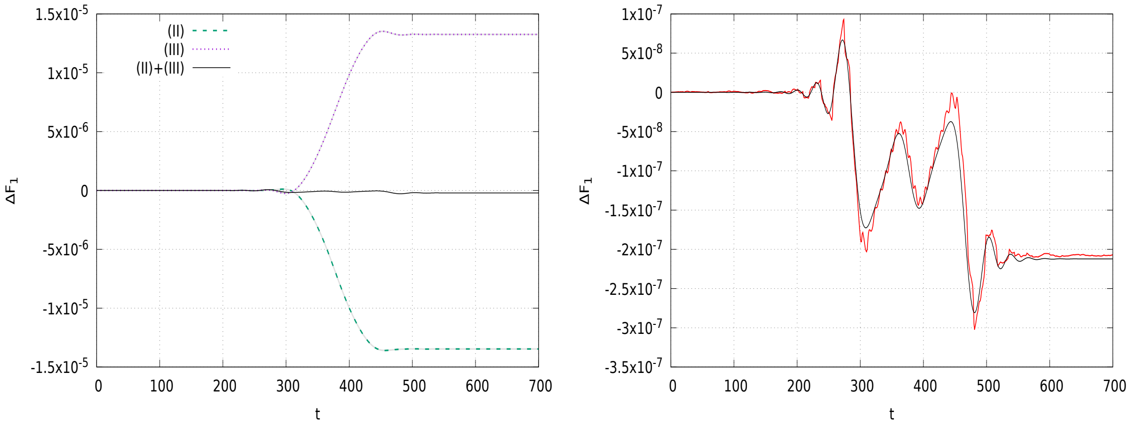

Figure 7: Evolution of ΔF_1 over a circulation of the resonant variables of an orbit in the resonance $\ell = (1, 1, 0)$ for $\varepsilon = 0.01$, by considering the Melnikov integrals whose phase is in the category (II) (dashed green line), in the category (III) (dotted purple line), and the contribution of the two categories together (black line). We notice the cancellations occurring between the Melnikov integrals in the first and second case, which produce a much smaller cumulative variation, represented also in the zoomed right panel.

middle of the homoclinic loop, the non patched estimate is more precise than the patched estimate, as also shown in the right panel of Fig. 6.

In Tables 1 and 2, for several different values of ε, we report the values $|\Delta F_1|_{Max}$ representing the maximum variation of F_1 in a full resonant libration, obtained from the numerical integration of the Hamilton equations for a swarm of 100 orbits with initial actions close to I_* in the two different resonances; $|\Delta F_1|_{NP}$ denotes the semi-analytic estimate of the maximum variation obtained by computing numerically the Melnikov integral whose phase is stationary or quasi-stationary; $|\Delta F_1|_P$ is analogous to $|\Delta F_1|_N$, but obtained with the patched formula (87); $|\Delta F_1|_{sa}$ is the value obtained using the asymptotic expansions of Lemmas 1 or 2. By comparing $|\Delta F_1|_{Max}$ with $|\Delta F_1|_N, |\Delta F_1|_{NP}$ we have a good agreement between the numerical integrations and the predictive model for all the values of ε (we notice that for a given ε, only one of the two values $|\Delta F_1|_N, |\Delta F_1|_{NP}$ is applicable), to within a factor 2 in variations over 6 orders of magnitude as ε varies between 0.0005 and 0.08). The values $|\Delta F_1|_{sa}$ are expected to be slightly less precise than $|\Delta F_1|_N, |\Delta F_1|_{NP}$, since they rely on the linear law (70) for the quasi-stationary cases, and do not take into account the patched formula (87); we expect that the errors can be more important for larger values of ε. Here we have an agreement within a factor 3 as ε varies between 0.0005 and 0.08, except in the interval $0.008 \leq \varepsilon \leq 0.02$, where the cancellations (as in Fig. 7) become important. Regarding the resonance $\ell = (1, 3, 0)$, we also have a very good match between the values $|\Delta F_1|_{Max}$ provided by the numerical experiments and the semi-analytic patched or the non-patched estimates $|\Delta F_1|_P$, $|\Delta F_1|_{NP}$. For the larger values of ε, the purely analytic estimate $|\Delta F_1|_{sa}$ is affected by the nearby crossing of the resonance $\ell = (1, 3, 0)$ by the higher order resonance $\ell = (25, -2, 0)$.

Diffusion and ballistic orbits. As discussed in Section 4, the long-term instability of an orbit may arise from a sequence of circulations/librations of S, σ, which produce very small jumps of F_1 and α, while the phases ϕ are treated as random variables. The random variation of the phases determines the random walk along the resonance in jumps of maximum amplitude estimated according to the theory of Section 3; for special initial conditions the sequence of jumps has the same sign, so that we have the orbits which move along the resonance with the largest speed (ballistic orbits). An illustration of this phenomenon is represented in Fig. 8 where we represent a ballistic orbit through a sequence of 14 circulations, which is the limit of the quadruple precision. The speed of the ballistic orbits numerically measured is in agreement with formula (83) (see Table 1). Note also the overall random walk nature of the jumps ΔF_1 for most other orbits nearby to the ballistic one. Since estimates on ΔF_j can be regarded as providing the one–step size in the random walk, they are crucial in modelling the diffusion process for a large measure of trajectories over times of practical interest in the applications.

Acknowledgments. This research has been supported by: the Research Committee of the Academy of Athens through grants 200/854 and 200/890 (01/09/2016-30/06/2018), and ERC project 677793 Stable and Chaotic Motions in the Planetary Problem (01/07/2018-30/06/2019) (R.P.); project CPDA149421/14 of the University of Padova "New Asymptotic Aspects of Hamiltonian Perturbation Theory" (M.G.). M. Guzzo acknowledges also the project MIUR-PRIN 20178CJA2B titled "New frontiers of Celestial Mechanics: theory and applications". C. Efthymiopoulos acknowledges the hospitality of the University of Padova, under the program Visiting Scientist. R. Paez wishes to thank Gabriella Pinzari for useful discussions. All the authors thank the anonymous referees for useful comments.

References

[1] V.I. Arnold, Instability of dynamical systems with several degrees of freedom, *Sov. Math. Dokl.* **6**, 581 (1964).

[2] G. Benettin, A. Carati, and G. Gallavotti, On the conservation of adiabatic invariants for a system of coupled rotators, *Physica D* **104**, 253 (1997).

[3] G. Benettin, F. Fassoànd M. Guzzo, Long-term stability of proper rotations of the Euler perturbed rigid body, *Commun. Math. Phys.* **250**, 133 (2004).

[4] G. Benettin and G. Gallavotti, Stability of motions near resonances in quasi–integrable Hamiltonian systems, *J. Stat. Phys.* **44**, 293 (1985).

[5] G. Benettin and A. Giorgilli, On the Hamiltonian interpolation of near to the identity symplectic mappings, with application to symplectic integration algorithms, *J. Stat. Phys.* **73**, 1117 (1994).

[6] P. Bernard, V. Kaloshin and K. Zhang, Arnold diffusion in arbitrary degrees of freedom and normally hyperbolic invariant cylinders, *Acta Math.* **217**, 1 (2016).

[7] M. Berti, L. Biasco, and P. Bolle, Drift in phase space: a new variational mechanism with optimal diffusion time, *J. Math. Pures Appl.* **82**, 613 (2003).

[8] M. Berti and P. Bolle, A functional analysis approach to Arnold diffusion, *Ann. Henri Poincaré C: Non Lin. Anal.* **19**, 395 (2002).

[9] U. Bessi, L. Chierchia and E. Valdinoci, Upper bounds on Arnold diffusion times via Mather theory, *J. Math. Pures Appl.* **80**, 105 (2001).

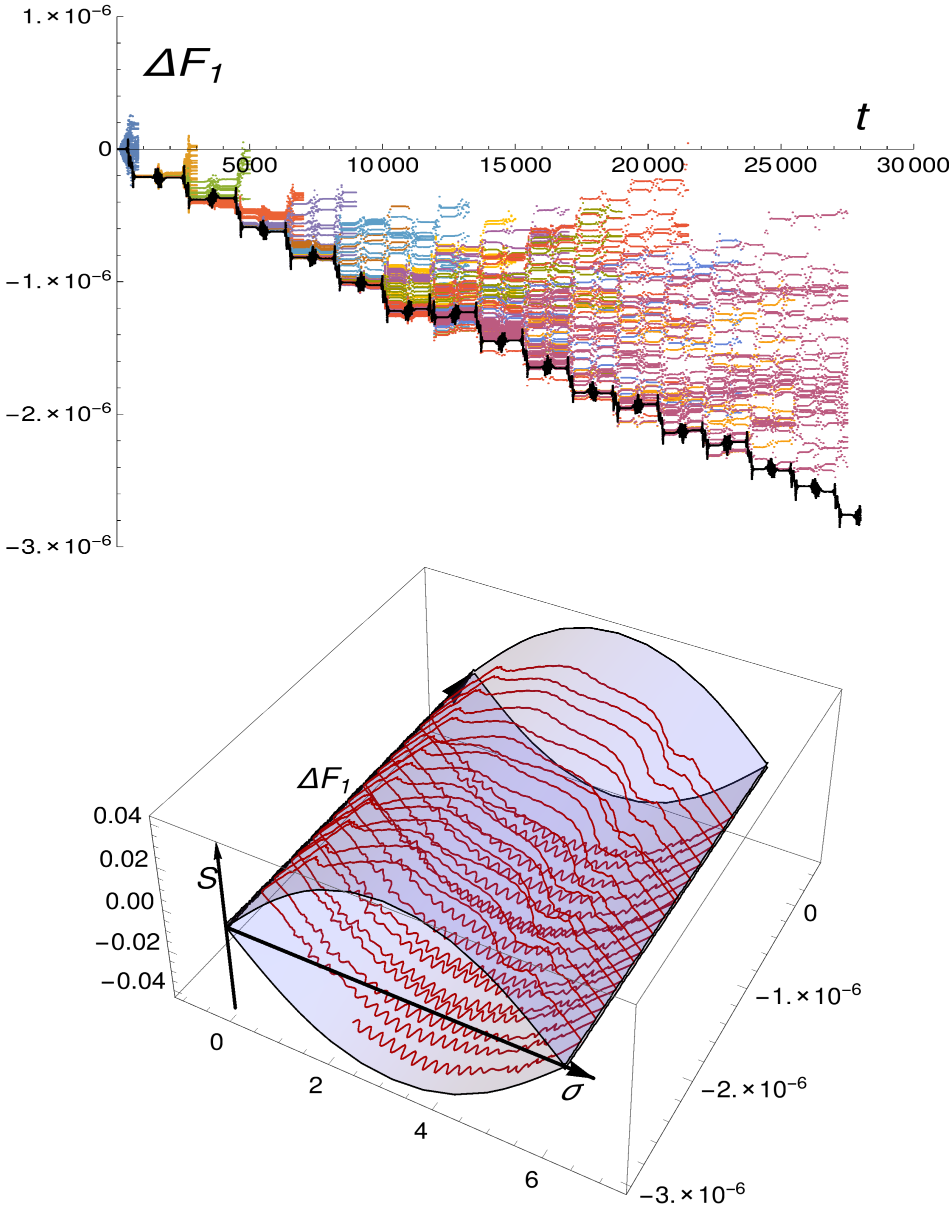

Figure 8: Top panel: $\Delta F_1(t)$ numerically computed for Hamiltonian (3) with $\varepsilon = 0.01$ and different sets of $\mathcal{N} = 16$ initial conditions in the resonance $\ell = (1, 1, 0)$ in a 1-dim grid around $(\hat{S}, \sigma, F, \phi) = (0, 0, F_*, \phi_*, 0)$ of some selected amplitude d^0 around ϕ_*. The initial conditions defined by $d^0 = 10^{-3}$ are numerically integrated until all the orbits undergo one complete circulation/libration. Then, we select ϕ_*^1 in the grid providing the largest value of $|\Delta F_1|$, and define a new set of $\mathcal{N}$ initial conditions in a 1-dim grid of amplitude $d^1 << d^0$ around $(S, \sigma, F, \phi) = (0, 0, F_*, \phi_*^1, 0)$, numerically integrated until all the orbits complete two circulations/librations. Due to precision limits, the iterative refinement of initial conditions is stopped after 14 iterations. Different sets are reported with different colors, clearly illustrating the random dispersion of values of F_1, as well as the orbits with largest variations. The fastest (ballistic) orbit is shown in black. The bottom panel shows its projection in the space S, σ, F_1, thus providing a non-schematic example of ballistic Arnold diffusion. The diffusion speed is of the same order as reported in Table 1.

[10] N. Bleistein and R.A. Handelman R.A., *Asymptotic Expansions of integrals*, Holt, Rinehart and Winston, New York (1975).

[11] A. Bounemoura and V. Kaloshin, Generic fast diffusion for a class of non-convex Hamiltonians with two degrees of freedom, *Mosc. Math. J.* **14**, 181 (2014).

[12] M. Capinski and M. Gidea, Diffusion, quantitative estimates and stochastic behavior in the three-body problem, *arXiv:1812.03665* (2018).

[13] A. Celletti and L. Chierchia, A Constructive Theory of Lagrangian Tori and Computer-assisted Applications, in *Dynamics Reported, New Series, Vol. 4*, Springer, Berlin (1995).

[14] A. Celletti and L. Chierchia, On the stability of realistic three-body problems, *Commun. Math. Phys.* **186**, 413 (1997).

[15] A. Celletti and L. Ferrara, An application of Nekhoroshev theorem to the restricted three-body problem, *Celest. Mech.Dyn. Astron.* **64**, 261 (1996).

[16] L. Chierchia and G. Gallavotti, Drift and diffusion in phase space. *Ann. Inst. Henri Poincaré* **60**, 1 (1994).

[17] L. Chierchia and G. Pinzari, The planetary N-body problem: Symplectic foliation, reductions and invariant tori, *Invent. math.* **186**, 1 (2012).

[18] B.V. Chirikov, An universal instability of many dimensional oscillator system, *Phys. Rep.* **52**, 263 (1979).

[19] C-Q. Cheng and J. Yan, Arnold diffusion in hamiltonian systems: a priori unstable case, *J. Differ. Geom.* **82**, 229 (2009).

[20] P.M. Cincotta, C. Efthymiopoulos, C.M. Giordano, and M.F. Mestre, Chirikov and Nekhoroshev diffusion estimates: Bridging the two sides of the river, *Physica D* **266**, 49 (2014).

[21] A. Delshams, R. de la Llave, and T.M. Seara, 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**, 844 (2006).

[22] A. Delshams, R. de la Llave, and T.M. Seara, Instability of high dimensional Hamiltonian Systems: Multiple resonances do not impede diffusion, *Adv. Math.* **294**, 689 (2016).

[23] A. Delshams, V. Gelfreich, A. Jorba, and T.M. Seara, Exponentially small splitting of separatrices under fast quasi-periodic forcing, *Commun. Math. Phys.* **189**, 35 (1997).

[24] A. Delshams, V. Kaloshin, A. De La Rosa A and T.M. Seara, Global Instability in the Restricted Planar Elliptic Three Body Problem, *Commun. Math. Phys.* **366**, 1173 (2019).

[25] C. Efthymiopoulos, N. Voglis and G. Contopoulos, 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, Amsterdam (1998).

[26] C. Efthymiopoulos and Z. Sándor, Optimized Nekhoroshev stability estimates for the Trojan asteroids with a symplectic mapping model of co-orbital motion, *Mon. Not. R. Astron. Soc.* **364**, 253 (2005).

[27] C. Efthymiopoulos, On the connection between the Nekhoroshev theorem and Arnold diffusion, *Celest. Mech. Dynam. Astron.* **102**, 49 (2008).

[28] C. Efthymiopoulos and M. Harsoula, The speed of Arnold diffusion, *Physica D* **251**, 19 (2013).

[29] A. Enciso., A. Luque and D. Peralta-Salas, D, Stationary phase methods and the splitting of separatrices, *Commun. Math. Phys.* **368**, 1297 (2019).

[30] A. Erdélyi, *Asymptotic Expansions*, Dover Publications, New York (1956).

[31] Cl. Froeschlé, M. Guzzo and E. Lega, Graphical Evolution of the Arnold Web: from order to chaos, *Science* **289**, 2108 (2000).

[32] Cl. Froeschlé, M. Guzzo and E. Lega, Local and global diffusion along resonant lines in discrete quasi–integrable dynamical systems. *Celest. Mech. Dyn. Astron.* **92**, 243 (2005).

[33] V. Gelfreich, C. Simó and A. Vieiro, Dynamics of 4D symplectic maps near a double resonance, *Physica D* **243**, 92 (2013).

[34] V. Gelfreich and A. Vieiro, Interpolating vector fields for near identity maps and averaging, *Nonlinearity* **31**, 4263 (2018).

[35] M. Gidea and J.P. Marco, Diffusion along chains of normally hyperbolic cylinders, *arXiv:1708.08314* (2017).

[36] C.M. Giordano and P.M. Cincotta, Chaotic diffusion of orbits in systems with divided phase–space, *Astron. Astrophys.* **423**, 745 (2004).

[37] A. Giorgilli, A. Delshams, E. Fontich, L. Galgani, and C. Simó, Effective Stability for a Hamiltonian System near an Elliptic Equilibrium Point, with an Application to the Restricted three Body Problem. *J. Diff. Eq.* **77**, 167 (1989).

[38] A. Giorgilli and C. Skokos, On the stability of the Trojan asteroids, *Astron. Astrophys.* **317**, 254 (1997).

[39] A. Giorgilli, U. Locatelli and M. Sansottera, Kolmogorov and Nekhoroshev theory for the problem of three bodies, *Cel. Mech. Dyn. Astron.* **104**, 159 (2009).

[40] M. Guzzo, L. Chierchia and G. Benettin, The steep Nekhoroshev's Theorem and optimal stability exponents (an announcement), *Rend. Lincei-Matematica App.* **25**, 293 (2014).

[41] M. Guzzo, L. Chierchia, and G. Benettin, The steep Nekhoroshev's Theorem, *Commun. Math. Phys.* **342**, 569 (2016).

[42] M. Guzzo, E. Lega, and Cl. Froeschlé, First numerical investigation of a conjecture by N.N. Nekhoroshev about stability in quasi-integrable systems, *Chaos* **21**, 3 (2011).

[43] M. Guzzo and E. Lega, The numerical detection of the Arnold web and its use for long-term diffusion studies in conservative and weakly dissipative systems, *Chaos* **23**, 023124 (2013).

[44] M. Guzzo and E. Lega, The Nekhoroshev theorem and the observation of long-term diffusion in Hamiltonian systems, *Regul. Chaotic Dyn.* **21**, 707 (2016).

[45] M. Guzzo, E. Lega, and Cl. Froeschlé, First Numerical Evidence of Arnold diffusion in quasi–integrable systems, *DCDS B* **5**, 687 (2005).

[46] M. Guzzo, E. Lega, and Cl. Froeschlé, A numerical study of Arnold diffusion in a priori unstable systems, *Commun. Math. Phys.* **290**, 557 (2009).

[47] M. Hénon, C. Heiles, The applicability of the third integral of motion: some numerical experiments, *Astron. J.* **69**, 73 (1964).

[48] V. Kaloshin and K. Zhang, Arnold diffusion for smooth convex systems of two and a half degrees of freedom, *Nonlinearity* **28**, 2699 (2015).

[49] V. Kaloshin and K. Zhang, A strong form of Arnold diffusion for two and a half degrees of freedom, *arXiv:1212.1150v3* (2018).

[50] T. Konishi and K. Kaneko, Diffusion in Hamiltonian chaos and its size dependence, *J. Phys. A: Math. Gen.* **23**, L715 (1990).

[51] J. Laskar, Frequency analysis for multi-dimensional systems. Global dynamics and diffusion. *Physica D* **67**, 257 (1993).

[52] E. Lega, M. Guzzo and Cl. Froeschlé, Detection of Arnold diffusion in Hamiltonian systems, *Physica D* **182**, 179 (2003).

[53] Ch. Lhotka, C. Efthymiopoulos and R. Dvorak, Nekhoroshev stability at L4 or L5 in the elliptic restricted three-body problem: application to Trojan asteroids, *Mon. Not. R. Astron. Soc.* **384**, 1165 (2008).

[54] P. Lochak, Canonical perturbation theory via simultaneous approximations, *Russ. Math. Surv.* **47**, 57 (1992).

[55] P. Lochak and A. Neishtadt, Estimates in the theorem of N. N. Nekhoroshev for systems with quasi-convex Hamiltonian, *Chaos* **2**, 495 (1992).

[56] R.S. MacKay and J. D. Meiss, *Hamiltonian Dynamical systems*, Adam Hilger IOP Publishing, Bristol (1987).

[57] N.N. Nekhoroshev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems - I, *Usp. Mat. Nauk* **32**, 5 (1977); *Russ. Math. Surv.* **32**, 1 (1977).

[58] N.N. Nekhoroshev, An exponential estimate of the time of stability of nearly integrable Hamiltonian systems - II, *Tr. Semin. Petrovsk.* **5**, 5 (1979); in *Topics in Modern Mathematics, Petrovskii Semin. 5, O.A. Oleinik ed.*, Consultant Bureau, New York (1985).

[59] A.I. Neishtdatd, The separation of motions in systems with rapidly oscillating rotating phase, *Prikl. Matem. Mekan.* **48**, 197 (1984); *PMM USSR* **48**, 133 (1984).

[60] G. Pinzari, On the co–existence of maximal and whiskered tori for the planetary three–body problem, *J. Math. Phys.* **59**, 052701 (2016).

[61] J. Pöschel, Nekhoroshev estimates for quasi-convex Hamiltonian systems, *Math. Z.* **213**, 187 (1993).

[62] M. Rudnev and S. Wiggins, Existence of exponentially small separatrix splittings and homoclinic connections between whiskered tori in weakly hyperbolic near-integrable Hamiltonian systems, *Physica D* **114**, 3 (1998).

[63] N. Todorović, M. Guzzo, E. Lega and Cl. Froeschlè, A numerical study of the stabilization effect of steepness, *Celest. Mech. Dyn. Astron.* **110**, 389 (2011).

[64] Treschev, D., Evolution of slow variables in a priori unstable Hamiltonian systems, *Nonlinearity* **17**, 1803 (2004).

[65] K. Zhang, Speed of Arnold diffusion for analytic Hamiltonian systems, *Invent. Math.* **186**, 255 (2011).