ARNOLD INSTABILITY FOR NEARLY–INTEGRABLE ANALYTIC HAMILTONIAN SYSTEMS

LUIGI CHIERCHIA
Department of Mathematics, Terza Università di Roma
via C. Segre 2, 00146 Roma, Italy
Internet: luigi@matrm3.mat.uniroma3.it

ABSTRACT

A notion of instability, which is proposed as a definition for the so–called "Arnold diffusion", for one–parameter families of nearly–integrable analytic Hamiltonian systems, is introduced. An example of unstable system is given.

1. Introduction

One of the central questions in the theory of nearly–integrable Hamiltonian systems concerns the "stability" of the "action" variables, or, equivalently, of the perturbed integrals of motions. In the late fifties KAM theory* gave a partial answer to that question, ensuring, under general conditions, "stability" for "most" initial data. In 1964, Arnold [2] pointed out a mechanism for "instability", illustrating it on an example containing two parameters. The name "Arnold diffusion" for related phenomena was introduced in the physics literature by B.V. Chirikov.

Here, we introduce, in the context of one–parameter families of real–analytic Hamiltonian systems, the definition of "Arnold instability", which captures the main features of the mechanism discovered by Arnold. After a brief review of some known results, we discuss an explicit example (containing only one parameter), of an Arnold–unstable system. The *content* of the rest of the paper is:

* A.N. Kolmogorov, V.I. Arnold, J.K. Moser; see, e.g., Ref. [1] and references therein.

2. Nearly–Integrable Analytic Hamiltonian Systems

Let $\mathcal{M}$ denote the Cartesian product of some open connected set $G \subset \mathbb{R}^N$ times the standard N–torus $\mathbb{T}^N \equiv \mathbb{R}^N/(2\pi\mathbb{Z}^N)$. On the "phase space" $\mathcal{M}$, endowed with the standard symplectic form $\sum_{i=1}^N dI_i \wedge d\varphi_i$, consider one–parameter families of Hamiltonian functions $H(I,\varphi;\varepsilon)$. *Throughout this paper we shall assume the following real–analyticity and near–integrability assumptions*: for some $\bar\varepsilon > 0$, (i) H is real-analytic on $\mathcal{M} \times (-\bar\varepsilon, \bar\varepsilon)$; (ii) $H(I,\varphi;0) \equiv H_0(I)$. The study of the trajectories (or "orbits") generated by such $H(I,\varphi;\varepsilon)$, namely, the solutions of the Hamilton equations

$$\dot I_i = -\frac{\partial H}{\partial \varphi_i} \ , \qquad \dot\varphi_i = \frac{\partial H}{\partial I_i} \ , \quad \left[i=1,...,N;\ \dot{()} \equiv \frac{d}{dt}()\right] \ , \tag{1}$$

was considered by H. Poincaré "*le problème général de la Dynamique*" (see Ref. [3], Tome I, chap. I, No. 13). For $\varepsilon = 0$, I_1,...,I_N are N independent integrals in involution and all trajectories are quasi–periodic[†]: this is equivalent to say that the system is *completely integrable.* The basic *stability problem* is here whether, for $\varepsilon \neq 0$, the evolution of the "action" variables, $I_1(t)$,...,$I_N(t)$, stays close, for all t, to their initial values $I_1(0)$,...,$I_N(0)$.

3. Definition of "Arnold Instability"

Let us denote phase points by $z = (I,\varphi)$ and by ϕ_H^t the Hamiltonian flow generated by H *i.e.* $z(t) \equiv \phi_H^t(z) \equiv (I(t),\varphi(t))$ solves Eq. (1) with initial data $z(0) \equiv z \equiv (I,\varphi)$. A set $\mathcal{T} \subset \mathcal{M}$ will be called *invariant for* H (or H–invariant) if it is closed and $\phi_H^t(\mathcal{T}) \subset \mathcal{T}$ for all $t \in \mathbb{R}$. We now give the following definitions (compare with Ref. [2]).

Definition 1. *An n–ple of trajectories $z^{(1)}(t)$,...,$z^{(n)}(t)$ is called a "chain of heteroclinic trajectories" if there are $(n+1)$ invariant sets $\mathcal{T}^{(1)}$,...,$\mathcal{T}^{(n+1)}$ such that, for $1 \le i \le n$,*

$$\lim_{t\to-\infty} dist\left(z^{(i)}(t), \mathcal{T}^{(i)}\right) = 0 \ , \qquad \lim_{t\to+\infty} dist\left(z^{(i)}(t), \mathcal{T}^{(i+1)}\right) = 0 \ .$$

The sets $\mathcal{T}^{(1)}$,...,$\mathcal{T}^{(n+1)}$ are said to be connected by the chain $z^{(1)}(t)$,...,$z^{(n)}(t)$.

Definition 2. *A chain $z^{(1)}(t)$,...,$z^{(n)}(t)$ is called a "transition chain" if, for any $\rho > 0$, there exists a trajectory $z(t)$ and a time $T > 0$ such that: $z(0)$ is in a ρ–neighbourhood of $\mathcal{T}^{(1)}$; $z(T)$ is in a ρ–neighbourhood of $\mathcal{T}^{(n+1)}$; for all $0 \le t \le T$, $z(t)$ is in a ρ–neighbourhood of $Z \equiv \bigcup_{i=1}^n \{z = z^{(i)}(t) :\ t \in \mathbb{R}\}$.*

---

† A function on $\mathcal{M}$ is an integral for H if it is constant on the trajectories generated by H; "independent" means "functionally independent"; two (smooth) functions f and g are in involutions if their Poisson bracket $\{f,g\} = \sum_{i=1}^N \frac{\partial f}{\partial I_i}\frac{\partial g}{\partial \varphi_i} - \frac{\partial g}{\partial I_i}\frac{\partial f}{\partial \varphi_i}$ vanishes; a trajectory is quasi–periodic if it is conjugated to a linear flow on a torus.

Definition 3. *A Hamiltonian system* $(\mathcal{M}, H)$ *with* $\mathcal{M}$ *and* H *as in §2 is called "Arnold unstable" if, for some* $0 < \varepsilon_0 \leq \bar{\varepsilon}$*, the following two conditions are fulfilled:*
(i) *for all* $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ *there exist invariant sets* $\mathcal{T}(\varepsilon)$ *and* $\mathcal{T}'(\varepsilon)$*, which are continuous in the Hausdorff metric**, and* $p_1(\mathcal{T}(0)) \neq p_1(\mathcal{T}'(0))$*, where* $p_1(I, \varphi) = I$ *denotes the projection onto the "action" variables;*
(ii) *for all* $\varepsilon \neq 0$ *in* $(-\varepsilon_0, \varepsilon_0)$ *there exists a transition chain connecting* $\mathcal{T}(\varepsilon)$ *with* $\mathcal{T}'(\varepsilon)$.

In other words, an Arnold–unstable system posses, for all $\varepsilon \neq 0$ small enough, trajectories along which (some of) the "action" variables undergo a variation of order 1 and, furthermore, such trajectories "shadow" long chains of heteroclinic connections.

Notice that the above instability mechanism is completely different from the "linear drift" occurring along so–called "superconductivity channels". The classical example (due to N. N. Nekhoroshev*) is the region $I_1 = -I_2$ for the Hamiltonian $\frac{1}{2}(I_1^2 - I_2^2) + \varepsilon\, sin(\varphi_1 - \varphi_2)$, which admits the trajectory $I_2(t) = \varepsilon t = -I_1(t)$, $\varphi_1(t) = \varphi_2(t) = -\frac{1}{2}\varepsilon t^2$.

4. A Brief Review of Some Known Results

Even though the above definition is strongly inspired by Arnold's paper [2], it is not known whether the Hamiltonian system considered there is (in a suitable sense) Arnold unstable or not. Let us be more specific on this point. In Ref. [2], Arnold considered on $\mathcal{M} \equiv \mathbb{R}^3 \times \mathbb{T}^3$ the *two–parameter* family of Hamiltonians given by

$$H = I_1 + \frac{1}{2}(I_2^2 + I_3^2) + \varepsilon(cos\,\varphi_3 - 1) + \varepsilon\mu(cos\,\varphi_3 - 1)(cos\,\varphi_1 + sin\,\varphi_2)\ . \tag{2}$$

Since $\dot{\varphi}_1 = 1$, it is $\varphi_1 = t_0 + t$ so that the above Hamiltonian describes, effectively, a system with two "degrees of freedom" depending explicitly (and periodically) on time†. The "resonant" 2–torus $\{I_1 = \text{const},\ I_2 = \text{const},\ I_3 = 0 = \varphi_3,\ (\varphi_1, \varphi_2) \in \mathbb{T}^2\}$ is invariant for H *for all* ε *and* μ; such a torus is run by the linear flow $(\varphi_1, \varphi_2) \to (\varphi_1 + t, \varphi_2 + \omega t)$ where ω is the constant value of I_2: using the frequency $(1, \omega)$ to parametrize the above invariant torus we shall denote it by $\mathcal{T}(\omega)$. Of course the fact the $\mathcal{T}(\omega)$ is invariant for all values of the parameters ε and μ is a very special feature of the particular perturbation chosen by Arnold.

** If A and B are closed subsets of $\mathcal{M}$, the Hausdorff metric $\rho(\cdot, \cdot)$ is defined by

$$\rho(A, B) = max\{\sup_{x \in A} dist\,(x, B)\ ,\ \sup_{x \in B} dist\,(x, A)\}\ .$$

* See, e.g., §3.3, chapter 5 of Ref. [1].

† We prefer the "autonomous" formulation which allows a unified treatment; one should, however, bear in mind that the "action" I_1 is an auxiliary variable introduced so as to make the system autonomous and is not really relevant from a dynamical point of view. In Ref. [2] φ_1, φ_2, t correspond here to φ_3, φ_2, φ_1 and I_1, I_2 to I_3, I_2.

Using the language introduced in §3, the main result in Ref. [2] concerning the system with Hamiltonian as in Eq. (2), can be formulated as follows

Theorem 1 (Arnold [2]) *For every $\varepsilon, r > 0$ there exists a $\mu_0 > 0$ such that for all $0 < \mu < \mu_0$ there are invariant tori $\mathcal{T}(\omega)$ and $\mathcal{T}(\omega')$, with $|\omega - \omega'| > r$, which are connected by a transition chain.*

Note that, for all ε, μ, ω, the projection onto the I_2–value of $\mathcal{T}(\omega)$ is just ω. The proof of this result is perturbative in μ keeping $\varepsilon > 0$ fixed. For $\mu = 0$, although one can write down explicitly all trajectories, the system is not "completely integrable" in the sense described in §2. In fact (for $\mu = 0$) the trajectories with initial values of (I_3, φ_3) on the "separatrix" $\frac{1}{2}I_3^2 + \varepsilon(cos\,\varphi_3 - 1) = 0$ (and $I_3 \neq 0$) are not quasi–periodic.

The arguments used in Ref. [2] indicate that μ_0 may be taken to be of order

$$exp\,\big(-\frac{1}{\varepsilon^a}\big)\ , \tag{3}$$

for some constant $a \geq \frac{1}{2}$, leaving, thus, open the following

Question *Does there exist a positive integer n such that the system with Hamiltonian as in Eq. (2) with $\mu = \varepsilon^n$ is Arnold unstable** ?*

A result analogous to that of Arnold has been recently obtained by Z. Xia [4] in the elliptic restricted three–body problem of Celestial Mechanics. Such a system contains many parameters and by "similar" we mean the following. First, transition chains are detected in parameter regions where, to a positive value of one of the perturbative parameter, say ε, corresponds some "reference" hyperbolic trajectory (analogous to the separatrix of the pendulum). Then, *keeping fixed* ε, perturbation theory in a second parameter (corresponding to μ in Arnold's example) is performed. Again, the argument indicates that the result holds for $0 < \mu < \mu_0(\varepsilon)$ with $\mu_0(\varepsilon)$ as in Eq. (3) and the question whether the elliptic restricted three–body problem is Arnold unstable in parameter regions where the various parameters are related by, say, power laws is open.

As of now the only explicit system, which has been proven to be Arnold unstable is the "eccentric D'Alembert model" of Celestial Mechanics considered in Ref. [5], which may be described as follows. Consider a planet modelled by a "homogeneous" rotational ellipsoid whose center of mass revolves on a Keplerian ellipse of eccentricity μ. The planet is subject to the gravitational attraction of a fixed star occupying one of the foci of the ellipse. If $(\theta_1, \theta_2, \theta_3)$ denote the Euler angle

** The choice $\mu = \varepsilon^n$ in Arnold's model is somehow "natural", not only because it makes H belong to the "Poincarè category" described in §2, but also in view of the fact that near a "simple resonance" the dynamics generated by a "general" Hamiltonian $h(I) + \varepsilon f(I, \varphi)$ is equivalent to that of a Hamiltonian $\tilde{h}(I;\varepsilon) + \varepsilon g(I, \varphi_N; \varepsilon) + \varepsilon^n \tilde{f}(I, \varphi; \varepsilon)$ (everything being real–analytic in a neighbourhood of, say, $I_0 \times \mathbb{T}^N$ and of $\varepsilon = 0$); for more informations see Ref. [1]. Furthermore, in case of models with several perturbation parameters, it seems natural (at least from the applicative point of view) to require that such parameters be related by a power law.

of the planet, the Lagrangian* of the system is given by

$$\frac{J_3}{2}(\dot\theta_2\, cos\,\theta_1+\dot\theta_3)^2+\frac{J_1}{2}(\dot\theta_1^2+\dot\theta_2^2\, sen^{\,2}\theta_1)+c\int_{E(t)}\frac{dx}{|x_0(t)+x|}$$

where J_3 and $J_2\equiv J_1$ are the inertia moments of the planet, c is a physical constant†, $E(t)$ is the region occupied by the planet at time t and $x_0(t)$ is the position of the center of mass of the planet at time t. Using the Andoyer variables (see Ref. [1]) , one can find symplectic coordinates on $\mathcal{M}\equiv G\times\mathbb{T}^3$ (G being some open set of $\mathbb{R}^3$) so that the dynamics generated by the above Lagrangian is equivalent to the dynamics generated by the Hamiltonian

$$H=\omega I_1+\varepsilon\frac{I_2^2}{2}+\frac{I_3^2}{2}+\varepsilon\mu f(I,\varphi;\mu) \tag{4}$$

for a suitable function f real–analytic (on $\mathcal{M}\times(-\mu_0\infty,\mu_0)$), whose explicit expression can be found in Appendix 14 of Ref. [5]; here ε is essentially the ratio between the polar and equatorial radii of the planet. In Ref. [5] it is proven the following

Theorem 2 *There exists $n_0\ge 1$ such that for all integers $n\ge n_0$ the "eccentric D'Alembert problem"** with $\mu=\varepsilon^n$ is Arnold unstable.*

We remark that in Ref. [1,2,4] there is no mathematical definition of "Arnold diffusion" (to be precise in Ref. [2] such a terminology does not even appear while in Ref. [1] appears only in a note by the translator). In Ref. [5], a genearal and detailed terminology (including notions such as "diffusion paths") is introduced in the context of perturbations of "a priori unstable systems", which, roughly speaking, are "integrable systems possessing separatrices". Examples of such systems include the above Arnold's model with ε kept fixed and positive and with μ as perturbation parameter, and the three–body problem in the parameter region considered by Xia.

Finally, we point out that Theorem 2 is differently formulated in Ref. [5] (as the notions given in §3 are introduced in this paper).

5. An Arnold–Unstable System (Result and Scheme of Proof)

5.1. Main Result

Let $\mathcal{M}\equiv\mathbb{R}^3\times\mathbb{T}^3$ and consider

$$H(I,\varphi;\varepsilon)\equiv\varepsilon\Big(\frac{I_1^2}{2}+\frac{I_2^2}{2}\Big)+\frac{I_3^2}{2}+\varepsilon(cos\,\varphi_3-1)+\varepsilon^n\big(cos(\varphi_1+\varphi_3)+cos(\varphi_2+\varphi_3)\big)\ , \tag{5}$$

* See, e.g., Ref. [1] for generalities.

† $c=\kappa m_0 m/V$ where κ is Newton's gravitational constant, m_0, m are the masses of the star and of the planet, and V is the volume of the planet.

** That is, the Hamiltonian system on $\mathcal{M}$ with H as in Eq. (4).

where n is an integer greater than 1. We shall prove

Theorem 3 *There exists $n_0 \geq 1$ such that for all integers $n \geq n_0$ the above Hamiltonian system $(\mathcal{M}, H)$ is Arnold unstable.*

Remark (i) In some sense $(\mathcal{M}, H)$ is the "simplest" Arnold–unstable system with a perturbation* exhibiting the main feature of a "general perturbation". In particular the invariant sets $\mathcal{T}(\varepsilon)$, which will be shown to be connected, for $\varepsilon \neq 0$, by heteroclinic trajectories, will be non–trivial continuations of invariant sets for $\varepsilon = 0$ (i.e., $\mathcal{T}(0)$ is not invariant for ϕ_H^t as it happens in Arnold's example).

(ii) What is really "not general" in the Hamiltonian in Eq. (5) is the integrable part $H_0 \equiv \frac{I_3^2}{2}$, which is highly degenerate (depending only on one of the three action variables). Note that also the "eccentric D'Alembert problem" features a similar (although less dramatic) degeneracy, which is actually quite common in problems arising in Celestial Mechanics.

(iii) *No example is known of an Arnold–unstable system on $G \times \mathbb{R}^N$* (G open set in $\mathbb{R}^N$) *with "non–degenerate" integrable part* (of the form, say, $H_0 \equiv \frac{1}{2}\sum_{i=1}^N I_i^2$).

5.2. Outline of Proof

The system with Hamiltonian

$$h(I, \varphi_3; \varepsilon) \equiv \varepsilon\Big(\frac{I_1^2}{2} + \frac{I_2^2}{2}\Big) + \frac{I_3^2}{2} + \varepsilon(\cos\varphi_3 - 1)$$

may be explicitly solved as it decouples into a bidimensional completely integrable system times a pendulum. The instability region will be sought in a neighbourhood of the (ε–dependent) separatrix of the pendulum. Since for ε small and positive†, $\varepsilon^n \ll \varepsilon$, it is natural to treat the full Hamiltonian system as a perturbation of the system $(\mathcal{M}, h)$. To make this remark precise, we replace, in a first moment, ε^n with an *auxiliary parameter* μ, and establish perturbative results in μ. Such result will be analytic in μ and we shall prove that for a suitable n, ε^n is inside the μ–analyticity domain allowing to recover the original system (by setting $\mu = \varepsilon^n$).

Therefore, consider the *modified system* on $\mathcal{M}$ with Hamiltonian

$$\widetilde{H}(I, \varphi; \varepsilon, \mu) \equiv h + \mu f \ , \qquad f \equiv \cos(\varphi_1 + \varphi_3) + \cos(\varphi_2 + \varphi_3) \ ,$$

so that $\widetilde{H}(I, \varphi; \varepsilon, \varepsilon^n) = H(I, \varphi; \varepsilon)$. For $\mu = 0$, $\varepsilon > 0$ and $\omega \in \mathbb{R}^2$, consider the h–invariant set

$$\mathcal{T}(\varepsilon, \omega) \equiv \{I_1 = \omega_1 \ , \quad I_2 = \omega_2 \ , \quad I_3 = \varphi_3 = 0 \ , \quad (\varphi_1, \varphi_2) \in \mathbb{T}^2\} \ ,$$

* In Eq. (5) the perturbation is $H_1 = \varepsilon\big\{\frac{I_1^2}{2} + \frac{I_2^2}{2} + (\cos\varphi_3 - 1) + \varepsilon^{n-1}\big(\cos(\varphi_1+\varphi_3) + \cos(\varphi_2+\varphi_3)\big)\big\}$, the "integrable part" is $H_0 = \frac{I_3^2}{2}$: $H = H_0 + H_1$. As already mentioned, having a "separation" (in terms of order in ε) may be achieved near a "simple resonance" by means of symplectic transformations.

† During the proof we shall always assume $\varepsilon > 0$. The case $\varepsilon < 0$ can be easily recovered by a simple translation in the φ_3 variable.

on which the flow is quasi–periodic with frequencies $\varepsilon\omega$:

$$\phi_h^t(I_1, I_2, 0, \varphi_1, \varphi_2, 0) = (I_1, I_2, 0, \varphi_1 + \varepsilon\omega_1 t, \varphi_2 + \varepsilon\omega_2 t, 0) \ .$$

These tori admit "asymptotic manifolds", called by Arnold "whiskers", given by

$$W(\varepsilon,\omega) \equiv \{I_1 = \omega_1 \ , \quad I_2 = \omega_2 \ , \quad \frac{1}{2}I_3^2 + \varepsilon(\cos\varphi_3 - 1) = 0 \ , \quad (\varphi_1, \varphi_2) \in \mathbb{T}^2\} \ .$$

Each h–trajectory $z(t)$ starting on $W(\varepsilon,\omega)$ approaches, as $t \to \pm\infty$, the torus $\mathcal{T}(\varepsilon,\omega)$; note that $W(\varepsilon,\omega) \supset \mathcal{T}(\varepsilon,\omega)$ and that if $z(0) \in W(\varepsilon,\omega)\backslash\mathcal{T}(\varepsilon,\omega)$, then $dist\,(z(t), \mathcal{T}(\varepsilon,\omega)) = O\big(exp(-|t|\sqrt{\varepsilon})\big)$. As a set, $\mathcal{T}(\varepsilon,\omega)$ is *independent of* ε (hence, in particular, $\mathcal{T}(\varepsilon,\omega)$ is continuous in $\varepsilon \in \mathbb{R}$); what changes with ε are the frequencies $\varepsilon\omega$. In the limit case $\varepsilon = 0$, $\mathcal{T}(0,\omega)$ is still an invariant 2–torus made up by fixed points. On the other hand, in the limit case $\varepsilon = 0$, the asymptotic manifolds $W(0,\omega)$ cease to exist in the sense that the set $W(0,\omega) \equiv \{I_1 = \omega_1,\ I_2 = \omega_2,\ I_3 = 0,\ \varphi \in \mathbb{T}^3\}$ is, again, an invariant set made up by fixed points hence a trajectory starting on $W(0,\omega)\backslash\mathcal{T}(0,\omega)$ *does not* approach $\mathcal{T}(0,\omega)$.

The proof is based on three separate steps, which we proceed to describe.

(1) Let, again, $\varepsilon > 0$. Roughly speaking, the first step consists in constructing the analytic continuation in μ of the tori $\mathcal{T}(\varepsilon,\omega)$ together with their "local" whiskers, for "many" values of the "rescaled frequencies" ω; by "local" we mean "in a (ε–dependent) neighbourhood of $\mathcal{T}(\varepsilon,\omega)$". If $\varepsilon > 0$ is small enough, the μ–analyticity domain contains a complex disk of radius $\mu_0 \equiv \varepsilon^m$ (for m suitable). The set of ω's for which the continuation holds (i.e. the set of "persistent tori") is a Cantor set of "local arbitrarly high density" and the whiskers can be smoothly interpolated. A somewhat more precise formulation is given in the following:

Proposition *Given $R > 0$ and an integer $s \geq 1$, there exist $\varepsilon_0 > 0$, an integer $d \geq 2$, and an open set $B \subset B_R^2$ (the bidimensional sphere of radius R) with the following properties. For all $0 < \varepsilon < \varepsilon_0$ and for all $m > max\{2, d/s\}$ there exists a Cantor set $\Omega_\varepsilon \subset B$ such that:*

(i) $\Omega_\varepsilon \subset \Omega_{\varepsilon'}$ whenever $\varepsilon' \leq \varepsilon$; if $\omega \in \Omega_\varepsilon$ then the ray $(1+a)\omega \in \Omega_\varepsilon$ for all $0 \leq a < R - |\omega|$; the Lebesgue measure of $B\backslash\Omega_\varepsilon$ can be estimated from above as

$$meas\ \big(B\backslash\Omega_\varepsilon\big) \leq \ const\ \varepsilon^{(ms-d)/2} \ ,$$

where (here and below) "const" denote constant(s) depending only on R.

(ii) For any $\omega \in \Omega_\varepsilon$, the h–invariant torus $\mathcal{T}(\varepsilon,\omega)$ can be analytically continued in μ, for $|\mu| \leq \mu_0$, with

$$\mu_0 \equiv \varepsilon^m \ ,$$

into a $\widetilde{H}$–invariant torus $\mathcal{T}_\mu(\varepsilon,\omega)$ $[\mathcal{T}_0(\varepsilon,\omega) = \mathcal{T}(\varepsilon,\omega)]$ on which the $\widetilde{H}$–flow is conjugated to the linear flow $t \to \varepsilon\omega t$ on $\mathbb{T}^2$.

(iii) For $|\mu| \leq \mu_0$, the local whiskers of $\mathcal{T}(\varepsilon,\omega)$ admit a μ–analytic continuation into $\widetilde{H}$–whiskers $W^\pm_{\mu,loc}(\varepsilon,\omega)$ coinciding with the set of points z in a $\sqrt{\varepsilon}$–neighbourhood

of $\mathcal{T}_\mu(\varepsilon,\omega)$ *such that*** $dist(\phi^t_{\widetilde H}(z),\mathcal{T}_\mu(\varepsilon,\omega))\to 0$ *as* $t\to\pm\infty$.

(iv) The family of manifolds $W^\pm_{\mu,loc}(\varepsilon,\omega)$ *can be embedded into a family of manifolds depending in a* C^1 *way on* $\omega\in B^2_R$.

(v) For $\bar\omega\in\Omega_\varepsilon$ *and* $|\mu|\le\mu_0$, *let**

$$\Omega_\varepsilon(\bar\omega,\mu)\equiv\{\omega\in\Omega_\varepsilon:\ \widetilde H(\mathcal{T}_\mu(\varepsilon,\omega);\varepsilon,\mu)=\widetilde H(\mathcal{T}_\mu(\varepsilon,\bar\omega);\varepsilon,\mu)\}\ .$$

Then, $\Omega_\varepsilon(\bar\omega,\mu)$ *can be embedded into a set* $\Gamma_\varepsilon(\bar\omega,\mu)$, *which is made up of* C^1 *curves,* $|\mu|$*–close to the connected components of the circle* $|\omega|=|\bar\omega|$ *lying in* B. *Furthermore, gaps on any connected piece of* $\Gamma_\varepsilon(\bar\omega,\mu)$ *are smaller than* $\varepsilon^{(ms-d)/2}$.

This result is essentially known and may be obtained using standard KAM techniques (references may be found in the bibliography of Ref. [5] where KAM techniques sufficient to deduce the above result are also discussed in detail). In section §6.1 we shall give some details on the above statement. The tori $\mathcal{T}(\varepsilon)$ appearing in Definition 3 of §3 will coincide with invariant tori $\mathcal{T}_{\varepsilon^n}(\varepsilon,\omega)$ with $\omega\in\Omega_\varepsilon(\bar\omega,\varepsilon^n)$ for a suitable $n\ge m$ and for given $\bar\omega\in\Omega_\varepsilon$. In step **(3)** below, we shall see that any two points of $\Omega_\varepsilon(\bar\omega,\varepsilon^n)$, belonging to the same connected component of $B\cap\{\omega\in B^2_R: |\omega|=|\bar\omega|\}$, may be connected by a transition chain and it is clear that there are (many) points ω,ω' on $\Omega_\varepsilon(\bar\omega,\varepsilon^n)$ such that $p_1(\mathcal{T}_{\varepsilon^n}(\varepsilon,\omega))\to I$ and $p_1(\mathcal{T}_{\varepsilon^n}(\varepsilon,\omega'))\to I'$ (as $\varepsilon\to 0$) with $I\ne I'$ (but $I_3=I'_3=0$ and $|I|=|I'|=|\bar\omega|$).

(2) The second step consists in studying *transversal intersections*† of stable and unstable whiskers. First, by "pushing with the flow" $\phi^{t_0}_{\widetilde H}$, with a fixed t_0, one can prolong $W^\pm_{\mu,loc}(\varepsilon,\omega)$ to larger stretches of local whiskers (preserving, obviously, all the smoothness properties listed above). For example, taking $t_0=O(\varepsilon^{-\frac12}\log\varepsilon^{-\frac12})$, and letting

$$W'^{\pm}_{\mu,loc}(\varepsilon,\omega)\equiv\phi^{t_0}_{\widetilde H}W^\pm_{\mu,loc}(\varepsilon,\omega) \tag{6}$$

the "extended local whiskers" $W'^\pm\equiv W'^\pm_{\mu,loc}(\varepsilon,\omega)$ are such that the projection onto the φ_3 variable of $W'^\pm$ contains the interval $(-\frac32\pi,\frac32\pi)$. We shall seek transveral intersections by looking into the "Poincarè section" $\mathcal{M}\cap\{\varphi_3=\pi\}$, which is natural because of the symmetry of the model#. *From now on, dropping the "prime",* $W^\pm_{\mu,loc}(\varepsilon,\omega)$ *will denote the extended local manifold* $W'^\pm_{\mu,loc}(\varepsilon,\omega)$ *defined in Eq. (6) with the above choice of* t_0.

A general observation: The whiskers $W^\pm_\mu$ are three dimensional submanifolds and lie into a five–dimensional ambient manifold, namely, the energy manifold $\{z\in$

** W^+ is called the stable whisker and W^- the unstable whisker; note that in Arnold's paper [2] it is adopted the opposite convention, that is, Y^+ there denotes what we call here the unstable whisker.

* If $\mathcal{T}$ is invariant for H, then $H(z)=H(z')$ for all $z,z'\in\mathcal{T}$ and we denote by $H(\mathcal{T})$ the common value.

† Two submanifold M_1, M_2 of $\mathcal{M}$ intersect transversally if for any $z\in M_1\cap M_2\ne\emptyset$ the direct sum of the tangent spaces T_zM_i span $T_z\mathcal{M}$.

\# We take this opportunity to remark that having chosen a "symmetric" model is not an essential restriction but it does simplify computations!

$\mathcal{M} : \widetilde{H}(z) = \widetilde{H}(\mathcal{T}_\mu(\varepsilon,\omega))\}$. In general, two n–dimensional submanifolds M_i lying in a $(2n-1)$–dimensional manifold intersect into a one-dimensional submanifold; and if the M_i's are invariant for some Hamiltonian flow, so is their intersection, which, therefore, is a Hamiltonian trajectory. On the contrary, in the "integrable limit" $\mu = 0$, we have a degenerate intersection (as the stable and unstable whiskers of $\mathcal{T}_0(\varepsilon,\omega)$ coincide): this is a pathotology of the "integrable" regime and it is natural to expect that "non–degenerate" perturbations* will remove such a pathology and that, indeed, $W^+_{\mu,loc}(\varepsilon,\omega)$ and $W^-_{\mu,loc}(\varepsilon,\omega)$ intersect transversally into a "homoclinic" trajectory, which approach, as $t \to \pm\infty$, the underlying torus $\mathcal{T}_\mu(\varepsilon,\omega)$.

To detect homoclinic intersections we observe that for $\mu = 0$ the whiskers may be represented as *graphs over the angles* φ having the representation I_1 =const, I_2 =const, $I_3 = \pm\sqrt{2\varepsilon(1-cos\,\varphi_3)}$. The property of being graphs is preserved under perturbation (not so the periodicity in φ_3, which is a reflection of the integrability). In fact, using the KAM step (1), one can easily check that

$$W^\pm_{\mu,loc}(\varepsilon,\omega) \cap \{\varphi_3 \in \big(-\frac{3}{2}\pi, \frac{3}{2}\pi\big) \ \ mod \ \ 2\pi\} = \{(I,\varphi) : I = I^\pm(\varphi,\mu,\varepsilon,\omega)$$
$$with\ (\varphi_1,\varphi_2) \in \mathbb{T}^2\ \ , \varphi_3 \in \mathbb{R}, |\varphi_3| < \frac{3}{2}\pi,\ \ \mu \in (-\mu_0,\mu_0)\ \ , \omega \in \Omega_\varepsilon\}\ \ ,$$

where $I^\pm$ is a suitable function, which is real–analytic (for a fixed $0 < \varepsilon < \varepsilon_0$) in φ and μ (on the above domains) and continuously differentiable in $\omega \in B^2_R$ (or, more precisely, any φ and μ–derivatives of $I^\pm$ is continuously differentiable in $\omega \in B^2_R$)†. Furthermore, *the derivatives* $\partial^j_\varphi \partial^k_\omega I^\pm$, *for integer 2–vectors* j *and* k *with* $|j| \le 2$ *and* $|k| \le 1$, *are uniformly bounded by a constant (depending only on* R*) on the domain* $(\varphi_1,\varphi_2) \in \mathbb{T}^2$, $|\varphi_3| < \frac{3}{2}\pi$, $\{\mu \in \mathbb{C} : |\mu| \le \mu_0\}$, $\omega \in B^2_R$.

Dropping, for the moment, in the notation the explicit dependence on ε and ω (which we shall keep fixed $0 < \varepsilon < \varepsilon_0$, $\omega \in \Omega_\varepsilon$), we note that to look for homoclinic intersections is equivalent to study the two–by–two system

$$I_1^-(\varphi_1,\varphi_2,\pi,\mu) = I_1^+(\varphi_1,\varphi_2,\pi,\mu)\ \ ,\quad I_2^-(\varphi_1,\varphi_2,\pi,\mu) = I_2^+(\varphi_1,\varphi_2,\pi,\mu)\ . \tag{7}$$

In fact, having fixed $\varphi_3 = \pi$ is equivalent to study the intersection within the Poincarè section mentioned above, which (as it is immediate to check) cuts transversally the intersection of $W^+_{\mu,loc}(\varepsilon,\omega)$ with $W^-_{\mu,loc}(\varepsilon,\omega)$; furthermore, energy conservation [i.e. the fact that the two whiskers $W^+_\mu(\varepsilon,\omega)$ and $W^-_\mu(\varepsilon,\omega)$ lie in the same energy manifold] and Eq. (7) imply also $I_3^-(\varphi_1,\varphi_2,\pi,\mu) = I_3^+(\varphi_1,\varphi_2,\pi,\mu)$. Now, *using simple symmetry arguments, one can show that* $(\varphi_1,\varphi_2) = (0,0)$ *is, for every* $\mu \in (-\mu_0,\mu_0)$*, a solution of Eq. (7).* Transversality, in this set–up, is equivalent to

$$\delta(\mu) \equiv det\ \ \frac{\partial \widetilde{F}(\varphi_1,\varphi_2;\mu)}{\partial(\varphi_1,\varphi_2)}\Bigg|_{(\varphi_1,\varphi_2)=(0,0)} \neq 0 \tag{8}$$

* Recall that, at the moment, the perturbation parameter is μ and not ε, which is kept fixed.

$\dagger$ Obviously, for $\omega \in B^2_R \backslash \Omega_\varepsilon$ the function $I^\pm$ has no dynamical meaning (i.e. is not related, in general, to any invariant manifold).

where

$$\widetilde{F}_1 \equiv I_1^-(\varphi_1,\varphi_2,\pi,\mu) - I_1^+(\varphi_1,\varphi_2,\pi,\mu) \ , \quad \widetilde{F}_2 \equiv I_2^-(\varphi_1,\varphi_2,\pi,\mu) - I_2^+(\varphi_1,\varphi_2,\pi,\mu) \ .$$

In order to proceed one needs to *evaluate the size of* δ and, especially, *one is interested in evaluating* $\delta(\varepsilon^{m_1})$ for some $m_1 \geq m$ [recall that δ is actually a function of (μ,ε,ω) and that $\delta(\varepsilon^{m_1})$ is short for the more proper notation $\delta(\varepsilon^{m_1},\varepsilon,\omega)$]. ¿From the above analysis it follows that $\delta(\mu)$ is analytic for $|\mu| \leq \mu_0 \equiv \varepsilon^m$ and that $\delta(0)=0$ i.e. $\delta(\mu) = \delta_1\mu + \delta_2\mu^2 + \cdots$, the series being convergent for $|\mu| \leq \mu_0$. It is clear that, in principle, one could compute δ_k using the strategy outlined above, which is based on KAM techniques. In practice, however, it is much simpler to derive a *direct algorithm*, which we shall briefly describe in §6.2 below. *The upshot is that* δ_1 *is not exponentially small in* ε, contrary to what happens in Arnold's example [where $\delta_1 \sim exp(-\varepsilon^{-1/2})$], but, as we shall see in §6.2, it is**

$$\delta_1 \sim \varepsilon^{d_0}$$

for a suitable number d_0 *depending only on* h *and* f. ¿From such relation and the analyticity properties in μ, it follows easily that δ *can be exatly evaluated and is of the size of* ε^{m_2} *for a suitable* m_2, which we now compute. In fact, since $\delta(\mu)$ is analytic in a disc of radius $\mu_0 = \varepsilon^m$ and can be bounded by an ε–independent constant, one has that $|\delta_k| \leq \ const\, \varepsilon^{-km}$. Thus, if one takes

$$\mu = \varepsilon^{m_1} \ , \qquad with\ m_1 > max\{2m + d_0, m\}$$

one has, for all $k \geq 2$,

$$|\delta_k \mu^k| \leq \ const\, \varepsilon^{k(m_1-m)} \leq \ const\, \varepsilon^{2(m_1-m)} < \ const\, \varepsilon^{d_0+m_1} \leq |\delta_1\mu|$$

(as above $const$'s denote positive constants, which depend only on R). Thus

$$\delta(\varepsilon^{m_1},\varepsilon,\omega) \sim \varepsilon^{m_2} \ , \qquad with \quad m_2 \equiv d_0 + m_1 \ .$$

We finally point out that *for the eccentric D'Alembert problem* δ_1 *is indeed exponentially small in* ε, $\delta_1 \sim exp(-\varepsilon^{-1/2})$, *while it is* δ_2 *which is a power of* ε (allowing to repeat the above argument).

(3) To construct transition chains we have to study *heteroclinic connections* i.e. orbits $z(t)$ approaching, as $t \to \pm\infty$, *different* invariant tori. In other words, we have to study intersections of $W^+_{\mu,loc}(\varepsilon,\omega)$ with $W^-_{\mu,loc}(\varepsilon,\omega')$ with $\omega \neq \omega'$. To do this, we fix $0 < \varepsilon < \varepsilon_0$, $\mu \equiv \varepsilon^{m_1}$, $\bar\omega \in \Omega_\varepsilon$, and let ω and $\omega+\alpha$ belong to $\Omega_\varepsilon(\bar\omega,\mu)$. Now, define, for $i = 1,2$,

$$F_i(\varphi_1,\varphi_2,\alpha) \equiv I_i^-(\varphi_1,\varphi_2,\pi,\varepsilon^{m_1},\varepsilon,\omega+\alpha) - I_i^+(\varphi_1,\varphi_2,\pi,\varepsilon^{m_1},\varepsilon,\omega) \ ,$$

so that $F(\varphi_1,\varphi_2,0) = \widetilde{F}(\varphi_1,\varphi_2;\varepsilon^{m_1})$. Since, by our choice of ω and α, the tori $\mathcal{T}_\mu(\varepsilon,\omega)$ and $\mathcal{T}_\mu,(\varepsilon,\omega+\alpha)$ lie in the same energy manifold [which is that of $\mathcal{T}_\mu(\varepsilon,\bar\omega)$], heteroclinic

** Here $a \sim x$ means that there exist two (x–independent) constants c_1, c_2 such that $c_1 x \leq a \leq c_2 x$ for all x small enough.

connections correspond to solutions $\varphi_1(\alpha)$, $\varphi_2(\alpha)$ satisfying $F(\varphi_1(\alpha), \varphi_2(\alpha), \alpha) = 0$. By the Implicit Function Theorem* and the properties listed in **(1)** and **(2)**, it follows that one has (transversal) heteroclinic intersections whenever

$$\omega, \omega+\alpha \in \Omega_\varepsilon(\bar{\omega}, \varepsilon^{m_1}) \ , \qquad \mathit{and} \quad |\alpha| \leq \mathit{const}\, \varepsilon^{m_2} \ .$$

To insure the existence of heteroclinic connections, we have to guarantee that the gaps on $\Omega_\varepsilon(\bar{\omega}, \varepsilon^{m_1})$ are smaller that ε^{m_2}. In view of point (v) of the Proposition in step **(1)**, it is sufficient to require that $(ms-d)/2 > m_2$. All conditions on m and s (which are still free) may be collected in

$$m > max\{2, \frac{d}{s}\} \ , \qquad ms - d > 2\, max\{m + d_0, 2(m+d_0)\}$$

and it is clear that such conditions may always be met choosing suitably m and s in terms of d and d_0. The number n_0 in Theorem 3, can be taken to be equal to m_1 which , in turn, may be fixed so that

$$max\{m, 2m + d_0\} < m_1 \equiv n_0 < \frac{ms-d}{2} - d_0 \ .$$

Thus, any two tori $\mathcal{T}(\varepsilon) \equiv \mathcal{T}_{\varepsilon^{n_0}}(\varepsilon, \omega)$ and $\mathcal{T}'(\varepsilon) \equiv \mathcal{T}_{\varepsilon^{n_0}}(\varepsilon, \omega')$, with ω, ω' belonging to $\Omega_\varepsilon(\bar{\omega}, \varepsilon^{n_0}) \cap \{$ connected component of $B\}$, may be connected by a chains of heteroclinic trajectories. Moreover, the tori $\mathcal{T}(\varepsilon)$ and $\mathcal{T}'(\varepsilon)$ may be chosen in such a way that, as $\varepsilon \to 0$, $p_1(\mathcal{T}(\varepsilon)) \to I$ and $p_1(\mathcal{T}'(\varepsilon)) \to I'$ with $I \neq I'$. Finally, from the KAM linearization around the invariant tori (see §6.1 below), it follows easily that such chains are *transition chain* concluding the proof of Theorem 3.

6. On Two Technical Points

6.1. Remarks on the KAM Construction

Here we discuss briefly the KAM Proposition in **(1)** of §5.2. First of all, we define explicitly the sets B and Ω_ε. If $s = 1$, we let $B \equiv B^2_R$, otherwise:

$$B \equiv \{\omega \in B^2_R : \ |\omega \cdot n| > aR \ , \ \forall\, n \in \mathbb{Z}^2 \ , \ 0 < |n| \leq s-1\}$$
$$\Omega_\varepsilon \equiv \{\omega \in B : \ |\omega \cdot n| \geq a\, \frac{\varepsilon^{(ms-d)/2}}{|n|^2} \ , \ \forall\, n \in \mathbb{Z}^2 \ , \ n \neq 0\} \ ,$$

where $a > 0$ is a suitable constant (depending on s) small enough and " $\cdot$ " detotes scalar product in $\mathbb{R}^2$. The set B is a disk of radius R minus a finite number of strips

* We use a standard Implicit Function Theorem, which guarantees continuous solutions $\alpha \in B^2_a \to \varphi(\alpha) \in B^2_b$ of $F(\varphi, \alpha) = 0$ [here $\varphi = (\varphi_1, \varphi_2)$] provided $F(0,0) = 0$, F and $\frac{\partial F}{\partial \varphi}$ are continuous in $B^2_b \times B^2_a$, the Jacobian $\frac{\partial F}{\partial \varphi}(0,0)$ has an inverse T and

$$\sup_{|\alpha| \leq a} |F(0,\alpha)| \leq \frac{1}{2} \|T\|^{-1}\, b \ , \qquad \sup_{|\varphi| \leq b, |\alpha| \leq a} \|I - T \frac{\partial F}{\partial \varphi}\| \leq \frac{1}{2} \ .$$

crossing the origin; the constant a may be chosen so that the measure of B is close as we like to πR^2. Ω_ε is a Cantor set in B (excluding, for all ε, the dense set of "resonant" ω's for which $\omega \cdot n = 0$ for some $n \in \mathbb{Z}^2$) that becomes of full measure as ε approaches 0.

The definition of B, as we shall see below, allows to perform $(s-1)$ steps of classical (Poincaré) perturbation theory, in a neighbourhood of "resonant" tori $\mathcal{T}(\varepsilon, \omega)$ for which $\omega \in B$. The definition of Ω_ε is the classical (Kolmogorov) way of defining the frequencies "simultaneously accessible" by KAM techniques (the ε–dependent constant appearing in the definition of Ω_ε will be explained below).

Next, we introduce symplectic "hyperbolic coordinates" for the pendulum $P \equiv \frac{1}{2} I_3^2 + \varepsilon(\cos\varphi_3 - 1)$. Such a pendulum has the origin $I_3 = 0 = \varphi_3$ as unstable (hyperbolic) fixed point. The quadratic part of P, given by $\frac{1}{2}(I_3^2 - (\sqrt{\varepsilon}\varphi_3)^2)$, is transformed, via the symplectic map

$$\begin{pmatrix} I_3 \\ \varphi_3 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \varepsilon^{1/4} & \varepsilon^{1/4} \\ -\frac{1}{\varepsilon^{1/4}} & \frac{1}{\varepsilon^{1/4}} \end{pmatrix} \begin{pmatrix} y \\ x \end{pmatrix} \tag{9}$$

into the "normal hyperbolic" form $\sqrt{\varepsilon} yx$. It is therefore clear that, near the origin, there exists a symplectic (real–analyitc) transformation S, whose linear part is given by Eq. (9), transforming the pendulum P into the normal form

$$P \circ S = g(yx, \sqrt{\varepsilon}) = \sqrt{\varepsilon}(yx) + O\big((yx)^2\big) \ .$$

The domain of definition of S (and hence of g) can be taken to be a 2–sphere centered at the origin of radius, say, $4b$ for a suitable $b \sim \varepsilon^{1/4}$. In the (I_3, φ_3)–coordinates, the sphere B^2_{4b} correspond, in a first approximation, to an ellipse with I_3–semiaxis of length $\sim \sqrt{\varepsilon}$ and with φ_3–semiaxis of length ~ 1. In the symplectic coordinates

$$\bar{I} \equiv (I_1, I_2) \ , \quad \bar{\varphi} \equiv (\varphi_1, \varphi_2) \ , \quad (y, x) \ ,$$

the Hamiltonian $\widetilde{H}$ takes the form

$$H_1(\bar{I}, \bar{\varphi}, y, x, \varepsilon, \mu) \equiv h_1 + \mu f_1 \equiv \frac{\varepsilon}{2} \bar{I}^2 + g(yx, \sqrt{\varepsilon}) + \mu f_1(\bar{\varphi}, y, x, \sqrt{\varepsilon}) \ ,$$

where f_1 is a linear combination of $\cos\varphi_i$ and $\sin\varphi_i$, with $i = 1, 2$, with coefficients which are real–analytic functions of $\xi = yx$. By classical perturbation theory, one can find a symplectic transformation μ–close to the identity, conjugating H_1 to a Hamiltonian

$$\begin{aligned} H_2(\bar{I}', \bar{\varphi}', y', x', \varepsilon, \mu) &\equiv h_2 + \mu^s f_2 \\ &\equiv h_1(\bar{I}', y'x', \sqrt{\varepsilon}) + \mu g_1(y'x', \sqrt{\varepsilon}, \mu) + \mu^s f_2(\bar{I}', \bar{\varphi}', y', x', \sqrt{\varepsilon}, \mu) \ . \end{aligned}$$

The "small divisors" appearing in such a symplectic transformation†, have the form $i\varepsilon \bar{I}' \cdot k + j\sqrt{\varepsilon} y'x'$ with $k \in \mathbb{Z}^2$, $j \in \mathbb{N}$, $|k| + j > 0$ and $|k| \le s-1$. Thus the domain of

† The transformation can be found, e.g., by solving up to order μ^{s-1} the Hamilton–Jacobi equation for a generating function of the form $\bar{I}' \cdot \bar{\varphi} + y'x +$ polynomial in μ of order $(s-1)$ vanishing at $\mu = 0$.

definition (and analyticity) of H_2 is given, for $0<\varepsilon<\varepsilon_0$, by $B\times \mathbb{T}^2\times B^2_{2b}\times B^1_{\bar\mu}$ and $\bar\mu$ may be choosen proportional to ε^2. In these coordinates, the tori $\mathcal{T}(\varepsilon,\omega)$ are given by $\bar I'=\omega$, $\bar\varphi'\in\mathbb{T}^2$, $y'=x'=0$ and the stable/unstable local whiskers are given by, respectively, $\bar I'=\omega$, $\bar\varphi'\in\mathbb{T}^2$, $|y'|\le 2b$, $x'=0$, and $\bar I'=\omega$, $\bar\varphi'\in\mathbb{T}^2$, $|x'|\le 2b$, $y'=0$.

At this point one can use a KAM "fast–converging" scheme to construct, simultaneously for $\omega\in\Omega_\varepsilon$, invariant tori with flow conjugated to $t\to\varepsilon\omega t$, and their local whiskers, which will be μ^s close to the local whiskers for h_2. As it is well known, the smallness requirement (or "KAM condition") necessary in order to carry out the scheme, is of the form

$$const\ \frac{\mu_0^s}{\varepsilon^d\gamma^2}<1\ , \tag{10}$$

where the analyticity domain in μ will contain the disk $|\mu|\le\mu_0$ and γ is the "rescaled Diophantine constant" associated to the persistent tori $\mathcal{T}(\varepsilon,\mu)$, i.e., ω is such that

$$|\omega\cdot n|\ge\frac{\gamma}{|n|^2}\ ,\qquad \forall\ n\neq 0\ , \tag{11}$$

[hence the small divisors appearing in the KAM scheme satisfy $|\varepsilon\omega\cdot n|\ge(\varepsilon\gamma)/|n|^2$]; d is a constant (depending only on the structure of h) which is greater or equal than two since, as already mentioned, the actual ("non–rescaled") Diophantine constant is $\varepsilon\gamma$ and such a constant appears quadratically in KAM conditions. All the frequencies ω satisfying the classical condition Eq. (11), may be taken care of simultaneously and the complementary set $\{\omega\in B:$ Eq. (11) does not hold $\}$ has a measure of the order of γ. Thus, choosing $\mu_0=\varepsilon^m$, with $m>max\{2,d/s\}$, we see that γ may be taken to be [in order to meet the KAM condition Eq. (10)] $\gamma=\ const\,\varepsilon^{(ms-d)/2}$.

The conclusion is that, with the above choices, one can construct C^3 functions ζ and λ,

$$\zeta:(\theta,(y,x),\omega,\mu)\in\mathbb{T}^2\times B^2_b\times B^2_R\times B^1_{\mu_0}\to\zeta(\theta,y,x,\omega,\mu)\in\mathcal{M}\ ,$$
$$\lambda:(\xi,\omega,\mu)\in B^1_b\times B^2_R\times B^1_{\mu_0}\to\lambda(\xi,\omega,\mu)\in\mathbb{R}\ ,$$

such that the following holds. For $\omega\in\Omega_\varepsilon$, ζ and λ are real–analytic in all their remaining variables; $\lambda(0,\omega,0)=\sqrt{\varepsilon}$; $\zeta(\theta,0,0,\omega,0)=(\omega,0,\theta,0)$;

$$\sup_{\mathbb{T}^2\times B^2_b\times B^2_R\times B^1_{\mu_0}}|\partial^j_\theta\partial^k_\omega\zeta|\le\ const\ ,\qquad\forall\ |j|\le 2\ ,\ |k|\le 1\ ;$$

for $\omega\in\Omega_\varepsilon$ and $|\mu|\le\mu_0$ one has

$$\mathcal{T}_\mu(\varepsilon,\mu)=\{z=\zeta(\theta,0,0,\omega,\mu):\ \theta\in\mathbb{T}^2\}\ ,$$
$$W^+_{\mu,loc}(\varepsilon,\omega)=\{z=\zeta(\theta,y,0,\omega,\mu);\ \theta\in\mathbb{T}^2\ ,|y|<b\}\ ,$$
$$W^-_{\mu,loc}(\varepsilon,\omega)=\{z=\zeta(\theta,0,x,\omega,\mu);\ \theta\in\mathbb{T}^2\ ,|x|<b\}\ .$$

Furthermore,

$$\phi^t_{\widetilde H}\zeta(\theta,y,x,\omega,\mu)=\zeta(\theta+\varepsilon\omega t,ye^{-\lambda(yx,\omega,\mu)t},xe^{\lambda(yx,\omega,\mu)t},\omega,\mu) \tag{12}$$

for all $\theta\in\mathbb{T}^2$ and all $(y,x)\in B^2_b$ such that $|ye^{-\lambda t}|^2+|xe^{\lambda t}|^2<b^2$.

This last "linearization property" implies that the tori $\mathcal{T}_\mu(\varepsilon,\omega)$ are transition tori in the sense of Arnold [2] [i.e. given any two points $z_\pm \in W^\pm_{\mu,loc}$ and given any neighbourhood U of z_+ the set $\{\phi^t U : t > 0\}$ crosses any manifold which intersects transversally $W^-_{\mu,loc}$ in z_-].

As for the final point (v) of the KAM Proposition, observe that from the definition of the set Ω_ε, it follows that the ray $(1+a)\omega$, with $0 \le a < R - |\omega|$, belongs to Ω_ε, for any $\omega \in \Omega_\varepsilon$. Also, the set $\Omega_\varepsilon(\bar\omega,\mu)$ coincides with

$$\Omega_\varepsilon(\bar\omega,\mu) \equiv \{\omega \in \Omega_\varepsilon :\ \widetilde{H}\big(\zeta(0,0,0,\omega,\mu)\big) = \widetilde{H}\big(\zeta(0,0,0,\bar\omega,\mu)\big)\}$$

and the set

$$\Gamma_\mu = \{\omega \in B :\ \widetilde{H}\big(\zeta(0,0,0,\omega,\mu)\big) = \widetilde{H}\big(\zeta(0,0,0,\bar\omega,\mu)\big)\}$$

is a finite union of C^1 curves, which are $|\mu|$–close [by the smoothness of ζ in $\omega \in B^2_R$ and the fact that $p_1\big(\zeta(0,0,0,\bar\omega,0)\big) = \bar\omega$] to $\{|\omega| = |\bar\omega|\}$ intersected with B. Thus, since Ω_ε has a measure not smaller than, approximately, $R^2\big(1-\varepsilon^{(ms-d)/2}\big)$, it is clear that, in general, $\Omega_\varepsilon \cap \Gamma_\varepsilon(\bar\omega,\varepsilon)$ cuts on $\Gamma_\varepsilon(\bar\omega,\varepsilon)$ a region of length not smaller than $|\bar\omega|\big(1-\varepsilon^{(ms-d)/2}\big)$.

6.2. Generalized Poincaré Integrals

Here we discuss with some detail *a recursive algorithm that allows to compute all the μ–derivatives of the function $\delta(\mu)$* introduced in Eq. (8). In particular, we shall prove that, for the Hamiltonian in Eq. (5), it is

$$\delta_1 = \frac{4}{\varepsilon}\big(1+O(\varepsilon)\big)\ , \qquad i.e.\quad d_0 = -1\ . \tag{13}$$

Actually, to prove Eq. (13) is rather easy, but we shall spend some time in order to discuss, in some generality, the above mentioned algorithm, which, we believe, is quite interesting in itself.

As already mentioned in the previous section, the derivatives of δ could be computed by using the KAM results, however we prefer, following seminal ideas of Poincaré**, a more direct approach.

Keeping fixed $0 < \varepsilon < \varepsilon_0$ and $\omega \in \Omega_\varepsilon$ [and using the above conventions: $\bar I \equiv (I_1, I_2)$, $\bar\varphi \equiv (\varphi_1,\varphi_2)$], we denote $\widetilde{H}$–trajectories starting at $W^\pm_\mu \cap \{\varphi_3 = \pi\}$ by

$$\begin{aligned} z^\pm(t;\bar\varphi,\mu) &\equiv \big(I^\pm(t;\bar\varphi,\mu),\varphi^\pm(t;\bar\varphi,\mu)\big) \equiv \big(\bar I^\pm(t;\bar\varphi,\mu),p^\pm(t;\bar\varphi,\mu),\bar\varphi^\pm(t;\bar\varphi,\mu),q^\pm(t;\bar\varphi,\mu)\big)\\ &\equiv \phi^t_{\widetilde{H}}\big(I^\pm(\bar\varphi,\pi,\mu),\bar\varphi,\mu\big) \end{aligned}$$

where $I^\pm(\varphi,\mu)$ are the functions introduced in **(2)** of §5.2 (in the first line of this definition, we have introduced different symbols for the same object, which is then defined in the second line). In particular, $I^\pm(0;\bar\varphi,\mu) = I^\pm(\bar\varphi,\pi,\mu)$, $\varphi^\pm(0;\bar\varphi,\mu) = (\bar\varphi,\pi)$. An important remark, based on the KAM analysis, is that, since trajectories on $W^\pm_\mu$ approach (exponentially fast) the invariant tori $\mathcal{T}_\mu(\varepsilon,\omega)$ on which the flow is conjugated to quasi–periodic, *the functions* $I^\pm(t)$, $\varphi^\pm(t)$ are "asymptotically (as*

** See also Ref. [2] and, for generalities and further references, Ref. [1].

* If not needed explicitly, we shall drop the dependence on $\bar\varphi$ and μ.

$t \to \pm\infty$) *quasi–periodic functions" with frequencies* $\varepsilon\omega$. Here by "f^σ asymptotically quasi–periodic", we mean that $f^\sigma = \psi + g$ with ψ quasi–periodic with frequencies $\varepsilon\omega$ and g converging to 0 *exponentially fast*, as $t \to \sigma\infty$.

The function $z^\pm(t) \equiv \big(I^\pm(t), \varphi^\pm(t)\big)$ is analytic in $|\mu| \le \mu_0$ and has a convergent power–series representation

$$I^\pm(t;\bar\varphi,\mu) = \sum_{k=0}^\infty I^{k\pm}(t,\bar\varphi)\mu^k \ , \qquad \varphi^\pm(t;\bar\varphi,\mu) = \sum_{k=0}^\infty \varphi^{k\pm}(t,\bar\varphi)\mu^k \ ,$$

where $\Big(I^{0\pm}(t), \varphi^{0\pm}(t)\Big) \equiv z^0(t)$ is the unperturbed h–trajectory given by

$$z^0(t) \equiv \big(\omega, p^0(t), \bar\varphi + \varepsilon\omega t, q^0(t)\big) \ ,$$

(p^0, q^0) being the separatrix motion of the pendulum $\frac{1}{2}I_3^2 + \varepsilon(\cos\varphi_3 - 1)$ starting at $\varphi_3 = \pi$:

$$q^0(t) = 4\,arctan\, e^{-\sqrt{\varepsilon}t} \ , \quad p^0(t) = \frac{2\sqrt{\varepsilon}}{cosh\sqrt{\varepsilon}t} \ , \quad cos\, q^0 = 1 - \frac{2}{(cos\sqrt{\varepsilon}t)^2} \ , \quad sin\, q^0 = 2\frac{sinh\sqrt{\varepsilon}t}{(cos\sqrt{\varepsilon}t)^2} \ .$$

(and, for definitess, we have choosen the upper branch, $I_3 > 0$).

Letting $(\bar I^{k\pm}, p^{k\pm}) \equiv I^{k\pm}$, $(\bar\varphi^{k\pm}, q^{k\pm}) \equiv \varphi^{k\pm}$ and inserting the power series into the $\widetilde H$–Hamilton equations and equating the coefficients of μ^k we get, for $k \ge 1$, the hierarchy of equations

$$\begin{aligned} \dot{\bar I}^{k\pm} &= \bar\Phi^{k\pm} \ , & \dot{\bar\varphi}^{k\pm} &= \varepsilon \bar I^{k\pm} \ , \\ \dot p^{k\pm} &= (\varepsilon\cos q^0)q^{k\pm} + \Phi_3^{k\pm} \ , & \dot q^{k\pm} &= p^{k\pm} \end{aligned} \tag{14}$$

where the functions $\Phi^{k\pm}(t;\bar\varphi) \equiv (\Phi_1^{k\pm}, \Phi_2^{k\pm}, \Phi_3^{k\pm}) \equiv (\bar\Phi^{k\pm}, \Phi_3^{k\pm})$ can be easy computed by Taylor expansion and depend only on $z^{j\pm}$ with $0 \le j \le k-1$. For example, $\bar\Phi^{1\pm}$, $\Phi_3^{1\pm}$, which are actually independent of the + or − sign, are given by

$$\begin{aligned} \bar\Phi^1(t,\bar\varphi) &\equiv -\frac{\partial f}{\partial\bar\varphi}\big(\bar\varphi + \varepsilon\omega t, q^0(t)\big) \equiv \Big(sin\,\big(\varphi_1 + \varepsilon\omega_1 t + q^0(t)\big), sin\,\big(\varphi_2 + \varepsilon\omega_2 t + q^0(t)\big)\Big) \ , \\ \Phi_3^1(t,\bar\varphi) &\equiv -\frac{\partial f}{\partial\varphi_3}\big(\bar\varphi + \varepsilon\omega t, q^0(t)\big) \equiv sin\,\big(\varphi_1 + \varepsilon\omega_1 t + q^0(t)\big) + sin\,\big(\varphi_2 + \varepsilon\omega_2 t + q^0(t)\big) \ . \end{aligned}$$

Observing that the fundamental solution (or "Wronskian") of the linearized pendulum equation $\dot p = (\varepsilon\cos q^0)q$, $\dot q = p$, is given by

$$W(t) = \begin{pmatrix} (1 - \frac{w(t)}{4}\frac{sinh\sqrt{\varepsilon}t}{cosh^2\sqrt{\varepsilon}t})\,cosh\sqrt{\varepsilon}t & -\sqrt{\varepsilon}\frac{sinh\sqrt{\varepsilon}t}{cosh^2\sqrt{\varepsilon}t} \\ \frac{w(t)}{4\sqrt{\varepsilon}} & \frac{1}{cosh\sqrt{\varepsilon}t} \end{pmatrix}, \qquad w(t) \equiv \frac{2\sqrt{\varepsilon}t + sinh\,2\sqrt{\varepsilon}t}{cosh\sqrt{\varepsilon}t} \ ,$$

integrating Eq. (14), one can express $z^{k\pm}(t)$ in terms of $z^0(t)$,...,$z^{(k-1)\pm}(t)$:

$$\begin{aligned} \bar I^{k\pm}(t) &= \bar I^{k\pm}(0) + \int_0^t \bar\Phi^{k\pm}(\tau)d\tau \ , \quad \bar\varphi^{k\pm}(t) = \varepsilon\bar I^{k\pm}(0)t + \varepsilon\int_0^t (t-\tau)\bar\Phi^{k\pm}(\tau)d\tau \ , \\ p^{k\pm}(t) &= w_{11}(t)q^{k\pm}(0) + w_{11}(t)\int_0^t w_{22}(\tau)\Phi_3^{k\pm}(\tau)d\tau - w_{12}(t)\int_0^t w_{21}(\tau)\Phi_3^{k\pm}(\tau)\,d\tau \ , \\ q^{k\pm}(t) &= w_{21}(t)q^{k\pm}(0) + w_{21}(t)\int_0^t w_{22}(\tau)\Phi_3^{k\pm}(\tau)d\tau - w_{22}(t)\int_0^t w_{21}(\tau)\Phi_3^{k\pm}(\tau)\,d\tau \ , \end{aligned} \tag{15}$$

where w_{ij} are the entries of $W(t)$ and we have used the fact that, for $k \geq 1$, $\bar{\varphi}^{k\pm}(0) = 0 = q^{k\pm}(0)$.

Now, *the key idea is to integrate from* 0 *to* $\pm\infty$ *and, using the fact that* $z^{k\pm}(t)$ *is aymptotically (as* $t \to \infty$*) a quasi–periodic function, to try to identify the "initial value"* $I^{k\pm}(0) \equiv \frac{1}{k!} \frac{d}{d\mu}\Big|_{\mu=0} I^{\pm}(\bar{\varphi}, \pi, \mu)$ *in terms of integrals of the* Φ*'s functions.*

In order to perform with ease improper integrations, we introduce the following *generalization of standard improper integration*†.

Let σ denote the $+$ or $-$ sign. Let $\mathcal{H}_\sigma$ be the class of measurable functions $f : \mathbb{R} \to \mathbb{C}$ such that there exist $a > 0 > b$ for which, given $t \in \mathbb{R}$, the function

$$u \to F(u,t) \equiv \int_{\sigma\infty}^{t} e^{-u\sigma\tau} f(\tau) d\tau$$

is analytic on the complex domain $\{u \in \mathbb{C} : Re\ u > a\}$ and admits an analytic continuation to $\{u \in \mathbb{C} : Re\ u < b\}$. If $f \in \mathcal{H}_\sigma$, we set

$$\mathcal{J}_\sigma^t f \equiv F(0,t) \ .$$

Note that if f is, say, piecewise continuous and $lim\,sup_{t\to\sigma\infty} e^{r\sigma t}|f(t)| < \infty$, for some $r > 0$, then $f \in \mathcal{H}_\sigma$ and $\mathcal{J}_\sigma^t f = \int_{\sigma\infty}^t f$. It is also easy to check that $f \equiv t^j e^{ct} \in \mathcal{H}_\sigma$ for any j, σ and any *non vanishing* complex number c. Instead, the polynomials are not contained in $\mathcal{H}_\sigma$. Nevertheless, we extend, by linearity, the operator $\mathcal{J}$ on $\widetilde{\mathcal{H}}_\sigma \equiv \mathcal{H}_\sigma \oplus$ ring of polynomials in t, by defining

$$\mathcal{J}_\sigma^t \tau^j = \frac{t^{j+1}}{(j+1)} \ .$$

Such an extension implies the following useful property of $\mathcal{J}$:

$$\mathcal{J}_\sigma^t f - \mathcal{J}_\sigma^s f = \int_s^t f(\tau) d\tau \ , \tag{16}$$

for any $f \in \widetilde{\mathcal{H}}_\sigma$ and any s, t.

An easy recursive argument (either based on the KAM analysis of §6.1 or direct) shows that $z^{k\pm}$*, and hence* $\Phi^{k\pm}$*, belong to* $\mathcal{H}_\pm$ *for all* $k \geq 1$. ¿From Eq. (15) and Eq. (16) it then follows

$$\bar{I}^{k\pm}(t) - \mathcal{J}_\pm^t \bar{\Phi}^{k\pm} = \bar{I}^{k\pm}(0) - \mathcal{J}_\pm^0 \bar{\Phi}^{k\pm} \ . \tag{17}$$

We, now, note the following two facts (which are easily checked): (i) The quasi–periodic average, $lim_{T\to\infty} (\sigma T)^{-1} \int_0^T f$, of a function $f \equiv \psi + g$, which is asymptotically (as $t \to \sigma\infty$) quasi–periodic, coincides with the quasi–periodic average of the limiting function ψ, and we shall denote such a number by $\langle f \rangle$; (ii) if f is asymptotically quasi–periodic with $\langle f \rangle = 0$, then ($f \in \mathcal{H}_\sigma$ and) so is $\mathcal{J}_\sigma^t f$. Thus, taking the quasi–periodic average in the first line of Eq. (14), one sees that both $\bar{\Phi}^{k\pm}$ and $\bar{I}^{k\pm}$,

† The operations discussed here are introduced in Ref. [5].

which are asymptotically quasi–periodic, have vanishing quasi–periodic average. Therefore, taking the quasi–periodic average in Eq. (17), we see that

$$\bar{I}^{\pm}(0) = \mathcal{J}^0_{\pm}\bar{\Phi}^{k\pm} \ , \qquad \textit{hence} \quad \bar{I}^{\pm}(t) = \mathcal{J}^t_{\pm}\bar{\Phi}^{k\pm} \ , \qquad \forall\ t\ .$$

With similar arguments one finds also that

$$\begin{aligned} p^{k\sigma}(t) &= w_{11}(t)\int_{\sigma\infty}^{t} w_{22}(\tau)\Phi_3^{k\sigma}(\tau)d\tau - w_{12}(t)\int_0^t w_{21}(\tau)\Phi_3^{k\sigma}(\tau)d\tau\ , \\ q^{k\sigma}(t) &= w_{21}(t)\int_{\sigma\infty}^{t} w_{22}(\tau)\Phi_3^{k\sigma}(\tau)d\tau - w_{22}(t)\int_0^t w_{21}(\tau)\Phi_3^{k\sigma}(\tau)d\tau\ . \end{aligned}$$

Let us now go back to the computation of δ_1 and let us introduce the following notation: if $f^{\pm}$ denotes *two* functions $f^+ \in \mathcal{H}_+$ and $f^- \in \mathcal{H}_-$, we let

$$\mathcal{J} f^{\pm} \equiv \mathcal{J}^0_- f^- - \mathcal{J}^0_+ f^+\ .$$

Thus, recalling the definition of $\bar{\Phi}^1$, we have

$$\begin{aligned} \bar{I}^{1-}(\bar{\varphi}) - \bar{I}^{1+}(\bar{\varphi}) &\equiv \frac{d}{d\mu}\Big|_{\mu=0} \left(\bar{I}^{1-}(\bar{\varphi},\pi,\mu) - \bar{I}^{1+}(\bar{\varphi},\pi,\mu)\right) \\ &= \Big(\mathcal{J}\, sin\left(\varphi_1 + \varepsilon\omega_1 t + q^0(t)\right), \mathcal{J}\, sin\left(\varphi_2 + \varepsilon\omega_2 t + q^0(t)\right)\Big)\ . \end{aligned}$$

A straightforward computation, based upon the definition of $\mathcal{J}$, symmetries** and residue theory, yields

$$\begin{aligned} \mathcal{J}\, sin(\varphi_i + \varepsilon\omega_i t + q^0) &= sin\,\varphi_i \mathcal{J}\Big(cos\,\varepsilon\omega_i t\, cos\, q^0 - sin\,\varepsilon\omega_i t\, sin\, q^0\Big) \\ &= sin\,\varphi_i \int_{-\infty}^{\infty} \{cos\,\varepsilon\omega_i t(cos\,q^0 - 1)\} - sin\,\varphi_i\int_{-\infty}^{\infty} sin\,\varepsilon\omega_i t\, sin\, q^0 \\ &= -2\, sin\,\varphi_i \int_{-\infty}^{\infty}\Big\{\frac{cos\,\varepsilon\omega_i t}{(cosh\sqrt{\varepsilon}t)^2} + \frac{sin\,\varepsilon\omega_i t\; senh\sqrt{\varepsilon}t}{(cosh\sqrt{\varepsilon}t)^2}\Big\} \\ &= -\pi\omega_i\, sin\,\varphi_i\Big(\frac{1}{sinh\,\frac{\pi\sqrt{\varepsilon}\omega_i}{2}} + \frac{1}{cosh\,\frac{\pi\sqrt{\varepsilon}\omega_i}{2}}\Big) \end{aligned}$$

from which Eq. (13) follows at once.

7. References

1. V.I. Arnold (Ed.) , *Encyclopaedia of Mathematical Sciences* **III** *"Dynamical Systems"* (Springer–Verlag, 1988) .
2. V. I. Arnold, *Instability of Dynamical Systems with Several Degrees of Freedom*, in *Sov. Math. Dokl.* **6** (1964) p. 581–585.
3. H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Céleste* Vol. 1–3 (Paris, Gauthier–Villars, 1892, 1893, 1899).
4. Z. Xia, *Arnold Diffusion in the Elliptic Restricted Three–Body Problem*, in *J. Dyn. Differ. Equations* **5**, No. 2, (1993), p. 219–240.
5. L. Chierchia and G. Gallavotti, *Drift and Diffusion in Phase Space*, in *Ann. H. Poincarè* (Physique Theorique), **60**, No. 1 (1994), p. 1–144.

** If $f \in \mathcal{H}_\sigma$ is an even/odd function of t, then $\mathcal{J}^t_\sigma f$ is odd/even.