Arnold diffusion: an overview through Dynamical Astronomy

Pablo M. Cincotta†

Facultad de Ciencias Astronómicas y Geofísicas, Universidad Nacional de La Plata, Paseo del Bosque, 1900 La Plata, Argentina

The aim of this work is to review the fundamental ideas behind the stability problem, emphasizing the differences between two well–known mechanisms that could lead to chaos, namely overlap of resonances and Arnold diffusion. Here we restrict the discussion to multidimensional autonomous Hamiltonian systems which are of major relevance in Dynamical Astronomy. Arnold diffusion is reviewed in a standard mathematical language, by means of different tools such as heuristic reasoning, graphic and geometrical considerations and numerical experiments. In this direction the pioneer work due to Chirikov (1979) is followed, but including additional notes, further examples and useful discussions that may well illuminate the understanding of Arnold diffusion. We also summarize the main difficulties when coping with this instability, from both the analytical and numerical sides of the problem. The discussion whether Arnold diffusion could play any role in the dynamical evolution of, for instance elliptical galaxies, is also included.
keywords: chaos–diffusion–instabilities–galaxies: dynamics

I. INTRODUCTION

The motion of a planet in the Solar System, an asteroid under the influence of the Sun and Jupiter, a star in a globular cluster or a star in a galaxy are examples of classical problems of Dynamical Astronomy. Each of them has its own difficulties. Let us consider an isolated elliptical galaxy. It contains about 10^{11} stars –with a certain mass spectrum– and gas. In addition, there is observational evidence that many elliptical galaxies rotate. Therefore, the dynamical equations of a test star (written in a rotating frame) should take into account the gravitational interaction among all these stars and the hydrodynamical response of the gas. The latter set of coupled differential equations is not easy to solve. In order to simplify the problem, let us introduce then, step by step, certain plausible assumptions. As the contribution of the gas to the total mass of the galaxy is small and the rotation is believed to be rather slow, we will neglect both effects. Thus, the galaxy is reduced to a purely non–rotating mechanical system. However, the field generated by 10^{11} stars may be a very complicated function of the position. Any attempt to solve the motion of a star under the influence of such a field should be done numerically by means of, for instance, a N-body code. To make any progress in the theoretical approach, further assumptions seems to be necessary. First we restrict the problem to that of the motion of a star in a smooth, mean gravitational field which we assume to be a known function of the position. For galaxies the latter is a good approximation. However, elliptical galaxies seem to be well represented by triaxial potentials where explicit orbits are, in general, unknown. Therefore an additional assumption has to be added. In this direction, we reduce even more the problem to that of a near–integrable Hamiltonian system. By *near–integrable* we mean, roughly, that the actual Hamiltonian is very close to a certain Hamiltonian whose orbits are known explicit functions. In spite of the latter oversimplifications, the remainder problem is not solved yet. Questions such as the stability of the motion are far from being completely understood. During the last forty years mathematicians (and some physicists too) were mainly devoted to developing the mathematical tools to cope with the so–called *stability problem*. Other researchers applied these mathematical tools in different scenarios in order to obtain information about the stability of the motion in physically interesting time–scales, for example, larger than a crossing time and less than a Hubble time in the case of galactic systems.

The stability problem is nowadays one of the main subjects of research in Hamiltonian dynamics and it is usually referred to (in physical or astronomical literature) as 'chaotic or stochastic motion' or simply as 'chaos'. Since the early eighties, an increasing number of papers involving intricate concepts related to chaos began to appear in the astrophysical literature, even though the main body of the theory is almost completely restricted to the mathematicians' domain. In this connection it should be understood that chaos had produced a significant impact in the dynamics of the Solar System while, just in recent years, it has been recognized that chaotic motion may play a fundamental role in, for example, the structure, evolution and dynamics of elliptical galaxies (see Udry and Pfenniger 1988, Merritt and Friedman 1996, Merritt and Valluri 1996, Valluri and Merritt 1998, Merritt 1999). It should be mentioned, however, that for fast rotating systems such as barred spiral galaxies, the presence of a large amount of chaos was observed many years ago (see Martinet 1973, Hasan and Norman 1990 and references therein). Nevertheless, the wide–spread idea was that chaos should be considered only for these special models for rotating systems.

The newcomer to the stability problem will find that almost all its references belong to the mathematical literature. Therefore, certain technical knowledge, for instance, on functional analysis, differential geometry or topology (which are not common in astrophysics) is required to learn the subject. Several advanced books in physics discuss extensively questions related to the stability of dynamical systems, for example, Rasband (1983, 1990), Gutzwiller (1990), Jackson (1991), Lichtenberg and Lieberman (1992) (LL92), Reichl (1992). This list is not intended to be complete; it just includes a few references of recent books with slightly different approaches. However, for those systems that we are mainly interested in, namely, autonomous Hamiltonian systems with more than two degrees of freedom, a detailed discussion is still lacking. In most of these books, the authors almost skip the subject and refer to Chirikov (1979) (CH79). In his pioneer work, Chirikov presents a complete review of this matter in a somewhat heuristic way and using a 'classical' or 'old' mathematical language. In its last section he discuss a distinctive property of N–dimensional autonomous Hamiltonian systems ($N > 2$), the so–called *Arnold diffusion*.

The Arnold diffusion (Arnold 1964, see however the remark at the beginning of Section VII) when seen as a global instability, seems to be closer to a theoretical conjecture rather than to a real physical process. There are several reasons for this. In the analytical domain, the key point is that there are so many unsolved mathematical details that make Arnold diffusion a controversial question. Although the paper by Lochak (1999) deals with the purely mathematical approach, even those readers without a strong mathematical background will realize the technical difficulties in coping with this problem. The scanty numerical evidence reveals that Arnold diffusion may work in certain (somewhat artificial) dynamical systems. In spite of this, there are many references to this instability in the astrophysical and Solar System literature. But in many cases there are differences between the physical interpretation of what Arnold diffusion is and how it works. Among others, Merritt and Valluri (1996) suggested that this mechanism (or what they assume to be Arnold diffusion) might be responsible for the slow mixing process at the latest stages of evolution of an elliptical galaxy. On the other hand, Udry and Pfenniger (1988) discuss the coexistence of different (large) stochastic regions in their 3–dimensional models, in contradiction to what is claimed to be the expected behavior if Arnold diffusion works.

Since the first report about Arnold diffusion more than thirty years ago, the literature about this instability written in an accessible mathematical language is scanty, Chirikov's work being an outstanding exception. Unfortunately, a complete understanding of Chirikov's review is not an easy task, particularly that section related to Arnold diffusion for multidimensional oscillator systems. Therefore, taking into account the actual relevance of chaos in Dynamical Astronomy, I would like to review here some topics of this subject within the framework of the standard language in astrophysics.

Though the present article does not introduce any new concept, it provides some additional notes, useful discussions, further examples and numerical simulations that may well help to clarify Chirikov's point of view on Arnold diffusion. Here we will not consider in detail those questions already discussed in either the above listed books or in Chirikov's review, where the reader interested in rigorous proofs will find an exhaustive list of (rather old but useful) mathematical references.

Below, we discuss Chirikov's approach to the stability problem. First we shall review the main results about both 'non–resonant' and 'resonant' perturbation theory and the geometry in action and frequency spaces. There follows a brief discussion about the transition from regular to stochastic motion. Then we come to the main aspect, Arnold diffusion. These issues will be addressed using the same tools as Chirikov displays, namely, graphic and geometric considerations, numerical simulations, simple models and mathematical formalism at the usual rigor in any astrophysical paper. Finally we discuss the possibility that Arnold diffusion could play some role in the dynamical evolution of a stellar system.

II. SET UP, DEFINITIONS AND NOTATION

The fundamental problem to cope with is that of a system governed by an N–dimensional autonomous Hamiltonian which we assume to have the following form:

$$H(\boldsymbol{I}, \boldsymbol{\theta}) = H_0(\boldsymbol{I}) + \epsilon V(\boldsymbol{I}, \boldsymbol{\theta}), \tag{1}$$

where:

$$\epsilon V(\boldsymbol{I}, \boldsymbol{\theta}) = \epsilon \sum_{\boldsymbol{m}} V_{\boldsymbol{m}}(\boldsymbol{I}) \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}). \tag{2}$$

Here $(\boldsymbol{I}, \boldsymbol{\theta})$ are the usual N–dimensional action–angle coordinates for the unperturbed Hamiltonian H_0, $\boldsymbol{m}$ is an N–dimensional integer vector, $V_{\boldsymbol{m}}$ are certain real functions and the so–called *perturbation parameter*, ϵ, is a real number. The equations of motion are: $\dot{\boldsymbol{\theta}} = \partial H/\partial \boldsymbol{I}$; $\dot{\boldsymbol{I}} = -\partial H/\partial \boldsymbol{\theta}$. The Hamiltonian (1)–(2) is often used to describe many astronomical systems. H_0 resembles to a certain toy model and the perturbation ϵV is added in order to obtain a somewhat more realistic problem. Thus, for example, some elliptical galaxies may be thought as the sum of a smooth spherical potential with or without a central cusp (simulating a mass concentration, a massive body or a black hole) plus a multipolar expansion. H_0 may include the spherical model and possibly the cusp while ϵV takes into account some terms of the multipolar expansion (see also Section VII). Let us assume that the transformation to action–angle coordinates for the unperturbed motion can be performed, which unfortunately, is not always the case for an arbitrary Hamiltonian. In fact, action–angle coordinates seem to be an unsuitable set of variables in Galactic Dynamics, since it is not clear if a global unperturbed Hamiltonian, like $H_0(\boldsymbol{I})$, actually exists for a given realistic stellar system. However, in the present discussion we assume that this is the case.

For $\epsilon = 0$ we have $H = H_0(\boldsymbol{I})$ and the motion is absolutely stable for any initial condition, since we have the complete set of N integrals of motion I_i (H_0 is cyclic in $\boldsymbol{\theta}$), which is equivalent to say that the Hamiltonian is completely integrable[1]. By *stable* we mean that the integrals are confined, for all times, to a very small neighborhood of their initial values (certainly, for $\epsilon = 0$, $\boldsymbol{I}$ is an exact integral and it does not change with time). The phase vector evolves linearly with time with a frequency vector $\boldsymbol{\omega}(\boldsymbol{I}) = \partial H_0/\partial \boldsymbol{I}$, where we assume that $det(\partial\omega_i/\partial I_j) \neq 0$ in order that $\boldsymbol{\omega}(\boldsymbol{I})$ be a one–to–one function.

In general the motion is quasiperiodic and proceeds densely over an N–dimensional torus. An N–torus (the Cartesian product of N circles) is completely specified by the initial conditions that fix the set of the N values of the integrals I_i or the frequencies ω_i. In other words, the coordinates that define a torus are the actions or the frequencies. A *resonance condition* for the frequencies is given by the equation $\boldsymbol{m} \cdot \boldsymbol{\omega}(\boldsymbol{I}) = \boldsymbol{m} \cdot \dot{\boldsymbol{\theta}} = 0$, for some integer vector $\boldsymbol{m}$, which determines the set of resonant values for the dynamical variables. Actually, we have an infinite set of vectors $\boldsymbol{I}$ or $\boldsymbol{\omega}$ that could fulfill the resonance condition for only one $\boldsymbol{m}$. Setting $N-1$ components of the actions (or frequencies), the remainder one is determined by the resonance condition. In this way we select one (fixed) resonant value $\boldsymbol{I}^r$ or $\boldsymbol{\omega}^r$. For any of these special values of the actions (or the frequencies), the orbit is not dense over the N–torus but, in general, it is on an $(N-1)$–torus. For the particular case of 2D–systems, the resonance condition implies that the orbit in configuration space closes on itself after a certain number of revolutions: we have a periodic orbit.

For instance, in a spherical system we can always reduce the spatial motion to a plane and then, fixing the energy and the angular momentum, we ensure that a given orbit proceeds in an unique 2–torus. The remaining initial conditions specify the position of the orbit on the torus. With the resonance condition we find the *resonant tori* (or the set of resonant values of the energy and the angular momentum) for the periodic orbits. These are essentially the main results for integrable systems (for a complete description see, for example, Rasband 1983, LL92 or Reichl 1992, while those readers interested in a more formal approach, the lectures by Giorgilli 1990 are recommended).

The nice picture given above for integrable systems is, unfortunately, unusual for real dynamical systems such as elliptical galaxies. A more realistic scenario could be the Hamiltonian (1)–(2) with $\epsilon \neq 0$. The presence of the perturbation (2) destroys, in general, the integrability of the Hamiltonian (due to the phase dependence of V), leading to a variation of the unperturbed integrals I_i. We can say that the stability of the motion breaks down when a large change in the actions takes place or, in Chirikov's language, when a 'gross' instability sets up.

To describe the motion of a star in the Hamiltonian (1)–(2), we apply a perturbation technique in order to obtain approximate solutions. As any perturbation theory requires a small parameter, we assume that $\epsilon \ll 1$. That is, we consider a near–integrable Hamiltonian system. The perturbation approach by the so–called *asymptotic series* means, roughly, that the variation of the unperturbed actions is computed via a power series in the perturbation parameter. Basically, we look for solutions to the perturbed motion by successive canonical transformations: $(\boldsymbol{I}, \boldsymbol{\theta}) \to (\boldsymbol{I}_1, \boldsymbol{\theta}_1) \to \cdots (\boldsymbol{I}_n, \boldsymbol{\theta}_n)$ chosen in such a way that at each step the perturbation term ϵV becomes independent of the phases at order $\epsilon^k, k = 1, \cdots, n$ and the 'actual' perturbation (i.e., the part of the perturbation that depends on the phases) becomes of order $\epsilon^{(k+1)}$. Therefore, in the first step, we look for a canonical transformation to kill the term of order ϵ in the perturbation. In the new variables $\boldsymbol{I}_1, \boldsymbol{\theta}_1$, the Hamiltonian (1)–(2) takes, in general, the following form

$$H'(\boldsymbol{I}_1, \boldsymbol{\theta}_1) = H_0'(\boldsymbol{I}_1) + \epsilon f(\boldsymbol{I}_1) + \epsilon^2 V'(\boldsymbol{I}_1, \boldsymbol{\theta}_1),$$

where, f is a certain function that depends on the canonical transformation and on the form of V; $\boldsymbol{I}_1, \boldsymbol{\theta}_1$ are corrections of order ϵ to the unperturbed values $\boldsymbol{I}, \boldsymbol{\theta}$, and V' has, by construction, zero average over the new phases. This is a standard procedure, particularly in Celestial Mechanics (further details about the asymptotic series technique may

[1] It is important to keep in mind that in general, the actions are known functions of the 'physical' integrals like energy, angular momentum, etc.

be found in, for example, CH79 or LL92[2]).

It is well known that the effect of a perturbation Fourier component (like in (2)) is stronger when the time variation of its phase, $\dot{\psi}_{\boldsymbol{m}} = \boldsymbol{m} \cdot \dot{\boldsymbol{\theta}}$, is slow. In the limit of constant phase we come to the resonance condition for the unperturbed frequencies. If we are far from a resonance (i.e., the initial conditions are chosen in such a way that $\boldsymbol{m} \cdot \boldsymbol{\omega}$ is far from zero for all $\boldsymbol{m}$), then it can be shown that the motion is stable. Briefly, we take advantage of the so–called *averaging method*: we neglect the oscillating part of the perturbation retaining only its average value. This principle (yet unproved) rests on the physical proposition that the oscillating part of the perturbation only causes small oscillations that do not affect the stability of the motion[3]. As an example, let us consider again the first canonical transformation which leads to the Hamiltonian H'. Averaging H' over the whole phase variables and due to the fact that, by construction, $\langle V' \rangle_{\boldsymbol{\theta}_1} = 0$ we arrive then at an average Hamiltonian where the perturbation (at first order) depends only on $\boldsymbol{I}_1$ and therefore, is integrable. The presence of the perturbation term only produces small changes of order ϵ in the unperturbed actions. Geometrically, the torus structure of the phase space for the unperturbed motion persists (locally) and the perturbation only produce a slight distortion of the unperturbed tori.

When we are near to a resonance condition, the asymptotic series technique does not work any longer due to the appearance of the so–called *small denominators* in the coefficients of the series (in the above given example, this happens to the correction of order ϵ in the unperturbed action, $\boldsymbol{I}_1$). These small denominators are the resonant values $\omega_{\boldsymbol{m}} = \boldsymbol{m} \cdot \boldsymbol{\omega}$ that may produce divergent series. It can be shown that the set $\omega_{\boldsymbol{m}}$ for all integer vectors is, in general, everywhere dense in phase space. Thus, to find initial conditions 'far from a resonance' is not an obvious task. In this connection, the so–called Diophantine condition for the frequency was introduced (see CH79 or Giorgilli 1990). However, throughout this paper, we assume that a resonance (or a set of resonances) is given a priori and the motion in its vicinity is the main subject of this work.

Let us consider the simple case of a single resonance. To be precise this means that the resonance condition is satisfied for one *resonant vector* $\boldsymbol{m}$ or, alternatively, that only one perturbing term is present in (2) and we chose initial conditions close to the resonance. We start discussing the geometric aspects of resonances in the action or frequency space (I–space and ω–space, respectively). In any of them, a torus is represented by a single point, since the 'position vector' is given by the N components of the action or the frequency. In the ω–space, the resonance condition $\boldsymbol{m} \cdot \boldsymbol{\omega} = 0$ has a very simple structure, just an $(N-1)$–dimensional plane where its normal is the resonant vector $\boldsymbol{m}$. In the I–space, $\boldsymbol{m} \cdot \boldsymbol{\omega}(\boldsymbol{I}) = 0$ leads to some other $(N-1)$–dimensional surface, whose local normal at the point $\boldsymbol{I} = \boldsymbol{I}^r$ is $(\partial [\boldsymbol{m} \cdot \boldsymbol{\omega}(\boldsymbol{I})] / \partial \boldsymbol{I})_{\boldsymbol{I}^r}$. From the conservation of the unperturbed energy we have the $(N-1)$–dimensional surface $H_0(\boldsymbol{I}) = E$ (in the I–space) and, since $\boldsymbol{\omega}(\boldsymbol{I})$ is a one–to–one function, we can also write $H_0(\boldsymbol{\omega}) = E$ (in the ω–space). Actually we should write $\mathcal{H}_0(\boldsymbol{\omega})$ since it will be a different function than H_0. In what follows we will consider only *convex* Hamiltonians which, for our purpose, means that the energy surfaces for the unperturbed Hamiltonian are everywhere convex. Then, the subspace defined by the intersection of both the resonant surface and the energy surface has dimension $N-2$. The latter subspace (in the I– or ω–space) is just a point for $N = 2$, a curve over the energy surface for $N = 3$, and so on.

By definition, the frequency vector $\boldsymbol{\omega}$ is normal to the energy surface in the I–space, since it is the I–gradient of H_0. The latter condition, together with the resonance equation, shows that the resonant vector lies in the tangent plane to the energy surface at $\boldsymbol{I} = \boldsymbol{I}^r$ (a simple sketch for $N = 3$ may help to visualize this particular geometry). Furthermore, a simple inspection of the equations of motion (for only one resonant perturbing term) shows that $\dot{\boldsymbol{I}}$ is parallel to the constant vector $\boldsymbol{m}$. This graphic picture of the dynamics allow us to conclude that the motion under a single resonant perturbation proceeds over the tangent plane to the energy surface at the point $\boldsymbol{I} = \boldsymbol{I}^r$ in the direction of the resonant vector. We may say then that as $\epsilon \to 0$, the resonant perturbation preserves the unperturbed energy.

The last example (a single resonance) also shows that the perturbation (2) only depends on a single phase: $\boldsymbol{m} \cdot \boldsymbol{\theta}$, the *resonant phase*. If we perform a canonical transformation in such a way that one of the N new phases, say ψ_1, is such that $\psi_1 = \boldsymbol{m} \cdot \boldsymbol{\theta}$, then the Hamiltonian is cyclic in the remainder $N-1$ phases $\psi_2, \cdots, \psi_N$ and $N-1$ new momenta, $p_2, \cdots, p_N$, are integrals of motion. This canonical transformation may be done through a linear transformation of the form: $\psi_i = \mu_{ik} \theta_k$; $I_i = I_i^r + p_k \mu_{ki}$, where μ_{ik} is some $N \times N$ matrix and $\mu_{1i} = m_i$, the resonant vector (the sum over repeated indexes is understood). Then, introducing this transformation into the Hamiltonian (1)–(2), expanding up to second order in p_i (which is assumed to be small) and averaging over the remainder non–resonant phases (i.e., those phase values for which $|\boldsymbol{m} \cdot \boldsymbol{\omega}|$ is large), we can reduce the Hamiltonian (1)–(2) to a one–dimensional one of the form (see CH79 for details, although we will come back to this point in Section V):

---

[2]See also the problem of convergence of the asymptotic series and the superconvergence of successive approximations.
[3]The stability of non–resonant motion is established by the KAM theory, which does not rely on averaging.

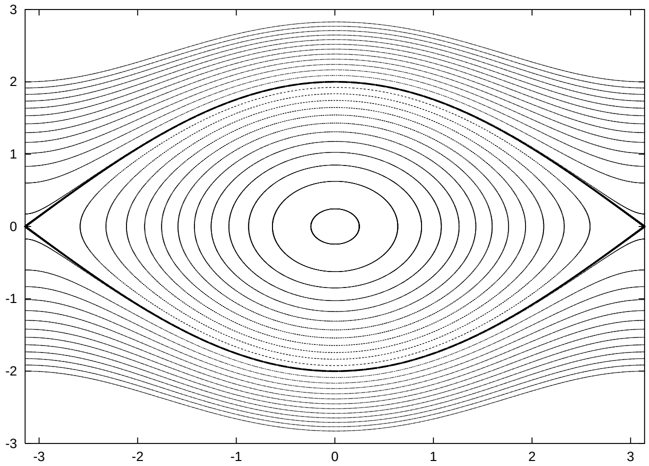


FIG. 1 Phase plane (p, ψ) of the pendulum $H = p^2/2 - \cos\psi$, where points at $\psi = -\pi, \pi$ should be identified. All curves are parametrized by the pendulum energy E; the solid line is the separatrix ($E = E_s = 1$).

$$H_r(p_1, \psi_1) \approx \frac{p_1^2}{2M} + \epsilon V_{\boldsymbol{m}}(\boldsymbol{I}^r)\cos\psi_1, \tag{3}$$

where H_r stands for the *resonant Hamiltonian* and M is the 'non–linear mass' given by: $1/M = m_i\,(\partial\omega_i/\partial I_k)_{\boldsymbol{I}^r}\,m_k$. To derive the resonant Hamiltonian certain restrictions to the matrix μ_{ik} are required (see Section V). The remaining $N-1$ momenta $p_k, k \geq 2$ are set equal to zero so that $\boldsymbol{I}^r$ is a point of the trajectory. Since H_r is independent of time, we have again the full set of N integrals of motion for the problem of a single resonant perturbation: $H_r, p_2, \cdots, p_N$. Consequently, the motion is stable.

The resonant Hamiltonian is nothing else than the Hamiltonian of a simple pendulum with mass M. We know the dynamics of this system from elementary mechanics (see Fig. 1). For small values of the energy, $H_r = E_r$, we have oscillations around the stable point and, for larger values of E_r, we have rotations. For the intermediate value $E_s = \epsilon V_{\boldsymbol{m}}$, we have *the separatrix*, a smooth curve which separates both kinds of motion. It is not difficult to show that the equation of the separatrix for the latter Hamiltonian is (see CH79, Rasband 1990):

$$p_1^s = \pm p_r\cos(\psi_1^s/2), \qquad \psi_1^s(t) = 4\arctan[e^{\omega_0(t-t_0)}] - \pi, \tag{4}$$

where $p_r = 2(\epsilon|MV_{\boldsymbol{m}}|)^{1/2}$, $\omega_0^2 \equiv \omega_0^2(\epsilon) = -\epsilon V_{\boldsymbol{m}}/M > 0$, and t_0 is the origin of t taken at $\psi_1 = 0$ (the point of stable equilibrium). The sign in p_1^s defines the upper or the lower branch of the separatrix. Equation (4) clearly shows that the motion along this trajectory has an asymptotic nature. Indeed, the separatrix joints smoothly for an infinite time the unstable points ($\psi_1 = \pm\pi$). From the first of (4) we see that in the neighborhood of a resonance, p_1 oscillates about the stable point with an amplitude $p_r \sim \sqrt{\epsilon}$. This value is just the maximum displacement of the actual value from the stable (resonant) point. Using the relationship between the old and new momenta (see Section V), we may find the amplitude of oscillation for the unperturbed actions

$$(\Delta\boldsymbol{I})^r \equiv (\boldsymbol{I} - \boldsymbol{I}^r)_{\max} = p_r\boldsymbol{m}.$$

We call *action width*, $(\Delta I)_r$, the magnitude given by $|(\Delta\boldsymbol{I})^r|$, where $|\cdot|$ is the usual Euclidean norm. As mentioned above, the change in the unperturbed action proceeds in the direction of the resonant vector and with an amplitude of the order $\sqrt{\epsilon}$. Linearizing $\boldsymbol{\omega}(\boldsymbol{I})$ about $\boldsymbol{I}^r$, we may write for the frequency amplitude,

$$(\Delta\boldsymbol{\omega})^r = p_r\left(\left(\boldsymbol{m}\cdot\frac{\partial}{\partial\boldsymbol{I}}\right)\boldsymbol{\omega}\right)_{\boldsymbol{I}^r}.$$

These two vectors $(\Delta\boldsymbol{I})^r$ and $(\Delta\boldsymbol{\omega})^r$ measure, in the $I-$ and ω–spaces, the maximum displacement of the system from the resonant point, respectively.

From the beginning we have assumed that we are dealing with a non–linear system. By *non–linear resonance*, we essentially mean that the oscillation frequency depends on the action (strictly speaking, $det(\partial\omega_i/\partial I_j) \neq 0$), while this is not true for linear resonances (where $\boldsymbol{\omega}$ is a constant vector). It is well known that when any trajectory enters into a domain of a linear resonance, the amplitude of oscillation grows indefinitely, leading to an instability. But within the domain of a non–linear resonance, the effect of the perturbation is to get the system out of resonance, leading to a kind of stabilization. This mechanism is known as *non–linear stabilization* (see CH79).

Let us look then at the scenario in the ω–space. The vector $(\Delta\boldsymbol{\omega})^r$ is, in general, not parallel to the resonant vector $\boldsymbol{m}$. Since the latter is normal to the resonance plane, it seems natural to define the *frequency width* as the component of $(\Delta\boldsymbol{\omega})^r$ normal to that plane,

$$(\Delta\omega)_r = (\Delta\boldsymbol{\omega})^r \cdot \frac{\boldsymbol{m}}{|\boldsymbol{m}|} = \frac{p_r}{M|\boldsymbol{m}|} \propto \sqrt{\frac{\epsilon V_{\boldsymbol{m}}}{M}}\,.$$

Thus, we see that the main parameter defining the non–linear condition is M. If $1/M = 0$, then the system does not get out of resonance and instability arises. The system behaves as a linear one, where the oscillation frequencies do not depend on the actions. In other words, the pendulum model for a non–linear resonance works only if the bob has a finite mass. Note that the condition $1/M \equiv m_i(\partial\omega_i/\partial I_j)m_j \neq 0$ is more restrictive than $det(\partial\omega_i/\partial I_j) \neq 0$.

From the above discussion we see that the effect of a resonant perturbation is larger than that for a non–resonant one. The latter produces a variation of order ϵ in the unperturbed integrals while, for the former, this variation is of the order of $\sqrt{\epsilon}$. From the geometrical point of view, we say that a single resonant perturbation produces (locally) a change in the topology of the phase space. Indeed, although the torus structure is preserved, the motion in the vicinity of a resonant torus proceeds, in general, over a chain of tori that twist around the latter. In the I– or ω–space, the resonance surface becomes actually a *resonance layer*, whose widths are $(\Delta I)_r$ and $(\Delta\omega)_r$, respectively (see the example given in Section VII). Clearly, this last picture is quite different from that for non–resonant perturbations (for further details see, for example, Rasband 1983, LL92).

III. THE TRANSITION TO CHAOS

The example discussed in the last section is quite restricted, since only one resonant perturbing term is retained. In order to describe a more realistic situation, it seems to be reasonable to include some other harmonics of the perturbation (2) (i.e., several vectors $\boldsymbol{m}$). Depending on the initial conditions, these additional terms may be killed by the application of the averaging method, but in this section we will consider the opposite case, when the system is moving within a region of the phase space where two or more resonances are present. Any attempt to describe the *resonance interaction* in a rigorous way is out of the scope of the present discussion. Nevertheless, Chirikov (CH79) developed a qualitative criterion called *overlap of non–linear resonances*, that may help us to understand the motion under the influence of several resonances. In order to give just a sketch of it, let us consider the case of only two resonances. Each resonance will determine its own domain in the phase space (as it was shown in the last section), but the motion in the vicinity of one resonance will be affected by the presence of the other. If the resonances are situated 'far enough' from each other, we may expect the motion to be confined to the neighborhood of one resonance or the other, depending on the initial conditions. The picture of a slight distortion of the pendulum model is then a fair approximation to the actual motion, and to give a qualitative criterion, we can neglect the effect of the perturbing resonance. Thus, each resonance has its own pendulum model, with small oscillations, separatrix and rotation about the resonant value $\boldsymbol{I}_1^r$ and $\boldsymbol{I}_2^r$ (see Fig. 2). On the other hand, if the two resonances are 'close enough' to each other, then it is reasonable to expect the motion not to be confined within one domain, and the trajectory may jump from one resonance domain to the other; i.e., the action could range from some neighborhood of $\boldsymbol{I}_1^r$ to some other neighborhood of $\boldsymbol{I}_2^r$. As Chirikov pointed out, this kind of motion seems to have nothing to do with any instability since $\boldsymbol{I}_1^r \approx \boldsymbol{I}_2^r$. But, as shown in many numerical experiments (see CH79 and LL92), the motion becomes irregular as if the system were dominated by a stochastic force (this is clearly seen in the surfaces of section of a given near–integrable potential; the classical example is the Hénon and Heiles model, 1964). Nevertheless, the equations of motion for the Hamiltonian (1)–(2) do not include stochastic forces like, for example, the Langevin equation (see Saslaw 1985, Ch. 3). This is the reason why the motion in question was called *stochastic instability*.

From these intuitive considerations we may infer that a plausible condition for the stochastic instability is that the separation between the resonances (in the I– or ω–space), is of the order of the resonance width; i.e., an overlap of resonances. Let us put it in another way, the overlap of resonances takes place when the unperturbed separatrix of one resonance touches the other. Let us consider, for instance, the ω–space. Let $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$ be two different resonant vectors such that $\boldsymbol{m}_1 \cdot \boldsymbol{\omega} \approx 0$ and $\boldsymbol{m}_2 \cdot \boldsymbol{\omega} \approx 0$; i.e., the initial conditions are chosen in such a way that $\boldsymbol{\omega}$ is close to any of the two fixed vectors $\boldsymbol{\omega}_1^r$ and $\boldsymbol{\omega}_2^r$, each of them belonging to its own resonance plane. Let $|\boldsymbol{\omega}_1^r - \boldsymbol{\omega}_2^r| = \Delta$. The resonances have a frequency width $(\Delta\omega)_{r1}$ and $(\Delta\omega)_{r2}$ respectively. Thus, the condition for the overlap of resonances

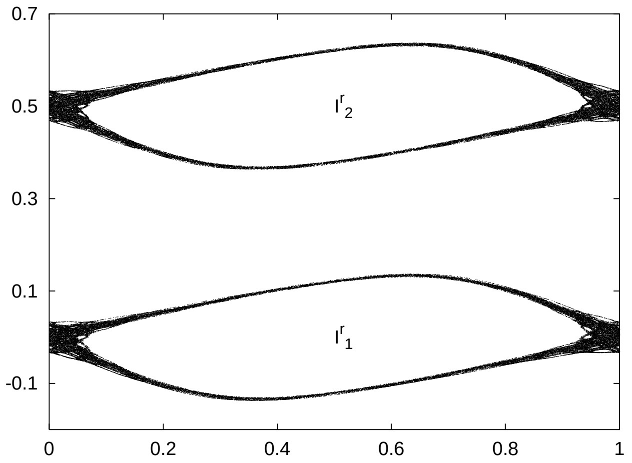


FIG. 2 Two resonances with their stochastic layers around the unperturbed pendulum separatrix. The plot corresponds to two orbits of the reduced Standard Map ($x \in (0,1)$, $y \in (-1,1)$), $\overline{y} = y + K' \sin(2\pi x)$, $\overline{x} = x + \overline{y}$, with $K = 2\pi K' = 0.659$. This map, for the given value of K, is an alternative representation of the motion of a pendulum under the effect of a low-to-moderate perturbation (see CH79).

may be formulated as

$$\gamma\left((\Delta\omega)_{r1}, (\Delta\omega)_{r2}\right) \approx \Delta,$$

where γ is the 'path' over the energy surface between the resonant points (see below). As $(\Delta\omega)_{ri} \sim \sqrt{\epsilon}$, then the latter condition gives an estimate of the so–called *stability border*. This means that if ϵ_{c} is the value of the perturbation parameter that satisfies the overlap condition, then for $\epsilon < \epsilon_{\mathrm{c}}$ we may expect the system to be confined within the domain of one resonance. That is, the motion is stable, as described in Section II. On the other hand, if $\epsilon \gtrsim \epsilon_{\mathrm{c}}$, then the stochastic instability arises; the resonances are connected and the motion proceeds over both domains. Let us take $N = 2$. The resonance condition leads, in the ω–space, to lines with different slopes (given by $\boldsymbol{m}_1$ and $\boldsymbol{m}_2$) passing through the origin. The energy surfaces are certain convex curves. Then, fixing the energy, the system is confined to that curve. The resonant (fixed) values $\boldsymbol{\omega}_1^r$ and $\boldsymbol{\omega}_2^r$ are the intersection points of the latter curve with the resonant lines. By assumption, the initial conditions are such that $\boldsymbol{\omega}_1^r \approx \boldsymbol{\omega}_2^r \approx \boldsymbol{\omega}^r$. Then, the separation between the two resonances is

$$\Delta \approx |\boldsymbol{\omega}^r| \frac{|\boldsymbol{m}_1 \times \boldsymbol{m}_2|}{|\boldsymbol{m}_1||\boldsymbol{m}_2|},$$

which is a function of the energy through $|\boldsymbol{\omega}^r|$. On the other hand, the frequency width may be put in the form

$$(\Delta\omega)_{ri} = \delta(\boldsymbol{m}_i)\sqrt{\epsilon}, \qquad \delta(\boldsymbol{m}) = \frac{2}{|\boldsymbol{m}|}\sqrt{|V_{\boldsymbol{m}}/M|}.$$

Then the stability border is given by:

$$(\delta(\boldsymbol{m}_1) + \delta(\boldsymbol{m}_2))\sqrt{\epsilon_{\mathrm{c}}} \approx \Delta.$$

Therefore, for $\epsilon < \epsilon_{\mathrm{c}}$ we can assure the stability of the motion since, for very small values of the perturbation, the size of each resonance is rather small and their domains do not overlap[4]. But as the perturbation increases, the domains become larger and they may overlap, leading to a connected region of stochastic motion. Clearly, the situation is much more complicated if we consider more than two resonances, but the qualitative picture is similar. In such a case, since the overlap includes more than two resonances, the connected region for the motion becomes larger and the unperturbed integrals may have a large variation, i.e., we have a gross instability.

---

[4]This statement is only true for the zeroth order approximation to the actual motion. If we consider the perturbed motion, higher order resonances appear. Therefore, the overlap criterion for the first order resonances is just a necessary but not sufficient condition for the stability of the motion.

From the beginning we have not considered the case of *intersection of resonance surfaces*, for which the analysis is rather complicated (see Section VII) but we may infer that a stochastic domain appears in the neighborhood of the intersection of both surfaces, due to the overlapping of resonance domains.

As we shall see in the next section, for $\epsilon < \epsilon_c$ the main effect of the interaction of two resonances is to produce a qualitative change in the separatrix of the perturbed resonance. This smooth curve becomes a layer, the so–called *stochastic layer*, since a stochastic behavior appears in the vicinity of the separatrix (see Fig. 2). All the invariant curves in a neighborhood of the separatrix disappear and, instead, a stochastic motion proceeds *across* the layer (i.e. in the direction of p_1, following the notation of equation (3)) of a given finite width. This kind of motion is quite different from that of an isolated resonance since, in the latter case, the motion follows a smooth curve in the (p_1, ψ_1) plane, as Fig. 1 shows (see next section). The stochastic layer is located at the edge of the resonant layer, then oscillations and rotations for the pendulum model are actually separated by a region of irregular, chaotic motion. Though the overlap criterion misses this fact as well as many other aspects, the results given by this simple and intuitive approach are in good agreement with numerical simulations (CH79). In fact, Chirikov showed that, in general, the overlap criterion provides a stability border that is of the same order as that obtained using a rigorous mathematical approach by the so–called KAM theory (see for example, CH79, LL92 or Reichl 1992). As we shall see later, the existence of a stochastic layer and the intersection of resonance surfaces when $N > 2$ is the key point in the discussion of Arnold diffusion.

The example considered above to derive ϵ_c is for a two–dimensional autonomous Hamiltonian system. For these systems, the resonance lines do not intersect (except at the origin). Therefore, the unique way to obtain a connected domain of stochastic motion is through an overlap of resonances. In other words, the motion becomes unstable, stochastic, for certain values of the perturbation parameter. This is just a consequence of the topology of the phase space for $N = 2$. Indeed, we may say that the 2–dimensional tori divide the 3–dimensional energy surface. This fact implies that any transition from one torus to another is only possible through all the intermediate tori between them. If $\epsilon < \epsilon_c$, most of the torus structure survives, leading to a picture similar to that for an integrable Hamiltonian. The tori act as a barrier for the stochastic motion and confine it to the stochastic layers. If we project the motion in question onto the I– or ω–space we have, as we mentioned above, an energy surface which is a curve where, every point on it represents a different ratio for the components of the frequency vector. A rational ratio corresponds to a resonance while, an irrational one, corresponds to a non–resonant orbit. If the system is confined within a resonance, it will move then along the energy curve and away from the resonant value by a distance which is about the amplitude of oscillation around this point (p_r). Although an instability is always present close to the separatrix (the stochastic layer), this region of stochastic motion is rather small, its width being exponentially small (see CH79 and next section). Therefore we conclude that, for $N = 2$ and $\epsilon < \epsilon_c$, the stability of the motion, in a broad sense, is preserved.

The story is quite different for $N > 2$. The results of the KAM theory do not guarantee the confinement of the system to a neighborhood of the resonant value. The tori no longer divide the energy surface; the resonant surfaces (in the I– or ω–space) do intersect, leading to a united network over the energy surface. Thus, an instability may occur, even for very small values of the perturbation. This peculiar instability, discovered by Arnold (1964) in a purely mathematical paper, seems to be an universal one, since it is always present despite the smallness of the perturbation. In order to give a picture of this geometry, let us consider the unperturbed Hamiltonian $H_0 = |\boldsymbol{I}|^2/2$, where $\boldsymbol{I}$ is a 3–dimensional vector, and the perturbation given by (2) with $\epsilon < \epsilon_c$. In this case $\boldsymbol{\omega}(\boldsymbol{I}) = \boldsymbol{I}$ and the I– and ω–spaces are identical. The energy surfaces are concentric spheres. Since the resonant surfaces are planes passing through the origin, we see that the intersection of any resonant plane with a given energy surface is a great circle. It is clear that these great circles intersect over the whole surface of the sphere leading to the so–called *Arnold Web* (see LL92, Fig. 6.3 for a sketch and Fig. 6 (left) in Section VII for a different example). The existence of this web does not depend on the strength of the perturbation. Let us fix the energy and consider initial conditions very close to a given resonance, where the motion is confined. Just to distinguish the actual resonance from the rest of the perturbing (non–resonant) terms, let us call it *guiding resonance*. Certainly, for other initial conditions, one of the perturbing terms may become the guiding resonance, while the former resonance will then play the same role as the rest of the perturbing terms. The guiding resonance surface leads, on the energy surface, to a certain great circle (in fact, to a resonant layer–see Fig. 6 (right)). The conservation of energy confines the motion to the sphere and, the resonance condition, restricts it to the resonant layer of the guiding resonance. Without perturbation (i.e, keeping only H_0 and the resonant term) the motion proceeds then over the tangent plane to the energy surface (at the resonant value) in the direction of the guiding resonant vector, the latter being normal to the resonance plane. Therefore, if the initial conditions are chosen near the separatrix of the resonance, then the presence of perturbing terms (besides the resonant one) gives rise to a motion *across* the resonant layer of the guiding resonance ('transversal to the great circle') and modifies the edges of the latter leading to the appearance of the stochastic layer. It follows then that close to the resonant value, the great circle looks like a spherical layer where the motion within its edges is stochastic. However, the energy and resonance constraints also allow for a motion in the remainder direction ('along the great circle'). We expect that the perturbation may also drive the stochastic motion *along* the stochastic layer of the guiding resonance.

Following the notation of equation (3), by motion along the stochastic layer we essentially mean that the components of the action, p_k, $k \neq 1$, will change with time, since now they are not exact integrals due to the dependence of the perturbation on several phases (Section V). During some time the stochastic motion may proceed along the stochastic layer of the guiding resonance. When the motion along this layer reaches a point of intersection with some other great circle (other resonant surface), a stochastic domain around the latter also appears, due to the overlap of both resonances. Although it has not been mathematically proved, it might be possible that the motion proceed now *along* this second layer: a new guiding resonance. Since the set of resonances do intersect over the whole energy surface, we conclude then that the Arnold web is actually a network of layers where the motion within it is stochastic (see however the example given in Section VII). On the other hand, the motion outside the web is regular and stable. If the initial conditions are chosen within any stochastic layer, then the stochastic motion might spread over the whole web through the intersecting points. From these qualitative considerations we then infer that it might be possible that all the regions on the energy surface where the motion is chaotic are connected. Therefore the properties of the stochastic component (as Lyapunov exponents, entropy, etc.) could be the same over the whole phase space accessible to the system (see Section VII).

As we have already said, for $N = 2$ and if the perturbation is small enough, the variation of the unperturbed integrals is rather small, just confined to the stochastic layer. Motion 'along' the layer does not exist due to the dimensionality (two) of the I– or ω–space. But, for $N > 2$, we have at least one more degree of freedom where the motion may proceed. In the latter case, large variations of the unperturbed integrals might occur. As the stochastic motion admits a diffusion–like description this instability was called Arnold diffusion (though Arnold had never used this word in his celebrated paper). These questions will be addressed in Section V.

IV. THE STOCHASTIC LAYER

In Section III we have described what happens to the separatrix of a non–linear resonance under the interaction with another resonance in a qualitative way. Now we derive a quantitative estimate of the effect of such an interaction.

Notice first that the motion in the vicinity of the separatrix is extremely unstable. Indeed, let us consider that only one term in (2) has $V_{\boldsymbol{m}} \neq 0$, so that the full Hamiltonian (1) reduces to the resonant Hamiltonian H_r given in (3), and take initial conditions such that $H_r \lesssim E_s$, i.e. the system oscillates near the separatrix of the resonance. When switching on to some other term $V_{\boldsymbol{m}'}$, slight (periodic) variations of H_r may cause the system to change drastically its motion, from oscillations to rotations and so on.

Next, let us recall that the motion on the separatrix has an infinite period. Measuring the distance from the separatrix by the relative energy $w = (H_r - E_s)/E_s$, it can be shown that the frequency, in a neighborhood of the separatrix, is (see CH79, LL92):

$$\omega(w) \approx \frac{\pi\omega_0}{\ln(32/|w|)} \rightarrow 0 \quad \text{as} \ w \rightarrow 0, \tag{5}$$

with ω_0 the small oscillation frequency given in (4).

Following Chirikov's approach, let us see now how the motion looks in the vicinity of the separatrix. The simplest way to do it is by means of the one–dimensional pendulum Hamiltonian acted upon by a periodic, time–dependent perturbation,

$$H(p,\psi,\tau) = H_r(p,\psi) + \mu V(\psi,\tau), \qquad \mu \ll 1, \tag{6}$$

where:

$$H_r = \frac{p^2}{2} - \omega_0^2 \cos\psi, \qquad \omega_0^2 = \epsilon|V_{\boldsymbol{m}}|, \tag{7}$$

is the Hamiltonian (3) with $M = 1$, and

$$\mu V = \mu\omega_0^2 \cos\psi \cos\tau = \frac{\mu\omega_0^2}{2}[\cos(\psi - \tau) + \cos(\psi + \tau)]. \tag{8}$$

The perturbation V depends on the perturbing phase $\tau(t) = \Omega(t - t_0) + \tau_0$, Ω being the perturbing frequency and $\tau_0 = \tau(t_0)$. As t_0 is the value of t when the pendulum crosses the equilibrium point, then τ_0 is the value of the phase of the perturbation when $\psi = 0$.

Let us now compute the change in the unperturbed energy of the pendulum, $H_r \approx E_s = \omega_0^2$, over a half period of oscillation or a period of rotation: $T(w) = \pi/\omega(w)$. To this end, we first calculate the time variation of H_r, to later integrate it over a whole period T. From the equations of motion for the Hamiltonian (6)–(8), one readily finds:

$$\dot{H}_r = -\mu p(t) \frac{\partial V}{\partial \psi} = \frac{\mu \omega_0^2}{2} p(t) \left[\sin\left(\psi(t) - \tau\right) + \sin\left(\psi(t) + \tau\right) \right], \tag{9}$$

which has to be integrated over $-T/2 \le t - t_0 \le T/2$. Let us notice that, though there are explicit formulae for the unperturbed values of $p(t)$ and $\psi(t)$, they involve elliptic functions or, eventually, Fourier series (see CH79), so that such expressions are not convenient for our purpose. However, since the motion in the vicinity of the separatrix is nearly the same as that on the separatrix itself (except for that $T \to \infty$), we can write $p(t) \approx p^s(t)$, $\psi(t) \approx \psi^s(t)$ and $-\infty < t < \infty$; p^s and ψ^s given by (4). Therefore, taking one branch of the separatrix (the plus sign in the first of (4)), setting $t_0 = 0$ and assuming that the *slow variable* is $(\psi - \tau)$, we can average over the *fast variable* $(\psi + \tau)$ to obtain:

$$\Delta H_r \approx \mu \omega_0^3 \int_{-\infty}^{\infty} dt \sin(\psi^s - \tau) \cos(\psi^s/2). \tag{10}$$

Using trigonometric relations, (10) can be recast as

$$\Delta H_r \approx \frac{\mu \omega_0^3}{2} \int_{-\infty}^{\infty} dt \left[\sin\left(\frac{3}{2}\psi^s - \tau\right) + \sin\left(\frac{1}{2}\psi^s - \tau\right) \right].$$

Further, replacing $\tau(t) = \Omega t + \tau_0$, expanding $\sin(n\psi^s - \tau)$ in terms of $(n\psi^s - \Omega t)$ and τ_0 ($n = 3/2, 1/2$) and taking into account the fact that $\psi^s(t)$ is an odd function and that the odd part does not contribute to the integral, we obtain:

$$\Delta H_r \approx -\frac{1}{2} \mu \omega_0^2 \sin \tau_0 \left[A_1\left(\frac{\Omega}{\omega_0}\right) + A_3\left(\frac{\Omega}{\omega_0}\right) \right], \tag{11}$$

where:

$$A_m(\lambda) = \int_{-\infty}^{\infty} d\hat{t} \cos\left(\frac{m}{2}\psi^s(\hat{t}) - \lambda \hat{t}\right), \tag{12}$$

with $\lambda = \Omega/\omega_0$, $\hat{t} = \omega_0 t$ and m an integer.

The integral in (12) is known as the *Melnikov–Arnold Integral* (MAI, hereafter) and its evaluation may be found in the Appendix of CH79 (though maybe the calculations done by Ferraz–Mello 1996 are more suitable for the non–expert reader). Actually, $A_m(\lambda)$ is defined as the 'mean value' of the improper integral (12), which in fact does not converge. A detailed computation of the MAI in the complex plane, shows that the integral, as a function of $\hat{t}$, oscillates with an amplitude of the order of $1/\lambda$ for $\hat{t} \gg 1$ and, therefore, the limit does not exist. But these periodic oscillations play no role in the problem of the stability of the motion so we can neglect them and retain only the aperiodic part. Besides, $\omega_0 \sim \sqrt{\epsilon}$ is small and $m \neq 0$ and then we can use the asymptotic value of the MAI for large positive λ:

$$A_m(\lambda) \approx \frac{4\pi(2\lambda)^{m-1}}{(m-1)!} e^{-\pi\lambda/2}, \qquad \lambda \gg m, \tag{13}$$

where the factorial should be replaced by the Gamma function, $\Gamma(m)$, for non–integer m.

Moreover, for negative λ we have

$$A_m(\lambda) = (-1)^m A_m(|\lambda|) e^{-\pi|\lambda|}, \qquad \lambda < 0, \tag{14}$$

which shows that, for large $|\lambda|$, $A_m(-|\lambda|) \ll A_m(|\lambda|)$. It is not difficult to show that the largest contribution to the MAI comes from $m\dot{\psi}^s\lambda > 0$. Also a recurrence relationship for different m values can be derived, namely

$$A_{m+1} = \frac{2\lambda}{m} A_m - A_{m-1} \approx \frac{2\lambda}{m} A_m, \tag{15}$$

where the approximation holds for large λ.

Therefore, for the upper branch of the separatrix ($p^s > 0$) and using (11), (13) and (15), we obtain that the variation of the unperturbed energy is given by

$$\Delta H_r \approx -\frac{\mu\omega_0\Omega}{2} A_2 \left(\frac{\Omega}{\omega_0}\right) \sin\tau_0 \approx -4\pi\mu\Omega^2 e^{-\frac{\pi\Omega}{2\omega_0}} \sin\tau_0 . \tag{16}$$

For the lower branch of the separatrix ($p^s < 0$), since $\dot{\psi}^s < 0$, it is enough to take $\lambda = \Omega/\omega_0$ with opposite sign in the first equality in (16), so (14) shows that its contribution to ΔH_r is negligible. Thus, actually (16) is the total change of H_r for a complete period of oscillation, i.e. over $2T$. But, since the perturbation is symmetric (it depends on $(\psi \pm \tau)$ both with the same amplitude, see (8)), then the upper and lower branches of the separatrix contribute in the same way to the MAI in each half period of oscillation. In other words, for $p^s > 0$ the fast angle is $(\psi + \tau)$ while, $(\psi - \tau)$ is the one for $p^s < 0$. Following the above discussion, we readily conclude that (16) is the variation of H_r in each period of motion. This variation depends on τ_0, the value of the perturbing phase when $\psi = 0$. After a period of motion T, τ changes to $\bar{\tau} \equiv \tau(t+T) = \Omega t + \Omega\pi/\omega + \tau_0 \equiv \Omega t + \bar{\tau}_0$, where, $\bar{\tau}_0 = \tau_0 + \lambda \ln(32/|\bar{w}|)$; λ being, as before, Ω/ω_0 and $\bar{w} = (\bar{H}_r - \omega_0^2)/\omega_0^2$ the relative energy of the pendulum after crossing the surface $\psi = 0$. Thus, in the variables (w, τ_0), the canonical mapping

$$\bar{w} = w + W \sin\tau_0, \qquad \bar{\tau}_0 = \tau_0 + \lambda \ln\frac{32}{|\bar{w}|} \mod(2\pi), \tag{17}$$

with

$$W = -4\pi\mu\lambda^2 e^{-\pi\lambda/2}, \qquad \lambda = \frac{\Omega}{\omega_0} \sim \frac{1}{\sqrt{\epsilon}} \gg 1,$$

describes the motion in the vicinity of the separatrix (details concerning this mapping can be found in Shevchenko 1998).

The mapping (17) is known as the separatrix mapping or *whisker mapping* (WM) in the terminology introduced by Arnold. He called whiskers the different branches of the separatrix, and *whiskered torus* the unstable points. Actually, the unstable (stable) point may be considered as a separate orbit, since for initial conditions $p_0 = 0, \psi_0 = \pm\pi$ ($p_0 = 0, \psi_0 = 0$), the system remains there for an infinite time. In the pendulum model, we have two whiskers (technically, the stable and unstable manifolds), arriving to and departing from the whiskered torus. Therefore, the whiskers are asymptotic trajectories that, for $t \to \pm\infty$, approach towards the unstable point (the whiskered torus). However, this picture is better seen in higher dimensions.

The WM describes the whiskers under a periodic perturbation. In absence of perturbation ($V = 0$) the first of (17) reduces to $\bar{w} = w$, giving a fixed point in the mapping. Depending on the value of w ($|w| \ll 1$), the system will rotate, oscillate or move along the separatrix. Therefore, for $V = 0$, the stochastic layer does not exist. The energy w is an integral and the motion on phase space, (p, ψ), proceeds along a smooth curve of constant energy (see Fig. 1). But for $V \neq 0$, w changes with time. The arriving whisker ($w \neq 0, \bar{w} = 0$) and the departing whisker ($w = 0, \bar{w} \neq 0$) no longer coincide. Indeed, while for $V = 0$ and $\bar{w} = w = 0$, the whiskers coincide, they split under a perturbation, the scale of this splitting being of the order of $|2W|$. The resulting motion in phase space becomes stochastic in a neighborhood of the separatrix, giving rise to the stochastic layer (see Fig. 2). The stochastic behavior of $p(t)$ is what we have called, in the previous section, the motion across the layer. That motion, however, is better described by $w(t)$.

Fig. 3 (left) displays an orbit in the WM using the variable $\hat{s} = w/W$ instead of w, for $\lambda = 8$ and $\mu \approx 3.5 \times 10^{-8}$. This figure shows that the variation of $\hat{s}$ is bounded, the width of the stochastic layer seems to be of the order of $|\hat{s}|_{\max} \sim 10$. We can also distinguish two regions; a central one, very close to the separatrix, that looks like ergodic[5]: $|\hat{s}| \lesssim 2$, and a external one: $2 \lesssim |\hat{s}| \lesssim 10$, where the phase space is shared between stochastic and regular motion. It is

[5]Roughly speaking, by ergodic we mean that the motion fills densely and uniformly some domain of the phase space.

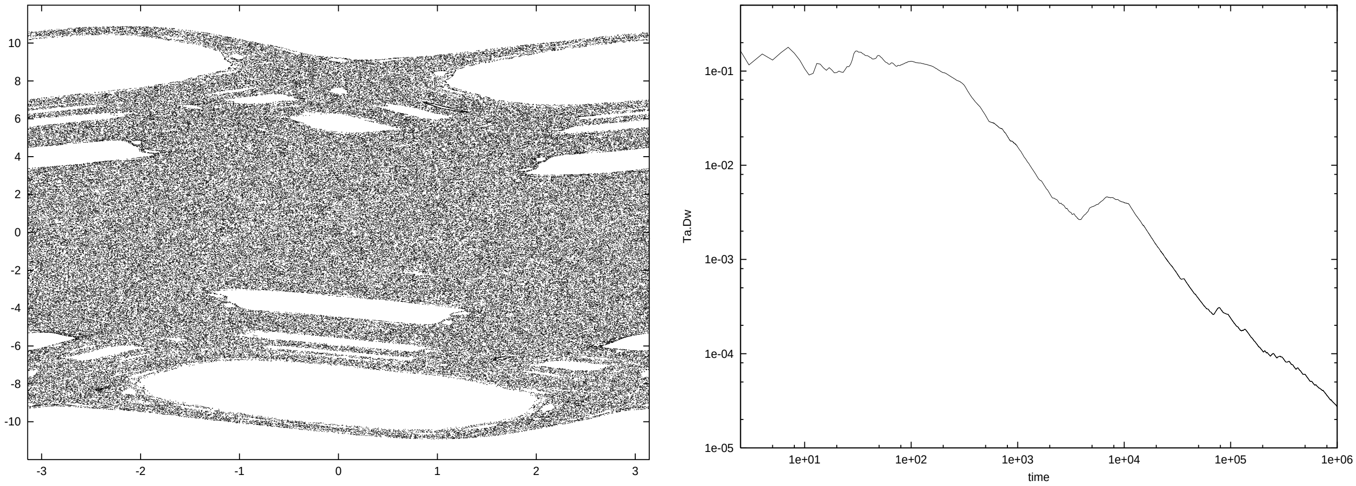


FIG. 3 Left, a sketch of the stochastic layer for the WM (17) in the $(\tau_0, \hat{s})$ plane, for $\lambda = 8$ and $\mu \approx 3.5 \times 10^{-8}$ $(W = 10^{-10})$, and $t = 4 \times 10^6$ iterations. Only 5% of the points were plotted. The separatrix equation in these variables is $\hat{s} = 0$, while $\hat{s} < 0$ and $\hat{s} > 0$ corresponds to oscillations and rotations, respectively. Right, time evolution of the diffusion coefficient (in units of W^2) for the same orbit (see text).

important to remark that the stability domains are due to resonances between the resonant phase (ψ) and the phase of the perturbation (τ) in a neighborhood of the separatrix. This type of resonances are, in some sense, different from that discussed above. Indeed, the *first level resonances*, are those between the original oscillations of the system and the perturbation, while the *second level resonances*, are those between the oscillations of the resonant phase in a vicinity of the separatrix and the same perturbation (see CH79). These second level resonances can also be seen (in a different space) in Fig. 2.

The numerical results obtained for this example have a theoretical support. Chirikov (CH79) showed that the width of the stochastic layer, w_s, is somewhere between $W\lambda \lesssim w_s \lesssim W(\lambda + 4)$, and the central region of *fast diffusion* is of the order of $|w| \lesssim w_s/4$.

Let us state clearly the proper meaning of diffusion in this example. After one iteration of the map, the relative energy changes in $\Delta w = W \sin \tau_0$ and after t iterations of the map, the total change in w is:

$$\Delta' w(t) = W \sum_{t'=0}^{t} \sin \tau_0(t').$$

If T_a denotes the mean period of motion within the stochastic layer (see next section), then tT_a is the physical time. The magnitude $\Delta' w(t)$ depends on the values of τ_0 within the interval $0 \leq t' \leq t$. From a statistical point of view, Fig. 3 (left) shows that while the system is confined within the central part of the stochastic layer, it seems plausible to assume that the values of $\tau_0(t')$ are nearly random. Furthermore, for $|w| \lesssim w_s/4$, we suppose that the behavior of τ_0 is ergodic, i.e., time–average and space–average coincide. Clearly this is not true for the external part of the layer. Then, while the system is moving very close to the separatrix, we may approximate the distribution ρ of the phase values by $\rho(\tau_0)d\tau_0 \approx d\tau_0/2\pi$. This is the so-called *limiting stochasticity approximation.* Then, for $|w| \lesssim w_s/4$ we have:

$$\overline{\Delta' w(t)} = \langle \Delta' w \rangle \approx 0, \qquad \overline{[\Delta' w(t)]^2} = \langle [\Delta' w]^2 \rangle \approx \frac{W^2}{2} t.$$

where the bar indicates time–average and $\langle \cdot \rangle$ denotes space–average. Thus, we define the diffusion rate as:

$$D_w \equiv \frac{\overline{[\Delta' w(t)]^2}}{tT_a},$$

whenever $\overline{\Delta' w(t)} = 0$; otherwise the variance of $\Delta' w(t)$ should be computed. Therefore, we arrive at

$$D_w \approx W^2/2T_a, \qquad |w| \lesssim w_s/4$$

for the value of the diffusion rate in a neighborhood of the separatrix[6].

As time increases, the system leaves the central part of the stochastic layer, where the successive values of τ_0 seem not to be independent of each other: the presence of stability islands prevents the free diffusion. Moreover, for very large values of t, the system reaches the edge of the stochastic layer where $[\Delta' w(t)]^2$ no longer changes. That is, since w is bounded, $[\Delta' w(t)]^2$ will reach a maximum value when $w \approx w_s$. However, we can still describe the motion as diffusion–like, but we assume that the *slow diffusion* in w is due to the presence of correlations between different values of $\tau_0(t')$. Thus, for the external part of the layer $\overline{[\Delta' w(t)]^2}$ does not grow linearly with time but in a slower way and reaches an asymptotic value when $t \to \infty$. Then we conclude that D_w has a finite value for short times, as the system moves within the central part of the stochastic layer (fast diffusion); takes smaller values for moderate times, while the system moves within the outer part of the layer (slow diffusion) and goes to zero for very large times, after the system reaches the border of the layer (the diffusion ceases). Fig. 3 (right) displays the time evolution of D_w for the orbit shown in the figure at the left that fills the stochastic layer.

The WM has been derived for a one–dimensional Hamiltonian acted upon by a periodic, time–dependent perturbation. For a time–independent formulation we only need to work on the extended phase space, introducing a new variable P as the conjugate momentum to the perturbing phase τ. Therefore the time–dependent Hamiltonian (6) becomes an autonomous two–dimensional one, where the unperturbed part is now $H_r + \Omega P$. Thus we have a two–dimensional model describing the interaction between two coupling resonances.

V. THE RATE OF ARNOLD DIFFUSION

In this Section we shall compute the rate of Arnold diffusion for the Hamiltonian given by equations (1)–(2). In other words, we will estimate the time–scale for the variation of the unperturbed integrals of motion for the former Hamiltonian. To this purpose we will follow the scheme and notation of CH79, Section 7.3. Let us take then initial conditions in such a way that the system is initially confined to move within a given resonance domain: the *guiding resonance*. Let us denote with V_g and $\boldsymbol{m}_g$ its associated amplitude and resonant vector, respectively, and let $\boldsymbol{\omega}^r = \boldsymbol{\omega}(\boldsymbol{I}^r)$ be the resonant frequency: $\boldsymbol{m}_g \cdot \boldsymbol{\omega}^r = 0$. Thus, the Hamiltonian (1)–(2) can be recast as

$$H = H_0(\boldsymbol{I}) + \epsilon V_g \cos(\boldsymbol{m}_g \cdot \boldsymbol{\theta}) + \epsilon V(\boldsymbol{I}, \boldsymbol{\theta}), \tag{18}$$

where $\epsilon < \epsilon_c$, $\boldsymbol{I}, \boldsymbol{\theta}$ are N–dimensional vectors ($N > 2$) and

$$\epsilon V = \epsilon \sum_{\boldsymbol{m} \neq \boldsymbol{m}_g} V_{\boldsymbol{m}}(\boldsymbol{I}) \cos(\boldsymbol{m} \cdot \boldsymbol{\theta}). \tag{19}$$

We introduce now the canonical transformation defined in Section II, $(\boldsymbol{I}, \boldsymbol{\theta}) \to (\boldsymbol{p}, \boldsymbol{\psi})$ such that: $\psi_i = \mu_{ik}\theta_k$; $I_i = I_i^r + p_k \mu_{ki}$, where μ_{ik} is a $N \times N$ matrix with $\mu_{1i} = m_{g_i}$ such that $\psi_1 = \boldsymbol{m}_g \cdot \boldsymbol{\theta}$ is the resonant phase. The action vector with components $(I_j - I_j^r)$ in the original (local) basis $\{\boldsymbol{u}_j, j = 1, \cdots, N\}$, has components p_j in the new (local) basis $\{\boldsymbol{\mu}_j, j = 1, \cdots, N\}$, which we construct taking advantage of the particular geometry of resonances in the action space (see Section III).

As already defined, $\boldsymbol{\mu}_1 = \boldsymbol{m}_g$ and since the vector $\boldsymbol{m}_g$ is orthogonal to the frequency vector $\boldsymbol{\omega}^r$ (due to the resonance condition), it seems natural to take $\boldsymbol{\mu}_2 = \boldsymbol{\omega}^r / |\boldsymbol{\omega}^r|$. The remainder vectors of the basis are $\boldsymbol{\mu}_k = \boldsymbol{e}_k$, $k = 3, \cdots, N$, being the vectors $\boldsymbol{e}_k$ orthonormal to each other and to $\boldsymbol{\mu}_2$. Let us take one of the $\boldsymbol{e}_k$, say $\boldsymbol{e}_s$, orthonormal to $(\partial[\boldsymbol{m}_g \cdot \boldsymbol{\omega}(\boldsymbol{I})]/\partial \boldsymbol{I})_{\boldsymbol{I}^r}$, the normal vector to the guiding resonance surface. Thus, in general, all the vectors $\boldsymbol{e}_k$ will be orthogonal also to $\boldsymbol{\mu}_1$, except $\boldsymbol{e}_s$. Geometrically, for $N = 3$, we have the following picture: the resonant vector $\boldsymbol{m}_g$ lies in the tangent plane to the energy surface at the resonant point $\boldsymbol{I}^r$, while the frequency vector is normal to that plane. Then the third vector, $\boldsymbol{e}$, lies in the intersection line between the tangent plane and the guiding resonance surface. In general $\boldsymbol{e}$ will not be orthogonal to $\boldsymbol{m}_g$ (unless $\boldsymbol{\omega}(\boldsymbol{I}) = \boldsymbol{I}$, which is the case for $H = \boldsymbol{I}^2/2$). Then, since $\boldsymbol{p} = p_i \boldsymbol{\mu}_i, i = 1, \cdots, 3$, we can say 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 (see below), and p_3 measures the departure from the resonant value along the guiding resonance layer, i.e., in the direction where Arnold diffusion proceeds. It follows then that in the case of a three–dimensional autonomous system, the diffusion is one–dimensional.

---

[6]For a different approach to diffusion process in phase space, see the example given by Saslaw (1985, Ch. 4) for the one–dimensional Fokker–Planck equation.

But for N degrees of freedom ($N > 2$), the subspace of intersection of the latter surfaces has dimension $N-2$ and, following CH79, we call this space the *diffusion surface.* The $N-2$ vectors $\boldsymbol{e}_k$ locally span (at the resonant value) a tangent plane to the diffusion surface called *the diffusion plane.* Then, in the new basis, the action may be written as: $\boldsymbol{p} = p_1\boldsymbol{m}_g + p_2\boldsymbol{\omega}^r/|\boldsymbol{\omega}^r| + \boldsymbol{q}$, where $\boldsymbol{q}$ is confined to the diffusion plane: $\boldsymbol{q} = q_k\boldsymbol{e}_k$, with $q_k = p_k$ for $k = 3, \cdots, N$.

As in Section II, we write the Hamiltonian (18) in terms of the new components of the action. Expanding up to second order in p_k, using the orthogonal properties of the new basis, recalling that ψ_1 is the resonant phase and neglecting the constant terms, we obtain for $k, l \neq 1$:

$$H(\boldsymbol{p},\boldsymbol{\psi}) \approx \frac{p_1^2}{2M_g} + \epsilon V_g \cos\psi_1 + |\boldsymbol{\omega}^r| p_2 + \frac{p_k p_l}{2M_{kl}} + \epsilon V(\boldsymbol{\psi}), \tag{20}$$

with:

$$\frac{1}{M_{kl}} = \mu_{ki}\frac{\partial\omega_i^r}{\partial I_j}\mu_{lj}, \qquad \frac{1}{M_g} \equiv \frac{1}{M_{11}} = m_{g_i}\frac{\partial\omega_i^r}{\partial I_j}m_{g_j}; \tag{21}$$

and where we have written $V(\boldsymbol{\psi})$ instead of $V(\boldsymbol{p},\boldsymbol{\psi})$ to retain order ϵ in the perturbation ($\boldsymbol{I} = \boldsymbol{I}^r$ or $\boldsymbol{p} = 0$).

In the absence of perturbation ($V = 0$), the components $p_k, k = 2, \cdots, N$ are integrals of motion, which we set equal to zero so that $\boldsymbol{I}^r$ is a point of the orbit. Then the Hamiltonian (20)–(21) reduces to:

$$H(\boldsymbol{p},\boldsymbol{\psi}) \approx H_1(p_1,\psi_1) + \epsilon V(\boldsymbol{\psi}), \tag{22}$$

where:

$$H_1 = \frac{p_1^2}{2M_g} + \epsilon V_g \cos\psi_1 \tag{23}$$

coincides with the resonant Hamiltonian H_r given in (3) and where the perturbing phases $\boldsymbol{\theta}$ in V must be written in terms of the new components ψ_k. Notice that the above Hamiltonian resembles the one used to understand the existence of the stochastic layer but, while Hamiltonian (6)–(8) represents a one–dimensional model acted upon by a time–dependent perturbation, Hamiltonian (22)–(23) describes an N–dimensional system.

To transform the phase variables, we take into account that the dot product is invariant under a change of basis. Recalling that $\psi_k = \mu_{kl}\theta_l$ 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 $\nu_k = m_i\mu_{ik}$. As we can readily see, while the m_k are integers, the quantities ν_k are, in general, real numbers.

As mentioned above, for $V = 0$ the p_k are integrals of motion and recalling that H_1 is also an unperturbed integral, we have the full set of N unperturbed integrals: H_1, p_2, q_k, $k = 3, \cdots, N$. But if we switch on the perturbation, the integrals will change with time. This variation is determined by the time dependence of $\varphi_{\boldsymbol{m}}$. Notice that in the case of the Hamiltonian (6)–(8) we have only one integral, H_r, and the time dependence of the perturbing phase is given beforehand: $\tau = \Omega t + \tau_0$. Thus, the next step is to calculate $\varphi_{\boldsymbol{m}}(t)$. This can be done using the equations of motion for the Hamiltonian (20)–(21), where $\dot\psi_j = \partial H/\partial p_j, j = 1, \cdots, N$. Performing the derivatives and replacing the unperturbed values for $p_2, q_k = 0, k = 3, \cdots, N$ we obtain:

$$\psi_j(t) \approx |\boldsymbol{\omega}^r| t\delta_{2j} + \frac{M_g}{M_{j1}}\psi_1(t) + \psi_{j_0},$$

where δ_{ij} is the familiar Kronecker delta and ψ_{j_0} is a constant. To get $\varphi_{\boldsymbol{m}}(t)$ we evaluate the dot product $\nu_i\psi_i$:

$$\varphi_{\boldsymbol{m}}(t) = \boldsymbol{m}\cdot\boldsymbol{\theta} = \boldsymbol{\nu}\cdot\boldsymbol{\psi} \approx \xi_{\boldsymbol{m}}\psi_1(t) + \omega_{\boldsymbol{m}} t + \beta_{\boldsymbol{m}}, \tag{24}$$

where:

$$\xi_{\boldsymbol{m}} = \nu_{k_{\boldsymbol{m}}}\frac{M_g}{M_{k1}} \text{ (sum over } k), \quad \omega_{\boldsymbol{m}} = \boldsymbol{m}\cdot\boldsymbol{\omega}^r = \nu_{2_{\boldsymbol{m}}}|\boldsymbol{\omega}^r|; \tag{25}$$

and $\beta_{\boldsymbol{m}}$ is a constant. The second of (25) can be obtained taking into account that $\boldsymbol{m}\cdot\boldsymbol{\omega}^r = (\nu_i\boldsymbol{\mu}_i).(\boldsymbol{\mu}_2|\boldsymbol{\omega}^r|)$ and using the fact that $\boldsymbol{\mu}_2$ is orthogonal to all $\boldsymbol{\mu}_i, i \neq 2$.

Now we are ready to compute the time variation of the unperturbed integrals. From (22), for $\dot p_k = -\partial H/\partial\psi_k$, $k \neq 1$ we easily find that:

$$\dot{\boldsymbol{p}}_*(t) \approx \epsilon\sum_{\boldsymbol{m}} \boldsymbol{\nu}_{\boldsymbol{m}} V_{\boldsymbol{m}} \sin\varphi_{\boldsymbol{m}}(t). \tag{26}$$

This equation holds for every component of the momentum $\boldsymbol{p}$, except for p_1, since the latter is not an unperturbed integral: $\boldsymbol{p}_* = (p_2, \cdots, p_N)$. Instead of p_1 we have H_1, thus we need to evaluate $\dot{H}_1$. Since H is an exact integral of motion, we may compute $\dot{H}_1$ using the relation $\dot{H}_1 = [H_1, H]$, where $[u,v] \equiv \partial u/\partial\psi_i \partial v/\partial p_i - \partial u/\partial p_i \partial v/\partial \psi_i$ denotes the Poisson bracket. Then it is straightforward to show that:

$$\dot{H}_1(t) \approx \epsilon \sum_{\boldsymbol{m}} \nu_{1_{\boldsymbol{m}}} V_{\boldsymbol{m}} \dot{\psi}_1(t) \sin\varphi_{\boldsymbol{m}}(t). \tag{27}$$

Next we have to integrate (26) and (27) to find the change of the unperturbed integrals over a certain time interval T. Thus if the initial conditions are such that the system is close to the separatrix of the guiding resonance, then we can again replace the motion in the vicinity of the separatrix by the motion laws on the unperturbed separatrix so that $T \to \infty$. Since the Hamiltonian H_1, given by (23), governs the unperturbed motion within the guiding resonance, ϵV_g is then the energy that corresponds to the separatrix. So we easily find its equations:

$$\dot{\psi}_1^x = 2\Omega_g \sin(\psi_1^x/2) \qquad \psi_1^x = 4\arctan\left[e^{\Omega_g(t-t_0)}\right],$$

where $\Omega_g(\epsilon) = (\epsilon|V_g/M_g|)^{1/2}$ is the small oscillation frequency for the resonant phase. This definition for the separatrix, which is convenient for future calculations, differs slightly from that given by (4). Indeed, now we assume that the stable point is $\psi_1^x = \pi$ so t_0 is the value of t when $\psi_1^x = \pi$. Besides, note that we have taken only the plus sign in $\dot{\psi}_1^x$. This means that the sign of $\dot{\psi}_1^x$ is given by the sign of ψ_1^x. In other words, we define the different branches in such a way that ψ_1^x and $\dot{\psi}_1^x$ have the same sign. For the upper branch, $0 \le \psi_1^x < 2\pi$ and for the lower one, $-2\pi \le \psi_1^x < 0$. Notice also that the relationship between this formula for the separatrix and the one given by (4) is $\psi_1^x(\hat{t}) = \psi_1^s(\hat{t}) \pm \pi$, where $\hat{t} = \Omega_g(t - t_0)$ and the $\pm$ sign corresponds to $\psi_1^s > 0$ and $\psi_1^s < 0$, respectively. If we consider both branches of the separatrix, then t_0 is the value of t when $\psi_1^x = \pm\pi$.

In order to compute $\Delta\boldsymbol{p}_*$ and ΔH_1, we have to perform the integrals:

$$\mathcal{I}_{\boldsymbol{m}} = \int_{-\infty}^{\infty} dt \sin\varphi_{\boldsymbol{m}}(t), \quad \mathcal{J}_{\boldsymbol{m}} = \int_{-\infty}^{\infty} dt \dot{\psi}_1(t) \sin\varphi_{\boldsymbol{m}}(t),$$

where $\varphi_{\boldsymbol{m}}$ is given by (24) with $\psi_1 = \psi_1^x$ and $\dot{\psi}_1 = \dot{\psi}_1^x$. Let us discuss in some detail the calculation of $\mathcal{I}_{\boldsymbol{m}}$. It is not difficult to show that for $\psi_1^x > 0$, $\varphi_{\boldsymbol{m}}$ can be written as:

$$\varphi_{\boldsymbol{m}}(\hat{t}) = \xi_{\boldsymbol{m}}\psi_1^s(\hat{t}) + (\omega_{\boldsymbol{m}}/\Omega_g)\hat{t} + \varphi_{\boldsymbol{m}}^0,$$

where:

$$\varphi_{\boldsymbol{m}}^0 \equiv \varphi_{\boldsymbol{m}}(t_0) = \xi_{\boldsymbol{m}}\psi_1^x(t_0) + \omega_{\boldsymbol{m}} t_0 + \beta_{\boldsymbol{m}}. \tag{28}$$

Expanding $\sin\{[\xi_{\boldsymbol{m}}\psi_1^s + (\omega_{\boldsymbol{m}}/\Omega_g)\hat{t}] + \varphi_{\boldsymbol{m}}^0\}$ and recalling that the term between square brackets is odd, we arrive at:

$$\mathcal{I}_{\boldsymbol{m}} = \frac{\sin\varphi_{\boldsymbol{m}}^0}{\Omega_g} \int_{-\infty}^{\infty} d\hat{t} \cos\left(\xi_{\boldsymbol{m}}\psi_1^s(\hat{t}) + \frac{\omega_{\boldsymbol{m}}}{\Omega_g}\hat{t}\right). \tag{29}$$

If we compare (29) with (12) and identifying $m/2 \to \xi_{\boldsymbol{m}}$, $\lambda_{\boldsymbol{m}} \to -\omega_{\boldsymbol{m}}/\Omega_g$, we obtain:

$$\mathcal{I}_{\boldsymbol{m}} = \frac{1}{\Omega_g} A_{2\xi_{\boldsymbol{m}}}(\lambda_{\boldsymbol{m}}) \sin\varphi_{\boldsymbol{m}}^0.$$

Since we have taken $\dot{\psi}_1^x > 0$, $\mathcal{I}_{\boldsymbol{m}}$ will be essentially different from zero when $\xi_{\boldsymbol{m}}\lambda_{\boldsymbol{m}} > 0$. If $\dot{\psi}_1^x < 0$, the same will happen for $\xi_{\boldsymbol{m}}\lambda_{\boldsymbol{m}} < 0$. Taking the absolute value of $\xi_{\boldsymbol{m}}$ and redefining $\lambda_{\boldsymbol{m}} = -\mathrm{sg}(\xi_{\boldsymbol{m}}\dot{\psi}_1^x)\omega_{\boldsymbol{m}}/\Omega_g$, where $\mathrm{sg}(\cdot)$ is the sign function, we can write

$$\mathcal{I}_{\boldsymbol{m}} = \frac{1}{\Omega_g} A_{2|\xi_{\boldsymbol{m}}|}(\lambda_{\boldsymbol{m}}) \sin\varphi_{\boldsymbol{m}}^0, \tag{30}$$

for both branches of the separatrix. In (30), $\varphi_{\boldsymbol{m}}^0$ is given by (28) with $\psi_1^x(t_0) = \pm\pi$. Whether $\mathcal{I}_{\boldsymbol{m}}$ has a finite or negligible value depends now on the sign of $\lambda_{\boldsymbol{m}}$.

The computation of $\mathcal{J}_{\boldsymbol{m}}$ is rather similar to that of $\mathcal{I}_{\boldsymbol{m}}$, except that, now the product $\dot{\psi}_1^x \sin\varphi_{\boldsymbol{m}}$, leads to two terms of the form:

$$\pm\cos[(\xi_{\boldsymbol{m}} \mp 1/2)\psi_1^x + \omega_{\boldsymbol{m}} t + \beta_{\boldsymbol{m}}].$$

It is clear that if, for instance $(\xi_{\boldsymbol{m}} + 1/2)$ is small, then the largest contribution to the integral comes from the term that includes $(\xi_{\boldsymbol{m}} - 1/2)$ and vice versa. In the limiting case in which $(\xi_{\boldsymbol{m}} + 1/2) = 0$ or $(\xi_{\boldsymbol{m}} - 1/2) = 0$, this term leads to free oscillations that do not contribute to a secular change in the integrals of motion. Thus we retain a single term, the one with the largest factor in ψ_1^x and we write it as:

$$-\mathrm{sg}(\xi_{\boldsymbol{m}})\cos[\mathrm{sg}(\xi_{\boldsymbol{m}})(|\xi_{\boldsymbol{m}}| + 1/2)\psi_1^x + \omega_{\boldsymbol{m}} t + \beta_{\boldsymbol{m}}].$$

Following the same procedure than that for $\mathcal{I}_{\boldsymbol{m}}$, for both branches of the separatrix we obtain:

$$\mathcal{J}_{\boldsymbol{m}} \approx -\mathrm{sg}(\xi_{\boldsymbol{m}}) A_{2|\xi_{\boldsymbol{m}}|+1}(\lambda_{\boldsymbol{m}}) \cos[\varphi_{\boldsymbol{m}}^0 + \mathrm{sg}(\xi_{\boldsymbol{m}}\psi_1^x)\pi/2]. \tag{31}$$

Now using (15) for $\lambda_{\boldsymbol{m}} \gg 1$ and evaluating the cosine in (31), it is straightforward to get:

$$\mathcal{J}_{\boldsymbol{m}} \approx -\frac{\omega_{\boldsymbol{m}}}{\xi_{\boldsymbol{m}}\Omega_g} A_{2|\xi_{\boldsymbol{m}}|}(\lambda_{\boldsymbol{m}}) \sin\varphi_{\boldsymbol{m}}^0 \approx -\frac{\omega_{\boldsymbol{m}}}{\xi_{\boldsymbol{m}}}\mathcal{I}_{\boldsymbol{m}}, \tag{32}$$

where for the second equality we have made use of (30).

Having computed $\mathcal{I}_{\boldsymbol{m}}$ and $\mathcal{J}_{\boldsymbol{m}}$, we are ready to evaluate the variation of the integrals over the interval T. The best choice for T depends on the form of the perturbation. To make things easier, we assume that the perturbation is symmetric as in the example given in Section IV (see CH79 for another choice). Then, since the perturbing terms come in pairs, one of the members of the pair works only during half a period of oscillation, while its contribution to the next half period is negligible. But for this second half, the other member of the pair contributes in the same way as the former does in the first half period. Then T, as in Section IV, is half a period of oscillation or a period of rotation of ψ_1 in the neighborhood of the separatrix of the guiding resonance and is given by

$$T(w) = \frac{1}{\Omega_g}\ln(32/|w|), \tag{33}$$

where $w = (H_1 - \epsilon V_g)/\epsilon V_g$ is the dimensionless energy relative to the separatrix. Thus, using (26), (27), (30) and (32), we find that the change in the unperturbed integrals over a period T is

$$\Delta\boldsymbol{P}_* \approx \frac{\epsilon}{\Omega_g}\sum_{\boldsymbol{m}} \boldsymbol{\nu}_{\boldsymbol{m}} Q_{\boldsymbol{m}} \sin\varphi_{\boldsymbol{m}}^0; \tag{34}$$

$$\Delta H_1 \approx -\frac{\epsilon|\boldsymbol{\omega}^r|}{\Omega_g}\sum_{\boldsymbol{m}} \frac{\nu_{1_{\boldsymbol{m}}}\nu_{2_{\boldsymbol{m}}}}{\xi_{\boldsymbol{m}}} Q_{\boldsymbol{m}} \sin\varphi_{\boldsymbol{m}}^0, \tag{35}$$

where $Q_{\boldsymbol{m}} = V_{\boldsymbol{m}} A_{2|\xi_{\boldsymbol{m}}|}(\lambda_{\boldsymbol{m}})$ and we have made use of the second equation in (25).

In a similar way as in Section IV, we will construct a mapping to describe not only the stochastic layer but also Arnold diffusion. We have already found the change in the integrals after a period of motion T. Let us now compute this variation for the phases $\varphi_{\boldsymbol{m}}^0$. As defined in (28), one immediately obtains

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

where $C_{\boldsymbol{m}}$ is a constant. For rotations we have $\Delta\psi_1^x(t_0) = \pm 2\pi$, depending on the sense of motion and, for oscillations, $\Delta\psi_1^x(t_0) = 0$, so that $C_{\boldsymbol{m}}$ takes the values $C_{\boldsymbol{m}} = \pm 2\pi\xi_{\boldsymbol{m}} \neq 0$ and 0 for rotations and oscillations, respectively. Rewriting (35) in terms of the dimensionless energy w instead of H_1, we arrive at the following mapping:

$$\bar{w} = w - \frac{|\boldsymbol{\omega}^r|}{\Omega_g}\sum_{\boldsymbol{m}} W_{\boldsymbol{m}} \sin\varphi_{\boldsymbol{m}}^0, \tag{36}$$

$$\bar{\varphi}^0_{\boldsymbol{m}} = \varphi^0_{\boldsymbol{m}} + \omega_{\boldsymbol{m}} T(\bar{w}) + C_{\boldsymbol{m}}, \tag{37}$$

where $W_{\boldsymbol{m}} = (\nu_{1_{\boldsymbol{m}}} \nu_{2_{\boldsymbol{m}}} Q_{\boldsymbol{m}})/(\xi_{\boldsymbol{m}} V_g)$ and the bar indicates, as before, the values of the variables after crossing the surface $\psi_1 = \pm\pi$.

The mapping given by (36)–(37) is, in some sense, similar to the WM defined in (17). One of the main differences is that the first equation in (17) includes only one perturbing term, while we have several in (36). Since the amplitudes $W_{\boldsymbol{m}}$ depend exponentially on $\lambda_{\boldsymbol{m}}$ (through the MAI), we may assume that one of them is much larger than the rest. Let that term be $W_{\boldsymbol{l}}$, i.e. $W_{\boldsymbol{l}} \gg W_{\boldsymbol{m}}$ for all $\boldsymbol{m} \neq \boldsymbol{l}$, then at this order, the mapping (36)–(37) reduces to the WM, which describes the motion (variation of w) across the stochastic layer and fix all its properties (width, diffusion rate, etc.). As $W_{\boldsymbol{l}}$ is responsible for the formation of the stochastic layer in a neighborhood of the separatrix of the guiding resonance, we call this perturbing term the *layer resonance*, as in CH79.

Recalling that the total energy (H) is constant and since the variation of w (or H_1) is essentially due to the largest term in (36), the remaining (smaller) terms in the mapping reflect the variation of the rest of the unperturbed integrals p_k. That is, all the terms in the mapping, except the largest one $W_{\boldsymbol{l}}$, are responsible for the time variation of the $(N-1)$–dimensional integral $\boldsymbol{P}_* \equiv (p_2, \boldsymbol{q})$, and variations in $\boldsymbol{q}$ imply motion along the stochastic layer, i.e., Arnold diffusion. Let us state this clearly: due to the presence of several perturbing terms in the Hamiltonian (18)–(19) (besides the resonant perturbation), the unperturbed energy H_0 is not preserved and the vector $p_2\boldsymbol{\mu}_2$, normal to the energy surface, takes into account its variation: $\dot{H}_0 \sim |\boldsymbol{\omega}^r|\dot{p}_2$. The remainder components $p_k = q_k, k = 3, \cdots, N$ lie on the diffusion plane and their variations are confined to the latter subspace. Thus Arnold diffusion is driven by the smaller terms of the perturbation, which, following CH79, we call *driving resonances.* Notice that, though we are using the term 'resonances' to refer to the perturbing terms, they are not taken as actual resonances, since we have assumed that $\boldsymbol{m} \cdot \boldsymbol{\omega}^r \neq 0$ for $\boldsymbol{m} \neq \boldsymbol{m}_g$.

For the sake of convenience, let us introduce the quantities: $v_{\boldsymbol{m}} = W_{\boldsymbol{m}}/W_{\boldsymbol{l}} \ll 1,\ r_{\boldsymbol{m}} = \omega_{\boldsymbol{m}}/\omega_{\boldsymbol{l}},\ s = w/w_s$, where w_s, is the width of the stochastic layer given by $w_s \approx |\boldsymbol{\omega}^r|\lambda_{\boldsymbol{l}} W_{\boldsymbol{l}}/\Omega_g$ (see Section IV) and $\lambda_{\boldsymbol{l}} = -\omega_{\boldsymbol{l}}/\Omega_g > 0$. While we deal with just one phase in the WM here we have several phases. The layer resonance phase $\varphi^0_{\boldsymbol{l}}$ behaves in a similar way as τ_0 in the WM (see Section IV), so it is useful to write all the phases in terms of $\varphi^0_{\boldsymbol{l}} \equiv \tau$. Then, using (37) and the above defined quantities, one readily finds that $\Delta\varphi^0_{\boldsymbol{m}} = r_{\boldsymbol{m}}\Delta\tau + C_{\boldsymbol{m}} - r_{\boldsymbol{m}} C_{\boldsymbol{l}}$, so after t iterations of the map, we obtain

$$\varphi^0_{\boldsymbol{m}}(t) = r_{\boldsymbol{m}}\tau(t) + b_{\boldsymbol{m}} t + d_{\boldsymbol{m}}, \quad \boldsymbol{m} \neq \boldsymbol{l}, \tag{38}$$

where: $b_{\boldsymbol{m}} = C_{\boldsymbol{m}} - r_{\boldsymbol{m}} C_{\boldsymbol{l}}$ and $d_{\boldsymbol{m}}$ is a constant. For $\boldsymbol{m} = \boldsymbol{l}$ and with the help of (33), the second term in (37) can be written in terms of $w = s w_s,\ |s| \leq 1$, as $\omega_{\boldsymbol{l}} T(\bar{w}) = \lambda_{\boldsymbol{l}} \ln|\bar{s}| - G'$, where $G' = \lambda_{\boldsymbol{l}} \ln(32/w_s)$. Recalling the value of w_s and separating the layer resonance in the sum in (36) (for which $r_{\boldsymbol{l}} = 1, b_{\boldsymbol{l}} = 0$), the mapping reduces to:

$$\bar{s} = s - \frac{1}{\lambda_{\boldsymbol{l}}}\sin\tau - \frac{1}{\lambda_{\boldsymbol{l}}}\sum_{\boldsymbol{m}\neq\boldsymbol{l}} v_{\boldsymbol{m}} \sin(r_{\boldsymbol{m}}\tau + b_{\boldsymbol{m}} t + d_{\boldsymbol{m}}), \tag{39}$$

$$\bar{\tau} = \tau + \lambda_{\boldsymbol{l}} \ln|\bar{s}| - G \mod(2\pi), \tag{40}$$

where $G = G' + C_{\boldsymbol{l}}$. The constant $C_{\boldsymbol{l}}$ is zero or $\pm 2\pi\xi_{\boldsymbol{l}} \sim \pm 2\pi|\boldsymbol{l}|/|\boldsymbol{m}_g| \ll G'$. As we note, the perturbation depends not only on the layer resonance phase, but also explicitly on time (unless $b_{\boldsymbol{m}} = 0$, which is the case for integers values of $\xi_{\boldsymbol{m}}$). The mapping (39)–(40) will allow us to infer something about the correlations of the driving resonance phases $\varphi^0_{\boldsymbol{m}}$. Indeed, from (39) we immediately get

$$\Delta' s(t) = \frac{1}{\lambda_{\boldsymbol{l}}}\sum_{t'=0}^{t}\left(\sin\tau(t') + \sum_{\boldsymbol{m}\neq\boldsymbol{l}} v_{\boldsymbol{m}} \sin\varphi^0_{\boldsymbol{m}}(t')\right).$$

For negligible values of $v_{\boldsymbol{m}}$, the mean square value of $\Delta' s(t)$ behaves similarly to $\Delta' w(t)$ in the WM. The successive values of τ are nearly random for small times while we have assumed that, for large times, the correlations among different values of τ reduce the diffusion across the layer. In the limit when $t \to \infty$, these correlations are such that the diffusion stops. Therefore, if the amplitude of the driving resonances are small enough (exponentially small with respect to that of the layer resonance), we may expect them to produce a slight distortion of the stochastic layer. It looks like this cannot be true if the driving resonance phases are correlated (see below and CH79 Section 7.2). Then, due to the smallness of all the $v_{\boldsymbol{m}}$, we may assume that the successive values of the driving resonance phases, $\varphi^0_{\boldsymbol{m}}$, are partially random, in the sense that:

$$\overline{\left(\sum_{t'=0}^{t} \sin\varphi^0_{\boldsymbol{m}}(t')\right)^2} \approx \frac{R}{2}t, \qquad t\to\infty, \tag{41}$$

with $R<1$ and where $R=1$ corresponds to the completely random case. This factor R, introduced by Chirikov (CH79), is called the *reduction factor* and takes into account the dependence of $\varphi^0_{\boldsymbol{m}}$ on τ. That is, for large times, the system moves within the external part of the stochastic layer, where the presence of stability islands leads to a slow diffusion. Chirikov called this approximation, *reduced stochasticity* in contrast to the limiting stochasticity approximation ($R=1$), where the reduction factor can be taken as the relative size of the central part of the stochastic layer $R\sim\frac{1}{4}$ (see CH79 for a more accurate determination). It must be added that the reduced stochasticity approximation is the central point in the discussion of Arnold diffusion. Indeed, only when the sum in (41) grows with time may the unperturbed integrals change along the stochastic layer. The validity of this assumption is however uncertain (see Section VII). Actually Chirikov extrapolates this approximation from a particular example (Arnolds' model), for which it seems to be a plausible assumption (see next section).

Let us define then the diffusion tensor on the diffusion plane as

$$D_{ij} = \frac{\overline{\Delta' q_i(t)\Delta' q_j(t)}}{tT_a}, \tag{42}$$

where T_a is the mean period of motion within the stochastic layer of the guiding resonance defined by

$$T_a(w_s) \approx \int_0^1 T(s)ds \approx \frac{1}{\Omega_g}\ln(32e/w_s),$$

with $w=sw_s$ and $T(w)$ given by (33) (see CH79 for details). The cross mean value in (42) means that we have to compute the total variation of the unperturbed integrals $q_k, k=3,\cdots,N$ over t periods of motion:

$$\Delta' q_k(t) = \sum_{t'=0}^{t} \Delta q_k(\varphi^0_{\boldsymbol{m}}(t')),$$

where $\Delta q_k(\varphi^0_{\boldsymbol{m}})$ is given by (34) for $k\neq 2$ and $\varphi^0_{\boldsymbol{m}}(t')$ by (38). Denoting with $\rho_{\boldsymbol{m}_i}=\nu_{\boldsymbol{m}_i}, i=3,\cdots,N$, it is straightforward to obtain:

$$\Delta' q_i(t)\Delta' q_j(t) \approx \frac{\epsilon^2}{\Omega_g}\sum_{\boldsymbol{m},\boldsymbol{m}'} \rho_{\boldsymbol{m}_i}\rho_{\boldsymbol{m}'_j} Q_{\boldsymbol{m}} Q_{\boldsymbol{m}'} \sum_{t',t''=0}^{t} \sin\varphi^0_{\boldsymbol{m}}(t')\sin\varphi^0_{\boldsymbol{m}'}(t''). \tag{43}$$

The latter equation reveals the existence of interference between different driving resonances. For instance, given a pair of driving resonances, evaluated at the same instant, the interference term has the form: $\cos[(r_{\boldsymbol{m}}\pm r_{\boldsymbol{m}'})\tau + (b_{\boldsymbol{m}}\pm b_{\boldsymbol{m}'})t + (d_{\boldsymbol{m}}\pm d_{\boldsymbol{m}'})]$. Though it seems difficult to cope with such terms when $\boldsymbol{m}\neq\boldsymbol{m}'$ and $t'\neq t''$, it is clear that the most important term is that for $r_{\boldsymbol{m}}=r_{\boldsymbol{m}'}$, $b_{\boldsymbol{m}}=b_{\boldsymbol{m}'}$ and $t'=t''$ (see CH79 for details). Therefore keeping only such terms in (43), by means of (41), (42), (43) and recalling that $Q_{\boldsymbol{m}}=V_{\boldsymbol{m}}A_{2|\xi_{\boldsymbol{m}}|}(\lambda_{\boldsymbol{m}})$, we obtain:

$$D_{ij} \approx \frac{\epsilon^2 R}{2T_a\Omega_g^2}\sum_{\boldsymbol{m}\neq\boldsymbol{l}} \rho_{\boldsymbol{m}_i}\rho_{\boldsymbol{m}_j} V^2_{\boldsymbol{m}} A^2_{2|\xi_{\boldsymbol{m}}|}(\lambda_{\boldsymbol{m}}). \tag{44}$$

From (44), using the asymptotic result (13) for $A_{2|\xi_{\boldsymbol{m}}|}$ ($\lambda_{\boldsymbol{m}}\gg 2|\xi_{\boldsymbol{m}}|$) and taking into account that $\omega_{\boldsymbol{m}}=\nu_{2_{\boldsymbol{m}}}|\boldsymbol{\omega}^r|$, we arrive at the following expression for the diffusion tensor

$$D_{ij} \approx \tilde{D}(\epsilon)\sum_{\boldsymbol{m}\neq\boldsymbol{l}} \frac{\rho_{\boldsymbol{m}_i}\rho_{\boldsymbol{m}_j}}{\nu_{2_{\boldsymbol{m}}}}\left|\frac{2\omega_{\boldsymbol{m}}}{\Omega_g}\right|^{4|\xi_{\boldsymbol{m}}|} \frac{V^2_{\boldsymbol{m}}}{\Gamma^2(2|\xi_{\boldsymbol{m}}|)} e^{-\frac{\pi|\omega_{\boldsymbol{m}}|}{\Omega_g}}, \tag{45}$$

where $\tilde{D}(\epsilon)=2\pi^2 R\epsilon^2/T_a|\boldsymbol{\omega}^r|^2$ and $\Omega_g\sim\sqrt{\epsilon}$. As we note the diffusion tensor is symmetric, then in general depends on $N-2$ coefficients. Whether Arnold diffusion spreads over the whole diffusion plane, it depends on the existence of the set of N independent vectors $\boldsymbol{m}$ (including $\boldsymbol{m}_g$). Therefore, at least three resonances – guiding, layer and driving– are required for the existence of Arnold diffusion. In the latter case, the diffusion will be confined to a one-dimensional subspace of the diffusion plane.

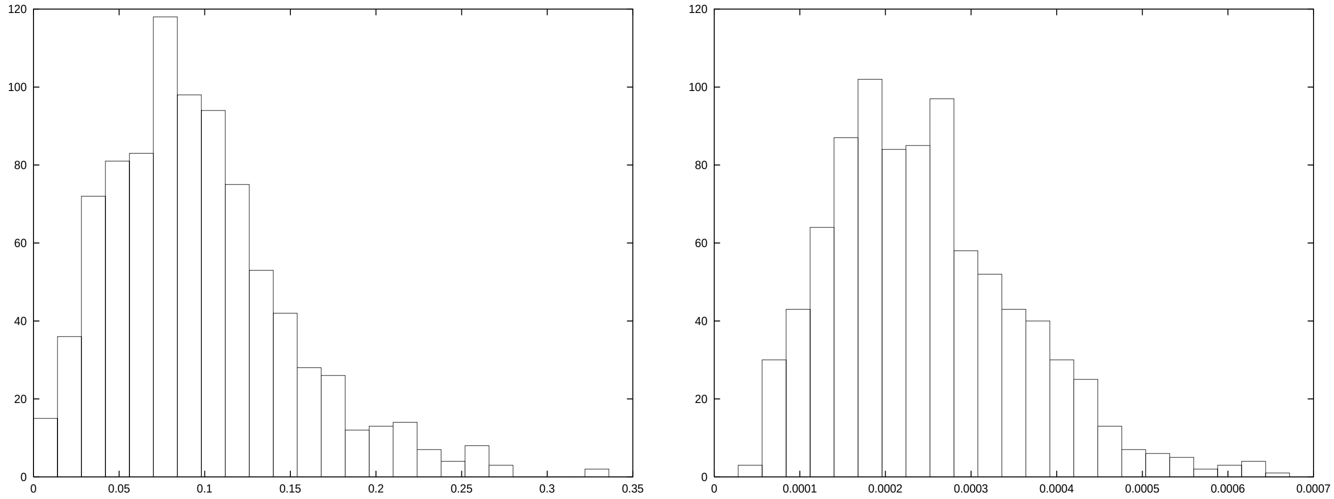


FIG. 4 Left, the distribution of the R values for ~ 1000 initial conditions computed using the mapping (46) for $\lambda = 12,\ v = 1.13\times10^{-3},\ r \approx 0.079,\ b \approx 3.260,\ d \approx 9.869,\ G \approx 70$. R is defined by equation (41). Right, the same but for L (see text).

VI. SOME NUMERICAL EXPERIMENTS

In this section we perform some numerical experiments in order to investigate the reduced stochasticity approximation by means of the mapping (39)–(40) for only one driving resonance, namely

$$\bar{s} = s - \frac{1}{\lambda}\sin\tau - \frac{1}{\lambda}v\sin(r\tau + bt + d), \qquad \bar{\tau} = \tau + \lambda\ln|\bar{s}| - G. \tag{46}$$

We also check up the hypothesis that the finite width of the stochastic layer is due to the existence of correlations between successive values of the layer resonance phase. It is worth mentioning that Arnold's example, for which diffusion along the layer has been rigorously proved, can be reduced to the mapping (46) with $b = 0$.

To study correlations in the driving resonance phase we just need to compute R by means of (41), while the correlations in the layer resonance phase will be measured by the quantity L defined as

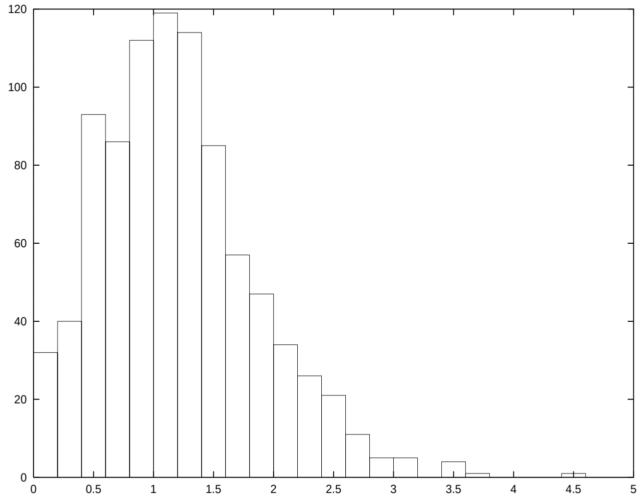
$$\overline{\left(\sum_{t'=0}^{t}\sin\tau(t')\right)^2} \approx \frac{L}{2}t, \qquad t\to\infty,$$

Though we deal with the simple case of one driving resonance, this mapping involves several free parameters. However, the definition of G and the smallness of the amplitude of the driving resonance impose the following restrictions $G \sim \lambda^2 \sim 1/\epsilon \gg 1, v \ll 1, r \ll 1$.

For given values of the parameters in (46), we iterate the mapping for 1000 initial conditions taken at random within the intervals $-1 \le s_0 \le 1$ and $0 \le \tau_0 \le 2\pi$. For each orbit, R and L are computed as the average of the 10 values obtained for every subinterval $\Delta t = 7\times10^5$, being then the total motion time $t = 7\times10^6$. From the whole set of random initial conditions, we have to remove those values of s_0 and τ_0 within a stability island of the stochastic layer (see Fig. 3). Indeed, for initial conditions within any such resonance domains, the stochastic layer does not exist; diffusion occurs neither across nor along the layer. Moreover, the presence of these resonances is responsible for the slow diffusion in the outer parts of the layer leading to the introduction of the reduction factor R. To eliminate initial conditions yielding to stable motion we skip those pairs (τ_0, s_0) for which the corresponding value of L satisfies the empirical condition: $L < L_0/10$, L_0 being the mode of the L distribution.

In Fig. 4 we show, as an example, the distribution of the ~ 1000 values of R and L for a given set of parameters, namely, $\lambda = 12,\ v = 1.13\times10^{-3},\ r \approx 0.079, b \approx 3.260,\ d \approx 9.869$ and $G \approx 70$. We clearly see that while the mean L is rather small, the computed mean value of R is finite, yet far from the rough theoretical estimation (~ 0.25). Notice, however, that both distributions are rather wide. In particular, the figure at the left shows that small values of R are much likely than large. So, it could happen that $R \to 0$ and then Arnold diffusion does not work.

Let us consider another example with a larger value of r, $r \approx 0.376$. The computed distribution for L resembles that presented in Fig. 4, while Fig. 5 shows that the distribution for R changes, but it is still non–negligible for small



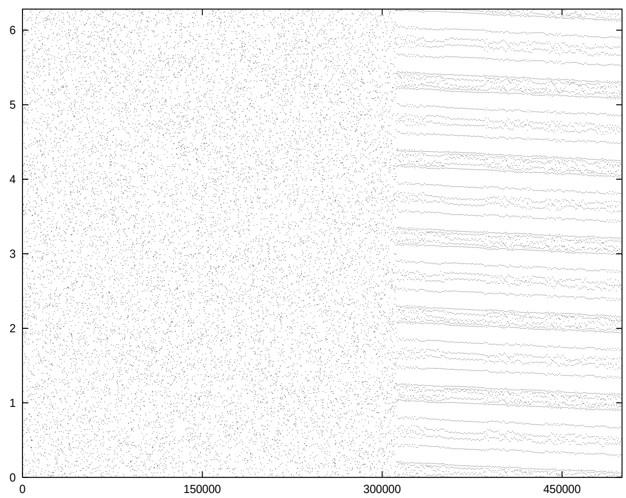


FIG. 5 (left) The distribution of the R values for $\sim$ 1000 initial conditions computed using the mapping (46) with the same parameters as in Fig. 4, but $r \approx 0.376$. FIG. 6 (right) The time evolution of φ for $\tau_0 \approx 4.601, s_0 \approx -0.567$ and the same parameters as in Fig. 4. Only 5% of the points were plotted.

R. Not only does the mean value of R increase, but also we observe very large fluctuations. These fluctuations may be due, in fact, to the presence of resonance domains in the outer part of the stochastic layer. Indeed, even though we took initial conditions out of the resonance domains, there are orbits that remain close to some islands for large periods of time ('sticking phenomena'), giving rise to such large fluctuations in the diffusion and may explain very low values of R. To guess something about the appearance of very large values of R, we computed the time evolution of the driving phase, $\varphi = r\tau + bt + d$, for one initial condition yielding a large R value ($R > 3$). Figure 6 displays the evolution of φ for $\tau_0 \approx 4.601, s_0 \approx -0.567$ during 5×10^5 iterations. For $t \lesssim 3 \times 10^5$, φ looks like ergodic, it covers uniformly the interval $(0, 2\pi)$. However, for $t > 3 \times 10^5$, the successive values of the phase present a strong correlation, φ varies within a finite set of values. That is, φ seems to be close to a high order resonance. The stochastic layer for this initial condition shows that, for large times, the orbit remains very close to the lower edge of the layer, $s \approx -1$, while τ covers completely the interval $(0, 2\pi)$ many times. This particular behavior leads to a much larger value of the sum (41) than that for the case of a nearly uniform distribution of φ (as for $t \lesssim 3 \times 10^5$).

Under the restrictions imposed to the reduced stochasticity approximation, the computed values of R and L satisfy, within fluctuations, the condition $L \ll R$ for different values of λ, v, r and G. The L values seem to be nearly independent of r, b and d, while the R values are more sensitive to those parameters (as shown in Figures 4 and 5). It should be also mentioned that while L increases with v, R decreases, but this variation is small for $v \lesssim 10^{-2}$. This behavior is similar to that found by Chirikov in his early experiments (CH79). Notice, however, that our simulations encompass smaller values of v and larger values of G ($G \gtrsim 25$), than Chirikov's. It seems that R weakly depends on b and d, provided that b is not too small. Another feature to point out is that, in a similar way as shown in Fig. 6 for large times, there should be correlations between successive values of the driving resonance, for certain values of r, b and G.

From the above results we may infer that, at least for the simple case of only one small perturbing term, the reduced stochasticity approximation is, statistically, a more or less realistic assumption. Keep in mind, however, that it is rather hard to test this hypothesis for the general case, since the mapping (39)–(40) involves several parameters.

VII. DISCUSSION

We start this section reviewing Arnold's work (Arnold 1964). In that note Arnold proved in a rigorous way and for a particular example, the existence of a motion along the stochastic layer of a given guiding resonance. Namely, he proved that in his model of 2.5 degrees of freedom, it is possible to find a trajectory in the vicinity of the separatrix of the guiding resonance that connects two points, p and q, separated by a distance $|p - q| \sim 1$, i.e. independent of the perturbation parameter. Clearly, the latter is much larger than the width of the layer, $w_s \sim \exp(-1/\sqrt{\epsilon})$. Arnold's proof rests on the existence of a chain of tori along the guiding resonance that may provide a path to the orbit. If these tori are very close to each other, then it is possible for an orbit to transit over that chain. Since every torus in the chain is labeled by an action value, we may have then a large variation in the unperturbed integrals.

This mechanism, that allows the existence of motion along the stochastic layer, is known (in mathematical literature) as *Arnold Mechanism* while the name *Arnold diffusion* is generally used (in physical literature) for referring to the global phase space instability (see Lochak 1999 and Giorgilli 1990 for details). Nevertheless, as far as I know, it is not possible to extend the latter mechanism to a generic Hamiltonian. One of the main difficulties is related to the construction of such a chain of tori. The result obtained in Section V for the diffusion tensor, is just an estimate of the time–scale for the motion along the stochastic layer (where the guiding resonance is determined by the vector $\boldsymbol{m}_g$), but we have not proved its existence. Indeed, we found that the change in the unperturbed integrals, in each period of motion, is exponentially small. In a somewhat implicit way (through the reduced stochasticity approximation), we assume that after many periods of motion, this variation is large (~ 1). However, from the mathematical side of the problem, the latter assumption is not well sustained.

On the other hand, nothing has been said about the mechanism under which the motion may fill up the whole web of stochastic layers, i.e., a global instability. In purely mathematical papers, the studies are restricted just to the above mentioned problem, without any attempt to describe the motion as a diffusion process or a global instability. In this direction one of the obstacles in the rigorous approach is, roughly speaking, the following. Let us suppose a system be confined to the stochastic layer of a given guiding resonance, such as in the discussion given in Section V. The diffusion along the layer (if it exists) is governed by the equations (34) for the unperturbed integrals q_k. This is true unless the system is close to some other resonance. In such a case, we have to solve beforehand the problem of a Hamiltonian where the resonant part of the perturbation includes two resonances (with amplitudes V_g and $V_{g'}$ for the resonant vectors $\boldsymbol{m}_g$ and $\boldsymbol{m}_{g'}$, respectively). But the problem of a *multiple resonance* is a quite difficult one since the unperturbed Hamiltonian H_1 is non–integrable. Indeed, the unperturbed Hamiltonian of, say, a double resonance can be reduced to that of two coupled pendulums, which is known to be non–integrable. However, if we want to show that the stochastic motion may spread over the whole web, the intersection of resonance surfaces, or a multiple resonance, is required. Therefore, it is clear that from purely mathematical point of view, the problem is far from being completely solved. To close this mathematical side of the discussion, it would be interesting to recall the mathematician's thought on these matters (Lochak 1999): '*...the global instability properties of near integrable Hamiltonian systems, thirty years after the pioneering work of V.I. Arnold, are far from well–understood. It could almost be said that little progress has been made, and new ideas are definitely called for.*'

Numerical studies on Arnold diffusion are rare. Besides the early simulations due to Chirikov (CH79) and Tennyson et al. (1979), an illustrative numerical evidence is given by Kaneko and Bagley (1985). Here, the authors present a visualization of Arnold diffusion for a very simple system, two coupled standard maps (where the associated phase space is 5–dimensional). In Figure 1 of that paper, they show the passage from one (half–integer) guiding resonance to another (integer) guiding resonance. However, as far as I know, numerical results about the existence of Arnold diffusion in more realistic dynamical systems are still lacking. The difficulties in performing numerical experiments are obvious. They come from the fact that multidimensional systems are not easy to deal with, and that exponentially small quantities and exponentially large times are involved in the calculations. Therefore, it turns out that the numerical evidence compiled up to now is not enough to regard Arnold diffusion as an experimental fact.

There is, at least, one more problem with Arnold diffusion for real dynamical systems such as galaxies. The time–scale for the manifestation of the instability is rather large, of the order of $\epsilon^2 \exp(-1/\sqrt{\epsilon})$ with $\epsilon \to 0$. For a generic perturbation, this estimate depends on the number of degrees of freedom, N. Indeed, from (45) we see that D strongly depends on the quantities $\omega_{\boldsymbol{m}}$. A detailed analysis using an analytical perturbation $V(\boldsymbol{\theta})$ (i.e. all the derivatives of $V(\boldsymbol{\theta})$ are analytic) reveals that $D \sim \epsilon^{1+p} \exp(-1/\epsilon^q)$ where $q \approx 1/2N$ and $p = q/\alpha$ for $\alpha \geq 1$ (see CH79). These estimates show that, if Arnold diffusion exists, it needs exponentially large times to connect different regions of chaotic motion. In fact, for real dynamical systems it seems very difficult to control time–scales like the latter. In Galactic Dynamics, only time–scales less than or of the order of the Hubble time have physical meaning. Since exponentially small values of the perturbation are necessary in order that Arnold's mechanism works, then the proper question is: Does the time–scale for Arnold diffusion have an upper bound less than the Hubble time?

Let us give some estimations. If T_0 is a typical dynamical time for a galaxy, then the Hubble time is about $10^2 T_0$ to $10^3 T_0$. The time scale for Arnold diffusion depends exponentially on $|\omega_{\boldsymbol{m}}/\Omega_g|$ which in fact is large, since it is of the order of the ratio between the orbital period ($\sim T_0$) and the libration period ($\ll T_0$). Thus 10^3 periods seems to be a rather short time for Arnold diffusion to work. In this direction let me discuss the work by Merritt and Valluri (1996) (also Merritt and Friedman 1996). They investigate orbits of stars in a triaxial galaxy with a density profile that is in good agreement with the observed surface brightness profile of early–type elliptical galaxies. The density law depends on few parameters. One of them is an exponent, which measures the steepness of the density as a function of the radius. In this way they distinguish between models with strong and weak central mass concentrations. Therefore we may say that, for the strong cusp model, the potential near the center is hard. In addition they also consider a central (spherical) black hole and use the mass of the latter as a parameter. They argue that in their models with a strong cusp, the *Arnold Web is filled after* $10^2 - 10^3 T_0$, while in those models with a weak cusp, the orbits behave as regular boxlike orbits for the same period of time. Moreover they remark the fact that the computed Lyapunov exponents are similar for different regions of the phase space where the motion is stochastic. They then infer that these regions are connected via Arnold diffusion. Even though it is not clear if these models may be approximated by a near–integrable well behaved Hamiltonian, let us assume that this is the case. Thus, one should think of a multipolar expansion of the triaxial potential. But notice that, in this way, 'the perturbation' (let say, the dependence on θ and φ) should not differ too much from one to another model when keeping the traxiality parameter constant. [Let us recall that the semiaxis ratio or the degree of flatness of the potential is the natural perturbation parameter to understand the transition from regular to stochastic motion (see Papaphilippou and Laskar 1996, 1998; Cincotta 1993, Cincotta et al. 1996, Cincotta and Simó 2000).] Nevertheless, for the experiments under consideration, this parameter is fixed and therefore the differences are in the strength of the cusp. But from this point of view the models differ in the integrable part of the Hamiltonian, that is the spherical one. So the problem should be stated as follows: how do orbits in different integrable potentials behave when we add a fixed, presumably non–integrable, perturbation? It turns out that the hardness of the unperturbed potential near the center, will be the key point in the disappearance of regular for boxlike orbits (which pass through the center) when we add some multipolar terms. In fact this is a well know result at least for 2D systems (see for instance, Miralda–Escudé and Shwarzschild 1989, Cincotta and Simó 2000).

It seems that those models with a strong cusp exhibit a large chaotic behavior due to an overlap of resonances, that is, when $\epsilon \gtrsim \epsilon_c$, but not strictly due to Arnold diffusion. The Arnold Web is relevant only when the resonance layers exist, i.e., when the resonances are well separated. In this direction Wachlin and Ferraz–Mello (1998) investigated the same model as Merritt and Friedman using the Frequency Map Analysis (Laskar, 1993), but only for the case of a weak cusp. In Figures 3a–3d of that paper it is clearly shown that even though they are dealing with a, 'small perturbation' (weak cusp), the resonances are present only for low values of the energy. For higher energies, almost all resonances are destroyed by overlap. Moreover, these figures show that the motion is chaotic in a relatively large region of the energy shell, in contradiction to the hypothesis for pure Arnold diffusion. A similar problem is studied by Valluri and Merritt (1998) also by means of the Frequency Map Analysis, and their results (see for instance, Fig. 9) allow for the same conclusion. The overlap of resonances may provide a more efficient mechanism to connect different regions of phase space that, in case of smaller perturbations, appear isolated. Indeed, the diffusion rate in this case is, roughly, of the order of the size of the perturbation (or some power of it, see Section IV) but not exponentially small as in (45). Small values of the diffusion rate can be explained in terms of the sticking phenomena, that reduces dramatically the diffusion, but it has nothing to do with Arnold diffusion.

To illustrate this let us consider the following toy model

$$\tilde{H}(\boldsymbol{p},\boldsymbol{q}) = \frac{\boldsymbol{p}^2}{2} + \frac{1}{4}(x^4+y^4+z^4) + \epsilon x^2(y+z). \tag{47}$$

The full dynamics of this 3D system is investigated by Cincotta and Giordano (2000) and Cincotta et al. (2001) by means of the *Mean Exponential Growth factor of Nearby Orbits* (MEGNO)–see Cincotta and Simó (2000), Cincotta et al. (2001). The advantage of this model is that the quartic oscillator can be easily written in terms of action variables and that the coordinates have a simple Fourier expansion. Indeed, (47) can be recast as

$$H(\boldsymbol{I},\boldsymbol{\theta}) = H_0(\boldsymbol{I}) + \epsilon V(\boldsymbol{I},\boldsymbol{\theta}), \tag{48}$$

where H_0 is given by

$$H_0(\boldsymbol{I}) = A\big(I_1^{4/3} + I_2^{4/3} + I_3^{4/3}\big), \tag{49}$$

with $A = (3\beta/2\sqrt{2})^{4/3}, \beta = \pi/2K(1/\sqrt{2})$ ($K(k)$ denotes the complete elliptic integral), and the perturbation admits of the Fourier expansion

$$\begin{aligned} V(\boldsymbol{I},\boldsymbol{\theta}) = \hat{V}_{12} &\sum_{n,m,k=1}^{\infty} \alpha_{nmk}\Big(\cos\big(2(n+m-1)\theta_1 \pm (2k-1)\theta_2\big) + \cos\big(2(n-m)\theta_1 \pm (2k-1)\theta_2\big)\Big) + \\ \hat{V}_{13} &\sum_{n,m,k=1}^{\infty} \alpha_{nmk}\Big(\cos\big(2(n+m-1)\theta_1 \pm (2k-1)\theta_3\big) + \cos\big(2(n-m)\theta_1 \pm (2k-1)\theta_3\big)\Big), \end{aligned} \tag{50}$$

the function $\hat{V}_{ij}$ and the coefficients α_{nmk} being

$$\hat{V}_{ij} \equiv \hat{V}(I_i,I_j) = C I_i^{2/3} I_j^{1/3}, \quad C = \frac{3\beta}{4}, \qquad \alpha_{nmk} = \alpha_n\alpha_m\alpha_k, \qquad \alpha_s = \frac{1}{\cosh\big((s-1/2)\pi\big)}, \qquad \frac{\alpha_{s+1}}{\alpha_s} \approx \frac{1}{23}.$$

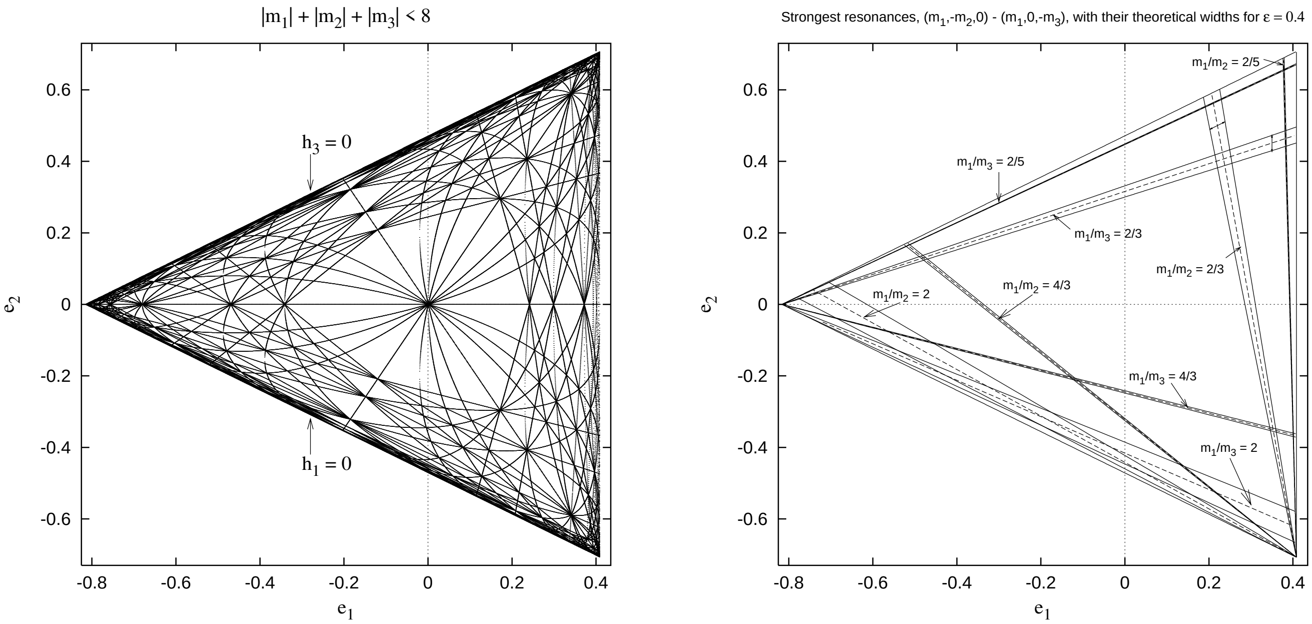


FIG. 6 (Left) Resonances of the unperturbed Hamiltonian (49) for $|m_1| + |m_2| + |m_3| < 8$ and energy $h = 1$ yielding the theoretical Arnold web on the energy surface. (Right) Strongest resonances and their theoretical widths. Arrows within resonances $(2, -3, 0)$ and $(2, 0, -3)$ indicate the direction in which $\boldsymbol{\Delta h}^r_{\boldsymbol{m}}$ oscillates about the corresponding resonant value.

We can split the perturbation in two, namely V_{xy} and V_{xz}, which, by introducing the integer vectors $\boldsymbol{l} = (l_1, l_2, 0), \boldsymbol{k} = (k_1, 0, k_3)$ and the new coefficients $\hat{\alpha}_{l_1 l_2}$ and $\hat{\alpha}_{k_1 k_3}$, read

$$V_{xy}(I_1, I_2; \theta_1, \theta_2) = \hat{V}_{12} \sum_{l_1, l_2} \hat{\alpha}_{l_1 l_2} \cos(l_1\theta_1 + l_2\theta_2),$$

$$V_{xz}(I_1, I_3; \theta_1, \theta_3) = \hat{V}_{13} \sum_{k_1, k_3} \hat{\alpha}_{k_1 k_3} \cos(k_1\theta_1 + k_3\theta_3).$$

The resonance structure of the unperturbed system on the energy surface and the theoretical widths of the principal resonances appearing in the perturbation are presented in Fig. 6. We take the energies in each degree of freedom, h_1, h_2, h_3, as action–like variables, so that the unperturbed Hamiltonian becomes $H_0 = h_1 + h_2 + h_3$ and the frequency vector is $\boldsymbol{\omega} = \sqrt{2}\beta\big(h_1^{1/4}, h_2^{1/4}, h_3^{1/4}\big)$. Then we pass to variables e_1, e_2, e_3, such that the e_3–axis is normal to the energy plane. In the figure at the right, the resonance widths are measured in terms of $\boldsymbol{\Delta h}$ instead of $\boldsymbol{\Delta I}$.

Fig. 7 displays the actual structure of action space (e–space) at different, low values of the perturbation parameter. The character of the motion (resonant, quasiperiodic and stochastic) is represented in gray scale, from white to black. We clearly distinguish resonances with their actual widths and the stochastic layers at their borders. We also note zones of stable and unstable motion around the resonance intersection. For increasing values of ϵ the overlap of the strongest resonances leads to a broad stochastic strip. The Arnold diffusion regime seems to be that of the smallest ϵ, when the stochastic layers are rather narrow and overlapping of resonances, if present, is negligible. We investigate diffusion along the stochastic layers, following an orbit with initial conditions in the stochastic domains and we also compute the diffusion coefficient. We note that even in the case of large perturbations, the diffusion is completely irrelevant over time–scales of the order of 10^9 periods. Only when the resonances labeled by the harmonics $(2, -1, 0)$ and $(2, 0, -1)$ overlap, the diffusion appears to be significant. Despite the fact that these two resonances intersect each other, only when a relatively large stochastic domain appears, the diffusion proceed along this zone. Recall that these resonances are very close one another, so that it could happen that two, very close, domains in action space that are not connected for some value of ϵ, at larger perturbations these domains appear connected by diffusion due not to Arnold diffusion but to the overlap of these resonances.

In fact, Chirikov recognized in more recent works that Arnold diffusion is a extremely slow process and it is confined to rather narrow stochastic layers, being then of almost no interest in real world. It is in this direction that Chirikov et al. (1985) suggested another mechanism, the so–called *modulational diffusion*, that may lead to a relatively fast diffusion along a resonance. By an abuse of language, they distinguish between “slow” and “fast” Arnold diffusion,

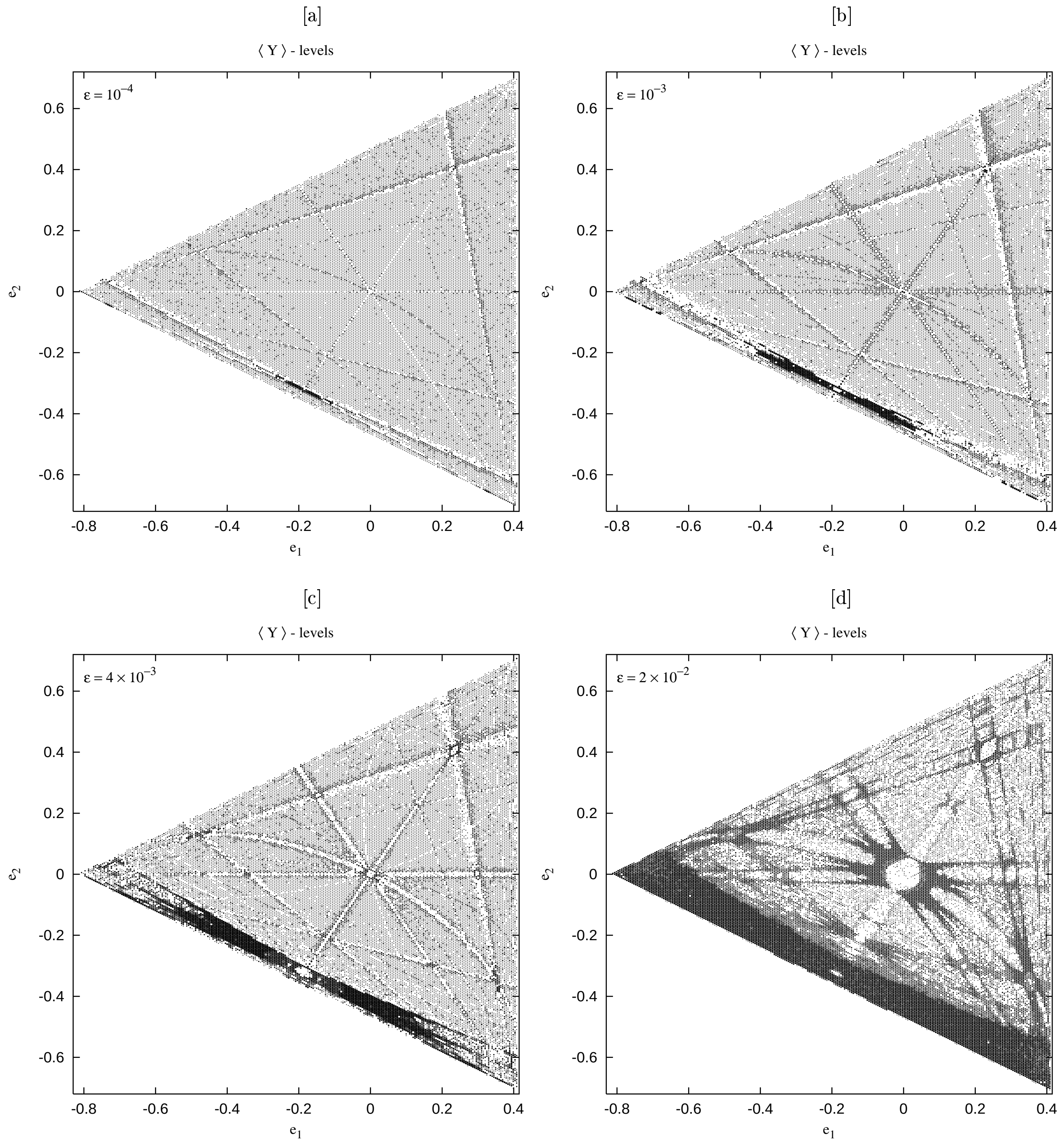


FIG. 7 MEGNO–levels ($\overline{Y}$) on the energy surface. The contour plot in gray scale means: white and light gray correspond to regular, stable motion while dark gray and black correspond to stochastic, unstable motion.

but in this second scenario, a complete overlap of several resonances is required in order to obtain a non–negligible diffusion coefficient.

Since we have neither any estimation of ϵ_c for the galactic models discussed above, nor a clear understanding of Arnold diffusion as a global instability, the question is still open. Nevertheless, let us think for a while about a real elliptical galaxy. In the present work we have neglected tidal effects, rotation, etc. We have further assumed that a galaxy, as a dynamical system, is represented by a near–integrable Hamiltonian. Even given these major simplifications, there is still a large list of unsolved questions related to Arnold diffusion: mathematical ones, involving the theory to support the instability, and physical ones, related to the manifestation of the instability in more or less actual systems in realistic times.

Summing up, Arnold diffusion looks like a suggestive mechanism to be taken into account in some applications, such as, for instance, the motion of charged particles in accelerators. Nevertheless, further advances in this field are required in order to illuminate whether Arnold diffusion could eventually play any role in Dynamical Astronomy.

ACKNOWLEDGEMENTS

I'm very grateful to S. Ferraz–Mello who introduced me into this fascinating subject and encouraged me to writing this paper; to C. Simó for his patience in explaining me, in an accessible way, some mathematical aspects and unsolved questions; and to P. Lochak, for reading a preliminary version of this manuscript and pointing out the works on modulational diffusion. Also I would like to thank A. Maciejewsky, I. Shevchenko, M. Valluri and my colleagues, J. Núñez, H. Vucetich, A. Brunini and, especially C. Giordano, for reading this manuscript and making helpful comments. The useful suggestions of an anonymous referee are also acknowledged. Finally, I would like to thank *CONICET-Argentina* and *FAPESP–Brazil* for financial support and the hospitality of the *Departament de Matemàtica Aplicada i Anàlisi de la Universitat de Barcelona* where this work was written and of the *Instituto Astronômico e Geofísico de la Universidade de São Paulo* where I spent a whole semester working on Chirikov's papers.

REFERENCES

[†]Electronic address: `pmc@fcaglp.unlp.edu.ar`

Arnold, V.I., 1964, Sov. Math.–Dokl. 5, 581.
Cincotta, P.M., 1993, Thesis, La Plata University.
Cincotta, P.M., Núñez, J.A. and Muzzio, J.C., 1996, ApJ 456, 274.
Cincotta, P.M. and Simó, C. 2000, A&AS, 147, 205.
Cincotta, P.M. and Giordano, C.M. 2000, in proceeding of X Brazilian Orbital Dynamics Colloquium, in press
Cincotta, P.M., Giordano, C.M. and Simó, C. 2001, preprint. To be submitted to Phys. D.
Chirikov, B.V., 1979, Phys. Rep. 52, 263 (CH79).
Chirikov, B.V., Lieberman, M., Shepelyansky, D. and Vivaldi, F., 1985, Phys. 14D, 289
Ferraz–Mello, S., 1996, unpublished.
Giorgilli, A, 1990, in 'Les Méthodes Modernes de la Méchanique Céleste', C. Froeschlé and D. Benest eds., pag. 249 (Paris: Frontièrs).
Gutzwiller, M.C., 1990, 'Chaos in Classical and Quantum Mechanics', (New York: Springer–Verlag).
Hasan, H. and Norman, C., 1990 ApJ, 361, 69.
Hénon, M. an Heiles, C., 1964, AJ, 69, 73.
Jackson, E.A., 1992, 'Perspectives of Nonlinear Dynamics', (Cambridge: Cambridge U.P.) (Vols. I and II).
Kaneko, K. and Bagley, R., 1985, Phys. Lett. 110A, 435.
Laskar, J., 1993, Phys. D 67, 257.
Lichtenberg, A.J. and Lieberman, M.A., 1992, 'Regular and Chaotic Dynamics', (New York: Springer–Verlag) (LL92).
Lochak, P., 1999, in 'Hamiltonian Systems with Three or More Degrees of Freedom', C. Simó ed., NATO ASI, pag 168 (Dordrecht: Kluwer).
Martinet, L., 1973, A&A 32, 329.
Merritt, D. and Friedman, T., 1996, ApJ, 460, 136.
Merritt, D. and Valluri, M., 1996, ApJ, 471, 82.
Merritt, D., 1999, PASP, 111, 129.
Miralda–Escudé, J. and Shwarzschild, M., 1989, ApJ, 339, 752
Papaphilippou, Y. and Laskar, J., 1996, A&A 307, 427.
Papaphilippou, Y. and Laskar, J., 1998, A&A 329, 451.
Rasband, N., 1983, 'Dynamics' (New York: John Wiley).

Rasband, N., 1990, ‘Chaotic Dynamics of Nonlinear Systems’, (New York: John Wiley).
Reichl, L. E., 1992, ‘The Transition to Chaos’, (New York: Springer–Verlag) (Vols. I and II).
Saslaw, W. C., 1985, ‘Gravitational Physics of Stellar and Galactic Systems’ (Cambridge: Cambridge U.P.).
Shevchenko, I., 1998, Phys. Scrip. 57, 185
Tennyson, J. L., Lieberman M. A. and Lichtenberg, A. J., 1979, in ‘Non–linear Dynamics and the Beam–Beam Interaction’, M. Month and J.C. Herrera eds., Am. Inst. Phys. Conf. Proc., pag. 57, 272.
Udry, S. and Pfenniger, D., 1988, A&A, 198, 135.
Valluri, M. and Merritt, D., 1998, ApJ, 506, 686.
Wachlin, F. C. and Ferraz–Mello, S., 1998, MNRAS, 298, 22.