Chirikov and Nekhoroshev diffusion estimates: bridging the two sides of the river

Pablo M. Cincotta,[1] Christos Efthymiopoulos,[2] Claudia M. Giordano,[1] and Martín F. Mestre[1]

[1] *Grupo de Caos en Sistemas Hamiltonianos, Facultad de Ciencias Astronómicas y Geofísicas,*
Universidad Nacional de La Plata and Instituto de Astrofísica de La Plata (CONICET), La Plata, Argentina
[2] *Research Center for Astronomy and Applied Mathematics, Academy of Athens, Greece*

We present theoretical and numerical results pointing towards a strong connection between the estimates for the diffusion rate along simple resonances in multidimensional nonlinear Hamiltonian systems that can be obtained using the heuristic theory of Chirikov and a more formal one due to Nekhoroshev. We show that, despite a wide-spread impression, the two theories are complementary rather than antagonist. Indeed, although Chirikov's 1979 review has thousands of citations, almost all of them refer to topics such as the resonance overlap criterion, fast diffusion, the Standard or Whisker Map, and not to the constructive theory providing a formula to measure diffusion along a single resonance. However, as will be demonstrated explicitly below, Chirikov's formula provides values of the diffusion coefficient which are quite well comparable to the numerically computed ones, provided that it is implemented on the so-called *optimal normal form* derived as in the analytic part of Nekhoroshev's theorem. On the other hand, Chirikov's formula yields unrealistic values of the diffusion coefficient, in particular for very small values of the perturbation, when used in the original Hamiltonian instead of the optimal normal form. In the present paper, we take advantage of this complementarity in order to obtain accurate theoretical predictions for the local value of the diffusion coefficient along a resonance in a specific 3DoF nearly integrable Hamiltonian system. Besides, we compute numerically the diffusion coefficient and a full comparison of all estimates is made for ten values of the perturbation parameter, showing a very satisfactory agreement.

keywords: Nekhoroshev's and Chirikov's diffusion theories– chaos—instabilities–dynamics–Arnold diffusion

I. INTRODUCTION

In the present paper we analyze theoretically, and study by a concrete numerical example, the connection between two different types of estimates regarding the speed of local diffusion along a resonance of a nearly-integrable multi-dimensional nonlinear Hamiltonian system. These are i) estimates based on the theory of diffusion developed by B. Chirikov [8], and ii) estimates based on normal forms and the construction of the Nekhoroshev theorem ([43], [2], [35], [44]).

The theory of Chirikov [8] relates the value of the diffusion coefficient D along a resonance, in a local domain of size $\epsilon^{1/2}$ around a simply resonant point $\boldsymbol{I}^r$ of the action space, with the magnitude, wavevector, and frequencies of the so-called driving harmonic terms in systems of the form

$$H(\boldsymbol{I}, \boldsymbol{\theta}) = H_0(\boldsymbol{I}) + \epsilon V(\boldsymbol{I}, \boldsymbol{\theta}), \qquad \boldsymbol{I} \in \mathrm{G} \subset \mathbb{R}^N, \quad \boldsymbol{\theta} \in \mathbb{T}^N, \quad \epsilon \ll 1, \tag{1}$$

where $(\boldsymbol{I}, \boldsymbol{\theta})$ are N–dimensional action–angle variables, H_0 is the integrable Hamiltonian and the perturbation has the form

$$\epsilon V(\boldsymbol{I}, \boldsymbol{\theta}) = \epsilon \sum_{\boldsymbol{m}} V_{\boldsymbol{m}}(\boldsymbol{I}) \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}), \qquad V_{\boldsymbol{m}} : \mathrm{G} \to \mathbb{R}, \qquad \boldsymbol{m} \in \mathbb{Z}^N/\{0\} \ . \tag{2}$$

The theory of Chirikov has been reviewed in the framework of applications to dynamical astronomy in [4]. Using notations and terminology relevant to the present paper, we may summarize its main points as follows:

1) We decompose the Hamiltonian (1) as

$$H(\boldsymbol{I}, \boldsymbol{\theta}) = Z(\boldsymbol{I}, \boldsymbol{\theta}) + R(\boldsymbol{I}, \boldsymbol{\theta}) \tag{3}$$

where:

i) The function $Z(\boldsymbol{I}, \boldsymbol{\theta})$ contains only resonant terms associated with the particular resonance of interest (called by Chirikov the 'guiding resonance'), and can be written as:

$$Z(\boldsymbol{I}, \boldsymbol{\theta}) = Z_0(\boldsymbol{I}) + \epsilon V_G \cos(\boldsymbol{m}_G \cdot \boldsymbol{\theta}) + \ldots \tag{4}$$

where V_G and $\boldsymbol{m}_G$ are the amplitude and harmonic vector of the main resonant term. After a change of basis, one requires furthermore that the function $Z_0(\boldsymbol{I})$ be at least quadratic in a so-called resonant action variable p_1, conjugate to a resonant angle $\psi_1 = \boldsymbol{m}_G \cdot \boldsymbol{\theta}$. Then, the dynamics in the variables (p_1, ψ_1) is given essentially by the pendulum dynamics (see section IV).

ii) The term $R(\boldsymbol{I}, \boldsymbol{\theta})$ contains harmonics of non-zero wavevectors $\boldsymbol{m}$ not parallel to $\boldsymbol{m}_G$. Denoting by $\mathcal{V}$ the set of all such vectors, one has

$$R(\boldsymbol{I}, \boldsymbol{\theta}) = \sum_{\boldsymbol{m} \in \mathcal{V}} V_{\boldsymbol{m}}(\boldsymbol{I}) \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}) \quad . \tag{5}$$

2) Estimates on the diffusion coefficient stem from examining how the various terms in (5) affect the time evolution of certain quantities, which represent exact integrals of motion of the Hamiltonian flow under the term $Z(\boldsymbol{I}, \boldsymbol{\theta})$, and approximate integrals of the full Hamiltonian flow. The most important quantities of the theory are: i) the energy of the pendulum part, and ii) the remaining action variables.

In a first approximation, one computes the per-period (of the pendulum) change of the values of the approximate integrals due to the various terms in $R(\boldsymbol{I}, \boldsymbol{\theta})$. In this evaluation, one approximates the time evolution of all angles in $R(\boldsymbol{I}, \boldsymbol{\theta})$ by the ones corresponding to the evolution under the unperturbed separatrix solution of the pendulum. One should consider phase correlations between the various angles due to stickiness phenomena in the outer parts of the separatrix-like chaotic layer. Integrating over the unperturbed separatrix solution for ψ_1 implies the use of Melnikov's integrals (see Appendix A).

3) Making the crucial assumption that the diffusion within the weakly chaotic layers in the resonance web has a *normal* character, the long-term variation of the approximate integrals can now be determined in terms of the per-step variation of the same quantities. The final outcome is a formula for the local value of the diffusion coefficient along the resonance $\boldsymbol{m_G}$ in the vicinity of the point $\boldsymbol{I^r}$. After some simplification, this formula reads

$$D \lesssim \frac{1}{\Omega_G} \sum_{\boldsymbol{m}} \epsilon^2 |\omega_{\boldsymbol{m}}| \, |V_{\boldsymbol{m}}(\boldsymbol{I^r})|^2 e^{\frac{-\pi|\omega_m|}{\Omega_G}} \quad , \tag{6}$$

where $\omega_{\boldsymbol{m}} \equiv \boldsymbol{m} \cdot \omega(\boldsymbol{I^r})$ and $\Omega_G = (\epsilon V_G / |M_G|)^{1/2}$, with M_G the nonlinear pendulum mass, defined in section IV. The inequality in (6) accounts for the fact that the amplitudes $|V_{\boldsymbol{m}}|$ are not in the *optimal form*, as will be discussed in detail along this paper. A variant of Chirikov's theory called the 'stochastic pump' model was developed in [50] and [34].

Regarding numerical implementations, in [48] the authors computed the diffusion coefficient in a particular 2.5DoF nearly–integrable Hamiltonian system, whose unperturbed part is a bidimensional quartic oscillator. They obtained a good agreement between theory and experiment as long as the perturbation parameter was larger than a certain bound. However, the system considered depends on two coupling parameters. A further example in the case of the so-called three body resonances in Solar System dynamics was provided in [3]. Agreement is again found beyond a certain bound in the perturbation.

Despite the large number of citations to Chirikov's report [8], by a systematic search we have been unable to identify other concrete quantitative applications of the same theory in the literature. In fact, most citations refer to the chaotic diffusion in the so-called resonance overlap regime, which occurs for sufficiently high values of ϵ.

Nevertheless, the main goal of Chirikov's theory is to characterize the diffusion in the resonance web of multidimensional systems in the weakly chaotic limit, where there is no substantial resonance overlap. In this limit, the diffusion is conjectured to share features encountered in the mechanism of *Arnold diffusion*, proposed by Arnold [1]. However, Arnold's model is also a specific case with two parameters that can be varied independently one from the other. In contrast, in generic Hamiltonian systems the normal form theory introduces a dependence of all parameters that renders hardly tractable to generalize the proof of the existence of Arnold's mechanism [36]. In fact, although the diffusion in the web of resonances in the weakly chaotic regime has been observed in many numerical experiments, for instance [15], [16], [17], [20], [49], [25], [51], [26], [27], [47], [11], [29], [21], [18], [23], [24], [30], [31], [32], [5], [40], [38], [13], not all these examples can be characterized as 'Arnold diffusion'. In the sequel we consider systems satisfying the definition given in [23], i.e. i) satisfying simultaneously the necessary conditions of the KAM and the Nekhoroshev theorems, and ii) being in the so-called 'Nekhoroshev regime'. The first unambiguous numerical detection of local and global Arnold diffusion for such systems was made in [29] and [23], respectively.

Our own main result presented below is the following: we will argue that, regarding the quantification of Arnold diffusion in such systems, Chirikov's and Nekhoroshev's theories meet and complement each other in an essential way, so that a proper implementation of Chirikov's theory requires computing first a so-called simply-resonant *normalized Hamiltonian function* which should be *optimal* in the Nekhoroshev sense. According to the Nekhoroshev theorem, the optimal normalized Hamiltonian function is computed via a recursive algorithm of canonical transformations,

starting from the Hamiltonian (1). Furthermore, this function has also the generic form of Eq.(3). However, the difference between the original and the optimal normalized Hamiltonian is that, in the latter case, all coefficients in the term $R(\boldsymbol{I}, \boldsymbol{\theta})$ are bounded by a size exponentially small in an inverse power of ϵ. Clearly, this affects also all the coefficients $V_{\boldsymbol{m}}(\boldsymbol{I})$ of the 'driving' resonances, whose values appear in Chirikov's Eq.(6). Working with a concrete numerical example, we then show that by using the optimal normalized Hamiltonian instead of the original one in Chirikov's formula, we can obtain precise estimates of the diffusion coefficient in the weakly chaotic limit. In fact, we compute such estimates and show their very satisfactory agreement with the values of the diffusion coefficient (for several values of ϵ) found by a purely numerical integration of ensembles of orbits.

We note in this context that the connection between the Chirikov and Nekhoroshev theories is addressed by Chirikov himself in subsection 7.4 of [8]. In this review, Chirikov makes a qualitative discussion of how the optimal exponents appearing in the exponential estimates of Nekhoroshev theory affect the estimates of the speed of diffusion found in his own theory. This is further substantiated in subsection 7.6 of [8], by an analysis of the effects of higher order resonant terms on the diffusion rate in the weak perturbation limit. Here, instead, we provide direct evidence, that our normal form computation has reached an optimal order, and we also determine directly the effects of every driving resonance in the optimal Hamiltonian function using the exact version of Eq.(6).

The main steps of our study are as follows: we first perform an optimal simply-resonant normal form computation using a computer algebraic program, in order to study the diffusion in the thin chaotic layer in a domain of size $\epsilon^{1/2}$ along a particular simple resonance chosen by fixing the values of the action variables $\boldsymbol{I^r} \equiv (I_1^r, I_2^r, I_3^r)$ in the so-called 'perturbed 3DoF quartic oscillator model':

$$\tilde{H}(\boldsymbol{y}, \boldsymbol{x}) = \tilde{H}_0(\boldsymbol{y}, \boldsymbol{x}) + \epsilon \tilde{V}(\boldsymbol{x}), \tag{7}$$

with

$$\tilde{H}_0(\boldsymbol{y}, \boldsymbol{x}) = \frac{1}{2}(y_1^2 + y_2^2 + y_3^2) + \frac{1}{4}(x_1^4 + x_2^4 + x_3^4), \quad \tilde{V}(\boldsymbol{x}) = x_1^2(x_2 + x_3) \ \ .$$

We express (7) in action angle variables $(\boldsymbol{I}, \boldsymbol{\theta})$ via a transformation $(\boldsymbol{y}, \boldsymbol{x}) \to (\boldsymbol{I}, \boldsymbol{\theta})$ described in section II. Besides our acquaintance with its properties, (see [38] and [40]), our choice of model is motivated by our aim to compare results found here with those found in [9], in which a different model was used [18]. Since extended studies of the diffusion in the weakly chaotic limit are available in this latter model as well (e.g. [23] and [29]), we obtain in this way some indications regarding how general our results are.

As explained above, the 'bridge' between normal forms and the Chirikov approach is established after computing a simply resonant optimal normalized Hamiltonian valid in a neighborhood of a simply-resonant point $\boldsymbol{I^r}$ in the action space of the model (7). We then compute numerically the contributions of all the Melnikov integrals of the driving resonant terms, which we identify as the terms (except for one, see section III) appearing in the the so-called remainder function of the optimal normalized Hamiltonian. In other words, we identify the function $R(\boldsymbol{I}, \boldsymbol{\theta})$ in Eq.(3) with the remainder function. Summing the values of the associated Melnikov integrals over all driving resonances we then arrive at a theoretical prediction for the value of the diffusion coefficient along the guiding resonance. We denote this value by D_C (C stands for 'Chirikov'), and we compute D_C as a function of ϵ for ten values of ϵ.

After computing $D_C(\epsilon)$ in the above way, we perform the following comparisons:

i) We compare $D_C(\epsilon)$ with the value of the diffusion coefficient $D(\epsilon)$ computed by numerical experiments, i.e. by integrating ensembles of orbits with initial conditions in the thin chaotic layer surrounding the resonance in the neighborhood of $\boldsymbol{I^r}$. This particular calculation reveals one more salient feature of the normal form method: by transforming our orbital data into 'good' action variables, obtained via a near-identity normalizing canonical transformation, we eliminate from the data all noisy behavior due to the so-called *deformation* effects (see e.g. [22], p. 63). This removal proves to be a crucial step allowing to measure the diffusion due only to the *drift*, i.e. the slow motion along the resonance, after an integration time of the order of $t \leq 10^7$. Had we relied, instead, for this computation on the original action variables, whose time evolution reflects a combination of both the deformation and the drift effects, we would require much longer integration times (between 10^9 and 10^{11} for the smaller values of ϵ; see also [28]). In addition, transforming the numerical data to good variables allows to reveal short-term features in the obtained diffusion curves which indicate up to what extent the diffusion can be considered as normal. In general, we find that the diffusion is indeed normal to a first approximation, but with secondary features representing possible deviations from the normal character. These, we aim to study in a future work.

ii) As in [9] and [13], we check whether our data indicate a power-law relation between the diffusion coefficient and the normal form remainder R. In the present study we compare the size of the remainder $||R||$ with both D and D_C. In both cases, we find a power-law of the form $D \sim ||R||^b$, or $D_C \sim ||R||^b$, with $b \approx 2.5$. This is somewhat smaller to the value $b \simeq 3$ found in [9] and in agreement with the results of [31]. Regarding this latter point, we note that considering Chirikov's approach in combination with an optimal normal form construction instead of the original Hamiltonian (1) leads to a simple argument of why the diffusion coefficient and the size of the remainder should be

related by a power-law of the above form. If in Eq.(6) we substitute the coefficients V_m of the original Hamiltonian by the coefficients f_m of the driving harmonics in the remainder of the optimal normalized Hamiltonian function, we arrive at the estimate

$$D_C \sim \sum_{\boldsymbol{m}} \epsilon^2 \frac{|\omega_{\boldsymbol{m}}|}{\Omega_G} f_{\boldsymbol{m}}^2 e^{\frac{-\pi|\omega_{\boldsymbol{m}}|}{\Omega_G}} \quad . \tag{8}$$

Here the sum is over wavevectors $\boldsymbol{m}$ labelling the driving resonances in the remainder function, while the relations $\omega_{\boldsymbol{m}} \equiv \boldsymbol{m} \cdot \omega(\boldsymbol{I^r})$ and $\Omega_G = (\epsilon V_G/|M_G|)^{1/2}$ still hold to the degree that the lowest order resonant term in the normal form and in the original Hamiltonian are practically the same. Now, from Nekhoroshev theory, for the largest $f_{\boldsymbol{m}}$ we have the estimate $f_{\boldsymbol{m}} \sim \exp[-(1/\epsilon^{1/2})^{\frac{1}{1+\tau}}]$, where τ is a positive constant (see [10] for a heuristic derivation of such estimates). The value of τ is determined by Diophantine bounds holding for the divisors $\boldsymbol{m} \cdot \omega(\boldsymbol{I^r})$ appearing in all terms of the remainder. These bounds are of the form $|\boldsymbol{m} \cdot \omega(\boldsymbol{I^r})| > \gamma(I^r)/|\boldsymbol{m}|^\tau$, where γ is a positive constant and $|\boldsymbol{m}|$ is the L_1 modulus of $\boldsymbol{m}$. However, at the optimal normalization we also have an estimate for the minimum possible size of the *wavevectors* (the so-called 'Fourier cut-off') of those terms in the remainder function containing the worst possible accumulation of divisors. This latter estimate reads $|\boldsymbol{m}| \approx \left(\epsilon^{1/2}\right)^{-1/(\tau+1)}$. Combining these two estimates, along with the $O(\epsilon^{1/2})$ scaling of Ω_G, and that $||R|| \sim f_{\boldsymbol{m}}$ we get:

$$D_C \sim \epsilon^\alpha f_{\boldsymbol{m}}^2 e^{-(C_2/\epsilon^{1/2})^{\frac{1}{1+\tau}}} \sim \epsilon^\alpha ||R||^{2+p} \quad , \tag{9}$$

with α an exponent that depends on τ and p on the value of the constant C_2. Since $C_2 \approx 1$, we also have $p \approx 1$ (see also [13]). In [9], the power-law relation between D and $||R||$ was measured without any reference to Chirikov's theory. Thus, Chirikov's theory seems to provide a suitable framework for their interpretation. Further results regarding the connection between normal forms and estimates based on the Melnikov method (as are so the Chirikov estimates) can be found in [42], [41] and [33].

The structure of the paper is as follows: in section II we briefly summarize the basic Hamiltonian model and choice of resonance in our study. In section III we discuss the normal form calculation. In section IV we present the estimates on D_C obtained via Chirikov's method. In section V we compare these results with numerical experiments, in which we show both the form of the diffusion curves found after transforming the numerical data to good normal form variables, as well as the main scalings, i.e. D_C versus D, and D_C versus R. Section VI is a summary of our main conclusions.

II. HAMILTONIAN MODEL AND CHOICE OF RESONANCE

A. Hamiltonian Model

Our model consists of the Hamiltonian function (7) expressed in action–angle variables. To find the latter, we consider first the one-dimensional quartic oscillator

$$\tilde{H}(y,x) = \frac{y^2}{2} + \frac{x^4}{4} \quad . \tag{10}$$

Let h be the total energy and a the associated oscillation amplitude, i.e. $h = a^4/4$. The solution $x(t)$ can be expressed in terms of the Jacobi elliptic cosine (cn) of modulus $k = 1/\sqrt{2}$. Using the Fourier series development of the Jacobi elliptic cosine, we have [8]:

$$x(t) = a\frac{\sqrt{2}\pi}{K_0} \sum_{n=1}^{\infty} \frac{1}{\cosh\left((n-1/2)\pi\right)} \cos\left((2n-1)\frac{\pi a t}{2K_0}\right),$$

where $K_0 \equiv K(1/\sqrt{2})$ denotes the complete elliptic integral of the first kind. Introducing the following constants:

$$\beta \equiv \frac{\pi}{2K_0} \approx 0.847213084793979, \qquad \alpha_n \equiv \frac{1}{\cosh\left((n-1/2)\pi\right)} \quad \text{and} \quad \omega \equiv \beta a, \tag{11}$$

we have

$$x(t) = 2^{3/2}\omega \sum_{n=1}^{\infty} \alpha_n \cos\left((2n-1)\omega t\right).$$

The quantity ω is the fundamental frequency of the motion. The coefficients α_n satisfy:

$$\frac{\alpha_{n+1}}{\alpha_n} \approx \frac{1}{23} \qquad \text{for } n \geq 1 \qquad \text{and} \qquad \alpha_1 \approx 0.4. \tag{12}$$

The relation between ω and h is:

$$\omega = \sqrt{2}\beta h^{1/4}. \tag{13}$$

Since, in action–angle variables $(I, \boldsymbol{\theta})$ we have $\omega(I) = \frac{\partial H(I)}{\partial I}$, by means of Eq. (13) we find

$$h = AI^{4/3}, \qquad \text{or equivalently} \quad I = \left(\frac{h}{A}\right)^{3/4},$$

where $A \equiv (3\beta/2\sqrt{2})^{4/3} \approx 0.867145326484821$. The dependence of the frequency on the action is given by:

$$\omega(I) = \frac{4}{3}AI^{1/3} \ . \tag{14}$$

The cartesian coordinates can be finally expressed in terms of action–angle variables via the equations:

$$x(I, \boldsymbol{\theta}) = \quad (3\beta I)^{1/3}\text{cn}\left(\frac{\theta}{\beta}, \frac{1}{\sqrt{2}}\right)$$

$$y(I, \theta) = \varrho\sqrt{2\left(AI^{4/3} - \frac{1}{4}[x(I,\theta)]^4\right)}, \tag{15}$$

where ϱ stands for the sign of y and its dependence on the angle is given by:

$$\varrho \equiv \begin{cases} 1 & \text{if } 0 \leq \theta < \pi, \\ -1 & \text{if } \pi \leq \theta < 2\pi. \end{cases} \tag{16}$$

In the numerical computations, we also use the inverse transformation that allows to express the action–angle variables in terms of the cartesian variables [40]:

$$I(y, x) = \left[\frac{1}{A}\left(\frac{1}{2}y^2 + \frac{1}{4}x^4\right)\right]^{3/4}, \tag{17}$$

$$\theta(y, x) = \begin{cases} \beta \ \text{cn}^{-1}\left(\frac{x}{[3\beta \ I(y,x)]^{1/3}}, \frac{1}{\sqrt{2}}\right) & \text{if } y \geq 0, \\ 2\pi - \beta \ \text{cn}^{-1}\left(\frac{x}{[3\beta \ I(y,x)]^{1/3}}, \frac{1}{\sqrt{2}}\right) & \text{if } y < 0. \end{cases} \tag{18}$$

Passing now to the 3DoF Hamiltonian (7), by means of Eq. (15), we obtain a similar equation for $x_j(I_j, \theta_j)$, $j = 1, 2, 3$ as the first in (15). In a similar fashion Eq. (14) can be easily extended to obtain the frequency vector in terms of the actions.

In action–angle variables the Hamiltonian is expressed as:

$$H(\boldsymbol{I}, \boldsymbol{\theta}) = H_0(\boldsymbol{I}) + \epsilon V(\boldsymbol{I}, \boldsymbol{\theta}), \tag{19}$$

where

$$H_0(\boldsymbol{I}) = A(I_1^{4/3} + I_2^{4/3} + I_3^{4/3}), \tag{20}$$

$$V(\boldsymbol{I}, \boldsymbol{\theta}) = 3\beta I_1^{2/3}\text{cn}^2\left(\frac{\theta_1}{\beta}, \frac{1}{\sqrt{2}}\right)\left[I_2^{1/3}\text{cn}\left(\frac{\theta_2}{\beta}, \frac{1}{\sqrt{2}}\right) + I_3^{1/3}\text{cn}\left(\frac{\theta_3}{\beta}, \frac{1}{\sqrt{2}}\right)\right].$$

This Hamiltonian system has been previously studied in [6], [21] and [39]. In the sequel we adopt a fixed value of the total energy:

$$h \equiv 0.485 \approx 1/4\beta^4, \tag{21}$$

which corresponds to a characteristic period (of the x_2, x_3 stable axial periodic orbits) very close to 2π.

The perturbing potential can be developed in a Fourier series as:

$$V(\boldsymbol{I}, \boldsymbol{\theta}) = \hat{V}_{12}(\boldsymbol{I}) \sum_{n,m,k=1}^{\infty} \alpha_{nmk} \left\{ \cos\left(2(n+m-1)\theta_1 \pm (2k-1)\theta_2\right) + \cos\left(2(n-m)\theta_1 \pm (2k-1)\theta_2\right) \right\} +$$
$$+ \hat{V}_{13}(\boldsymbol{I}) \sum_{n,m,k=1}^{\infty} \alpha_{nmk} \left\{ \cos\left(2(n+m-1)\theta_1 \pm (2k-1)\theta_3\right) + \cos\left(2(n-m)\theta_1 \pm (2k-1)\theta_3\right) \right\},$$

(22)

with $\alpha_{nmk} \equiv \alpha_n \alpha_m \alpha_k \approx \alpha_1^3 / 23^{n+m+k-3}$ and $\hat{V}_{1j}(\boldsymbol{I}) \equiv 2^{5/2} 3\beta^4 I_1^{2/3} I_j^{1/3}$, and where the $\pm$ sign means that both terms are included in the series.

In [39] it is shown how to group all the coefficients α_{nmk} associated to the same trigonometric function into a single coefficient $\alpha_{\boldsymbol{m}}$ that satisfies $\mathcal{O}(\alpha_1^3/23^2) \le \mathcal{O}(\alpha_{\boldsymbol{m}}) \le \mathcal{O}(\alpha_1^3)$, in such a way that Eq. (22) can be rewritten as:

$$V(\boldsymbol{I}, \boldsymbol{\theta}) = \hat{V}_{12}(\boldsymbol{I}) \sum_{\boldsymbol{m} \in \mathcal{Y}} \alpha_{\boldsymbol{m}} \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}) + \hat{V}_{13}(\boldsymbol{I}) \sum_{\boldsymbol{m} \in \mathcal{Z}} \alpha_{\boldsymbol{m}} \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}) + \mathcal{O}(\alpha_1^3/23^3),$$

(23)

where $\mathcal{Y}$ and $\mathcal{Z}$ denote the subsets of wavevectors whose third and second components, respectively, are zero.

B. Choice of resonance

On applying the resonance condition, $\boldsymbol{m} \cdot \boldsymbol{\omega}(\boldsymbol{I}) = 0$, with $\boldsymbol{m} \in \mathbb{Z}^3/\{\boldsymbol{0}\}$, to the unperturbed Hamiltonian, we get

$$m_1 I_1^{1/3} + m_2 I_2^{1/3} + m_3 I_3^{1/3} = 0 \quad .$$

(24)

Thus, no resonant vector $\boldsymbol{m}$ can have all three components of the same sign. Besides the resonant vectors with two components equal to zero correspond to harmonics with null amplitude in the perturbation term of the original Hamiltonian.

A point $\boldsymbol{I^r}$ in the action space is called a resonant point with respect to the resonance wavevector $\boldsymbol{m}$ if the three components (I_1^r, I_2^r, I_3^r) satisfy Eq. (24). By construction, every resonance wavevector is tangent, at $\boldsymbol{I^r}$, to the unperturbed energy surface, denoted as $\mathcal{I}_0$.

There are 12 resonant vectors of $\mathcal{O}(\epsilon)$ whose Fourier coefficient $\alpha_{\boldsymbol{m}}$, is at most of $\mathcal{O}(\alpha_1^3/23^2)$ and they are grouped in the following set:

$$\mathcal{V}_r(\epsilon, 1/23^2) = \big\{ (2,-1,0), (2,-3,0), (2,-5,0), (4,-1,0), (4,-3,0), (6,-1,0), (2,0,-1), (2,0,-3), (2,0,-5),$$
$$(4,0,-1), (4,0,-3), (6,0,-1) \big\}$$

The main resonances at order $\mathcal{O}(\epsilon^2, 1/23^2)$ were obtained in [39]. Figure 1 shows a part of the resonant structure as projected in the action plane (I_1, I_2), for the given energy h, and $\epsilon = 0.012$. The plot is obtained by computing the time evolution of 1000×1000 orbits with initial conditions $\theta_1 = \theta_2 = \theta_3 = \pi/2$, and I_1, I_2 chosen in the domain $(I_1, I_2) \in [0.2, 0.4] \times [0, 0.2]$. The value of I_3 is obtained by solving the constant energy condition. In fact, after transforming to cartesian variables, we perform all numerical calculations in these variables, and back-transform, when needed, to action-angle variables. Along with the original equations, the variational equations of motion are integrated up to a time $t = 10^4$. We compute then, for each orbit, the value of the so-called Smaller Alignment Index (SALI, [46]), which is an indicator yielding the degree of regular or chaotic character of the associated orbit. That is, the evolution in time of two different initial deviation vectors computing the norms of the difference d- (parallel alignment index) and the sum d+ (antiparallel alignment index) of the two vectors is followed, the time evolution of the smaller alignment index reflecting the chaotic or ordered nature of the orbit. The use of chaotic indicators is known to yield an efficient method of depicting the resonant structure in the action space of multidimensional Hamiltonian systems (see [19] for the use of the FLI indicator, [7] for the use of the MEGNO indicator, and [37] for the use of the APLE indicator in this framework). In the present case, using a color map of the values of the SALI indicator obtained for the various orbits in our grid, we obtain a clear representation of the web of resonances as well as the various chaotic layers appearing around each resonance.

In our study of diffusion, we focus on initial conditions (ICs) in the thin chaotic layer around the (guiding) resonance given by the wavevector:
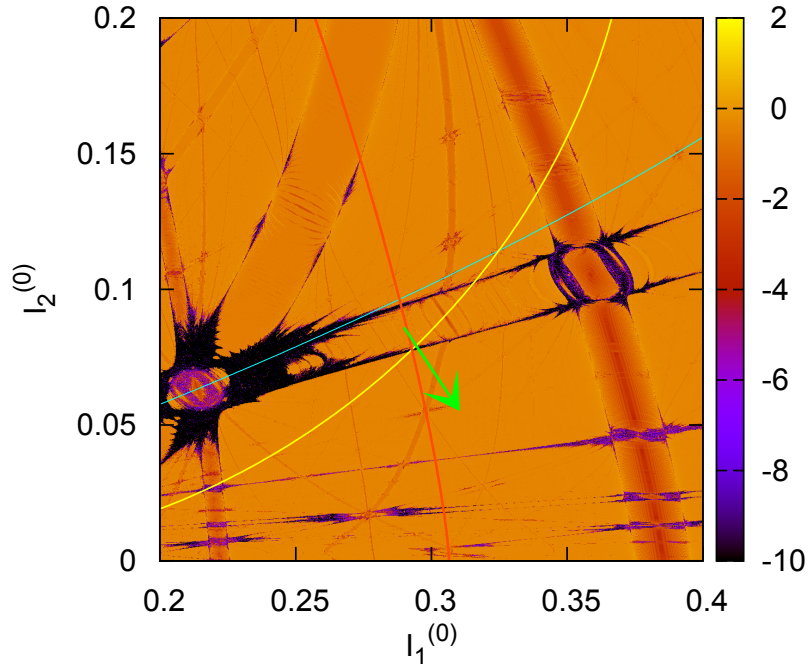
$$\boldsymbol{m}_G \equiv (2,-3,0) \quad .$$

(25)

FIG. 1: SALI map in action space, for $\epsilon = 0.012$. The color palette is such that the more chaotic orbits appear in black while the regular ones in yellow. The green arrow starts at the point $\boldsymbol{I^r}$ and it is parallel to the vector $(2, -3, 0)$.

This resonance is projected close to the center of the plane (I_1, I_2) of Fig.1. Two more conspicuous and wide resonances cross transversally our resonance of interest. These are the resonances $(2, -1, -1)$, crossing $\boldsymbol{m_G}$ in the left part of the plot, and $(2, 0, -2)$, crossing $\boldsymbol{m_G}$ in the right part of the plot. Many more resonances of smaller width, transverse to $\boldsymbol{m_G}$, are visible in the same plot. We have drawn the center of the (unperturbed) high order resonances $(6, -7, -1)$ in cyan, $(6, -4, -3)$ in yellow and $(8, 0, -7)$ in red. These resonances will play a significant role in Chirikov's formulation of the diffusion along the guiding resonance.

In order to avoid as much as possible using initial conditions overlapping with the domains of important resonance crossings, we chose to study the diffusion in the weakly chaotic layer surrounding the exactly resonant point

$$\boldsymbol{I^r} = (0.29, 0.08592592592592592, 0.434838361446344)$$

of the action space. The associated resonant frequencies are

$$\boldsymbol{\omega^r} \equiv \boldsymbol{\omega}(\boldsymbol{I^r}) = (0.7652969051118440, 0.5101979367412294, 0.8759377456886241).$$

We consider the following set of values for the perturbation parameter:

$$\epsilon \in \mathcal{E} = \{0.003, 0.005, 0.007, 0.008, 0.010, 0.012, 0.013, 0.015, 0.018, 0.020\}.$$

In section IV we present the results on estimates of the diffusion coefficient after an implementation of Chirikov's formula (8) on the data obtained by a local simply-resonant normal form construction valid in a domain that contains the region covered by our numerical orbits, around the above value of $\boldsymbol{I^r}$. Then, in section V, we compare these estimates with the results found by numerical integration of ensembles of orbits in the domain of interest.

III.   NORMAL FORM CONSTRUCTION

In computing a resonant normal form for the dynamics in our domain of interest, we used the same method as in [9]. The main steps of the method are the following:

i) *Expansion around the center*: considering a union of polydisks $|I_i'| < \rho$, where ρ is a positive constant, and setting $I_i' \equiv I_i - I_i^r$, where $\boldsymbol{I^r}$ is the central resonant value in consideration, we perform an expansion of H_0 as a Taylor series

$$H_0 = H_0^r + \boldsymbol{\omega^r} \cdot \boldsymbol{I'} + \sum_{i=1}^{3}\sum_{j=1}^{3}\frac{1}{2}M_{ij}^r I_i' I_j' + \dots \tag{26}$$

where $\boldsymbol{\omega^r} = \nabla_{\boldsymbol{I}}H_0(\boldsymbol{I^r})$, while M_{ij}^r are the entries of the Hessian matrix of H_0 at $\boldsymbol{I^r}$, denoted by M^r. Also, writing the perturbation as

$$V(\boldsymbol{I}, \boldsymbol{\theta}) = \sum_{\boldsymbol{m}} h_{\boldsymbol{m}}(\boldsymbol{I})\exp(i\boldsymbol{m}\cdot\boldsymbol{\theta}) \tag{27}$$

in a domain where all three angles satisfy $0 \le Re(\theta_i) < 2\pi$, $|Im(\theta_i)| < \sigma$ for some positive constant σ, we expand all the coefficients $h_{\boldsymbol{m}}$ around $\boldsymbol{I^r}$, namely

$$h_{\boldsymbol{m}} = h_{\boldsymbol{m}}^r + \nabla_{\boldsymbol{I}}h_{\boldsymbol{m}}^r \cdot \boldsymbol{I'} + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}h_{\boldsymbol{m},ij}^r I_i' I_j' + \dots \tag{28}$$

Both series (28) and (26) have a common domain of convergence around $\boldsymbol{I^r}$.

ii) *Action rescaling and book-keeping*: We re-scale all action variables according to

$$J_i = \epsilon^{-1/2}(I_i - I_i^r) = \epsilon^{-1/2}I_i', \quad i = 1,2,3 \tag{29}$$

so that all actions J_i are $O(1)$ quantities in the domain of interest. Since the transformation (29) is not canonical, we multiply the Hamiltonian function by $\epsilon^{-1/2}$ to restore correctness of the Hamiltonian dynamics in the re-scaled action variables. Thus, the new Hamiltonian reads: $\mathcal{H}(\boldsymbol{J},\boldsymbol{\theta}) = \epsilon^{-1/2}H(\boldsymbol{I^r} + \epsilon^{1/2}\boldsymbol{J},\boldsymbol{\theta})$.

We then split the Hamiltonian $\mathcal{H}(\boldsymbol{J},\boldsymbol{\theta})$ in terms of a similar order of smallness. In order to do so, we take into account the fact that the Fourier harmonics $\cos(\boldsymbol{m}\cdot\boldsymbol{\theta})$ in the Hamiltonian (19) have amplitudes whose scaling is given essentially by Eq.(12). This implies an exponential decay factor $\sim e^{-\sigma|\boldsymbol{m}|}$ for a harmonic of order $|\boldsymbol{m}|$, where $\sigma = 0.5\ln(23)$. Taking this fact into account we divide all harmonics in groups of a similar order of smallness, by introducing an integer constant

$$K' = \left[-\frac{1}{2\sigma} <\ln(\epsilon)>\right] \tag{30}$$

where $<\ln(\epsilon)>$ denotes the average value of $\ln(\epsilon)$ in the domain of values of ϵ considered in the present study, namely from $\epsilon = 0.003$ to $\epsilon = 0.02$. In practice, we take $K' = 2$. Then, we re-write the Hamiltonian using a so-called *book-keeping factor* λ, whose numerical value is $\lambda = 1$, as

$$\mathcal{H}(\boldsymbol{J},\boldsymbol{\theta}) = \boldsymbol{\omega^r}\cdot\boldsymbol{J} + \lambda\epsilon^{1/2}\sum_{i=1}^{3}\sum_{j=1}^{3}\frac{1}{2}M_{ij}^r J_i J_j + \dots + \sum_{\boldsymbol{m}}\left(\lambda^{1+[|\boldsymbol{m}|/K']}\epsilon^{1/2}h_{\boldsymbol{m}}^r \right. \tag{31}$$

$$\left. + \ \lambda^{2+[|\boldsymbol{m}|/K']}\epsilon\nabla_{\boldsymbol{I}}h_{\boldsymbol{m}}^r\cdot\boldsymbol{J} + \lambda^{3+[|\boldsymbol{m}|/K']}\frac{\epsilon^{3/2}}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}h_{\boldsymbol{m},ij}^r J_i J_j + \dots\right)\exp(i\boldsymbol{m}\cdot\boldsymbol{\theta}) \ .$$

Setting $Z_0 = \boldsymbol{\omega^r}\cdot\boldsymbol{J}$, the Hamiltonian (31) takes the form

$$\mathcal{H}(\boldsymbol{J},\boldsymbol{\theta}) \equiv H^{(0)}(\boldsymbol{J},\boldsymbol{\theta}) = Z_0 + \sum_{s=1}^{\infty}\lambda^s H_s^{(0)}(\boldsymbol{J},\boldsymbol{\theta};\epsilon^{1/2}) \tag{32}$$

where the superscript (0) denotes, as usually, the original Hamiltonian, and the functions $H_s^{(0)}$ are given by

$$H_s^{(0)} = \sum_{\mu=1}^{s}\epsilon^{\mu/2}\sum_{|\boldsymbol{m}|=K'(s-\mu)}^{K'(s-\mu+1)-1}H_{\mu,\boldsymbol{m}}^{(0)}(\boldsymbol{J})\exp(i\boldsymbol{m}\cdot\boldsymbol{\theta}) \tag{33}$$

where $H_{\mu,\boldsymbol{m}}^{(0)}(\boldsymbol{J})$ are polynomials containing terms of degree $\mu - 1$ or μ in the action variables $\boldsymbol{J}$. Precisely, we have:

$$H_{\mu,\boldsymbol{m}}^{(0)}(\boldsymbol{J}) = \sum_{\mu_1=0}^{\mu-1} \sum_{\mu_2=0}^{\mu-1-\mu_1} \sum_{\mu_3=0}^{\mu-1-\mu_1-\mu_2} \frac{1}{\mu_1!\mu_2!\mu_3!} \frac{\partial^{\mu-1} h_{\boldsymbol{m}}(\boldsymbol{I^r})}{\partial^{\mu_1} I_1 \partial^{\mu_2} I_2 \partial^{\mu_3} I_3} J_1^{\mu_1} J_2^{\mu_2} J_3^{\mu_3}$$

if $|\boldsymbol{m}| > 0$, or

$$H_{\mu,\boldsymbol{m}}^{(0)}(\boldsymbol{J}) = \sum_{\mu_1=0}^{\mu} \sum_{\mu_2=0}^{\mu-\mu_1} \sum_{\mu_3=0}^{\mu-\mu_1-\mu_2} \frac{1}{\mu_1!\mu_2!\mu_3!} \frac{\partial^{\mu} H_0(\boldsymbol{I^r})}{\partial^{\mu_1} I_1 \partial^{\mu_2} I_2 \partial^{\mu_3} I_3} J_1^{\mu_1} J_2^{\mu_2} J_3^{\mu_3}$$

$$+ \sum_{\mu_1=0}^{\mu-1} \sum_{\mu_2=0}^{\mu-1-\mu_1} \sum_{\mu_3=0}^{\mu-1-\mu_1-\mu_2} \frac{1}{\mu_1!\mu_2!\mu_3!} \frac{\partial^{\mu-1} h_0(\boldsymbol{I^r})}{\partial^{\mu_1} I_1 \partial^{\mu_2} I_2 \partial^{\mu_3} I_3} J_1^{\mu_1} J_2^{\mu_2} J_3^{\mu_3}$$

if $\boldsymbol{m} = 0$.

iii) *Resonant module:* After choosing the vector of the guiding resonance according to Eq. (25), we define the resonant module as the set of all harmonics satisfying

$$\mathcal{M} \equiv \{\boldsymbol{m} = 0 \text{ or } \boldsymbol{m}//\boldsymbol{m_G}\} \quad . \tag{34}$$

The set $\mathcal{M}$ includes the wavevectors of all possible terms appearing in the normal form.

Hamiltonian normalization: we perform Hamiltonian normalization using a computer-algebraic program written by one of us (C.E.) in Fortran. In this, we generate canonical transformations using the method of Lie generating functions. For a review of the advantages of this method from a computational point of view, see [10].

The normalization is performed in steps $r = 1, 2, ...$, according to the recursive formula

$$H^{(r)} = \exp(L_{\chi_r}) H^{(r-1)} \tag{35}$$

where χ_r is the r-th step generating function defined by the homological equation

$$\{\omega^r \cdot \boldsymbol{J}^{(r)}, \chi_r\} + \lambda^r \tilde{H}_r^{(r-1)}(\boldsymbol{J}^{(r)}, \boldsymbol{\theta}^{(r)}) = 0 \tag{36}$$

and $\tilde{H}_r^{(r-1)}(\boldsymbol{J}^{(r)}, \boldsymbol{\theta}^{(r)})$ denotes all terms of $H^{(r-1)}$ which do not belong to the resonant module $\mathcal{M}$. The operator L_χ is the Poisson bracket $L_\chi \equiv \{\cdot, \chi\}$.

Remainder and optimal normalization order: After r normalization steps, the transformed Hamiltonian $H^{(r)}$ has the form

$$H^{(r)}(\boldsymbol{\theta}, \boldsymbol{J}) = Z^{(r)}(\boldsymbol{\theta}, \boldsymbol{J}; \lambda, \epsilon) + R^{(r)}(\boldsymbol{\theta}, \boldsymbol{J}; \lambda, \epsilon) \tag{37}$$

where $Z^{(r)}(\boldsymbol{J}^{(r)}, \boldsymbol{\theta}^{(r)}; \lambda, \epsilon)$ and $R^{(r)}(\boldsymbol{J}^{(r)}, \boldsymbol{\theta}^{(r)}; \lambda, \epsilon)$ are the normal form and the remainder respectively. The normal form is a finite expression which contains terms up to order r in the book-keeping parameter λ, while the remainder is a convergent series containing terms of order λ^{r+1} and beyond. Since in the computer we can only store a finite number of remainder terms, we probe numerically the convergence of the remainder function in the domain of interest in the following way: writing the remainder in the form

$$R^{(r)}(\boldsymbol{J}^{(r)}, \boldsymbol{\theta}^{(r)}) = \sum_{s=r+1}^{\infty} \lambda^s \sum_{|\boldsymbol{m}|} R_{\boldsymbol{m},s}^{(r)}(\boldsymbol{J}^{(r)}) \exp(i\boldsymbol{m} \cdot \boldsymbol{\theta}^{(r)}) \tag{38}$$

we define the truncated norms

$$||R^{(r)}(\xi)||_{\leq p} = \sum_{s=r+1}^{p} \sum_{|\boldsymbol{m}|} |\mathcal{R}_{\boldsymbol{m},s}^{(r)}(\xi)| \tag{39}$$

where $\mathcal{R}_{\boldsymbol{m},s}^{(r)}(\xi) = R_{\boldsymbol{m},s}^{(r)}(J_1 = J_2 = J_3 = \xi)$, and ξ denotes a distance from the central point $\boldsymbol{I}^r$ in the action space, in re-scaled units, at which the diffusion is measured. By plotting the values of $||R^{(r)}(\xi)||_{\leq p}$ versus p we can have a

numerical indication of whether the remainder function converges at ξ after a certain value of p. Since the distance from the center to the chaotic layer is constant in the re-scaled action variables, by plots like in Fig. 1 we have estimated the value of $\xi = 0.07$.

Figure 2 (a) shows a calculation of this type for $\epsilon = 0.01$. The middle curve, which corresponds to the normalization order $r = 6$, shows the value of $||R^{(r)}(\xi)||_{\leq p}$ as a function of p for $p = 7, \ldots, 22$. Clearly, after $p = 9$ the cumulative sum (39) shows no further substantial variation, which indicates that the remainder series converges after just three consecutive terms $p = 7, 8$ and 9. The lower and upper curves show now the same effect for the normalization orders $r = 11$ and $r = 14$, respectively. The main effect to note is that the estimated remainder value $||R^{(r)}(\xi)||_{\leq 22}$ found for $r = 14$ is larger than the one for $r = 11$, implying that the *optimal* normalization order r_{opt} is below $r = 14$. Fig.2 (b) shows, precisely, the asymptotic character of the above normalization, showing $||R^{(r)}(\xi)||_{\leq 22}$ against the normalization order r for various values of ϵ as indicated in the figure. We observe that in the considered range of values of ϵ the optimal normalization order turns to be always 10 or 11. In fact, it is known from basic theory that the optimal order of normalization is in general an increasing function of $1/\epsilon$. However, depending on the number–theoretical properties of the frequencies $\boldsymbol{\omega}^r$, the increase may occur by abrupt steps (see e.g. [12]). These steps are due to the fact that the smallest possible divisor $|\boldsymbol{m} \cdot \boldsymbol{\omega}^r|$ may remain invariant for long transient intervals of values of $|\boldsymbol{m}|$, before eventually being forced to follow the 'envelope' provided by the Diophantine inequality $|\boldsymbol{m} \cdot \boldsymbol{\omega}^r| > \gamma/|\boldsymbol{m}|^\tau$. This phenomenon appears, precisely, for our present choice of resonant frequencies.
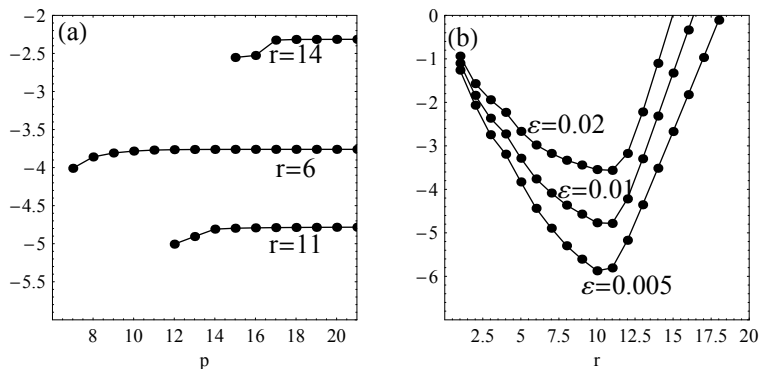

FIG. 2: (a) Values of $||R^{(r)}(\xi)||_{\leq p}$ as a function of p for three normalization orders, $r = 6, 11, 14$. (b) Values of the remainder as a function of the normalization order r for three values of ϵ (right panel).

From the normal form calculation we retain three sets of data used in subsequent calculations:

1) The value of the optimal remainder $||R^{(r_{opt})}||$ as a function of ϵ (found from the minima of all curves as in Fig.2 (b)). This is given in the Table I.

ϵ	$||R^{(r_{opt})}||$
0.020	2.4×10^{-4}
0.018	1.6×10^{-4}
0.015	7.7×10^{-5}
0.013	4.4×10^{-5}
0.012	3.3×10^{-5}
0.010	1.6×10^{-5}
0.008	7.1×10^{-6}
0.007	4.4×10^{-6}
0.005	1.3×10^{-6}
0.003	2.4×10^{-7}

TABLE I: The value of the optimal remainder $||R^{(r_{opt})}||_{\leq 22}$ as a function of ϵ for all values $\epsilon \in \mathcal{E}$ according to the simply resonant normal form calculation performed as exposed above.

2) The form of the normalized Hamiltonian at the optimal normalization order $r_{opt} = 10$, including the remainder terms up to the order 22. This is transformed below in the basis used in Chirikov's theory in order to compute the amplitudes and wavevectors of the guiding resonances.

3) The normalizing canonical transformation yielding the old canonical variables as functions of the new canonical variables. This transformation is provided directly by the composition of the computed Lie generating functions via the equations:

$$q_{new} \;=\; \exp(-L_{\chi_1})\exp(-L_{\chi_2})...\exp(-L_{\chi_r})q_{old}$$

(40)

where q_{old} and q_{new} refer to anyone of the three action or angle variables before and after implementing the canonical transformations. In the case of the old action variables, we first compute the values of the re-scaled actions $\boldsymbol{J} = \epsilon^{-1/2}(\boldsymbol{I} - \boldsymbol{I^r})$ from the values of the original actions $\boldsymbol{I}$ which are available by our numerical data. Then, we 'pass' the values of the actions $\boldsymbol{J}$ to the transformation (40).

IV.   DIFFUSION ESTIMATES USING CHIRIKOV'S THEORY

We will now use the data of the normal form computation exposed in the previous section, in order to obtain estimates of the diffusion coefficient using the theory of Chirikov. The reader is deferred to [8] and [4] for a detailed presentation of this theory.

As a preliminary step, we re-express the optimal normalized Hamiltonian functions found in the above section, for each value of ϵ in the considered set $\mathcal{E}$, into a function expressed in non-scaled action variables $(I_j - I_j^r) = \epsilon^{1/2}J_j$, $j = 1, 2, 3$. This is done by the back transform

$$H^{(r_{opt})}(\boldsymbol{\theta}, \boldsymbol{I} - \boldsymbol{I^r}) = \epsilon^{1/2}\left[Z^{(r_{opt})}(\boldsymbol{\theta}, \epsilon^{-1/2}(\boldsymbol{I} - \boldsymbol{I^r}); \lambda, \epsilon) + R^{(r_{opt})}(\boldsymbol{\theta}, \epsilon^{-1/2}(\boldsymbol{I} - \boldsymbol{I^r}); \lambda, \epsilon) \right] \quad .$$

(41)

The lowest order terms in the above expression are of the form

$$H^{(r_{opt})} = \boldsymbol{\omega_G} \cdot (\boldsymbol{I} - \boldsymbol{I^r}) + O\left((\boldsymbol{I} - \boldsymbol{I^r})^2\right) + ... + \epsilon[V_G + O\left((\boldsymbol{I} - \boldsymbol{I^r})\right) + ...](\cos(\boldsymbol{m_G} \cdot \boldsymbol{\theta}) + ...$$

(42)

The guiding resonance is the one given by $\boldsymbol{m_G} = (2, -3, 0)$. For the constant V_G we find the numerical value $V_G \simeq 0.005259$.

Following Chirikov's formulation, we perform a 'change of basis', i.e. define three new fundamental directions in the action space and re-write the action variables in terms of components in these new directions. This will be done through a canonical transformation and, to this end, we define the vectors

$$\boldsymbol{\mu}_1 = \boldsymbol{m_G}, \qquad \boldsymbol{\mu}_2 = \boldsymbol{\omega}^r/|\boldsymbol{\omega}^r|$$

$$\boldsymbol{\mu}_3 = (\boldsymbol{n}^r \wedge \boldsymbol{\omega}^r)/|\boldsymbol{n}^r \wedge \boldsymbol{\omega}^r| = (0.6769019893644146, 0.2005635524042710, -0.7082216872148703),$$

where

$$\boldsymbol{n}^r = (\partial[\boldsymbol{m_G} \cdot \boldsymbol{\omega}(\boldsymbol{I})]/\partial \boldsymbol{I})_{\boldsymbol{I}^r} = (1.759303230142170, -5.937648401729823, 0),$$

is a vector normal to the guiding resonance surface at the point $\boldsymbol{I}^r$.

Geometrically, we have that $\boldsymbol{\mu}_1$ lies in the tangent plane to $\mathcal{I}_0$ at the point $\boldsymbol{I}^r$, $\boldsymbol{\mu}_2$ is normal to that plane and $\boldsymbol{\mu}_3$ is simultaneously orthogonal to $\boldsymbol{n}^r$ and to $\boldsymbol{\mu}_2$, i.e. it is tangent, at $\boldsymbol{I}^r$, to the intersection between the guiding resonance surface and $\mathcal{I}_0$. These three vectors are linearly independent if and only if $\boldsymbol{m_G}$ is not perpendicular to $\boldsymbol{n}^r$. A way to ensure this geometrical condition is to assume that $\mathcal{I}_0$ is convex at the point $\boldsymbol{I}^r$. This is easily checked to be true in our specific example.

Let $\Upsilon \in \mathbb{R}^{3 \times 3}$ be the matrix whose $i-$th row is the vector $\boldsymbol{\mu}_i$, for $i = 1, 2, 3$, and let G be a generating function given by:

$$G(\boldsymbol{p}, \boldsymbol{\theta}) \equiv \sum_{j=1}^{3}\left(I_j^r + \sum_{k=1}^{3} p_k \Upsilon_{kj} \right)\theta_j.$$

The associated canonical transformation $(\boldsymbol{I}, \boldsymbol{\theta}) \rightarrow (\boldsymbol{p}, \boldsymbol{\psi})$ can be written explicitly as:

$$\boldsymbol{\psi} = \Upsilon\boldsymbol{\theta},$$

$$\boldsymbol{p} = \Upsilon^{-T}(\boldsymbol{I} - \boldsymbol{I^r}),$$

(43)

where $\Upsilon^{-T} \equiv (\Upsilon^T)^{-1}$ denotes the inverse matrix of Υ's transpose. This transformation consists of a translation of the origin to the resonant action, followed by a change of base in the action space.

The new momenta, p_k ($k = 1, 2, 3$), are the components of the vector $(\boldsymbol{I} - \boldsymbol{I}^r)$ in Chirikov's base such that, p_1 measures the deviation of the actual motion from the resonant point *across* the guiding resonance layer, p_2 gives the unperturbed energy variation H_0, and p_3 measures the departure from the resonant value *along* the guiding resonance layer, i.e., in the direction along which we will measure the diffusion coefficient. The conjugate angle of p_1, $\psi_1 = \boldsymbol{m}_G \cdot \boldsymbol{\theta}$, is the resonant angle.

After the transformation (43), the truncated Hamiltonian (up to order 22) takes the form:

$$
\begin{aligned}
H(\boldsymbol{p}, \boldsymbol{\psi}) \;=\; & \frac{p_1^2}{2M_G} + |\boldsymbol{\omega}^r| p_2 + \sum_{l=1}^{3} \sum_{k+l>2}^{3} \frac{p_k p_l}{2M_{kl}} \\
& + \; \epsilon [V_G + V_1(\boldsymbol{p}) + V_2(\boldsymbol{p}) + \ldots + V_{r_{opt}}(\boldsymbol{p})] \cos \psi_1 + \epsilon \sum_{q=2}^{q_{opt}} U_q(\boldsymbol{p})] \cos(q\psi_1) \\
& + \; \epsilon \sum_{\boldsymbol{m}} [V_{0,\boldsymbol{m}} + V_{1,\boldsymbol{m}}(\boldsymbol{p}) + \ldots + V_{20,\boldsymbol{m}}(\boldsymbol{p}) \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}(\boldsymbol{\psi})) \quad .
\end{aligned}
\tag{44}
$$

In the above expression:

i) We have already computed V_G, while

$$
\frac{1}{M_{kl}} = \sum_{i=1}^{3} \sum_{j=1}^{3} \Upsilon_{ki} \frac{\partial \omega_i^r}{\partial I_j} \Upsilon_{lj}, \qquad \frac{1}{M_G} \equiv \frac{1}{M_{11}} = \sum_{i=1}^{3} \sum_{j=1}^{3} m_{g_i} \frac{\partial \omega_i^r}{\partial I_j} m_{g_j} \quad .
\tag{45}
$$

We find the following values:

$$
\begin{aligned}
M_G \;&=\; 4.68789151245171606 \times 10^{-2}, \\
M_{12}^{-1} = M_{21}^{-1} \;&=\; -1.3250433064004110 \\
M_{13}^{-1} = M_{31}^{-1} \;&=\; 0, \\
M_{22}^{-1} \;&=\; 0.95805552130252458, \\
M_{23}^{-1} = M_{32}^{-1} \;&=\; 0.19026669807696181 \\
M_{33}^{-1} \;&=\; 0.81946117880019043.
\end{aligned}
$$

ii) The functions V_s and $V_{s,\boldsymbol{m}}$ are homogeneous polynomial of degree s in the variables p_i, $i = 1, 2, 3$.

iii) The functions $U_q(\boldsymbol{p})$ are polynomial in the variables p_i.

iv) The coefficients $V_{s,\boldsymbol{m}}$ are much smaller in size than the coefficients V_s or U_q, since the former belong to the remainder, while the latter belong to the normal form.

The Hamiltonian (44) does not have precisely the form required for the implementation of Chirikov's formulae. We thus proceed in obtaining an approximate form of the Hamiltonian, by implementing a number of simplifications as follows:

i) By construction, the initial conditions of all the orbits are taken to lie nearly exactly on the so called 'plane of fast drift' (see section V), which corresponds to setting initially $p_2 \simeq p_3 \simeq 0$. Furthermore, if we neglect the effect of the remainder, p_2 and p_3 are preserved quantities under the normal form dynamics. We thus set $p_2 = p_3 = 0$ in all estimates for the coefficients V_s, U_q, or $V_{s,\boldsymbol{m}}$.

ii) In contrast, p_1 is subject to oscillations in time, since, according to Eq.(44), it is a pendulum action variable. For the separatrix half-width we find the estimate

$$
\Delta p_1 = 2\sqrt{\epsilon |M_G V_G|} \simeq 0.0314 \epsilon^{1/2} \quad .
\tag{46}
$$

Since the orbits actually evolve in a thin separatrix-like layer, instead of an exact separatrix solution, we may assume that the time evolution of $p_1(t) \approx \Delta p_1 \cos(\psi_1/2)$ does not differ much from the evolution of the momentum along an oscillation solution with period $2\pi/\Omega_G$ and amplitude Δp_1. Then, performing the average over one period of motion, for all odd powers we have $< p_1(t)^n > \simeq 0$, while for the even orders we use the approximation:

$$
< p_1(t)^n > \simeq \frac{1}{2^n} \frac{n!}{(n/2)!(n/2)!} \Delta p_1^n \quad .
\tag{47}
$$

Then, we estimate the numerical values of all coefficients $V_{s,\boldsymbol{m}}(p_1; p_2 = p_3 = 0)$ in Eq.(44), setting $V_{s,\boldsymbol{m}}(p_1; p_2 = p_3 = 0) = 0$ if s is odd, and using the expression (47), with Δp_1 given by Eq.(46), if s is even. In practice, the remainder contains hundreds of thousands of terms, most of which, after the above substitutions, are found to be of negligible size. We thus impose a size cut-off limit and in subsequent calculations keep only the remainder terms of size larger than 10^{-10}.

iii) Finally, we ignore all the normal form terms $V_s(\boldsymbol{p})$, as well as $U_q(\boldsymbol{p})$. In fact, one can check that these terms introduce corrections of the order 10^{-4} of the leading normal form term, i.e. $V_G \cos(\psi_1)$.

With the above simplifications, the Hamiltonian resumes finally an approximate form suitable for the implementation of Chirikov's formulae, namely

$$H(\boldsymbol{p}, \boldsymbol{\psi}) \;=\; \frac{p_1^2}{2M_G} + |\boldsymbol{\omega}^r| p_2 + \sum_{l=1}^{3} \sum_{k+l>2}^{3} \frac{p_k p_l}{2M_{kl}} + \epsilon V_G \cos \psi_1 + \epsilon \sum_{\boldsymbol{m}} \tilde{V}_{\boldsymbol{m}} \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}(\boldsymbol{\psi})) \tag{48}$$

where the coefficients $\tilde{V}_{\boldsymbol{m}}$ have now constant values.

Under the form (48), the Hamiltonian lends now itself to the computation of Chirikov's diffusion coefficient as follows:

For all amplitudes of the remainder, $\tilde{V}_{\boldsymbol{m}}$, we compute the coefficients (see appendix A)

$$W_{\boldsymbol{m}} = \frac{4\pi \nu_1(\boldsymbol{m}) \nu_2(\boldsymbol{m}) \tilde{V}_{\boldsymbol{m}} (2|\boldsymbol{m} \cdot \boldsymbol{\omega}^r|/\Omega_G)^{2|\xi_{\boldsymbol{m}}|-1}}{\xi_{\boldsymbol{m}} V_G \Gamma(2|\xi_{\boldsymbol{m}}|)} \exp\left(\frac{-\pi |\boldsymbol{m} \cdot \boldsymbol{\omega}^r|}{2\Omega_G}\right) \tag{49}$$

where

$$\nu_k(\boldsymbol{m}) = \sum_{i=1}^{3} m_i \Upsilon_{ik}, \quad k = 1, 2, 3, \quad \xi_{\boldsymbol{m}} = \sum_{k=1}^{3} \frac{\nu_k(\boldsymbol{m})}{M_{k1}} \quad .$$

The coefficients $W_{\boldsymbol{m}}$ are used in the computation of the per-period variation of the pendulum energy integral (see appendix A). According to Chirikov, we first isolate the perturbing term $\tilde{V}_{\boldsymbol{m}}$ yielding the largest value of $W_{\boldsymbol{m}}$, and we call this term 'layer' resonance, i.e. the resonance mainly responsible for the formation of the thin separatrix chaotic layer at the border of the guiding resonance. In our data, we found that the layer resonance is in every case associated to the harmonic $\boldsymbol{m}_l = (6, -7, -1)$. This amplitude will be hereafter denoted by W_l.

All other terms $\tilde{V}_{\boldsymbol{m}}$ except for the layer one are called 'driving' resonances. These are the harmonics whose average time variation causes a variation of the values of all the normal form integrals. For the ten values of ϵ considered, we computed the ratio of the layer resonance amplitude and that of the largest term of the driving resonances. These ratios are shown in Table II.

ϵ	$v_{\boldsymbol{m}} = W_l/W_{\boldsymbol{m}}$
0.020	7.6
0.018	8.1
0.015	9.0
0.013	9.6
0.012	9.9
0.010	10.3
0.008	10.4
0.007	10.2
0.005	8.6
0.003	2.0

TABLE II: Amplitude of the layer resonance with respect to the largest driving resonance.

For all values of ϵ, the largest driving resonance corresponds to the harmonic $(6, -4, -3)$, except for $\epsilon = 0.003$, for which the leading driving resonance is $(8, 0, -7)$. Table III shows the most significant driving resonances for each value of ϵ, ordered by the size of the corresponding amplitude $W_{\boldsymbol{m}}$.

The results shown in Table II are in agreement with the assumption that $W_l \gg W_{\boldsymbol{m}}$ for almost all $\boldsymbol{m} \neq \boldsymbol{l}$. For $\epsilon = 0.003$, we obtain that $W_l \sim W_{\boldsymbol{m}}$, and although Chirikov [8] pointed out that the approximation $v_{\boldsymbol{m}} = W_{\boldsymbol{m}}/W_l \sim 1$ should be sufficient to justify all estimates, this seems not to be true. For $\epsilon = 0.003$ the amplitudes $W_{\boldsymbol{m}}$ of three leading

ϵ	Leading Driving Resonances
0.020	$(6,-4,-3);\ (6,-2,-4);\ (6,-6,-2);\ (4,-1,-3);\ (-2,-2,3);\ (4,-3,-2)$
0.018	$(6,-4,-3);\ (6,-2,-4);\ (4,-1,-3);\ (6,-6,-2);\ (-2,-2,3);\ (4,-3,-2)$
0.015	$(6,-4,-3);\ (6,-2,-4);\ (4,-1,-3);\ (6,-6,-2);\ (-2,-2,3);\ (8,0,-7)$
0.013	$(6,-4,-3);\ (6,-2,-4);\ (4,-1,-3);\ (6,-6,-2);\ (-2,-2,3);\ (8,0,-7)$
0.012	$(6,-4,-3);\ (6,-2,-4);\ (4,-1,-3);\ (6,-6,-2);\ (-2,-2,3);\ (8,0,-7)$
0.010	$(6,-4,-3);\ (6,-2,-4);\ (4,-1,-3);\ (8,0,-7);\ (6,-6,-2);\ (-2,-8,3)$
0.008	$(6,-4,-3);\ (6,-2,-4);\ (4,-1,-3);\ (8,0,-7);\ (-2,-8,3);\ (8,-7,-3)$
0.007	$(6,-4,-3);\ (6,-2,-4);\ (8,0,-7);\ (4,-1,-3);\ (-2,-8,3);\ (8,-7,-3)$
0.005	$(6,-4,-3);\ (6,-2,-4);\ (8,0,-7);\ (-2,-8,3);\ (6,-3,7);\ (4,-1,-3)$
0.003	$(8,0,-7);\ (6,-4,-3);\ (6,-2,-4);\ (10,-3,-7);\ (4,-1,-3);\ (-2,-2,3)$

TABLE III: Principal driving resonances ordered by their amplitude $W_{\boldsymbol{m}}$ for each ϵ value.

driving resonances are very similar to the one of the layer resonance, so it seems difficult to distinguish between layer and driving resonances. It is interesting to see from Table III how certain harmonics increase their importance in the perturbation as ϵ changes from larger to lower values. Notice that the driving resonances $(6,-4,-3); (6,-2,-4)$ and $(4,-1,3)$ are always present as leading terms for all values of the perturbation parameter. Anyway, the amplitudes of the smaller driving resonances shown in Table III are of the order of $\sim 10^{-1}$ or $\sim 10^{-2}W_{\boldsymbol{m}}$ of the leading one. This difference is even larger for $\epsilon = 0.003$ where $W_{\boldsymbol{m}}$ corresponding to the harmonic $(-2,-2,3)$ is $\sim 3 \times 10^{-3}$ times the $W_{\boldsymbol{m}}$ of $(8,0,-7)$.

Excluding the layer resonance, the scalar diffusion coefficient along the direction of the vector $\boldsymbol{\mu}_3$ can be computed by Chirikov's formula

$$D(\boldsymbol{I}^r;\boldsymbol{\mu}_3) \approx \frac{2\pi^2 F\epsilon^2}{T_a|\boldsymbol{\omega}^r|^2}\sum_{\boldsymbol{m}\neq\boldsymbol{l}}\frac{\nu_3(\boldsymbol{m})^2}{\nu_2(\boldsymbol{m})^2}\left|\frac{2\omega_{\boldsymbol{m}}}{\Omega_G}\right|^{|4|\xi_{\boldsymbol{m}}|}\frac{V_{\boldsymbol{m}}^2}{\Gamma^2(2|\xi_{\boldsymbol{m}}|)}e^{-\frac{\pi|\omega_{\boldsymbol{m}}|}{\Omega_G}} \quad , \tag{50}$$

where

$$T_a(w_s) \approx \frac{1}{\Omega_G}\ln\left(\frac{32e}{w_s}\right), \qquad T_a \sim \left(\frac{1+\ln\epsilon}{\sqrt{\epsilon}}\right)+\frac{1}{\epsilon} \qquad \Omega_G(\epsilon)=\sqrt{\epsilon|V_G|/|M_G|}, \qquad \omega_{\boldsymbol{m}}=\boldsymbol{m}\cdot\boldsymbol{\omega}^r=\nu_2(\boldsymbol{m})|\boldsymbol{\omega}^r|,$$

$\Gamma(x)$ denotes the Gamma function and w_s is the width of the guiding resonance chaotic layer (see appendix A). In Eq.(50), F is a positive constant called the 'reduction factor', which accounts for correlations between the phases representing the initial conditions of the various iterates of the orbits in the chaotic layer, as described by the separatrix mapping. The true value of F is unknown, but for simplicity we set $F = 1$ (see Section V). Although Chirikov [8] suggested, by a crude theoretical estimate, the value $F \sim 1/3$, we find the introduction of such a "reduction factor" rather unnecessary in the regime examined in the present paper, in which theoretical estimates can be compared to numerical ones only up to an order of magnitude agreement.

Implementing Eq.(50) in our data we obtained the following values for D_C in terms of ϵ:

We should emphasize that the obtained values for D_C rest under the assumption of normal diffusion, for which the unperturbed (original or transformed) actions vary as a linear power of t. This is in fact a strong assumption. As discussed below, our numerical results indicate that the assumption of normal diffusion is an adequate first approximation. However, there are additional features in all our obtained diffusion curves, whose quantitative description goes beyond the assumptions of Chirikov's theory. At any rate, as it will be shown in the next section, the diffusion coefficients computed numerically under the assumption of normal diffusion agree very well with the diffusion coefficients reported in table IV. To this comparison we now turn our attention.

V. NUMERICAL RESULTS. COMPARISON OF ALL ESTIMATES

In the present section, we will present the results of numerical simulations of ensembles of orbits in our system, aiming to determine the diffusion coefficient D along the simple resonance, as a function of ϵ, in a purely numerical way. In what follows, we present computations using two different sets of variables, namely the canonical variables in the original Hamiltonian, and those arising after an optimal canonical transformation. To separate in notation the former from the latter, we use the superscript (0) to denote the old canonical variables, i.e., the angles $\boldsymbol{\theta}^{(0)}$, and the

ϵ	D_C
0.020	8.4×10^{-11}
0.018	3.0×10^{-11}
0.015	5.0×10^{-12}
0.013	1.3×10^{-12}
0.012	5.9×10^{-13}
0.010	1.2×10^{-13}
0.008	1.9×10^{-14}
0.007	6.9×10^{-15}
0.005	6.7×10^{-16}
0.003	2.4×10^{-17}

TABLE IV: Diffusion coefficient value for $\epsilon \in \mathcal{E}$. Note the significant decrease of D_C for $\epsilon = 0.003$. See text.

actions $\boldsymbol{I^{(0)}}$ (or $\boldsymbol{p^{(0)}}$ in Chirikov's basis). We use bar, or non-bar, symbols to denote all ensemble quantities defined over the set of the old, or the new, action variables respectively.

A. Statistical quantities

Using the results from the integration of ensembles of orbits, we measure the time variation of the *variance* of the old and new action variables, $p_i^{(0)}$ and p_i respectively, for $i = 1, 2, 3$. The statistical quantities presented below are *ensemble* averages computed numerically over a finite number, N_p, of test particles. The initial conditions of these test particles are chosen in a small domain in the action space, all of them having the same total energy.

Let $p_j(t, i)$ be the value of the j–th component of the vector $\boldsymbol{p}$, at the time t, associated to the i–th particle. Thus, the (instantaneous) mean value of this component is given by:

$$\mu_j(t) \equiv \langle p_j(t) \rangle \equiv \frac{1}{N_p} \sum_{i=1}^{N_p} p_j(t, i),$$

and the corresponding variance is:

$$\sigma_j^2(t) \equiv \langle (p_j(t) - \mu_j(t))^2 \rangle \equiv \frac{1}{N_p} \sum_{i=1}^{N_p} (p_j(t, i) - \mu_j(t))^2.$$

Analogously, we can define the mean and the variance of the old actions:

$$\bar{\mu}_j(t) \equiv \langle p_j^{(0)}(t) \rangle, \qquad \bar{\sigma}_j^2(t) \equiv \langle (p_j^{(0)}(t) - \bar{\mu}_j(t))^2 \rangle. \tag{51}$$

Diffusion processes are commonly characterized by a power law relationship of the form $\sigma^2(t) = c\, t^\eta$, with $c > 0$. If $\eta = 1$ we have normal diffusion, while in case of $\eta < 1$, the phenomenon is called subdiffusion, or when $\eta > 1$ it is called superdiffusion. In the normal diffusion case, within some time interval $[t_0, t_f]$, it is possible to define a numerical diffusion coefficient, D. In this work, the diffusion coefficient is associated to σ_3^2 through a least square fit of the ansatz:

$$\sigma_3^2(t) = Dt + \rho, \tag{52}$$

where D and ρ are the fitted parameters.

B. The ensembles

For each value of $\epsilon \in \mathcal{E}$, we consider one ensemble of $N_p = 10^3$ particles. The ICs of the ensembles are chosen inside the chaotic layer of the guiding resonance. More specifically, they are located in a neighborhood of the unperturbed separatrix of the simple pendulum model associated to the resonant action ($\boldsymbol{I^r}$). All the ICs belong to the same plane

used to compute the SALI maps (as in Fig.1), i.e. they satisfy $\theta_i^{(0)} = \pi/2$ ($i = 1, 2, 3$) and $H(\boldsymbol{I}^{(0)}, \boldsymbol{\theta}^{(0)}) = h$ so that they are solved from: $H_0(\boldsymbol{I}^{(0)}) \equiv H(\boldsymbol{I}^{(0)}, \pi/2, \pi/2, \pi/2) = h$. They are selected at random inside a square of size 4×10^{-12}.

Fig. 3 display the SALI map in a neighborhood of $\boldsymbol{I}^r$ together with the location of the initial ensemble for ϵ equal to 0.015 and 0.003, respectively. The left panel of Fig. 3 shows a magnification in the neighborhood of the ensembles of ICs (green square) for the smallest perturbation parameter.
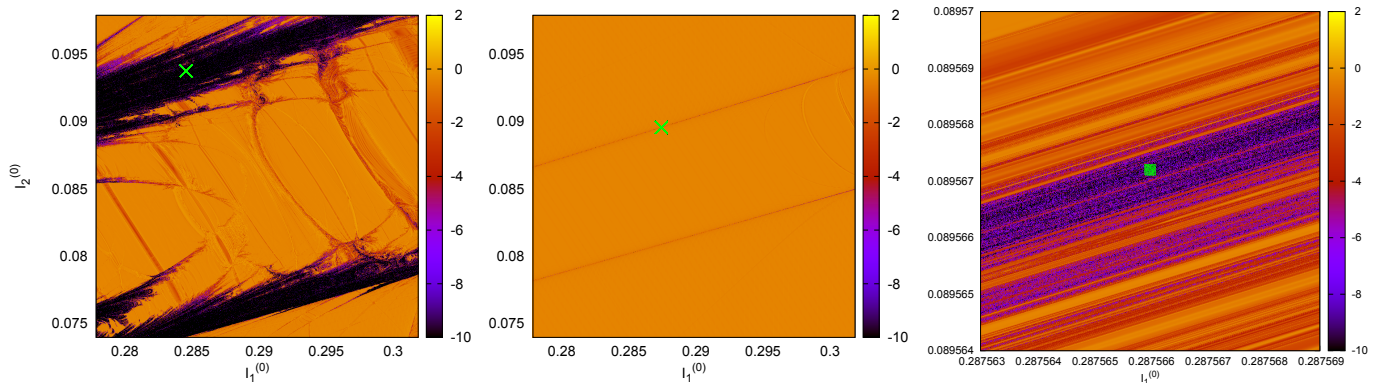
FIG. 3: Location of the ICs for $\epsilon = 0.015$ (left panel) and $\epsilon = 0.003$ (middle and right panels). The green square in the third plot shows the size and shape of the ensemble.

C. The measurements

The numerical integrations of the trajectories were performed with a 8th order symplectic integrator called S8b and elaborated by Teloy, Freiburg, as mentioned in [45]. We use a double precision arithmetic in cartesian variables considering a time step $\Delta t_{int} = 10^{-2}$. For all $\epsilon \in \mathcal{E}$, the ensembles were numerically evolved up to the time 10^7.

Before computing the statistical parameters for all the ϵ values, we qualitatively show the dynamics of these ensembles in the original action space, for $\epsilon = 0.012$ and $\epsilon = 0.015$. We use a double section technique, applied in [29] among others. While integrating the test particles we consider the surface of section $x_1 = 0$ ($y_1 > 0$) and collect those points at which the intersecting orbit also satisfies the condition $x_2^2 + x_3^2 \leq \delta^2$ ($y_2 > 0$, $y_3 > 0$), with $\delta = 0.002$. In terms of angle variables, the double section is equivalent to $\theta_1 = \pi/2$ with both θ_2 and θ_3 belonging to a certain neighborhood of $\pi/2$, whose size decays to zero with δ.

In Fig. 4 we plot, for the smaller perturbation parameter, all the intersections with the double section starting at $t = 0$ up to four final times: $t = 10^5$, 5×10^5, 10^6, 5×10^6.

Similarly, in Fig. 5 we plot, for a larger perturbation parameter, all the intersections with the double section. There we can see how the ensemble expands along the stochastic layer of the guiding resonance.

For the whole set $\mathcal{E}$, we have computed the evolution of the variance of the three components of $\boldsymbol{p}^{(0)}$, as given in Eq. (51). The left panel of Fig. 6 displays $\bar{\sigma}_j^2(t)$ for $j = 1, 2, 3$, for $\epsilon = 0.015$ in colors red, green and blue, respectively. We can see that both $\bar{\sigma}_1^2(t)$ and $\bar{\sigma}_2^2(t)$ are bounded quantities within this time interval, as expected. We also notice that, on average, $\bar{\sigma}_3^2(t)$ has a secular growth. The initial variance is $\bar{\sigma}_3^2(0) \approx 7.2 \times 10^{-15}$ and has a relatively large jump that starts at $t \approx 900$. The behavior for $\epsilon = 0.005$ is shown in the right panel of Fig. 6. For the rest of the ϵ values, the observed behaviors are qualitatively similar to the above exposed ones.

From now on we work only with the variance in the $\boldsymbol{\mu}_3$ direction, comparing its time evolution as computed using the original action variables or the ones corresponding to the optimal canonical transformation. Figs. 7(a), 7(c) and 7(e) show, respectively for $\epsilon = 0.015$, 0.010 and 0.005, the values of $\bar{\sigma}_3^2(t)$, in blue, and $\sigma_3^2(t)$, in black. Figs. 7(b), 7(d) and 7(f) show only the value of $\sigma_3^2(t)$ for the same perturbation parameters.

In Fig. 7(a) the main observation is that the variance $\bar{\sigma}_3^2$ computed in the original canonical variables exhibits significant fluctuations, while σ_3^2 (in the new canonical variables) evolves in a smooth linear way without any initial jump. Despite this difference, it is remarkable that, for $\epsilon = 0.015$, both quantities exhibit a quite similar average slope in time. Now let us focus on $\epsilon = 0.010$ (Fig. 7(c)). We see that $\bar{\sigma}_3^2$ still has a measurable average slope and equal to the one of σ_3^2. However, the relative difference between the two quantities is higher than in the previous case. In fact, the fluctuations in $\bar{\sigma}_3^2$ are such that the slow systematic time variation is just visible in these variables, up to $t = 10^7$.

This picture changes dramatically for smaller values of ϵ, like 0.005. Inspecting the right panel of Fig. 6, we cannot distinguish any measurable secular growth in $\bar{\sigma}_3^2$. In fact, the growth exists, but it is completely "hidden" by the large variations due to the 'deformation' effects. Thus, in practice the diffusion rate associated to $p_3^{(0)}$ cannot be experimentally measured. Figs. 7(e) and 7(f) show that the variance in the new variable p_3 is many orders of magnitude smaller than $\bar{\sigma}_3^2$. Furthermore, all fluctuations due to deformation effects are absorbed by the normalizing transformation, and do not show up in the time evolution of σ_3^2. This allows to identify and measure the diffusion rate, in this case using *only* the new canonical variables.

We have not introduced yet any assumption about the dependence of $\sigma_3^2(t)$ on time. The closeness to normal diffusion can be graphically estimated in Fig. 8, displaying the values of $\sigma_3^2(t)$ in logarithmic scale $\forall \epsilon \in \mathcal{E}$. We have fitted power laws, according to the ansatz given at the end of Subsec. V A, for each element of $\mathcal{E}$, obtaining the values of the exponent η shown in Table V. These values are not far from $\eta \approx 1$. Thus, from Figs. 7(b), 7(d) and 7(f), as well as the corresponding plots for the rest of the values of ϵ, that present a similar behavior, we conclude that the assumption of normal diffusion is a good first approximation. However, in some cases we find values appreciably smaller than 1, i.e., a slight sub-diffusive behavior. This we propose as a subject for future study.

ϵ	0.003	0.005	0.007	0.008	0.010	0.012	0.013	0.015	0.018	0.020
η	0.75	0.87	0.91	0.95	0.91	1.10	1.18	0.98	0.89	0.86

TABLE V: Power law behavior: values of exponent for the ansatz $\sigma_3^2(t) = ct^\eta$ for the interval $t \in [10^4, 10^7]$.

We should note at this point that, although Chirikov's estimates of the theoretical diffusion coefficient concern time-averages as in (A12), the assumption that time-averages and space-averages are similar is used along his analytical derivation of D_C. In fact, the aim of introducing the reduction factor F to take into account heuristically the correlations of the driving phases at the borders of the stochastic layer, but assuming that the time-variance grows linearly with time. However, a numerical derivation of D is in practice only possible considering ensemble averages. In the lack of sufficient information about the correlations of the driving phases, we will then simply set $F = 1$ in the comparison of D with D_C below.

The fitted values of the numerical diffusion coefficients D, computed according to Eq. (52) using two different time
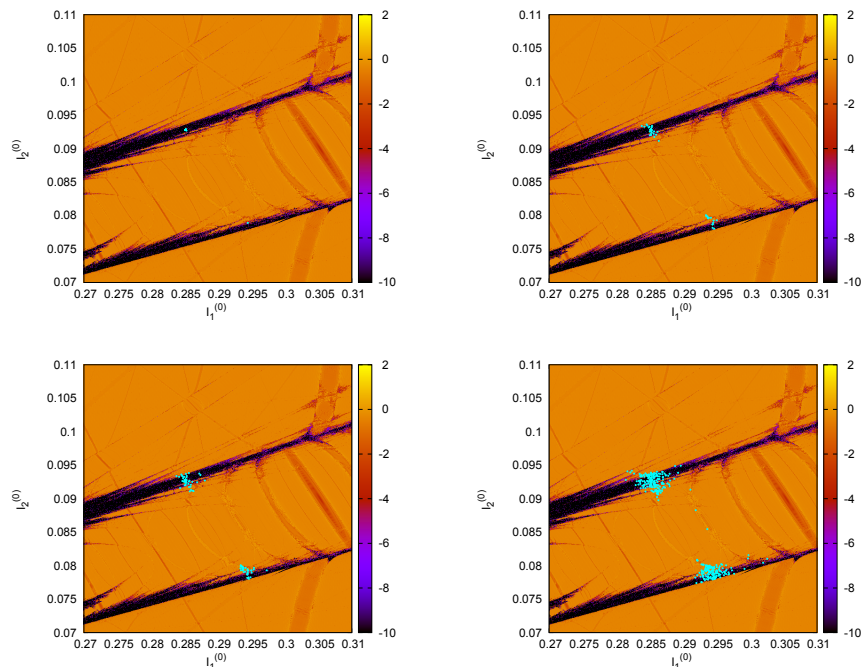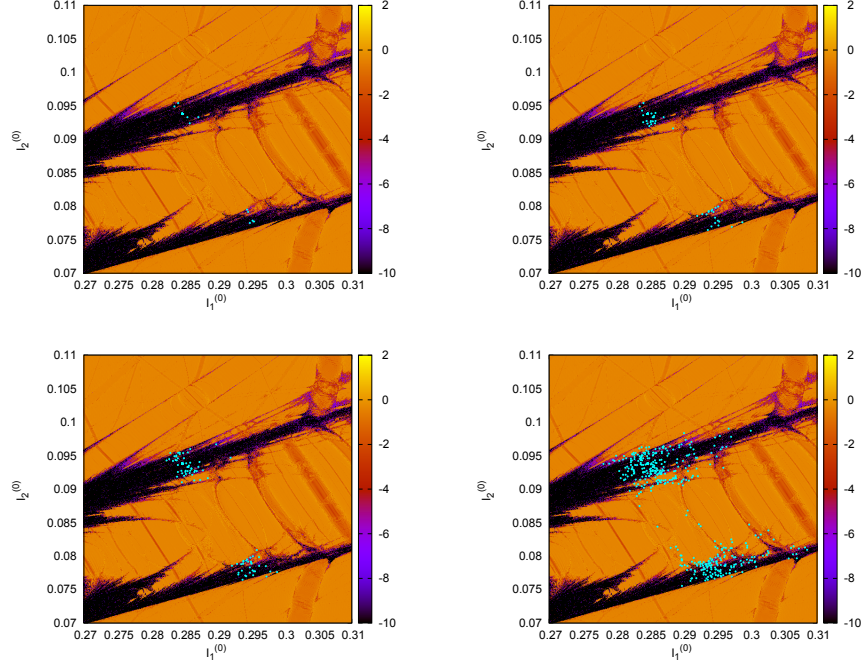

FIG. 4: Intersection of the trajectories with the double section defined by $x_1 = 0$, $(x_2^2 + x_3^2)^{1/2} \leq 0.002$ and $y_i > 0$ ($i = 1, 2, 3$), projected onto the $[I_1, I_2]$ plane for times: $t \leq 10^5$ (top–left), $t \leq 5 \times 10^5$ (top–right), $t \leq 10^6$ (bottom–left) and $t \leq 5 \times 10^6$ (bottom–right). The data corresponds to $\epsilon = 0.012$.
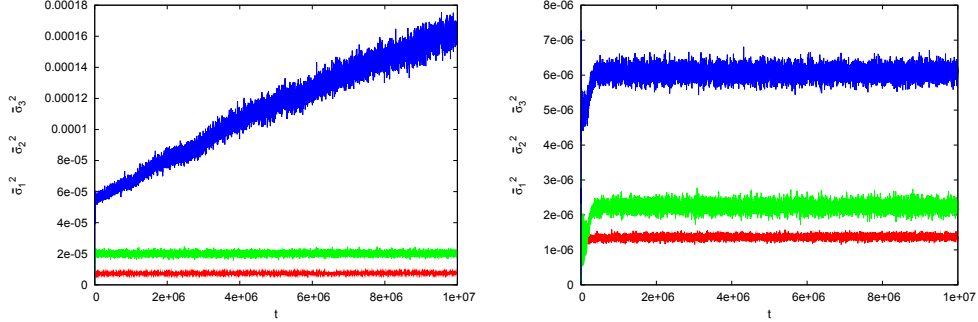
FIG. 5: Intersection of the trajectories with the double section defined by $x_1 = 0$, $(x_2^2 + x_3^2)^{1/2} \leq 0.002$ and $y_i > 0$ ($i = 1, 2, 3$), projected onto the $[I_1, I_2]$ plane for times: $t \leq 10^5$ (top–left), $t \leq 5 \times 10^5$ (top–right), $t \leq 10^6$ (bottom–left) and $t \leq 5 \times 10^6$ (bottom–right). The data corresponds to $\epsilon = 0.015$.

FIG. 6: The left panel displays the variance evolution in the original actions for $\epsilon = 0.015$ along the three directions of the new basis: $\bar{\sigma}_j^2(t)$, for $j = 1, 2, 3$, in colors red, green and blue, respectively. $\bar{\sigma}_1^2(t)$ and $\bar{\sigma}_2^2(t)$ are bounded quantities within this time interval while $\bar{\sigma}_3^2(t)$ presents a secular growth. The right panel displays, for $\epsilon = 0.005$, the values of $\bar{\sigma}_j^2(t)$, for $j = 1, 2, 3$, in colors red, green and blue, respectively. The three quantities appear to be bounded in the considered time interval.

intervals: $[0, 10^7]$ and $[10^6, 10^7]$, are shown in the second and third columns of Table VI, respectively. Both computations of D yield quite similar values. In the fourth column we include the theoretical estimation D_C, given in table IV, for comparison. In Fig. 9–left we superpose all different estimates of the diffusion coefficient in semilogarithmic scale. The values of D fitted with $0 \leq t \leq 10^7$ and D_C, are displayed in colors red and green, respectively. Both the theoretical and the numerical coefficients have nearly the same functional behavior with respect to ϵ. Moreover, we notice that $0.1 \lesssim D/D_C \lesssim 3.4$ for $\epsilon \in \mathcal{E}$, which implies that the theoretical and numerical estimates agree rather well; the lower bound 0.1 corresponding to $\epsilon = 0.003$ as can be seen from Figure 9–middle. The dots in this figure correspond to the values $D = D(D_C)$ in logarithmic scale, for $\epsilon \in \mathcal{E}$. Considering only parameter values: $0.005 \leq \epsilon \leq 0.015$, we made a least square fit of the ansatz $\log(D) = q \log(D_C) + r$, obtaining an exponent $q = 1.08333$. Thus, we conclude that the relation $D \sim D_C$ essentially holds true. Notice that while performing the fit of the coefficients we explicitly discarded the values for $\epsilon = 0.003, 0.015, 0.020$. In fact, in the case of the lowest ϵ value, there are several terms in the
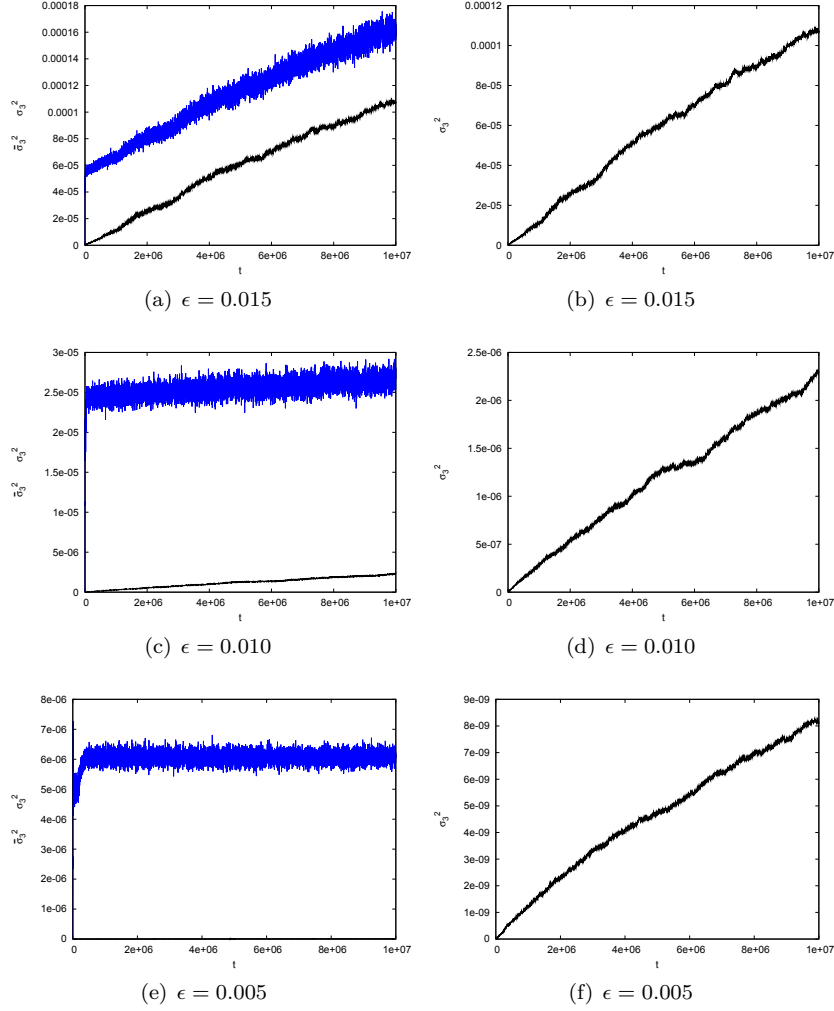
FIG. 7: Variance evolution in the original and optimal actions for $\epsilon = 0.015$, 0.010 and 0.005. The left hand plots display both $\bar{\sigma}_3^2(t)$ and $\sigma_3^2(t)$, in colors blue and black, respectively. The right hand plots display only $\sigma_3^2(t)$. We observe a transition in which diffusion can be measured in both families of actions obtaining the same average result ($\epsilon = 0.015$), towards a state in which diffusion is undetectable in the old actions but precisely measurable in the correct ones ($\epsilon = 0.005$). The case of $\epsilon = 0.010$ is qualitatively similar to the one for $\epsilon = 0.015$ but is close to the limit of undetectability in the old variables.

perturbation of the very same order of magnitude, so that the layer resonance, that should have a coefficient W_l much larger than the ones of the leading driving resonances, is actually not well defined accordingly to Chirikov's theory (see Table II and related discussion). Meanwhile, for the largest ϵ values, an overlap between the layer resonance and some very close high order resonances is observed, since the variance in p_1, σ_1, presents at large times, a slight increasing behavior, indicating that the width of the stochastic layer is not bounded, in contradiction with Chirikov's formulation.

As a final estimate, we have performed several fits of the ansatz D, $D_C \propto ||R^{opt}||^b$, considering different subsets of $\mathcal{E}$. Although the results depend on the chosen subset, the value $b \approx 2.5$ represents the mean situation. Fig. 9–right shows the corresponding behavior of D_C (red) and D (black). The green line corresponds to the function $||R_{opt}||^{2.5}$, and it is displayed as a 'guide to the eye' law for comparison with the laws found for various estimates of the diffusion coefficient.
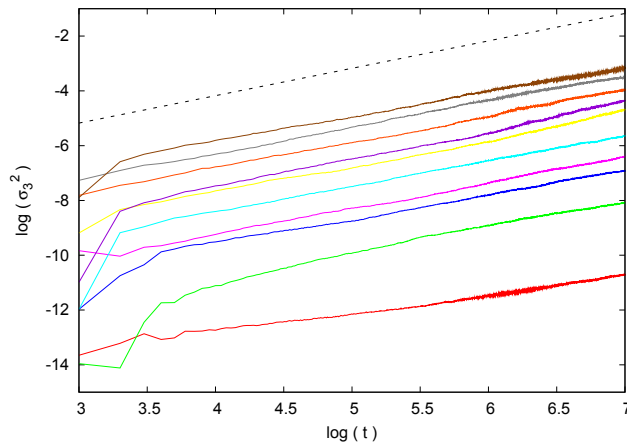
FIG. 8: Variance evolution for all the parameter values in logarithmic scale, for $t \in [10^3, 10^7]$. The relation ϵ–color is: 0.003–red, 0.005–green, 0.007–blue, 0.008–magenta, 0.010–cyan, 0.012–yellow, 0.013–violet, 0.015–orange, 0.018–grey, 0.003–brown. The slopes have been fitted according to ansatz $\sigma_3^2(t) = ct^\eta$, for the interval $t \in [10^4, 10^7]$, obtaining the values of η given in Table V. The dashed curve corresponds to a straight line of unitary slope plotted for comparison. For most of ϵ values the diffusion is close to be normal.

ϵ	$D\ (0 < t < 10^7)$	$D\ (10^6 < t < 10^7)$	D_C
0.020	6.7×10^{-11}	7.0×10^{-11}	8.4×10^{-11}
0.018	3.5×10^{-11}	3.5×10^{-11}	3.0×10^{-11}
0.015	1.1×10^{-11}	1.1×10^{-11}	5.0×10^{-12}
0.013	4.2×10^{-12}	4.2×10^{-12}	1.3×10^{-12}
0.012	1.9×10^{-12}	1.9×10^{-12}	5.9×10^{-13}
0.010	2.3×10^{-13}	2.3×10^{-13}	1.2×10^{-13}
0.008	4.0×10^{-14}	4.0×10^{-14}	1.9×10^{-14}
0.007	1.3×10^{-14}	1.3×10^{-14}	6.9×10^{-15}
0.005	8.9×10^{-16}	8.9×10^{-16}	6.7×10^{-16}
0.003	2.1×10^{-18}	2.1×10^{-18}	2.4×10^{-17}

TABLE VI: Diffusion coefficients for $\epsilon \in \mathcal{E}$. The second and third columns give the values of the numerical coefficient, D, using the time intervals $[t_0, t_f]$ equal to $[0, 10^7]$ and $[10^6, 10^7]$, respectively. In the fourth column we have added the theoretical estimation, D_C, given in table IV.

VI. CONCLUSIONS

In the present paper, we presented theoretical and numerical results pointing towards an important connection between the estimates for the diffusion rate along simple resonances in multidimensional nonlinear Hamiltonian systems that can be obtained using i) the theory of Chirikov [8], and ii) the theory of Nekhoroshev [43]. We emphasized that, despite a common impression, the two theories are complementary rather than antagonist. In fact, we exploited this complementarity in order to obtain accurate theoretical predictions for the value of the diffusion coefficient along a resonance in a particular numerical example. Our main conclusions can be summarized as follows:

1) The theory of Chirikov *requires*, as a starting point, the construction of a simply-resonant normal form valid in local domains of the action space, which has to be optimal in the Nekhoroshev sense. In particular, the so-called driving terms of Chirikov's theory are identified with the remainder terms of the optimal normal form construction. Then, for small enough nonlinear perturbations ϵ, the size of the driving terms turns to be of order $O(\exp(-1/\epsilon^a))$ (for some exponent $a > 0$).

2) We constructed the above optimal normal form in the so-called 3D quartic oscillator model [6], using a computer-algebraic program in order to implement the normalization algorithm suggested in [9]. We were able to reach the optimal normalization order, at which the size of the normal form remainder becomes the least possible, and also to
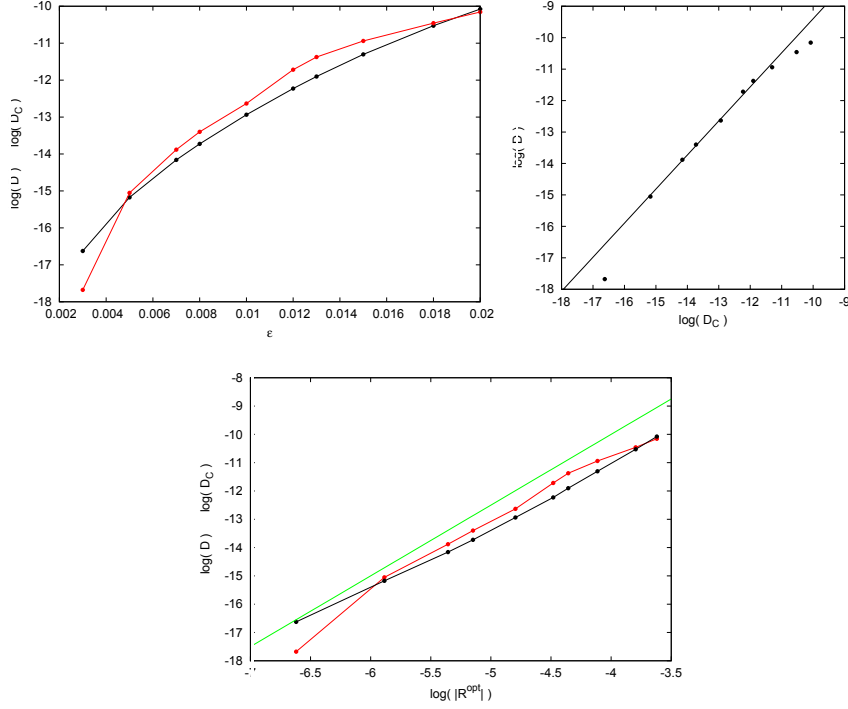
FIG. 9: The left side panel shows the numerical and theoretical estimates of the diffusion coefficients extracted from table VI: D ($0 \leq t \leq 10^7$) in red and D_C in black. The middle panel shows the comparison between the numerical and the theoretical diffusion coefficients. The line is the least square fit of the ansatz $\log(D) = q\log(D_C) + r$, with $q = 1.08333$. The right side panel shows the diffusion coefficients versus the norm of the optimal remainder. The green line corresponds to the functions $||R_{opt}||^{2.5}$ and is displayed for comparison with the diffusion coefficients. It can be seen that both coefficients have an approximate functional trend of the shape $D, D_C \sim ||R_{opt}||^{2.5}$, in black and red respectively.

observe the expected asymptotic character of the normal form series.

3) We used the computed expression for the optimal normalized Hamiltonian in order to transform all data in the basis of action variables suggested by Chirikov (see [4]). In this way, we identified the main resonant terms in the latter theory corresponding to i) the layer resonance, and ii) the driving resonances. With the above information at hand, we finally implemented Chirikov's formulae and computed a theoretical estimate D_C for the diffusion coefficient along the resonance.

4) We compared the values of D_C with a purely numerical measurement of the diffusion coefficient using ensembles of orbits integrated along the resonance's chaotic layer up to a quite long time. We found that for small perturbation values ϵ, the diffusion can only be measured after subtracting, from the numerical orbital data, the so-called 'deformation' effects. This requires transforming the data in new canonical variables arising by the same normalizing transformation that leads to the construction of the optimal normal form.

5) We compared the theoretical prediction D_C with the numerical value of the diffusion coefficient D for various values of ϵ. We found a quite satisfactory agreement, so that essentially one has $D_C \propto D$. The coefficient of proportionality depends on the so-called reduction factor of Chirikov's theory. In the lack of sufficient information about the value of the reduction factor, we simply set it equal to unity, which proves to be an adequate approximation.

6) We pointed out that in the framework of Chirikov's theory one obtains a power-law relation between D and the size of the optimal remainder $||R_{opt}||$, i.e. we have $D \sim ||R_{opt}||^{2+p}$, where p is a positive constant of order unity. This is in agreement with the analysis made in [13]. We applied this relation by a power-law fitting on our numerical data. We found that the heuristic law $D \sim ||R_{opt}||^{2.5}$ is adequately precise for all practical purposes. This latter fact allows to estimate directly the value of the diffusion coefficient using only normal form data.

7) Finally, we made a preliminary study of the character of the diffusion along a resonance in the weakly chaotic regime. Although to a first approximation the diffusion can be characterized as normal, we found secondary features in all diffusion curves, that we attribute to the passage of some orbits of our ensemble from crossing points with secondary resonances. Furthermore, we tried power-law fitting of all the diffusion curves $\sigma^2 \sim t^w$, where σ^2 is the variance of the ensemble in the action variables suggested by Chirikov. We found values of w close to 1, but with a slight preference towards values smaller than unity. Thus, we conclude that the overall effect of secondary resonances

is to render the chaotic spreading slightly sub-diffusive. This subject, however, necessitates a focused study that is proposed as a future subject.

Acknowledgements

[1] V. I. Arnold, Instability of dynamical systems with several degrees of freedom, Sov. Math.-Dokl. 5 (1964) 581-585.

[2] G. Benettin, L. Galgani, A. Giorgilli, A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems, Celestial Mechanics and Dynamical Astronomy 37 (1985) 1-25.

[3] F. Cachucho, P. M. Cincotta, S. Ferraz–Mello, Chirikov diffusion in the asteroidal three-body resonance (5,-2,-2), Celestial Mechanics and Dynamical Astronomy 108 (2010) 35-58.

[4] P. M. Cincotta, Arnold diffusion: an overview through dynamical astronomy, New Astronomy Reviews 46 (2002) 13-39.

[5] P. M. Cincotta, C. M. Giordano, Diffusion in multidimensional Hamiltonian Systems. Application to Galactic Dynamics, in P.M. Cincotta, C.M. Giordano, & C. Efthymiopoulos (eds), Proceedings of the 3rd La Plata School on Astronomy and Geophysics, WSAAA, (2012) 185-206 .

[6] P. M. Cincotta, C. M. Giordano, C. Simó, Phase space structure of multi-dimensional systems by means of the mean exponential growth factor of nearby orbits, Physica D 182 (2003) 151-178.

[7] P. M. Cincotta, C. Simó, C., Simple tools to study global dynamics in non-axisymmetric galactic potentials - I, A & AS 147 (2000) 205-228.

[8] B. V. Chirikov, A universal instability of many-dimensional oscillator systems, Physics Reports 52 (1979) 263-379.

[9] C. Efthymiopoulos, On the connection between the Nekhoroshev theorem and Arnold diffusion, Celestial Mechanics and Dynamical Astronomy 102 (2008) 49-68.

[10] C. Efthymiopoulos, Canonical Perturbation Theory, stability and Diffusion in Hamiltonian systems: Applications to Dynamical Astronomy, in P.M. Cincotta, C.M. Giordano, & C. Efthymiopoulos (eds), Proceedings of the 3rd La Plata School on Astronomy and Geophysics, WSAAA, (2012) 3-143.

[11] C. Efthymiopoulos, C. Kalapotharakos, N. Voglis, Special features in galactic dynamics in Topics in Gravitational Dynamics, Lect. Notes Phys. 729, D. Benest, C. Froeschlé & E. Lega eds., Springer, Berlin, (2007) 297-389.

[12] C. Efthymiopoulos, A. Giorgilli, G. Contopoulos, Nonconvergence of formal integrals: II. Improved estimates for the optimal order of truncation, JPhyA 37 (2004) 10831-10858.

[13] C. Efthymiopoulos, M. Harsoula, The speed of Arnold diffusion, Physica D, 251 (2013) 19-38.

[14] S. Ferraz-Mello, Canonical Perturbation Theories. Degenerate Systems and Resonance, Springer, New York, 2007.

[15] C. Froeschlé, Numerical Study of Dynamical Systems with Three Degrees of Freedom. I. Graphical Displays of Four-Dimensional Sections, Astron. Astrophys. 4 (1970a) 115-128.

[16] C. Froeschlé, Numerical Study of Dynamical Systems with Three Degrees of Freedom. II. Numerical Displays of Four-Dimensional Sections, Astron. Astrophys. 5 (1970b) 177-183.

[17] C. Froeschlé, Numerical Study of a Four-Dimensional Mapping, Astron. Astrophys. 16 (1972) 172-189.

[18] C. Froeschlé, M. Guzzo, E. Lega, Local and global diffusion along resonant lines in discrete quasi-integrable dynamical systems, Celestial Mechanics and Dynamical Astronomy 92 (2005) 243-255.

[19] C. Froeschlé, E. Lega, On the structure of the Hamiltonian systems. The Fast Lyapunov Indicator: a new very sensitive tool, Bulletin of the American Astronomical Society 32 (2000) 1115.

[20] C. Froeschlé, I.P. Scheidecker, Numerical Study of a Four-Dimensional Mapping. II, Astron. Astrophys. 22 (1973) 431-436.

[21] C. M. Giordano, P. M. Cincotta, Chaotic diffusion of orbits in systems with divided phase space, Astron. Astrophys. 423 (2004) 745-753.

[22] A. Giorgilli, Notes on exponential stability of Hamiltonian systems, in Dynamical Systems. Part I: Hamiltonian Systems and Celestial Mechanics, Pubblicazioni della Classe di Scienze, Scuola Normale Superiore, Pisa, (2002).

[23] M. Guzzo, E. Lega, C. Froeschlé, First numerical evidence of global Arnold diffusion in quasi–integrable systems, Disc. Cont. Dyn. Syst. B, 5 (2005) 687-698.

[24] M. Guzzo, E. Lega, C. Froeschlé, Diffusion and stability in perturbed non-convex integrable systems, Nonlinearity, 19 (2006) 1049-1067.

[25] K. Kaneko, T. Konishi, Diffusion in Hamiltonian dynamical systems with many degrees of freedom, Phys. Rev. A, 40 (1989) 6130-6133.

[26] J. Laskar, Frequency analysis for multi-dimensional systems. Global dynamics and diffusion, Physica D, 67 (1993) 257-281.

[27] H.S. Dumas, J. Laskar, Global dynamics and long-time stability in Hamiltonian systems via numerical frequency analysis Phys. Rev. Let., 70 (1993) 2975-2979.

[28] E. Lega, C. Froeschlé, On the Relationship between Fast Lyapunov Indicator and Periodic Orbits for Symplectic Mappings, in Dynamics of Natural and Artificial Celestial Bodies, Halina Pretka-Ziomek, Edwin Wnuk, P. Kenneth Seidelmann & David Richardson eds., Kluwer Academic Publishers, Dordrecht, (2002) 129-147.

[29] E. Lega, M. Guzzo, C. Froeschlé, Detection of Arnold diffusion in Hamiltonian systems, Physica D 182 (2003) 179-187.

[30] E. Lega, M. Guzzo, C. Froeschlé, Measure of the exponential splitting of the homoclinic tangle in four-dimensional symplectic mappings, Celestial Mechanics and Dynamical Astronomy 104 (2009) 191-204.

[31] E. Lega, M. Guzzo, C. Froeschlé, A numerical study of the size of the homoclinic tangle of hyperbolic tori and its correlation with Arnold diffusion in Hamiltonian systems, Celestial Mechanics and Dynamical Astronomy 107 (2010) 129-147.

[32] E. Lega, M. Guzzo, C. Froeschlé, A numerical study of the hyperbolic manifolds in a priori unstable systems. A comparison with Melnikov approximations, Celestial Mechanics and Dynamical Astronomy 107 (2010) 115-127.

[33] E. Lega, M. Guzzo, C. Froeschlé, Numerical studies of hyperbolic manifolds supporting diffusion in symplectic mappings, EPJ. Special Topics 186 (2010) 3-31.

[34] M. A. Lieberman, Arnold diffusion in Hamiltonian systems with three degrees of freedom, Annals of the New York Academy of Sciences, 357 1 (1980) 119-142.

[35] P. Lochak, Canonical perturbation theory via simultaneous approximation, Russian Mathematical Surveys 47 (1992) 59-140.

[36] P. Lochak, in C. Simó (ed.), Hamiltonian Systems with Three or More Degrees of Freedom, NATO ASI Series C, 533, Kluwer, Dordrecht, (1999)168.

[37] G. Lukes-Gerakopoulos, N. Voglis, C. Efthymiopoulos, The production of Tsallis entropy in the limit of weak chaos and a new indicator of chaoticity, Physica A 387 (2008) 1907-1925.

[38] M. F. Mestre, Difusión caótica en sistemas Hamiltonianos casi–integrables, PhD Thesis, Facultad de Ciencias Astronómicas y Geofísicas– Universidad Nacional de La Plata, (2012).

[39] M. F. Mestre, P. M. Cincotta, C. M. Giordano, Testing the accuracy of the overlap criterion IJNLM 44 (2009) 180-189.

[40] M.F. Mestre, P. M. Cincotta, C. M. Giordano, Diffusion measurements in a 3DoF Hamiltonian flow, in P.M. Cincotta, C.M. Giordano, & C. Efthymiopoulos (eds), Proceedings of the 3rd La Plata School on Astronomy and Geophysics, WSAAA, (2012) 319-344.

[41] A. Morbidelli, A. Giorgilli, On the role of high order resonances in normal forms and in separatrix splitting, Physica D 102 (1997) 195-207.

[42] A.I. Neishtadt, The separation of motions in systems with rapidly rotating phase, J. Appl. Math. Mech., (1984) 48 133-139.

[43] N. N. Nekhoroshev, An Exponential Estimate of the Time of Stability of Nearly-Integrable Hamiltonian Systems, Russian Mathematical Surveys 32 (1977) 1-65.

[44] J. Pöshel, Nekhoroshev estimates for quasi-convex Hamiltonian systems, Math. Z., 213, (1993) 187-216.

[45] Ch. Schlier, A. Seiter, High-order symplectic integration: an assessment, CoPhC 130 (2000) 176-189.

[46] Ch. Skokos, Alignment indices: a new, simple method for determining the ordered or chaotic nature of orbits, JPhA 34 (2001) 10029-10043.

[47] Ch. Skokos, G. Contopoulos, C. Polymilis, Structures in the Phase Space of a Four Dimensional Symplectic Map, Celestial Mechanics and Dynamical Astronomy 65 (1997) 223-251.

[48] B. V. Chirikov, J. Ford, J. F. Vivaldi, Some numerical studies or Arnold diffusion in a simple model, American Institute of Physics Conference Series 57 (1980) 323-340.

[49] J. L. Tennyson, Resonance transport in near-integrable systems with many degrees of freedom Physica D 5 (1982) 123-135.

[50] J. L. Tennyson, M. A. Lieberman, A. J. Lichtenberg, Diffusion in near-integrable Hamiltonian systems with three degrees of freedom, American Institute of Physics Conference Series 57 (1980) 272-301.

[51] B. P. Wood, A. J. Lichtenberg, M. A. Lieberman, Arnold diffusion in weakly coupled standard maps, Phys. Rev. A 42 (1990) 5885-5893.

Appendix A: Chirikov's main derivations

We start from the expression for the Hamiltonian (48), given in section IV namely,

$$H(\boldsymbol{p}, \boldsymbol{\psi}) \;=\; \frac{p_1^2}{2M_G} + |\boldsymbol{\omega}^r|p_2 + \sum_{l=1}^{3} \sum_{k+l>2}^{3} \frac{p_k p_l}{2M_{kl}} + \epsilon V_G \cos\psi_1 + \epsilon \sum_{\boldsymbol{m}} \tilde{V}_{\boldsymbol{m}} \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}(\boldsymbol{\psi}))$$

where the coefficients $\tilde{V}_{\boldsymbol{m}}$ have constant values.

In absence of perturbation ($\tilde{V}_{\boldsymbol{m}} = 0$), the components $p_k, k = 2, 3$ are local integrals of motion, whose value is equal to zero if $\boldsymbol{I}^r$ is a point of the orbit. Then, the Hamiltonian reduces to:

$$H(\boldsymbol{p}, \boldsymbol{\psi}) \approx H_1(p_1, \psi_1) + \epsilon \tilde{V}(\boldsymbol{\psi}), \tag{A1}$$

where:

$$H_1 = \frac{p_1^2}{2M_G} + \epsilon V_G \cos \psi_1 \tag{A2}$$

is the pendulum Hamiltonian for the guiding resonance, and the perturbing phases $\boldsymbol{\theta}$ in $\tilde{V}$ are written in terms of the new components ψ_k.

To transform the phase variables, we take into account that the dot product is invariant under a change of basis. Recalling that $\psi_k = \sum_l \Upsilon_{kl}\theta_l$ then, if $\boldsymbol{\nu}$ denotes the vector $\boldsymbol{m}$ in the new basis, we have: $\varphi_{\boldsymbol{m}} \equiv \boldsymbol{m} \cdot \boldsymbol{\theta} = \boldsymbol{\nu} \cdot \boldsymbol{\psi}$, where $\nu_k = \sum_i m_i \Upsilon_{ik}$. As we can readily see, while the m_k are integers, the quantities ν_k are, in general, real numbers.

As mentioned above, for $\tilde{V} = 0$ the p_k are local integrals of motion and recalling that H_1 is also an unperturbed integral, we have the full set of three local integrals: H_1, p_2, p_3. But if we switch on the perturbation, they will change with time. This variation is determined by the time dependence of $\varphi_{\boldsymbol{m}}$. To get $\varphi_{\boldsymbol{m}}(t)$ we evaluate the dot product $\boldsymbol{\nu} \cdot \boldsymbol{\psi}$:

$$\varphi_{\boldsymbol{m}}(t) = \boldsymbol{m} \cdot \boldsymbol{\theta} = \boldsymbol{\nu} \cdot \boldsymbol{\psi} \approx \xi_{\boldsymbol{m}}\psi_1(t) + \omega_{\boldsymbol{m}}t + \beta_{\boldsymbol{m}}, \tag{A3}$$

where:

$$\xi_{\boldsymbol{m}} = \sum_{k=1}^{3} \frac{\nu_k(\boldsymbol{m})}{M_{k1}}, \quad \omega_{\boldsymbol{m}} = \boldsymbol{m} \cdot \boldsymbol{\omega}^r = \nu_2(\boldsymbol{m})|\boldsymbol{\omega}^r|; \tag{A4}$$

and $\beta_{\boldsymbol{m}}$ is a constant.

It can be found that the change in the unperturbed integrals over a half period of oscillation T, given by

$$T(w) = \frac{1}{\Omega_G} \ln(32/|w|), \qquad w = \frac{H_1}{\epsilon V_G} - 1, \tag{A5}$$

is

$$\Delta p_i \approx \frac{\epsilon}{\Omega_G} \sum_{\boldsymbol{m}} \nu_i(\boldsymbol{m})Q_{\boldsymbol{m}}\sin\varphi_{\boldsymbol{m}}^0; \qquad i = 2, 3, \tag{A6}$$

while

$$\Delta H_1 \approx -\frac{\epsilon|\boldsymbol{\omega}^r|}{\Omega_G} \sum_{\boldsymbol{m}} \frac{\nu_1(\boldsymbol{m})\nu_2(\boldsymbol{m})}{\xi_{\boldsymbol{m}}}Q_{\boldsymbol{m}}\sin\varphi_{\boldsymbol{m}}^0. \tag{A7}$$

The variation of the integrals depends on $Q_{\boldsymbol{m}} = \tilde{V}_{\boldsymbol{m}}A_{2|\xi_{\boldsymbol{m}}|}(\lambda_{\boldsymbol{m}})$ where $A_m(\lambda)$ denotes the Melnikov–Arnold integral given by

$$A_m(\lambda) = \int_{-\infty}^{\infty} d\hat{t} \cos\left(\frac{m}{2}\psi^s(\hat{t}) - \lambda\hat{t}\right), \qquad \psi^s(\hat{t}) = 4\arctan\left(e^{\hat{t}}\right),$$

whose asymptotic value for large λ, is

$$A_m(\lambda) \approx \frac{4\pi(2\lambda)^{m-1}}{(m-1)!}e^{-\pi\lambda/2}, \qquad \lambda \gg m. \tag{A8}$$

The factorial should be replaced by the Gamma function, $\Gamma(m)$, for non–integer m. For details, see for instance [14], [4].

Having already attained the change in the integrals after a period of motion T, we need to compute the variation for the phases $\varphi_{\boldsymbol{m}}^0$ which are the quantities (A3) evaluated in the separatrix, over the same time interval, in order to obtain a map describing Arnold diffusion. We have

$$\Delta\varphi_{\boldsymbol{m}}^0 = \xi_{\boldsymbol{m}}\Delta\psi_1^s(t_0) + \omega_{\boldsymbol{m}}\Delta t_0 = \omega_{\boldsymbol{m}}T(w) + C_{\boldsymbol{m}},$$

where $C_{\boldsymbol{m}}$ is a constant. Rewriting (A7) in terms of the dimensionless energy w instead of H_1, we arrive at the following map:

$$\bar{w} = w - \frac{|\boldsymbol{\omega^r}|}{\Omega_G} \sum_{\boldsymbol{m}} W_{\boldsymbol{m}} \sin \varphi_{\boldsymbol{m}}^0, \tag{A9}$$

$$\bar{\varphi}_{\boldsymbol{m}}^0 = \varphi_{\boldsymbol{m}}^0 + \omega_{\boldsymbol{m}} T(\bar{w}) + C_{\boldsymbol{m}}, \tag{A10}$$

where

$$W_{\boldsymbol{m}} = \frac{\nu_1(\boldsymbol{m})\nu_2(\boldsymbol{m})Q_{\boldsymbol{m}}}{\xi_{\boldsymbol{m}} V_G}$$

is the very same coefficient given by (49). In the above map the bar indicates the values of the variables after crossing the surface $\psi_1 = \pm\pi$.

The mapping given by (A9)–(A10) is, in some sense, similar to the Whisker or Separatrix map, which has the following expression

$$\bar{w} = w + W \sin \tau_0, \qquad \bar{\tau}_0 = \tau_0 + \omega_{\boldsymbol{l}} T(\bar{w}) \mod(2\pi), \tag{A11}$$

where W is a perturbation parameter like the $W_{\boldsymbol{m}}$'s, $\omega_{\boldsymbol{l}} = \boldsymbol{m_l} \cdot \boldsymbol{\omega}$ and τ_0 is the phase of the perturbation. This map describes the motion in the vicinity of the separatrix of the pendulum under a perturbation.

A well known result of this mapping is that the separatrix becomes a chaotic layer of width, w_s. In other words, the change of the pendulum energy under a small perturbation turns out to be bounded. Experimentally, this bound, w_s, is due to the strong correlations of the phases τ_0 for large times (we refer to [8] for details).

Taking into account the similarities of the mappings (A9) and (A11), Chirikov argues that the largest term in (A9) leads to the so-called layer resonance. In fact this conjecture seems to be true, considering the results given in Table II.

Finally, the scalar diffusion coefficient along the direction of the vector $\boldsymbol{\mu_3}$ is defined as

$$D = \frac{\overline{\Delta p_3(t)^2}}{T_a}, \tag{A12}$$

where $\overline{X(t)}$ denotes time average and T_a is the mean period of motion within the stochastic layer of the guiding resonance defined by

$$T_a(w_s) \approx \int_0^{w_s} T(w)dw \approx \frac{1}{\Omega_G} \ln(32e/w_s),$$

with $w_s \approx |\boldsymbol{\omega^r}||W_{\boldsymbol{l}}|\omega_{\boldsymbol{l}}|/\Omega_G^2$ and $T(w)$ given by (A5) (see CH79 for details). Therefore, from (A6) for $i = 3$ and (A12) the formula for Chirikov's diffusion coefficient given in (50) can be derived.