



0.4

ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Stability and diffusion in Hamiltonian systems via analytical and variational perturbative methods

CANDIDATE

Luca Biasco

SUPERVISORS

Prof. Antonio Ambrosetti

Prof. Massimiliano Berti

Thesis submitted for the degree of *Doctor Philosophiae*

Academic Year 2001-2002

**SISSA - SCUOLA
INTERNAZIONALE
SUPERIORE
STUDI AVANZATI**

TRIESTE
Via Beirut 2-4

TRIESTE

Ringraziamenti

Desidero ringraziare innanzitutto i miei relatori Antonio Ambrosetti e Massimiliano Berti; il primo per la grande esperienza e conoscenza matematica che ha messo a mia disposizione e i preziosi consigli, il secondo per le tante cose che ho imparato da lui (e non solo di matematica), ma soprattutto per la sua amicizia.

Ringrazio inoltre Philippe Bolle per aver lavorato insieme a me e a Massimiliano con profitto matematico e divertimento.

Un ringraziamento speciale va all'amico Luigi Chierchia per l'affetto con cui mi ha seguito in tutti questi anni di matematica e senza il quale di certo tutto questo non sarebbe stato possibile.

Ancora un grazie a tutti miei colleghi della SISSA e in particolare ai miei compagni di ufficio Veronica, Beppolone e Mario.

Infine ricordo tutti gli amici che mi hanno spronato nelle difficoltà e la mia famiglia a cui questa tesi è dedicata.

Contents

Introduction	3
I Preliminaries	23
1 Notations, norms and Normal Form Lemma	25
2 Complex action-angle variables	26
3 The D'Alembert Hamiltonian planetary model	41
3.1 Hamiltonian formalism for the D'Alembert planetary model	42
3.2 Linear analysis and the effective Hamiltonian	46
II Total stability: a model problem for properly-degenerate Hamiltonian Systems with two degrees of freedom	51
4 The total stability Theorems	53
5 Technical lemmas	56
6 Proofs of the Theorems	61
III Exponential stability: the resonant D'Alembert model of Celestial Mechanics	71
7 The exponential stability Theorem	73
8 Proof of Theorem in the case $(p, q) \neq (1, 1), (2, 1)$	78
8.1 Step 1 (linear change of variables)	78
8.2 Step 2 (time averaging)	80
8.3 Step 3 (averaging over $\hat{\varphi}_1$)	81
8.4 Step 4 (energy conservation)	82
9 Proof of Theorem in the case $(p, q) = (1, 1), (2, 1)$	85
IV Topological instability: a new variational mechanism with optimal diffusion time	91
10 Statement of the main results	93
11 The variational setting and the finite dimensional reduction	95

12 The approximation of the reduced functional	98
13 Ergodization times	105
14 The unperturbed pseudo-diffusion orbit	107
15 The diffusion orbit	112
16 The stability result and the optimal time	117
 V Appendix	 127
A Proof of lemma 11.1.	129
B Proof of Theorem 13.2.	133

Introduction

As pointed out with particular emphasis by H. Poincaré (see [92]), one of the main problem in Dynamical Systems concerns *the stability of action variables in nearly-integrable Hamiltonian systems*. Actually the study of this problem has nowadays great interest in many fields of mathematics, such as PDE's, Ergodic Theory and Differential Geometry, etc...

Roughly speaking, the problem is whether a very small perturbation of an “integrable” (i.e., completely stable) Hamiltonian system can give rise to an “appreciable” instability of the action variables (which, in general, measures some physically interesting quantities).

Originally this question arose in famous problems of Celestial Mechanics such as the *n-body problem* or the *D'Alembert planetary model*.

The *n-body problem* deals, as it is well known, with the motion of $n - 1$ planets around a fixed star which can be regarded as a perturbation of an integrable system of $n - 1$ non-interacting point masses running on Keplerian ellipses around a fixed center of Newtonian attraction. In this setting the action variables represent the lengths of the semi-axes of the above-mentioned ellipses.

The *D'Alembert planetary model* is a Hamiltonian model for a rotational oblate planet, revolving periodically on a given Keplerian ellipse of small eccentricity around a fixed star occupying one of the foci of the ellipse; the planet is subject only to the gravitational attraction of the star. This model can be also regarded as a perturbation of the integrable system obtained considering the case in which the planet is completely spherical. In this contest, the action variable represents the inclination of the planet polar axis with respect to the plane containing the Keplerian ellipse.

We observe that, with regard to the two mentioned problems, from a physical point of view, an “appreciable” variation, for example, of the distance Earth–Sun or of the inclination of the Earth polar axis would produce sensible (and probably inauspicious) effects on the Life on the Earth (weather, length of the seasons, glaciations...).

It has to be noted that, notwithstanding the efforts of Poincaré himself (followed by Birkhoff, Kolmogorov, Siegel, Arnold, Moser, Herman,...), and the great success of powerful, more modern techniques in Perturbation Theory (such as averaging theory, KAM and Nekhoroshev theory, see [8] for general information), *the action-stability problem for general nearly-integrable Hamiltonian systems remains essentially open*.

Let us be more precise. Consider a (real-analytic) nearly-integrable, one-parameter family of Hamiltonian functions

$$H(I, \varphi; \varepsilon) = h(I) + \varepsilon f(I, \varphi; \varepsilon) \quad (0.1)$$

where (I, φ) are standard symplectic variables in a $2d$ -dimensional symplectic manifold¹ \mathcal{M} , called “phase space” and $\varepsilon \geq 0$ is a small perturbative parameter. We will take \mathcal{M} of the form $\mathcal{M} := \Omega \times \mathbb{T}^d$, with Ω an open set in \mathbb{R}^d and $\mathbb{T}^d := \mathbb{R}^d / 2\pi\mathbb{Z}$; (I, φ) are called *action-angle variables*.

The *action-stability problem* is, then, to give upper bounds on the quantity $|I(t) - I_0|$, where $(I(t), \varphi(t))$ denotes the H -flow at time t of the initial datum (I_0, φ_0) , and “total stability” means that $\sup_{t \in \mathbb{R}} |I(t) - I_0|$ goes to zero when ε goes to zero, for any $(I_0, \varphi_0) \in \mathcal{M}$.

The KAM Theorem

If $\varepsilon = 0$ the manifold \mathcal{M} is foliated by d -dimensional invariant tori $\mathcal{T}_{I_0} := \{I \equiv I_0, \varphi \in \mathbb{T}^d\}$. The motion on each torus is quasi-periodic with frequency $\omega(I_0) := h'(I_0) \in \mathbb{R}^d$. A torus \mathcal{T}_{I_0} is called *non-resonant* if $\omega(I_0)$ is rationally independent. Each phase trajectory on such a torus fills it densely. The set of non-resonant tori form a set of full measure.

The celebrated KAM Theorem (see [8] for general information), in the fix energy version, may be formulated as it follows: *if h is isoenergetically non-degenerate, then for ε sufficiently small, most non-resonant invariant tori do not disappear but are only slightly deformed, so that in the phase space of the perturbed system there are invariant tori densely filled with quasi-periodic phase curves winding around them, with a number of frequencies equal to the number of degrees of freedom d . These invariant tori form a majority on each manifold of the energy.*

We recall that h is called *isoenergetically non-degenerate* on the energy level $h \equiv E$ if

$$\det \begin{pmatrix} \partial_I^2 h(I) & \partial_I h(I) \\ \partial_I h(I) & 0 \end{pmatrix} \neq 0 \quad (0.2)$$

for $I \in \{h \equiv E\}$.

Thus, if condition (0.2) is satisfied, KAM theory yields “metric” stability (i.e., stability for the majority of initial data) and this implies “total stability” (i.e. for all initial

¹ A manifold \mathcal{M} is called *symplectic* if it is endowed with a *symplectic structure* i.e. a closed alternate and non-degenerate 2-form ω ; in suitable local coordinates $(I, \varphi) \in \mathcal{M}$, ω takes the form $\omega := dI \wedge d\varphi := \sum_i dI_i \wedge d\varphi_i$. Variables for which $\omega = dI \wedge d\varphi := \sum_i dI_i \wedge d\varphi_i$ are called *standard*. In the following we will use standard variables only. A change of variable which preserves ω is called *symplectic*. Given $(I_0, \varphi_0) \in \mathcal{M}$, we call $(I(t), \varphi(t)) := \Phi_H^t(I_0, \varphi_0)$, the Hamiltonian flow generated by the Hamiltonian H i.e. the solution of the Hamilton equations:

$$\frac{dI}{dt}(t) = -\partial_\varphi H(I(t), \varphi(t)) , \quad \frac{d\varphi}{dt}(t) = \partial_I H(I(t), \varphi(t)) , \quad I(0) = I_0 , \quad \varphi(0) = \varphi_0 .$$

data and all times) in systems with two degrees of freedom: in such a case, the three-dimensional energy surfaces are separated by a multitude of two-dimensional invariant KAM tori and trajectories are trapped in-between these tori allowing only for a small (with ε) variation of the action variables (for any time and for any initial data).

Remark 0.1 *If condition (0.2) does not hold, then “linear drift” occurring along so-called “superconductivity channels” may appear as a consequence of some “resonance phenomena”. Consider in fact the following example due to N.N.Nekoroshev (see [90] and remark 4.1 (i)): the Hamiltonian system governed by $H := I_1^2/2 - I_2^2/2 + \varepsilon \sin(\varphi_1 - \varphi_2)$. Such a system admits the trajectory $I_2(t) = \varepsilon t = -I_1(t)$, $\varphi_1(t) = \varphi_2(t) = -\varepsilon t^2/2$. Hence in a time T the actions have a drift of order 1, namely $|I_1(t) - I_1(0)| = |I_2(t) - I_2(0)| = 1$.*

Stability and instability for the action variables: some open problems

V.I. Arnold, in 1963, in one of the fundamental paper of KAM theory [9], conjectured that the general feature of nearly-integrable Hamiltonian dynamics with more than two degrees-of-freedom (i.e., with phase space of dimension greater than four) is “action-instability”. Arnold called it “topological instability” and formulated the following **Conjecture (Arnold²)**: “A typical case in many-dimensional problems of perturbation theory is topological instability: through an arbitrary small neighborhood of any point there pass phase trajectories along which the “slow variables³” drift away from the initial values by a quantity of order one”.

In any case, KAM Theorem does not completely solve the problem of action-(in)stability for the Hamiltonian system (0.1). For example, it leaves open the following questions:

- (1) what can we say for two degrees of freedom degenerate systems?
- (2) if action-instability takes place, what is the minimal time (depending on ε) needed for it to happen?
- (3) given h , for which perturbations f does action-instability take place?

In this thesis we will study some examples from (1),(2),(3). We first briefly discuss what is known about these problems and later we explain our contributions. In any case we observe that the problem (3) (and the connected question (2)) is more interesting and it is still deeply studied.

Problem 1: total stability for two degrees of freedom degenerate systems

We observe that to investigate degenerate systems is interesting not only from a mathematical point of view but also from the physical one. In fact, a typical feature in Celestial Mechanics (and this was the field from which Poincaré himself originally took motivations to look up at the action-stability problem) is that the unperturbed system

² Compare [8] pg. 189 from which the citation is taken.

³ Namely the “action variables”.

is properly degenerate, i.e., the unperturbed Hamiltonian function h in (0.1) does not depend upon all action variables. In such a case the non-degeneracy condition (0.2) is obviously strongly violated.

Properly degenerate models coming from Celestial Mechanics, as the spatial restricted three-body, have been studied, for example, in [72], [14]. Another typical example of a properly degenerate system studied in Classical Mechanics is the problem of the fast rotations of a symmetric rigid body (see [12], [13], [11]).

In any case, in [9], Arnold proved the following result (compare also [8], Chapter 5, Section 3). Consider a nearly-integrable (real-analytic) Hamiltonian system with two degrees of freedom governed by

$$H(I, \varphi; \varepsilon) \equiv H_0(I; \varepsilon) + \varepsilon^2 H_1(I, \varphi) \equiv H_{00}(I_1) + \varepsilon H_{01}(I) + \varepsilon^2 H_1(I, \varphi) , \quad (0.3)$$

where $(I, \varphi) = (I_1, I_2, \varphi_1, \varphi_2) \in \Omega \times \mathbb{T}^2$, and $\Omega \subset \mathbb{R}^2$. We say that the “perturbation removes the degeneracy⁴” on the energy level $H^{-1}(E)$, if

$$\frac{\partial H_{00}}{\partial I_1}(I) \neq 0 , \quad \frac{\partial^2 H_{01}}{\partial I_2^2}(I) \neq 0 , \quad \forall I \in H_0^{-1}(E) . \quad (0.4)$$

Theorem 0.1 ([9]) *If, in a (real-analytic) properly degenerate system with two degrees of freedom, the perturbation removes the degeneracy (i.e., condition (0.4) holds), then, for all ε small enough, total stability holds (i.e., for all initial data on the given energy level, the values of the action-variables stay forever near their initial values).*

Notwithstanding the previous results the problem of the action-stability for the Hamiltonian (0.3), when (0.4) does not hold, is still open (see [36]).

On the other hand it is interesting (also for physical reasons, see below) to take up the action-stability problem for properly degenerate Hamiltonian system with two degrees of freedom allowing the intermediate system H_{01} to depend also on the angle φ_1 . Hence we want to consider also real-analytic, properly-degenerate systems with two degrees of freedom described by nearly-integrable, real-analytic Hamiltonians given by

$$H(I, \varphi; \varepsilon) \equiv H_{00}(I_1) + \varepsilon H_{01}(I, \varphi_1) + \varepsilon^a H_1(I, \varphi) , \quad 0 < \varepsilon \ll 1 , \quad a > 1 . \quad (0.5)$$

We observe also that the dependence of H_{01} upon the angle φ_1 (that is, on the angle conjugated to the non-degenerate action I_1), besides being motivated by classical examples, is the only significant angle-dependence one wants to take into account in connection with the problems considered here⁵.

⁴ Or, more precisely, that “the intermediate term H_{01} removes the degeneracy”.

⁵ In general, in fact, a Hamiltonian function of the form $H_{00}(I_1) + \varepsilon H_{01}(I, \varphi_2) + \varepsilon^2 H_1(I, \varphi)$ will be trivially unstable as the following example shows. Let $H_{01} = \frac{I_2^2}{2} - (1 + \cos \varphi_2)$ and $H_1 = 0$. Then, one has $\sup_t |I_2(t) - I_2(0)| = 2$, for any $\varepsilon > 0$ and for any motion with $(I_2(0), \varphi_2(0))$ belonging to the (open) separatrix of the pendulum H_{01} . Moreover, these hyperbolic motions would be persistent under non-vanishing perturbations H_1 .

As we pointed out before, the interest for such systems stems again from Celestial Mechanics: compare, for example, with the above-mentioned “D’Alembert planetary model”, which we are, now, going to illustrate more in details.

The D’Alembert planetary model

As we said before, the D’Alembert planetary model is a Hamiltonian model for a rotational planet (or satellite) with polar radius slightly smaller than the equatorial one, whose center of mass revolves periodically on a given Keplerian ellipse of small eccentricity around a fixed star (or major body) occupying one of the foci of the ellipse; the planet is subject only to the gravitational attraction of the star. This is a Hamiltonian system with two degrees of freedom depending periodically on time (the period being the “year” of the planet). Many planets and satellites of the Solar System are observed in a nearly exact spin-orbit resonance, i.e., the ratio between the period of revolution around the major body and the period of rotation around the spin axis of the planet is (nearly) rational. Therefore it is of particular interest to investigate the stability in regions of space surrounding such resonances.

In formulae, the above system is governed by a real-analytic Hamiltonian of the form⁶

$$H_{\varepsilon, \mu} := \frac{I_1^2}{2} + \omega(pI_1 - qI_2 + qI_3) + \varepsilon F_0(I_1, I_2, \varphi_1, \varphi_2) + \varepsilon \mu F_1(I_1, I_2, \varphi_1, \varphi_2, \varphi_3; \mu) , \quad (0.6)$$

where: $(I, \varphi) \in A \times \mathbb{T}^3$ are standard symplectic coordinates; the domain $A \subset \mathbb{R}^3$ is given by

$$A \equiv \left\{ |I_1| < r\varepsilon^\ell , \quad |I_2 - \bar{J}_2| < r , \quad I_3 \in \mathbb{R} \right\} , \quad (0.7)$$

with $0 < \ell < 1/2$, $r > 0$; \bar{J}_2 is a fixed “reference datum” (avoiding certain singularities); ε and μ are two small parameters (measuring, respectively, the oblateness of the planet and the eccentricity of the Keplerian ellipse); p and q are two positive co-prime integers, which identify the spin-orbit resonance (the planet, in the unperturbed regime, revolves q times around the star and p times around its spin axis); ωq is the frequency of the Keplerian motion; the action I_1 measures the displacement from the exact resonance: in these units, $I_1 = 0$ corresponds exactly to a $p : q$ spin-orbit resonance. In fact, $\bar{J}_1 + I_1$ (where $\bar{J}_1 := p\omega$) and I_2 are (in suitable physical units), respectively, the absolute value and the projection onto the unit normal to the ecliptic plane of the angular momentum of the planet, while I_3 is an artificially introduced variable canonically conjugated to time. We also observe that, since the planet is rotational, the absolute value of the projection L of the angular momentum of the planet onto its polar axis is constant. If the parameters \bar{J}_i , L and the constant r are assumed to satisfy

$$L + 3r\varepsilon^\ell < \bar{J}_1 , \quad |\bar{J}_2| + 3r(\varepsilon^\ell + 1) < \bar{J}_1 , \quad (0.8)$$

then the functions F_i are real-analytic functions in all their arguments, computable via Legendre expansions in the eccentricity μ from the Lagrangian expression of the gravitational (Newtonian) potential (see below and [49] for explicit computations).

⁶ See part I below.

We immediately observe that the Hamiltonian (0.6) is strongly degenerate. This fact is even more clear after the (natural) symplectic linear change of variables⁷ $(I, \varphi) \rightarrow (I', \varphi')$, $(I'_1 := I_1, I'_2 := I_2, I'_3 := pI_1 - qI_2 + qI_3)$, by which the unperturbed part, $H_{0,\mu}$, becomes simply $I_1'^2/2 + \omega I_3'$, which is even properly-degenerate.

Moreover the previous change of variables clearly shows the appearance of *three (well separated) time scales* for the evolution of the angles φ' 's, namely (for ε small and $\mu < 1$)

$$\dot{\varphi}'_1 = O(\varepsilon^\ell) \gg O(\sqrt{\varepsilon}) , \quad \dot{\varphi}'_2 = O(\varepsilon) , \quad \dot{\varphi}'_3 = \omega = O(1) . \quad (0.9)$$

We will see in part III below that, after taking $\mu := \varepsilon^c$ (with $c > 0$), we can perform another symplectic change of variables (which is $\varepsilon^{c'}$ -near to the identity with $c' > 0$) $(I', \varphi') \rightarrow (\hat{I}, \hat{\varphi})$, averaging the “fast” angle φ_3 . After the previous change of variables the motion of the planet is governed, *up to an ε -exponentially small term*, by the two degree of freedom⁸ Hamiltonian⁹

$$H_D(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon, \mu) \equiv \frac{\hat{I}_1^2}{2} + \varepsilon \bar{H}_{01}(\hat{I}_2, \hat{\varphi}_1) + \varepsilon^a H_1^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon) , \quad (0.10)$$

which is actually of the form (0.5). We point out that the term $\bar{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ does really depend on φ_1 only in the cases $(p, q) = (1, 1), (2, 1)$, which are however very interesting in our solar system (as it is well known the Moon has $(p, q) = (1, 1)$, revolving one time around the Earth while it makes one turn on itself.).

Problem 2: the Nekhoroshev Theorem and exponential stability for the D'Alembert model

The most important result regarding the time of stability for the Hamiltonian (0.1) is the Nekhoroshev Theorem. In [90] N.N. Nekhoroshev proved exponential stability for the action variables, namely that if (0.1) verifies some non-degeneracy conditions (namely h is “steep”, see [90],[91]) then $|I(t) - I(0)| \leq \varepsilon^b$ for all $|t| \leq \exp(1/\varepsilon^a)$ where a, b are suitable positive constants. Hence, if the actions I undergo a drift of order one in a time¹⁰ T_d (called “instability time” or “diffusion time”), then it must be $T_d \geq \exp(1/\varepsilon^a)$.

In particular we deal with the problem of exponential stability for the D'Alembert model (0.6). From a physical point of view, one is interested in knowing the variation of the angle between the normal to the ecliptic plane and the polar axis of the planet. For example, in the case of the Earth, an “appreciable” increasing of this angle would cause a modification of the weather with warmer summer and colder winter (see also remark 7.2). On the other hand, from a mathematical point of view, the variation of the above-mentioned angle is related to the variation of the action variables for the Hamiltonian

⁷ See (8.2).

⁸ In fact the dependence on φ'_3 appears only at exponentially small terms: so the action I'_3 is “constant” up to exponentially large times and can be disregarded.

⁹ Where \bar{H}_{01} and $H_1^{(1)}$ are defined in (8.6), (8.12) and $a := 1 + \min\{\ell, c\} > 1$.

¹⁰ Namely $|I(T_d) - I(0)| \geq \text{const.}$ for a certain $T_d > 0$.

(0.6). However, due to the strong degeneracy of the model, the previous Nekhoroshev Theorem cannot be directly applied.

Problem 3: topological instability (Arnold diffusion)

We first point out that J. Mather has recently announced the solution of the above-mentioned Arnold's conjecture for a *generic* class of perturbations¹¹ (see, for Mather's Theory, [81], [82], [83], [84], [85], [64]).

In any case, the level of generality proposed is much higher than the one required on usual constructive proofs based on concrete examples. However, it could be not completely obvious to verify, in concrete cases, that the perturbation f belongs to the generic class proposed by Mather. Moreover it has to be noted that the proof of this statement is not yet available. Finally we remark that, regarding problem (2), it could be interesting to find what kind of estimates on the diffusion time can be derived by Mather's techniques.

The classical approach to prove the existence of topological instability usually follows the scheme proposed by Arnold in his famous paper [10].

As suggested by normal form theory near simple resonances, the Hamiltonian model

¹¹ For completeness we state the abstract of the talk given by Mather in [86] (see also [87]):
Consider a small perturbation of an integrable Hamiltonian:

$$H_\varepsilon := h(I) + \varepsilon P(\theta, I, t, \varepsilon)$$

where $I = (I_1, \dots, I_n) \in B^n$, a ball of radius 2 in \mathbb{R}^n ; $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{T}^n$; $t \in \mathbb{T}$ (i.e., the system is time-periodic of period 1); and $0 < |\varepsilon| < \varepsilon_0$ (i.e. ε is a small parameter). We study the solutions of Hamilton's equations $\dot{\theta} = \partial_I H$, $\dot{I} = -\partial_\theta H$.

The oscillation $\text{osc} I$ along a trajectory is defined by $\text{osc} I := \sup \|I(t_1) - I(t_0)\|$. A particular perturbation H_ε is said to exhibit Arnold diffusion if there exists a trajectory for which $\text{osc} \geq 1$. In this talk, we announced the existence of a large class of perturbations which exhibit Arnold diffusion, in the case that $h'' > 0$.

Let E be a topological vector space, We will say that a subset W of E is a "cusp-residual" if the following holds:

- 1) There exists an open dense subset U of E such that $v \in U$ and $\lambda > 0$ imply $\lambda v \in U$.
- 2) There exists an open subset V of U such that if $\gamma : [0, \delta_0] \rightarrow E$ is a C^1 -curve, $\gamma(0) = 0$ and $\gamma'(0) \in U$, then there exists $0 < \delta \leq \delta_0$ such that $\gamma((0, \delta)) \subset V$.
- 3) W is an open dense subset of V .

In the following we assume r is a large integer, ∞ , or ω and $h'' > 0$. We assume h is C^r .

Theorem There exists a cusp-residual set W in the space of C^r -perturbations of h such that any perturbation in W exhibits Arnold diffusion.

considered by Arnold is¹²

$$H(I, \varphi, p, q, t; \varepsilon, \mu) := \frac{I^2}{2} + \frac{p^2}{2} + \varepsilon(\cos q - 1) + \varepsilon\mu f(\varphi, q, t) \quad (\mathcal{AH})$$

where $I, p \in \mathbb{R}$, $\varphi, q, t \in \mathbb{T}$; while ε and μ are small parameters. Taking $f := (\cos q - 1)(\sin \varphi + \cos t)$, Arnold proved topological instability for μ exponentially small w.r.t. $\sqrt{\varepsilon}$. The diffusion time turns out to be exponentially long with respect to ε (as subsequently proved by [26]).

The mechanism proposed in [10] to prove the existence of action–instability and thereafter become classical as “Arnold diffusion¹³”, is based on the construction of *transition chains*. One first remarks that for $\mu = 0$, the Hamiltonian system associated to H admits a continuous family of 2-dimensional partially hyperbolic invariant tori $\mathcal{T}_{\bar{I}} = \{(\varphi, t) \in \mathbb{T}^2, I = \bar{I}, q = p = 0\}$ possessing 3-dimensional stable and unstable manifolds $W_0^s(\mathcal{T}_{\bar{I}}) = W_0^u(\mathcal{T}_{\bar{I}}) = \{(\varphi, t) \in \mathbb{T}^2, I = \bar{I}, (p^2/2) + \varepsilon(\cos q - 1) = 0\}$.

Next one tries to construct, for $\mu \neq 0$, transition chains, namely chains of perturbed partially hyperbolic tori \mathcal{T}_I^μ close to $\mathcal{T}_{\bar{I}}$, connected one to another by heteroclinic orbits. Finally one proves, through a “shadowing argument”, the existence of a true diffusion orbit close to the transition chain.

By the previous scheme one is then faced with the following 3 steps:

- (i) **splitting problem.** Prove the persistence, for $\mu \neq 0$ small enough, of such hyperbolic tori \mathcal{T}_I^μ and show that the perturbed stable and unstable manifolds $W_\mu^s(\mathcal{T}_I^\mu)$ and $W_\mu^u(\mathcal{T}_I^\mu)$ “split” and intersect transversally, giving a quantitative measure of these “splitting angles”.
- (ii) **construction of a transition chain.** This is a difficult task since the surviving perturbed tori \mathcal{T}_I^μ are separated by the gaps appearing in KAM constructions: two perturbed invariant tori \mathcal{T}_I^μ and $\mathcal{T}_{I'}^\mu$ could be too distant one from the other one, forbidding the existence of a heteroclinic intersection between $W_\mu^u(\mathcal{T}_I^\mu)$ and $W_\mu^s(\mathcal{T}_{I'}^\mu)$. This is the well-known **gap problem**.
- (iii) **shadowing argument:** to prove the existence of a true diffusion orbit, shadowing the transition chain, for which the action variables I undergo a drift of $O(1)$ in a certain diffusion time T_d .

In [10] the splitting of stable and unstable manifolds is proved assuming μ exponentially small with respect to ε in order to justify straightforwardly the Poincaré-Melnikov

¹² Hamiltonian H can be written in an *autonomous* (i.e. independent on time) way as in (0.1), introducing the angular variable $\varphi_3 = t$ and its conjugated action I_3 and renaming $I_1 = I$, $\varphi_1 = \varphi$, $I_2 = p$, $\varphi_2 = q$, in the following way

$$H(I, \varphi, ; \varepsilon, \mu) := I_1^2/2 + I_2^2/2 + I_3 + \varepsilon(\cos \varphi_2 - 1) + \varepsilon\mu f(\varphi_1, \varphi_2, \varphi_3).$$

¹³ Such a term was introduced, in the physical literature, in [51].

approximation. The gap problem is bypassed by the peculiar choice of the perturbation $(\cos q - 1)(\sin \varphi + \cos t)$ whose gradient vanishes on the unperturbed tori \mathcal{T}_I , leaving them *all* invariant also for $\mu \neq 0$. The splitting of stable and unstable manifolds is proved assuming μ exponentially small with respect to ε in order to justify straightforwardly the Poincaré-Melnikov approximation. Finally we mention that the shadowing argument of [10] is based on geometrical techniques.

After thirty years from Arnold seminal work [10], attention to Arnold diffusion was renewed in 1994 by Chierchia and Gallavotti [49] followed by several papers approaching the problem from different points of view (see e.g. [48], [26], [27], [28], [66], [69], [30], [23], [24] and references therein).

In [49] topological instability is proved for a class of perturbations not preserving the unperturbed tori: extending Arnold's analysis, they dealt with step (ii) showing that, if the perturbation f is a trigonometric polynomial in the angles, then, in some regions of the phase space, the *density* of perturbed invariant tori is high enough to allow the construction of a transition chain.

Regarding problem (i) we point out that, up to now, Arnold diffusion has been proved only for μ exponentially small w.r.t. ε and no results are available for $\mu = O(\varepsilon^p)$ for some positive power p . This is a very difficult problem since the Melnikov function is exponentially small w.r.t. some power of ε , and then the naive Poincaré Melnikov expansion provides a valid measure of the splitting only for μ exponentially small w.r.t. ε . For this reason the case $\mu = O(\varepsilon^p)$ has only been proved in simpler models, starting from the pioneering paper [74], which deals with a rapidly forced pendulum. This method has been further refined in [71] and [70]. Different approaches to the splitting problem have been developed in [67], [68], by means of trees techniques and in [63], [60] by normal form theory. We also quote the more recent papers [96]-[79], [23] and [94]-[95].

We finally mention that the shadowing problem has been extensively studied in the last years by both geometrical (see e.g. [49], [80], [66], [55], [56]) and variational (see e.g. [40], [26],[27], [28], [42], [30], [23], [24]) techniques.

Main results

We now explain our contributions regarding problems (1),(2),(3).

Problem 1: total stability for two degrees of freedom degenerate systems

We consider Hamiltonians of the form (0.6). We first study the Hamiltonian

$$H_0(I, \varphi_1; \varepsilon) := H_{00}(I_1) + \varepsilon H_{01}(I, \varphi_1) , \quad (0.11)$$

which, regarded as a one-degree-of-freedom system in the (I_1, φ_1) variables, is still integrable exhibiting, in general, the typical features of a one-degree-of-freedom dimensional system (phase space regions foliated by invariant circles of possibly different homotopy,

stable/unstable equilibria, separatrices, etc.). A natural approach (which we shall, in fact, follow) is to introduce action–angle variables for the one–degree–of–freedom Hamiltonian $H_0(I, \varphi_1; \varepsilon)$ (regarding I_2 as a dumb parameter) and then to apply KAM techniques trying to confine all motions among KAM tori (as in the non–degenerate case). The problem with this approach is that the action–angle variable for the (I_1, φ_1) system are singular in any neighborhood of the separatrix (and stable equilibria) and is exactly near separatrices where one expects the motion to become “chaotic” and where, in principle, drift of order one in the I_2 variable is conceivable¹⁴ even in the two–degrees–of–freedom (properly degenerate) case considered here. Therefore a careful analysis near these “singular phase space regions” is needed and arguments different from KAM theory have to be used to control the displacement of the action variable in such singular regions. Clearly (see Remark 0.1) regions where the non–degeneracy assumption fails need a separate discussion: in fact, in such zones (and in the non convex case), we can not exclude a “possibly non–chaotic–drift” of the I_2 action.

To avoid “extra” technical difficulties we consider *model problems*, namely, we shall let

$$H_{00} := \frac{I_1^2}{2}, \quad H_{01} := H_{01}^{(\sigma)} := \sigma \frac{I_2^2}{2} - (1 + \cos \varphi_1), \quad (0.12)$$

with σ equal either $+1$ or -1 ; the phase space will be taken to be $\mathcal{M}_{R_0} := B_{R_0}^2 \times \mathbb{T}^2$ where $B_{R_0}^2$ denotes a ball of radius R_0 around the origin.

Remark 0.2 *These model problems are intended to capture the main features of “general” properly degenerate systems with two degrees of freedom as, for example, the exponential approximation (0.10) to the D’Alembert Hamiltonian. This is the reason for considering both the convex and the non convex¹⁵ case in (0.12), corresponding, respectively to $\sigma = 1$ and $\sigma = -1$.*

We can now state our result (see Theorem 4.3) about problem (1) which is a simple corollary of Theorem 4.4.

Theorem 0.2 *Let $H^{(\sigma)}(I, \varphi; \varepsilon) := H(I, \varphi; \varepsilon)$ as in (0.5), (0.12), and \mathcal{M}_{R_0} be as in (4.5), (4.6). Assume $a > 3/2$ and choose*

$$0 < R < R_0 \quad \text{and} \quad 0 < b < \min \left\{ \frac{1}{4}, \frac{a-1}{4}, \frac{1}{3} \left(a - \frac{3}{2} \right) \right\}. \quad (0.13)$$

Then, there exists $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, the ϕ_H^t –evolution $(I(t), \varphi(t))$ of an initial datum (I_0, φ_0) satisfies

$$|I(t)| < R_0, \quad |I(t) - I_0| < \varepsilon^b, \quad \forall t \in \mathbb{R}, \quad (0.14)$$

¹⁴ Better: “compatible with energy conservation”.

¹⁵ With respect to the D’Alembert model, regions in which the Hamiltonian (0.10) is non convex correspond to unperturbed situations in which the spin axis of the planet is nearly orthogonal to ecliptic plane (i.e., to the plane containing the Keplerian ellipse), as it results from (8.7): this is the observed situation for most planets in the Solar system.

where, in the case $\sigma = 1$, (I_0, φ_0) is an arbitrary point in the phase space \mathcal{M}_R , while, in the “non-convex” case $\sigma = -1$, (I_0, φ_0) belongs to $\mathcal{M}_R \setminus \mathcal{N}_*$, \mathcal{N}_* being an open region whose measure does not exceed $\varepsilon^{2/3}$.

Problem 2: exponential stability for the D’Alembert planetary model

Now we pass to the problem of action–stability in the D’Alembert planetary model. Notwithstanding the strong degeneracies of the model, stability results *à la* Nekhoroshev (i.e. for times which are exponentially long in the perturbative parameters) hold (see also Theorem 7.1).

Theorem 0.3 *Let $c > 0$, $0 < \ell < 1/2$ and $0 < \gamma_0 < \min\{c, \ell\}$. Assume that (0.8) holds and that*

$$\bar{J}_1 \neq \sqrt{3}L. \quad (0.15)$$

Then, there exist $\varepsilon_0, \gamma_i, \omega > 0$ such that, if $0 \leq \varepsilon \leq \varepsilon_0$ and $0 \leq \mu \leq \varepsilon^c$, then

$$|I(t) - I(0)| < \gamma_3 r \varepsilon^{\gamma_1}, \quad \forall \quad |t| < T(\varepsilon) := \frac{\gamma_5}{\omega \varepsilon^{\gamma_4}} \exp\left(\frac{\gamma_2}{\varepsilon^{\gamma_0}}\right), \quad (0.16)$$

where $(I(t), \varphi(t))$ denotes the evolution of an initial datum¹⁶ $(I(0), \varphi(0)) \in A \times \mathbb{T}^3$ for the D’Alembert Hamiltonian (0.6).

The reason for which it is possible to prove stability notwithstanding the properly-degeneracy of the model, is mainly related to the appearance of the above mentioned three well separated times scales in (0.9). In particular, the “non-degeneracy” assumption (0.15) is made in order to really have $\dot{\varphi}'_2 = O(\varepsilon)$, i.e., in order to have, $\frac{1}{\varepsilon} \partial_{I'_2} H_{\varepsilon, \mu}|_{I'_1=0, I'_2=\bar{J}_2} \neq 0$.

As we said before, after averaging the “fast” angle φ_3 , we have that I'_3 is stable for exponentially long time¹⁷ and the motion of I_1 and I_2 is $\varepsilon^{c'}$ -near to motion of \hat{I}_1 and \hat{I}_2 , which is governed, up to an ε -exponentially small term, by the two degree of freedom Hamiltonian (0.10). From (0.10), by energy conservation, it is simple to see that \hat{I}_1 (and hence I_1) is stable, being its variation bounded by $O(\sqrt{\varepsilon})$. So the only non-trivial stability is related to \hat{I}_2 (namely I_2). At this point the further analysis strongly depends on the dependence (in the case $(p, q) = (1, 1)(2, 1)$) or independence (in the other cases) of the term \bar{H}_{01} upon $\hat{\varphi}_1$. In the second case we can exploit the separation of scales between $\dot{\hat{\varphi}}_1$ and $\dot{\hat{\varphi}}_2$: roughly speaking, $\hat{\varphi}_1$ is fast with respect to $\hat{\varphi}_2$ and, therefore, the dependence upon $\hat{\varphi}_1$ can be removed up to exponentially small terms; finally using energy conservation we obtain stability also for \hat{I}_2 and hence for I_2 .

To carry out an analogous strategy in the first case, we have to, first, put the two-dimensional integrable system $\frac{\hat{I}_2^2}{2} + \varepsilon \bar{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ into action–angle variables. Moreover for our purposes, we need detailed information on the analyticity domains of this symplectic change of variables, which becomes singular as $\varepsilon \rightarrow 0$. After this (elementary but technical)

¹⁶ A is defined in (0.7).

¹⁷ See footnote 8.

analysis, we prove the Theorem using jointly normal form theory and energy conservation arguments, fully exploiting the separation of scales described above.

Problem 3: a new variational method

(*Preliminary observations*) As proved in [18] we provide a new mechanism to produce diffusion orbits *not based* on the existence of transition chains of tori: we avoid the KAM construction of the perturbed hyperbolic tori, proving directly the existence of a drifting orbit as a local minimum of the action functional. At the same time our variational approach achieves the optimal diffusion time. Our diffusion time estimate is the optimal one as a consequence of a general stability result, proved via classical perturbation theory.

As in [49] we deal with a perturbation which is a trigonometric polynomial in the angles and our diffusion orbits will not connect any two arbitrary frequencies of the action space, even if we manage to connect more frequencies than in [49], proving the drift also in some regions of the phase space where transition chains might not exist. Clearly if the perturbation is chosen as in Arnold's example we can drift in all the phase space with no restriction.

We will assume, as in Arnold's paper, the parameter μ to be small enough in order to validate the so called Poincaré-Melnikov approximation, when the first order expansion term in μ for the splitting, the so called Poincaré-Melnikov function, is the dominant one. For this reason, we will fix the "Lyapunov exponent" of the pendulum $\varepsilon := 1$, considering the so called "a-priori unstable" case. Actually our variational shadowing technique is not restricted to the a-priori unstable case, but could allow, in the same spirit of [22], [23] and [24], once a "splitting condition" is somehow proved, to get diffusion orbits with the best diffusion time, in terms of some measure of the splitting (see remark 0.3 below).

(*Our model*) We consider nearly integrable *non-isochronous* Hamiltonian systems defined by

$$\mathcal{H}_\mu = \frac{I^2}{2} + \frac{p^2}{2} + (\cos q - 1) + \mu f(I, \varphi, p, q, t), \quad (0.17)$$

where $(\varphi, q, t) \in \mathbb{T}^d \times \mathbb{T}^1 \times \mathbb{T}^1$ are the angle variables, $(I, p) \in \mathbb{R}^d \times \mathbb{R}^1$ are the action variables and $\mu \geq 0$ is a small real parameter. The Hamiltonian system associated with \mathcal{H}_μ writes

$$\dot{\varphi} = I + \mu \partial_I f, \quad \dot{I} = -\mu \partial_\varphi f, \quad \dot{q} = p + \mu \partial_p f, \quad \dot{p} = \sin q - \mu \partial_q f. \quad (\mathcal{S}_\mu)$$

The perturbation f is a real trigonometric polynomial of order N in φ and t , namely¹⁸

$$f(I, \varphi, p, q, t) = \sum_{|(n,l)| \leq N} f_{n,l}(I, p, q) e^{i(n \cdot \varphi + lt)}. \quad (0.18)$$

For simplicity, even if it is not really necessary, we assume f to be a purely spatial perturbation, namely $f(\varphi, q, t) = \sum_{0 \leq |(n,l)| \leq N} f_{n,l}(q) \exp(i(n \cdot \varphi + lt))$. The functions $f_{n,l}$ are assumed to be smooth.

¹⁸ Where $\bar{f}_{n,l}(I, p, q) = f_{-n,-l}(I, p, q)$ for all $(n, l) \in \mathbb{Z}^d \times \mathbb{Z}$ with $|(n, l)| \leq N$, and \bar{z} denotes the complex conjugate of $z \in \mathbb{C}$.

Define the “resonant web” \mathcal{D}_N formed by the frequencies ω “resonant with the perturbation”

$$\mathcal{D}_N := \left\{ \omega \in \mathbb{R}^d \mid \exists (n, l) \in \mathbb{Z}^{d+1} \text{ s.t. } 0 < |(n, l)| \leq N \text{ and } \omega \cdot n + l = 0 \right\} \quad (0.19)$$

and the Poincaré-Melnikov primitive

$$\Gamma(\omega, \theta_0, \varphi_0) := - \int_{\mathbb{R}} \left[f(\omega t + \varphi_0, q_0(t), t + \theta_0) - f(\omega t + \varphi_0, 0, t + \theta_0) \right] dt,$$

where $q_0(t) := 4 \arctan(\exp t)$ is the separatrix of the unperturbed pendulum equation $\ddot{q} = \sin q$ satisfying $q_0(0) = \pi$.

(Our diffusion Theorem) The next Theorem (see Theorem 10.1) states that, for any connected component $\mathcal{C} \subset \mathcal{D}_N^c$, $\omega_I, \omega_F \in \mathcal{C}$, there exists a solution of (\mathcal{S}_μ) connecting a $O(\mu)$ -neighborhood of ω_I in the action space to a $O(\mu)$ -neighborhood of ω_F , in the time-interval $T_d = O((1/\mu)|\log \mu|)$.

Theorem 0.4 *Let \mathcal{C} be a connected component of \mathcal{D}_N^c , $\omega_I, \omega_F \in \mathcal{C}$ and let $\gamma : [0, L] \rightarrow \mathcal{C}$ be a smooth embedding such that $\gamma(0) = \omega_I$ and $\gamma(L) = \omega_F$. Assume that, for all $\omega := \gamma(s)$ ($s \in [0, L]$), $\Gamma(\omega, \cdot, \cdot)$ possesses a non-degenerate local minimum $(\theta_0^\omega, \varphi_0^\omega)$. Then $\forall \eta > 0$ there exists $\mu_0 = \mu_0(\gamma, \eta) > 0$ and $C = C(\gamma) > 0$ such that $\forall 0 < \mu \leq \mu_0$ there exists a solution $(I_\mu(t), \varphi_\mu(t), p_\mu(t), q_\mu(t))$ of (\mathcal{S}_μ) and two instants $\tau_1 < \tau_2$ such that $I_\mu(\tau_1) = \omega_I + O(\mu)$, $I_\mu(\tau_2) = \omega_F + O(\mu)$ and*

$$|\tau_2 - \tau_1| \leq \frac{C}{\mu} |\log \mu|. \quad (0.20)$$

Moreover $\text{dist}(I_\mu(t), \gamma([0, L])) < \eta$ for all $\tau_1 \leq t \leq \tau_2$.

In addition, the above result still holds for any perturbation $\mu(f + \mu \tilde{f})$ with any smooth $\tilde{f}(\varphi, q, t)$.

We can also build diffusion orbits approaching the boundaries of \mathcal{D}_N at distances as small as a certain power of μ : see for a precise statement Theorem 15.1.

(Our improvements w.r.t. previously known results) Theorem 0.4 improves the corresponding result in [49] which enables to connect two frequencies ω_I and ω_F belonging to the same connected component $\mathcal{C} \subset \mathcal{D}_{N_1}^c$ for $N_1 = 14dN$ and with $\text{dist}\{\{\omega_I, \omega_F\}, \mathcal{D}_{N_1}\} = O(1)$. Such restrictions of [49] in connecting the action space through diffusion orbits arise because transition chains could not exist in all $\mathcal{C} \subset \mathcal{D}_N^c$ (see remark 11.2). Unlikely, given any two frequencies $\omega_I, \omega_F \in \mathcal{C} \subset \mathcal{D}_N^c$ and any path, lying in \mathcal{D}_N^c , joining them, our method enables to find a trajectory of (0.17) running along this path and joining ω_I and ω_F .

Theorem 0.4 also improves the known estimates on the diffusion time. The first estimate obtained by geometrical method in [49], is $T_d = O(\exp(1/\mu^2))$. In [80]-[55]-[56], still by geometrical methods, and in [30], by means of Mather’s theory, the diffusion time has been proved to be just polynomially long in the splitting μ (the splitting angles between

the perturbed stable and unstable manifolds $\mathcal{W}_\mu^{s,u}(\mathcal{T}_\omega^\mu)$ at a homoclinic point are, by classical Poincaré-Melnikov theory, $O(\mu)$. We note that the variational method proposed by Bessi in [26] had already given, in the case of perturbations preserving all the unperturbed tori, the diffusion time estimate $T_d = O(1/\mu^2)$. For isochronous systems the estimate on the diffusion time was $T_d = O(\exp(1/\mu))$ in [66] and $T_d = O((1/\mu)|\log \mu|)$ has been obtained in [22]-[23]. Very recently, in [56], the diffusion time (in the non isochronous case) has been estimated as $T_d = O((1/\mu)|\log \mu|)$ by a method which uses “hyperbolic periodic orbits”; however the result of [56] is of local nature: the previous estimate holds only for diffusion orbits shadowing a transition chain close to some torus run with diophantine flow.

In Theorem 0.6 below we will conclude this historical quest for the minimal diffusion time T_d showing the optimality of our estimate $T_d = O((1/\mu)|\log \mu|)$.

(*Our method*) We now briefly discuss the method of proof of Theorem 0.4 (and Theorem 0.5 below). It relies on a finite dimensional reduction of Lyapunov-Schmidt type, variational in nature, introduced by Ambrosetti and Badiale in [1] and later extended in [22],[23] and [24] to the problem of Arnold diffusion for isochronous systems. We also mention that the approach of [1], originated in [4], has been later on also extended with success for studying a huge variety of bifurcation problems of variational nature concerning nonlinear elliptic equations (see e.g. [5], [25], [6]).

The diffusion orbit of Theorem 0.4 is found as a local minimum of the action functional. In order to find critical points of the action functional, we evaluate it on suitable *pseudo-diffusion orbits*, whose (p, q) variables move along the separatrices of the pendulum. Pseudo-diffusion orbits are true solutions of (S_μ) except, possibly, at some instants $\theta_1 < \dots < \theta_i < \dots < \theta_k$ (for a suitable $k \in \mathbb{N}$ which represents the number of “bumps”), where they are glued continuously at the section $\{q \equiv \pi \bmod 2\pi\}$; namely $(\varphi_\mu(\theta_i^+), q_\mu(\theta_i^+)) = (\varphi_\mu(\theta_i^-), q_\mu(\theta_i^-)) = (\varphi_i, \pi)$ for suitable φ_i , but the speeds $(\dot{\varphi}_\mu(\theta_i^\pm), \dot{q}_\mu(\theta_i^\pm)) = (I_\mu(\theta_i^\pm), p_\mu(\theta_i^\pm))$ may have a jump.

The *reduced action functional*, obtained evaluating the action functional on the pseudo-diffusion orbits, is defined on the (finite) $k(1+d)$ -dimensional space parametrized by $\lambda = (\theta_1, \varphi_1, \dots, \theta_k, \varphi_k)$. It turns out (see lemma 11.3) that, after this finite-dimensional reduction, critical points of the reduced action functional correspond to smooth pseudo-orbits (i.e. pseudo-orbits with no jumps for the speeds at any θ_i), namely to true solutions of (S_μ) .

The time interval $T_i = \theta_{i+1} - \theta_i$ is heuristically the time required to perform a single transition during which the rotators can exchange $O(\mu)$ -energy, i.e. the action variables vary of $O(\mu)$. During each transition we can exchange only $O(\mu)$ -energy because the Melnikov contribution in the perturbed functional is $O(\mu)$. Hence in order to exchange $O(1)$ energy the number of transitions required will be $k = O(1/\mu)$.

We underline that the question of finding the optimal time and the mechanism for which we can avoid the construction of transition chains of tori are deeply connected. We are going to explain this essential fact.

A first arising question is the following one: if there are no KAM tori, how can we approximate the action functional in order to find a minimum? Namely, what are the correct objects to use (instead of perturbed hyperbolic KAM tori) in order to construct pseudo-orbits?

The key idea is to approximate the action functional using a suitable *chain of unperturbed solutions*. Namely, we construct our pseudo-orbits, “bifurcating” them from a chain of unperturbed solutions, joined together, at every θ_i , by continuity.

Actually, the unperturbed solutions of (S_0) , which we will use, are defined in the following way¹⁹: $(\omega_i(t - \theta_i) + \varphi_i, Q_{T_i}(t - \theta_i))$, for $t \in [\theta_i, \theta_{i+1}]$, where²⁰ $\omega_i := \varphi_{i+1} - \varphi_i / (\theta_{i+1} - \theta_i)$ and Q_{T_i} is the unique T_i -periodic solution of the pendulum $-\ddot{Q}_{T_i} + \sin Q_{T_i} = 0$ with positive energy. We will chose (θ_i, φ_i) in such a way that they are δ -close to the minima of the Poincaré-Melnikov primitive (let us say $(0, 0) \equiv \text{mod } 2\pi\mathbb{Z}^{1+d}$ for simplicity), where δ is a small, but *independent on μ* , constant.

Then we fix $\bar{\omega}_i \in \mathbb{R}^d$ such that: (a) $\bar{\omega}_i \bar{T}_i \equiv 0 \text{ mod } 2\pi\mathbb{Z}^d$, for suitable $\bar{T}_i \in 2\pi\mathbb{Z}$; (b) $|\bar{\omega}_{i+1} - \bar{\omega}_i| = O(\rho\mu)$, with ρ a suitable constant small as we wish; (c) $|\bar{\omega}_k - \bar{\omega}_1| \geq \text{const}$. We will search $\lambda = (\theta_1, \varphi_1, \dots, \theta_k, \varphi_k)$ in a set verifying, not only the above-mentioned condition $|(\theta_i, \varphi_i)| \leq \delta \text{ mod } 2\pi\mathbb{Z}^{1+d}$, but also $T_i - \bar{T}_i = O(\delta)$ and $\varphi_{i+1} - \varphi_i - \bar{\omega}_i \bar{T}_i = O(\delta)$. We observe that, with this positions, we have $|\omega_i - \bar{\omega}_i| = O(\delta/\bar{T}_i)$.

We will find (see lemma 11.1) our pseudo-orbits from the above-defined chain of unperturbed solutions, by the Implicit Function Theorem, provided every T_i is sufficiently large, namely $T_i \geq \text{const} \cdot |\log \mu|$.

With the above choice for T_i , we have $|\omega_i - \bar{\omega}_i| = O(1/|\log \mu|)$.

We point out that, if we are able to find a $\tilde{\lambda}$ which minimizes the reduced action functional, then we have found a diffusion orbit, since $I_\mu(\tilde{\theta}_1) \approx \omega_1 \approx \bar{\omega}_1$ and $I_\mu(\tilde{\theta}_k) \approx \omega_k \approx \bar{\omega}_k$.

Roughly speaking, by the estimates of the Implicit Function Theorem, the following approximation holds (see lemma 12.5)

$$\text{reduced action functional} = \text{kinetic part} + \mu \times \text{Melnikov contributions} + \text{rest} \quad (0.21)$$

where *kinetic part* is the kinetic part of the rotators, namely²¹ $\sum_i (\varphi_{i+1} - \varphi_i)^2 / 2(\theta_{i+1} - \theta_i)$ and *rest* $= O(\rho\mu + \mu^2 \max_i T_i)$.

If it is possible, as it is in fact, to take $T_i = O(|\ln \mu|)$ (actually we only need $T_i \ll 1/\mu$), then, on such “short” time interval, the *rest* is negligible.

However, the fact that it is possible to perform a single transition in a very short time interval like $T_i = O(|\ln \mu|)$ is not obvious at all. In the approach of Bessi [26] the time to perform a single transition, in the example of Arnold, is $O(1/\mu)$. This transition time arises in order to ensure that the variations of the *kinetic part* of the action functional associated with the rotators are small compared with $\mu \times \text{Melnikov contributions}$. Unfortunately this time is too long to use a simple approximation of the functional, because, as we have seen

¹⁹ In the (φ, q) variables. The solution for the (I, p) can be simply found deriving with respect to time.

²⁰ Namely the straight line connecting φ_i with φ_{i+1} .

²¹ We should consider also some border terms (see (12.15) and lemma 12.5).

before, the error should be $O(\mu)$, namely of the same order as $\mu \times$ Melnikov contributions, possibly destroying the existence of a minimum.

The key observation that enables us to perform a single transition in a very short time interval, improving substantially the arguments of [26], concerns the behavior of the “gradient flow” of the unperturbed action functional of the rotators. Roughly speaking, we have the following:

- (i) in the directions which are orthogonal to the “unperturbed gradient flow”, the kinetic part is convex and it becomes as a “potential well”, forcing the solution to remain close to the unperturbed pseudo-solution;
- (ii) instead, in the direction which is parallel to the “unperturbed gradient flow”, the kinetic part is almost “flat” (being invariant for the flow) and the minimum of the Poincaré-Melnikov primitive is preserved.

We point out that (i) implies a sort of a-priori estimate satisfied by the minimal diffusion orbits (see remark 15.1). In fact it results that, if $\tilde{\omega}_i := (\tilde{\varphi}_{i+1} - \tilde{\varphi}_i) / (\tilde{\theta}_{i+1} - \tilde{\theta}_i)$ are the frequencies of the minimal orbit, then $|\tilde{\omega}_i - \omega_i| = O(\sqrt{\mu}/|\log \mu|)$, which is much better than $|\tilde{\omega}_i - \bar{\omega}_i| = O(1/|\log \mu|)$, which would hold in general. We think that this estimate (see also (15.18)) is interesting in itself.

In this way we can show that the variations of the action of the rotators (namely the kinetic part) are small enough, even on time intervals $T_i \ll 1/\mu$, and do not “destroy” the minimum of the Poincaré-Melnikov primitive.

Summarizing the two main differences of our approach with respect to Bessi’s one are: (1) the new variational argument described above for controlling the oscillations of the action functional of the rotators; (2) the construction of the pseudo-diffusion orbits by the I.F.T. and not by minimization, obtaining sharp estimates for the reduced action functional.

(*Ergodization*) When trying to build a pseudo-diffusion orbit which performs single transitions in very short time intervals we encounter another difficulty linked with the ergodization time. The time to perform a single transition T_i must be long enough to settle, at each instant θ_i , the projection (θ_i, φ_i) of the pseudo-orbit on the torus \mathbb{T}^{1+d} sufficiently close to the minimum of the Poincaré-Melnikov function, i.e. the homoclinic point (in our method it is sufficient to arrive just $O(1)$ -close, independently of μ , to the homoclinic point). This necessary request creates some difficulty since our pseudo-diffusion orbit may arrive $O(\mu)$ -close in the action space to resonant hyperplanes of frequencies whose linear flow does not provide a dense enough net of the torus. The way in which this problem is overcome is discussed in section 14: we observe a phenomenon of “stabilization close to resonances” which forces the time for some single transitions to increase. Anyway the total time required to cross these (finite number of) resonances is still $T_d = O((1/\mu) \log(1/\mu))$, see (14.13) and the proof of Theorem 10.1. This discussion enables us to prove optimal fast action-instability in large regions of the phase space and allows to improve the local diffusion results of [56].

We need therefore some results on the ergodization time of the torus for linear flows possibly resonant but only at a “sufficiently high order”. We present these results in section 13. We point out that the main result of this section, Theorem 13.2, implies as corollaries Theorems B and D of [47], see remark 13.1. It is of independent interest and could possibly improve the other results of [47].

(*The Arnold’s example*) As a byproduct of the techniques developed we have the following result (see Theorem 10.2) concerning “Arnold’s example” [10] where $\mathcal{T}_\omega := \{I = \omega, \varphi \in \mathbb{T}^d, p = q = 0\}$ are, for all $\omega \in \mathbb{R}^d$, even for $\mu \neq 0$, invariant tori of (\mathcal{S}_μ) .

Theorem 0.5 *Let $f(\varphi, q, t) := (1 - \cos q)\tilde{f}(\varphi, t)$. Assume that for some smooth embedding $\gamma : [0, L] \rightarrow \mathbb{R}^d$, with $\gamma(0) = \omega_I$ and $\gamma(L) = \omega_F$, $\forall \omega := \gamma(s)$ ($s \in [0, L]$), $\Gamma(\omega, \cdot, \cdot)$ possesses a non-degenerate local minimum $(\theta_0^\omega, \varphi_0^\omega)$. Then $\forall \eta > 0$ there exists $\mu_0 = \mu_0(\gamma, \eta) > 0$, and $C = C(\gamma) > 0$ such that $\forall 0 < \mu \leq \mu_0$ there exists a heteroclinic orbit (η -close to γ) connecting the invariant tori \mathcal{T}_{ω_I} and \mathcal{T}_{ω_F} . Moreover the diffusion time T_d needed to go from a μ -neighborhood of \mathcal{T}_{ω_I} to a μ -neighborhood of \mathcal{T}_{ω_F} is bounded by $(C/\mu)|\log \mu|$ for some constant C .*

Remark 0.3 *Consider the a-priori stable Arnold’s example Hamiltonian (\mathcal{AH})*

$$\bar{H}(\bar{I}, \bar{\varphi}, \bar{p}, \bar{q}, t) := \bar{I}^2/2 + \bar{p}^2/2 + \varepsilon(\cos \bar{q} - 1) + \varepsilon\mu(\cos \bar{q} - 1)[\sin \bar{\varphi} + \cos t]. \quad (0.22)$$

After rescaling²², we have that, if $(\bar{I}(t), \bar{\varphi}(t), \bar{p}(t), \bar{q}(t))$ is a solution of (0.22), then $\bar{I}(t) = \sqrt{\varepsilon}I(\sqrt{\varepsilon}t)$, where $(I(t), \varphi(t), p(t), q(t))$ is a solution of the a-priori unstable Hamiltonian

$$H(I, \varphi, p, q, t) := I^2/2 + p^2/2 + (\cos q - 1) + \mu(\cos q - 1)[\sin \varphi + \cos t]. \quad (0.23)$$

Hence, in order to prove topological instability for (0.22), the fact that $\bar{I}(T_d) - \bar{I}(0)$ is order 1 in ε is equivalent to $I(\sqrt{\varepsilon}T_d) - I(0) = O(1/\sqrt{\varepsilon})$. The oscillations of the Poincaré-Melnikov function in the Hamiltonian (0.23), for frequencies $\omega = O(1/\sqrt{\varepsilon})$ (which we must deal with), turn out to be exponentially small in $\sqrt{\varepsilon}$ (namely $O(\exp(-c_1/\sqrt{\varepsilon}))$). In order to use Theorem 0.4, we have to find the values of μ_0 and C such that we have a diffusion orbit performing a distance of $O(1/\sqrt{\varepsilon})$. Obviously μ_0 and C will depend on the parameter $\sqrt{\varepsilon}$. It is simple to see that the dependence²³ of C on $\sqrt{\varepsilon}$ is $C = O(1/\sqrt{\varepsilon})$. The standard way in order to make negligible the rests in the development of the action functional, so that the Melnikov part in (0.21) dominates, is to take μ_0 exponentially small w.r.t. $\sqrt{\varepsilon}$ (namely $O(\exp(-c_2/\sqrt{\varepsilon}))$ with $c_2 > c_1$). Finally we are able to prove diffusion for the a-priori unstable Hamiltonian (0.22), finding a diffusion time T_d exponentially long w.r.t. $\sqrt{\varepsilon}$ as in [26].

²² Perform former a $\sqrt{\varepsilon}$ rescaling in the action variable $(\bar{I}, \bar{\varphi}, \bar{p}, \bar{q}) = (I/\sqrt{\varepsilon}, \varphi, p/\sqrt{\varepsilon}, q)$, which casts the Hamiltonian (0.22) in

$$\tilde{H}(\bar{I}, \bar{\varphi}, \bar{p}, \bar{q}, t) := \sqrt{\varepsilon} \left[\bar{I}^2/2 + \bar{p}^2/2 + (\cos \bar{q} - 1) + \mu(\cos \bar{q} - 1)[\sin \bar{\varphi} + \cos t] \right]$$

and later a $\sqrt{\varepsilon}$ rescaling of the time.

²³ Practically C depends linearly on the length of the embedding γ in Theorem 0.4.

We conclude this remark briefly mention two possible developments: (i) try to combine our method with the one used in [23] in order to deal with the very difficult and interesting problem of taking $\mu = \varepsilon^p$ for p positive; (ii) in order to find non existence results for KAM tori as in [29], try to use our method (which looks like more sensible).

(Optimality of our diffusion time estimate) We now state our stability result (see Theorem 10.3).

Theorem 0.6 *Let $f(I, \varphi, p, q, t)$ be as in (0.18), where the $f_{n,l}$ ($|(n, l)| \leq N$) are analytic functions. Then $\forall \kappa, \bar{\tau}, \tilde{\tau} > 0$ there exist $\mu_1, \kappa_0 > 0$ such that $\forall 0 < \mu \leq \mu_1$, for any solution $(I(t), \varphi(t), p(t), q(t))$ of (\mathcal{S}_μ) with $|I(0)| \leq \bar{\tau}$ and $|p(0)| \leq \tilde{\tau}$, there results*

$$|I(t) - I(0)| \leq \kappa \quad \forall t \text{ such that } |t| \leq \frac{\kappa_0}{\mu} \ln \frac{1}{\mu}. \quad (0.24)$$

Actually the proof of Theorem 0.6 contains much more information: in particular the stability time (0.24) is sharp only for orbits lying close to the separatrices. On the other hand the orbits lying far away from the separatrices are much more stable, namely exponentially stable in time according to Nekhoroshev type time estimates, see (16.4) and (16.11). Indeed the diffusion orbit of Theorem 0.4 is found close to some pseudo-diffusion orbit whose (q, p) variables move along the separatrices of the pendulum.

We now sketch the proof of Theorem 0.6. First we prove stability in the region “far from the separatrices of the pendulum” $\mathcal{E}_1 := \{(I, \varphi, q, p) \mid |E(q, p)| \geq \mu^c\}$, where $E(q, p) := p^2/2 + (\cos q - 1)$ and c is a suitable positive constant. In \mathcal{E}_1 we can write Hamiltonian \mathcal{H}_μ in action-angle variables (I, φ, Q, P, t) where $Q := Q(q, p)$ and $P := P(q, p)$ are the action-angle variables of the pendulum $E(q, p)$, i.e. $E(q(Q, P), p(Q, P)) := K(P)$. In these new variables the new Hamiltonian writes $\mathcal{H}_1 := I^2/2 + K(P) + \mu f_1(\varphi, Q, t, P, I)$, that, by Proposition 2.1, is analytic with an analyticity radius $r_1 \approx \mu^c$ (when μ goes to zero, the region \mathcal{E}_1 approximate closer and closer to the separatrices and the analyticity estimate deteriorates). It turns out that \mathcal{H}_1 is steep (actually for E positive it is even *quasi-convex*, see [93]) and then, for $c > 0$ small enough, we can apply the Nekhoroshev Theorem as proved in [90]. In this way we obtain exponential stability in the whole region \mathcal{E}_1 , i.e.

$$|I(t) - I(0)| \leq \text{const.} \mu^b \quad \forall |t| \leq T := \text{const.} \frac{1}{\mu} \exp\left(\frac{1}{\mu}\right)^a$$

for two constants $a, b > 0$. Finally we study the behavior of an orbit close to the separatrices of the pendulum, namely in the region \mathcal{E}_1^c . Roughly speaking, such an orbit will spend alternatively a time $T_S = O(|\ln \mu|)$ into a small $O(1)$ -neighborhood of $(p, q) = (0, 0)$ and a time $T_F = O(1)$ outside. In this second case we directly obtain $\Delta I := \int_s^{s+T_F} \dot{I}_\mu = O(\mu T_F) = O(\mu)$. Then we prove, roughly speaking, that $\Delta I = O(\mu)$ also in the first case. This result is obtained first writing the pendulum in hyperbolic variables in a small neighborhood of the origin and then performing one step of classical perturbation theory (with a resonant normal form) and an analysis of the resonances of Nekhoroshev-type.

Scheme of the thesis

Part I

We discuss some preliminaries regarding Normal Form Theory. We construct complex action–angle variables for a pendulum with non–constant gravity as presented in [33], [35]. We study the Hamiltonian formalism for D’Alembert planetary model finding the “effective Hamiltonian” recalling the results showed in [34].

Part II

We prove total stability for the action variables in the Hamiltonian (0.5) as proved in [33].

Part III

We prove stability of the action variables of the planetary D’Alembert model for exponentially long time in the perturbative parameters. This result was announced in [32] and proved in [35].

Part IV

We prove the existence of action–instability with diffusion time $T_d = O((1/\mu) \log(1/\mu))$ for the non-isochronous, nearly integrable, a-priori unstable Hamiltonian (0.17). We also prove that our estimate of the diffusion time T_d is optimal as a consequence of a general stability result derived from classical perturbation theory. This results were announced in [18] and proved in [19].

Appendix

We prove some technical lemmas used in part IV. In particular we prove a result on the ergodization time of the torus for linear flows possibly resonant but only at a “sufficiently high order”.

List of problems

We close this introduction with a list of problems regarding:

(a) Properly–degenerate Hamiltonian systems in two degrees of freedom

a.1 Find general conditions on H_0 given in (0.11) under which Theorem 0.2 holds (see [36]).

a.2 Extend the example in Remark 0.1 to properly–degenerate systems²⁴. Such examples may indicate a possible route to $O(1)$ –drift of action variables, in properly degenerate systems, *different* from Arnold Diffusion, as we have discussed above with the presence of some “resonance phenomena” (see also point (b.2) below).

²⁴ In particular in remark 4.1 (i) we give such an example with

$$H(I_1, I_2, \varphi_1, \varphi_2; \varepsilon) := H_{00}(I_1) + \varepsilon H_{01}(I_2) + \varepsilon^2 H_{1,j}(\varphi_1, \varphi_2) := \frac{I_1^2}{2} - \varepsilon \frac{I_2^2}{2} + \varepsilon^2 e^{-j} \sin(\varphi_1 - j\varphi_2)$$

when $\varepsilon := \varepsilon_j := j^{-2}$. It could be interesting to extend it to ε –independent perturbations H_1 or to $H_0 := H_{00} + \varepsilon H_{01}$ dependent also on φ_1 .

(b) D'Alembert planetary model

b.1 The major open problem in this context is certainly to prove action–instability (see remark 7.1 where the problem of Arnold diffusion in this context is considered) and, eventually, to compare the diffusion time with the stability times given in Theorem 0.3.

b.2 Prove (or disprove) the non–existence of “instability channels” for the two–degrees–of–freedom system (0.10) obtained by disregarding the exponentially small terms in (0.6).

b.3 Refine the techniques used in the proof of Theorem 0.3 so as to apply the stability result to some concrete problem arising, for example, in the Solar System using astronomical data and compare the obtained result with the existing estimates on the life time for the Solar System.

b.4 Generalize the results of Theorem 0.3 to, e.g., the case of a non–rotational planet (in which case, the system has one more degree of freedom), the case of a non–fixed star, etc.

(c) Possible applications of our variational methods for proving action–instability²⁵

c.1 Suitably refine our variational method in order to prove the existence of drifting orbits in the whole action space and then to prove such results for generic analytic perturbations too.

c.2 Try to apply these methods to infinite dimensional Hamiltonian systems (PDE, wave equations, etc...), where the existence of “transition chains of infinite dimensional hyperbolic tori” is quite far for being proved.

²⁵ This method is a further step of a research line, started in [22]–[23] and [24], for finding new mechanisms to prove Arnold diffusion.

Part I

Preliminaries

1 Notations, norms and Normal Form Lemma

We shall use the following notations: if $\emptyset \neq D \subset \mathbb{R}^d$ and $\rho := (\rho_1, \rho_2, \dots, \rho_d)$ with $0 < \rho_j \leq \infty$ for $1 \leq j \leq d$, we denote

$$D_\rho := \{I = (I_1, \dots, I_d) \in \mathbb{C}^d : |I_j - \bar{I}_j| < \rho_j, j = 1, \dots, d, \text{ for some } \bar{I} \in D\};$$

\mathbb{T}_σ^d denotes the complex set $\{z \in \mathbb{C}^d : |\text{Im} z_j| < \sigma, j = 1, \dots, d\}$ (thought of as a complex neighborhood of \mathbb{T}^d).

We shall work in the Banach space $\mathcal{H}_\mathbb{R}(D_\rho \times \mathbb{T}_\sigma^d)$ of functions f real-analytic on $D_\rho \times \mathbb{T}_\sigma^d$ having finite norm

$$\|f\|_{\rho, \sigma} := \sum_{k \in \mathbb{Z}^d} \sup_{I \in D_\rho} |\hat{f}_k(I)| e^{|k|\sigma},$$

$\hat{f}_k(I)$ being the Fourier coefficients of the periodic function $\varphi \rightarrow f(I, \varphi)$. Notice that if $f \in \mathcal{H}_\mathbb{R}(D_\rho \times \mathbb{T}_\sigma^d)$ and $\rho_j = \infty$ for some $1 \leq j \leq d$, then, by Liouville Theorem, f does not depend on I_j .

We shall use the following standard result from normal form theory; see [93] for the proof with $\rho_1 = \dots = \rho_d$; (for the trivial modifications in the case of different analyticity radii, see [33] or [34]).

Lemma 1.1 (Normal Form Lemma) *Let $H := H(I, \varphi) := h(I) + f(I, \varphi)$ be a real-analytic Hamiltonian on $D \times \mathbb{T}^d$ belonging to $\mathcal{H}_\mathbb{R}(D_\rho \times \mathbb{T}_\sigma^d)$ for certain $\rho := (\rho_1, \rho_2, \dots, \rho_d)$ with $0 < \rho_j \leq \infty$, $1 \leq j \leq d$, and $\sigma > 0$. Let $\rho_0 := \min_{1 \leq j \leq d} \rho_j$. Let Λ a sub-lattice of \mathbb{Z}^d , $\alpha > 0$, $K \in \mathbb{N}$ with $K\sigma \geq 6$. Suppose that*

$$|h'(I) \cdot k| \geq \alpha, \quad \forall k \in \mathbb{Z}^d \setminus \Lambda, \quad 0 < |k| \leq K, \quad \forall I \in D_\rho, \quad (1.1)$$

and that the following condition is satisfied:

$$\|f\|_{\rho, \sigma} =: \eta \leq \frac{\alpha \rho_0}{2^{10} K}. \quad (1.2)$$

Then, there exist a real-analytic symplectic transformation

$$\phi : (J, \psi) \in D_{\rho/2} \times \mathbb{T}_{\sigma/6}^d \mapsto (I, \varphi) = \phi(J, \psi) \in D_\rho \times \mathbb{T}_\sigma^d$$

such that

$$H \circ \phi(J, \psi) = h(J) + \sum_{k \in \Lambda, |k| \leq K} f_k(J) e^{ik \cdot \psi} + g(J, \psi) + f_*(J, \psi) \quad (1.3)$$

where $f_* := f_*(J, \psi)$, $g := g(J, \psi) := \sum_{k \in \Lambda} g_k(J) e^{ik \cdot \psi}$ belong to $\mathcal{H}_\mathbb{R}(D_{\rho/2} \times \mathbb{T}_{\sigma/6}^d)$. Moreover the following estimates hold:

$$\begin{aligned} \|g\|_{\rho/2, \sigma/6} &\leq \frac{2^{11}}{\alpha \rho_0 \sigma} \eta^2 \leq \frac{2}{K\sigma} \eta \leq \frac{1}{3} \eta, \\ \|f_*\|_{\rho/2, \sigma/6} &\leq \eta e^{-K\sigma/6}, \\ |I - J| &\leq \frac{2^5}{\alpha \sigma} \eta \leq \frac{\rho_0}{2^7}, \quad |\varphi - \psi| \leq \frac{2^6}{\alpha \rho_0} \eta \leq \frac{\sigma}{2^5}, \quad \forall (J, \psi) \in D_{\rho/2} \times \mathbb{T}_{\sigma/6}^d. \end{aligned} \quad (1.4)$$

Remark 1.1 If we have $h := h(I) := \hat{h}(I_1, \dots, I_{d-1}) + \omega I_d$ and $f := f(I_1, \dots, I_{d-1}, \varphi)$, then the symplectic transformation ϕ preserves the form of the Hamiltonian and has the form:

$$\begin{cases} I_j = \tilde{I}_j(J_1, \dots, J_{d-1}, \psi) , & \varphi_j = \tilde{\varphi}_j(J_1, \dots, J_{d-1}, \psi) , \\ I_d = J_d + \tilde{I}_d(J_1, \dots, J_{d-1}, \psi) , & \varphi_d = \psi_d , \end{cases} \quad (1 \leq j \leq d-1) ,$$

and, also, f_* and g do not depend on J_d .

2 Complex action-angle variables

In this Section we will construct complex action-angle variables for the “suspended pendulum” with energy $E(p, I, q) := p^2/2 - \varepsilon k(I)(1 + \cos q)$, where p is the action conjugated to the angle $q \in \mathbb{T}$ and $\varepsilon k(I)$ is a small gravity varying with the parameter I according to the values of the strictly positive function $k(I)$.

Of course such action-angle variables will be *singular* near the separatrices and near elliptic periodic orbits (corresponding to the equilibria of the pendulum) and therefore a careful blow-up analysis near the singularities, as $\varepsilon \rightarrow 0$, is needed.

Furthermore we also need to study the *complex* analytic continuation of the action-angle variables since we will apply later Normal Form and KAM theory in real-analytic class. Notwithstanding the enormous literature on elliptic integrals, such information does not seem to be available. So, in order to perform this blow-up in analytic class, a certain amount of straightforward (although rather lengthy) computations are needed. This will be done below.

Throughout this section, we shall denote z_1 and z_2 , respectively, the real and imaginary part of a complex number $z = z_1 + iz_2$.

In the following we will choose the positive branch of the square root i.e. if $z = |z|e^{i\alpha}$ with $\alpha \in (-\pi, \pi)$, then we define the analytic function $\sqrt{z} := \sqrt{|z|}e^{i\alpha/2}$. We also define $\ln z := \ln |z| + i\alpha$ and $\arccos z := -i \ln(z + i\sqrt{1 - z^2})$.

We shall use the following notation: if A, B are two strictly positive functions we shall say $A \sim B$ if there exist positive constants c^+, c^- so that $c^- A \leq B \leq c^+ A$ pointwise. For example,

$$\sqrt{A+B} - \sqrt{A} = \frac{B}{\sqrt{A+B} + \sqrt{A}} \sim \frac{B}{\sqrt{A+B}}. \quad (2.1)$$

Obviously, “ \sim ” is transitive. Also, if A, B and C are strictly positive, then $A \sim B$ implies $A + C \sim B + C$.

We need the following elementary lemma:

Lemma 2.1 *Let $x_1, x_2 \geq 0$. Then $\sqrt{x_1 \pm ix_2} = w_1 \pm iw_2$ and $(x_1 \pm ix_2)^{-1/2} = y_1 \mp iy_2$ where*

$$\begin{aligned} w_1(x_1, x_2) &:= \frac{1}{\sqrt{2}} \sqrt{x_1 + \sqrt{x_1^2 + x_2^2}}, & w_2(x_1, x_2) &:= \frac{1}{\sqrt{2}} \sqrt{-x_1 + \sqrt{x_1^2 + x_2^2}}, \\ y_1(x_1, x_2) &:= \frac{\sqrt{x_1 + \sqrt{x_1^2 + x_2^2}}}{\sqrt{2}\sqrt{x_1^2 + x_2^2}}, & y_2(x_1, x_2) &:= \frac{\sqrt{-x_1 + \sqrt{x_1^2 + x_2^2}}}{\sqrt{2}\sqrt{x_1^2 + x_2^2}}. \end{aligned}$$

We observe that, for x_1 fixed, w_1 (resp. y_1) is increasing (resp. decreasing) for $x_2 \geq 0$; y_2 is also increasing but only for $x_2 \leq \sqrt{3}x_1$. Moreover if $x_1 \geq x_2$ the following estimates hold

$$\begin{aligned} \sqrt{x_1} \leq w_1 \leq \sqrt{x_1 + x_2} \leq \sqrt{2}\sqrt{x_1}, & \quad \frac{1}{3} \frac{x_2}{\sqrt{x_1}} \leq w_2 \leq \frac{1}{\sqrt{2}} \frac{x_2}{\sqrt{x_1}}, \\ \frac{1}{\sqrt{2}} \frac{1}{\sqrt{x_1}} \leq \frac{1}{\sqrt{x_1 + x_2}} \leq y_1 \leq \frac{\sqrt{2}}{\sqrt{x_1}}, & \quad \frac{1}{4} \frac{x_2}{x_1^{3/2}} \leq y_2 \leq \frac{1}{\sqrt{2}} \frac{x_2}{x_1^{3/2}}. \end{aligned} \quad (2.2)$$

Finally $(x_1 \pm ix_2)^{-3/2} = z_1 \mp iz_2$ where $z_1 = y_1(y_1^2 - 3y_2^2)$, $z_2 = y_2(3y_1^2 - y_2^2)$ and y_1, y_2 are as above. Furthermore, if $y_1 \geq 2y_2$ then $y_1^2 - 3y_2^2 \sim y_1^2$, $3y_1^2 - y_2^2 \sim y_1^2$ and

$$z_1 \sim \frac{1}{(|x_1| + x_2)^{3/2}}, \quad z_2 \sim \frac{x_2}{(|x_1| + x_2)^{5/2}} \quad (2.3)$$

Proposition 2.1 *Let $k(I)$ real-analytic on $\Delta^0 := [\alpha, \beta]$ for $\alpha < \beta$ with analytic extension on $\Delta_{r_2}^0$ for $r_2 > 0$. Let*

$$\bar{k} := \max_{\bar{I}_1 \in \Delta^0} k(\bar{I}_1), \quad \hat{k} := \min_{\bar{I}_1 \in \Delta^0} k(\bar{I}_1), \quad \bar{k}' := \max_{I \in \Delta_{r_2}^0} |k'(I)|,$$

and suppose that $\hat{k} > 0$. Let $\varepsilon > 0$, $\eta > 0$, $R_0 \geq 2r_1 > 0$, $0 < s_1 \leq 1$, $s_2 > 0$, $D^0 := [-R_0, R_0]$,

$$E(p, I, q) := p^2/2 - \varepsilon k(I)(1 + \cos q)$$

and²⁶

$$\begin{aligned} \mathcal{M}^+ &:= \{(p, I, q, \varphi) \in [0, R_0] \times \Delta^0 \times \mathbb{T}^2 : \eta \leq E(p, I, q) \leq R_0^2/2\} \\ \mathcal{M}^- &:= \{(p, I, q, \varphi) \in D^0 \times \Delta^0 \times \mathbb{T}^2 : -2\varepsilon k(I) + \eta \leq E(p, I, q) \leq -\eta\}. \end{aligned}$$

Then, there exist positive constants c_1, c_2 (sufficiently large and depending only on $\bar{k}, \hat{k}, \bar{k}'$) and $c_3, c_4, c_5, c_6, \varepsilon_0$ (sufficiently small) such that if $\varepsilon \leq \varepsilon_0$, $R_0 \geq c_1\sqrt{\varepsilon}$, $r_1 \geq c_2\sqrt{\varepsilon}$, $\eta \leq c_3\varepsilon$ and

$$\rho_1 := c_4 \frac{\eta}{\sqrt{\varepsilon}}, \quad \rho_2 := \min\left\{c_5 \frac{\eta}{\varepsilon \ln(\varepsilon/\eta)}, r_2\right\}, \quad \sigma_1 := c_6 \frac{s_1}{\ln(\varepsilon/\eta)}, \quad \sigma_2 := \frac{s_2}{2},$$

²⁶ An analogous statement holds for $p \in [-R_0, 0]$. Of course, we are assuming that the various parameters are chosen so that $\mathcal{M}^\pm \neq \emptyset$.

then the following holds. There exist two real-analytic symplectic transformations ϕ^\pm and two real-analytic functions $E^\pm : \Omega_{(\rho_1, \rho_2)}^\pm \rightarrow \mathbb{C}^2$ so that

$$\begin{aligned} \phi^\pm : (P, J, Q, \psi) &\in \Omega_{(\rho_1, \rho_2)}^\pm \times \mathbb{T}_{\sigma_1} \times \mathbb{T}_{\sigma_2} \rightarrow (p, I, q, \varphi) \in D_{r_1}^0 \times \Delta_{r_2}^0 \times \mathbb{T}_{s_1} \times \mathbb{T}_{s_2}, \\ p &= p^\pm(P, J, Q), \quad I = J, \quad q = q^\pm(P, J, Q), \quad \varphi = \varphi^\pm(P, J, Q, \psi), \\ E^\pm(P, J) &= E(p^\pm(P, J, Q), J, q^\pm(P, J, Q)), \\ \phi^\pm(\Omega^\pm \times \mathbb{T}^2) &= \mathcal{M}^\pm, \end{aligned} \tag{2.4}$$

where:

$$\begin{aligned} \Omega^\pm &:= \{(P_1, I_1) \in \mathbb{R}^2 \quad \text{s.t.} \quad P_1 \in D^\pm(I_1), \quad I_1 \in \Delta^0\}, \\ D^+(I_1) &:= (P^+(\eta, I_1), P^+(R_0^2/2, I_1)), \quad D^-(I_1) := (P^-(-2\epsilon k(I_1) + \eta, I_1), P^-(-\eta, I_1)), \\ P^+ &:= \frac{\sqrt{2}}{\pi} \int_0^\pi \sqrt{g(E, I, \theta)} d\theta, \quad P^- := \frac{2\sqrt{2}}{\pi} \int_0^{\psi_0(E, I)} \sqrt{g(E, I, \theta)} d\theta, \end{aligned}$$

and

$$g(E, I, \theta) := E + \epsilon k(I)(1 + \cos \theta), \quad \psi_0(E, I) := \arccos(-1 - E/\epsilon k(I)).$$

Moreover, the following estimates hold for $(P, I) \in \Omega_{(\rho_1, \rho_2)}^\pm$:

$$|\partial_P E^\pm(P, I)| \geq \frac{1}{2\sqrt{2}\pi} \frac{\sqrt{\epsilon k_1(I)}}{\ln\left(1 + \sqrt{\frac{\epsilon k_1(I)}{|E_1^\pm(P, I)|}}\right)} \geq \frac{1}{4\pi} \frac{\sqrt{\epsilon \hat{k}}}{\ln\left(1 + \sqrt{\frac{\epsilon \hat{k}}{\eta}}\right)}, \tag{2.5}$$

$$|\partial_I E^\pm(P, I)| \leq \frac{\sqrt{4\sqrt{6}\pi}|k'(I)|\epsilon}{\ln\left(1 + \sqrt{\frac{\epsilon k_1(I)}{|E_1^\pm(P, I)|}}\right)}. \tag{2.6}$$

Finally, since $\partial_I E^\pm(P, I) = -\partial_I P^\pm(E^\pm(P, I), I)[\partial_E P^\pm(E^\pm(P, I), I)]^{-1}$, we can write, for real P and I ,

$$\partial_I E^\pm(P, I) = -\epsilon k(I)[1 + Y^\pm(P, I)] \quad \text{with} \quad |Y^\pm(P, I)| < 1 \tag{2.7}$$

where

$$Y^\pm(P, I) := \left[\int_0^{\psi^\pm} \frac{1}{\sqrt{g(E^\pm(P, I), I, \theta)}} d\theta \right]^{-1} \int_0^{\psi^\pm} \frac{\cos \theta}{\sqrt{g(E^\pm(P, I), I, \theta)}} d\theta$$

with $\psi^+ := \pi$ and $\psi^- := \psi_0(E^-(P, I), I)$.

In particular if $k(I) \equiv 1$ we have that $p = p^\pm(P, Q)$, $q = q^\pm(P, Q)$, $\varphi = \psi$, $P^\pm = P^\pm(E)$ and $E^\pm = E^\pm(P)$ and the following estimates on P^\pm and its derivatives hold

$$\begin{aligned} P_1^\pm(E) &\sim \frac{\tilde{E}_1}{\sqrt{\epsilon}}, \quad P_2^\pm(E) \sim \frac{E_2}{\sqrt{\epsilon}} \ln\left(1 + \sqrt{\frac{\epsilon}{|E_1|}}\right), \\ \dot{P}_1^\pm(E) &\sim \frac{1}{\sqrt{\epsilon}} \ln\left(1 + \sqrt{\frac{\epsilon}{|E_1|}}\right), \quad \dot{P}_2^\pm(E) \sim \frac{E_2}{|E_1|\sqrt{\epsilon}}, \\ \ddot{P}_1^\pm(E) &\sim \frac{1}{|E_1|\sqrt{\epsilon}}, \quad \ddot{P}_2^\pm(E) \sim \frac{E_2}{E_1^2\sqrt{\epsilon}}. \end{aligned} \tag{2.8}$$

where with “ \cdot ” we have denoted the derivative with respect to E .

Proof

• (First step: estimates on the action domains)

Let $\bar{k}_1 := \sup_{I \in \Delta_{\rho_2}^0} |k_1(I)|$, $\bar{k}_2 := \sup_{I \in \Delta_{\rho_2}^0} |k_2(I)|$, $\bar{a}_1 := \bar{k}_1(1 + \cosh s_1) + \bar{k}_2 \sinh s_1$ and $\bar{a}_2 := \bar{k}_2(1 + \cosh s_1) + \bar{k}_1 \sinh s_1$.

For suitable $c_7, c_8 > 0$ small enough, we define $E_2^*(E_1) := c_7 \eta \ln^{-1}(1 + \sqrt{\frac{\varepsilon}{|E_1|}})$, $\bar{E} := c_8 r_1(R_0 + r_1) + R_0^2/2$, and, for $\bar{I}_1 \in \Delta^0$, we define the following domains²⁷

$$\begin{aligned} \mathcal{E}^+ &:= \mathcal{E}^+(\bar{I}_1) := \{E_1 + iE_2, \text{ s.t. } \eta/2 \leq E_1 \leq \bar{E}; |E_2| \leq E_2^*(E_1)\}, \\ \mathcal{E}^- &:= \mathcal{E}^-(\bar{I}_1) := \{E_1 + iE_2, \text{ s.t. } -2\varepsilon k(\bar{I}_1) + \eta/2 \leq E_1 \leq -\eta/2; |E_2| \leq E_2^*(E_1)\}. \end{aligned}$$

Let $F(p) := p^2/2$. We claim that

$$|I - \bar{I}_1| \leq \rho_2, \bar{I}_1 \in \Delta^0, E \in \mathcal{E}^+ \cup \mathcal{E}^-(\bar{I}_1), \theta \in \mathbb{T}_{s_1} \implies g(E, I, \theta) \in F(D_{r_1}^0). \quad (2.9)$$

It is immediate to see that $F(D_{r_1}^0) \supseteq \hat{\mathcal{E}}$, where

$$\begin{aligned} \hat{\mathcal{E}} &:= \{-r_1^2 \leq 2E_1 \leq R_0^2 - r_1^2, |E_2| \leq r_1 \sqrt{2E_1 + r_1^2}\} \\ &\quad \cup \{R_0^2 - r_1^2 \leq 2E_1 \leq (R_0 + r_1)^2, |E_2| \leq \hat{E}_2(E_1)\} \end{aligned}$$

and $\hat{E}_2(E_1) := [-E_1 + (R_0 + r_1)^2/2]R_0/(R_0 + r_1)$.

Next, we define

$$\tilde{\mathcal{E}} := \{-2\varepsilon \bar{k}_1 \leq E_1 \leq 2\varepsilon, |E_2| \leq E_2^0\} \cup \{2\varepsilon < E_1 \leq \bar{E}, |E_2| \leq 2c_7 \eta \sqrt{E_1/\varepsilon}\},$$

where $E_2^0 := \max\{E_2^*(2\varepsilon), E_2^*(2\varepsilon \bar{k}_1)\}$, and $\bar{\mathcal{E}} := \bar{\mathcal{E}}^{(1)} \cup \bar{\mathcal{E}}^{(2)} \cup \bar{\mathcal{E}}^{(3)}$ with:

$$\begin{aligned} \bar{\mathcal{E}}^{(1)} &:= \{-2\varepsilon \bar{k}_1 - \varepsilon \bar{a}_1 \leq E_1 \leq 2\varepsilon - \varepsilon \bar{a}_1, |E_2| \leq E_2^0 + \varepsilon \bar{a}_2\}, \\ \bar{\mathcal{E}}^{(2)} &:= \{2\varepsilon - \varepsilon \bar{a}_1 < E_1 \leq \bar{E} - \varepsilon \bar{a}_2, |E_2| \leq 2c_7 \eta \sqrt{(E_1 + \varepsilon \bar{a}_1)/\varepsilon} + \varepsilon \bar{a}_1\}, \\ \bar{\mathcal{E}}^{(3)} &:= \{\bar{E} - \varepsilon \bar{a}_1 < E_1 \leq \bar{E} + \varepsilon \bar{a}_1, |E_2| \leq 2c_7 \eta \sqrt{\bar{E}/\varepsilon} + \varepsilon \bar{a}_2\}. \end{aligned}$$

We observe that obviously $\tilde{\mathcal{E}} \subseteq \bar{\mathcal{E}}$; moreover (recalling the definitions of \bar{a}_1, \bar{a}_2)

$$|I - \bar{I}_1| \leq \rho_2, \bar{I}_1 \in \Delta^0, E \in \tilde{\mathcal{E}}, \theta \in \mathbb{T}_{s_1} \implies g(E, I, \theta) \in \bar{\mathcal{E}}.$$

We observe that $E_1 \geq 2\varepsilon$ implies $\ln^{-1}(1 + \sqrt{\varepsilon/E_1}) \leq 2\sqrt{E_1/\varepsilon}$ and hence $\tilde{\mathcal{E}} \supseteq \mathcal{E}^+ \cup \mathcal{E}^-(\bar{I}_1)$.

Now we prove $\bar{\mathcal{E}} \subseteq \hat{\mathcal{E}}$ which will imply (2.9), since $F(D_{r_1}^0) \supseteq \hat{\mathcal{E}} \supseteq \bar{\mathcal{E}} \supseteq \tilde{\mathcal{E}} \supseteq \mathcal{E}^+ \cup \mathcal{E}^-(\bar{I}_1)$.

It is simple to see that $\bar{\mathcal{E}} \subseteq \hat{\mathcal{E}}$ is implied by the following conditions:

²⁷ In the positive energy case \mathcal{E}^+ does not really depend on \bar{I}_1 .

- (1) $(-2\varepsilon\bar{k}_1 - \varepsilon\bar{a}_1) + i(E_2^0 + \varepsilon\bar{a}_2) \in \hat{\mathcal{E}}$, which is implied by $r_1\sqrt{2(-2\varepsilon\bar{k}_1 - \varepsilon\bar{a}_1) + r_1^2} \geq E_2^0 + \varepsilon\bar{a}_2$
- (2) $(\bar{E} + \varepsilon\bar{a}_1) + i(2c_7\eta\sqrt{\bar{E}/\varepsilon} + \varepsilon\bar{a}_2)$, which is implied by $\hat{E}_2(\bar{E} + \varepsilon\bar{a}_1) \geq 2c_7\eta\sqrt{\bar{E}/\varepsilon} + \varepsilon\bar{a}_2$
- (3) if $(R_0^2 - r_1^2)/2 > 2\varepsilon - \bar{a}_1$ then $r_1\sqrt{2E_1 + r_1^2} \geq 2c_7\eta\sqrt{\bar{E}/\varepsilon} + \varepsilon\bar{a}_2$ for all $2\varepsilon - \bar{a}_1 \leq E_1 \leq \min\{(R_0^2 - r_1^2)/2, \bar{E} - \varepsilon\bar{a}_1\}$.

Defining $\tilde{k}_1 := \max\{1, \bar{k}_1\}$, one sees that (1) holds provided

$$c_2 \geq \max\{\sqrt{4c_7c_3\tilde{k}_1} + \sqrt{2\bar{a}_2}, 2\sqrt{2\bar{k}_1 + \bar{a}_1}\}.$$

If $c_8 \leq 1/2$, we have that $\hat{E}_2(\bar{E} + \varepsilon\bar{a}_1) \geq R_0(R_0r_1/2 - \varepsilon\bar{a}_1)/(R_0 + r_1)$ and hence (2) holds provided $c_2 \geq 16\sqrt{2}c_7c_3$ and $c_1c_2 \geq 4\bar{a}_1 + 8\bar{a}_2$.

Finally, conditions $c_2 \geq \sqrt{2\bar{a}_2}$ and $c_2 \geq 2c_7c_3$ imply (3).

- (Second step: estimates on the action derivatives in the positive energy case)

In the following we put $\epsilon := \varepsilon k_1(I)$. We observe that, for $\theta \in [0, \pi]$, $2\tilde{g}/\pi^2 \leq g_1 \leq \tilde{g}$, where $\tilde{g}(E_1, I; \theta) := E_1 + \varepsilon k_1(I)(\pi - \theta)^2$. The following estimates hold²⁸

$$\begin{aligned} \frac{1}{\sqrt{\epsilon}} \ln \left(1 + \sqrt{\frac{\epsilon}{E_1}} \right) &\leq \int_0^\pi \frac{d\psi}{\sqrt{\tilde{g}(\psi)}} = \int_0^\pi \frac{d\psi}{\sqrt{E_1 + \epsilon\psi^2}} = \frac{1}{\sqrt{\epsilon}} \int_0^a \frac{dy}{\sqrt{1 + y^2}} \\ &= \frac{1}{\sqrt{\epsilon}} \operatorname{arcsinh}(a) \leq \frac{2\pi}{\sqrt{\epsilon}} \ln \left(1 + \sqrt{\frac{\epsilon}{E_1}} \right); \\ \frac{1}{E_1\sqrt{E_1 + \epsilon}} &\leq \int_0^\pi \frac{d\psi}{(\tilde{g}(\psi))^{3/2}} = \int_0^\pi \frac{d\psi}{(E_1 + \epsilon\psi^2)^{3/2}} \\ &= \frac{1}{E_1\sqrt{\epsilon}} \int_0^a \frac{dy}{(1 + y^2)^{3/2}} = \frac{\pi}{E_1^{3/2}\sqrt{1 + \pi^2\epsilon/E_1}} \leq \frac{\pi}{E_1\sqrt{E_1 + \epsilon}}; \end{aligned} \quad (2.10)$$

where $a := \pi\sqrt{\epsilon/E_1}$.

Next, we prove that $\forall |I - \bar{I}_1| \leq \rho_2$, $\bar{I}_1 \in \Delta^0$, $E \in \mathcal{E}^+$ $g_1(E, I, \theta) \geq 2|g_2(E, I, \theta)|$. In fact we first have that if $c_5 \leq 1/8\bar{k}'$ then $g_1(E, I, \theta) \geq E_1/2$ and, hence, we have only to prove that $E_1/4 \geq E_2^*(E_1) + 2\bar{k}'c_5\eta \ln^{-1}(\varepsilon/\eta)$. Taking $c_5 \leq c_7/4\bar{k}'$ we have only to verify that

$$|E_1| \geq 6E_2^*(E_1). \quad (2.11)$$

It is easy to see that the previous inequality is verified for $c_7 \leq 1/36$.

Consider, now, $I = \bar{I}_1 \in \Delta^0$ real and $g_2 = E_2 \geq 0$. Using Lemma 2.1, $g_1 \geq \tilde{g}$ and (2.10) we have

$$P_2^+(E, \bar{I}_1) \geq \frac{\sqrt{2}}{3\pi} E_2 \int_0^\pi \frac{d\psi}{\sqrt{\tilde{g}(\psi)}} \geq \frac{\sqrt{2}}{3\pi} E_2 \frac{1}{\sqrt{\epsilon}} \ln \left(1 + \sqrt{\frac{\epsilon}{E_1}} \right). \quad (2.12)$$

²⁸Use $\ln(1 + t) \leq \operatorname{arcsinh}(t) = \ln(t + \sqrt{1 + t^2}) \leq 2\ln(1 + t)$

Using Lemma 2.1, $2\tilde{g}/\pi^2 \leq g_1 \leq \tilde{g}$ and (2.10) we have

$$\frac{1}{2\pi\sqrt{\epsilon}} \ln \left(1 + \sqrt{\frac{\epsilon}{E_1}} \right) \leq \partial_E P_1^+(E, I) \leq \frac{\sqrt{2}\pi}{\sqrt{\epsilon}} \ln \left(1 + \sqrt{\frac{\epsilon}{E_1}} \right). \quad (2.13)$$

Using Lemma 2.1, $2\tilde{g}/\pi^2 \leq g_1$, (2.10) and the fact that $|g_2| \leq |E_2| + 2\epsilon\bar{k}_2$ we have

$$|\partial_E P_2^+(E, I)| \leq (|E_2| + 2\epsilon\bar{k}_2) \frac{\pi^4}{4\sqrt{2} E_1 \sqrt{E_1 + \epsilon}}. \quad (2.14)$$

We observe that²⁹

$$\frac{|E_2|}{E_1} \leq \frac{c_7 \eta}{E_1 \ln(1 + \sqrt{\epsilon/E_1})} \leq c_7 \ln(1 + \sqrt{\epsilon/E_1}) \leq \frac{c_7}{\sqrt{k_1(I)}} \ln(1 + \sqrt{\epsilon/E_1}). \quad (2.15)$$

Let us proceed by proving that

$$\epsilon \bar{k}' \rho_2 \pi^3 = \bar{k}' c_5 \pi^3 \eta \leq \eta \ln(1 + \sqrt{2\epsilon/\eta}) \leq 2E_1 \ln(1 + \sqrt{\epsilon/E_1}). \quad (2.16)$$

In fact, last inequality holds because the function $E_1 \ln(1 + \sqrt{\epsilon/E_1})$ is increasing and attains minimum for $E_1 = \eta/2$; the first inequality is proved, if $c_3 \leq 1/8$ and using $k_1(I) \geq \hat{k}/2$, if $\bar{k}' c_5 \pi^3 \leq \ln(1 + 2\sqrt{2}\sqrt{\hat{k}})$, which is verified if $c_5 \leq (\pi^3 \sqrt{2\bar{k}}')^{-1} \min\{1, \hat{k}\}$. Using (2.13), (2.14), (2.15), (2.16) we have

$$|\partial_E P^+(E, I)| \leq \frac{2\sqrt{2}\pi}{\sqrt{\epsilon}} \ln \left(1 + \sqrt{\frac{\epsilon}{E_1}} \right). \quad (2.17)$$

It remains to estimate

$$|\partial_I P^+(E, I)| = \frac{\sqrt{2}}{2\pi} \epsilon |k'(I)| \left| \int_0^\pi \frac{1 + \cos \theta}{\sqrt{g(E, I; \theta)}} d\theta \right|.$$

We observe that

$$\left| \int_0^\pi \frac{1 + \cos \theta}{\sqrt{g(E, I; \theta)}} d\theta \right| = \left| \int_0^2 \frac{\sqrt{x}}{\sqrt{2-x} \sqrt{E + \epsilon k(I)x}} dx \right| \leq \int_0^2 F_1(x) dx,$$

where $F_1(x) := \sqrt{x}(\sqrt{2-x}\sqrt{E_1 + \epsilon x})^{-1}$. In order to estimate last integral we split it as

$$\int_0^2 F_1(x) dx = \int_0^1 F_1(x) dx + \int_1^2 F_1(x) dx.$$

We have

$$\int_0^1 F_1(x) dx \leq \int_0^1 \frac{\sqrt{x}}{\sqrt{E_1 + \epsilon x}} dx \leq \int_0^1 \frac{\sqrt{x}}{\sqrt{E_1 x + \epsilon x}} dx = \frac{1}{\sqrt{E_1 + \epsilon x}},$$

²⁹ We use the fact that $x \ln^2(1 + \sqrt{\epsilon/E_1}) \geq \eta$ if $x \geq \eta/2$ and $c_3 \leq 1/8$.

$$\int_1^2 F_1(x) dx \leq \frac{\sqrt{2}}{\sqrt{E_1 + \epsilon}} \int_1^2 \frac{1}{\sqrt{2-x}} dx = \frac{2\sqrt{2}}{\sqrt{E_1 + \epsilon}},$$

which implies

$$|\partial_I P^+(E, I)| \leq \frac{2\sqrt{2}}{\pi} |k'(I)| \frac{\epsilon}{\sqrt{E_1 + \epsilon}} \leq 2\sqrt{6} \frac{|k'(I)|}{\sqrt{k_1(I)}} \sqrt{\epsilon}. \quad (2.18)$$

We, now, prove that

$$P^+(\mathcal{E}^+) \supseteq (D^+(\bar{I}_1))_{2\rho_1} \quad \forall \bar{I}_1 \in \Delta^0. \quad (2.19)$$

Since $P_1^+(E_1 + iE_2, \bar{I}_1) = \frac{\sqrt{2}}{\pi} \int_0^\pi w_1(E_1 + \epsilon k(\bar{I}_1)(1 + \cos \theta), E_2) d\theta$, we have by Lemma 2.1 that P_1^+ is an increasing function for $E_2 \geq 0$. Hence, in order to prove (2.19), we have to prove the following estimates, $\forall E_1 + iE_2 \in \mathcal{E}^+, \bar{I}_1 \in \Delta^0$:

- (i) $P_2^+(E_1 + iE_2^*(E_1), \bar{I}_1) \geq 2\rho_1$,
- (ii) $P_1^+(\eta/2 + iE_2^*(\eta/2), \bar{I}_1) \leq P_1^+(3\eta/4, \bar{I}_1)$,
- (iii) $P_1^+(\eta, \bar{I}_1) - P_1^+(3\eta/4, \bar{I}_1) \geq 2\rho_1$,
- (iv) $P_1^+(\bar{E}, \bar{I}_1) - P_1^+(R_0^2/2, \bar{I}_1) \geq 2\rho_1$.

If $c_4 \leq \frac{\sqrt{2}}{6\pi} c_7 \min\{1, \frac{1}{k}\}$ we obtain (i), since, from (2.12),

$$P_2^+(E_1 + iE_2^*(E_1), \bar{I}_1) \geq c_7 \frac{\sqrt{2}\eta}{3\pi\sqrt{\epsilon}} \frac{\ln(1 + \sqrt{k(\bar{I}_1)}\sqrt{\epsilon/E_1})}{\sqrt{k(\bar{I}_1)} \ln(1 + \sqrt{\epsilon/E_1})} \geq c_7 \frac{\sqrt{2}\eta}{3\pi\sqrt{\epsilon}} \min\{1, \frac{1}{k}\}.$$

Inequality (ii) follows from

$$\begin{aligned} P_1^+(\eta/2 + iE_2^*(\eta/2), \bar{I}_1) &\leq (\sqrt{2}/\pi) \int_0^\pi \sqrt{\eta/2 + E_2^*(E_1) + \epsilon k(\bar{I}_1)(1 + \cos \theta)} d\theta \\ &\leq (\sqrt{2}/\pi) \int_0^\pi \sqrt{3\eta/4 + \epsilon k(\bar{I}_1)(1 + \cos \theta)} d\theta = P_1^+(3\eta/4, \bar{I}_1). \end{aligned}$$

Using (2.13) and the fact that $\eta \leq \epsilon/8$, we have³⁰

$$P_1^+(\eta, \bar{I}_1) - P_1^+(3\eta/4, \bar{I}_1) \geq \frac{1}{8\pi} \frac{\eta}{\sqrt{\epsilon}} \left[\frac{1}{\sqrt{k}} \ln(1 + 2\sqrt{k}) \right] \geq \frac{1}{8\pi} \frac{\eta}{\sqrt{\epsilon}} \min\{1, \frac{1}{k}\},$$

which implies (iii), provided $c_4 \leq \frac{1}{16\pi} \min\{1, \frac{1}{k}\}$.

Again, from (2.13), we have

$$P_1^+(\bar{E}, \bar{I}_1) - P_1^+(R_0^2/2, \bar{I}_1) \geq \frac{c_8}{\pi} r_1(R_0 + r_1) \frac{1}{\sqrt{\bar{E}}} \min \left\{ 1, 2\sqrt{\frac{\bar{E}}{\epsilon k(\bar{I}_1)}} \right\}.$$

³⁰ Use $\ln(1 + 2x)/x \geq \min\{1, 1/x\}$.

Distinguishing the two cases for the minimum, we see that (iv) holds, provided $c_8 c_2 \geq \sqrt{2} \pi c_4 c_3$ and $c_8 c_2 c_1 \geq \sqrt{k} \pi c_4 c_3$.

Now we consider the case in which $K(I) \equiv 1$ in order to prove (2.8). The estimate on P_1^+ is trivial, the ones on P_2^+ and \dot{P}_1^+ are consequence of³¹ (2.12) and (2.13) respectively; for the other ones we have, using (2.2) and (2.3),

$$\begin{aligned}
-\ddot{P}_1^+(E) &\sim \int_0^\pi \frac{d\psi}{(\tilde{g}(\psi))^{3/2}} \sim \int_0^\pi \frac{d\psi}{(E_1 + \varepsilon(\pi - \psi)^2)^{3/2}} \\
&= \frac{1}{E_1 \sqrt{\varepsilon}} \int_0^\pi \sqrt{\frac{\varepsilon}{E_1}} \frac{dy}{(1 + y^2)^{3/2}} = \frac{\pi}{E_1^{3/2} \sqrt{1 + \pi^2 \varepsilon / E_1}} \sim \frac{1}{E_1 \sqrt{E_1 + \varepsilon}}, \\
-\dot{P}_2^+(E) &\sim E_2 \int_0^\pi \frac{d\psi}{(\tilde{g}(\psi))^{3/2}} \sim \frac{E_2}{E_1 \sqrt{E_1 + \varepsilon}}, \\
-\ddot{P}_2^+(E) &\sim E_2 \int_0^\pi \frac{d\psi}{(\tilde{g}(\psi))^{5/2}} \sim E_2 \int_0^\pi \frac{d\psi}{(E_1 + \varepsilon(\pi - \psi)^2)^{5/2}} \\
&= \frac{E_2}{E_1 \sqrt{\varepsilon}} \int_0^\pi \sqrt{\frac{\varepsilon}{E_1}} \frac{dy}{(1 + y^2)^{5/2}} \sim \frac{E_2}{E_1^2 \sqrt{E_1 + \varepsilon}}
\end{aligned}$$

- (Third step: estimates on the action derivatives in the negative energy case)

Defining $\tilde{E} := E + 2\varepsilon k(I)$ and using the substitution $\xi := 1 + \varepsilon k(I)(\cos \theta - 1)/\tilde{E}$ we obtain

$$\begin{aligned}
P^-(E, I) &= \frac{2\sqrt{2}}{\pi} \int_0^1 \frac{\tilde{E} \sqrt{\xi}}{\sqrt{1 - \xi} \sqrt{\tilde{E} \xi - E}} d\xi, \\
\partial_E P^-(E, I) &= \frac{\sqrt{2}}{\pi} \int_0^1 \frac{1}{\sqrt{\xi} \sqrt{1 - \xi} \sqrt{\tilde{E} \xi - E}} d\xi, \\
\partial_I P^-(E, I) &= \frac{\sqrt{2} k'(I)}{\pi k(I)} \int_0^1 \frac{\sqrt{\tilde{E} \xi - E}}{\sqrt{\xi} \sqrt{1 - \xi}} d\xi, \\
\partial_{EE}^2 P^-(E, I) &= \frac{\sqrt{2}}{2\pi} \int_0^1 \frac{\sqrt{1 - \xi}}{\sqrt{\xi} (\tilde{E} \xi - E)^{3/2}} d\xi.
\end{aligned}$$

We define $\tilde{E} \xi - E = x_1 + i x_2$ where $x_1 := 2\varepsilon \xi - E_1(1 - \xi)$ and $x_2 := 2\varepsilon k_2(I) \xi - E_2(1 - \xi)$. Using that $|E_1| \geq |E_2|$ and that, if $c_5 c_3 \leq \hat{k}/(2\bar{k}')$, we have $k_1(I) \geq |k_2(I)|$, we obtain that $x_1 \geq |x_2|$.

We observe also that in order to perform the previous change of variables $\theta = \arccos(-1 + (\xi - 1)\tilde{E}/\varepsilon k(I))$ we have to verify that the argument of \arccos is well defined³². For any $I \in \Delta_{\rho_2}^0$, we take $\bar{I}_1 \in \Delta^0$ with $|I - \bar{I}_1| \leq \rho_2$. We have $-E_1 \leq 2\varepsilon k(\bar{I}_1) - \eta/2$. We have

³¹ The other direction of the estimate in (2.12) is completely analogous.

³² We define $\arccos(z_1 + i z_2)$ in the complementary of the set $\{z_1 \in (-\infty, -1] \cup [1 + \infty)\}$

to prove that, defining $y := k_2^2(I)/k_1^2(I)$ with $0 \leq y \leq 1$, if $E_2 = E_1 k_2(I)/k_1(I)$ then $[2\epsilon(k(\bar{I}_1) + k_2(I)) - \eta/2]x + [-2\epsilon k_1 + \eta/2] \geq 0$, which is verified provided $c_5 \leq 1/(4\bar{k}')$.

In the following, in order to estimate the derivatives of P^- , we set $b := 2\epsilon/|E_1| \geq 1$ and we will use Lemma 2.1.

For $I = \bar{I}_1 \in \Delta^0$, we have

$$\begin{aligned} P_2^-(E, \bar{I}_1) &= \frac{2\sqrt{2}}{\pi} \int_0^1 (\tilde{E}_1 y_2 + E_2 y_1) \frac{\sqrt{\xi}}{\sqrt{1-\xi}} d\xi \geq \int_0^1 \frac{\sqrt{2\epsilon} E_2}{\pi(2\epsilon\xi + |E_1|(1-\xi))^{3/2}} \frac{\sqrt{\xi}}{\sqrt{1-\xi}} d\xi \\ &\geq \frac{E_2}{2\pi\sqrt{\mu}} \int_0^b \frac{\sqrt{\xi}}{(2\epsilon\xi + |E_1|)^{3/2}} d\xi \geq \frac{E_2}{6\pi\sqrt{\mu}} \ln\left(1 + \sqrt{\frac{\epsilon}{|E_1|}}\right). \end{aligned} \quad (2.20)$$

Furthermore,³³

$$\begin{aligned} \partial_E P_1^-(E, I) &= \frac{\sqrt{2}}{\pi} \int_0^1 \frac{y_1(x_1, x_2)}{\sqrt{\xi}\sqrt{1-\xi}} d\xi \leq \frac{2}{\pi} \int_0^1 \frac{1}{\sqrt{\xi}\sqrt{2\epsilon\xi + |E_1|/2}} d\xi + \frac{2}{\pi} \int_0^1 \frac{1}{\sqrt{\xi}\sqrt{\epsilon}} d\xi \\ &= \frac{\sqrt{2}}{\pi\sqrt{\mu}} \int_0^b \frac{1}{\sqrt{t}\sqrt{1+t}} dt + \frac{\sqrt{2}}{\pi\sqrt{\mu}} \leq \frac{12}{\pi\sqrt{\mu}} \ln\left(1 + \sqrt{\frac{\epsilon}{|E_1|}}\right). \end{aligned} \quad (2.21)$$

On the other hand,

$$\begin{aligned} \partial_E P_1^-(E, I) &\geq \frac{1}{\pi} \int_0^{1/2} \frac{1}{\sqrt{\xi}\sqrt{2\epsilon\xi + |E_1|}} d\xi = \frac{1}{\sqrt{2\pi}\sqrt{\mu}} \int_0^{b/2} \frac{1}{\sqrt{t}\sqrt{1+t}} dt \\ &\geq \frac{1}{2\pi\sqrt{\mu}} \ln\left(1 + \sqrt{\frac{\epsilon}{|E_1|}}\right). \end{aligned} \quad (2.22)$$

Using the fact that $|x_2| \leq 2\epsilon\bar{k}_2 + |E_2|$ and the estimate

$$|\partial_E P_2^-(E, I)| \leq \frac{1}{\pi} \int_0^1 \frac{|x_2|}{\sqrt{\xi}\sqrt{1-\xi} x_1^{3/2}} d\xi \leq \frac{6}{\pi|E_1|\sqrt{\mu}} + \frac{2}{\epsilon^{3/2}},$$

we obtain that, as in the positive energy case,

$$|\partial_E P^-(E, I)| \leq \frac{2\sqrt{2}\pi}{\sqrt{\epsilon}} \ln\left(1 + \sqrt{\frac{\epsilon}{|E_1|}}\right). \quad (2.23)$$

Since $x_1 \geq |x_2|$, one has $\sqrt{|x_1 + ix_2|} \leq \sqrt{2}\sqrt{x_1}$. From the previous inequality we conclude that

$$|\partial_I P^-(E, I)| \leq \frac{2|k'(I)|}{\pi|k_1(I)|} \int_0^1 \frac{\sqrt{x_1}}{\sqrt{\xi}\sqrt{1-\xi}} d\xi \leq 2\sqrt{6} \frac{|k'(I)|}{\sqrt{k_1(I)}} \sqrt{\epsilon}. \quad (2.24)$$

Finally differentiating the equality $E^\pm(P^\pm(E, I), I) = E$ with respect to E and I we obtain respectively $\partial_P E^\pm(P, I) = [\partial_E P^\pm(E^\pm(P, I), I)]^{-1}$ and

$$\partial_I E^\pm(P, I) = -\partial_I P^\pm(E^\pm(P, I), I) [\partial_E P^\pm(E^\pm(P, I), I)]^{-1},$$

³³ Use $\int_0^b \frac{1}{\sqrt{t}\sqrt{1+t}} dt \leq 4\ln(1 + \sqrt{b})$.

which, by (2.13), (2.17), (2.18), (2.22), (2.23), (2.24), imply (2.5) and (2.6) .

Next, we prove that

$$P^-(\mathcal{E}^-(\bar{I}_1)) \supseteq (D^-(\bar{I}_1))_{2\rho_1} \quad \forall \bar{I}_1 \in \Delta^0. \quad (2.25)$$

Since $P_1^-(E_1 + iE_2, \bar{I}_1) = \frac{2\sqrt{2}}{\pi} \int_0^1 (\tilde{E}_1 y_1 - E_2 y_2) \sqrt{\xi/1 - \xi} d\xi$, we have by Lemma 2.1 that (being $y_1, -y_2$ decreasing) P_1^- is a decreasing function (for $E_2 \geq 0$). Hence, in order to prove (2.25), it is enough to prove the following estimates, $\forall \bar{I}_1 \in \Delta^0, E_1 + iE_2 \in \mathcal{E}^-(\bar{I}_1)$:

- (i) $P_2^-(E_1 + iE_2^*(E_1), \bar{I}_1) \geq 2\rho_1$,
- (ii) $P_1^-(-\eta/2 + iE_2^*(\eta/2), \bar{I}_1) - P_1^-(-\eta, \bar{I}_1) \geq 2\rho_1$,
- (iii) $P_1^-(-2\varepsilon k(\bar{I}_1) + \eta, \bar{I}_1) - P_1^-(-2\varepsilon k(\bar{I}_1) + \eta/2, \bar{I}_1) \geq 2\rho_1$.

If $c_4 \leq \frac{1}{12\pi} c_7 \min\{1, \frac{1}{k}\}$ we obtain (i) because of:

$$P_2^-(E_1 + iE_2^*(E_1), \bar{I}_1) \geq c_7 \frac{\eta}{6\pi\sqrt{\varepsilon}} \frac{\ln(1 + \sqrt{k(\bar{I}_1)}\sqrt{\varepsilon/E_1})}{\sqrt{k(\bar{I}_1)}\ln(1 + \sqrt{\varepsilon/E_1})} \geq c_7 \frac{\eta}{6\pi\sqrt{\varepsilon}} \min\{1, \frac{1}{k}\}.$$

Since

$$\begin{aligned} & P_1^-(-\eta/2 + iE_2^*(\eta/2), \bar{I}_1) - P_1^-(-\eta, \bar{I}_1) \\ & \geq P_1^-(-\eta/2, \bar{I}_1) - P_1^-(-\eta, \bar{I}_1) - |P_1^-(-\eta/2 + iE_2^*(\eta/2), \bar{I}_1) - P_1^-(-\eta/2, \bar{I}_1)|, \end{aligned}$$

using (2.22) and (2.23), if $c_7 \leq 1/(32\pi^2)$, we have

$$\begin{aligned} P_1^-(-\eta/2 + iE_2^*(\eta/2), \bar{I}_1) - P_1^-(-\eta, \bar{I}_1) & \geq \left(\frac{1}{4\pi} - c_7 4\pi\right) \frac{\eta}{\sqrt{\varepsilon}} \ln\left(1 + \sqrt{\frac{\varepsilon}{E_1}}\right) \\ & \geq \frac{1}{8\pi} \frac{\eta}{\sqrt{\varepsilon}} \min\{1, \frac{1}{k}\}, \end{aligned}$$

which (exactly as in the positive energy case) yields (ii).

From (2.22) we have

$$P_1^-(-2\varepsilon k(\bar{I}_1) + \eta, \bar{I}_1) - P_1^-(-2\varepsilon k(\bar{I}_1) + \eta/2, \bar{I}_1) \geq \frac{\eta}{4\pi\sqrt{\mu}} \ln(1 + 1/\sqrt{2}) \geq \frac{1}{8\sqrt{2}\pi\sqrt{k}},$$

which yields (iii), provided $c_4 \leq 1/(16\sqrt{2}\pi\sqrt{k})$.

We now consider the case $k(I) \equiv 1$ and prove (2.8) also in the negative case. As before the estimate on P_1^- is trivial, the ones on P_2^- and \dot{P}_1^- are consequence of³⁴ (2.20) and (2.21), (2.22) respectively; for the other ones we have, using again (2.2) and (2.3),

³⁴ As before the other direction of the estimate in (2.20) is completely analogous.

$$\begin{aligned}\dot{P}_2^-(E) &\sim E_2 \int_0^1 \frac{\sqrt{1-\xi}}{\sqrt{\xi}(\tilde{E}_1\xi - E_1)^{3/2}} d\xi, \\ \ddot{P}_2^-(E) &\sim \int_0^1 \frac{\sqrt{1-\xi}}{\sqrt{\xi}(\tilde{E}_1\xi - E_1)^{3/2}} d\xi, \quad \ddot{P}_2^-(E) \sim \int_0^1 \frac{(1-\xi)^{3/2}}{\sqrt{\xi}(\tilde{E}_1\xi - E_1)^{5/2}} d\xi. \quad (2.26)\end{aligned}$$

If $-2\varepsilon < E_1 < -\varepsilon$ (since, in such case, $\tilde{E}_1\xi - E_1 \sim \varepsilon$, $-E_1 \sim \varepsilon$) we have

$$\dot{P}_2^-(E) \sim \frac{E_2}{\varepsilon^{3/2}}, \quad \ddot{P}_1^-(E) \sim \frac{1}{\varepsilon^{3/2}}, \quad \ddot{P}_2^-(E) \sim \frac{E_2}{\varepsilon^{5/2}}. \quad (2.27)$$

The case $-\varepsilon < E_1 < 0$ (i.e. $\tilde{E}_1 \sim \varepsilon$) is a bit more complicate and it is convenient to break up the integrals in (2.26) as $\int_0^1 = \int_0^{1/2} + \int_{1/2}^1$. The latter integrals are easier to handle since if $1/2 \leq \xi \leq 1$ then $\sqrt{\xi} \sim 1$ and $\tilde{E}_1\xi - E_1 \sim \varepsilon$ and therefore the estimates in (2.27) follow. As for the other integrals, since $0 \leq \xi \leq 1/2$, one has $1 - \xi \sim 1$. Substituting $t = \frac{\tilde{E}_1}{-E_1}\xi$ (so that $\tilde{E}_1\xi - E_1 = -E_1(t+1)$) and denoting $a = -\tilde{E}_1/2E_1$, in view of the estimates in (2.27) and of the estimates done in the integrals over $(1/2, 1)$, we obtain

$$\begin{aligned}\dot{P}_2^-(E) &\sim \frac{E_2}{\varepsilon^{3/2}} + \frac{E_2}{-E_1\sqrt{\varepsilon}} \int_0^a \frac{1}{\sqrt{t}(t+1)^{3/2}} dt \sim E_2 \left(\frac{1}{\varepsilon^{3/2}} + \frac{1}{-E_1\sqrt{\varepsilon}} \right) \sim \frac{E_2}{-E_1\sqrt{\varepsilon}}, \\ \ddot{P}_1^-(E) &\sim \frac{1}{\varepsilon^{3/2}} + \frac{1}{-E_1\sqrt{\varepsilon}} \int_0^a \frac{1}{\sqrt{t}(t+1)^{3/2}} dt \sim \frac{1}{\varepsilon^{3/2}} + \frac{1}{-E_1\sqrt{\varepsilon}} \sim \frac{1}{-E_1\sqrt{\varepsilon}}, \\ \ddot{P}_2^-(E) &\sim \frac{E_2}{\varepsilon^{5/2}} + \frac{E_2}{(-E_1)^2\sqrt{\varepsilon}} \int_0^a \frac{1}{\sqrt{t}(t+1)^{5/2}} dt \sim \frac{E_2}{\varepsilon^{3/2}} + \frac{E_2}{(-E_1)^2\sqrt{\varepsilon}} \sim \frac{E_2}{(-E_1)^2\sqrt{\varepsilon}},\end{aligned}$$

from which (2.8) follows.

- (Fourth step: construction of the symplectic transformation)

We will find our symplectic transformation using the generating function $S(E, I, q) := \sqrt{2} \int_0^q \sqrt{g(E, I, \theta)} d\theta$. We note that in order to well define S we have to take into account the presence of the square root. In particular we are interested in the definition of the functions

$$\chi^\pm(E, I, q) := \frac{\partial_E S(E, I, q)}{\partial_E P^\pm(E, I)}, \quad \xi^\pm(E, I, q) := \frac{\partial_I S(E, I, q)}{\partial_I P^\pm(E, I)}.$$

Let $\mathcal{T}^+ := \mathbb{C}$ and $\mathcal{T}^- := \{q \in \mathbb{C} \text{ s.t. } |q_1| < \pi\}$ and define³⁵

$$\tilde{\mathcal{D}}^\pm(E, I) := \{q \in \mathcal{T}^\pm \text{ s.t. } g(E, I, q) \notin (-\infty, 0]\}.$$

For any $E \in \mathcal{E}^\pm$, $I \in \Delta_{\rho_2}^0$, the functions

$$S(E, I, q), \quad \partial_E S(E, I, q), \quad \partial_I S(E, I, q), \quad \chi^\pm(E, I, q), \quad \xi^\pm(E, I, q),$$

³⁵ If $a, b \in \mathbb{C}$ we denote $(a, b) := \{z = a + t(b - a), \text{ with } t \in (0, 1)\}$ (and, analogously, for $[a, b)$, $(a, b]$, $[a, b]$); symbols like $(a, \alpha\infty)$, with $\alpha \in \mathbb{C}$ and $|\alpha| = 1$, (or $[a, \alpha\infty)$, $(\alpha\infty, \beta\infty)$, etc) denote lines: $(a, \alpha\infty) := \{z = a + \alpha t, \text{ with } t > 0\}$.

are analytic in q on $\tilde{\mathcal{D}}^\pm(E, J)$. So the functions $\chi^\pm(E(p, I, q), I, q)$ and $\xi^\pm(E(p, I, q), I, q)$ are analytic in p, q on the disconnected set³⁶

$$\begin{aligned}\tilde{\mathcal{D}}^\pm &:= \tilde{\mathcal{D}}^\pm(I) := \{(p, q) \in \mathbb{C} \times \mathcal{T}^\pm \text{ s.t. } g(E(p, I, q), I, q) \notin (-\infty, 0]\} \\ &= \{(p, q) \in \mathbb{C} \times \mathcal{T}^\pm \text{ s.t. } p^2/2 \notin (-\infty, 0]\} = \{(p, q) \in \mathbb{C} \times \mathcal{T}^\pm \text{ s.t. } p_1 \neq 0\}.\end{aligned}$$

Our next step will be to define both $\tilde{\chi}^\pm(p, I, q) := \chi^\pm(E(p, I, q), I, q)$ and $\tilde{\xi}^\pm(p, I, q) := \xi^\pm(E(p, I, q), I, q)$ for all $(p, q) \in \mathbb{C} \times \mathcal{T}^\pm$. We set

$$\tilde{\chi}^\pm(p, I, q) := \begin{cases} \frac{1}{\sqrt{2} \partial_E P^\pm(E(p, I, q), I)} \int_0^q \frac{d\theta}{\sqrt{g(E(p, I, q), I, \theta)}}, & \text{if } p_1 > 0, \\ \pi - \frac{1}{\sqrt{2} \partial_E P^\pm(E(p, I, q), I)} \int_0^q \frac{d\theta}{\sqrt{g(E(p, I, q), I, \theta)}}, & \text{if } p_1 < 0, \end{cases}$$

and

$$\tilde{\xi}^\pm(p, I, q) := \begin{cases} \frac{\epsilon}{\sqrt{2} \partial_I P^\pm(E(p, I, q), I)} \int_0^q \frac{1 + \cos \theta}{\sqrt{g(E(p, I, q), I, \theta)}} d\theta, & \text{if } p_1 > 0, \\ \pi - \frac{\epsilon}{\sqrt{2} \partial_I P^\pm(E(p, I, q), I)} \int_0^q \frac{1 + \cos \theta}{\sqrt{g(E(p, I, q), I, \theta)}} d\theta, & \text{if } p_1 < 0, \end{cases}$$

which are well defined and analytic for $(p, q) \in \tilde{\mathcal{D}}^\pm$. Notice that, in the positive energy case, there are no problems with the definition of $\tilde{\chi}^+$ and $\tilde{\xi}^+$, and we note that

$$\tilde{\chi}^+(p, I, q + 2\pi) = \tilde{\chi}^+(p, I, q) + 2\pi \quad \text{and} \quad \tilde{\xi}^+(p, I, q + 2\pi) = \tilde{\xi}^+(p, I, q) + 2\pi. \quad (2.28)$$

In the negative energy case we proceed differently. We define

$$F_E(E, I, p) := - \int_0^p \frac{1}{\epsilon \sqrt{\hat{g}(E, I, z)}} dz, \quad F_I(E, I, p) := - \int_0^p \frac{-E + z^2}{\epsilon \sqrt{\hat{g}(E, I, z)}} dz,$$

where $\hat{g}(E, I, p) := 1 - (-1 - E/\epsilon + p^2/2\epsilon)^2$ is analytic on the complex domain $\hat{\mathcal{D}}^-(E, I) := \{p \in \mathbb{C} \text{ s.t. } \hat{g}(E, I, p) \notin (-\infty, 0]\}$. Then, $F_E(E(p, I, q), I, p)$ and $F_I(E(p, I, q), I, p)$ are well defined and analytic on³⁷

$$\begin{aligned}\hat{\mathcal{D}}^- &:= \hat{\mathcal{D}}^-(I) := \{(p, q) \in \mathbb{C} \times \mathcal{T}^- \text{ s.t. } \hat{g}(E(p, I, q), I, p) \notin (-\infty, 0]\} \\ &= \{(p, q) \in \mathbb{C} \times \mathcal{T}^- \text{ s.t. } 1 - \cos^2 q \notin (-\infty, 0]\} = \{(p, q) \in \mathbb{C} \times \mathcal{T}^- \text{ s.t. } q_1 \neq 0\}.\end{aligned}$$

³⁶ We see that $\tilde{\mathcal{D}}^\pm$ does not really depend on I .

³⁷ Also in this case $\hat{\mathcal{D}}^-$ does not really depend on I .

We now split the integral in the definition of S , $\partial_E S$ and $\partial_I S$ as $\int_0^q = \int_0^{\psi_0} + \int_{\psi_0}^q$ and in the second integral we perform the change of variable $\theta = \arccos(-1 - E/\epsilon + z^2/2\epsilon)$. Then, defining³⁸

$$\hat{\chi}^-(p, I, q) := \begin{cases} \pi/2 + \left(\partial_E P^-(E(p, I, q), I)\right)^{-1} F_E(E(p, I, q), I, p), & \text{if } q_1 > 0, \\ -\pi/2 - \left(\partial_E P^-(E(p, I, q), I)\right)^{-1} F_E(E(p, I, q), I, p), & \text{if } q_1 < 0, \end{cases}$$

and

$$\hat{\xi}^-(p, I, q) := \begin{cases} \pi/2 + \left(\partial_I P^-(E(p, I, q), I)\right)^{-1} F_I(E(p, I, q), I, p), & \text{if } q_1 > 0, \\ -\pi/2 - \left(\partial_I P^-(E(p, I, q), I)\right)^{-1} F_I(E(p, I, q), I, p), & \text{if } q_1 < 0, \end{cases}$$

we have $\forall I \in \Delta_{\rho_2}^0, \forall (p, q) \in \tilde{\mathcal{D}}^- \cap \hat{\mathcal{D}}^-$

$$\tilde{\chi}^-(p, I, q) \equiv \hat{\chi}^-(p, I, q) \bmod 2\pi \quad \text{and} \quad \tilde{\xi}^-(p, I, q) \equiv \hat{\xi}^-(p, I, q) \bmod 2\pi. \quad (2.29)$$

Using (2.29), we can finally define³⁹ $\bar{\chi}^-, \bar{\xi}^- : (p, q) \in \tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^- \rightarrow \mathbb{C}/2\pi\mathbb{Z}$

$$\bar{\chi}^-(p, I, q) := \begin{cases} \tilde{\chi}^-(p, I, q), & \text{if } (p, q) \in \tilde{\mathcal{D}}^-, \\ \hat{\chi}^-(p, I, q), & \text{if } (p, q) \in \hat{\mathcal{D}}^-, \end{cases}$$

and

$$\bar{\xi}^-(p, I, q) := \begin{cases} \tilde{\xi}^-(p, I, q), & \text{if } (p, q) \in \tilde{\mathcal{D}}^-, \\ \hat{\xi}^-(p, I, q), & \text{if } (p, q) \in \hat{\mathcal{D}}^-, \end{cases}$$

where, on $\tilde{\mathcal{D}}^-$,

$$\chi^-(E(p, I, q), I, q) \equiv \bar{\chi}^-(p, I, q) \quad \text{and} \quad \xi^-(E(p, I, q), I, q) \equiv \bar{\xi}^-(p, I, q).$$

Moreover we finally extend by periodicity the definition of $\bar{\chi}^-(p, I, q)$ and $\bar{\xi}^-(p, I, q)$ on all $\{q \in \mathbb{C} \text{ s.t. } q_1 \neq \pi + 2k\pi, k \in \mathbb{Z}\} = \cup_{k \in \mathbb{Z}} (2k\pi + \mathcal{T}^-)$ in the following way: if $q \in 2k\pi + \mathcal{T}^-$ we define $\bar{\chi}^-(p, I, q) := \bar{\chi}^-(p, I, q - 2k\pi)$ and $\bar{\xi}^-(p, I, q) := \bar{\xi}^-(p, I, q - 2k\pi)$.

Now we are able to construct our symplectic transformation. Since $\partial P_E^\pm \neq 0$, by the Implicit Function Theorem, there exists $E^\pm = E^\pm(P, J)$ such that

$$P^\pm(E^\pm(P, J), J) \equiv P. \quad (2.30)$$

³⁸ We note that $\hat{\chi}^-$ and $\hat{\xi}^-$ are analytic on $\hat{\mathcal{D}}^-$.

³⁹ We observe that $\tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^-$ is an open set and that its complementary set $(\tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^-)^c = \{(p, q) \in \mathbb{C} \times \mathcal{T}^-, \text{ s.t. } p_1 = 0, q_1 = 0\}$ does not interest our analysis. In fact, if $(p, q) \in (\tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^-)^c$, then $p = ip_2$ and $q = iq_2$ and, hence, we have $E(ip_2, I, iq_2) = -p^2/2 - \epsilon k(I)(1 + \cosh q_2)$ and $E_1(ip_2, I, iq_2) = -p_2^2/2 - \epsilon k_1(I)(1 + \cosh q_2) \leq -2\epsilon k_1(I) < -2\epsilon k(\bar{I}_1) + \eta/2$ where $|I - \bar{I}_1| \leq \rho_2$ (and we have used the fact that $c_5 \leq 1/4\bar{k}'$). We conclude that $E(ip_2, I, iq_2) \notin \mathcal{E}^-(\bar{I}_1)$.

Let $S^\pm(P, J, q) := S(E^\pm(P, J), J, q)$; we , then, define the following generating functions (which depend on the new actions and on the old angles): $G^\pm(P, J, q, \varphi) := J\varphi + S^\pm(P, J, q)$. Our symplectic transformation ϕ^\pm is implicitly defined by

$$\begin{cases} p = \partial_q G^\pm = \partial_q S^\pm(P, J, q), & Q = \partial_P G^\pm = \partial_P S^\pm(P, J, q), \\ I = \partial_\varphi G^\pm = J, & \psi = \partial_J G^\pm = \varphi + \partial_J S^\pm(P, J, q), \end{cases}$$

We want to express $(\phi^\pm)^{-1}$ as a function of the old variables (p, I, q, φ) . We immediately have $J = I$ and $P = P^\pm(E(p, I, q), I)$. Differentiating (2.30) with respect to J and P we have, respectively, $\partial_E P^\pm \partial_J E^\pm + \partial_J P^\pm = 0$ and $\partial_E P^\pm \partial_P E^\pm = 1$. Now, we can express the new angles as functions of the old variables:

$$\begin{cases} Q = Q^\pm(p, I, q) := \bar{\chi}^\pm(p, I, q), \\ \psi = \psi^\pm(p, I, q, \varphi) := \varphi - \partial_I P^\pm(E(p, I, q), I) [Q^\pm(p, I, q) - \bar{\xi}^\pm(p, I, q)]. \end{cases}$$

We observe that Q^- and ψ^- are 2π -periodic in q by definition of $\bar{\chi}^-$ and $\bar{\xi}^-$; by (2.28) we deduce that ψ^+ is 2π -periodic in q too and $Q^+(p, I, q + 2\pi) = Q^+(p, I, q) + 2\pi$.

• (Fifth step: estimate on the angle analyticity radius)

We first study the analyticity radius in Q . Fix $I \in \Delta_{\rho_2}^0$ and $\bar{I}_1 \in \Delta^0$ with $|I - \bar{I}_1| \leq \rho_2$. We must prove that $\forall P_* \in D_{\rho_1}^\pm$ and $\forall Q_* \in \mathbb{T}_{\sigma_1}$ there exist p_*^\pm and q_*^\pm such that $Q^\pm(p_*^\pm, I, q_*^\pm) = Q_*$. So it is sufficient to prove that $\forall E_* \in \mathcal{E}^\pm(\bar{I}_1)$ we have $\chi^\pm(E_*, I, \mathbb{T}_{s_1}) \supseteq \mathbb{T}_{\sigma_1}$.

We first consider the positive energy case.

We observe that we have $\chi^+(E, I, 0) = 0$, $\chi^+(E, I, \pm\pi) = \pm\pi$. Let us first consider the case $I = \bar{I}_1 \in \Delta^0$, $E = E_1 \in \mathcal{E}^+$. In such a case

$$\chi^+(E_1, \bar{I}_1, (-\pi, \pi)) = (-\pi, \pi), \quad \chi^+(E_1, \bar{I}_1, (0, \pm i\infty)) = (0, \pm is^+(E_1, \bar{I}_1)),$$

and

$$\chi^+(E_1, \bar{I}_1, (\pm\pi, \pm\pi \pm i\psi_0(E_1, \bar{I}_1))) = (\pm\pi, \pm\pi \pm is^+(E_1, \bar{I}_1)),$$

where

$$s^\pm(E_1, \bar{I}_1) := \sqrt{2}(\partial_E P^\pm)^{-1} \int_0^\infty \frac{d\theta}{\sqrt{E_1 + \varepsilon k(\bar{I}_1)(1 + \cosh \theta)}}. \quad (2.31)$$

In fact, it is $\chi^+(E_1, \bar{I}_1, \mathcal{D}(E_1, \bar{I}_1)) = \mathbb{T}_{s^+(E_1, \bar{I}_1)}$. We will prove that

$$\chi^+(E_1, \bar{I}_1, \mathbb{T}_{s_1} \cap \tilde{\mathcal{D}}^+(E_1, \bar{I}_1)) \supseteq \mathbb{T}_{\sigma^+(E_1, \bar{I}_1, s_1)} \quad (2.32)$$

where, for $s > 0$,

$$\sigma^\pm(E, I, s) := \inf_{t \in (-\pi, \pi)} \chi_2^\pm(E, I, t + is). \quad (2.33)$$

Observe that $\partial_E P^+ \in \mathbb{R}$ and $g = g_1 - ig_2$, where⁴⁰ $g_1 = E_1 + \varepsilon k(\bar{I}_1)(1 + \cos q_1 \cosh q_2)$ and $g_2 = \varepsilon k(\bar{I}_1) \sin q_1 \sinh q_2$. Splitting the integral $\int_0^{t+is} = \int_0^{is} + \int_{is}^{t+is}$, we have $\int_0^{is} 1/\sqrt{g} =$

⁴⁰ For symmetry reasons we can consider $t, q_1 \geq 0$.

$i \int_0^s 1/\sqrt{g_1}$ and, using the notation of Lemma 2.1, we obtain $\text{Im} \int_{is}^{is+t} 1/\sqrt{g} = \int_0^t y_2(g_1, g_2)$, which, since $y_2 > 0$, attains its minimum at $t = 0$. Collecting all these informations we have that

$$\sigma^+(E_1, \bar{I}_1, s) = \chi_2^+(E_1, \bar{I}_1, is) := \sqrt{2}(\partial_E P^+)^{-1} \int_0^s \frac{d\theta}{\sqrt{E_1 + \varepsilon k(\bar{I}_1)(1 + \cosh \theta)}}.$$

It is easy to see that

$$\int_0^s \frac{d\theta}{\sqrt{E_1 + \varepsilon k(\bar{I}_1)(1 + \cosh \theta)}} \geq c_9 \frac{1}{\sqrt{\varepsilon k(\bar{I}_1)}} \ln \left(1 + s \sqrt{\frac{\varepsilon k(\bar{I}_1)}{E_1 + 2\varepsilon k(\bar{I}_1)}} \right).$$

Thus, by (2.17), we get

$$\sigma^+(E_1, \bar{I}_1, s) \geq c_{10} \ln \left(1 + s \sqrt{\frac{\varepsilon k(\bar{I}_1)}{E_1 + 2\varepsilon k(\bar{I}_1)}} \right) \ln^{-1} \left(1 + \sqrt{\frac{\varepsilon k(\bar{I}_1)}{E_1}} \right),$$

which implies that, $\forall \eta/2 \leq E_1 \leq \bar{E}$ and $\bar{I}_1 \in \Delta^0$,

$$\sigma^+(E_1, \bar{I}_1, s) \geq c_{11} \frac{s}{\ln(\varepsilon/\eta)}.$$

In the general case, using the estimates on χ^+ and its derivatives⁴¹, one has that, if $E = E_1 + iE_2 \in \mathcal{E}^+$ and $\bar{I} \in \Delta_{\rho_2}^0$, then $\sigma^+(E, I, s) \geq c_{12}\sigma^+(E_1, \bar{I}_1, s) \geq c_{13} \frac{s}{\ln(\varepsilon/\eta)}$. Taking $s := s_1$ and $c_6 \leq c_{13}$, we have the claim concerning the form of σ_1 .

We now pass to the negative energy case. As before fix E and I and observe that $\chi^-(E, I, 0) = 0$ and $\chi^-(E, I, \pm\psi_0(E, I)) = \pm\pi/2$. Consider first the case $I = \bar{I}_1 \in \Delta^0$ and $E = E_1 \in \mathcal{E}^-(\bar{I}_1)$. We find $\chi^-(E_1, \bar{I}_1, (-\psi_0(E_1, \bar{I}_1), \psi_0(E_1, \bar{I}_1))) = (-\pi/2, \pi/2)$ and $\chi^-(E_1, \bar{I}_1, (0, \pm i\infty)) = (0, \pm is^-(E_1, \bar{I}_1))$, where $s^-(E_1, \bar{I}_1)$ was defined in (2.31). It is simple to see that we have

$$\left\{ \bar{\chi}^-(p, \bar{I}_1, q) \text{ s.t. } (p, q) \in \tilde{\mathcal{D}}^- \cup \hat{\mathcal{D}}^-, E(p, \bar{I}_1, q) = E_1, |q_2| < s_1 \right\} \supseteq \left\{ |Q_2| < \sigma^-(E_1, \bar{I}_1, s_1) \right\}$$

which is analogous to (2.32). The estimate on σ_1 for the general case $E \in \mathcal{E}^-(\bar{I}_1)$ and $I \in \Delta_{\rho_2}^0$ with $\bar{I}_1 \in \Delta^0$, $|I - \bar{I}_1| \leq \rho_2$, follows exactly as in the positive energy case.

We now briefly discuss the analyticity radius in the angle ψ . Observing that, as it is simple to see, $|\bar{\chi}^\pm|, |\bar{\xi}^\pm| \leq c_{14}$ and remembering (2.18) and (2.24), we see that, $|\psi^\pm(P, I, q, \varphi) - \varphi| \leq c_{15}\sqrt{\varepsilon}$. Hence, if ε is sufficiently small, we can take $\sigma_2 = s_2/2$. The proof of Proposition 2.1 is now complete. \blacksquare

⁴¹ See Appendix B of [33].

3 The D'Alembert Hamiltonian planetary model

In this section we consider the Hamiltonian version of the D'Alembert model for the planetary spin/orbit problem. It will be one of the physical motivation of the analysis developed in Part II and, especially, the time stability of this model will be studied in Part III below. The model may be described as follows.

Let a planet be modelled by a rotational ellipsoid slightly flattened along the symmetry axis (called “north–south” direction); assume that the center of mass of such planet revolves on a slightly eccentric Keplerian ellipse around a *fixed* star occupying one of the foci of the ellipse: the planet is subject to the gravitational attraction of the star and the problem is to study the relative position of the planet and, most notably, the time evolution of its angular momentum.

Such model may be described using Hamiltonian formalism (see subsection 3.1 below) using action–angle symplectic variables. The Hamiltonian system describing the D'Alembert model results to be a two–degrees–of–freedom system depending explicitly and periodically on time (the period being the year of the planet); furthermore such Hamiltonian system is nearly–integrable (with *two* smallness parameters: the flatness of the planet and the eccentricity of the Keplerian ellipse) and *properly degenerate*⁴².

In particular, we are interested in studying the D'Alembert model in the vicinity of a *spin/orbit resonance*, i.e., in a phase space region where the period of revolution of the planet around the star (the “year”) and the period of the rotation of the planet around its spin axis (the “day”) are in a close–to–exact rational relation. If such rational relation is p/q (p and q positive, co–prime integers) we shall speak of a $p : q$ (spin/orbit) resonance.

The degeneracy of the system implies that the time variable (better: the angle–variable corresponding to time, i.e., the so–called “mean anomaly”) may be considered a *fast variable* with respect to the two (“symplectic”) angles describing the relative position of the planet. Thus the explicit dependence of the system upon time may be *averaged out*, as we will do in Part III below. In fact we will see that, *up to an exponentially small term*, the “effective” Hamiltonian results to be exactly the average of the whole Hamiltonian over the “mean anomaly”.

In subsection 3.2 below we will show that *near a $p : q$ spin/orbit resonance with (p, q) different from $(1, 1)$ and $(2, 1)$, the intermediate system (and hence the “effective” Hamiltonian) is independent of any angle variable*: thus the integrable system obtained dropping (besides the exponential remainder) the higher order term is a completely integrable system with phase space entirely foliated by (maximal) invariant curves. On the other hand, *near a $p : q$ spin/orbit resonance with (p, q) equal to $(1, 1)$ or $(2, 1)$, the intermediate system (and hence the “effective” Hamiltonian) does depend on one (and only one) angle variable*: in such “exceptional” case, the system obtained by dropping the higher order terms is still integrable (being, effectively, a one–degree–of–freedom system) but its phase space presents a structure similar to that of a standard pendulum (i.e.,

⁴²Roughly speaking, “properly degenerate” means that in the integrable limit (i.e., when the perturbative parameters are set to zero) the Hamiltonian does not depend on the action–variables in a “general” way.

elliptic and hyperbolic equilibria, separatrices, invariant curves of different homotopy).

Thus, the effective Hamiltonian associated to the 1 : 1 or 2 : 1 spin-orbit resonance exhibits instability phase-space zones that are not present in the general case, a phenomenon which may be, perhaps, exploited in the understanding of the exceptional role played by such resonances in our Solar system and in its evolution.

We also mention that such peculiarity of the 1 : 1 and 2 : 1 spin/orbit resonance is *intrinsic* in the model and does not depend upon the particular variables used.

3.1 Hamiltonian formalism for the D'Alembert planetary model

In this subsection we revisit briefly the Hamiltonian version of the planetary D'Alembert model as presented, e.g., in [49].

Consider an oblate planet \mathcal{P} of mass $m_{\mathcal{P}}$ modelled by a rotational ellipsoid slightly flattened along the symmetry axis (“north-south axis”); assume that its center of mass revolves on a *Keplerian orbit* (of small eccentricity) around a *fixed star* of mass m_S occupying one of the foci of the ellipse⁴³.

We start by writing down the Lagrangian of this system. Let (i, j, k) be an orthogonal fixed basis with “origin” on the star so that: i is the unit versor pointing towards the aphelion (or, equivalently, pointing towards the other focus of the ellipse); j is the versor in the ecliptic plane (i.e., the plane containing the Keplerian ellipse) orthogonal to i oriented according to the motion of the center of mass of the planet (i.e., if $x_{\mathcal{P}}(t)$ denotes the position at time t of the center of mass of \mathcal{P} and if $x_{\mathcal{P}}$ passes at time t_0 at the aphelion, then⁴⁴ $j \cdot \dot{x}_{\mathcal{P}}(t_0) > 0$); $k = i \times j$ is the unit normal to ecliptic plane⁴⁵. Let (i_1, i_2, i_3) be a co-moving frame with “origin” in the center of mass of the planet so that: i_1 and i_2 determine the “equatorial plane” and i_3 points towards the “north pole⁴⁶”; clearly i_1 and i_2 may be interchanged; for simplicity, we shall assume that i_3 is never orthogonal nor parallel to the ecliptic plane and that the basis (i_1, i_2, i_3) is such that

$$0 < i_3 \cdot k < 1. \quad (3.1)$$

Let, now, $\theta \equiv (\theta_1, \theta_2, \theta_3)$ denote the Euler angles of the planet, namely, if n denotes a unit vector identifying the *equatorial node* on the ecliptic (i.e., the line obtained as intersection between the ecliptic plane and the equatorial plane), then

$$\theta_1 = \text{angle}(i, n), \quad \theta_2 = \text{angle}(i_3, k), \quad \theta_3 = \text{angle}(n, i_1). \quad (3.2)$$

⁴³In other words, we assume that the motion of the star is not influenced by the form of the planet.

⁴⁴ $a \cdot b$ denotes the standard inner product in \mathbb{R}^n (here $n = 3$); and \dot{a} denotes the time derivative of a .

⁴⁵Here, “ \times ” denotes the standard “vector” (or “external”) skew-symmetric product in \mathbb{R}^3 . Informally, an observer “standing” on the ecliptic in the position identified by k would see the center of mass of \mathcal{P} revolve “counter-clockwise”.

⁴⁶Recall that we are assuming the the planet is a rotational ellipsoid; thus the “equatorial plane” is the plane identified by the maximal circle of the ellipsoid and the “north-south” axis is the line orthogonal to the equatorial plane.

Then, if $\mathcal{I}_1 = \mathcal{I}_2$ and \mathcal{I}_3 denotes the inertia moments of the planet, γ denotes the gravitational constant, $x_{\mathcal{P}}(t)$ denotes, as above, the position at time t of the center of mass of \mathcal{P} and $\mathcal{P}(t)$ denotes the space region occupied at time t by the planet, then *the Lagrangian describing the above model is given by*

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}\mathcal{I}_3(\dot{\theta}_1 \cos \theta_2 + \dot{\theta}_3)^2 + \frac{1}{2}\mathcal{I}_1(\dot{\theta}_2^2 + \dot{\theta}_1^2 \sin^2 \theta_2) \\ & + \gamma \frac{m_{\mathcal{P}}m_S}{Vol\mathcal{P}} \int_{\mathcal{P}(t)} \frac{dx}{|x_{\mathcal{P}}(t) + x|} . \end{aligned} \quad (3.3)$$

Thanks to a well known result by Andoyer and Deprit (see, e.g., [8], [65]), *the Legendre transform of \mathcal{L} is equivalent, in suitable physical units, to the following Hamiltonian function*⁴⁷

$$\begin{aligned} H_{\varepsilon,\mu}(I, \varphi) \equiv & \frac{(\bar{J}_1 + I_1)^2}{2} + \bar{\omega}(I_3 - I_2) \\ & + \varepsilon F_0(I_1, I_2, \varphi_1, \varphi_2) + \varepsilon \mu F_1(I_1, I_2, \varphi_1, \varphi_2, \varphi_3; \mu) , \end{aligned} \quad (3.4)$$

where:

- a) \bar{J}_1 is constant parameter, which may be interpreted as a “reference datum” in a neighborhood of which the system will be studied;
- b) ε and μ are two *small* non-negative parameters measuring, respectively, the flatness of the planet and the eccentricity of the Keplerian orbit described by the center of mass of the planet;
- c) $(I, \varphi) \equiv (I_1, I_2, I_3, \varphi_1, \varphi_2, \varphi_3) \in A \times \mathbb{T}^3$ are standard symplectic coordinates⁴⁸; the domain $A \subset \mathbb{R}^3$ is given by

$$A \equiv \left\{ |I_1| < d , \quad |I_2 - \bar{J}_2| < d , \quad I_3 \in \mathbb{R} \right\} , \quad (3.5)$$

where d is a suitable fixed (and small) positive number while \bar{J}_2 is fixed “reference datum” (verifying, together with \bar{J}_1 , certain assumptions spelled out below);

- d) $2\pi/\bar{\omega}$ is the period of the Keplerian motion (“year of the planet”);
- e) the function F_0 is a trigonometric polynomial given by

$$F_0 = \sum_{\substack{j \in \mathbb{Z} \\ |j| \leq 2}} c_j \cos(j\varphi_1) + d_j \cos(j\varphi_1 + 2\varphi_2) , \quad (3.6)$$

where c_j and d_j are functions of $(\bar{J}_1 + I_1, I_2)$ listed in the following item;

⁴⁷See [49].

⁴⁸The symbol \mathbb{T}^n denotes the standard n -dimensional flat torus $\mathbb{R}^n/(2\pi\mathbb{Z}^n)$.

f) let

$$\begin{aligned}\kappa_1 &\equiv \kappa_1(I_1) \equiv \frac{L}{\bar{J}_1 + I_1}, & \kappa_2 &\equiv \kappa_2(I_1, I_2) \equiv \frac{I_2}{\bar{J}_1 + I_1}, \\ \nu_1 &\equiv \nu_1(I_1) \equiv \sqrt{1 - \kappa_1^2}, & \nu_2 &\equiv \nu_2(I_1, I_2) \equiv \sqrt{1 - \kappa_2^2};\end{aligned}\quad (3.7)$$

where L is a real parameter; the parameters \bar{J}_i , L and the constant d are assumed to satisfy

$$L + d < \bar{J}_1, \quad |\bar{J}_2| + 2d < \bar{J}_1; \quad (3.8)$$

in this way $0 < \kappa_i < 1$ (and the ν_i 's are well defined on the domain A). Then, the functions c_j and d_j are defined by

$$\begin{aligned}c_0(I_1, I_2) &\equiv \frac{1}{4} (2\kappa_1^2 \nu_2^2 + \nu_1^2 (1 + \kappa_2^2)), \\ d_0(I_1, I_2) &\equiv -\frac{\nu_2^2}{4} (2\kappa_1^2 - \nu_1^2), \\ c_{\pm 1}(I_1, I_2) &\equiv \frac{\kappa_1 \kappa_2 \nu_1 \nu_2}{2}, \\ d_{\pm 1}(I_1, I_2) &\equiv \mp \frac{(1 \pm \kappa_2) \kappa_1 \nu_1 \nu_2}{2}, \\ c_{\pm 2}(I_1, I_2) &\equiv -\frac{\nu_1^2 \nu_2^2}{8}, \\ d_{\pm 2}(I_1, I_2) &\equiv -\frac{\nu_1^2 (1 \pm \kappa_2)^2}{8}.\end{aligned}\quad (3.9)$$

g) the function F_1 is a convergent series in μ of trigonometric polynomials (with increasing degrees); for example $F_1|_{\mu=0} \equiv F_1^0$ is given by

$$\begin{aligned}F_1^0 &= \sum_{\substack{j \in \mathbb{Z} \\ |j| \leq 2}} (-3) c_j \cos(j\varphi_1 + \varphi_3) \\ &\quad + \frac{d_j}{2} \left\{ \cos(j\varphi_1 + 2\varphi_2 + \varphi_3) - 7 \cos(j\varphi_1 + 2\varphi_2 - \varphi_3) \right\}.\end{aligned}$$

Remark 3.1 (i) Since I_3 appears only linearly with coefficient $\bar{\omega}$, the angle φ_3 corresponds to time t and $H_{\varepsilon, \mu}$ is actually a two-degrees-of-freedom Hamiltonian depending explicitly on time in a periodic way (with period $2\pi/\bar{\omega}$).

(ii) The *physical interpretation* of the action-variables I_1 , I_2 , the parameter L and the angles φ_i , which are closely related to (but do not coincide with) the Andoyer canonical variables, is the following. In suitable physical units, the variable $\bar{J}_1 + I_1$ corresponds to the absolute value of the angular momentum of the planet; the variable I_2 corresponds to the absolute value of the projection of the angular momentum of the planet onto the

direction k orthogonal to the *ecliptic* plane and L corresponds to the absolute value of the projection of the angular momentum of the planet in the direction i_3 of the polar axis of the planet (and, because of the symmetry of the planet, is a constant of the motion). In formulae, if K_P denotes the angular momentum of the planet, then:

$$\bar{J}_1 + I_1 = |K_P|, \quad I_2 = K_P \cdot k, \quad L = K_P \cdot i_3 = \text{const}. \quad (3.10)$$

To describe the angles φ_i let us introduce two more relevant “nodes”: let m be a versor in the direction of the line of intersection (“node”) of the ecliptic plane with the “angular momentum plane” (i.e., the plane orthogonal to the angular momentum of the planet); let, also, n_0 be a versor in the direction of the line of intersection (“node”) of the equatorial plane with the angular momentum plane. Then: φ_3 is the so-called “mean anomaly” and is proportional to time, as seen above; φ_1 is the angle between the nodes m and n_0 ; φ_2 is the *difference* between the angle between m and i and φ_3 . In formulae:

$$\varphi_3 = \text{const} + \bar{\omega}t, \quad \varphi_1 = \text{angle}(m, n_0), \quad \varphi_2 = \text{angle}(m, i) - \varphi_3. \quad (3.11)$$

(iii) Under our assumptions (i.e., that $0 < d \ll 1$), the *average over the angles of $H_{\varepsilon,0}$* is given by

$$\frac{(\bar{J}_1 + I_1)^2}{2} + \bar{\omega}(I_3 - I_2) + \varepsilon \frac{1}{4} \left\{ (2 - \bar{\nu}_1^2) - (2 - 3\bar{\nu}_1^2) \frac{I_2^2}{\bar{J}_1^2} + O(d) \right\}, \quad (3.12)$$

where $\bar{\nu}_1 \equiv \nu_1(0) = \sqrt{1 - (L/\bar{J}_1)^2}$. The number $\bar{\nu}_1$ is the so-called *Euler nutation constant*. By (ii) we see that $\bar{\nu}_1 \ll 1$ corresponds to rotations of the planet with spin axis nearly parallel to the polar axis (a case common, for example, in the Solar System). In such a case the *average over the angles of $H_{\varepsilon,0}$* is not a convex function of the action variables (I_1, I_2) . This lack of convexity is quite a common feature in Celestial Mechanics and is exhibited, for example, also in three-body-problems.

We are interested in studying the above system in a neighborhood of a day/year (or “spin/orbit”) resonance. Since the daily rotation is measured by the angle φ_1 and since in the unperturbed situation ($\varepsilon = 0$ and $I_1 = 0$) $\varphi_1 = \varphi_1^0 + \bar{J}_1 t$, we see that *an approximate day/year resonance corresponds to take the “reference datum” \bar{J}_1 (which, in our units, coincides with the daily frequency) in a rational relation with the year frequency $\bar{\omega}$, i.e., $\bar{J}_1 = \frac{p}{q} \bar{\omega}$ with p and q co-prime positive integers; we shall speak in such a case of a “ $p : q$ spin/orbit-resonance”.*

Setting

$$\bar{J}_1 \equiv \frac{p}{q} \bar{\omega}, \quad \omega \equiv \frac{\bar{\omega}}{q}, \quad (3.13)$$

we see that *the dynamics near a $p : q$ spin/orbit resonance is described by the hamiltonian*

$$H_{\varepsilon,\mu}(I, \varphi) \equiv \frac{I_1^2}{2} + \omega(pI_1 - qI_2 + qI_3) + \varepsilon F_0(I_1, I_2, \varphi_1, \varphi_2) + \varepsilon \mu F_1(I_1, I_2, \varphi_1, \varphi_2, \varphi_3; \mu), \quad (3.14)$$

(where we have omitted the constant term $\bar{J}_1^2/2$).

Finally, to make the analysis perturbative, we shall take as action-variable domain an ε -dependent subset of A :

- h) the domain of definition A introduced in item c) above will, from here on, be replaced by its subset

$$A_\varepsilon \equiv \left\{ |I_1| < r\varepsilon^\ell, \quad |I_2 - \bar{J}_2| < r, \quad I_3 \in \mathbb{R} \right\}, \quad (3.15)$$

where $0 \leq \ell < 1/2$, $r > 0$. The parameters \bar{J}_i , L and the constant r are assumed to satisfy

$$L + 3r\varepsilon^\ell < \bar{J}_1, \quad |\bar{J}_2| + 3r(\varepsilon^\ell + 1) < \bar{J}_1, \quad (3.16)$$

so that $0 < \kappa_i < 1$ and the ν_i 's are well defined on the domain A .

The Hamiltonian $H_{\varepsilon,\mu}$ in (3.14) will be called the “resonant D’Alembert Hamiltonian” and, in the rest of this thesis, we shall consider only the resonant D’Alembert Hamiltonian defined on the domain $A \times \mathbb{T}^3$.

3.2 Linear analysis and the effective Hamiltonian

The appearance of the linear combination $(pI_1 - qI_2 + qI_3)$ in the D’Alembert Hamiltonian $H_{\varepsilon,\mu}$ suggests to look for a linear symplectic (ε -independent) change of variables casting $H_{\varepsilon,\mu}$ in a simpler and more informative form. Calling

$$\Phi_L : (I', \varphi') \rightarrow (I, \varphi) = \Phi_L(I', \varphi') \quad (3.17)$$

such a linear change of variables, it is quite natural to set⁴⁹

$$I'_3 \equiv pI_1 - qI_2 + qI_3. \quad (3.18)$$

Besides (3.18) we shall also require the following condition, which needs a little explanation (given below):

$$\int_0^{2\pi} F_0 \circ \Phi_L(I', \varphi') \frac{d\varphi'_3}{2\pi} = \text{function depending on } I' \quad (3.19)$$

and at most on one angle .

The idea beyond these conditions is the following. The unperturbed frequencies of the transformed Hamiltonian (i.e., $\nabla_{I'} H_{0,0} \circ \Phi_L$) are given by⁵⁰

$$I_1(I') \nabla_{I'} I_1(I') + (0, 0, \omega) = (0, 0, \omega) + O(\varepsilon^\ell).$$

⁴⁹Clearly, the choice of the index 3 is arbitrary.

⁵⁰Recall that in our domain A_ε , (3.15), I_1 has been taken of order ε^ℓ . The precise quantitative analysis will be described in the next section, where we will also assume that $\mu \leq \varepsilon^c$ for some $c > 0$.

This implies that φ'_1 and φ'_2 are “slow” angles, while φ'_3 is a “fast” angle so that φ'_3 “averages out” (see part III for a precise mathematical statement) leaving an “effective Hamiltonian” given by

$$\begin{aligned} H_{\text{eff}} &\equiv \int_0^{2\pi} H_{\varepsilon,0} \circ \Phi_L(I', \varphi') \frac{d\varphi'_3}{2\pi} \\ &= \frac{I_1(I')^2}{2} + \omega I'_3 + \varepsilon \int_0^{2\pi} F_0 \circ \Phi_L(I', \varphi') \frac{d\varphi'_3}{2\pi}, \end{aligned} \quad (3.20)$$

which, in view of (3.19), is a one-degree-of-freedom Hamiltonian (and hence integrable): H_{eff} depends, possibly, on all the actions I'_i but, because of (3.19), *on at most one angle* (φ'_1 or φ'_2). In the case H_{eff} depends explicitly on one angle, say φ'_1 , then the actions I'_2 and I'_3 are just parameters for the dynamics generated by H_{eff} .

The rest of this section is devoted to find linear symplectic diffeomorphisms, Φ_L , of $A \times \mathbb{T}^3$ satisfying (3.19) and the upshot will be that *if $p : q$ is different from $1 : 1$ or $2 : 1$, then H_{eff} depends only on the action variables while in the other cases H_{eff} depends explicitly on one angle also*: in the first case the phase portrait of the integrable system associated to H_{eff} is entirely foliated by (homotopically non trivial) invariant curves while in the latter case there are also hyperbolic equilibria, separatrices and curves with different topology (exactly as in the phase portrait of the standard pendulum).

The linear symplectic diffeomorphism Φ_L has a generating function given, up to an arbitrary (and meaningless) plus or minus sign, by⁵¹

$$\begin{aligned} S(I, \varphi') &\equiv MI \cdot \varphi', \quad \text{with } M \in SL(3, \mathbb{Z}), \\ I &= M^{-1}I', \quad \varphi = M^T \varphi'. \end{aligned} \quad (3.21)$$

The relation (3.18) means that M has the form

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ p & -q & q \end{pmatrix} \quad (3.22)$$

with integers a, \dots, f to be determined. Thus, by (3.21) and (3.22), we have that

$$\varphi_1 = a\varphi'_1 + d\varphi'_2 + p\varphi'_3, \quad \varphi_2 = b\varphi'_1 + e\varphi'_2 - q\varphi'_3, \quad \varphi_3 = c\varphi'_1 + f\varphi'_2 + q\varphi'_3. \quad (3.23)$$

By e) above and (3.23), we find

$$\begin{aligned} \int_0^{2\pi} F_0(I_1, I_2, \varphi(\varphi')) \frac{d\varphi'_3}{2\pi} &= \int_0^{2\pi} \left[\sum_{|j| \leq 2} c_j \cos j(a\varphi'_1 + d\varphi'_2 + p\varphi'_3) \right. \\ &\quad \left. + d_j \cos \left((aj + 2b)\varphi'_1 + (dj + 2e)\varphi'_2 + (pj - 2q)\varphi'_3 \right) \right] \frac{d\varphi'_3}{2\pi} \\ &= c_0 + \sum_{\substack{|j| \leq 2 \\ pj = 2q}} d_j \cos \left((aj + 2b)\varphi'_1 + (dj + 2e)\varphi'_2 \right). \end{aligned} \quad (3.24)$$

⁵¹ $SL(3, \mathbb{Z})$ denotes the group of real (3×3) matrices with integer entries and determinant one; the superscript T denotes matrix transposition.

If (p, q) is different from $(1, 1)$ and $(2, 1)$ we see that there are no integers j with $|j| \leq 2$ such that $pj = 2q$, so that, in this case, the sum in the last line of (3.24) is absent and we have that H_{eff} depends only on the action variables and is given by

$$H_{\text{eff}}(I(I'), \varphi(\varphi')) = \frac{I_1(I')^2}{2} + \omega I'_3 + \varepsilon c_0(I_1(I'), I_2(I')) . \quad (3.25)$$

Next we show that when (p, q) is equal to $(1, 1)$ or $(2, 1)$, then H_{eff} cannot be as in (3.25) and it must depend explicitly on one angle (φ'_1 or φ'_2).

Let us consider first the case $(p, q) = (1, 1)$. In this case, $pj = 2q$ means $j = 2$ and (3.24) implies that

$$\int_0^{2\pi} F_0(I_1, I_2, \varphi(\varphi')) \frac{d\varphi'_3}{2\pi} = c_0 + d_2 \cos(2(a+b)\varphi'_1 + 2(d+e)\varphi'_2) . \quad (3.26)$$

Thus, H_{eff} independent on angles means

$$a + b = 0 = d + e ,$$

a relation which makes the first two columns of the matrix M proportional (one is the opposite of the other) and this implies that the determinant of M would vanish.

The case $(p, q) = (2, 1)$ is similar: $pj = 2q$ means $j = 1$ and (3.24) implies that

$$\int_0^{2\pi} F_0(I_1, I_2, \varphi(\varphi')) \frac{d\varphi'_3}{2\pi} = c_0 + d_1 \cos((a+2b)\varphi'_1 + (d+2e)\varphi'_2) . \quad (3.27)$$

Thus, H_{eff} independent on angles means

$$a + 2b = 0 = d + 2e ,$$

a relation which, as above, makes the first two columns of the matrix M be one the opposite of the other, implying, again, the vanishing of the determinant of M .

Remark 3.2 In what follows we shall make particular (and “convenient”) choices for the matrix M (and hence for the symplectic transformation Φ_L), but one should bear in mind that in doing this there is quite a bit of freedom but that the physical relevant quantities (such as H_{eff}) are essentially intrinsic.

In the case $p = 1, 2$ and $q = 1$, by the above analysis, we see that (3.19) is satisfied *provided*

$$\text{either } a + pb = 0 \quad \text{or} \quad d + pe = 0 .$$

We then take $d = 0 = e$ and⁵²

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ p & -1 & 1 \end{pmatrix} , \quad (p = 1, 2) , \quad (3.28)$$

⁵²In [49], where it is studied the resonant D'Alembert Hamiltonian when $(p, q) = (2, 1)$, it is taken $M = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ p & -1 & 1 \end{pmatrix}$.

leading to the linear symplectic transformation

$$\Phi_L : (I', \varphi') \rightarrow \begin{cases} I = (I'_1, pI'_1 + I'_2 - I'_3, I'_2) \\ \varphi = (\varphi'_1 + p\varphi'_3, -\varphi'_3, \varphi'_2 + \varphi'_3) \end{cases}, \quad (p = 1, 2). \quad (3.29)$$

In the new coordinates the Hamiltonian becomes

$$\begin{aligned} H_{\varepsilon, \mu} \circ \Phi_L &= \frac{(I'_1)^2}{2} + \omega I'_3 + \varepsilon F_0(I'_1, pI'_1 + I'_2 - I'_3, \varphi'_1 + p\varphi'_3, -\varphi'_3) \\ &\quad + \varepsilon \mu F_1(I'_1, pI'_1 + I'_2 - I'_3, \varphi'_1 + p\varphi'_3, -\varphi'_3, \varphi'_2 + \varphi'_3; \mu) \\ &\equiv H_{00}(I'_1, I'_3) + \varepsilon G_0(I', \varphi') + \varepsilon \mu G_1(I', \varphi'; \mu); \end{aligned} \quad (3.30)$$

and the averaged resonant D'Alembert Hamiltonian is (recall (3.26) and (3.27))

$$\begin{aligned} H_{\text{eff}}(I', \varphi'_1; \varepsilon) &\equiv \int_0^{2\pi} H_{\varepsilon, 0} \circ \Phi_L(I', \varphi') \frac{d\varphi'_3}{2\pi} \\ &= \frac{(I'_1)^2}{2} + \omega I'_3 \\ &\quad + \varepsilon \left\{ c_0(I'_1, pI'_1 + I'_2 - I'_3) + d_{j_p}(I'_1, pI'_1 + I'_2 - I'_3) \cos(j_p \varphi'_1) \right\} \\ &\equiv H_{00}(I'_1, I'_3) + \varepsilon H_{01}(I', \varphi'_1), \end{aligned} \quad (3.31)$$

where $j_1 \equiv 2$ and $j_2 \equiv 1$.

Let us turn to the case in which (p, q) is different from $(1, 1)$ and $(2, 1)$. In this case, as discussed above, (3.19) is always satisfied and H_{eff} does not depend on angles. To make a particular choice, let a and b be integers such that

$$aq + bp = 1. \quad (3.32)$$

In view of an elementary algebraic identity⁵³, such (infinitely many) integers always exist and we shall fix the ones that minimize the sum $|a| + |b|$. Then, we define

$$M = \begin{pmatrix} a & b & -b \\ 0 & 1 & 0 \\ p & -q & q \end{pmatrix}, \quad (p, q) \neq (1, 1), (2, 1), \quad (3.33)$$

leading to the linear symplectic transformation

$$\Phi_L : (I', \varphi') \rightarrow \begin{cases} I = (qI'_1 + bI'_3, I'_2, -pI'_1 + I'_2 + aI'_3) \\ \varphi = (q\varphi'_1 - p\varphi'_3, \varphi'_2 + \varphi'_3, b\varphi'_1 + a\varphi'_3) \end{cases}. \quad (3.34)$$

In the new coordinates the resonant D'Alembert Hamiltonian becomes

$$\begin{aligned} H_{\varepsilon, \mu} \circ \Phi_L &= \frac{(qI'_1 + bI'_3)^2}{2} + \omega I'_3 \\ &\quad + \varepsilon F_0(qI'_1 + bI'_3, I'_2, q\varphi'_1 - p\varphi'_3, \varphi'_2 + \varphi'_3) \\ &\quad + \varepsilon \mu F_1(qI'_1 + bI'_3, I'_2, q\varphi'_1 - p\varphi'_3, \varphi'_2 + \varphi'_3, b\varphi'_1 + a\varphi'_3; \mu) \\ &\equiv H_{00}(I'_1, I'_3) + \varepsilon G_0(I', \varphi') + \varepsilon \mu G_1(I', \varphi'; \mu), \end{aligned} \quad (3.35)$$

⁵³The so-called "Bezout identity", which is an immediate consequence of the Euclidean algorithm.

and, in this case, the averaged resonant D'Alembert Hamiltonian is simply

$$\begin{aligned}
H_{\text{eff}}(I'; \varepsilon) &\equiv \int_0^{2\pi} H_{\varepsilon,0} \circ \Phi_L(I', \varphi') \frac{d\varphi'_3}{2\pi} \\
&= \frac{(qI'_1 + bI'_3)^2}{2} + \omega I'_3 + \varepsilon c_0(qI'_1 + bI'_3, I'_2) \\
&\equiv H_{00}(I'_1, I'_3) + \varepsilon H_{01}(I') .
\end{aligned} \tag{3.36}$$

Remark 3.3 Notice that we are using a unified notation for different objects (such as G_i or H_{00} or H_{01}), which, in fact, depend explicitly on the resonance (p, q) .

Part II

**Total stability: a model problem for
properly-degenerate Hamiltonian
Systems with two degrees of freedom**

4 The total stability Theorems

In this part we will consider a (real-analytic) nearly-integrable, one-parameter family of Hamiltonian functions $H(I, \varphi; \varepsilon) = h(I) + \varepsilon f(I, \varphi)$ where $(I, \varphi) \in \Omega \times \mathbb{T}^2$, with $\Omega \subset \mathbb{R}^2$, are standard symplectic “action-angle” variables and ε is a small parameter.

As pointed out in the introduction a typical feature in Celestial Mechanics is that the unperturbed system is *properly degenerate*, i.e., the unperturbed Hamiltonian function $h(I)$ does not depend upon all action variables. In such a case it is obviously strongly violated the non-degeneracy KAM condition that is the fact that, at fixed energy, the (unperturbed) map between action variables on fixed energy surface and the frequency map viewed in projective space is a diffeomorphism.

However, as we said in the introduction, if one considers a nearly-integrable (real-analytic) Hamiltonian system with two degrees of freedom governed by

$$H(I, \varphi; \varepsilon) := H_0(I; \varepsilon) + \varepsilon^2 H_1(I, \varphi) := H_{00}(I_1) + \varepsilon H_{01}(I) + \varepsilon^2 H_1(I, \varphi) , \quad (4.1)$$

in [9] (compare also [8], Chapter 5, Section 3), Arnold proved the following Theorem (see Theorem 0.1):

Theorem 4.1 *If the “perturbation removes the degeneracy”⁵⁴ on the energy level $H^{-1}(E)$, namely*

$$\frac{\partial H_{00}}{\partial I_1}(I) \neq 0 , \quad \frac{\partial^2 H_{01}}{\partial I_2^2}(I) \neq 0 , \quad \forall I \in H_0^{-1}(E) , \quad (4.2)$$

then, for all ε small enough, total stability holds; namely for all initial data on the given energy level, the values of the action-variables stay forever near their initial values.

Remark 4.1 (i) If condition (4.2) is violated “instability channels” may appear as suggested by the following example (which is a trivial modification of an example due to N.N. Nekhoroshev [90]). Let

$$H_{00}(I_1) + \varepsilon H_{01}(I_2) := \frac{I_1^2}{2} - \varepsilon \frac{I_2^2}{2} , \quad (4.3)$$

and notice that (the first inequality in) condition (4.2) is violated on each energy level crossing the axis $\{I_1 = 0\}$ (in particular is violated at $E = 0$). Then, one can construct a sequence $\varepsilon_j \downarrow 0$ and a sequence of perturbations $H_{1,j}(\varphi)$ with $\sup_{|\operatorname{Im} \varphi_i| \leq 1} |H_{1,j}(\varphi)|$ uniformly

bounded such that

$$I_\varepsilon(t) := e^{-1/\sqrt{\varepsilon}} \left(-\varepsilon^2 t, \varepsilon^{3/2} t \right) , \quad \varphi_\varepsilon(t) := e^{-1/\sqrt{\varepsilon}} \left(-\varepsilon^2 \frac{t^2}{2}, -\varepsilon^{5/2} \frac{t^2}{2} \right) , \quad (4.4)$$

⁵⁴Or, more precisely, that “the intermediate term H_{01} removes the degeneracy”.

is a solution of the Hamilton equation associated to $H_{00}(I_1) + \varepsilon H_{01}(I_2) + \varepsilon^2 H_{1,j}(\varphi)$ when $\varepsilon = \varepsilon_j$. In fact, it is enough to take

$$\varepsilon_j := j^{-2}, \quad H_{1,j}(\varphi) := e^{-j} \sin(\varphi_1 - j\varphi_2).$$

Notice that a displacement of order one of the action variables $I_\varepsilon(t)$ with respect to their initial value $I_\varepsilon(0) = (0, 0)$ occurs in the exponentially long time $\sim \exp(1/\sqrt{\varepsilon_j})/\varepsilon_j^2$.

(ii) Condition (4.2) is violated, at $E = 0$, also by the “convex” Hamiltonian $H_0 := \frac{I_1^2}{2} + \varepsilon \frac{I_2^2}{2}$, ($\varepsilon > 0$). However, in such a case, $H_0^{-1}(0)$ consists only of one point and exploiting convexity (and using energy conservation arguments), it is not difficult to show that, also on the energy level $E = 0$, total stability holds for $\varepsilon > 0$ small enough. It is therefore clear that “convexity” (or, more in general, “steepness”) should play a fundamental role in this business.

Properly degenerate systems with two degrees of freedom of the form (4.1), are, in general, “more integrable” than non-degenerate systems, as A.I. Nejshtadt proved in 1981:

Theorem 4.2 ([89]) *Assume that a (real-analytic) properly degenerate system with two degrees of freedom satisfies condition (4.2) together with $\frac{\partial H_{01}}{\partial I_2} \neq 0$. Then the measure of the set of unperturbed tori that disappear when $\varepsilon > 0$ is exponentially small (i.e. $O(\exp(-\text{const}/\varepsilon))$ rather than $O(\sqrt{\varepsilon})$ as in general nondegenerate systems). Furthermore the deviation of a perturbed torus from the unperturbed one is of $O(\varepsilon)$ (rather than $O(\sqrt{\varepsilon})$).*

We take up the action–stability problem for properly degenerate Hamiltonian system with two degrees of freedom allowing the intermediate system H_{01} to depend also on the angle φ_1 . Thus, we shall consider real-analytic, properly-degenerate systems with two degrees of freedom described by nearly-integrable, real-analytic Hamiltonians given by

$$H(I, \varphi; \varepsilon) := H_{00}(I_1) + \varepsilon H_{01}(I, \varphi_1) + \varepsilon^a H_1(I, \varphi), \quad 0 < \varepsilon \ll 1, \quad a > 1. \quad (4.5)$$

The interest for such systems stems again from Celestial Mechanics. In fact, as we have seen before, the planetary D’Alembert model is governed, up to an exponentially small term, by the Hamiltonian (0.10) which is of the form (4.5).

We want to study *model problems* able to capture the main features of “general” properly degenerate systems with two degrees of freedom and, in particular, the features of the above-mentioned Hamiltonian (0.10). As we explained in the introduction, in order to avoid “extra” technical difficulties, we shall take

$$H_{00} := \frac{I_1^2}{2}, \quad H_{01} := H_{01}^{(\sigma)} := \sigma \frac{I_2^2}{2} - (1 + \cos \varphi_1), \quad (4.6)$$

with σ equal either $+1$ or -1 ; the phase space will be taken to be $\mathcal{M}_{R_0} := B_{R_0}^2 \times \mathbb{T}^2$ where $B_{R_0}^2$ denotes a ball of radius R_0 around the origin.

We can now state our main results. Denote, as above, by $(I(t), \varphi(t)) := \phi_H^t(I_0, \varphi_0)$ the time t evolution of the initial data $(I(0), \varphi(0)) := (I_0, \varphi_0)$ governed by the Hamiltonian H . We shall prove the following

Theorem 4.3 Let $H^{(\sigma)}(I, \varphi; \varepsilon) := H(I, \varphi; \varepsilon)$ and \mathcal{M}_{R_0} be as in (4.5), (4.6). Assume $a > 3/2$ and choose

$$0 < R < R_0 \quad \text{and} \quad 0 < b < \min \left\{ \frac{1}{4}, \frac{a-1}{4}, \frac{1}{3} \left(a - \frac{3}{2} \right) \right\}. \quad (4.7)$$

Then, there exists $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, the ϕ_H^t -evolution $(I(t), \varphi(t))$ of an initial datum (I_0, φ_0) satisfies

$$|I(t)| < R_0, \quad |I(t) - I_0| < \varepsilon^b, \quad \forall t \in \mathbb{R}, \quad (4.8)$$

where, in the case $\sigma = 1$, (I_0, φ_0) is an arbitrary point in the phase space \mathcal{M}_R , while, in the “non-convex” case $\sigma = -1$, (I_0, φ_0) belongs to $\mathcal{M}_R \setminus \mathcal{N}_*$, \mathcal{N}_* being an open region whose measure does not exceed $\varepsilon^{2/3}$.

This theorem will be a simple corollary of the following result, which describes the distribution and density of KAM tori. Let $H_{\mathcal{P}}$ denote the pendulum Hamiltonian⁵⁵

$$H_{\mathcal{P}} := H_{\mathcal{P}}(I_1, \varphi_1; \varepsilon) := \frac{I_1^2}{2} - \varepsilon(1 + \cos \varphi_1). \quad (4.9)$$

Theorem 4.4 Let the hypotheses and choices of Theorem 4.3 hold and let $\mathcal{M}^{(\sigma)} := \mathcal{M}_R \setminus \mathcal{N}^{(\sigma)}$ where the sets $\mathcal{N}^{(\sigma)} := \mathcal{N}^{(\sigma)}(\varepsilon, b)$ are defined by

$$\begin{aligned} \mathcal{N}^{(1)} &:= \left\{ (I, \varphi) : |H_{\mathcal{P}}| < \varepsilon^{1+2b} \text{ or } H_{\mathcal{P}} < -2\varepsilon + \varepsilon^{1+2b} \right\} \cup \left\{ (I, \varphi) : |I_2| < R\varepsilon^b \right\} \\ \mathcal{N}_* &:= \left\{ (I, \varphi) : c \varepsilon^{\frac{2}{3} + \frac{4}{3}b} < H_{\mathcal{P}} < \frac{\varepsilon^{\frac{2}{3}}}{c} \right\}, \\ \mathcal{N}^{(-1)} &:= \mathcal{N}^{(1)} \cup \mathcal{N}_*, \end{aligned} \quad (4.10)$$

$0 < c < 1$ being a suitable constant. Fix q such that

$$0 < q < a - \frac{3}{2} - 3b. \quad (4.11)$$

Then, there exists $\varepsilon_0 > 0$ such that, for all $0 < \varepsilon < \varepsilon_0$, the following holds. Apart from a small dense subset of measure $O(\exp(-1/\varepsilon^q))$, the region $\mathcal{M}^{(\sigma)}$, is filled up by two-dimensional, real-analytic $H^{(\sigma)}$ -invariant tori; each of these tori is $O(\exp(-1/\varepsilon^q))$ -close to an unperturbed torus $\{(I_1, \varphi_1) : H_{\mathcal{P}} = E\} \times \{(I_2, \varphi_2) \text{ s.t. } I_2 = \text{const}\}$ in $\mathcal{M}^{(\sigma)}$. Furthermore, for any motion $(I(t), \varphi(t))$ in $\mathcal{M}^{(\sigma)}$, the displacement of $I(t)$ from its initial value I_0 is bounded, for all times t , by $\sqrt{\varepsilon}$.

Remark 4.2 (i) By simple energy-conservation argument one sees immediately that $|I_1(t) - I_1(0)| < \text{const } \sqrt{\varepsilon}$ for any motion $(I(t), \varphi(t))$ in \mathcal{M}_R ; thus the “stability” statement in (7.6) concerns actually only the I_2 action variable.

⁵⁵ $H_{\mathcal{P}}$ is a standard mathematical pendulum having the stable equilibrium in $(0, 0)$ with energy -2ε , the unstable equilibrium in $(0, \pm\pi)$ with energy 0 (hence the separatrix as well has energy 0).

(ii) The discarded region $\mathcal{N}^{(\sigma)}$ is a (“elementary”) set small with ε . If we replace $\mathcal{N}^{(\sigma)}$ by a small set of order one (say, $\{(I_1, \varphi_1) : |H_{\mathcal{P}}| < \delta\} \times \{(I_2, \varphi_2) : |I_2| < \delta\}$ for a fixed $0 < \delta \ll 1$), then the displacement of $I(t)$ from its initial value I_0 is bounded by ε .

(iii) In the two-degrees-of-freedom case considered here, as mentioned above, the 2-dimensional KAM tori constructed in Theorem 4.4 (which fill, up to an exponentially small set, the region $\mathcal{M}^{(\sigma)}$) *separate* the three-dimensional energy levels. Thus, the topological “trapping” argument may be applied leading to stability, for all times, of the action variables in $\mathcal{M}^{(\sigma)}$. Then, an elementary *energy-conservation argument* implies action stability in \mathcal{M}_R or in $\mathcal{M}_R \setminus \mathcal{N}_*$ (according to whether $\sigma = 1$ or $\sigma = -1$).

(iv) In the case $a = 2$ one can take any $0 < b < 1/6$ and $q < \frac{1}{2} - 3b$.

(v) Theorem 4.3 and Theorem 4.4 may be viewed as extensions, in the model cases considered here, of, respectively, Theorem 4.1 and 4.2.

The main ideas of the proof are the following. We consider first the Hamiltonian

$$H_0 := H_{00} + \varepsilon H_{01} = H_{\mathcal{P}} + \sigma I_2^2/2, \quad (4.12)$$

in which we introduce action-angle variables for the pendulum $H_{\mathcal{P}}$ and try to apply KAM techniques in order to confine all motions among KAM tori. Obviously the action-angle variables for the pendulum $H_{\mathcal{P}}$ are singular in any neighbourhood of the separatrix (and stable equilibria) and it is exactly near separatrices where one expects the motion to become “chaotic” and where, in principle, drift of order one in the I_2 variable is conceivable⁵⁶. Therefore a careful analysis near these “singular phase space regions” is needed and arguments different from KAM theory have to be used to control the displacement of the action variable in such singular regions. Obviously, as discussed in Remark 4.1, regions where the non-degeneracy assumption fails need a separate discussion: in fact, in the non convex case, we are not able to exclude a “possibly non-chaotic-drift” of the I_2 action in the regions $\mathcal{N}^{(\sigma)}$.

5 Technical lemmas

The construction of KAM tori in $\mathcal{M}^{(\sigma)}$ is based on the following three lemmata: the first lemma provides (real-analytic) action-angle variable for the pendulum slightly away from the separatrix and the stable equilibrium; the second lemma is a “normal form lemma”; the third lemma is a “iso-energetic” KAM theorem. For general information about normal forms, KAM theory, etc, we refer to, e.g., [8] and references therein.

In the following we shall use the following notations: if $A \subset \mathbb{R}^d$ and $r > 0$, we denote by A_r the subset of points in \mathbb{C}^d at distance less than r from A ; \mathbb{T}_s^d denotes the complex set $\{z \in \mathbb{C}^d : |\operatorname{Im} z_j| < s \text{ for all } j\}$ (thought of as a complex neighborhood of \mathbb{T}^d). If $f(I, \varphi)$ is

⁵⁶Better: “compatible with energy conservation”.

a real analytic function on $A_r \times \mathbb{T}_s^d$ we let $\|f\|_{r,s}$ denote the following norm⁵⁷

$$\|f\|_{r,s} := \sum_{k \in \mathbb{Z}^d} \sup_{I \in A_r} |\hat{f}_k(I)| e^{|k|s}, \quad (5.1)$$

$\hat{f}_k(I)$ being the Fourier coefficients of the periodic function $\varphi \rightarrow f(I, \varphi)$.

Lemma 5.1 (Real-analytic action-angle variables for the pendulum)

Let $D^0 := [-R_0, R_0]$, let $E_0 := H_{\mathcal{P}}(R_0, 0) = R_0^2/2$, let $0 < \eta < \varepsilon/32$ and define

$$\begin{aligned} \mathcal{M}_p^+ &:= \mathcal{M}_p^+(\eta, \varepsilon) := \left\{ (I_1, \varphi_1) \in D^0 \times \mathbb{T} : I_1 > 0, \eta < H_{\mathcal{P}}(I_1, \varphi_1) < E_0 - \varepsilon \right\}, \\ \mathcal{M}_p^- &:= \mathcal{M}_p^-(\eta, \varepsilon) := \left\{ (I_1, \varphi_1) \in D^0 \times \mathbb{T} : -2\varepsilon + \eta < H_{\mathcal{P}}(I_1, \varphi_1) < -\eta \right\}. \end{aligned} \quad (5.2)$$

Then, for all $r_* < R_0/2$ and s_* positive, there exist positive numbers r_0 and s_0 , closed intervals $D^\pm \subset \mathbb{R}$, symplectic transformations ϕ^\pm real-analytic on $D^\pm \times \mathbb{T}$ and functions h^\pm real-analytic on D^\pm such that

$$\phi^\pm : (\hat{I}_1, \hat{\varphi}_1) \in D_{r_0}^\pm \times \mathbb{T}_{s_0} \rightarrow \phi^\pm(\hat{I}_1, \hat{\varphi}_1) \in D_{r_*}^0 \times \mathbb{T}_{s_*}, \quad (5.3)$$

$$\phi^\pm(D^\pm \times \mathbb{T}) = \mathcal{M}^\pm(\eta, \varepsilon), \quad (5.4)$$

$$H_{\mathcal{P}} \circ \phi^\pm(\hat{I}_1, \hat{\varphi}_1) = h^\pm(\hat{I}_1), \quad \forall (\hat{I}_1, \hat{\varphi}_1) \in D_{r_0}^\pm \times \mathbb{T}_{s_0}. \quad (5.5)$$

The analyticity radii r_0 and s_0 may be taken to be

$$r_0 := c r_* \frac{\eta}{\sqrt{\varepsilon}}, \quad s_0 := c s_* \frac{1}{\ln(\varepsilon/\eta)}, \quad (5.6)$$

where $0 < c < 1$ is a suitable (universal) constant. Furthermore, the functions h^\pm satisfy, for all $\hat{I}_1 \in D_{r_0}^\pm$, the following bounds

$$\eta \leq \text{Re} h^+(\hat{I}_1) \leq E_0 - \varepsilon, \quad -2\varepsilon + \eta \leq \text{Re} h^-(\hat{I}_1) \leq -\eta, \quad (5.7)$$

$$\frac{dh^\pm}{d\hat{I}_1}(\hat{I}_1) = \frac{\alpha^\pm}{\pi_1^\pm}, \quad (5.8)$$

$$\frac{d^2 h^\pm}{d\hat{I}_1^2}(\hat{I}_1) = \pm \beta^\pm \frac{\pi_2^\pm}{(\pi_1^\pm)^3}, \quad (5.9)$$

where

$$\begin{aligned} \pi_1^\pm &:= \pi_1^\pm(\hat{I}_1) := \frac{1}{\sqrt{\varepsilon}} \ln \left(1 + \sqrt{\frac{\varepsilon}{|\text{Re} h^\pm(\hat{I}_1)|}} \right), \\ \pi_2^\pm &:= \pi_2^\pm(\hat{I}_1) := \frac{1}{|\text{Re} h^\pm(\hat{I}_1)|} \frac{1}{\sqrt{|\text{Re} h^\pm(\hat{I}_1)| + \varepsilon}}, \end{aligned} \quad (5.10)$$

⁵⁷ The specific choice of norm will play no role in the sequel; obviously if f is a real-analytic function on \mathbb{T}_s^d , $\|f\|_s$ stands for $\sum_{k \in \mathbb{Z}^d} |\hat{f}_k| e^{|k|s}$, \hat{f}_k being the Fourier coefficients of f , while, if f is a real-analytic function on A_r , then $\|f\|_r = \sup_{I \in A_r} |f(I)|$.

$\alpha^\pm := \alpha^\pm(\hat{I}_1)$, $\beta^\pm := \beta^\pm(\hat{I}_1)$ are real-analytic functions such that

$$d_1 \leq \frac{\operatorname{Re}(\alpha^\pm)}{R_0}, \operatorname{Re}(\beta^\pm) \leq d_2; \quad \left| \frac{\operatorname{Im}(\alpha^\pm)}{R_0} \right|, |\operatorname{Im}(\beta^\pm)| \leq d_0 \quad (5.11)$$

for suitable (universal) constants $0 < d_1 < d_2$ and $0 < d_0 < d_1/10$. An identical statement holds if, in the definition of \mathcal{M}_p^+ , one replaces “ $I_1 > 0$ ” with⁵⁸ “ $I_1 < 0$ ”.

Lemma 5.2 (Normal forms) *Let $\hat{\Delta}$ and $\hat{\Delta}'$ be two subsets of \mathbb{R} and consider a Hamiltonian function $H(\hat{I}, \hat{\varphi}) := h(\hat{I}) + f(\hat{I}, \hat{\varphi})$ real-analytic on $\widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}} := (\hat{\Delta}_{\hat{r}_1} \times \hat{\Delta}'_{\hat{r}_2}) \times \mathbb{T}_{\hat{s}}^2$ for some $\hat{r}_2 \geq \hat{r}_1 > 0$ and $\hat{s} > 0$. Assume that there exist $K \geq 6/\hat{s}$ and $\alpha > 0$ such that*

$$|\omega(\hat{I}) \cdot k| \geq \alpha, \quad \forall k \in \mathbb{Z}^2, \quad 0 < |k| \leq K, \quad \forall \hat{I} \in \hat{\Delta}_{\hat{r}_1} \times \hat{\Delta}'_{\hat{r}_2}, \quad (5.12)$$

where $\omega(\hat{I}) := \nabla h(\hat{I})$. Assume also that⁵⁹

$$\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \leq \frac{\alpha \hat{r}_1}{2^{10} K}. \quad (5.13)$$

Then, there exist a real-analytic symplectic transformation

$$\Phi : (J, \psi) \in \widehat{W}_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \rightarrow \Phi(J, \psi) \in \widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}}$$

such that

$$H \circ \Phi(J, \psi) = h(J) + g(J) + f_*(J, \psi) \quad (5.14)$$

with⁶⁰

$$\begin{aligned} \|g - f_0\|_{\hat{r}_1/2, \hat{r}_2/2} &\leq \frac{1}{3} \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}}, \\ \|f_*\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} &\leq \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \exp(-K\hat{s}/6), \\ \|\Phi(J, \psi) - (J, \psi)\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} &\leq \hat{c} \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}}, \end{aligned} \quad (5.15)$$

where $\hat{c} > 0$ is a suitable constant.

The next lemma is a typical statements from KAM theory (see [8] for generalities). It is an iso-energetic KAM theorem (i.e. a KAM theorem on fixed energy levels).

Let, as above, $\omega(J)$ denote the gradient $\nabla h(J)$, let $h''(J)$ denote the Hessian matrix of h . We recall that a vector $\omega \in \mathbb{R}^d$ is said to be (γ, τ) -Diophantine if

$$|\omega \cdot k| \geq \frac{\gamma}{|k|^\tau}, \quad \forall k \in \mathbb{Z}^d \setminus \{0\}, \quad (5.16)$$

for some $\gamma > 0$ and⁶¹ $\tau > 0$.

⁵⁸By symmetry, the interval D^+ in the case $I_1 < 0$ is just the opposite of the interval D^+ in the case $I_1 > 0$.

⁵⁹Adapt the norms in (5.1) and in the footnote 57 in the obvious way replacing A_r by $\hat{\Delta}_{\hat{r}_1} \times \hat{\Delta}'_{\hat{r}_2}$ (and replacing the subscript “ r ” in the norms by “ \hat{r}_1, \hat{r}_2 ”).

⁶⁰ f_0 is the zero-Fourier coefficient of f , i.e., the average of $f(\hat{I}, \hat{\varphi})$ over \mathbb{T}^2 .

⁶¹Necessarily $\tau \geq d - 1$ by a theorem of Liouville. Also, (5.16) (with $|k| = 1$) implies that $\gamma \leq \min_i |\omega_i|$.

Lemma 5.3 (Iso-energetic KAM theorem) *Let $D \subset \mathbb{R}^d$ be a bounded domain and consider a Hamiltonian $H(J, \psi) := h(J) + f(J, \psi)$ real-analytic on the domain $W_{r,s} := D_r \times \mathbb{T}_s^d$ for some $r > 0$ and $s > 0$. Assume that $\|h''\|_r > 0$ and that the $(d+1) \times (d+1)$ matrix*

$$U := \begin{pmatrix} h''(J) & \omega(J) \\ \omega(J) & 0 \end{pmatrix} \quad (5.17)$$

is invertible on D_r . Given $E \in \mathbb{R}$ (such that $h^{-1}(E) \neq \emptyset$) and given

$$0 < \gamma < \min_{i, J \in D} |\omega_i(J)| \quad \text{and} \quad \tau \geq d-1, \quad (5.18)$$

denote

$$\mathcal{D} = \{J \in D : h(J) = E \text{ and } \omega(J) \text{ is } (\gamma, \tau) \text{-Diophantine}\}. \quad (5.19)$$

Then, if $\|f\|_{r,s}$ is small enough, for each $J \in \mathcal{D}$, there exists a unique d -dimensional, real-analytic, invariant torus $\mathcal{T} \subset H^{-1}(E)$ which is a graph over the angle ψ , which is close to the torus $\{J\} \times \mathbb{T}^d$ and on which the H -flow is analytically conjugated to the translation $\theta \rightarrow \theta + \omega(J)(1 + \kappa)t$, κ being a small real number. More precisely, let A , F and G be positive numbers such that

$$A \geq \|h''\|_r, \quad F \geq A \|f\|_{r,s} \gamma^{-2}, \quad G \geq \max\{A \|U^{-1}\|_r, 1\}, \quad (5.20)$$

let $0 < \bar{s} < s$ and let

$$C := \max\left\{1, \frac{\gamma (s - \bar{s})^{c_1}}{c_2 A r |\ln F|^{c_3}}\right\}, \quad \widehat{F} := C c_4 \frac{1}{(s - \bar{s})^{c_5}} G^{c_6} F |\ln F|^{c_7}, \quad (5.21)$$

where the $c_i > 1$ are suitable constants depending only upon τ and d . If $\widehat{F} \leq 1$, then, for each $J \in \mathcal{D}$, there exists a unique invariant torus $\mathcal{T} \subset H^{-1}(E)$ satisfying the following properties:

- (i) $\mathcal{T} = \{(\mathcal{J}(\psi), \psi) : \psi \in \mathbb{T}_s^d\}$ with \mathcal{J} real-analytic on \mathbb{T}_s^d and $|\mathcal{J}(\psi) - J| \leq r\widehat{F}$ for all $\psi \in \mathbb{T}_s^d$;*
- (ii) there exist real-analytic functions on \mathbb{T}_s^d , u, v and a smooth function $\kappa : D_r \rightarrow \mathbb{C}$ (real for real J) such that*

$$\max\{r^{-1}\|v - J\|_{\bar{s}}, \|u\|_{\bar{s}}, |\kappa|\} \leq \widehat{F};$$

the map $\theta \in \mathbb{T}_s^d \rightarrow (v(\theta), \theta + u(\theta))$ is a real-analytic embedding whose real image is the torus $\mathcal{T} : \mathcal{T} = \{(v(\theta), \theta + u(\theta)), \theta \in \mathbb{T}^d\}$; on the torus \mathcal{T} the H -flow, ϕ^t , linearizes: denoting $\omega_ := (1 + \kappa(J)) \omega(J)$, one has*

$$\phi^t(v(\theta), \theta + u(\theta)) = (v(\theta + \omega_* t), \theta + \omega_* t + u(\theta + \omega_* t));$$

- (iii) if $\tau > d-1$ and $\hat{\gamma} := \left(\text{const.} \frac{\|h'\|_r^d}{\min_{D_r} |\det U|}\right) \gamma$, then*

$$\text{meas}(H^{-1}(E) \setminus \{\text{tori satisfying (i) and (ii)}\}) \leq \hat{\gamma}.$$

Remark 5.1 As mentioned above, in the case of two degrees-of-freedom ($d = 2$) considered in this part, the above KAM tori *separate* the three-dimensional energy levels forming *barriers* for the motion; any two KAM tori (with equal energy) bound an invariant region in corresponding energy level. More precisely, let $[a_1, b_1] \times [a_2, b_2] \subset D$ with $a_i < b_i$. Then, because of (5.18), we can take as coordinates for the three-dimensional energy level $H^{-1}(E)$ either of the action variables⁶² plus the angles ψ . Take first as coordinates (J_1, ψ_1, ψ_2) and fix $\bar{J}_1 \in [a_1 + \delta, a_2 - \delta]$ where $\delta := 2 \max\{r\hat{F}, \hat{\gamma}\}$ ($\hat{\gamma}$ measures the complement of the surviving KAM tori and $r\hat{F}$ the maximal oscillation of the graph of each KAM torus). Then, by (i) and (iii) in Lemma 5.3, it follows that there exist two tori \mathcal{T}' and \mathcal{T}'' so that $\sup_{\psi} \mathcal{J}'_1 < \bar{J}_2 < \inf_{\psi} \mathcal{J}''_1$ and $0 < \inf_{\psi} \mathcal{J}''_1 - \sup_{\psi} \mathcal{J}'_1 \leq O(\delta)$. The same reasoning applies to \bar{J}_2 . Hence, if $(J(t), \psi(t)) := \phi^t(\bar{J}, \bar{\psi})$ (for any $\bar{\psi}$) one has that $\sup_t |J(t) - \bar{J}| \leq O(\delta)$.

Remark 5.2 (On the proofs of the lemmata)

(i) Lemma 5.1 is a subcase of Proposition 2.1 (whit $k(I) := 1$). For completeness we have rewrite it in the notation of this Section. In particular we observe that (5.8), (5.9) and (5.10) are direct consequences of (2.8).

We mention also that for our main purpose (i.e., total stability of action variables) it would be enough to apply a iso-energetic KAM theorem in *smooth* class (since all we need is a topological “trapping argument”); however a *quantitative* version of such a theorem (necessary for our task) is not available in literature and providing the details for its proof would be certainly much longer (and far less elementary) than the proof of Lemma 5.1.

(ii) Lemma 5.2 derives directly from Lemma 1.1 with $d := 2$, $D := \hat{\Delta} \times \hat{\Delta}'$, $\rho := (\hat{r}_1, \hat{r}_2)$, $\sigma := \hat{s}$, $\rho_0 := \hat{r}_1$, $\Lambda := \{0\}$, $\phi := \Phi$, and substituing g with $g + f_0$.

(iii) Lemma 5.3 is by now rather standard. In fact it is easy, under an extra “non-degeneracy condition” satisfied in our application⁶³, to derive the iso-energetic KAM theorem directly by the standard one by means of a standard Implicit Function Theorem. Alternatively, one can find a very detailed version, e.g., in [58]. For these reasons we shall omit the proof of Lemma 5.3. In our application the exact values of the constants c_i are not needed; however we can prove Lemma 5.3 with the following constants:

$$\begin{aligned} c_1 &= \tau + 1, & c_2 &= 2 \cdot 6^{\tau+1}, & c_3 &= c_1, & c_4 &= d \cdot 2^{10}, \\ c_5 &= 2(\tau + 1), & c_6 &= 2, & c_7 &= 2(\tau + 1). \end{aligned}$$

Also, in our case, it will be $C = 1$.

⁶²Furthermore, the map $J_1 \rightarrow \alpha_1(J_1) = \omega_1(J_1, J_2^0(J_1))/\omega_2(J_1, J_2^0(J_1))$, where J_2^0 is such that $h(J_1, J_2^0(J_1)) = E$, is a diffeomorphism:

$$\frac{d\alpha_1}{dJ_1} = - \frac{\det U}{\omega_2^3} \Big|_{(J_1, J_2^0(J_1))};$$

(and a completely symmetric statement holds interchanging the indices 1 and 2).

⁶³Namely, the invertibility of the Hessian h'' on D_r , which is the usual nondegeneracy condition in the standard KAM theorem.

6 Proofs of the Theorems

We first prove Theorem 4.4 (Theorem 4.3 will be a simple corollary of it). Since most of the arguments are identical for both models $\sigma = 1$ and $\sigma = -1$, we shall usually do not indicate explicitly the dependence upon σ . The only point where the two models differ is in the estimates regarding the iso-energetical non-degeneracy (see Lemma 6.1 below).

Proof of Theorem 4.4

The *first step* is to use Lemma 5.1 to put H_0 in (4.12) into action-angle variables. Let R be as in (4.7) and assume that H_1 in (4.5) is analytic on $B_{r_1} \times \mathbb{T}_{s_1}$ where B denotes here $B_{R_0}^2(0)$, and $0 < r_1 < R/2$, $s_1 > 0$. Since, in our case, $H_{\mathcal{P}}$ is an entire function we can choose, in Lemma 5.1 the parameters

$$r_* := r_1, \quad s_* := s_1. \quad (6.1)$$

Let b and q be as in (4.7) and (respectively) (4.11), let

$$\lambda = 1 + 2b, \quad (6.2)$$

and let q_0 be a number such that

$$q < q_0 < a - \frac{3}{2} - 3b. \quad (6.3)$$

Notice that with such choices the followig relations hold:

$$\lambda > 1, \quad 0 < b < \lambda - \frac{1}{2}, \quad b + \lambda + q_0 + \frac{1}{2} < a. \quad (6.4)$$

We also set

$$\eta := \varepsilon^\lambda \quad (6.5)$$

so that r_0 and s_0 in Lemma 5.1 become

$$r_0 = c r_1 \varepsilon^{\lambda-1/2}, \quad s_0 = \frac{c}{\lambda-1} \frac{1}{\ln \varepsilon^{-1}} s_1. \quad (6.6)$$

Let D^0 , D^\pm and ϕ^\pm be as in Lemma 5.1 and let

$$D := [-R_1, R_1] \subset D^0, \quad R_1 := \frac{R_0 + R}{2}. \quad (6.7)$$

Now, define $D^\pm(\sigma) \subset \mathbb{R}$ as follows:

$$D^-(\sigma) := D^-, \quad D^+(1) := D^+, \quad D^+(-1) \times \mathbb{T} := (\phi^+)^{-1}(\mathcal{M}_*^+), \quad (6.8)$$

where

$$\mathcal{M}_*^+ := \mathcal{M}_*^+(\eta, \varepsilon) := \mathcal{M}_p^+ \setminus \mathcal{R}_* := \mathcal{M}_p^+ \setminus \left\{ (I_1, \varphi_1) : c \varepsilon^{\frac{2}{3} + \frac{4}{3}b} \leq H_{\mathcal{P}} \leq \frac{\varepsilon^{\frac{2}{3}}}{c} \right\}, \quad (6.9)$$

with a suitable small positive constant c to be fixed later. Denoting $J = (J_1, J_2)$, $\psi = (\psi_1, \psi_2)$, $\hat{I} = (\hat{I}_1, \hat{I}_2)$, $\hat{\varphi} = (\hat{\varphi}_1, \hat{\varphi}_2)$, then, by Lemma 5.1, we have

$$\begin{aligned} \hat{\phi}^\pm : (\hat{I}, \hat{\varphi}) \in (D^\pm(\sigma)_{r_0} \times D_{r_1}) \times \mathbb{T}_{s_0}^2 &\rightarrow (I, \varphi) \in B_{r_1} \times \mathbb{T}_{s_1}^2, \\ \text{where } (I_1, \varphi_1) &:= \hat{\phi}^\pm(\hat{I}_1, \hat{\varphi}_1), \quad (I_2, \varphi_2) := (\hat{I}_2, \hat{\varphi}_2). \end{aligned} \quad (6.10)$$

In the symplectic coordinates $(\hat{I}, \hat{\varphi})$ the Hamiltonian H in (4.5) takes the form

$$H^\pm(\hat{I}, \hat{\varphi}; \varepsilon) := H \circ \hat{\phi}^\pm(\hat{I}, \hat{\varphi}) = h^\pm(\hat{I}_1) + \varepsilon \sigma \frac{\hat{I}_2^2}{2} + \varepsilon^a H_1^\pm(\hat{I}, \hat{\varphi}; \varepsilon), \quad (6.11)$$

where h^\pm is as in Lemma 5.1 and $H_1^\pm := H_1 \circ \hat{\phi}^\pm$; hence

$$\|H_1^\pm\|_{r_0, r_1, s_0} \leq \|H_1\|_{r_1, s_1}. \quad (6.12)$$

The *second step* is to apply the normal form lemma (Lemma 5.2), in a suitable phase space region, to the Hamiltonian H^\pm : in such a way we shall be able to put H^\pm in a normal form of the type appearing in (5.14)–(5.15), to meet the (stringent) KAM condition, $\hat{F} \leq 1$, in the KAM theorem (Lemma 5.3) and to give a “good” estimates on the measure of the KAM tori. We therefore set⁶⁴

$$\begin{aligned} h(\hat{I}) &:= h^\pm(\hat{I}_1) + \varepsilon \sigma \frac{\hat{I}_2^2}{2}, \quad f := \varepsilon^a H_1^\pm, \\ \hat{r}_1 &:= r_0 = cr_1 \varepsilon^{\lambda - \frac{1}{2}}, \quad \hat{r}_2 := \frac{R_1}{10} \varepsilon^b, \quad \hat{s} := s_0 = \frac{c}{\lambda - 1} \frac{1}{\ln \varepsilon^{-1}} s_1, \\ \hat{\Delta} &:= D^\pm(\sigma), \quad \hat{\Delta}' := \{\hat{I}_2 \in \mathbb{R} : R_1 \varepsilon^b \leq |\hat{I}_2| \leq R_1\}, \\ \widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}} &:= \{\hat{I} \in D^\pm(\sigma)_{\hat{r}_1} \times \hat{\Delta}'_{\hat{r}_2}\} \times \mathbb{T}_{\hat{s}}^2. \end{aligned} \quad (6.13)$$

Notice that the second relation in (6.4) implies that $\hat{r}_1 \ll \hat{r}_2$ for ε small. Define also

$$K := \frac{1}{\varepsilon^{q_0} \ln \varepsilon^{-1}}, \quad (6.14)$$

where q_0 is as in (6.4). Let us, now, estimate α in (5.12). Denote by $\omega^\pm(\hat{I}) := ((h^\pm)'(\hat{I}_1), \varepsilon \sigma \hat{I}_2)$. Then, for any $k \in \mathbb{Z}^2 \setminus \{0\}$ with $|k| \leq K$, by (5.7)–(5.11) and the choice of η , we find

$$|\omega^\pm(\hat{I}) \cdot k| \geq \begin{cases} |(h^\pm)'| - 2\varepsilon R_1 K \geq \kappa_1 R_1 \frac{\sqrt{\varepsilon}}{\ln \varepsilon^{-1}}, & \text{if } k_1 \neq 0, \\ \frac{R_1}{2} \varepsilon^{b+1}, & \text{if } k_1 = 0, \end{cases} \quad (6.15)$$

for a suitable constant⁶⁵ κ_1 and provided $\varepsilon > 0$ is small enough. We can therefore take

$$\alpha := \frac{R_1}{2} \varepsilon^{1+b}. \quad (6.16)$$

⁶⁴ $I_2 = 0$ is a singularity (resonance): we therefore have to stay a bit away from it.

⁶⁵From here on, κ_i denote suitable constants depending, possibly, on $\lambda, a, b, c, q_i, s_1$ and r_1 .

We can now check (5.13). Since, by (6.12),

$$\|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} := \varepsilon^a \|H_1^\pm\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \leq \varepsilon^a \|H_1\|_{r_1, s_1} , \quad (6.17)$$

because of the choices of α , \hat{r}_1 , \hat{s} and K (see (6.16), (6.6), we find (5.6), (6.5) and (6.14)),

$$\frac{\alpha \hat{r}_1}{2^8 K} = \frac{c R_1 r_1 \varepsilon^{b+\lambda+q_0+1/2} \ln \varepsilon^{-1}}{2^9} . \quad (6.18)$$

Thus, in view of the choice of the various parameters made in (6.4), (5.13) is satisfied for $\varepsilon > 0$ small enough. Thus, by Lemma 5.2, there exist a real-analytic symplectic transformation

$$\Phi^\pm : (J, \psi) \in \widehat{W}_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \rightarrow \Phi^\pm(J, \psi) \in \widehat{W}_{\hat{r}_1, \hat{r}_2, \hat{s}} \quad (6.19)$$

such that

$$H^\pm \circ \Phi^\pm(J, \psi) = h^\pm(J_1) + \varepsilon \sigma \frac{J_2^2}{2} + g^\pm(J) + H_*^\pm(J, \psi) \quad (6.20)$$

with (recall (5.15), (6.17), (6.14))

$$\begin{aligned} \|g^\pm - \varepsilon^a (H_1^\pm)_0\|_{\hat{r}_1/2, \hat{r}_2/2} &\leq \frac{1}{4} \varepsilon^a \|H_1\|_{r_1, s_1} , \\ \|H_*^\pm\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} &\leq \|f\|_{\hat{r}_1, \hat{r}_2, \hat{s}} \exp(-K \hat{s}/6) \leq \|H_1\|_{r_1, s_1} \exp\left(\frac{-\kappa_2}{\varepsilon^{q_0} (\ln \varepsilon^{-1})^2}\right) , \\ \|\Phi^\pm(J, \psi) - (J, \psi)\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} &\leq \hat{c} \varepsilon^a \|H_1\|_{r_1, s_1} , \end{aligned} \quad (6.21)$$

for a suitable $\kappa_2 > 0$ (and ε small enough). Thus, if we pick a q_1 so that

$$q < q_1 < q_0 , \quad (6.22)$$

we have that, for all $\varepsilon > 0$ small enough,

$$\|H_*^\pm\|_{\hat{r}_1/2, \hat{r}_2/2, \hat{s}/6} \leq \|H_1\|_{r_1, s_1} \exp\left(-\frac{1}{\varepsilon^{q_1}}\right) . \quad (6.23)$$

Third step. In order to apply the KAM theorem (Lemma 5.3) we set:

$$\begin{aligned} h(J) &= h^\pm(J_1) + \varepsilon \sigma \frac{J_2^2}{2} + g^\pm(J) := h_*^\pm(J) , \quad f(J, \psi) = H_*^\pm(J, \psi) , \\ r &= \kappa_3 r_1 \varepsilon^{\lambda-\frac{1}{2}} , \quad s = \kappa_3 s_1 \frac{1}{\ln \varepsilon^{-1}} , \quad \bar{s} = \frac{s}{2} , \\ D &= D^\pm(\sigma) \times \widehat{D}' , \quad W_{r,s} = D_r \times \mathbb{T}_s^2 , \end{aligned} \quad (6.24)$$

where κ_3 is a suitable constant such that⁶⁶

$$r \leq \frac{\hat{r}_1}{4} , \quad s \leq \frac{\hat{s}}{6} .$$

⁶⁶Recall (6.19) and that $\hat{r}_1 < \hat{r}_2$. The factor $1/4$ is included in order to bound derivatives of g^\pm (and hence of h'') via Cauchy estimates. We recall the statement concerning Cauchy estimates in our context: if $g(J)$ is a function analytic on $D_r \times D'_{r'}$, then for any integers p_1, p_2 and for any $0 < c < 1$

$$\left\| \frac{\partial^{p_1+p_2} g}{\partial J_1^{p_1} \partial J_2^{p_2}} \right\|_{cr, cr'} \leq \text{const.} (p_1! p_2!) \frac{\|g\|_{r, r'}}{r^{p_1} r'^{p_2}} \frac{1}{(1-c)^{p_1+p_2}} .$$

Obviously the norm relative to the domain $W_{r,s}$ will again be denoted $\|\cdot\|_{r,s}$ but beware that the sup-norms in the action variables are taken on different domains according to whether $\sigma = 1$ or $\sigma = -1$ (recall (6.8) and (6.9): in the case $\sigma = -1$ the set \mathcal{R}_* has to be discarded). The estimates on $\|(h_*^\pm)''\|$ and on⁶⁷ $\|U^{-1}\|$ require some computations, which we collect in the following lemma. Recall that from (6.2) and (4.7) there follows that b and λ satisfy

$$b < \frac{1}{4}, \quad \lambda < \frac{a+1}{2}. \quad (6.25)$$

Lemma 6.1 *There exists⁶⁸ $C_0 > 0$ such that, for all $\varepsilon > 0$ small enough,*

$$\|(h_*^\pm)''\|_r \leq \frac{C_0}{\varepsilon^{\lambda-1}(\ln \varepsilon^{-1})^3}, \quad \|U^{-1}\|_r \leq \frac{C_0}{\varepsilon^\lambda \ln \varepsilon^{-1}}, \quad (6.26)$$

where $\|\cdot\|_r$ denotes the sup-norm on D_r defined in (6.24), (6.8), (6.9), (6.13).

Proof First, we need estimates on the derivatives of g^\pm . From (6.21) there follows $\|g^\pm\|_{\hat{r}_1/2, \hat{r}_2/2} \leq \frac{5}{4}\varepsilon^a \|H_1\|_{r_1, s_1}$; whence, by Cauchy estimates⁶⁹,

$$\begin{aligned} \left\| \frac{\partial g^\pm}{\partial J_1} \right\|_r &\leq \kappa_5 \frac{\|H_1\|_{r_1, s_1}}{r_1} \varepsilon^{a-\lambda+\frac{1}{2}}, \quad \left\| \frac{\partial g^\pm}{\partial J_2} \right\|_r \leq \kappa_5 \frac{\|H_1\|_{r_1, s_1}}{r_1} \varepsilon^{a-b}, \\ \left\| \frac{\partial^2 g^\pm}{\partial J_1^2} \right\|_r &\leq \kappa_5 \frac{\|H_1\|_{r_1, s_1}}{r_1^2} \varepsilon^{a-2\lambda+1}, \quad \left\| \frac{\partial^2 g^\pm}{\partial J_2^2} \right\|_r \leq \kappa_5 \frac{\|H_1\|_{r_1, s_1}}{r_1^2} \varepsilon^{a-2b}, \\ \left\| \frac{\partial^2 g^\pm}{\partial J_1 \partial J_2} \right\|_r &\leq \kappa_5 \frac{\|H_1\|_{r_1, s_1}}{r_1^2} \varepsilon^{a-\lambda-b+\frac{1}{2}}, \end{aligned} \quad (6.27)$$

with a suitable constant $\kappa_5 > 0$. By (6.4) and (6.25), one has

$$\begin{aligned} a - \lambda + \frac{1}{2} &> 1, \quad a - b > \frac{3}{2}, \quad a - 2\lambda + 1 > 0, \\ a - 2b &> \frac{5}{4}, \quad a - \lambda - b + \frac{1}{2} > 1. \end{aligned} \quad (6.28)$$

The symmetric matrix U has the form

$$U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{12} & u_{22} & u_{23} \\ u_{13} & u_{23} & 0 \end{pmatrix} \quad (6.29)$$

⁶⁷Recall the definition of the matrix U in Lemma 5.3.

⁶⁸From here on, C_i denote suitable constants depending, possibly, on $\lambda, n, a, b, q_i, E_0, R_1, r_1$ and $\|H_1\|_{r_1, s_1}$.

⁶⁹It is exactly in order to get the estimates (6.27) that we kept track of the different complex extension sizes in the variables J_1 and J_2 .

where (recall (6.24), (6.20) and (5.5))

$$\begin{aligned} u_{11} &= \frac{\partial^2 h_*^\pm}{\partial J_1^2} = (h^\pm)'' + \frac{\partial^2 g^\pm}{\partial J_1^2}, & u_{12} &= \frac{\partial^2 h_*^\pm}{\partial J_1 \partial J_2} = \frac{\partial^2 g^\pm}{\partial J_1 \partial J_2}, \\ u_{13} &= \frac{\partial h_*^\pm}{\partial J_1} = \frac{\partial h^\pm}{\partial J_1} + \frac{\partial g^\pm}{\partial J_1}, & u_{22} &= \frac{\partial^2 h_*^\pm}{\partial J_2^2} = \varepsilon \sigma + \frac{\partial^2 g^\pm}{\partial J_2^2}, \\ u_{23} &= \frac{\partial h_*^\pm}{\partial J_2} = \varepsilon \sigma J_2 + \frac{\partial g^\pm}{\partial J_2}. \end{aligned} \quad (6.30)$$

Since (recall the estimates in Lemma 5.1)

$$C_1 \varepsilon^{\lambda-1} (\ln \varepsilon^{-1})^3 \leq \frac{(\pi_1^\pm)^3}{\pi_2^\pm} \leq C_2, \quad C_1 \leq \pi_1^\pm \leq C_2 \frac{\ln \varepsilon^{-1}}{\sqrt{\varepsilon}}, \quad \frac{C_1}{\sqrt{\varepsilon}} \leq \pi_2^\pm \leq \frac{C_2}{\varepsilon^{\lambda+\frac{1}{2}}}, \quad (6.31)$$

for suitable constants $C_i > 0$, by (6.27) and (6.28), we see that there exists a $\bar{q}_1 > 0$ such that, for all $J_2 \in \widehat{D}'_r$, the following asymptotics hold⁷⁰

$$\begin{aligned} u_{11} &= \pm \beta^\pm \frac{\pi_2^\pm}{(\pi_1^\pm)^3} \left(1 + O(\varepsilon^{\bar{q}_1})\right), & u_{12} &= O(\varepsilon^{\frac{1}{2}+\bar{q}_1}), & u_{22} &= \varepsilon \sigma \left(1 + O(\varepsilon^{\bar{q}_1})\right), \\ u_{13} &= \frac{\alpha^\pm}{\pi_1^\pm} \left(1 + O(\varepsilon^{\bar{q}_1})\right), & u_{23} &= \varepsilon \sigma J_2 \left(1 + O(\varepsilon^{\bar{q}_1})\right). \end{aligned} \quad (6.32)$$

From these relations there follows immediately that

$$\|(h_*)''\|_r \leq \frac{C_3}{\varepsilon^{\lambda-1} (\ln \varepsilon^{-1})^3}. \quad (6.33)$$

Let us, now, write the matrix U^{-1} as follows

$$U^{-1} = \frac{1}{\delta} \begin{pmatrix} u_1 & u_2 & u_3 + u_4 \\ u_2 & 1 & u_5 + u_6 \\ u_3 + u_4 & u_5 + u_6 & u_7 + u_8 \end{pmatrix} \quad (6.34)$$

where

$$\begin{aligned} \delta &:= \frac{u_{11}}{u_{13}^2} u_{23}^2 + u_{22} - 2 \frac{u_{12} u_{23}}{u_{13}}, & u_1 &:= \left(\frac{u_{23}}{u_{13}}\right)^2, \\ u_2 &:= -\frac{u_{23}}{u_{13}}, & u_3 &:= \frac{u_{22}}{u_{13}}, & u_4 &:= -\frac{u_{12} u_{23}}{u_{13}^2}, & u_5 &:= \frac{u_{11} u_{23}}{u_{13}^2}, \\ u_6 &:= -\frac{u_{12}}{u_{13}}, & u_7 &:= \left(\frac{u_{12}}{u_{13}}\right)^2, & u_8 &:= -\frac{u_{11} u_{22}}{u_{13}^2}. \end{aligned} \quad (6.35)$$

Observe that from the above asymptotics (6.32) it follows

$$\frac{u_{11}}{u_{13}^2} = \pm \frac{\beta^\pm}{(\alpha^\pm)^2} \pi_2^\pm \left(1 + O(\varepsilon^{\bar{q}_2})\right), \quad (6.36)$$

⁷⁰Obviusly, $x = O(\varepsilon^c)$ means that there exists a positive constant d such that, for all ε small enough, $|x| \leq d\varepsilon^c$.

for some $\bar{q}_2 > 0$; we also recall that

$$\pi_2^\pm = \frac{1}{|E_\pm| \sqrt{\varepsilon + |E_\pm|}} , \quad E_\pm := \operatorname{Re}(h^\pm(J_1)) , \quad (6.37)$$

where $-2\varepsilon + \varepsilon^\lambda \leq E_- \leq -\varepsilon^\lambda$, $\varepsilon^\lambda \leq E_+ \leq E_0$. Notice that, from (6.31) and (6.32), it follows also that

$$\sup_{i, J \in D_r} |u_i| \leq \frac{C_3}{\varepsilon^{\lambda-1} \ln \varepsilon^{-1}} . \quad (6.38)$$

Thus, it remains to estimate $1/|\delta|$. From (6.35), (6.32) and (6.36), one sees that there exist a complex number $z := z_1 + iz_2$ with $z_1 > 0$ and $|z_2| < z_1/10$ such that⁷¹

$$\delta = \varepsilon \left(\pm z \pi_2^\pm \varepsilon (\operatorname{Re} J_2)^2 + \sigma + O(\varepsilon^{\bar{q}_3}) \right) , \quad \operatorname{Re} \delta = \varepsilon \left(\pm z_1 \pi_2^\pm \varepsilon (\operatorname{Re} J_2)^2 + \sigma + O(\varepsilon^{\bar{q}_3}) \right) \quad (6.39)$$

for a suitable $\bar{q}_3 > 0$. Let us consider the two different signs separately. In the “plus” case, we have to distinguish whether $\sigma = 1$ or $\sigma = -1$. When $\sigma = 1$, since $z_1 \pi_2^+ \varepsilon (\operatorname{Re} J_2)^2 > 0$, one has

$$|\delta| \geq |\operatorname{Re} \delta| = \varepsilon \left(z_1 \pi_2^+ \varepsilon (\operatorname{Re} J_2)^2 + 1 + O(\varepsilon^{\bar{q}_3}) \right) \geq \frac{\varepsilon}{2} \quad (6.40)$$

for $\varepsilon > 0$ small enough. Let now $\sigma = -1$ and notice that⁷²

$$\begin{aligned} \pi_2^+(E_+) &\leq \pi_2^+(\varepsilon^{\frac{2}{3}}/c) \leq \frac{c}{\varepsilon} , & \forall E_+ &\geq \frac{\varepsilon^{\frac{2}{3}}}{c} , \\ \pi_2^+(E_+) &\geq \pi_2^+(c\varepsilon^{\frac{2}{3}+\frac{4}{3}b}) \geq \frac{1}{2c\varepsilon^{1+2b}} , & \forall E_+ &\leq c\varepsilon^{\frac{2}{3}+\frac{4}{3}b} . \end{aligned} \quad (6.41)$$

Choose

$$c := \frac{1}{16} \min \left\{ z_1 R_1^2 , \frac{1}{z_1 R_1^2} \right\} . \quad (6.42)$$

Thus, in the region $E_+ \geq \varepsilon^{2/3}/c$, one has

$$|\delta| \geq |\operatorname{Re} \delta| \geq \varepsilon \left(1 - 4z_1 c R_1^2 + O(\varepsilon^{\bar{q}_3}) \right) \geq \frac{\varepsilon}{2} ;$$

in the region $E_+ \leq c\varepsilon^{\frac{2}{3}+\frac{4}{3}b}$, one has

$$|\delta| \geq |\operatorname{Re} \delta| \geq \left(\frac{z_1 R_1^2}{8c} - 1 + O(\varepsilon^{\bar{q}_3}) \right) \geq \frac{\varepsilon}{2} .$$

Let us turn now to the “minus” sign case and notice that $\varepsilon^\lambda \leq |E_-| \leq 2\varepsilon$ and $\pi_2^- \geq \kappa_6/\varepsilon^{3/2}$ with a suitable $\kappa_6 > 0$. Hence (recalling (6.32) and the assumption $b < 1/4$)

$$|\delta| \geq C_4 \pi_2^- \varepsilon^{2(1+b)} - C_5 \varepsilon \geq C_4 \kappa_6 \varepsilon^{\frac{1}{2}+2b} - C_5 \varepsilon \geq C_6 \varepsilon^{\frac{1}{2}+2b} , \quad (6.43)$$

⁷¹Use also that, for $J_2 \in \widehat{D}'_r$, $|\operatorname{Im} J_2|/|\operatorname{Re} J_2| \leq \operatorname{const}$. $\varepsilon^{\lambda-\frac{1}{2}-b} < \sqrt{\varepsilon}$ by (6.4).

⁷²By (6.37) π_2^+ is a decreasing function of E_+ . Recall also (6.9), that $c < 1$ and that ε is small.

where C_4 , C_5 and C_6 are suitable positive constants. Thus, since $\frac{1}{2} + 2b < 1$, we see that (in all cases)

$$|\delta^{-1}| \leq \frac{C_7}{\varepsilon}, \quad (6.44)$$

with a suitable $C_7 > 0$. This bound together with (6.38) leads to the estimates on $\|U^{-1}\|$ given in (6.26), completing the proof of the lemma. ■

We proceed to estimating the parameters appearing in the statement of Lemma 5.3. From (6.30), (6.32) and (5.10) there follows that

$$\begin{aligned} \left| \frac{\partial h_*^\pm}{\partial J_1} \right| &:= |u_{13}| \geq C_8 \frac{\sqrt{\varepsilon}}{\ln \varepsilon^{-1}}, \\ \left| \frac{\partial h_*^\pm}{\partial J_2} \right| &:= |u_{23}| \geq C_8 \varepsilon^{1+b}, \end{aligned} \quad (6.45)$$

for a suitable $C_8 > 0$ so that $\min_{i,J} \left| \frac{\partial h_*^\pm}{\partial J_i} \right| \geq C_8 \varepsilon^{1+b}$. We next choose $\gamma \ll C_8 \varepsilon^{1+b}$. Since the norm of H_*^\pm is exponentially small with ε , we can choose also γ exponentially small with ε : we let, in fact, for a suitable $\gamma_0 > 0$,

$$\gamma := \gamma_0 \exp \left(-\frac{1}{\varepsilon^{q_2}} \right), \quad \text{with } q < q_2 < q_1. \quad (6.46)$$

Therefore, in view of (6.23), (6.26) and (6.46), we can take⁷³

$$A := \frac{C_0}{\varepsilon^{\lambda-1} (\ln \varepsilon^{-1})^3}, \quad F := \exp \left(-\frac{1}{2\varepsilon^{q_1}} \right), \quad G := \frac{C_9}{\varepsilon^{2\lambda-1} (\ln \varepsilon^{-1})^4}, \quad (6.47)$$

for a suitable $C_9 > 0$. Next, we show that C in (5.21) is one in our case. By (6.24), (6.26), (6.46) and (6.47), we see that (for a suitable $C_{10} > 0$)

$$\frac{\gamma(s - \bar{s})^{c_1}}{c_2 A r |\ln F|^{c_3}} = C_{10} \frac{\varepsilon^{c_3 q_1} \exp \left(-\frac{1}{\varepsilon^{q_2}} \right)}{(\ln \varepsilon^{-1})^{c_1-2}};$$

which implies that $C = 1$ for ε small enough. Therefore, recalling the definition (5.21) of \hat{F} , we can take, for a suitable $C_{11} > 0$ (see (6.47) and (6.24)) and for $\varepsilon > 0$ small enough,

$$\hat{F} \leq C_{11} \exp \left(-\frac{1}{\varepsilon^{q_2}} \right), \quad (6.48)$$

which obviously will be smaller than one for any $\varepsilon > 0$ small enough. Thus, under conditions (6.4), (6.25) and (6.46), Lemma 5.3 can be applied to the Hamiltonian (6.20) showing the existence of KAM tori in each energy level of $W_{r,s}$ apart from a small set of measure bounded by⁷⁴ $O(\hat{\gamma}) \leq O(\exp(-1/\varepsilon^q))$. Thus (recall Remark 5.1), the motions

⁷³Recall the definitions of F and G given in (5.20).

⁷⁴Provided τ is chosen strictly larger than one; the constant $\hat{\gamma}$ is defined in (iii) of Lemma 5.3 and, in view of Lemma 6.1, is related to γ by a power of ε .

starting in $W_{r,s}$ have action variables $O(\exp(-1/q))$ -close to their initial values for all times. In the original coordinates (I, φ) , the measure of the complementary of the KAM tori is again bounded by $O(\exp(-1/q))$; the KAM tori fill up the region $\mathcal{M}^{(\sigma)}$ with the exception of a set of measure $O(\exp(-1/q))$. In view of (6.21), the displacement of the KAM tori from the corresponding unperturbed ones is $O(\varepsilon^a)$ while the oscillation of the graph of the tori may be bounded by $O(\sqrt{\varepsilon})$. Repeating the argument in Remark 5.1 we find that, denoting $(I(t), \varphi(t))$ the ϕ^t evolution of (I_0, φ_0) with⁷⁵ $(I_0, \varphi_0) \in \mathcal{M}^{(\sigma)}$,

$$|I(t) - I_0| < C_{12}\sqrt{\varepsilon}, \quad \forall t, \quad (6.49)$$

(provided $\varepsilon > 0$ is small enough).

This concludes the proof of Theorem 4.4. \blacksquare

Proof of Theorem 4.3

We proceed to show that Theorem 4.4 and energy conservation imply (7.6) in \mathcal{M}_R when $\sigma = 1$ and in $\mathcal{M}_R \setminus \mathcal{N}_*$ when $\sigma = -1$ (recall the definition of \mathcal{N}_* in (4.10)).

In view of the oscillations of the KAM tori in the region $\mathcal{M}^{(\sigma)}$ we shall consider slightly smaller sets $\widetilde{\mathcal{M}}^{(\sigma)} \subset \mathcal{M}^{(\sigma)}$. To define such sets we let $\widetilde{\mathcal{N}}_* := \mathcal{N}_*$ and:

$$\begin{aligned} \widehat{\mathcal{N}}_* &:= \left\{ (I, \varphi) : 2c \varepsilon^{\frac{2}{3} + \frac{4}{3}b} < H_{\mathcal{P}} < \frac{\varepsilon^{\frac{2}{3}}}{2c} \right\}, \\ \widetilde{\mathcal{N}}^{(1)} &:= \left\{ (I, \varphi) : |H_{\mathcal{P}}| < 2\varepsilon^{1+2b} \text{ or } H_{\mathcal{P}} < -2\varepsilon + 2\varepsilon^{1+2b} \right\}, \\ \widetilde{\mathcal{N}}^{(2)} &:= \left\{ (I, \varphi) : |I_2| < 2R\varepsilon^b \right\}, \\ \widetilde{\mathcal{M}}^{(1)} &:= \mathcal{M}_R \setminus (\widetilde{\mathcal{N}}^{(1)} \cup \widetilde{\mathcal{N}}^{(2)}), \quad \widetilde{\mathcal{M}}^{(-1)} := \widetilde{\mathcal{M}}^{(1)} \setminus \widetilde{\mathcal{N}}_*. \end{aligned} \quad (6.50)$$

Remark 6.1 Because of Theorem 4.4 (and, hence, because of the confinement due to the presence of two-dimensional KAM tori in three-dimensional energy levels), the smaller sets $\widetilde{\mathcal{M}}^{(\sigma)}$ have the property that $\bigcup_{t \in \mathbb{R}} \phi_{\sigma}^t(\widetilde{\mathcal{M}}^{(\sigma)}) \subset \mathcal{M}^{(\sigma)}$ (where ϕ_{σ}^t denotes the $H^{(\sigma)}$ -flow). In particular, in the case $\sigma = -1$, a trajectory cannot cross the region $\widehat{\mathcal{N}}_*$ (a fact that could also be checked directly by energy conservation since $\frac{2}{3} + \frac{4}{3}b < 1$).

Denote by $z(t) := (I(t), \varphi(t))$ the motion with initial data $z_0 := (I_0, \varphi_0)$ governed by $H^{(\sigma)}$ in \mathcal{M}_R (if $\sigma = 1$) or $\mathcal{M}_R \setminus \mathcal{N}_*$ (if $\sigma = -1$). Let us consider the different cases which may occur.

(i) If $z_0 \in \widetilde{\mathcal{M}}^{(\sigma)}$ then, as remarked above, $z(t)$ does not leave $\mathcal{M}^{(\sigma)}$ where (6.49) (and hence (7.6)) holds.

⁷⁵Recall that $R < R_1 < R_0$ and that ε will be small compared also to $(R_0 - R)$.

(ii) If $z(t) \in \widetilde{\mathcal{N}}^{(1)}$ for $|t| < T$ for some $T > 0$, then, by energy conservation, (7.6) holds⁷⁶ for $|t| < T$.

(iii) If $z(t) \in \widetilde{\mathcal{N}}^{(2)}$ for $|t| < T$ then (7.6) (trivially) holds for $|t| < T$.

(iv) By (ii) and (iii) (7.6) holds until $z(t) \in \widetilde{\mathcal{N}}^{(1)} \cup \widetilde{\mathcal{N}}^{(2)}$. But if $z(t)$ leaves $\widetilde{\mathcal{N}}^{(1)} \cup \widetilde{\mathcal{N}}^{(2)}$ and enters the region $\widetilde{\mathcal{M}}^{(\sigma)}$, then, by (i), (7.6) holds again. ■

⁷⁶In fact, calling $E_p(t) = H_p(I_1(t), \varphi_1(t))$, if $z(t) \in \widetilde{\mathcal{N}}^{(1)}$ for $|t| < T$, then $|E_p(t) - E_p(0)| \leq O(\varepsilon^\lambda)$ for all $|t| \leq T$ (recall that $\lambda = 1 + 2b$ and that $a > \lambda$). Thus, by energy conservation, there follows that

$$\frac{I_2(t)^2 - I_2(0)^2}{2} + \frac{E_p(t) - E_p(0)}{\varepsilon} = O(\varepsilon^{a-1}) ,$$

for all $|t| < T$. Therefore, $I_2(t)^2 - I_2(0)^2 = O(\varepsilon^{\lambda-1})$ and (7.6) follows.

Part III

Exponential stability: the resonant D'Alembert model of Celestial Mechanics

7 The exponential stability Theorem

In this part we shall consider the resonant “day/year” (or “spin/orbit”) planetary *D’Alembert model*, which, together with the three-body problem, may be considered among the oldest and most intriguing problem in Celestial Mechanics (compare, e.g., [73]); in particular we shall address the problem of the (exponentially) long time stability for such a model and will show how to obtain results *à la* Nekhoroshev for it.

In particular we shall prove that *for small oblateness and eccentricity and for any motion starting near an exact spin/orbit resonance, the Hamiltonian evolution of the action variables stay close to their initial values for exponentially long times.*

Before giving a more precise formulation of this stability result, let us briefly mention its connection with Arnold diffusion.

Remark 7.1 *In [49], it was claimed that the planetary D’Alembert model, near a resonance (2:1), has an instability region where the variable J_2 undergoes a variation of order one (i.e., independent of the perturbative parameters) in finite time, provided ε and $\mu = \varepsilon^c$ (for a suitable $c > 1$) are positive and small enough. The proof of this claim proposed in [49] contained an algebraic error (see the Erratum in [49]) and, even though such error has been corrected ([67]) and several technical progresses, in such direction, have been obtained (see, e.g., [68], [69], [94]), a complete proof of the above claim is still missing. As well known, proving Arnold diffusion for analytic problems is an extremely difficult problem; for the D’Alembert model an extra (technical) problem (related to the degeneracies present in the model) might come from the particular slowness of the expected instability: compare comment 7.1 below.*

Let us proceed to formulate, in a more precise way, our main result. We briefly recall Section 3. In that section we stated that the planetary D’Alembert model near an exact ($p:q$) resonance may be described (in suitable physical units) by the real-analytic Hamiltonian (3.14) namely

$$H_{\varepsilon,\mu} := \frac{I_1^2}{2} + \omega(pI_1 - qI_2 + qI_3) + \varepsilon F_0(I_1, I_2, \varphi_1, \varphi_2) + \varepsilon\mu F_1(I_1, I_2, \varphi_1, \varphi_2, \varphi_3; \mu) , \quad (7.1)$$

where: $(I, \varphi) \in A \times \mathbb{T}^3$ are standard symplectic coordinates and the domain $A \subset \mathbb{R}^3$ is given by⁷⁷

$$A := \left\{ |I_1| < r\varepsilon^\ell , \quad |I_2 - \bar{J}_2| < r , \quad I_3 \in \mathbb{R} \right\} , \quad (7.2)$$

with $0 < \ell < 1/2$, $r > 0$, \bar{J}_2 is a fixed “reference datum” (avoiding certain singularities); ε and μ are two small parameters (measuring, respectively, the oblateness of the planet and the eccentricity of the Keplerian ellipse); p and q are two positive co-prime integers, which identify the spin-orbit resonance (the planet, in the unperturbed regime, revolves q times

⁷⁷ See (3.15).

around the star and p times around its spin axis); ωq is the frequency of the Keplerian motion; the action I_1 measures the displacement from the exact resonance: in these units, $I_1 = 0$ corresponds exactly to a $p : q$ spin-orbit resonance. In fact, $\bar{J}_1 + I_1$, with $\bar{J}_1 := p\omega$, and I_2 are (in suitable physical units), respectively, the absolute value and the projection onto the polar axis of the planet of the angular momentum of the planet, while I_3 is an artificially introduced variable canonically conjugated to time. The functions F_i are real-analytic functions in all their arguments. While the explicit form of F_1 is not important in the sequel, and, in fact, *our result holds for any function F_1 real-analytic and bounded on A* , the explicit form of F_0 plays a major rôle in the following analysis. We recall the form of F_0 given in (3.6)

$$F_0(I_1, I_2, \varphi_1, \varphi_2) = \sum_{\substack{j \in \mathbb{Z} \\ |j| \leq 2}} c_j \cos(j\varphi_1) + d_j \cos(j\varphi_1 + 2\varphi_2) , \quad (7.3)$$

where c_j and d_j are suitable functions listed in (3.9). As in Section 3 we introduce the real parameter L which corresponds to the projection of the angular momentum of the planet onto the polar axis of the planet and, since the planet is rotational, it turns out to be a constant of the motion. As assumed in (3.16) the parameters \bar{J}_i , L and the constant r are assumed to satisfy (3.16) i.e.

$$L + 3r\varepsilon^\ell < \bar{J}_1 , \quad |\bar{J}_2| + 3r(\varepsilon^\ell + 1) < \bar{J}_1 , \quad (7.4)$$

so that c_j, d_j are well defined on the domain A .

With the above positions, for the motions governed by the $(p:q)$ -resonant D'Alembert Hamiltonian $H_{\varepsilon, \mu}$, there holds the following

Theorem 7.1 *Let $c > 0$, $0 < \ell < 1/2$ and $0 < \gamma_0 < \min\{c, \ell\}$. Assume that (7.4) holds and that*

$$\bar{J}_1 \neq \sqrt{3}L \quad (7.5)$$

(which is equivalent to⁷⁸ $\nu_1(0) \neq \frac{2}{3}$). Then, there exist $\varepsilon_0, \gamma_i > 0$ such that, if $0 \leq \varepsilon \leq \varepsilon_0$ and $0 \leq \mu \leq \varepsilon^c$, then

$$|I(t) - I(0)| < \gamma_3 r \varepsilon^{\gamma_1} , \quad \forall \quad |t| < T(\varepsilon) := \frac{\gamma_5}{\omega \varepsilon^{\gamma_4}} \exp\left(\frac{\gamma_2}{\varepsilon^{\gamma_0}}\right) , \quad (7.6)$$

where $(I(t), \varphi(t))$ denotes the $H_{\varepsilon, \mu}$ -evolution of an initial datum $(I(0), \varphi(0)) \in A \times \mathbb{T}^3$.

Remark 7.2 *From a physical point of view, one is interested in knowing the variation of the angle α_1 between the normal to the ecliptic plane and the polar axis of the planet. It results that $\alpha_1 \in [\alpha_2 - \alpha_3, \alpha_2 + \alpha_3]$ where α_2 is the angle between the normal to the ecliptic plane and the angular momentum of the planet and α_3 is the angle between the angular momentum of the planet and its polar axis. Hence it is clear that if we prove stability (for a certain amount of time) for the action variables I_1 and I_2 (I_3 has no physical meaning)*

⁷⁸ See the definition of ν_1 given in (3.7).

we have proved that the angles α_2 and α_3 are nearly constant (in the above amount of time). Now in the Solar System it is a common fact that planets have angular momentum nearly parallel to their polar axis (so it is, for example, for the Earth), namely α_3 is very small. Hence as a corollary of the action-stability we should prove the stability (for a certain amount of time) of the inclination of the planets with respect to the normal to the ecliptic plane.

Let us make a few comments on the above Theorem.

7.1 Standard Nekhoroshev estimates for a system with d degrees of freedom yield a Nekhoroshev exponent (i.e., the exponent of ε in the exponential part of $T(\varepsilon)$), which, in our case is $\gamma_0) 1/(2d)$; compare, e.g., [93]. In our case, taking ℓ close to $1/2$ and $c \geq \ell$, we see that $\gamma_0 \sim 1/2$, getting a *better* Nekhoroshev exponent with respect to the general non-degenerate case (“better” means that the stability time is longer, for small ε). This fact (as mentioned above) makes, *a priori*, Arnold diffusion particularly slow in the present example. This phenomenon is due to the appearance of *three well separated time scales*, as discussed in the next item.

7.2 After a natural, symplectic linear change of variables $(I, \varphi) \rightarrow (I', \varphi')$, ($I'_1 := I_1$, $I'_3 := pI_1 - qI_2 + qI_3$), the unperturbed part, $H_{0,\mu}$, becomes simply $I_1'^2/2 + \omega I_3'$, showing, in a more clear way, the proper-degeneracy of the system ($H_{0,\mu}$, in these variables, does not depend upon $I_2' = I_2$, which is the physically interesting variable). The appearance of *three (well separated) time scales* for the evolution of the angles φ' 's is also more evident in these variables: in fact (for ε small and $\mu < 1$)

$$\dot{\varphi}'_1 = O(\varepsilon^\ell) \gg O(\sqrt{\varepsilon}), \quad \dot{\varphi}'_2 = O(\varepsilon), \quad \dot{\varphi}'_3 = \omega = O(1). \quad (7.7)$$

The “non-degeneracy” assumption (7.5) is made in order to really have $\dot{\varphi}'_2 = O(\varepsilon)$, i.e., in order to have, $\frac{1}{\varepsilon} \partial_{I_2'} H_{\varepsilon,\mu}|_{I_1'=0, I_2'=J_2} \neq 0$.

7.3 An immediate consequence of the above different time scales is the following. Having only one frequency (namely ω) of order one suggests to perform *averaging* over φ'_3 ; this can be done at any order of ε and, in fact, using “*normal form*” theory, we shall do it at an exponential order. This implies immediately that the quantity $|I_3'(t) - I_3'(0)|$ is *exponentially small* with (a suitable power of) $1/\varepsilon$ for times which are *exponentially long* with (a suitable power of) $1/\varepsilon$ (at least, as long as the other action variables remain in their domain of definition). This shows that, up to such exponentially long times, the resonant D'Alembert Hamiltonian behaves effectively as a two-degree-of-freedom system (see, also, [34]). Furthermore, besides the just mentioned “super stability” of I_3' (which, after all, has a limited physical interest), we see immediately, by (approximate) energy conservation⁷⁹, that also I_1' is stable (again: provided I_2' stays in its domain of definition); in this case, in fact, one checks immediately that $|I_1(t) - I_1(0)| \leq \text{const. } \sqrt{\varepsilon}$. This comment shows that – as mentioned above – the only non-trivial stability is related to $I_2' = J_2$, which (not by chance) is the most relevant physical quantity.

⁷⁹ “Approximate” because, for the obtained two-degree-of-freedom system, energy conservation holds only up to exponentially long times.

7.4 It is remarkable (particularly in view of the special occurrence, in our Solar System, of the resonances (1:1) and (2:1)) that *the analysis of the resonant D'Alembert Hamiltonian is different according to whether the resonance $(p:q)$ is considered with $(p, q) = (1, 1), (2, 1)$ or with $(p, q) \neq (1, 1), (2, 1)$* . In fact, the case $(p, q) = (1, 1), (2, 1)$ will turn out to be significantly more difficult, at a technical level, than the other cases, as briefly explained in the next item **7.5**.

We have formerly discussed the apparency of this difference between the case $(p, q) = (1, 1), (2, 1)$ and the cases $(p, q) \neq (1, 1), (2, 1)$ in subsection 3.2. In particular we observe that, after the above exponential averaging in φ'_3 (calling \hat{I} and $\hat{\varphi}$ the new averaged symplectic variables, which are close to the variables I' and φ'), the D'Alembert Hamiltonian takes the form

$$\frac{\hat{I}_1^2}{2} + \omega \hat{I}_3 + \varepsilon \overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1) + \varepsilon^a H_1^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon) + O\left(\exp(1/\varepsilon^{\text{const}})\right), \quad (7.8)$$

with $a := 1 + \min\{c, \ell\}$, $H_1^{(1)}$ real-analytic and bounded by a constant and the “effective Hamiltonian” \overline{H}_{01} given by

$$\overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1) := \frac{1}{2\pi} \int_0^{2\pi} F_0(0, \hat{I}_2, \hat{\varphi}_1 + p\hat{\varphi}_3, \hat{\varphi}_2 - q\hat{\varphi}_3) d\hat{\varphi}_3. \quad (7.9)$$

As we have discussed in section 3, a straightforward computation of this average using the explicit expression for F_0 given in (7.3), yields

$$\overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1) = \begin{cases} c_0(0, \hat{I}_2), & \text{if } (p, q) \neq (1, 1), (2, 1) \\ c_0(0, \hat{I}_2) + d_{j_p}(0, \hat{I}_2) \cos(j_p \hat{\varphi}_1), & \text{if } (p, q) = (1, 1) \text{ or } (2, 1) \end{cases} \quad (7.10)$$

with $j_1 := 2$ and $j_2 := 1$: thus, in the case $(p, q) \neq (1, 1), (2, 1)$ the effective Hamiltonian \overline{H}_{01} is independent of $\hat{\varphi}$, while in the cases $(p, q) = (1, 1)$ or $(2, 1)$, it depends explicitly upon $\hat{\varphi}_1$.

7.5 In the case \overline{H}_{01} in (7.8) is independent of $\hat{\varphi}$, one can exploit the separation of scales between $\hat{\varphi}_1$ and $\hat{\varphi}_2$ (which remain of the same order of $\dot{\varphi}'_1$ and $\dot{\varphi}'_2$): in a suitable ε -dependent \hat{I} -domain (i.e., excluding an ε -dependent small neighborhood of $\{\hat{I}_1 = 0\}$) the variable $\hat{\varphi}_1$ is fast with respect to the variable $\hat{\varphi}_2$ and, therefore, the dependence upon $\hat{\varphi}_1$ (again, by means of “averaging” or “normal form” theory) can be removed up to exponentially small terms. Optimizing the various choices and using energy conservation arguments will allow us to obtain the stability claim in all action variables for the case $(p, q) \neq (1, 1)$ or $(2, 1)$.

To carry out an analogous strategy in the case \overline{H}_{01} in (7.8) does depend upon $\hat{\varphi}_1$, one has to, first, put the two-dimensional integrable system $\frac{\hat{I}_1^2}{2} + \varepsilon \overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ into action-angle variables. Moreover for our purposes, we need *detailed information on the analyticity domains* of this symplectic change of variables, which becomes singular as $\varepsilon \rightarrow 0$. This analysis was carried out in Proposition 2.1.

7.6 Apart from the complex analysis technicalities mentioned in the previous item, our approach consists in a careful joint use of normal form theory and energy conservation arguments fully exploiting the separation of scales described in item **7.2**. We remark that *the integrable part of the resonant D'Alembert Hamiltonian*, i.e., the term $\frac{\hat{I}_2^2}{2} + \omega \hat{I}_3 + \varepsilon \overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ in (7.8), *is a non convex function of \hat{I}* ; compare [33] and [34] where it is shown that the coefficient of \hat{I}_2^2 in $c_0(0, \hat{I}_2)$, in typical regions of phase space, is *negative*. This explains the reason why the arguments that we use, in order to obtain action-confinement, are somewhat more related to the original approach of Nekhoroshev [90] rather than to simpler arguments based on convexity, as introduced in [15] (and used, also, in [93]).

7.7 For the D'Alembert model, a different approach to exponential stability is, in principle, possible. Namely, one may use KAM theory to establish the existence of maximal tori for the two-dimensional system in (7.8) obtained disregarding the exponentially small terms (and the term $\omega \hat{I}_3$) concluding the entrapment of the \hat{I}_2 variable up to exponentially long times (for which the dynamics of the two-dimensional approximation coincides, essentially, with the full dynamics). Furthermore, using *smooth, iso-energetic* KAM theory, this approach might avoid the delicate complex analysis mentioned in item **7.5**. We decided not to follow this approach since in order to apply smooth KAM one needs to control (in the two-dimensional case) the C^r -norm of the Hamiltonian with $r > 4$, a fact which would involve to have a precise control, as $\varepsilon \rightarrow 0$, of the first five derivatives of the symplectic diffeomorphism casting $\frac{\hat{I}_2^2}{2} + \varepsilon \overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ into action-angle variables: these computations might not be simpler than the computations performed in Proposition 2.1. A second reason is that the two-dimensional Hamiltonian obtained from (7.8) is in general *not a convex function of \hat{I}* and iso-energetic KAM does not hold in the whole four-dimensional phase space. Indeed, instability channels (along which the action variables may drift away by a quantity of order one in exponentially long times – compare [33]) may, in general, appear. *We believe that, in the D'Alembert case, such channels do not appear*: it would be interesting to prove this fact, since the presence of channels might lead to a topological instability for the D'Alembert problem based on a mechanism entirely different from Arnold diffusion.

7.8 Finally, we mention that the estimates provided in the proof of Theorem 7.1 allow to compute explicitly all the constants involved. For example, let us indicate, here, for the case $(p, q) \neq (1, 1), (2, 1)$, the exact expression of the constants γ_i and ε_0 involved in the Theorem. Let $a := 1 + \min\{c, \ell\}$ and choose $1/2 < b \leq (a - \gamma_0)/2$. Then the time of stability can be taken to be⁸⁰

$$T(\varepsilon) = \frac{eSr}{576\overline{M}_2^{(2)}} \frac{1}{\varepsilon^{a-2b}} \exp(C_2\varepsilon^{-\gamma_0}). \quad (7.11)$$

where $S, C_2, \overline{M}_2^{(2)}$ are defined, respectively, in (8.4), (8.18), (8.24) below. Furthermore, for any time $|t| \leq T(\varepsilon)$, we obtain

$$|I_1(t) - I_1(0)| \leq \frac{5}{4}r\varepsilon^b \quad \text{and} \quad |I_1(t)| \leq \frac{5}{4}r\varepsilon^\ell; \quad (7.12)$$

⁸⁰ Here e is the Neper constant $e := \sum_k (k!)^{-1}$.

defining $\gamma_1 := (2b - 1)/2$, $\gamma_6 := \sqrt{C_3} + 2^{-6}$, with C_3 is defined in (8.31) below, we also obtain

$$|I_2(0)| \leq \gamma_6 r \varepsilon^{\gamma_1} \implies |I_2(t) - I_2(0)| \leq (1 + \sqrt{2}) \gamma_6 r \varepsilon^{\gamma_1} \leq \frac{r}{8} \quad (7.13)$$

$$|I_2(0)| > \gamma_6 r \varepsilon^{\gamma_1} \implies |I_2(t) - I_2(0)| \leq \frac{\gamma_6^2 r^2}{|\hat{I}_2(0)|} \varepsilon^{2b-1} \leq \frac{r}{8}. \quad (7.14)$$

An explicit expression for ε_0 may be immediately obtained by looking at the conditions (8.10), (8.11), (8.18), (8.22), (8.32), (8.36) below.

8 Proof of Theorem in the case $(p, q) \neq (1, 1), (2, 1)$

In this section we present the proof of Theorem 7.1 in the case $(p, q) \neq (1, 1)$ or $(2, 1)$. In such a case the function $\bar{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ is simply $c_0(0, \hat{I}_2)$. Bare in mind, however, that subsection 8.1, 8.2, and 8.3 hold for the general case and, therefore, in these subsections we maintain the general notation $\bar{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$.

In what follows, we shall assume that, for any $0 \leq \varepsilon \leq \bar{\varepsilon}$, the Hamiltonian⁸¹ $H_{\varepsilon, \mu}$ belongs to $\mathcal{H}_{\mathbb{R}}(A_R \times \mathbb{T}_s^3)$, where

$$R := (r\varepsilon^\ell, r, \infty), \quad (8.1)$$

$s > 0$ (and $0 < \bar{\varepsilon} < 1$). We shall also denote M_0 and M_1 (ε -independent) upper bounds on, respectively, $\|F_0\|_{R,s}$ and $\|F_1\|_{R,s}$.

8.1 Step 1 (linear change of variables)

Let ϕ_0 be the following linear symplectic map:

$$\phi_0(I', \varphi') := \left((I'_1, I'_2, -\frac{p}{q}I'_1 + I'_2 + \frac{1}{q}I'_3), (\varphi'_1 + p\varphi'_3, \varphi'_2 - q\varphi'_3, q\varphi'_3) \right). \quad (8.2)$$

Then, ϕ_0 casts the Hamiltonian $H_{\varepsilon, \mu}$ into the form

$$\begin{aligned} H^{(0)}(I', \varphi'; \varepsilon, \mu) &:= H_{\varepsilon, \mu} \circ \phi_0(I', \varphi') \\ &:= \frac{I'^2_1}{2} + \omega I'_3 + \varepsilon G_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3) + \varepsilon \mu G_1(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3; \mu), \end{aligned} \quad (8.3)$$

which belongs to $\mathcal{H}_{\mathbb{R}}(A_R \times \mathbb{T}_S)$ with

$$S := c's, \quad c' := \min\{1/(1+p), 1/(1+q)\}, \quad (8.4)$$

and

$$\|G_0\|_{R,S} \leq M_0, \quad \|G_1\|_{R,S} \leq M_1.$$

⁸¹Recall the definitions of $H_{\varepsilon, \mu}$ and A given above.

Moreover:

$$G_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3) := H_{01}(I'_1, I'_2, \varphi'_1) + \tilde{G}_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3)$$

with

$$\int_0^{2\pi} \tilde{G}_0(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3) d\varphi'_3 = 0 ,$$

and

$$H_{01}(I'_1, I'_2, \varphi'_1) := \begin{cases} c_0(I'_1, I'_2) , & \text{if } (p, q) \neq (1, 1), (2, 1) , \\ c_0(I'_1, I'_2) + d_{j_p}(I'_1, I'_2) \cos(j_p \varphi'_1) , & \text{if } (p, q) = (1, 1), (2, 1) , \end{cases}$$

with $j_1 := 2$ and $j_2 := 1$.

Obviously, since ϕ_0 depends upon p and q , also the functions G_i and H_{01} depend upon p and q , but *we shall not indicate such dependence in the notation*.

We remark that, in general, ϕ_0 is *not* a diffeomorphism of $\mathbb{R}^3 \times \mathbb{T}^3$ (since the induced map on \mathbb{T}^3 has determinant equal to q); this fact, however, does not affect the following analysis.

If $a := 1 + \min\{c, \ell\}$, using the fact that $|I'_1| < 2r\varepsilon^\ell$ and $\mu \leq \varepsilon^c$, one see that $H^{(0)}$ has the following form:

$$\frac{I_1'^2}{2} + \omega I'_3 + \varepsilon \bar{H}_{01}(I'_2, \varphi'_1) + \varepsilon \bar{G}_0(I'_2, \varphi'_1, \varphi'_2, \varphi'_3) + \varepsilon^a H_2^{(0)}(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3; \varepsilon)$$

where $\bar{G}_0(I'_2, \varphi) := \tilde{G}_0(0, I'_2, \varphi)$,

$$\int_0^{2\pi} \bar{G}_0(I'_2, \varphi'_1, \varphi'_2, \varphi'_3) d\varphi'_3 = 0 , \quad (8.5)$$

and (recall (3.9))

$$\bar{H}_{01}(I'_2, \varphi'_1) := H_{01}(0, I'_2, \varphi'_1) := \begin{cases} \bar{c}_0(I'_2) , & \text{if } (p, q) \neq (1, 1), (2, 1) , \\ \bar{c}_0(I'_2) + \bar{d}_{j_p}(I'_2) \cos(j_p \varphi'_1) , & \text{if } (p, q) = (1, 1), (2, 1) , \end{cases} \quad (8.6)$$

where

$$\bar{c}_0(I'_2) := c_{00} + c_{02} \frac{I_2'^2}{2}, \quad c_{00} := \frac{1}{4}(2 - \bar{\nu}_1^2), \quad c_{02} := \frac{1}{\bar{J}_1^2} \left(\frac{3}{2} \bar{\nu}_1^2 - 1 \right) \quad (8.7)$$

and

$$\begin{aligned} \bar{d}_1(I'_2) &:= \bar{d}_1(0, I'_2) := -\frac{1}{2} \bar{\kappa}_1 \bar{\nu}_1 \sqrt{1 - \frac{I_2'^2}{\bar{J}_1^2}} \left(1 + \frac{I_2'}{\bar{J}_1} \right) , \\ \bar{d}_2(I'_2) &:= \bar{d}_2(0, I'_2) := -\frac{1}{8} \bar{\nu}_1^2 \left(1 + \frac{I_2'}{\bar{J}_1} \right)^2 , \end{aligned} \quad (8.8)$$

where

$$\bar{\kappa}_1 := \kappa_1(0) := \frac{L}{J_1}, \quad \bar{\nu}_1 := \nu_1(0) := \sqrt{1 - \bar{\kappa}_1^2}. \quad (8.9)$$

The function $H_2^{(0)}(I'_1, I'_2, \varphi'_1, \varphi'_2, \varphi'_3; \varepsilon)$ belongs to $\mathcal{H}_{\mathbb{R}}(A_R \times \mathbb{T}_S^3)$ and

$$\|\varepsilon \bar{H}_{01} + \varepsilon \bar{G}_0 + \varepsilon^a H_2^{(0)}\|_{R,S} = \|\varepsilon G_0 + \varepsilon^{1+c} G_1\|_{R,S} \leq \varepsilon(M_0 + \varepsilon^c M_1).$$

8.2 Step 2 (time averaging)

Here, we shall remove, up to exponentially small terms, the (fast) dependence upon φ'_3 . To do this, we shall apply the Normal Form Lemma with $d := 3$, $(I, \varphi) := (I', \varphi')$, $H := H^{(0)}$, $h := I_1'^2/2 + \omega I_3'$, $f := \varepsilon \bar{H}_{01} + \varepsilon \bar{G}_0 + \varepsilon^a H_2^{(0)} = \varepsilon G_0 + \varepsilon^{1+c} G_1$, $D := A$, $\rho := R$, $\rho_0 := r\varepsilon^\ell$, $\sigma := S$, $\Lambda := \{(k_1, k_2, 0) \text{ s.t. } k_1, k_2 \in \mathbb{Z}\}$, $\alpha := \omega/2$, $K := \omega/(4r\varepsilon^\ell)$. The condition $K\sigma \geq 6$ is implied by

$$\varepsilon \leq (\omega S/24r)^{1/\ell}. \quad (8.10)$$

Condition (1.2) becomes

$$\varepsilon(M_0 + \varepsilon^c M_1) \leq 2^{-9} r^2 \varepsilon^{2\ell}$$

which is verified, for example, if

$$\varepsilon \leq \left(\frac{r^2}{2^9(M_0 + M_1)} \right)^{1/(1-2\ell)}. \quad (8.11)$$

Hence, for ε small enough, we can apply the Normal Form Lemma, finding a real-analytic symplectic transformation

$$\phi_1 : (\hat{I}, \hat{\varphi}) \in A_{R/2} \times \mathbb{T}_{S/6}^3 \mapsto (I', \varphi') \in A_R \times \mathbb{T}_S^3,$$

such that

$$\begin{aligned} H^{(1)}(\hat{I}, \hat{\varphi}; \varepsilon, \mu) &:= H^{(0)} \circ \phi_1(\hat{I}, \hat{\varphi}; \varepsilon) \\ &:= \frac{\hat{I}_1^2}{2} + \omega \hat{I}_3 + \varepsilon \bar{H}_{01}(\hat{I}_2, \hat{\varphi}_1) + \varepsilon^a H_1^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon) + H_*^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}; \varepsilon), \end{aligned} \quad (8.12)$$

and such that the following bounds hold. For any $(\hat{I}, \hat{\varphi}) \in A_{R/2} \times \mathbb{T}_{S/6}^3$,

$$|I' - \hat{I}| \leq \frac{2^6}{\omega S} \varepsilon(M_0 + \varepsilon^c M_1) \leq \left(\frac{r}{8S\omega} \right) r \varepsilon^{2\ell} \leq \frac{1}{2^7} r \varepsilon^\ell, \quad (8.13)$$

and

$$\|H_1^{(1)}\|_{R/2, S/6} \leq M_1^{(1)}, \quad \|H_*^{(1)}\|_{R/2, S/6} \leq M_*^{(1)} := \varepsilon(M_0 + \varepsilon^c M_1) \exp(-C_1 \varepsilon^{-\ell}), \quad (8.14)$$

where

$$M_1^{(1)} := \left(M_1 + \frac{8r}{\omega S} (M_0 + M_1) \right), \quad C_1 := \frac{\omega S}{24r}. \quad (8.15)$$

8.3 Step 3 (averaging over $\hat{\varphi}_1$)

Here, we shall exploit the fact that, for (p, q) different from $(1, 1)$ and from $(2, 1)$, the Hamiltonian $\bar{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ is independent of the angles, allowing to treat the angle $\hat{\varphi}_1$ as a “fast” angle in a suitable domain \hat{A} . Consider, therefore, the Hamiltonian

$$\hat{H}^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon) := \frac{\hat{I}_1^2}{2} + \varepsilon \bar{c}_0(\hat{I}_2) + \varepsilon^a H_1^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon), \quad (8.16)$$

and let

$$\hat{A} := \{\hat{I}_1 \in (-\frac{5}{4}r\varepsilon^\ell, -\frac{1}{2}r\varepsilon^b) \cup (\frac{1}{2}r\varepsilon^b, \frac{5}{4}r\varepsilon^\ell), \quad |\hat{I}_2 - \bar{J}_2| < \frac{5}{4}r\}. \quad (8.17)$$

In order to apply the Normal Form Lemma, we let b, γ_0 be as in 7.8 and let: $d := 2$, $(I, \varphi) := (\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2)$, $h := \hat{I}_1^2/2 + \varepsilon \bar{c}_0(\hat{I}_2)$, $f := \varepsilon^a H_1^{(1)}$, $D := \hat{A}$, $\rho := (r\varepsilon^b/4, r/4)$, $\rho_0 := r\varepsilon^b/4$, $\sigma := S/6$, $\Lambda := \{(0, k_2) \text{ s.t. } k_2 \in \mathbb{Z}\}$, $\alpha := r\varepsilon^b/8$, $K := r^2(2^{15}M_1^{(1)}\varepsilon^{\gamma_0})^{-1}$. With such positions, we see that we can apply Lemma 1.1, provided

$$\varepsilon \leq \min \left\{ \left[\frac{2^{12}M_1^{(1)}}{|c_{02}|r(\bar{J}_2 + 2r)} \right]^{1/(1-b-\gamma_0)}, \quad C_2^{1/\gamma_0} \right\}, \quad C_2 := \frac{r^2 S}{9 \cdot 2^{17} M_1^{(1)}}. \quad (8.18)$$

Under such condition, we can find a real-analytic symplectic transformation

$$\hat{\phi}_2 : (\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2) \in \hat{A}_{(r\varepsilon^b/8, r/8)} \times \mathbb{T}_{S/36}^2 \mapsto (\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2) \in \hat{A}_{(r\varepsilon^b/4, r/4)} \times \mathbb{T}_{S/6}^2$$

such that

$$\begin{aligned} \hat{H}^{(2)}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) &:= H^{(1)} \circ \hat{\phi}_2(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) \\ &:= \frac{\tilde{I}_1^2}{2} + \varepsilon \bar{c}_0(\tilde{I}_2) + \varepsilon^a H_1^{(2)}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_2; \varepsilon) + \hat{H}_*^{(2)}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) \end{aligned} \quad (8.19)$$

with

$$\begin{aligned} \|H_1^{(2)}\|_{(r\varepsilon^b/8, r/8), S/36} &\leq M_1^{(2)} := \frac{4}{3}M_1^{(1)}, \\ \|\hat{H}_*^{(2)}\|_{(r\varepsilon^b/8, r/8), S/36} &\leq \hat{M}_*^{(2)} := \varepsilon^a M_1^{(1)} \exp(-C_2 \varepsilon^{-\gamma_0}), \end{aligned}$$

and

$$|\hat{I}_1 - \tilde{I}_1|, |\hat{I}_2 - \tilde{I}_2| \leq \frac{3 \cdot 2^9 M_1^{(1)}}{Sr} \varepsilon^{a-b} \leq \frac{r\varepsilon^b}{2^9}, \quad \forall (\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2) \in \hat{A}_{(r\varepsilon^b/8, r/8)} \times \mathbb{T}_{S/36}^2. \quad (8.20)$$

Extend such symplectic transformation on $\hat{A}_{(r\varepsilon^b/8, r/8)} \times \mathbb{C} \times \mathbb{T}_{S/36}^3$ by setting

$$\phi_2(\tilde{I}_1, \tilde{I}_2, \tilde{I}_3, \tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3; \varepsilon) := (\hat{\phi}_2(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon), \tilde{I}_3, \tilde{\varphi}_3).$$

In this way, denoting $(\tilde{I}, \tilde{\varphi}) = (\tilde{I}_1, \tilde{I}_2, \tilde{I}_3, \tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3)$, we see that

$$\begin{aligned} H^{(2)}(\tilde{I}, \tilde{\varphi}; \varepsilon) &:= H^{(1)} \circ \phi_2(\tilde{I}, \tilde{\varphi}; \varepsilon) \\ &:= \frac{\tilde{I}_1^2}{2} + \omega \tilde{I}_3 + \varepsilon \bar{c}_0(\tilde{I}_2) + \varepsilon^a H_1^{(2)}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_2; \varepsilon) + H_*^{(2)}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}; \varepsilon) \end{aligned} \quad (8.21)$$

with

$$\|H_*^{(2)}\|_{(r\varepsilon^b/8, r/8, \infty), S/36} \leq M_*^{(2)} := \hat{M}_*^{(2)} + M_*^{(1)}.$$

In order to simplify the calculus of the constants we assume that

$$\varepsilon \leq \min \left\{ \left(\frac{C_1}{2C_2} \right)^{1/(\ell-\gamma_0)}, (\ell C_1)^{2/\ell} \right\}. \quad (8.22)$$

Using (8.22) it is simple to prove that

$$\varepsilon^{-1} \exp(C_1 \varepsilon^{-\ell}) \geq \varepsilon^{-a} \exp(C_2 \varepsilon^{-\gamma_0}). \quad (8.23)$$

In fact, using (8.22) and the fact that $a < 3/2$, it is sufficient to prove that $\exp(C_1 \varepsilon^{-\ell}) \geq \varepsilon^{-1}$ which is guaranteed⁸² again by (8.22).

By (8.23) we, also, obtain

$$M_*^{(2)} \leq \bar{M}_2^{(2)} \varepsilon^a \exp(-C_2 \varepsilon^{-\gamma_0}), \quad \bar{M}_2^{(2)} := \left(M_1 + \left(1 + \frac{8r}{\omega S} \right) (M_0 + M_1) \right). \quad (8.24)$$

8.4 Step 4 (energy conservation)

We are, now, in the position of concluding the proof of Theorem 7.1 for the case $(p, q) \neq (1, 1), (2, 1)$. The arguments we shall use, here, are based on energy conservation. However, such arguments, are not completely straightforward because we have to keep track of domains (recall that the variables $(\tilde{I}_1, \tilde{I}_2)$ are not defined in a neighborhood of the origin) and also because we shall freely use different sets of variables.

- (Energy conservation for the Hamiltonians $H^{(1)}$ and $H^{(2)}$)

Denote by $\hat{z}(t) := (\hat{I}(t), \hat{\varphi}(t))$ and $\tilde{z}(t) := (\tilde{I}(t), \tilde{\varphi}(t))$ the solutions of the Hamilton equations associated, respectively, to the Hamiltonians $H^{(1)}$ in (8.12) and $H^{(2)}$ in (8.21), with respective initial data $\hat{z}(0) := (\hat{I}(0), \hat{\varphi}(0))$ and $\tilde{z}(0) := (\tilde{I}(0), \tilde{\varphi}(0))$. Furthermore, if $F = F(\tilde{I}, \tilde{\varphi})$, denote $\hat{\Delta}_t F := F(\hat{z}(t)) - F(\hat{z}(0))$ and $\tilde{\Delta}_t F := F(\tilde{z}(t)) - F(\tilde{z}(0))$. Then, conservation of energy for the Hamiltonians in (8.12) and (8.21) yields⁸³:

$$\begin{aligned} \varepsilon c_{02} \left[\hat{I}_2(0) \hat{\Delta}_t \hat{I}_2 + \frac{1}{2} (\hat{\Delta}_t \hat{I}_2)^2 \right] &+ \left[\hat{I}_1(0) \hat{\Delta}_t \hat{I}_1 + \frac{1}{2} (\hat{\Delta}_t \hat{I}_1)^2 \right] \\ &+ \omega \hat{\Delta}_t \hat{I}_3 + \varepsilon^a \hat{\Delta}_t H_1^{(1)} + \hat{\Delta}_t H_*^{(1)} = 0, \end{aligned} \quad (8.25)$$

and

$$\begin{aligned} \varepsilon c_{02} \left[\tilde{I}_2(0) \tilde{\Delta}_t \tilde{I}_2 + \frac{1}{2} (\tilde{\Delta}_t \tilde{I}_2)^2 \right] &+ \left[\tilde{I}_1(0) \tilde{\Delta}_t \tilde{I}_1 + \frac{1}{2} (\tilde{\Delta}_t \tilde{I}_1)^2 \right] \\ &+ \omega \tilde{\Delta}_t \tilde{I}_3 + \varepsilon^a \tilde{\Delta}_t H_1^{(2)} + \tilde{\Delta}_t H_*^{(2)} = 0. \end{aligned} \quad (8.26)$$

⁸² Setting $x := \varepsilon^{-\ell}$ and $y := 1/\ell C_1$ we have to prove that $e^x \geq x^y$. This is obvious if $y \leq 1$; if $y > 1$ it is true if, for example, $x \geq y^2$.

⁸³ Recall (8.6) and observe that for any numbers x, y , one has $\frac{x^2}{2} - \frac{y^2}{2} = \frac{1}{2}(x-y)^2 + y(x-y)$.

- (*A-priori exponential estimates for the drift of \tilde{I}_1 , \hat{I}_3 and \tilde{I}_3*)

For⁸⁴ $0 \leq t \leq T(\varepsilon)$ we have directly by Hamilton equations, (8.23) and Cauchy estimates⁸⁵

$$|\tilde{\Delta}_t \tilde{I}_3| = |\hat{\Delta}_t \hat{I}_3| \leq \sup_{0 \leq \tau \leq t} |\partial_{\hat{\varphi}_3} H_2^{(1)}(\hat{z}(\tau); \varepsilon)| t \leq \frac{6}{eS} M_2^{(1)} t \leq \frac{M_0 + M_1}{96 \overline{M}_2^{(2)}} r \varepsilon^{2b}, \quad (8.27)$$

$$|\tilde{\Delta}_t \tilde{I}_1| \leq \sup_{0 \leq \tau \leq t} |\partial_{\hat{\varphi}_1} H_2^{(2)}(\tilde{z}(\tau); \varepsilon)| t \leq \frac{36}{eS} M_2^{(2)} t \leq \frac{1}{16} r \varepsilon^{2b}. \quad (8.28)$$

- Consider, now, (real) initial positions

$$(\hat{I}_1(0), \hat{I}_2(0)) \in \{|\hat{I}_1| \leq (1 + 2^{-7}) r \varepsilon^\ell\} \times \{|\hat{I}_2 - \bar{J}_2| \leq (1 + 2^{-7}) r\},$$

and let us consider, separately, two cases:

$$(i) \quad |\hat{I}_1(t)| < \frac{1}{2} r \varepsilon^b, \quad \forall 0 \leq t \leq T(\varepsilon);$$

$$(ii) \quad \exists 0 \leq t^* < T(\varepsilon) \quad \text{s.t.} \quad |\hat{I}_1(t)| < \frac{1}{2} r \varepsilon^b \quad \forall 0 \leq t < t^* \quad \text{and} \quad |\hat{I}_1(t^*)| \geq \frac{1}{2} r \varepsilon^b.$$

- (*Case (i) and stability of \hat{I}_2*)

Consider case (i): by (8.27) and (8.25), we see that, until

$$|\hat{I}_2(t) - \bar{J}_2| \leq 3r/2, \quad (8.29)$$

we have

$$|2\hat{I}_2(0)\hat{\Delta}_t \hat{I}_2 + (\hat{\Delta}_t \hat{I}_2)^2| \leq C_3 r^2 \varepsilon^{2b-1} \quad (8.30)$$

where we can take⁸⁶

$$C_3 := \frac{1}{|c_{02}|} \left(\frac{1}{4} + \frac{M_0 + M_1}{48 \overline{M}_2^{(2)}} \frac{\omega}{r} + \frac{16}{3} \frac{M_1^{(1)}}{r^2} + 4 \frac{\overline{M}_2^{(2)}}{r^2} \right). \quad (8.31)$$

We need, at this point, an elementary estimate (whose trivial proof is left to the reader):

Lemma 8.1 *Let $y, y_0 \in \mathbb{R}$ and $C > 0$ and suppose that*

$$|2y_0 y + y^2| \leq C^2.$$

Then:

⁸⁴ We shall consider only positive times since negative times are treated in a completely analogous way.

⁸⁵ “Cauchy estimates” allow to bound derivatives of analytic functions in terms of their sup-norm on larger domains; with our choice of norms, Cauchy estimates take the following form. Consider a 2π -periodic function $f(\varphi) := \sum_{k \in \mathbb{Z}} f_k e^{ik \cdot \varphi}$, analytic on \mathbb{T}_s with $\|f\|_s := \sum_k |f_k| e^{|k|s}$, then $\max_{\varphi \in \mathbb{T}} |\partial_\varphi f(\varphi)| \leq \frac{1}{\varepsilon s} \|f\|_s$. In fact for all $0 < \sigma < s$ we have $\max_{\varphi \in \mathbb{T}} |\partial_\varphi f(\varphi)| \leq \|\partial_\varphi f\|_{s-\sigma} = \sum_k (|k| e^{-|k|\sigma}) |f_k| e^{|k|s} \leq \frac{1}{\varepsilon \sigma} \sum_k |f_k| e^{|k|s}$ and taking the sup over $\sigma < s$ of the right hand side, we have the thesis.

⁸⁶ Recall (7.5), which implies $c_{02} \neq 0$.

(1) if $|y_0| \leq C$ then $|y| \leq |y_0| + \sqrt{y_0^2 + C^2} \leq (1 + \sqrt{2})C$,

(2) if $|y_0| > C$ then⁸⁷ $|y| \leq C^2|y_0|^{-1} \leq C$.

Let us now assume that

$$\varepsilon \leq \left(\frac{1}{20\sqrt{C_3}} \right)^{2/(2b-1)} \quad (8.32)$$

and let us apply the estimates of Lemma 8.1 to (8.30) with $C := \sqrt{C_3}r\varepsilon^\gamma$, $y_0 := \hat{I}_2(0)$ and $y := \hat{\Delta}\hat{I}_2$. Then:

$$|\hat{I}_2(0)| \leq \sqrt{C_3}r\varepsilon^\gamma \implies |\hat{I}_2(t) - \hat{I}_2(0)| \leq (1 + \sqrt{2})\sqrt{C_3}r\varepsilon^\gamma \leq \left(\frac{1}{8} - \frac{1}{2^8} \right)r \quad (8.33)$$

$$|\hat{I}_2(0)| > \sqrt{C_3}r\varepsilon^\gamma \implies |\hat{I}_2(t) - \hat{I}_2(0)| \leq \frac{C_3r^2}{|\hat{I}_2(0)|}\varepsilon^{2b-1} \leq \left(\frac{1}{8} - \frac{1}{2^8} \right)r, \quad (8.34)$$

which in particular imply (8.29).

- (Case (ii) and stability of \hat{I}_1)

If (ii) occurs, then, by (8.34), we have that

$$(1 + 2^{-7}) \geq |\hat{I}_1(t^*)| \geq r\varepsilon^b/2, \quad |\hat{I}_2(t^*) - \bar{J}_2| \leq 5r/4.$$

Then, by (8.20), we can find

$$(\tilde{I}_1^*, \tilde{I}_2^*, \tilde{\varphi}_1^*, \tilde{\varphi}_2^*) \in \hat{A}_{(\frac{r\varepsilon^b}{2^9}, \frac{r}{2^9})} \times \mathbb{T}^2 \text{ s.t. } \hat{\phi}(\tilde{I}_1^*, \tilde{I}_2^*, \tilde{\varphi}_1^*, \tilde{\varphi}_2^*) = (\hat{I}_1(t^*), \hat{I}_2(t^*), \hat{\varphi}_1(t^*), \hat{\varphi}_2(t^*)).$$

Now, as in (8.28), we have

$$|\tilde{I}_1(t) - \tilde{I}_1(t^*)| \leq r\varepsilon^{2b}/16 \quad (8.35)$$

hence, using (8.20),

$$\begin{aligned} |\hat{I}_1(t) - \hat{I}_1(0)| &\leq |\hat{I}_1(t) - \tilde{I}_1(t)| + |\tilde{I}_1(t) - \tilde{I}_1(t^*)| + |\tilde{I}_1(t^*) - \tilde{I}_1(0)| + |\tilde{I}_1(0) - \hat{I}_1(0)| \\ &\leq \left(1 + \frac{1}{16} + \frac{1}{2^8}\right)r\varepsilon^b \\ |\hat{I}_1(t)| &\leq \left(1 + \frac{1}{16} + \frac{1}{2^6}\right)r\varepsilon^\ell. \end{aligned}$$

Finally, using (8.13), we have (7.12), provided the following condition is, also, satisfied

$$\varepsilon \leq \left(\frac{r}{2^{10}\omega S(M_0 + M_1)} \right)^{1/1-b}. \quad (8.36)$$

- (Stability of \tilde{I}_2)

In order to prove stability for the I_2 -variable, we can apply (8.26) until

$$|\tilde{I}_1(t) - \tilde{I}_1^*| \leq \left(\frac{1}{8} - \frac{1}{2^9} \right)r\varepsilon^b \quad \text{and} \quad |\tilde{I}_2(t) - \tilde{I}_2^*| \leq \left(\frac{1}{8} - \frac{1}{2^9} \right)r \quad (8.37)$$

⁸⁷ We set $x := C^2y_0^{-2}$ and we have used that $\sqrt{1+x} - 1 \leq x/2$ and $1 - \sqrt{1-x} \leq x$ for $0 \leq x \leq 1$.

obtaining again (as for (8.30))

$$|2\tilde{I}_2^*(\tilde{I}_2(t) - \tilde{I}_2^*) + (\tilde{I}_2(t) - \tilde{I}_2^*)^2| \leq C_3 r^2 \varepsilon^{2b-1}. \quad (8.38)$$

We prove the first inequality in (8.37) using (8.20) and (8.35). As in case (i) we use Lemma 8.1 with $C := \sqrt{C_3} r \varepsilon^{\gamma_1}$, $y_0 := \tilde{I}_2^*$ and $y := \tilde{I}_2(t) - \tilde{I}_2^*$. Using again (8.32) we have that

$$|\tilde{I}_2^*| \leq \sqrt{C_3} r \varepsilon^{\gamma_1} \implies |\tilde{I}_2(t) - \tilde{I}_2^*| \leq (1 + \sqrt{2}) \sqrt{C_3} r \varepsilon^{\gamma_1} \leq \left(\frac{1}{8} - \frac{1}{2^8}\right) r, \quad (8.39)$$

$$|\tilde{I}_2^*| > \sqrt{C_3} r \varepsilon^{\gamma_1} \implies |\tilde{I}_2(t) - \tilde{I}_2^*| \leq \frac{C_3 r^2}{|\tilde{I}_2^*|} \varepsilon^{2b-1} \leq \left(\frac{1}{8} - \frac{1}{2^8}\right) r, \quad (8.40)$$

which, in particular, imply the second condition in (8.37).

• (Conclusion)

Finally (8.33), (8.34), (8.39), (8.40), (8.13), (8.20) imply (7.13) and (7.14), concluding the proof of Theorem 7.1 in the case $(p, q) \neq (1, 1), (2, 1)$. ■

9 Proof of Theorem in the case $(p, q) = (1, 1), (2, 1)$

We now turn to the special cases $(p, q) = (1, 1)$ or $(2, 1)$, in which case the Hamiltonian $\overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$ depends explicitly on $\hat{\varphi}_1$; see (8.6). In order to carry out the analogous of step 3 above, we have, first, to introduce action-angle variables for the two-dimensional integrable system $\frac{\hat{I}_1^2}{2} + \varepsilon \overline{H}_{01}(\hat{I}_2, \hat{\varphi}_1)$, which may be viewed as a “suspended pendulum” (with potential $\cos \hat{\varphi}_1$ or $\cos 2\hat{\varphi}_1$) having a small gravity varying with a second action-variable. Of course such action-angle variables will be *singular* near the separatrices and near elliptic periodic orbits (corresponding to the equilibria of the pendulum) and therefore a careful blow-up analysis near the singularities, as $\varepsilon \rightarrow 0$, is needed. The results were collected in Proposition 2.1.

We recall that the Hamiltonian (8.12) has, in the case $(p, q) = (1, 1)$ or $(2, 1)$, the form

$$H^{(1)} = \frac{\hat{I}_1^2}{2} + \omega \hat{I}_3 - \varepsilon k_p(\hat{I}_2)(1 + \cos j_p \hat{\varphi}_1) + \varepsilon h_p(\hat{I}_2) + \varepsilon^a H_1^{(1)} + H_*^{(1)}, \quad (9.1)$$

where

$$k_p(\hat{I}_2) := -\bar{d}_{j_p}(\hat{I}_2), \quad h_p(\hat{I}_2) := \bar{c}_0(\hat{I}_2) + k_p(\hat{I}_2), \quad j_1 = 2, \quad j_2 = 1. \quad (9.2)$$

In this subsection ξ_i will denote positive (ε -independent) constants and we will take ε as small as we need.

Choose $1 < \lambda \leq a - \gamma_0$ (here λ corresponds to $2b$). From (8.14) we deduce that

$$|\hat{I}_3(t) - \hat{I}_3(t_0)| \leq \xi_1 \varepsilon^\lambda, \quad \forall 0 \leq t_0 \leq t \leq T_1(\varepsilon) := \xi_3 \exp(-\xi_2/\varepsilon^\ell). \quad (9.3)$$

In order to prove stability in the other actions we state the following elementary Lemma concerning the conservation of energy.

Lemma 9.1 *Let $H := H(I, t; \mu) := h(I) + \mu f(t) \in \mathbb{R}$, $I, t, \mu \in \mathbb{R}$ and assume that h is analytic and not identically constant and that $|f(t)| \leq 1$ for all $t \in \mathbb{R}$. Fix $r_0 > 0$. Then, there exist $0 < \mu_0, v \leq 1$ and $c > 0$ such that, if for some continuous function $I(t) := I(t; \mu)$ with $|I_0| := |I(0)| \leq r_0$ $H(I(t), t; \mu) \equiv 0$, then for all $0 \leq \mu \leq \mu_0$ we have $|I(t) - I_0| \leq c\mu^v$.*

Proof Being h analytic we have that, if $N := \{|I| \leq 2r_0 \text{ s.t. } h'(I) = 0\}$, then $\#N < \infty$. Hence, there exists $p_* \in \mathbb{N}$ such that $\forall I_0 \in N$ there exist $1 < p_0 \leq p_*$ for which⁸⁸ $h^{(p_0)}(I_0) \neq 0$ and $h^{(p)}(I_0) = 0 \forall 1 \leq p \leq p_0$.

There exist $b_0 > 0$ and $0 \leq r'_0 \leq r_0$ such that $\forall I_0 \in N, |I_0| \leq r_0$ we have that $|h^{(p_0)}(\tilde{I})| \geq b_0, \forall |\tilde{I} - I_0| \leq r'_0$. We claim that the Lemma holds with $v \leq 1/p_*$, $c \leq (2p_0!/b_0)^{1/p_0}$ and $\mu_0 \leq (r'_0/c)^{p_*}$. In fact, by Taylor's formula, $\forall |I - I_0| \leq r'_0, \exists |I_* - I_0| \leq r'_0$ such that $h(I) - h(I_0) = h^{(p_0)}(I_*)(I - I_0)^{p_0}/p_0!$ and hence $|I(t) - I_0|^{p_0} \leq 2p_0!\mu/b_0$.

On the other hand, if $|I_0| \leq r_0$ with $I_0 \in \{|I| \leq 2r_0, \text{ s.t. } |I - I_1| \geq r'_0/2, \forall I_1 \in N\} =: M$, then, defining $m := \min_M |h'| > 0$, the Lemma holds with $v \leq 1, c \leq 2/m, \mu_0 \leq r'_0 m/4$, since $2\mu \geq |h(I) - h(I_0)| \geq m|I - I_0|$. ■

We consider first the case $(p, q) = (2, 1)$; the case $(p, q) = (1, 1)$, being analogous, will be considered later. For brevity we will omit the dependence on $p = 2$ in the formulas. In the Hamiltonian (9.1) we analyze first the following part, which represents a pendulum with a small gravity depending on a parameter:

$$E := E(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1) := E(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1; \varepsilon) := \frac{\hat{I}_1^2}{2} - \varepsilon k(\hat{I}_2)(1 + \cos \hat{\varphi}_1), \quad (9.4)$$

where $k(\hat{I}_2) = k_2(\hat{I}_2)$. We denote $E(t) := E(\hat{I}_1(t), \hat{I}_2(t), \hat{\varphi}_1(t))$.

We claim that if $0 \leq t_0 \leq t \leq T_1(\varepsilon)$ then

$$|E(t) - E(t_0)| \leq 4\varepsilon^\lambda \implies |\hat{I}_1(t) - \hat{I}_1(t_0)| \leq \xi_4 \sqrt{\varepsilon}, \quad |\hat{I}_2(t) - \hat{I}_2(t_0)| \leq \varepsilon^{\xi_5}. \quad (9.5)$$

In fact, if $|E(t) - E(t_0)| \leq 4\varepsilon^\lambda$ then the variable \hat{I}_1 may vary, at most, by order $\sqrt{\varepsilon}$ and using (9.3), we can apply the energy conservation to the Hamiltonian (9.1) obtaining that

$$|h(\hat{I}_2(t)) - h(\hat{I}_2(t_0))| \leq \xi_7 \varepsilon^{\lambda-1}.$$

Hence, by the fact that $h(\cdot)$ is a non constant analytic function (as it is immediate to verify), using Lemma 9.1 we get (9.5).

Now we want to apply Proposition 2.1 to the pendulum (9.4). We set $(p, I, q, \varphi) := (\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2)$, $(P, J, Q, \psi) = (\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2)$, $k := k_2$, $\Delta^0 := [\bar{I}_2 - \frac{5}{4}r, \bar{I}_2 + \frac{5}{4}r]$, $r_2 := r/4$, $\eta := \varepsilon^\lambda$, $R_0 := \frac{5}{4}r\varepsilon^\ell$, $r_1 := r\varepsilon^\ell/4$, $s_1 := s_2 := S/6$. If ε is sufficiently small we can apply Proposition 2.1 transforming the Hamiltonian

$$\hat{H}^{(1)} := \hat{H}^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon) := \frac{\hat{I}_1^2}{2} - \varepsilon k(\hat{I}_2)(1 + \cos \hat{\varphi}_1) + \varepsilon h(\hat{I}_2) + \varepsilon^a H_1^{(1)}$$

⁸⁸ We denote with $h^{(p)}$ the p -th derivative of h with respect to I .

into the new Hamiltonians

$$\tilde{H}^{(1)\pm} := \hat{H}^{(1)} \circ \phi^\pm(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) = E^\pm(\tilde{I}_1, \tilde{I}_2) + \varepsilon h(\tilde{I}_2) + \varepsilon^a H_1^{(1)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2)$$

which belongs to $\mathcal{H}_{\mathbb{R}}(\Omega_{(\rho_1, \rho_2)}^\pm \times \mathbb{T}_{\sigma_1} \times \mathbb{T}_{\sigma_2})$, where⁸⁹

$$\rho_1 := \xi_8 \varepsilon^{\lambda-1/2}, \quad \rho_2 := \xi_9 \varepsilon^{\lambda-1} \ln^{-1}(1/\varepsilon), \quad \sigma_1 := \xi_{10} S \ln^{-1}(1/\varepsilon), \quad \sigma_2 := \xi_{11} S.$$

We now perform the analogous of the Step 3 in § 2.4. In order to apply the Normal Form Lemma, we take ε sufficiently small and we set⁹⁰ $d := 2$, $(I, \varphi) := (\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2)$, $h := E^\pm(\tilde{I}_1, \tilde{I}_2) + \varepsilon h(\tilde{I}_2)$, $f := \varepsilon^a H_1^{(1)\pm}$, $D := \Omega^\pm$, $\rho := (\rho_1, \rho_2)$, $\rho_0 := \rho_1$, $\sigma := \sigma_1$, $\Lambda := \{(0, k_2) \text{ s.t. } k_2 \in \mathbb{Z}\}$, $\alpha := \xi_{12} \sqrt{\varepsilon} \ln^{-1} \varepsilon^{-1}$, $K := \xi_{13} \varepsilon^{-\gamma_0}$. So we find two real-analytic symplectic transformations $\hat{\phi}_1^\pm$ such that

$$\begin{aligned} \tilde{H}^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) &:= \tilde{H}^{(1)\pm} \circ \hat{\phi}_2^\pm(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) := \\ &E^\pm(\tilde{I}_1, \tilde{I}_2) + \varepsilon h(\tilde{I}_2) + \varepsilon^a H_1^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) + \tilde{H}_2^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) \end{aligned}$$

belongs to $\mathcal{H}_{\mathbb{R}}(\Omega_{(\rho_1, \rho_2)/2}^\pm \times \mathbb{T}_{\sigma_1/6} \times \mathbb{T}_{\sigma_2/6})$ and

$$|\tilde{I}_1 - \tilde{I}_1|, |\tilde{I}_2 - \tilde{I}_2| \leq \rho_1/2^7, \quad \|\tilde{H}_2^{(2)\pm}\| \leq \xi_{14} \varepsilon^a \exp(-\xi_{15} S/\varepsilon_0^\gamma). \quad (9.6)$$

Now we *complete* our two symplectic transformations defining

$$\phi_2^\pm(\tilde{I}_1, \tilde{I}_2, \tilde{I}_3, \tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3; \varepsilon) := (\hat{\phi}_1^\pm \circ \phi^\pm(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon), \tilde{I}_3, \tilde{\varphi}_3)$$

so that

$$\begin{aligned} H^{(2)\pm}(\tilde{I}, \tilde{\varphi}; \varepsilon) &:= H^{(1)} \circ \phi_2^\pm(\tilde{I}, \tilde{\varphi}; \varepsilon) := \\ &E^\pm(\tilde{I}_1, \tilde{I}_2) + \varepsilon h(\tilde{I}_2) + \omega \tilde{I}_3 + \varepsilon^a H_1^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon) + H_2^{(2)\pm}(\tilde{I}_1, \tilde{I}_2, \tilde{\varphi}_1, \tilde{\varphi}_2; \varepsilon). \end{aligned} \quad (9.7)$$

belongs to $\mathcal{H}_{\mathbb{R}}(\Omega_{(\rho_1, \rho_2)/2}^\pm \times \mathbb{C} \times \mathbb{T}_{\sigma_1/6} \times \mathbb{T}_{\sigma_2/6} \times \mathbb{T}_{S/6})$ with

$$\|H_2^{(2)\pm}\| \leq \xi_{16} \varepsilon^a \exp(-\xi_{17} S/\varepsilon_0^\gamma).$$

We now perform the analogous of Step 4 in § 2.5. Let

$$\tilde{\Omega}^\pm := \Omega_{(1+2^{-\gamma})(\rho_1, \rho_2)}^\pm \cap \mathbb{R}^2.$$

From the form of the Hamiltonian (9.7) we deduce that $\forall (\tilde{I}_1(t_0), \tilde{I}_2(t_0)) \in \tilde{\Omega}^\pm$

$$|\tilde{I}_1(t) - \tilde{I}_1(t_0)| \leq \xi_{18} \varepsilon^{\xi_{19}}, \quad \forall 0 \leq t_0 \leq t \leq T_2(\varepsilon) := \xi_{20} \exp(-\xi_{21}/\varepsilon^{\gamma_0}) < T_1(\varepsilon). \quad (9.8)$$

⁸⁹ We have $\xi_8 = c_4$, $\xi_9 = c_5/(\lambda - 1)$, $\xi_{10} = c_6/6(\lambda - 1)$, $\xi_{11} = 1/12$.

⁹⁰ We can make such a choice of α , using (2.5) and (2.6).

Since $\tilde{I}_3(t) - \tilde{I}_3(t_0) = \hat{I}_3(t) - \hat{I}_3(t_0)$, from (9.3) and (9.8) we deduce, using the energy conservation, that

$$\left| [E^\pm(\tilde{I}_1(t_0), \tilde{I}_2(t)) + \varepsilon h(\tilde{I}_2(t))] - [E^\pm(\tilde{I}_1(t_0), \tilde{I}_2(t_0)) + \varepsilon h(\tilde{I}_2(t_0))] \right| \leq \xi_{22} \varepsilon^{\xi_{23}}. \quad (9.9)$$

We now prove that $G^\pm(\cdot) := E^\pm(\tilde{I}_1(t_0), \cdot) + \varepsilon h(\cdot)$ are non constant analytic functions. From (8.7) and (9.2), it follows that

$$\frac{dG^\pm}{dy}(y) = \varepsilon [c_{02}y - k'(y)Y^\pm(\tilde{I}_1(t_0), y)]. \quad (9.10)$$

Now we observe that by (9.2), (8.8), (8.9), $k(y)$ is effectively defined and analytic for all $|y| < \bar{I}_1$ and the same is true for Y^\pm . Thus, from the fact that $\lim_{y \rightarrow (-\bar{I}_1)^+} k'(y) = 0$ (as it follows differentiating (8.8)) and that $|Y^\pm| \leq 1$, by (9.10) we deduce that

$$\lim_{y \rightarrow (-\bar{I}_1)^+} \frac{dG^\pm}{dy}(y) = -\varepsilon \bar{I}_1 c_{02},$$

which is different from 0 by (8.7) and the non-degeneracy assumption (7.5). This proves that G^\pm are non constant analytic functions. Finally, using (9.9), we can apply Lemma 9.1 and find $\xi_{24}, \xi_{25} > 0$ such that

$$|\tilde{I}_2(t) - \tilde{I}_2(0)| \leq \xi_{25} \varepsilon^{\xi_{24}}. \quad (9.11)$$

We remark that, in principle, ξ_{24}, ξ_{25} found with Lemma 9.1, depend on $\tilde{I}_1(t_0)$ but, since we work in compact sets $\tilde{\Omega}^\pm$, we can take them independent on $\tilde{I}_1(t_0)$.

We have proved stability for $0 \leq t_0 \leq t \leq T_2(\varepsilon)$ and $(\tilde{I}(t_0), \tilde{\varphi}(t_0)) \in \tilde{\Omega}^\pm \times \mathbb{R} \times \mathbb{T}^3$. By (9.6) this implies stability for $(\tilde{I}(t_0), \tilde{\varphi}(t_0)) \in \Omega^\pm \times \mathbb{R} \times \mathbb{T}^3$. By (2.4) this is equivalent to prove stability for $(\hat{I}(t_0), \hat{\varphi}(t_0)) \in M^\pm \times \mathbb{R} \times \mathbb{T}^3$ where

$$\begin{aligned} M^+ &:= \{(\hat{I}, \hat{\varphi}) \text{ s.t. } \varepsilon^\lambda \leq E(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1) \leq R_0^2/2, \hat{I}_2 \in \Delta^0\} \\ M^- &:= \{(\hat{I}, \hat{\varphi}) \text{ s.t. } -2\varepsilon k(\hat{I}_2) + \varepsilon^\lambda \leq E(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1) \leq -\varepsilon^\lambda, \hat{I}_2 \in \Delta^0\}. \end{aligned}$$

Using (9.5) and (9.3) it is immediate to prove stability for $(\hat{I}(0), \hat{\varphi}(0)) \in M \times \mathbb{R} \times \mathbb{T}^3$ and $0 \leq t \leq T_2(\varepsilon)$, where

$$M := \{(\hat{I}, \hat{\varphi}) \text{ s.t. } E(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1) \leq R_0^2/2, \hat{I}_2 \in \Delta^0\}.$$

Observing that $M \supset (A_{\frac{5}{4}R} \cap \mathbb{R}^3) \times \mathbb{T}^3$, by (8.13) and the fact that ϕ_0 is linear, we finally obtain (7.6). This finishes the proof in the case $(p, q) = (2, 1)$. ■

It remains to consider the case $(p, q) = (1, 1)$. The Hamiltonian (9.1) becomes

$$H^{(1)} = \frac{\hat{I}_1^2}{2} - \varepsilon k(\hat{I}_2)(1 + \cos 2\hat{\varphi}_1) + F \quad (9.12)$$

where $F := F(\hat{I}, \hat{\varphi}; \varepsilon) := \omega \hat{I}_3 + \varepsilon h_p(\hat{I}_2) + \varepsilon^a H_1^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}_1, \hat{\varphi}_2; \varepsilon) + H_*^{(1)}(\hat{I}_1, \hat{I}_2, \hat{\varphi}; \varepsilon)$. Next, we perform the following linear change of variables $\hat{I}_1^* := \hat{I}_1/2$, $\hat{I}_2^* := \hat{I}_2$, $\hat{I}_3^* := \hat{I}_3$, $\hat{\varphi}_1^* := 2\hat{\varphi}_1$, $\hat{\varphi}_2^* := \hat{\varphi}_2$, $\hat{\varphi}_3^* := \hat{\varphi}_3$, casting the Hamiltonian (9.12) into the form

$$H_*^{(1)} := H_*^{(1)}(\hat{I}^*, \hat{\varphi}^*; \varepsilon) := 4 \left[E(\hat{I}_1^*, \hat{I}_2^*, \hat{\varphi}_1^*; \varepsilon) + \frac{1}{4} F(2\eta_1^*, \hat{\varphi}_1^*/2) \right] \quad (9.13)$$

with

$$E(\hat{I}_1^*, \hat{I}_2^*, \hat{\varphi}_1^*; \varepsilon) := \frac{(\hat{I}_1^*)^2}{2} - \varepsilon \frac{k(\hat{I}_2^*)}{4} (1 + \cos \hat{\varphi}_1^*) . \quad (9.14)$$

For ease of notation, we have omitted in F the $(\hat{I}_2^*, \hat{I}_3^*, \hat{\varphi}_2^*, \hat{\varphi}_3^*; \varepsilon)$ -dependence, which, here, plays no rôle. We now apply⁹¹ Proposition 2.1 to E defined in (9.14), finding two symplectic change of variables $\hat{I}_1^* := p^\pm(P, Q)$, $\hat{\varphi}_1^* := q^\pm(P, Q)$, putting the Hamiltonian (9.13) in the form

$$H_{**}^{(1)} := 4[E^\pm(P) + F_*^\pm(P, Q)] , \quad (9.15)$$

where $F_*^\pm(P, Q) := \frac{1}{4} F(2p^\pm(P, Q), q^\pm(P, Q)/2)$. We note that the functions p^\pm are both 2π -periodic in Q . The function q^- , and hence F^- , is 2π -periodic in Q too; so we can define $(\check{I}_1, \check{\varphi}_1) := (P, Q)$ and proceed exactly as for $(p, q) = (2, 1)$, applying the Normal Form Lemma and the subsequent arguments.

The positive energy case is different: in fact $q^+(P, Q + 2\pi) = q^+(P, Q) + 2\pi$ so that F_*^+ is only 4π -periodic in Q , but in order to apply the Normal Form Lemma we need a 2π -periodic function. We, therefore, define another linear change of variables $P_0 := 2P$, $Q_0 := Q/2$, so that (9.15) becomes

$$H_{***}^{(1)} := 4[E^+(P_0/2) + F_{**}^+(P_0, Q_0)] . \quad (9.16)$$

where $F_{**}^+(P_0, Q_0) := F_*^+(P_0/2, 2Q_0)$ which is 2π -periodic in Q_0 . Therefore, we may define $(\check{I}_1, \check{\varphi}_1) := (P_0, Q_0)$ and proceed again as in the case $(p, q) = (2, 1)$. The proof of Theorem 7.1 is, now, complete. ■

⁹¹ Again we will omit the dependence on the variables $(\hat{I}_2^*, \hat{I}_3^*, \hat{\varphi}_2^*, \hat{\varphi}_3^*; \varepsilon)$.

Part IV

Topological instability: a new variational mechanism with optimal diffusion time

10 Statement of the main results

In this part we perform a mechanism to produce topological instability which is *alternative* to the one used first by Arnold [10] and become classical to prove action-instability in nearly-integrable Hamiltonian systems⁹². Indeed this method is *not based on the existence of a transition chain of tori*. In particular we avoid the KAM construction of the perturbed hyperbolic tori, proving directly the existence of a drifting orbit as a local minimum of an action functional. At the same time our variational approach achieves the optimal diffusion time, as a consequence of a general stability result, proved via classical perturbation theory.

We will consider nearly integrable *non-isochronous* Hamiltonian systems defined by

$$\mathcal{H}_\mu = \frac{I^2}{2} + \frac{p^2}{2} + (\cos q - 1) + \mu f(\varphi, q, t), \quad (10.1)$$

where $(\varphi, q, t) \in \mathbb{T}^d \times \mathbb{T}^1 \times \mathbb{T}^1$ are the angle variables, $(I, p) \in \mathbb{R}^d \times \mathbb{R}^1$ are the action variables and $\mu \geq 0$ is a small real parameter. The Hamiltonian system associated with \mathcal{H}_μ writes

$$\dot{\varphi} = I, \quad \dot{I} = -\mu \partial_\varphi f, \quad \dot{q} = p, \quad \dot{p} = \sin q - \mu \partial_q f. \quad (\mathcal{S}_\mu)$$

As in [49], the perturbation f is assumed to be a real trigonometric polynomial of order N in φ and t , namely⁹³

$$f(I, \varphi, p, q, t) = \sum_{|(n,l)| \leq N} f_{n,l}(I, p, q) e^{i(n \cdot \varphi + lt)}. \quad (10.2)$$

The unperturbed Hamiltonian system (\mathcal{S}_0) is completely integrable and in particular the energy $I_i^2/2$ of each rotator is a constant of the motion. Now, for $\mu \neq 0$, we want to exchange $O(1)$ -energy among the rotators and find the minimal time for which this exchange appens.

Let us define the “resonant web” \mathcal{D}_N , formed by the frequencies ω “resonant with the perturbation”

$$\mathcal{D}_N := \left\{ \omega \in \mathbb{R}^d \mid \exists (n, l) \in \mathbb{Z}^{d+1} \text{ s.t. } 0 < |(n, l)| \leq N \text{ and } \omega \cdot n + l = 0 \right\} = \bigcup_{0 < |(n,l)| \leq N} E_{n,l} \quad (10.3)$$

where $E_{n,l} := \{ \omega \in \mathbb{R}^d \mid \omega \cdot n + l = 0 \}$. Let us also consider the Poincaré-Melnikov primitive

$$\Gamma(\omega, \theta_0, \varphi_0) := - \int_{\mathbb{R}} \left[f(\omega t + \varphi_0, q_0(t), t + \theta_0) - f(\omega t + \varphi_0, 0, t + \theta_0) \right] dt,$$

where $q_0(t) = 4 \arctan(\exp t)$ is the separatrix of the unperturbed pendulum equation $\ddot{q} = \sin q$ satisfying $q_0(0) = \pi$.

⁹² See also the Introduction

⁹³ $\bar{f}_{n,l}(I, p, q) = f_{-n,-l}(I, p, q)$ for all $(n, l) \in \mathbb{Z}^d \times \mathbb{Z}$ with $|(n, l)| \leq N$ where \bar{z} denotes the complex conjugate of $z \in \mathbb{C}$.

The next Theorem (see Theorem 0.4) states that, for any connected component $\mathcal{C} \subset \mathcal{D}_N^c$, $\omega_I, \omega_F \in \mathcal{C}$, there exists a solution of (\mathcal{S}_μ) connecting a $O(\mu)$ -neighborhood of ω_I in the action space to a $O(\mu)$ -neighborhood of ω_F , in the time-interval $T_d = O((1/\mu)|\log \mu|)$.

Theorem 10.1 *Let \mathcal{C} be a connected component of \mathcal{D}_N^c , $\omega_I, \omega_F \in \mathcal{C}$ and let $\gamma : [0, L] \rightarrow \mathcal{C}$ be a smooth embedding such that $\gamma(0) = \omega_I$ and $\gamma(L) = \omega_F$. Assume that, for all $\omega := \gamma(s)$ ($s \in [0, L]$), $\Gamma(\omega, \cdot, \cdot)$ possesses a non-degenerate local minimum $(\theta_0^\omega, \varphi_0^\omega)$. Then $\forall \eta > 0$ there exists $\mu_0 = \mu_0(\gamma, \eta) > 0$ and $C = C(\gamma) > 0$ such that $\forall 0 < \mu \leq \mu_0$ there exists a solution $(I_\mu(t), \varphi_\mu(t), p_\mu(t), q_\mu(t))$ of (\mathcal{S}_μ) and two instants $\tau_1 < \tau_2$ such that $I_\mu(\tau_1) = \omega_I + O(\mu)$, $I_\mu(\tau_2) = \omega_F + O(\mu)$ and*

$$|\tau_2 - \tau_1| \leq \frac{C}{\mu} |\log \mu|. \quad (10.4)$$

Moreover $\text{dist}(I_\mu(t), \gamma([0, L])) < \eta$ for all $\tau_1 \leq t \leq \tau_2$.

In addition, the above result still holds for any perturbation $\mu(f + \mu \tilde{f})$ with any smooth $\tilde{f}(\varphi, q, t)$.

We can also build diffusion orbits approaching the boundaries of \mathcal{D}_N at distances as small as a certain power of μ : see for a precise statement Theorem 15.1.

As a byproduct of the techniques developed we have the following result (see Theorem 0.5), which is proved in section 15, concerning “Arnold’s example” [10] where $\mathcal{T}_\omega := \{I = \omega, \varphi \in \mathbb{T}^d, p = q = 0\}$ are, for all $\omega \in \mathbb{R}^d$, even for $\mu \neq 0$, invariant tori of (\mathcal{S}_μ) .

Theorem 10.2 *Let $f(\varphi, q, t) := (1 - \cos q)\tilde{f}(\varphi, t)$. Assume that for some smooth embedding $\gamma : [0, L] \rightarrow \mathbb{R}^d$, with $\gamma(0) = \omega_I$ and $\gamma(L) = \omega_F$, $\forall \omega := \gamma(s)$ ($s \in [0, L]$), $\Gamma(\omega, \cdot, \cdot)$ possesses a non-degenerate local minimum $(\theta_0^\omega, \varphi_0^\omega)$. Then $\forall \eta > 0$ there exists $\mu_0 = \mu_0(\gamma, \eta) > 0$, and $C = C(\gamma) > 0$ such that $\forall 0 < \mu \leq \mu_0$ there exists a heteroclinic orbit (η -close to γ) connecting the invariant tori \mathcal{T}_{ω_I} and \mathcal{T}_{ω_F} . Moreover the diffusion time T_d needed to go from a μ -neighbourhood of \mathcal{T}_{ω_I} to a μ -neighbourhood of \mathcal{T}_{ω_F} is bounded by $(C/\mu)|\log \mu|$ for some constant C .*

Our next stability result (see Theorem 0.6) proves the optimality of our estimate (10.4) on the diffusion time.

Theorem 10.3 *Let $f(I, \varphi, p, q, t)$ be as in (10.2), where the $f_{n,l}$ ($|(n, l)| \leq N$) are analytic functions. Then $\forall \kappa, \bar{r}, \tilde{r} > 0$ there exist $\mu_1, \kappa_0 > 0$ such that $\forall 0 < \mu \leq \mu_1$, for any solution $(I(t), \varphi(t), p(t), q(t))$ of (\mathcal{S}_μ) with $|I(0)| \leq \bar{r}$ and $|p(0)| \leq \tilde{r}$, there results*

$$|I(t) - I(0)| \leq \kappa \quad \forall t \text{ such that } |t| \leq \frac{\kappa_0}{\mu} \ln \frac{1}{\mu}. \quad (10.5)$$

Actually the proof of Theorem 10.3 contains much more information: in particular the stability time (10.5) is sharp only for orbits lying close to the separatrices. On the other hand the orbits lying far away from the separatrices are much more stable, namely

exponentially stable in time according to Nekhoroshev type time estimates, see (16.4) and (16.11). Indeed the diffusion orbit of Theorem 10.1 is found close to some pseudo-diffusion orbit whose (q, p) variables move along the separatrices of the pendulum.

This part is organized as follows: in section 11 we perform the finite dimensional reduction and we define the variational setting. In section 12 we provide a suitable development of the reduced action functional. In section 13 we prove the new results on the ergodization time. In section 14 we define the unperturbed pseudo-orbit. In section 15 we prove the existence of the diffusion orbit. In section 16 we prove the stability result, that is to say the optimality of our diffusion time.

Notations: In the following we shall use the notation $a(z_1, \dots, z_k) = O(b(\mu))$ will mean that, for a suitable positive constant $C(\gamma, f) > 0$, $|a(z_1, \dots, z_p)| \leq C(\gamma, f)|b(\mu)|$.

11 The variational setting and the finite dimensional reduction

Since the perturbation $f(\varphi, q, t)$ is purely spatial,⁹⁴ system (\mathcal{S}_μ) reduces to the second order system

$$\ddot{\varphi} = -\mu \partial_\varphi f(\varphi, q, t), \quad -\ddot{q} + \sin q = \mu \partial_q f(\varphi, q, t) \quad (11.1)$$

with associated Lagrangian

$$\mathcal{L}_\mu(\varphi, \dot{\varphi}, q, \dot{q}, t) = \frac{\dot{\varphi}^2}{2} + \frac{\dot{q}^2}{2} + (1 - \cos q) - \mu f(\varphi, q, t). \quad (11.2)$$

Using the Contraction Mapping Theorem we will prove in lemma 11.1 that, near the unperturbed solutions $(\omega(t-\theta) + \varphi_0, q_0(t-\theta))$ living on the stable and unstable manifolds of the unperturbed tori \mathcal{T}_ω , there exist, for μ small enough, solutions of the perturbed system (11.1) which connect the sections $\{\varphi = \varphi^+, q = -\pi, t = \theta^+\}$ and $\{\varphi = \varphi^-, q = \pi, t = \theta^-\}$ (under some assumptions). The diffusion orbit will be a chain of such connecting orbits.

We first introduce a few definitions and notations. For $\lambda := (\theta^+, \theta^-, \varphi^+, \varphi^-) \in \mathbb{R}^2 \times \mathbb{R}^{2d}$ with $\theta^+ < \theta^-$ we define $T_\lambda := \theta^- - \theta^+$ and the “mean frequency” $\omega_\lambda \in \mathbb{R}^d$ as $\omega_\lambda := \frac{\varphi^- - \varphi^+}{\theta^- - \theta^+}$. The “small denominator” of a frequency $\omega \in \mathbb{R}^d$ is defined by

$$\beta(\omega) := \beta_N(\omega) := \min_{0 < |(n, l)| \leq N} |n \cdot \omega + l|. \quad (11.3)$$

$\beta(\omega)$ measures how close the frequency ω lies to the resonant web \mathcal{D}_N defined in (10.3). We use the abbreviation β_λ for $\beta(\omega_\lambda)$. We shall always assume through this paper that ω stays in a fixed bounded set containing the curve γ .

⁹⁴We will develop all the computations for f . All the next arguments remain unchanged if the perturbation is $f + \mu \tilde{f}$, see the proof of Theorem 10.1.

For T large enough, there exists a unique T -periodic solution Q_T of the pendulum equation, of small positive energy with $Q_T(0) = -\pi$, $Q_T(T) = \pi$. Moreover Q_T satisfies $\forall t \in [0, T/2) \cup (T/2, T]$,

$$|\partial_T Q_T(t)| \leq K_1 e^{-K_2(T-t)} \quad , \quad |\partial_T(Q_T(T - \cdot))(t)| \leq K_1 e^{-K_2(T-t)}$$

and

$$|Q_T(t) - q_\infty(t)| + |\dot{Q}_T(t) - \dot{q}_\infty(t)| \leq K_1 e^{-K_2 T} \quad , \quad |\dot{Q}_T(t)| \leq K_1 \max\{e^{-K_2 t}, e^{-K_2(T-t)}\}, \quad (11.4)$$

for some positive constants K_1 and K_2 , where q_∞ is defined by

$$q_\infty(t) = q_0(t) - 2\pi \text{ if } t \in [0, T/2), \quad q_\infty(t) = q_0(t - T) \text{ if } t \in (T/2, T].$$

Lemma 11.1 *There exists $\mu_2 > 0$ and constants $C_0, C_1, \bar{c}, c_1 > 0$ such that $\forall 0 < \mu \leq \mu_2$, $\forall \lambda = (\theta^+, \theta^-, \varphi^+, \varphi^-)$ such that $C_0 \beta_\lambda^2 > \mu$ and $C_1 |\ln \mu| \leq T_\lambda \leq C_0 \beta_\lambda / \mu$ there exists a unique solution $(\varphi_\mu(t), q_\mu(t)) := (\varphi_{\mu, \lambda}(t), q_{\mu, \lambda}(t))$ of (11.1), defined for $t \in (\theta^+ - 1, \theta^- + 1)$, satisfying $\varphi_\mu(\theta^\pm) = \varphi^\pm$, $q_\mu(\theta^\pm) = \mp \pi$ and*

$$\begin{aligned} (i) \quad & |\varphi_\mu(t) - \bar{\varphi}(t)| \leq \bar{c}\mu(1 + c_1 \mu T_\lambda^2) / \beta_\lambda^2, \quad |\dot{\varphi}_\mu(t) - \omega| \leq \bar{c}\mu / \beta_\lambda, \\ (ii) \quad & |q_\mu(t) - Q_{T_\lambda}(t - \theta^+)| \leq \bar{c}\mu, \quad |\dot{q}_\mu(t) - \dot{Q}_{T_\lambda}(t - \theta^+)| \leq \bar{c}\mu, \end{aligned} \quad (11.5)$$

where $\bar{\varphi}(t) := \omega_\lambda(t - \theta^+) + \varphi^+$. Moreover $\varphi_{\mu, \lambda}(t)$, $\dot{\varphi}_{\mu, \lambda}(t)$, $q_{\mu, \lambda}(t)$ and $\dot{q}_{\mu, \lambda}(t)$ are C^1 functions of (t, λ) .

The proof of lemma 11.1 is given in the Appendix.

Remark 11.1 *Roughly, the meaning of the above estimates is the following.*

1) We have imposed $C_1 |\ln \mu| < T_\lambda := \theta^- - \theta^+$ so that by (11.4), on such intervals of time, the periodic solution Q_{T_λ} is $O(\mu)$ close to “separatrices” q_∞ of the unperturbed pendulum.

2) Estimate (ii) implies that for $t \approx (\theta^+ + \theta^-)/2$ the perturbed solution q_μ may have $O(\mu)$ oscillations around the unstable equilibrium of the pendulum $q = 0$, mod 2π , which is exactly what one expects perturbing with a general f . On the contrary for the class of perturbations considered in [10] as $f(\varphi, q, t) = (1 - \cos q)f(\varphi, t)$ preserving all the invariant tori, estimate (ii) can be improved, getting $\max\{|q_\mu(t) - Q_{T_\lambda}(t - \theta^+)|, |\dot{q}_\mu(t) - \dot{Q}_{T_\lambda}(t - \theta^+)|\} = O(\mu \max\{\exp(-C|t - \theta^+|), \exp(-C|t - \theta^-|)\})$.

3) For $\beta_\lambda \approx \sqrt{\mu}$ estimate (i) becomes meaningless: for a mean frequency ω_λ such that $n \cdot \omega_\lambda + l \approx \sqrt{\mu}$ for some $0 < |(n, l)| \leq N$ the perturbed transition orbits φ_μ are no more well-approximated by the straight lines $\bar{\varphi}(t) := \varphi^+ + \omega_\lambda(t - \theta^+)$.

Remark 11.2 *Let us define $\mathcal{D}_N^\beta := \{\omega \in \mathbb{R}^d \mid |\omega \cdot n + l| > \beta, \quad \forall 0 < |(n, l)| \leq N\}$. In [49] it is proved that hyperbolic invariant tori \mathcal{T}_ω^μ of system (S_μ) exist for Diophantine frequencies $\omega \in \mathcal{D}_{N_1}^{\beta_1}$, for some $\beta_1 = O(1)$ and some $N_1 = O(dN) > N$, namely avoiding more “resonances with the trigonometric polynomial f ” than just N . The presence of such “resonant hyperplanes $E_{n, l}$ ” for $N < |(n, l)| < N_1$ may be reflected in estimate (i) by the term μT_λ^2 . However such term, for our purposes, can be ignored. From this point of view lemma 11.1 could perhaps be interpreted as the first iterative step for looking at invariant hyperbolic tori in the perturbed system bifurcating from the unperturbed one’s.*

By lemma 11.1, for $0 < \mu \leq \mu_2$, we can define on the set

$$\Lambda_\mu := \left\{ \lambda = (\theta^+, \theta^-, \varphi^+, \varphi^-) \mid C_0 \beta_\lambda^2 > \mu, \ C_1 |\ln \mu| \leq T_\lambda \leq \frac{C_0 \beta_\lambda}{\mu} \right\},$$

the Lagrangian action functional $G_\mu : \Lambda_\mu \rightarrow \mathbb{R}$ as

$$G_\mu(\lambda) = G_\mu(\theta^+, \theta^-, \varphi^+, \varphi^-) := \int_{\theta^+}^{\theta^-} \mathcal{L}_\mu(\varphi_\mu(t), \dot{\varphi}_\mu(t), q_\mu(t), \dot{q}_\mu(t), t) dt. \quad (11.6)$$

We have

Lemma 11.2 G_μ is differentiable and (with the abbreviations φ, q for φ_μ, q_μ)

$$\nabla_{\varphi^+} G_\mu(\lambda) = -\dot{\varphi}(\theta^+), \quad \partial_{\theta^+} G_\mu(\lambda) = \frac{1}{2} |\dot{\varphi}(\theta^+)|^2 + \frac{1}{2} \dot{q}^2(\theta^+) + \cos q(\theta^+) - 1 + \mu f(\varphi^+, \pi, \theta^+)$$

$$\nabla_{\varphi^-} G_\mu(\lambda) = \dot{\varphi}(\theta^-), \quad \partial_{\theta^-} G_\mu(\lambda) = -\left(\frac{1}{2} |\dot{\varphi}(\theta^-)|^2 + \frac{1}{2} \dot{q}^2(\theta^-) + \cos q(\theta^-) - 1 + \mu f(\varphi^-, \pi, \theta^-) \right).$$

PROOF. By lemma 11.1 the map $(\lambda, t) \mapsto (\varphi_{\mu, \lambda}(t), \dot{\varphi}_{\mu, \lambda}(t), q_{\mu, \lambda}(t), \dot{q}_{\mu, \lambda}(t))$ is C^1 on the set $\{(\lambda, t) \in \Lambda_\mu \times \mathbb{R} \mid \theta^+ \leq t \leq \theta^-\}$. Hence G_μ is differentiable and

$$\begin{aligned} \partial_{\theta^+} G_\mu(\lambda) &= -\mathcal{L}_\mu(\varphi^+, \dot{\varphi}(\theta^+), -\pi, \dot{q}(\theta^+), \theta^+) \\ &+ \int_{\theta^+}^{\theta^-} \dot{\varphi}(s) \cdot \partial_{\theta^+} \dot{\varphi}(s) + \dot{q}(s) \partial_{\theta^+} \dot{q}(s) ds + \int_{\theta^+}^{\theta^-} \sin q(s) \partial_{\theta^+} q(s) \\ &- \mu \partial_{\varphi} f(\varphi(s), q(s), s) \cdot \partial_{\theta^+} \varphi(s) - \mu \partial_q f(\varphi(s), q(s), s) \partial_{\theta^+} q(s) ds. \end{aligned}$$

Integrating by parts and using that $(q_{\mu, \lambda}, \varphi_{\mu, \lambda})$ satisfies (11.1) in (θ^+, θ^-) , we obtain

$$\partial_{\theta^+} G_\mu(\lambda) = -\mathcal{L}_\mu(\varphi^+, \dot{\varphi}(\theta^+), -\pi, \dot{q}(\theta^+), \theta^+) + \left[\dot{q}(s) \partial_{\theta^+} q(s) + \dot{\varphi}(s) \cdot \partial_{\theta^+} \varphi(s) \right]_{\theta^+}^{\theta^-}.$$

Now $q_{\mu, \lambda}(\theta^+) = -\pi$ for all λ hence $\dot{q}(\theta^+) + \partial_{\theta^+} q(\theta^+) = 0$. Similarly we get $\dot{\varphi}(\theta^+) + \partial_{\theta^+} \varphi(\theta^+) = 0$, $\partial_{\theta^+} q(\theta^-) = 0$, $\partial_{\theta^+} \varphi(\theta^-) = 0$. As a consequence

$$\partial_{\theta^+} G_\mu(\lambda) = \frac{1}{2} |\dot{\varphi}|^2(\theta^+) + \frac{1}{2} \dot{q}^2(\theta^+) + (\cos q(\theta^+) - 1) + \mu f(\varphi^+, \pi, \theta^+).$$

The other partial derivatives are computed in the same way. ■

For $\beta > 0$ fixed, denoting $\lambda_i = (\theta_i, \theta_{i+1}, \varphi_i, \varphi_{i+1})$, we define on the set

$$\begin{aligned} \Lambda_{\mu, k} := \Lambda_{\mu, k}^\beta &:= \left\{ \lambda = (\theta_1, \dots, \theta_k, \varphi_1, \dots, \varphi_k) \in \mathbb{R}^k \times \mathbb{R}^{kd} \right. \\ &\left. \text{s.t. } \forall 1 \leq i \leq k-1, \quad \lambda_i \in \Lambda_\mu, \quad \beta_{\lambda_i} \geq \beta \right\}, \end{aligned}$$

the reduced action functional $\mathcal{F}_\mu : \Lambda_{\mu, k} \rightarrow \mathbb{R}$ as

$$\begin{aligned} \mathcal{F}_\mu(\lambda) &= \omega_I \varphi_1 - \frac{|\omega_I|^2}{2} \theta_1 + \mu \Gamma^u(\omega_I, \theta_1, \varphi_1) + \mu F(\omega_I, \theta_1, \varphi_1) + \sum_{i=1}^{k-1} G_\mu(\lambda_i) \\ &- \omega_F \varphi_k + \frac{|\omega_F|^2}{2} \theta_k + \mu \Gamma^s(\omega_F, \theta_k, \varphi_k) - \mu F(\omega_F, \theta_k, \varphi_k) \end{aligned} \quad (11.7)$$

where

$$\Gamma^u(\omega, \theta_0, \varphi_0) := - \int_{-\infty}^0 \left[f(\omega t + \varphi_0, q_0(t), t + \theta_0) - f(\omega t + \varphi_0, 0, t + \theta_0) \right] dt, \quad (11.8)$$

$$\Gamma^s(\omega, \theta_0, \varphi_0) := - \int_0^{+\infty} \left[f(\omega t + \varphi_0, q_0(t), t + \theta_0) - f(\omega t + \varphi_0, 0, t + \theta_0) \right] dt, \quad (11.9)$$

are called resp. the unstable and the stable Poincaré-Melnikov primitive, and

$$F(\omega, \theta_0, \varphi_0) := -f_{0,0}\theta_0 - \sum_{0 < |(n,l)| \leq N} f_{n,l} \frac{e^{i(n \cdot \varphi_0 + l\theta_0)}}{i(n \cdot \omega + l)}, \quad (11.10)$$

$f_{n,l} := f_{n,l}(0)$ being the Fourier coefficients of $f(\varphi, 0, t)$.

Critical points of the “reduced action functional” \mathcal{F}_μ give rise to diffusion orbits whose action variables I go from a small neighbourhood of ω_I to a small neighbourhood of ω_F , as stated in lemma 11.3 below. The “boundary terms” $\omega_I \varphi_1 - \frac{|\omega_I|^2}{2} \theta_1 + \mu \Gamma^u(\omega_I, \theta_1, \varphi_1) + \mu F(\omega_I, \theta_1, \varphi_1)$ and $-\omega_F \varphi_k + \frac{|\omega_F|^2}{2} \theta_k + \mu \Gamma^s(\omega_F, \theta_k, \varphi_k) - \mu F(\omega_F, \theta_k, \varphi_k)$ have been added also to enable us to find critical points of \mathcal{F}_μ w.r.t. all the variables (including $\theta_1, \varphi_1, \theta_k, \varphi_k$). More precisely, for $\lambda = (\theta, \varphi) \in \Lambda_{\mu,k}$ we define the pseudo diffusion solutions $(\varphi_{\mu,\lambda}, q_{\mu,\lambda})$ on the interval $[\theta_1, \theta_k]$ by

$$(\varphi_{\mu,\lambda}(t), q_{\mu,\lambda}(t)) := (\varphi_{\mu,\lambda_i}(t), q_{\mu,\lambda_i}(t) + 2\pi(i-1)) \quad \text{for } t \in [\theta_i, \theta_{i+1}],$$

where $(\varphi_{\mu,\lambda_i}(t), q_{\mu,\lambda_i}(t))$ are given by lemma 11.1. The pseudo diffusion solutions $(\varphi_{\mu,\lambda}, q_{\mu,\lambda})$ are then continuous functions which are true solutions of the equations of motion (11.1) on each interval (θ_i, θ_{i+1}) , but the time derivatives $(\dot{\varphi}_{\mu,\lambda}, \dot{q}_{\mu,\lambda})$ may undergo a jump at time θ_i . We have

Lemma 11.3 *If $\tilde{\lambda} = (\tilde{\theta}, \tilde{\varphi}) \in \Lambda_{\mu,k}$ is a critical point of \mathcal{F}_μ , then $(\varphi_{\mu,\tilde{\lambda}}(t), q_{\mu,\tilde{\lambda}}(t))$ is a solution of (11.1) in the time interval $(\tilde{\theta}_1, \tilde{\theta}_k)$. Moreover $\dot{\varphi}_\mu(\tilde{\theta}_1) = \omega_I + O(\mu)$, $\dot{\varphi}_\mu(\tilde{\theta}_k) = \omega_F + O(\mu)$, i.e. $(\varphi_{\mu,\tilde{\lambda}}, q_{\mu,\tilde{\lambda}})$ is a diffusion orbit between ω_I and ω_F with diffusion time $T_d = |\tilde{\theta}_k - \tilde{\theta}_1|$.*

PROOF. By lemma 11.2 if $\nabla_{\varphi_i} \mathcal{F}_\mu(\tilde{\lambda}) = 0$, then for $2 \leq i \leq k-1$, $\dot{\varphi}_{\mu,\tilde{\lambda}}(\tilde{\theta}_i^-) = \dot{\varphi}_{\mu,\tilde{\lambda}}(\tilde{\theta}_i^+)$ and $\dot{\varphi}_{\mu,\tilde{\lambda}}(\tilde{\theta}_1) = \omega_I + O(\mu)$, $\dot{\varphi}_{\mu,\tilde{\lambda}}(\tilde{\theta}_k) = \omega_F + O(\mu)$. Moreover, if $\nabla_{\varphi_i} \mathcal{F}_\mu(\tilde{\lambda}) = 0$ and $\partial_{\theta_i} \mathcal{F}_\mu(\tilde{\lambda}) = 0$ then (for $2 \leq i \leq k-2$), $\dot{q}_{\mu,\tilde{\lambda}}^2(\tilde{\theta}_i^+) = \dot{q}_{\mu,\tilde{\lambda}}^2(\tilde{\theta}_i^-)$. Now, by lemma 11.1 and (11.4), $\dot{q}_{\mu,\tilde{\lambda}}(\tilde{\theta}_i^\pm) = \dot{q}_0(0) + O(\mu)$. Hence $\dot{q}_{\mu,\tilde{\lambda}}(\tilde{\theta}_i^+) = \dot{q}_{\mu,\tilde{\lambda}}(\tilde{\theta}_i^-)$ and the proof is complete. ■

12 The approximation of the reduced functional

In order to prove the existence of critical points of the reduced action functional \mathcal{F}_μ thanks to the properties of the Poincaré-Melnikov primitives $\Gamma(\omega, \cdot, \cdot)$ we need an appropriate

expression of \mathcal{F}_μ , see lemma 12.5. We shall express \mathcal{F}_μ as the sum of a function whose definition contains the $\Gamma(\omega, \cdot, \cdot)$ (for which we can prove the existence of critical points) and of a remainder whose derivatives are so small that it cannot destroy the critical points of the first function.

The first lemma gives an approximation of G_μ (defined in (11.6)).

Lemma 12.1 *For $0 < \mu \leq \mu_3$, for $\lambda \in \Lambda_\mu$ we have*

$$G_\mu(\lambda) = \frac{1}{2} \frac{|\varphi^- - \varphi^+|^2}{(\theta^- - \theta^+)} + \mu \Gamma^s(\omega_\lambda, \theta^+, \varphi^+) + \mu \Gamma^u(\omega_\lambda, \theta^-, \varphi^-) - \mu \int_{\theta^+}^{\theta^-} f(\bar{\varphi}(t), 0, t) dt + R_0(\mu, \lambda) \quad (12.1)$$

where

$$\nabla_\lambda R_0(\mu, \lambda) = O\left(\frac{\mu^2(1 + \mu T_\lambda^2)}{\beta_\lambda^2} T_\lambda\right). \quad (12.2)$$

PROOF. By lemma 11.1, we can write $\varphi_{\mu,\lambda}(t) = \bar{\varphi}(t) + v_{\mu,\lambda}(t)$, $q_{\mu,\lambda}(t) = Q_{T_\lambda}(t - \theta^+) + w_{\mu,\lambda}(t)$, where $v_{\mu,\lambda}(\theta^+) = v_{\mu,\lambda}(\theta^-) = 0$, $\|\dot{v}_{\mu,\lambda}\|_{L^\infty(\theta^+, \theta^-)} = O(\mu/\beta_\lambda)$, $\|v_{\mu,\lambda}\|_{L^\infty(\theta^+, \theta^-)} = O((\mu/\beta_\lambda^2)(1 + \mu T_\lambda^2))$ and $w_{\mu,\lambda}(\theta^+) = w_{\mu,\lambda}(\theta^-) = 0$, $\|\dot{w}_{\mu,\lambda}\|_{L^\infty(\theta^+, \theta^-)} + \|w_{\mu,\lambda}\|_{L^\infty(\theta^+, \theta^-)} = O(\mu)$.

In the following, in order to avoid cumbersome notation, we shall use the abbreviations v, w, Q for $v_{\mu,\lambda}, w_{\mu,\lambda}, Q_{T_\lambda}(\cdot - \theta^+)$, the dependency w.r.t. λ and μ being implicit. We have

$$\begin{aligned} G_\mu(\lambda) &= \int_{\theta^+}^{\theta^-} \frac{1}{2} |\dot{\bar{\varphi}}(t)|^2 + \dot{\bar{\varphi}}(t) \cdot \dot{v}(t) + \frac{1}{2} |\dot{v}(t)|^2 + \frac{1}{2} \dot{Q}^2(t) + \dot{Q}(t) \dot{w}(t) + \frac{1}{2} \dot{w}^2(t) \\ &\quad + \int_{\theta^+}^{\theta^-} [1 - \cos(Q(t) + w(t))] - \mu f(\bar{\varphi}(t) + v(t), Q(t) + w(t), t) dt. \end{aligned}$$

Now since $v(\theta^+) = v(\theta^-) = 0$ and $w(\theta^+) = w(\theta^-) = 0$, $\int_{\theta^+}^{\theta^-} \dot{\bar{\varphi}}(t) \cdot \dot{v}(t) dt = \int_{\theta^+}^{\theta^-} \omega_\lambda \cdot \dot{v}(t) dt = 0$ and $\int_{\theta^+}^{\theta^-} \dot{Q}(t) \dot{w}(t) dt = \int_{\theta^+}^{\theta^-} -\ddot{Q}(t) w(t) dt = \int_{\theta^+}^{\theta^-} -(\sin Q(t)) w(t) dt$. As a result, $G_\mu(\lambda) = G_\mu^0(\lambda) + R_1(\lambda)$, where

$$G_\mu^0(\lambda) = \int_{\theta^+}^{\theta^-} \frac{1}{2} |\dot{\bar{\varphi}}|^2 + \frac{1}{2} \dot{Q}^2 + (1 - \cos Q) - \mu f(\bar{\varphi}, Q, t),$$

$$R_1(\lambda) = \int_{\theta^+}^{\theta^-} \frac{1}{2} |\dot{v}|^2 + \frac{1}{2} \dot{w}^2 + (\cos Q - \cos(Q + w) - w \sin Q) - \mu f(\bar{\varphi} + v, Q + w, t) + \mu f(\bar{\varphi}, Q, t).$$

We shall first prove that $|\nabla R_1| = O\left(\frac{\mu^2(1 + \mu T_\lambda^2)}{\beta_\lambda^2} T_\lambda\right)$. We have $\partial_{\theta^+} R_1 = r_s 1 + r_2 + r_3 + r_4 +$

$r_5 + r_6$ where

$$\begin{aligned}
r_1 &:= \int_{\theta^+}^{\theta^-} \dot{v} \cdot \frac{d}{dt}(\partial_{\theta^+} v) - \mu \partial_{\varphi} f(\bar{\varphi} + v, Q + w, t) \cdot (\partial_{\theta^+} v), \\
r_2 &:= \int_{\theta^+}^{\theta^-} \dot{w} \frac{d}{dt}(\partial_{\theta^+} w) + [\sin(Q + w) - \sin Q - \mu \partial_q f(\bar{\varphi} + v, Q + w, t)](\partial_{\theta^+} w), \\
r_3 &:= \int_{\theta^+}^{\theta^-} (-\sin Q + \sin(Q + w) - w \cos Q) \partial_{\theta^+} Q, \\
r_4 &:= \mu \int_{\theta^+}^{\theta^-} [\partial_{\varphi} f(\bar{\varphi}, Q, t) - \partial_{\varphi} f(\bar{\varphi} + v, Q + w, t)] \cdot \partial_{\theta^+} \bar{\varphi}, \\
r_5 &:= \mu \int_{\theta^+}^{\theta^-} [\partial_q f(\bar{\varphi}, Q, t) - \partial_q f(\bar{\varphi} + v, Q + w, t)] \partial_{\theta^+} Q, \\
r_6 &:= -\frac{1}{2} |\dot{v}(\theta^+)|^2 - \frac{1}{2} \dot{w}(\theta^+)^2.
\end{aligned}$$

Now v and w satisfy

$$\begin{cases} -\ddot{v}(t) &= \mu \partial_{\varphi} f(\bar{\varphi}(t) + v(t), Q(t) + w(t), t) \\ -\ddot{w}(t) + \sin(Q(t) + w(t)) &= \mu \partial_q f(\bar{\varphi}(t) + v(t), Q(t) + w(t), t) + \sin Q(t). \end{cases}$$

Moreover, deriving w.r.t. θ^+ the equality $v(\theta^+) = 0$ we obtain that $(\partial_{\theta^+} v)(\theta^+) = -\dot{v}(\theta^+)$. Similarly $(\partial_{\theta^+} w)(\theta^+) = -\dot{w}(\theta^+)$, $(\partial_{\theta^+} v)(\theta^-) = 0$ and $(\partial_{\theta^+} w)(\theta^-) = 0$. Therefore an integration by parts gives $r_1 = |\dot{v}(\theta^+)|^2$, $r_2 = \dot{w}(\theta^+)^2$ hence $|r_1| + |r_2| = O(\mu^2/\beta^2)$.

By the properties of Q_T , $\partial_{\theta^+} Q$ is bounded in the interval $[\theta^+, \theta^-]$ by a constant independent of λ . Moreover $-\sin Q(t) + \sin(Q(t) + w(t)) - w(t) \cos Q(t) = O(w(t)^2)$. Therefore $r_3 = O(\mu^2 T)$.

We have also, for some positive constant c ,

$$|r_4| + |r_5| \leq c\mu T \left[\sup_{t \in [\theta^+, \theta^-]} |\partial_{\theta^+} Q(t)| + |\partial_{\theta^+} \bar{\varphi}(t)| \right] \left[\sup_{t \in [\theta^+, \theta^-]} (|v(t)| + |w(t)|) \right].$$

Since $\partial_{\theta^+} \bar{\varphi}$ is bounded independently of λ , we have by lemma 11.1

$$|r_4| + |r_5| = O\left(\frac{\mu^2(1 + \mu T_{\lambda}^2)}{\beta_{\lambda}^2} T_{\lambda}\right).$$

Still by lemma 11.1, $r_6 = O(\mu^2/\beta^2)$. The estimate of the other derivatives of R_1 is obtained in the same way.

indent We now develop $G_{\mu}^0(\lambda)$ as

$$\begin{aligned}
G_{\mu}^0(\lambda) &= \frac{1}{2} \frac{|\varphi^- - \varphi^+|^2}{(\theta^- - \theta^+)} + \mu \Gamma^s(\omega_{\lambda}, \theta^+, \varphi^+) + \mu \Gamma^u(\omega_{\lambda}, \theta^-, \varphi^-) - \mu \int_{\theta^+}^{\theta^-} f(\bar{\varphi}(t), 0, t) dt \\
&+ R_2(\lambda) + R_3(\lambda),
\end{aligned}$$

where

$$R_2(\lambda) = \int_{\theta^+}^{\theta^-} \frac{1}{2} \dot{Q}^2(t) + (1 - \cos Q(t)) dt = \int_0^{T_{\lambda}} \frac{1}{2} \dot{Q}_{T_{\lambda}}^2(t) + (1 - \cos Q_{T_{\lambda}}(t)) dt, \quad (12.3)$$

$$R_3(\lambda) = \int_{\theta^+}^{\theta^-} -\mu \left[(f(\bar{\varphi}(t), Q(t), t) - f(\bar{\varphi}(t), 0, t)) \right] dt - \mu \Gamma^s(\omega_\lambda, \theta^+, \varphi^+) - \mu \Gamma^u(\omega_\lambda, \theta^-, \varphi^-).$$

There remains to prove estimate (12.2) for ∇R_2 and ∇R_3 . By (12.3) $\partial_{\varphi^\pm} R_2 = 0$ and $\partial_{\theta^+} R_2(\lambda) = -\partial_{\theta^-} R_2(\lambda)$ is the energy of the T_λ -periodic solution Q_{T_λ} of the pendulum equation. Now this energy is $O(e^{-c_2 T_\lambda})$. Hence (provided C_1 is large enough) $|\nabla R_2(\lambda)| = O(\mu^2)$.

In order to estimate the derivatives of R_3 , let us define $g(\varphi, q, t) := f(\varphi, q, t) - f(\varphi, 0, t)$. We have

$$R_3(\lambda) = \int_{\theta^+}^{\theta^-} -\mu g(\bar{\varphi}(t), Q(t), t) dt - \mu \Gamma^s(\omega_\lambda, \theta^+, \varphi^+) - \mu \Gamma^u(\omega_\lambda, \theta^-, \varphi^-) = \mu(a_3(\lambda) + b_3(\lambda))$$

where

$$a_3(\lambda) := - \int_0^{T_\lambda/2} g(\omega_\lambda t + \varphi^+, Q_{T_\lambda}(t), t + \theta^+) dt + \int_0^\infty g(\omega_\lambda t + \varphi^+, q_0(t), t + \theta^+) dt,$$

$$b_3(\lambda) := - \int_{-T_\lambda/2}^0 g(\omega_\lambda t + \varphi^-, Q_{T_\lambda}(t + T_\lambda), t + \theta^-) dt + \int_{-\infty}^0 g(\omega_\lambda t + \varphi^-, q_0(t), t + \theta^-) dt.$$

We have

$$\begin{aligned} a_3(\lambda) &= - \int_0^{T_\lambda/2} [g(\omega_\lambda t + \varphi^+, Q_{T_\lambda}(t), t + \theta^+) - g(\omega_\lambda t + \varphi^+, q_0(t), t + \theta^+)] \\ &\quad + \int_{T_\lambda/2}^\infty g(\omega_\lambda t + \varphi^+, q_0(t), t + \theta^+) dt. \end{aligned}$$

Recalling that $\sup_{t \in (0, T/2)} |\partial_T Q_T(t)| = O(e^{-c_2 T})$, $\sup_{t \in (0, T/2)} |Q_T(t) - q_0(t)| = O(e^{-c_2 T})$, it is easy to see that the derivatives of the first integral are $O(T_\lambda e^{-c_2 T_\lambda}) = O(\mu)$ (still provided C_1 is large enough). Moreover, using that $(|g(\omega_\lambda t + \varphi^+, q_0(t), t)| + |\partial_\varphi g(\omega_\lambda t + \varphi^+, q_0(t), t)| + |\partial_t g(\omega_\lambda t + \varphi^+, q_0(t), t)|) = O(q_0(t) - 2\pi) = O(e^{-c_2 t})$ for $t \in (T_\lambda/2, +\infty)$, we find that the derivatives of the second integral are $O(\mu)$ as well. Hence $|\nabla a_3(\lambda)| = O(\mu)$. The same estimate holds for b_3 . We then conclude that $\nabla R_3(\lambda) = O(\mu^2)$, which completes the proof of lemma 12.1. ■

In section 15 we will look for a critical point of \mathcal{F}_μ in the set

$$\begin{aligned} E &:= \left\{ \lambda = (\theta_1, \dots, \theta_k, \varphi_1, \dots, \varphi_k) \in \mathbb{R}^k \times \mathbb{R}^{kd} \right. \\ &\quad \left. \text{s.t. } \theta_i = \bar{\theta}_i + b_i, \quad \varphi_i = \bar{\varphi}_i + a_i, \quad |b_i| \leq 2\pi, \quad |a_i| \leq 2\pi \right\}, \end{aligned} \quad (12.4)$$

where $k, \bar{\varphi}_i, \bar{\theta}_i$ will be defined in section 14. It will result that $E \subset \Lambda_{\mu, k}$ (for some $\beta > 0$ depending on the curve γ). In particular, for all $\lambda \in E$

$$C_1 |\ln \mu| \leq \theta_{i+1} - \theta_i < \frac{C_0 \beta_i}{\mu}, \quad \forall i = 1, \dots, k-1, \quad (12.5)$$

where $\beta_i := \beta_{\lambda_i} := \beta(\omega_i)$ and $\omega_i := \omega_{\lambda_i} := (\varphi_{i+1} - \varphi_i)/(\theta_{i+1} - \theta_i)$. Moreover we will assume (see (14.8))

$$|\bar{\omega}_{i+1} - \bar{\omega}_i| \leq \rho\mu \quad \text{where } \bar{\omega}_i := \frac{\bar{\varphi}_{i+1} - \bar{\varphi}_i}{\bar{\theta}_{i+1} - \bar{\theta}_i} \quad (1 \leq i \leq k-1), \quad \omega_0 := \omega_I, \quad \omega_k := \omega_F \quad (12.6)$$

and $\rho > 0$ is a small constant to be chosen later (see (15.3)). For the time being, assuming (12.5) and (12.6), we want to give a suitable expression of \mathcal{F}_μ in E . By lemma 12.1, for $\lambda \in E$, we have

$$\begin{aligned} \mathcal{F}_\mu(\lambda) &= \sum_{i=1}^{k-1} \frac{1}{2} \frac{|\varphi_{i+1} - \varphi_i|^2}{\theta_{i+1} - \theta_i} + \omega_I \varphi_1 - \omega_F \varphi_k - \frac{|\omega_I|^2}{2} \theta_1 + \frac{|\omega_F|^2}{2} \theta_k \\ &+ \sum_{i=1}^k \mu \left(\Gamma^u(\omega_{i-1}, \theta_i, \varphi_i) + \Gamma^s(\omega_i, \theta_i, \varphi_i) \right) + \mu F(\omega_I, \theta_1, \varphi_1) \\ &- \sum_{i=1}^k \mu \int_{\theta_i}^{\theta_{i+1}} f(\omega_i(t - \theta_i) + \varphi_i, 0, t) dt - \mu F(\omega_F, \theta_k, \varphi_k) + \sum_{i=1}^{k-1} R_0(\mu, \lambda_i), \end{aligned} \quad (12.7)$$

where $|\nabla_\lambda R_0(\mu, \lambda)|$ satisfies (12.2). We shall write \mathcal{F}_μ in an appropriate form thanks to the following lemmas. The first one says how close the “mean frequencies” ω_i are to the unperturbed $\bar{\omega}_i$.

Lemma 12.2 *Let $\lambda = (\theta_1, \dots, \theta_k, \varphi_1, \dots, \varphi_k)$ belong to E . Then*

$$|\omega_i - \bar{\omega}_i| = O\left(\frac{1}{\theta_{i+1} - \theta_i}\right) = O\left(\frac{1}{|\ln \mu|}\right). \quad (12.8)$$

Moreover

$$\Gamma^u(\omega_{i-1}, \theta_i, \varphi_i) + \Gamma^s(\omega_i, \theta_i, \varphi_i) = \Gamma(\bar{\omega}_i, \theta_i, \varphi_i) + R_4(\lambda_i), \quad \text{where } \nabla R_4 = O(1/|\ln \mu|). \quad (12.9)$$

PROOF. Set $\Delta\theta_i := \theta_{i+1} - \theta_i$, $\Delta a_i := a_{i+1} - a_i$ and $\Delta b_i := b_{i+1} - b_i$. By an elementary computation we get $\omega_i - \bar{\omega}_i = -\bar{\omega}_i \Delta b_i / \Delta\theta_i + \Delta a_i / \Delta\theta_i$. By the definition of E and (12.5), estimate (12.8) follows.

From the definition of Γ^u, Γ^s and the exponential decay of q_0 it results that $\partial_\omega \Gamma^{u,s}$ is bounded by a uniform constant, as well as its partial derivatives. Hence (12.9) is a straightforward consequence of (12.8) and of (12.6). ■

Lemma 12.3 *For $0 < \mu \leq \mu_4$*

$$\mu F(\omega_I, \theta_1, \varphi_1) - \sum_{i=1}^k \mu \int_{\theta_i}^{\theta_{i+1}} f(\omega_i(t - \theta_i) + \varphi_i, 0, t) dt - \mu F(\omega_F, \theta_k, \varphi_k) = \sum_{i=1}^k R_5^i(\mu, \lambda_{i-1}, \lambda_i), \quad (12.10)$$

where, for all i ⁹⁵

$$\nabla R_5^i(\mu, \theta_{i-1}, \varphi_{i-1}, \theta_i, \varphi_i, \theta_{i+1}, \varphi_{i+1}) = O\left(\frac{\mu}{\beta_{i-1}^2(\theta_i - \theta_{i-1})} + \frac{\mu}{\beta_i^2(\theta_{i+1} - \theta_i)} + \frac{\mu|\beta_i - \beta_{i-1}|}{\beta_{i-1}\beta_i}\right). \quad (12.11)$$

⁹⁵In the cases $i = 1, i = k$ we only have $R_5^1 = R_5^1(\mu, \theta_1, \varphi_1, \theta_2, \varphi_2)$ and $R_5^k = R_5^k(\mu, \theta_{k-1}, \varphi_{k-1}, \theta_k, \varphi_k)$.

PROOF. We have

$$\begin{aligned} - \int_{\theta_i}^{\theta_{i+1}} f(\varphi_i + \omega_i(t - \theta_i), 0, t) dt &= F(\omega_i, \theta_{i+1}, \varphi_{i+1}) - F(\omega_i, \theta_i, \varphi_i) \\ &= (F(\omega_i, \theta_{i+1}, \varphi_{i+1}) - F(\omega_{i-1}, \theta_i, \varphi_i)) + (F(\omega_{i-1}, \theta_i, \varphi_i) - F(\omega_i, \theta_i, \varphi_i)), \end{aligned}$$

where $F(\omega, \cdot, \cdot)$ is defined in (11.10). We obtain

$$\mu F(\omega_I, \theta_1, \varphi_1) - \sum_{i=1}^{k-1} \mu \int_{\theta_i}^{\theta_{i+1}} f(\varphi_i + \omega_i(t - \theta_i), 0, t) dt - \mu F(\omega_F, \theta_k, \varphi_k) = \sum_{i=1}^k R_5^i$$

where

$$\begin{aligned} R_5^i &:= R_5^i(\mu, \theta_{i-1}, \varphi_{i-1}, \theta_i, \varphi_i, \theta_{i+1}, \varphi_{i+1}) := \mu (F(\omega_{i-1}, \theta_i, \varphi_i) - F(\omega_i, \theta_i, \varphi_i)) \\ &= -\mu \sum_{0 < |(n,l)| \leq N} f_{n,l} \frac{e^{i(n \cdot \varphi_i + l \theta_i)}}{i} \left(\frac{1}{(n \cdot \omega_{i-1} + l)} - \frac{1}{(n \cdot \omega_i + l)} \right) \end{aligned}$$

Now we prove (12.11). Let us consider for example $\partial_{\theta_i} R_5^i$. We have

$$\begin{aligned} \partial_{\theta_i} R_5^i &= \mu \partial_{\theta_i} (F(\omega_{i-1}, \theta_i, \varphi_i) - F(\omega_i, \theta_i, \varphi_i)) \\ &= \mu \left(\partial_{\omega} F(\omega_{i-1}, \theta_i, \varphi_i) \cdot \frac{-\omega_{i-1}}{(\theta_i - \theta_{i-1})} - \partial_{\omega} F(\omega_i, \theta_i, \varphi_i) \cdot \frac{\omega_i}{(\theta_{i+1} - \theta_i)} \right) \\ &\quad - \mu \left(\sum_{0 < |(n,l)| \leq N} f_{n,l} l e^{i(n \cdot \varphi_i + l \theta_i)} \left(\frac{1}{(n \cdot \omega_{i-1} + l)} - \frac{1}{(n \cdot \omega_i + l)} \right) \right), \end{aligned} \quad (12.12)$$

where

$$\partial_{\omega} F(\omega, \theta_0, \varphi_0) = \sum_{0 < |(n,l)| \leq N} f_{n,l} \frac{n e^{i(n \cdot \varphi_0 + l \theta_0)}}{i(n \cdot \omega + l)^2}, \quad (12.13)$$

Estimate (12.11) follows immediately from (12.12) and (12.13). The other partial derivatives of R_5^i can be estimated similarly. ■

Finally, to get a suitable expression of \mathcal{F}_μ , we find convenient to introduce coordinates $(b, c) \in \mathbb{R}^{(1+d)k}$ defined by (12.4) and

$$c_i = a_i - \bar{\omega}_i b_i, \quad \forall i = 1, \dots, k, \quad (12.14)$$

(we are just performing a linear change of coordinates adapted to the direction of the unperturbed flow at each i -transition $(b_i, a_i) = b_i(1, \bar{\omega}_i) + (0, c_i)$).

Lemma 12.4 *We have*

$$\begin{aligned} \sum_{i=1}^{k-1} \frac{1}{2} \frac{|\varphi_{i+1} - \varphi_i|^2}{(\theta_{i+1} - \theta_i)} + \omega_I \varphi_1 - \omega_F \varphi_k - \frac{|\omega_I|^2}{2} \theta_1 + \frac{|\omega_F|^2}{2} \theta_k = \\ \frac{1}{2} \sum_{i=1}^{k-1} \frac{|c_{i+1} - c_i|^2}{\Delta \bar{\theta}_i + (b_{i+1} - b_i)} + \sum_{i=1}^k R_6^i(\mu, \theta_i, \varphi_i, \theta_{i+1}, \varphi_{i+1}), \end{aligned} \quad (12.15)$$

where $\Delta\bar{\theta}_i := \bar{\theta}_{i+1} - \bar{\theta}_i$ and⁹⁶

$$\nabla R_6^i(\mu, \theta_{i-1}, \varphi_{i-1}, \theta_i, \varphi_i, \theta_{i+1}, \varphi_{i+1}) = O(\Delta\bar{\omega}_i) = O(\rho\mu). \quad (12.16)$$

PROOF. Let $\{\gamma_i\}_{i=1, \dots, k-1}$ be defined by $\varphi_{i+1} - \varphi_i = \bar{\omega}_i(\theta_{i+1} - \theta_i) + \gamma_i$. We can write $\omega_I\varphi_1 - \omega_F\varphi_k$ as

$$\begin{aligned} \omega_I\varphi_1 - \omega_F\varphi_k &= \sum_{i=1}^{k-1} \left((\bar{\omega}_{i-1} - \bar{\omega}_i)\varphi_i - \bar{\omega}_i(\varphi_{i+1} - \varphi_i) \right) + \varphi_k(\bar{\omega}_{k-1} - \omega_F) \\ &= \sum_{i=1}^{k-1} \left((\bar{\omega}_{i-1} - \bar{\omega}_i)\varphi_i - |\bar{\omega}_i|^2(\theta_{i+1} - \theta_i) - \bar{\omega}_i\gamma_i \right) + \varphi_k(\bar{\omega}_{k-1} - \omega_F). \end{aligned} \quad (12.17)$$

We can also write

$$-\frac{|\omega_I|^2}{2}\theta_1 + \frac{|\omega_F|^2}{2}\theta_k = \sum_{i=1}^{k-1} \left(\left(\frac{|\bar{\omega}_i|^2}{2} - \frac{|\bar{\omega}_{i-1}|^2}{2} \right)\theta_i + \frac{|\bar{\omega}_i|^2}{2}(\theta_{i+1} - \theta_i) \right) + \left(\frac{|\omega_F|^2}{2} - \frac{|\bar{\omega}_{k-1}|^2}{2} \right)\theta_k, \quad (12.18)$$

$$\sum_{i=1}^{k-1} \frac{1}{2} \frac{|\varphi_{i+1} - \varphi_i|^2}{(\theta_{i+1} - \theta_i)} = \sum_{i=1}^{k-1} \frac{|\bar{\omega}_i|^2}{2}(\theta_{i+1} - \theta_i) + \frac{1}{2} \frac{|\gamma_i|^2}{(\theta_{i+1} - \theta_i)} + \bar{\omega}_i\gamma_i. \quad (12.19)$$

Summing (12.17), (12.18) and (12.19) we get

$$\begin{aligned} \sum_{i=1}^{k-1} \frac{1}{2} \frac{|\varphi_{i+1} - \varphi_i|^2}{(\theta_{i+1} - \theta_i)} + \omega_I\varphi_1 - \omega_F\varphi_k - \frac{|\omega_I|^2}{2}\theta_1 + \frac{|\omega_F|^2}{2}\theta_k &= \sum_{i=1}^{k-1} \frac{1}{2} \frac{|\gamma_i|^2}{(\theta_{i+1} - \theta_i)} + \\ \sum_{i=1}^{k-1} \left(\frac{|\bar{\omega}_i|^2}{2} - \frac{|\bar{\omega}_{i-1}|^2}{2} \right)\theta_i + (\bar{\omega}_{i-1} - \bar{\omega}_i)\varphi_i + \varphi_k(\bar{\omega}_{k-1} - \omega_F) &+ \left(\frac{|\omega_F|^2}{2} - \frac{|\bar{\omega}_{k-1}|^2}{2} \right)\theta_k. \end{aligned} \quad (12.20)$$

Substituting $\bar{\varphi}_i + a_i$ for φ_i and $\bar{\theta}_i + b_i$ for θ_i , we get $\gamma_i = (a_{i+1} - a_i) - \bar{\omega}_i(b_{i+1} - b_i)$. Moreover the non constant terms in the right handside of (12.20) (i.e. those depending on a_i, b_i) are the first one and

$$\sum_{i=1}^k (\bar{\omega}_{i-1} - \bar{\omega}_i)a_i + \left(\frac{|\bar{\omega}_i|^2}{2} - \frac{|\bar{\omega}_{i-1}|^2}{2} \right)b_i =: \sum_{i=1}^k R^i(\mu, \theta_i, \varphi_i)$$

with $\nabla R^i(\mu, \theta_i, \varphi_i) = O(\Delta\bar{\omega}_i)$. Finally, expressing γ_i in terms of (b_i, c_i) we get $\gamma_i = (a_{i+1} - a_i) - \bar{\omega}_i(b_{i+1} - b_i) = (c_{i+1} - c_i) + b_{i+1}\Delta\bar{\omega}_i$ and then from (12.20), developing the square, we get (12.16). ■

From (12.7) lemmas 12.2, 12.3 and 12.4 we obtain the expression of \mathcal{F}_μ in the new coordinates (b, c) required to apply the variational argument of section 15.

Lemma 12.5 *There exists $\mu_5, C_2 > 0$ such that $\forall 0 < \mu \leq \mu_5$, if*

$$\beta_i \geq C_2 \max \{ \mu^{1/2}(\theta_{i+1} - \theta_i)^{1/2}, \mu(\theta_{i+1} - \theta_i)^{3/2}, (\theta_{i+1} - \theta_i)^{-1/2} \} \quad (12.21)$$

⁹⁶For $i = k$ we have $R_6^k = R_6^k(\mu, \theta_k, \varphi_k)$.

then

$$\begin{aligned}\mathcal{F}_\mu(b, c) &= \frac{1}{2} \sum_{i=1}^{k-1} \frac{|c_{i+1} - c_i|^2}{\Delta \bar{\theta}_i + (b_{i+1} - b_i)} + \mu \sum_{i=1}^k \Gamma(\bar{\omega}_i, \bar{\theta}_i + b_i, \bar{\varphi}_i + \bar{\omega}_i b_i + c_i) + R_7(b, c), \\ R_7(b, c) &:= \sum_{i=1}^k R_7^i(\mu, b_{i-1}, c_{i-1}, b_i, c_i, b_{i+1}, c_{i+1}),\end{aligned}\tag{12.23}$$

where⁹⁷

$$|\nabla R_7^i| \leq C_2 \rho \mu. \tag{12.24}$$

PROOF. It is easy to see that (12.6), (12.8) and (12.21) imply (provided μ is small enough) that

$$\frac{\beta_{i-1}}{2} \leq \beta_i \leq 2\beta_{i-1}, \quad |\beta_i - \beta_{i-1}| = O\left(\frac{1}{\theta_i - \theta_{i-1}} + \frac{1}{\theta_{i+1} - \theta_i} + \mu\right). \tag{12.25}$$

Noting that $\partial_{c_i} = \partial_{\varphi_i}$ and $\partial_{b_i} = \bar{\omega}_i \partial_{\varphi_i} + \partial_{\theta_i}$, estimate (12.24) follows from (12.2), (12.9), (12.11), (12.25) and (12.16). ■

13 Ergodization times

In order to define $\bar{\varphi}_i, \bar{\theta}_i$ ($1 \leq i \leq k$) we need some results, stated in this section, on the ergodization time of the torus $\mathbb{T}^l := \mathbb{R}^l / \mathbb{Z}^l$ for linear flows possibly resonant but only at a “sufficiently high level”,

Let $\Omega \in \mathbb{R}^l$; it is well known that, if $\Omega \cdot p \neq 0, \forall p \in \mathbb{Z}^l \setminus \{0\}$, then the trajectories of the linear flow $\{\Omega t + A\}_{t \in \mathbb{R}}$ are dense on \mathbb{T}^l for any initial point $A \in \mathbb{T}^l$. It is also intuitively clear that the trajectories of the linear flow $\{\Omega t + A\}_{t \in \mathbb{R}}$ will make an arbitrarily fine δ -net ($\delta > 0$) if Ω is resonant only at a sufficiently high level, namely if $\Omega \cdot p \neq 0, \forall p \in \mathbb{Z}^l$ with $0 < |p| \leq M(\delta)$ for some large enough $M(\delta)$. Let us make more precise and quantitative these considerations.

For any $\Omega \in \mathbb{R}^l$ define the ergodization time $T(\Omega, \delta)$ required to fill \mathbb{T}^l within $\delta > 0$ as

$$T(\Omega, \delta) = \inf \left\{ t \in \mathbb{R}_+ \mid \forall x \in \mathbb{R}^l, d(x, A + [0, t]\Omega + \mathbb{Z}^l) \leq \delta \right\},$$

where d is the Euclidean distance and A some point of \mathbb{R}^l . $T(\Omega, \delta)$ is clearly independent of the choice of A . Above and in what follows, $\inf E$ is equal to $+\infty$ if E is empty. For $R > 0$ let

$$\alpha(\Omega, R) = \inf \left\{ |p \cdot \Omega| \mid p \in \mathbb{Z}^l, p \neq 0, |p| \leq R \right\}.$$

Theorem 13.1 $\forall l \in \mathbb{N}$ there exists a positive constant a_l such that, $\forall \Omega \in \mathbb{R}^l, \forall \delta > 0$, $T(\Omega, \delta) \leq (\alpha(\Omega, a_l/\delta))^{-1}$. Moreover $T(\Omega, \delta) \geq (1/4)\alpha(\Omega, 1/4\delta)^{-1}$.

⁹⁷In the cases $i = 1, i = k$ we have $R_7^1 = R_7^1(\mu, \theta_1, \varphi_1, \theta_2, \varphi_2)$ and $R_7^k = R_7^k(\mu, \theta_{k-1}, \varphi_{k-1}, \theta_k, \varphi_k)$.

In the above Theorem α^{-1} is equal to 0 if $\alpha = +\infty$ and to $+\infty$ if $\alpha = 0$.

Remark 13.1 Assume that Ω is a C - τ Diophantine vector, i.e. there exist $C > 0$ and $\tau \geq l - 1$ such that $\forall k \in \mathbb{Z}^l$ $|k \cdot \Omega| \geq C/|k|^\tau$. Then $\alpha(\Omega, R) \geq C/R^\tau$ and so $T(\Omega, \delta) \leq a_l^\tau/C\delta^\tau$. This estimate was proved in Theorem D of [47]. Also Theorem B of [47] is an easy consequence of Theorem 13.1.

Theorem 13.1 is a direct consequence of more general statements, see Theorem 13.2 and remark 13.2. Let us introduce first some notations. Let Λ be a lattice of \mathbb{R}^l , i.e. a discrete subgroup of \mathbb{R}^l such that \mathbb{R}^l/Λ has finite volume. For all $\Omega \in \mathbb{R}^l$ we define

$$T(\Lambda, \Omega, \delta) = \inf \left\{ t \in \mathbb{R}_+ \mid \forall x \in \mathbb{R}^l \ d(x, [0, t]\Omega + \Lambda) \leq \delta \right\}$$

($T(\Lambda, \Omega, \delta)$ is the time required to have a δ -net of the torus \mathbb{R}^l/Λ endowed with the metric inherited from \mathbb{R}^l). For $R > 0$, let

$$\Lambda^* = \{p \in \mathbb{R}^l \mid \forall \lambda \in \Lambda, p \cdot \lambda \in \mathbb{Z}\} \quad \text{and} \quad \Lambda_R^* = \{p \in \Lambda^* \mid 0 < |p| \leq R\}$$

(Λ^* is a lattice of \mathbb{R}^l which is conjugated to Λ). We define

$$\alpha(\Lambda, \Omega, R) = \inf \left\{ |p \cdot \Omega| \mid p \in \Lambda_R^* \right\}.$$

The following result holds:

Theorem 13.2 $\forall l \in \mathbb{N}$ there exists a positive constant a_l such that, for all lattice Λ of \mathbb{R}^l , $\forall \Omega \in \mathbb{R}^l$, $\forall \delta > 0$, $T(\Lambda, \Omega, \delta) \leq (\alpha(\Lambda, \Omega, a_l/\delta))^{-1}$.

Remark 13.2 It is fairly obvious that $T(\Lambda, \Omega, \delta) \geq (1/4)\alpha(\Lambda, \Omega, 1/4\delta)^{-1}$. Indeed, assume that $\Lambda_{1/4\delta}^* \neq \emptyset$ and let $p \in \Lambda_{1/4\delta}^*$ be such that $p \cdot \Omega = \alpha := \alpha(\Lambda, \Omega, 1/4\delta)$. Let $x \in \mathbb{R}^l$ satisfy $p \cdot x = 1/2$. Then $\forall t \in [0, 1/4\alpha)$, $\forall \lambda \in \Lambda$,

$$|x - (t\Omega + \lambda)| \geq \frac{|p \cdot (x - t\Omega - \lambda)|}{|p|} \geq 4\delta |p \cdot x - tp \cdot \Omega - p \cdot \lambda|,$$

and $p \cdot x - p \cdot \lambda \in (1/2) + \mathbb{Z}$, whereas $|tp \cdot \Omega| = t\alpha < 1/4$. Hence $|x - (t\Omega + \lambda)| > \delta$.

In the next section we will apply Theorem 13.1 when $\Omega = (\omega, 1) \in \mathbb{R}^{d+1}$. The proof of Theorem 13.2 is given in the Appendix. We could give an explicit expression of a_l . However it is not useful for our purpose and the constants a_l which can be derived from our proof are certainly far from being optimal.

14 The unperturbed pseudo-diffusion orbit

Consider the set Q_M of “non-ergodizing frequencies”

$$Q_M := \left\{ \omega \in \mathbb{R}^d \mid \exists (n, l) \in \mathbb{Z}^{d+1} \text{ with } 0 < |(n, l)| \leq M, \text{ and } \omega \cdot n + l = 0 \right\} = \bigcup_{h \in S_M} E_h$$

where $S_M := \{h = (n, l) \in (\mathbb{Z}^d \setminus \{0\}) \times \mathbb{N} \mid 0 < |h| \leq M, h \neq jh', \forall j \in \mathbb{Z}, h' \in (\mathbb{Z}^d \setminus \{0\}) \times \mathbb{N}\}$ and $E_h = E_{n,l} := \{\omega \in \mathbb{R}^d \mid (\omega, 1) \cdot h = \omega \cdot n + l = 0\}$. By Theorem 13.1 (or Theorem 13.2, with $\Lambda = 2\pi\mathbb{Z}^{d+1}$), for $\delta > 0$, if ω belongs to

$$Q_M^c = \left\{ \omega \in \mathbb{R}^d \mid \omega \cdot n + l \neq 0, \forall 0 < |(n, l)| \leq M \right\}, \quad (14.1)$$

with $M = 8\pi a_{d+1}/\delta$, then the flow of $(\omega, 1)$ provides a $\delta/4$ -net of the torus \mathbb{T}^{d+1} .

Moreover if $\omega \notin Q_M$ then for all $(n, l) \in \mathbb{Z}^d \setminus \{0\} \times \mathbb{Z}$,

$$|n \cdot \omega + l| = |n| \text{dist}(\omega, E_{n,l}) \geq \text{dist}(\omega, E_{n,l}) \geq \text{dist}(\omega, Q_M) > 0. \quad (14.2)$$

By Theorem 13.1 (or Theorem 13.2), we deduce from (14.2) the estimate

$$T((\omega, 1), \delta/4) \leq \frac{2\pi}{\text{dist}(\omega, Q_M)} \quad (14.3)$$

which measures the divergence of the ergodization time $T((\omega, 1), \delta)$ as ω approaches the set Q_M .

Definition 14.1 *Given $M > 0$, a connected component \mathcal{C} of \mathcal{D}_N^c and $\omega_I, \omega_F \in \mathcal{C}$, we say that an embedding $\gamma \in C^2([0, L], \mathcal{C})$ is a Q_M -admissible connecting curve between ω_I and ω_F if the following properties are satisfied:*

- (a) $\gamma(0) = \omega_I$, $\gamma(L) = \omega_F$, $|\dot{\gamma}(s)| = 1 \forall s \in (0, L)$,
- (b) $\forall h = (n, l) \in S_M$, $\forall s \in [0, L]$ such that $\gamma(s) \in E_h$, $n \cdot \dot{\gamma}(s) \neq 0$.

Condition (b) means that for all $h \in S_M$, $\gamma([0, L])$ may intersect E_h transversally only. It is easy to see that condition (b) implies that $\mathcal{I}(\gamma) = \{s \in [0, L] \mid \gamma(s) \in Q_M\}$ is finite and that there exists $\nu > 0$ such that for all $s \in \mathcal{I}(\gamma)$, for all $h = (n, l) \in S_M$ such that $\gamma(s) \in E_h$, $|\dot{\gamma}(s) \cdot n|/|n| \geq \nu$.

If a curve α is not admissible we can always find “close to it” an admissible one γ . Indeed the following lemma holds.

Lemma 14.1 *Let $M > 0$, \mathcal{C} be a connected component of \mathcal{D}_N^c , $\omega_I, \omega_F \in \mathcal{C}$ and let $\alpha \in C^2([0, L_0], \mathcal{C})$ be an embedding with $\alpha(0) = \omega_I$ and $\alpha(L_0) = \omega_F$. Then, $\forall \eta > 0$, there exists a curve γ , Q_M -admissible between ω_I and ω_F , satisfying $\text{dist}(\gamma(s), \alpha([0, L_0])) < \eta$, $\forall s \in [0, L]$.*

PROOF. First it is easy to see that there exists an embedding $\alpha_1 : [0, L_1] \rightarrow \mathcal{C}$ such that $\alpha_1(0) = \omega_I, \alpha_1(L_1) = \omega_F$, $\text{dist}(\alpha_1(s), \alpha([0, L_0])) \leq \eta/4$ and $\forall h = (n, l) \in S_M, \omega_I \notin E_h$ (resp. $\omega_F \notin E_h$) or $\dot{\alpha}_1(0) \cdot n \neq 0$ (resp. $\dot{\alpha}_1(L_1) \cdot n \neq 0$).

Let $r > 0, \nu_1 > 0$ be such that $\forall s \in [0, r] \cup [L_1 - r, L_1], \forall h = (n, l) \in S_M, \text{dist}(\alpha_1(s), E_h) \geq \nu_1$ or $|\dot{\alpha}_1(s) \cdot n| \geq \nu_1$. Let $\phi : [0, L_1] \rightarrow [0, 1]$ be a smooth function such that $\phi(0) = \phi(L_1) = 0$ and $\forall s \in [r, L_1 - r] \phi(s) = 1$.

We shall prove that for all $\varepsilon > 0$ there exists $\omega_\varepsilon \in \mathbb{R}^d, |\omega_\varepsilon| < \varepsilon$, such that $\forall h = (n, l) \in S_M$, for all $s \in [r, L_1 - r]$ such that $\alpha_1(s) \in E_h + \omega_\varepsilon, \dot{\alpha}_1(s) \cdot n \neq 0$. For $h = (n, l) \in S_M$, let $\mathcal{J}_h = \{s \in [r, L_1 - r] \mid n \cdot \dot{\alpha}_1(s) = 0\}$ and $\mathcal{V}_h = \{\alpha_1(s) - u \mid s \in \mathcal{J}_h, u \in E_h\}$. Let $\psi_h : [r, L_1 - r] \times E_h \rightarrow \mathbb{R}^d$ be defined by $\psi_h(s, u) = \alpha_1(s) - u$. $D\psi_h(s, u)$ is singular iff $s \in \mathcal{J}_h$. Therefore \mathcal{V}_h is the set of the critical values of ψ_h and by Sard's lemma, $\text{meas}(\mathcal{V}_h) = 0$. Hence for all $\varepsilon > 0$ there exists $\omega_\varepsilon \in \mathbb{R}^d$ such that $|\omega_\varepsilon| < \varepsilon, \omega_\varepsilon \notin \mathcal{V}_h$ for all $h \in S_M$. Our claim follows.

Now we can define $\alpha_2 : [0, L_1] \rightarrow \mathcal{C}$ by $\alpha_2(s) = \alpha_1(s) - \phi(s)\omega_\varepsilon$. It is easy to check that, provided ε is small enough, α_2 is an embedding which satisfies condition (b). γ is obtained from α_2 by a simple time reparametrization. ■

If $\Gamma(\alpha(s), \cdot, \cdot)$ possesses, for each s , a non-degenerate local minimum $(\theta_0^{\alpha(s)}, \varphi_0^{\alpha(s)})$, then, by the Implicit Function Theorem, along any curve γ sufficiently close to α , $\Gamma(\gamma(s), \cdot, \cdot)$ possesses local minima $(\theta_0^{\gamma(s)}, \varphi_0^{\gamma(s)})$ such that

$$D_{(\theta, \varphi)}^2 \Gamma(\gamma(s), \theta_0^{\gamma(s)}, \varphi_0^{\gamma(s)}) > \lambda \text{Id}, \quad \forall s \in [0, L], \quad (14.4)$$

for some constant $\lambda > 0$ depending on α . Therefore, by the above lemma, it is enough to prove the existence of drifting orbits along admissible curves γ . Property (14.4) will be used in lemma 15.1.

Given a Q_M -admissible curve γ , let us call s_1^*, \dots, s_r^* the elements of $\mathcal{I}(\gamma)$, and $\omega_1^* = \gamma(s_1^*), \dots, \omega_r^* = \gamma(s_r^*)$ the corresponding frequencies. Since, $\forall m = 1, \dots, r, (\theta_0^{\omega_m^*}, \varphi_0^{\omega_m^*})$ is a nondegenerate local minimum of $\Gamma(\omega_m^*, \cdot, \cdot)$, there is a neighborhood W_m of ω_m^* such that, $\forall \omega \in W_m, \Gamma(\omega, \cdot)$ admits a nondegenerate local minimum $(\theta_0^\omega, \varphi_0^\omega)$, the map $\omega \mapsto (\theta_0^\omega, \varphi_0^\omega)$ being Lipschitz-continuous on W_m . Therefore we shall assume without loss of generality that for all $m = 1, \dots, r$,

$$\forall (\omega, \omega') \in (W_m \cap \gamma([0, L]))^2 \quad |(\theta_0^\omega, \varphi_0^\omega) - (\theta_0^{\omega'}, \varphi_0^{\omega'})| \leq K|\omega - \omega'|. \quad (14.5)$$

It is easy to prove that, if γ is an admissible curve, there exists $d_0 > 0$ such that

- (*) $\{s \in [0, L] \mid \text{dist}(\gamma(s), Q_M) \leq d_0\}$ is the union of a finite number of disjoint intervals $[S_1, S'_1], \dots, [S_r, S'_r]$; for all $m = 1, \dots, r$ each interval $[S_m, S'_m]$ intersects $\mathcal{I}(\gamma)$ at a unique point s_m^* and $\gamma([S_m, S'_m]) \subset W_m$. Moreover $(s \mapsto \text{dist}(\gamma(s), Q_M))$ is decreasing on $[S_m, s_m^*)$, increasing on $(s_m^*, S'_m]$, and $\text{dist}(\gamma(s), Q_M) \geq (\nu/2)|s - s_m^*|$ for all $s \in [S_m, S'_m]$.

Now we are able to define the “unperturbed transition chain”: for some small constant $\rho > 0$ which will be specified later we choose $k \in \mathbb{N}$ and $k + 1$ “intermediate frequencies”

$$\omega_I =: \bar{\omega}_0, \bar{\omega}_1, \dots, \bar{\omega}_{k-1}, \bar{\omega}_k := \omega_F$$

with $\bar{\omega}_i := \gamma(s_i)$ for certain $0 =: s_0 < s_1 < \dots < s_{k-1} < s_k := L$ verifying

$$\frac{\rho\mu}{2} \leq s_{i+1} - s_i \leq \rho\mu, \quad \forall i = 0, \dots, k-1. \quad (14.6)$$

By (14.6) there results that

$$\frac{L}{\rho\mu} \leq k \leq \frac{2L}{\rho\mu}, \quad (14.7)$$

moreover it follows from (a) that

$$|\bar{\omega}_{i+1} - \bar{\omega}_i| \leq \rho\mu, \quad \forall i = 0, \dots, k-1. \quad (14.8)$$

This condition has been used before in lemma 12.4. Given k time instants $\bar{\theta}_1 := \theta_0^{\bar{\omega}_1} < \bar{\theta}_2 < \dots < \bar{\theta}_i < \dots < \bar{\theta}_k$, we define the $\{\bar{\varphi}_i\}_{i=1, \dots, k}$ by the iteration formula

$$\bar{\varphi}_1 = \varphi_0^{\bar{\omega}_1}, \quad \bar{\varphi}_{i+1} = \bar{\varphi}_i + \bar{\omega}_i(\bar{\theta}_{i+1} - \bar{\theta}_i). \quad (14.9)$$

The choice of the instants $\{\bar{\theta}_i\}_{i=1, \dots, k}$ is specified in the next lemma: the main request is that $(\bar{\theta}_i, \bar{\varphi}_i)$ must arrive δ -close mod $2\pi\mathbb{Z}^{d+1}$, to the local minimum point $(\theta_0^{\bar{\omega}_i}, \varphi_0^{\bar{\omega}_i})$ of the Poincaré-Melnikov primitive $\Gamma(\bar{\omega}_i, \cdot, \cdot)$, see (14.11)-(14.12). From (14.3) we derive that if $\bar{\omega}_i$ is $1/|\ln \mu|$ far from the set Q_M of “non-ergodizing frequencies” we can reach this goal for “short” time intervals $\bar{\theta}_{i+1} - \bar{\theta}_i \approx |\ln \mu|$. In order to cross the set Q_M of “non-ergodizing frequencies” we need to use longer time intervals $\bar{\theta}_{i+1} - \bar{\theta}_i \approx 1/\text{dist}(Q_M, \bar{\omega}_i)$ if $\sqrt{\mu}/|\ln \mu| < \text{dist}(Q_M, \bar{\omega}_i) < 1/|\ln \mu|$. When the $\bar{\omega}_i$ are “close” (less than $\sqrt{\mu}/|\ln \mu|$ -distant) to the set of non-ergodizing hyperplanes Q_M we choose again $\bar{\theta}_{i+1} - \bar{\theta}_i \approx |\ln \mu|$. We also estimate in (14.13) the total time $\bar{\theta}_k - \bar{\theta}_1 = \sum_{i=1}^k \bar{\theta}_{i+1} - \bar{\theta}_i$.

Lemma 14.2 $\forall \delta > 0$ there exists $\mu_6 > 0$ such that $\forall 0 < \mu \leq \mu_6$ there exist $\{\bar{\theta}_i\}_{i=1, \dots, k}$ with $\bar{\theta}_1 = \theta_0^{\bar{\omega}_1}$ satisfying,

- (i) if $\text{dist}(\bar{\omega}_i, Q_M) > \frac{\sqrt{\mu}}{|\ln \mu|}$ then

$$\max \left\{ C_1 |\ln \mu|, \frac{2\pi}{\text{dist}(\bar{\omega}_i, Q_M)} \right\} < \bar{\theta}_{i+1} - \bar{\theta}_i < 2 \max \left\{ C_1 |\ln \mu|, \frac{2\pi}{\text{dist}(\bar{\omega}_i, Q_M)} \right\}, \quad (14.10)$$

where $M = 8\pi a_{d+1}/\delta$;

- (ii) if $\text{dist}(\bar{\omega}_i, Q_M) \leq \frac{\sqrt{\mu}}{|\ln \mu|}$ then $C_1 |\ln \mu| < \bar{\theta}_{i+1} - \bar{\theta}_i < 2C_1 |\ln \mu|$,

and such that

$$\text{dist}((\bar{\theta}_i, \bar{\varphi}_i), (\theta_0^{\bar{\omega}_i}, \varphi_0^{\bar{\omega}_i}) + 2\pi\mathbb{Z}^{d+1}) < \delta, \quad \forall i = 1, \dots, k, \quad (14.11)$$

where $\bar{\varphi}_1, \dots, \bar{\varphi}_k$ are defined by (14.9). Equivalently, $\forall i = 1, \dots, k$, there exist $h_i \in \mathbb{Z}^{d+1}$ and $\chi_i \in \mathbb{R}^{d+1}$ such that

$$(\bar{\theta}_i, \bar{\varphi}_i) = (\theta_0^{\bar{\omega}_i}, \varphi_0^{\bar{\omega}_i}) + 2\pi h_i + \chi_i \quad \text{with} \quad |\chi_i| < \delta. \quad (14.12)$$

Moreover there exists a constant $K(\gamma)$ such that

$$\bar{\theta}_k - \bar{\theta}_1 \leq K(\gamma) \frac{|\ln \mu|}{\rho\mu}. \quad (14.13)$$

PROOF. Let $\mu_6 > 0$ be so small that $\sqrt{\mu_6}/|\ln \mu_6| < d_0$ and $\sqrt{|\ln \mu_6|} \geq 32\sqrt{C_1}/(\nu\sqrt{\delta\rho})$.

Let us define $(\bar{\theta}_1, \bar{\varphi}_1) := (\theta_0^{\bar{\omega}_1}, \varphi_0^{\bar{\omega}_1})$. Assume that $(\bar{\theta}_1, \dots, \bar{\theta}_i)$ has been defined. If $\text{dist}(\bar{\omega}_i, Q_M) > \sqrt{\mu}/|\ln \mu|$ then by (14.3) there certainly exists $(\bar{\theta}_{i+1}, \bar{\varphi}_{i+1})$ satisfying (14.9), (14.10), such that

$$\text{dist}\left((\bar{\theta}_{i+1}, \bar{\varphi}_{i+1}), (\theta_0^{\bar{\omega}_{i+1}}, \varphi_0^{\bar{\omega}_{i+1}}) + 2\pi\mathbb{Z}^{d+1}\right) < \delta/4.$$

We now consider the case in which $\bar{\omega}_i$ is close to some “non-ergodizing” hyperplanes of Q_M . If $\text{dist}(\bar{\omega}_{i-1}, Q_M) > \sqrt{\mu}/|\ln \mu|$ and $\text{dist}(\bar{\omega}_i, Q_M) \leq \sqrt{\mu}/|\ln \mu|$ we proceed as follows. We have $\bar{\omega}_i = \gamma(s_i)$, with $s_i \in [S_q, S'_q]$ for some q , $1 \leq q \leq r$. Moreover, by property (*) there exists $p^* \in \mathbb{N}$ such that $\{j \in \{1, \dots, k\} \mid s_j \in [S_q, S'_q] \text{ and } \text{dist}(\bar{\omega}_j, Q_M) \leq \sqrt{\mu}/|\ln \mu|\} = \{i, \dots, i+p^*-1\}$, and $s_i \leq s_q^* \leq s_{i+p^*-1}$. We shall use the abbreviations s^* for s_q^* , and ω^* for ω_q^* . We claim that

$$1 \leq p^* \leq p := \left\lfloor \frac{\sqrt{\delta}}{4\sqrt{C_1\rho\mu}|\ln \mu|} \right\rfloor. \quad (14.14)$$

In fact, by (14.6) and (*)

$$\frac{\nu\rho}{4}\mu(p^*-1) \leq \frac{\nu}{2}[(s_{i+p^*-1} - s^*) + (s^* - s_i)] \leq \text{dist}(\bar{\omega}_{i+p^*-1}, Q_M) + \text{dist}(\bar{\omega}_i, Q_M) \leq 2\frac{\sqrt{\mu}}{|\ln \mu|}$$

Hence $p^* \leq 8(\nu\rho\sqrt{\mu}|\ln \mu|)^{-1}$, which implies (14.14), by the choice of μ_6 .

Now we can define the $\bar{\theta}_{i+1}, \dots, \bar{\theta}_{i+p^*}$. The flow of $(\omega^*, 1)$, as any linear flow on a torus, has the following property : there exists $T^*(\omega^*, \delta) > 0$ (abbreviated as T^*) such that any time interval of length T^* contains t satisfying $\text{dist}((t\omega^*, t), 2\pi\mathbb{Z}^{d+1}) \leq \delta/4$.

Therefore (provided $C_1|\ln \mu_6| > T^*$) we can define $\bar{\theta}_{i+1}, \dots, \bar{\theta}_{i+p^*}$ such that

$$C_1|\ln \mu| \leq \bar{\theta}_{i+j+1} - \bar{\theta}_{i+j} \leq 2C_1|\ln \mu|, \quad \text{dist}\left((\bar{\theta}_{i+j}, \bar{\varphi}_{i+j}), (\bar{\theta}_i, \bar{\varphi}_i) + 2\pi\mathbb{Z}^{d+1}\right) \leq \delta/4, \quad (14.15)$$

where $\bar{\varphi}_{i+j} = \bar{\varphi}_i + \omega^*(\bar{\theta}_{i+j} - \bar{\theta}_i)$. For $1 \leq j \leq p^*$, let

$$\bar{\varphi}_{i+j} = \bar{\varphi}_i + \sum_{q=1}^j \bar{\omega}_{i+q-1}(\bar{\theta}_{i+q} - \bar{\theta}_{i+q-1}). \quad (14.16)$$

We now check that for all $j = 1, \dots, p^*$, $(\bar{\theta}_{i+j}, \bar{\varphi}_{i+j})$, as defined in (14.15) and (14.16), satisfy estimate (14.11), namely

$$\text{dist}_T\left((\bar{\theta}_{i+j}, \bar{\varphi}_{i+j}), (\theta_0^{\bar{\omega}_{i+j}}, \varphi_0^{\bar{\omega}_{i+j}})\right) := \text{dist}\left((\bar{\theta}_{i+j}, \bar{\varphi}_{i+j}), (\theta_0^{\bar{\omega}_{i+j}}, \varphi_0^{\bar{\omega}_{i+j}}) + 2\pi\mathbb{Z}^{d+1}\right) \leq \delta. \quad (14.17)$$

We have by (14.16) that

$$\text{dist}_T\left((\bar{\theta}_{i+j}, \bar{\varphi}_{i+j}), (\bar{\theta}_i, \bar{\varphi}_i)\right) \leq \text{dist}_T\left((\bar{\theta}_{i+j}, \bar{\varphi}_{i+j}), (\bar{\theta}_i, \bar{\varphi}_i)\right)$$

$$\begin{aligned}
& + \left| \sum_{q=1}^j (\bar{\omega}_{i+q-1} - \omega^*) (\bar{\theta}_{i+q} - \bar{\theta}_{i+q-1}) \right| \\
& \leq \delta/4 + 2C_1 |\ln \mu| \sum_{q=1}^{p^*} |s_{i+q-1} - s^*| \quad (\text{by (14.15) and (a)}) \\
& \leq \delta/4 + 2C_1 |\ln \mu| p^* (s_{i+p^*-1} - s_i) \\
& \leq \delta/4 + 2C_1 |\ln \mu| p^2 \rho \mu \leq 3\delta/8,
\end{aligned}$$

by (14.6) and (14.14). Therefore, by (14.5),

$$\begin{aligned}
\text{dist}_T((\bar{\theta}_{i+j}, \bar{\varphi}_{i+j}), (\theta_0^{\bar{\omega}_{i+j}}, \varphi_0^{\bar{\omega}_{i+j}})) & \leq \frac{3\delta}{8} + \text{dist}_T((\bar{\theta}_i, \bar{\varphi}_i), (\theta_0^{\bar{\omega}_i}, \varphi_0^{\bar{\omega}_i})) + K|\bar{\omega}_{i+j} - \bar{\omega}_i| \\
& \leq \frac{3\delta}{8} + \frac{\delta}{4} + K\rho\mu p < \delta
\end{aligned}$$

by (14.14), provided μ_6 has been chosen small enough.

There remains to prove (14.13). By (*) we can write

$$A_m := \left\{ s \in [S_m, S'_m] \mid \frac{\sqrt{\mu}}{|\ln \mu|} \leq \text{dist}(\gamma(s), Q_M) \leq \frac{1}{2C_1 |\ln \mu|} \right\} = [U_m, V_m] \cup [V'_m, U'_m],$$

with $S_m < U_m < V_m < s_m^* < V'_m < U'_m < S'_m$ (in the case when $\omega^* = \omega_{I,F}$, A_m is just an interval). Moreover, by (a), $s_m^* - V_m, V'_m - s_m^* \geq \sqrt{\mu}/|\ln \mu|$. Define $A := \cup_{m=1}^r A_m$. We have $\bar{\theta}_k - \bar{\theta}_1 = \sigma_0 + \sum_{m=1}^r \sigma_m$, where

$$\sigma_0 := \sum_{1 \leq i \leq k-1, s_i \notin A} (\bar{\theta}_{i+1} - \bar{\theta}_i), \quad \sigma_m := \sum_{1 \leq i \leq k-1, s_i \in A_m} (\bar{\theta}_{i+1} - \bar{\theta}_i).$$

For $s_i \notin A$, $\bar{\theta}_{i+1} - \bar{\theta}_i \leq 2C_1 |\ln \mu|$, hence $\sigma_0 \leq 2C_1 k |\ln \mu| \leq 4C_1 L \ln \mu / (\rho\mu)$. For $i \in A_m$, $\bar{\theta}_{i+1} - \bar{\theta}_i \leq 4\pi (\text{dist}(\bar{\omega}_i, Q_M))^{-1} \leq 8\pi / (\nu |s_i - s_m^*|)$ by (*), and hence, using that by (14.6) $s_{i+1} \geq s_i + \rho\mu/2$,

$$\sigma_m \leq \frac{8\pi}{\nu} \sum_{1 \leq i \leq k-1, s_i \in A_m} \frac{1}{|s_i - s_m^*|} \leq \frac{16\pi}{\nu\rho\mu} \sum_{1 \leq i \leq k-1, s_i \in A_m} \frac{s_{i+1} - s_i}{|s_i - s_m^*|}.$$

Estimating the above sum with an integral we easily get

$$\sigma_m \leq \frac{8\pi}{\nu(s_m^* - V_m)} + \frac{16\pi}{\nu\rho\mu} \int_{U_m}^{V_m} \frac{ds}{s_m^* - s} + \frac{8\pi}{\nu(V'_m - s_m^*)} + \frac{16\pi}{\nu\rho\mu} \int_{V'_m}^{U'_m} \frac{ds}{s - s_m^*}.$$

(14.13) can be easily deduced by the bound on $s_m^* - V_m, V'_m - s_m^*$. ■

In the next section we will prove the existence of a diffusion orbit (φ_μ, q_μ) close to the “unperturbed pseudo-diffusion orbit” $(\bar{\varphi}(t), \bar{q}(t)) : (\bar{\theta}_1, \bar{\theta}_k) \rightarrow \mathbb{R}^{d+1}$ defined, for $t \in [\bar{\theta}_i, \bar{\theta}_{i+1}]$, as $\bar{\varphi}(t) := \bar{\varphi}_i + \bar{\omega}_i(t - \bar{\theta}_i)$ and $\bar{q}_{[\bar{\theta}_i, \bar{\theta}_{i+1}]} := Q_{\bar{\theta}_{i+1} - \bar{\theta}_i}(\cdot - \bar{\theta}_i) \pmod{2\pi}$.

15 The diffusion orbit

We need the following property of the Melnikov function $\tilde{\Gamma}(\omega, \cdot, \cdot)$ defined w.r.t. to the variables (b, c) by

$$\tilde{\Gamma}(\omega, b, c) := \Gamma(\omega, \theta_0^\omega + b, \varphi_0^\omega + b\omega + c).$$

Lemma 15.1 *Assume that $\Gamma(\omega, \cdot, \cdot)$ possesses a non-degenerate local minimum in the point $(\theta_0^\omega, \varphi_0^\omega)$. Then there exist $r > 0$, $\bar{b} > 0$, $\nu_j > 0$ ($j = 1, 2$) depending only on γ such that $\forall \omega = \gamma(s)$, $s \in [0, L]$*

- (i) $\partial_c \tilde{\Gamma}(\omega, b, c) \cdot c \geq \nu_2 > 0$ or $|\partial_b \tilde{\Gamma}(\omega, b, c)| \geq \nu_1 > 0$ for $|c| = r$, $|b| \leq \bar{b}$,
- (ii) $\partial_b \tilde{\Gamma}(\omega, b, c) \times \text{sign}(b) \geq \nu_1 > 0$ for $|c| \leq r$ and $b = \pm \bar{b}$.

PROOF. We can assume that (14.4) is satisfied. Since $\Gamma(\omega, \cdot, \cdot)$ possesses a non-degenerate minimum in $(\theta_0^\omega, \varphi_0^\omega)$, $\tilde{\Gamma}(\omega, b, c)$ possesses in $(0, 0)$ a non degenerate minimum. Hence we write $\tilde{\Gamma}(\omega, b, c)$, up to a constant, as $\tilde{\Gamma}(\omega, b, c) = Q_2(b, c) + Q_3(b, c)$ where $Q_2(b, c) =: \beta_\omega b^2/2 + (\alpha_\omega \cdot c)b + (\gamma_\omega c \cdot c)/2$ is a positive definite quadratic form ($\beta_\omega \in \mathbb{R}$, $\alpha_\omega \in \mathbb{R}^d$, $\gamma_\omega \in \text{Mat}(d \times d)$) and $Q_3 = O(|b|^3 + |c|^3)$. More precisely, by (14.4), there exists $\varepsilon > 0$ such that $\beta_\omega > \varepsilon$, and $d_\omega(c) := \beta_\omega(\gamma_\omega c \cdot c) - (\alpha_\omega \cdot c)^2 > \varepsilon|c|^2$ for all $\omega \in \gamma([0, L])$. In addition, by the smoothness of Γ and the fact that $\omega = \gamma(s)$ lives in a compact subset of \mathbb{R}^d , there exists a constant M such that, $\forall \omega \in \gamma([0, L])$, $|\alpha_\omega| + |\beta_\omega| + |\gamma_\omega| \leq M$, $|\nabla Q_3(b, c)| \leq M(b^2 + |c|^2)$.

We have $\partial_b Q_2(b, c) = \beta_\omega b + \alpha_\omega \cdot c$ and $\partial_c Q_2(b, c) \cdot c = b\alpha_\omega \cdot c + (\gamma_\omega c \cdot c)$.

Let us define $\bar{\nu}_1 := \inf_{\omega \in \gamma([0, L])} \varepsilon/(4|\alpha_\omega|) > 0$ and $\bar{\nu}_2 := \inf_{\omega \in \gamma([0, L])} \varepsilon/(4\beta_\omega) > 0$. Then consider $\nu_1 := \bar{\nu}_1 r$, $\nu_2 = \bar{\nu}_2 r^2$ and $\bar{b} := r \sup_{\omega \in \gamma([0, L])} (3\bar{\nu}_1 + |\alpha_\omega|)/\beta_\omega$, $r \in (0, 1]$. We now prove that, provided $r > 0$ has been chosen sufficiently small, conditions (i) and (ii) are satisfied with the above choice of the constants. Indeed if $(|\alpha_\omega \cdot c| + 2\bar{\nu}_1 r)/\beta_\omega \leq |b| \leq \bar{b}$ and $|c| \leq r$ then $\partial_b \tilde{\Gamma}(\omega, b, c) \cdot \text{sign}(b) \geq \beta_\omega |b| - |\alpha_\omega \cdot c| - |\partial_b Q_3(b, c)| \geq 2\bar{\nu}_1 r - O(r^2) \geq \nu_1$ for r sufficiently small. In particular this proves (ii). On the other hand if $|b| < (|\alpha_\omega \cdot c| + 2\bar{\nu}_1 r)/\beta_\omega$ and $|c| = r$ then

$$\begin{aligned} \partial_c \tilde{\Gamma}(\omega, b, c) \cdot c &= b(\alpha_\omega \cdot c) + (\gamma_\omega c \cdot c) + \partial_c Q_3(b, c) \cdot c \geq (\gamma_\omega c \cdot c) - |b(\alpha_\omega \cdot c)| + O(r^3) \\ &\geq \frac{\varepsilon r^2 + (\alpha_\omega \cdot c)^2 - |\alpha_\omega \cdot c|(|\alpha_\omega \cdot c| + 2\bar{\nu}_1 r)}{\beta_\omega} + O(r^3) \\ &\geq \frac{\varepsilon - 2\bar{\nu}_1 |\alpha_\omega|}{\beta_\omega} r^2 + O(r^3) \geq \frac{\varepsilon}{2\beta_\omega} r^2 - O(r^3) \geq 2\bar{\nu}_2 r^2 + O(r^3). \end{aligned}$$

Hence (i) is satisfied for r small enough. ■

The partial derivatives of $\tilde{\Gamma}$ are Lipschitz-continuous w.r.t. (b, c) uniformly in $\omega \in \gamma([0, L])$. Therefore, by lemma 15.1, there exists $\delta > 0$ such that, $\forall \eta \in \mathbb{R}$ with $|\eta| \leq \delta$, $\forall \xi \in \mathbb{R}^d$ with $|\xi| \leq \delta$, $\forall \omega \in \gamma([0, L])$,

$$\begin{aligned} \partial_c \tilde{\Gamma}(\omega, b + \eta, c + \xi) \cdot c &\geq 3\nu_2/4 > 0 \quad \text{or} \quad |\partial_b \tilde{\Gamma}(\omega, b + \eta, c + \xi)| \geq 3\nu_1/4 > 0 \quad (15.1) \\ \text{for } |c| = r, |b| &\leq \bar{b}, \end{aligned}$$

$$\partial_b \tilde{\Gamma}(\omega, b + \eta, c + \xi) \times \text{sign}(b) \geq 3\nu_1/4 > 0 \quad \text{for } |c| \leq r \text{ and } b = \pm \bar{b}. \quad (15.2)$$

Moreover let us fix $\rho > 0$ such that

$$\rho \leq \min\{\nu_1/2, \nu_2/r\}/(6C_2), \quad (15.3)$$

where C_2 appears in (12.24). These are the positive constants (δ, ρ) that we use in order to define, for $0 < \mu < \mu_6$, $\bar{w}_i, \bar{\theta}_i, \bar{\varphi}_i$ by lemma 14.2.

Since $\gamma([0, L])$ is a compact subset of \mathcal{D}_N^c , $\inf_{s \in [0, L]} \beta(\gamma(s)) > 0$ and, by the choice of $\bar{\theta}_i$, for μ small enough (12.21) is satisfied. Therefore, by lemma 12.5 and (14.12), there exists $\mu_7 > 0$ such that, $\forall 0 < \mu \leq \mu_7$,

$$\mathcal{F}_\mu(b, c) = \frac{1}{2} \sum_{i=1}^{k-1} \frac{|c_{i+1} - c_i|^2}{\Delta \bar{\theta}_i + (b_{i+1} - b_i)} + \mu \sum_{i=1}^k \tilde{\Gamma}(\bar{w}_i, \eta_i + b_i, \xi_i + c_i) + R_7, \quad (15.4)$$

where $|\eta_i| \leq \delta$, $|\xi_i| \leq \delta$, R_7 is given by (12.23) and satisfies (12.24).

We minimize the functional \mathcal{F}_μ on the closure of

$$W := \{(b, c) := (b_1, c_1, \dots, b_k, c_k) \in \mathbb{R}^{(d+1)k} \mid |b_i| < \bar{b}, |c_i| < r, \forall i = 1, \dots, k\}.$$

Since \bar{W} is compact, \mathcal{F}_μ attains its minimum in \bar{W} , say at (\tilde{b}, \tilde{c}) . By lemma 11.3 the existence of the diffusion orbit will be proved once we show that $(\tilde{b}, \tilde{c}) \in W$, see lemma 15.3. Let us define for $i = 1, \dots, k-1$

$$w_i := w_i(b, c) := \frac{c_{i+1} - c_i}{\theta_{i+1} - \theta_i} = \frac{c_{i+1} - c_i}{\Delta \bar{\theta}_i + (b_{i+1} - b_i)},$$

and $w_0 = w_k = 0$. From (14.9) and (12.14), w_i can be written as

$$w_i = \frac{\varphi_{i+1} - \varphi_i}{(\theta_{i+1} - \theta_i)} - \bar{w}_i - \frac{\Delta \bar{w}_i b_{i+1}}{(\theta_{i+1} - \theta_i)} = (\omega_i - \bar{w}_i) + O\left(\frac{\mu}{|\ln \mu|}\right). \quad (15.5)$$

By the expression of \mathcal{F}_μ in (15.4) we have, for all $i = 1, \dots, k$,

$$\partial_{c_i} \mathcal{F}_\mu(b, c) = w_{i-1} - w_i + \mu \partial_c \tilde{\Gamma}(\bar{w}_i, \eta_i + b_i, \xi_i + c_i) + R_i \quad (15.6)$$

$$\partial_{b_i} \mathcal{F}_\mu(b, c) = \frac{1}{2} (|w_i|^2 - |w_{i-1}|^2) + \mu \partial_b \tilde{\Gamma}(\bar{w}_i, \eta_i + b_i, \xi_i + c_i) + S_i \quad (15.7)$$

where $R_i := \partial_{c_i} R_7$, $S_i := \partial_{b_i} R_7$ satisfy, by (12.24) and (15.3)

$$|R_i|, |S_i| \leq \frac{\mu}{2} \min\left\{\frac{\nu_1}{2}, \frac{\nu_2}{r}\right\}. \quad (15.8)$$

By (15.6)-(15.7), a way to see critical points of \mathcal{F}_μ is to show that the terms $w_{i-1} - w_i$ and $|w_i|^2 - |w_{i-1}|^2$ are small w.r.t the $O(\mu)$ -contribution provided by the Melnikov function. By (12.8) $|\omega_i - \bar{w}_i| = O(1/(\theta_{i+1} - \theta_i))$ and hence, using (15.5), an estimate for each w_i separately is given by $w_i = O(1/|\bar{\theta}_{i+1} - \bar{\theta}_i|) + O(\mu/|\ln \mu|)$. Hence each $|w_i|$ is $O(\mu)$ -small if the time to make a transition $|\bar{\theta}_{i+1} - \bar{\theta}_i| = O(1/\mu)$, as in [26]. These time intervals are too large to obtain the approximation for the reduced action functional \mathcal{F}_μ given in lemma 12.5 and (15.4). Therefore we need more refined estimates: the proof of Theorem 10.1 (and Theorem 10.2) relies on the following crucial property for $\tilde{w}_i := w_i(\tilde{b}, \tilde{c})$, satisfied by the minimum point (\tilde{b}, \tilde{c}) .

Lemma 15.2 *We have (for $i = 1, \dots, k$),*

$$i) \quad |\tilde{w}_i - \tilde{w}_{i-1}| = O(\mu), \quad ii) \quad |\tilde{w}_i| = O\left(\frac{\sqrt{\mu}}{\sqrt{|\ln \mu|}}\right). \quad (15.9)$$

PROOF. Estimate (15.9) – i) is a straightforward consequence of (15.6) and (15.8) if $|\tilde{c}_i| < r$, since in this case $\partial_{c_i} \mathcal{F}_\mu(\tilde{b}, \tilde{c}) = 0$. We now prove that (15.9)i) holds also if $|\tilde{c}_i| = r$ for some i . Indeed if $|\tilde{c}_i| = r$ then

$$\partial_{c_i} \mathcal{F}_\mu(\tilde{b}, \tilde{c}) = \alpha_\mu \tilde{c}_i \quad \text{for some} \quad \alpha_\mu \leq 0 \quad (15.10)$$

(since (\tilde{b}, \tilde{c}) is a minimum point) and then by (15.6), (15.10) and (15.8) we deduce

$$\tilde{w}_{i-1} - \tilde{w}_i = \alpha_\mu \tilde{c}_i + O(\mu). \quad (15.11)$$

Let us decompose \tilde{w}_{i-1} and \tilde{w}_i in the “radial” and “tangent” directions to the ball $S_i = \{|b_i| \leq \bar{b}, |c_i| \leq r\}$:

$$\tilde{w}_{i-1} = a_i \tilde{c}_i + u_i \quad \text{with} \quad u_i \cdot \tilde{c}_i = 0 \quad (15.12)$$

$$-\tilde{w}_i = a'_i \tilde{c}_i + u'_i, \quad \text{with} \quad u'_i \cdot \tilde{c}_i = 0. \quad (15.13)$$

Since $|\tilde{c}_{i-1}| \leq |\tilde{c}_i| = r$, $|\tilde{c}_{i+1}| \leq |\tilde{c}_i| = r$, there results that

$$a_i r^2 = \tilde{w}_{i-1} \cdot \tilde{c}_i \geq 0 \quad \text{and} \quad a'_i r^2 = -\tilde{w}_i \cdot \tilde{c}_i \geq 0, \quad (15.14)$$

so that $a_i, a'_i \geq 0$. Summing (15.12) and (15.13) and using (15.11) we obtain

$$(a_i + a'_i) \tilde{c}_i + (u_i + u'_i) = O(\mu) + \alpha_\mu \tilde{c}_i,$$

with $a_i, a'_i, -\alpha_\mu \geq 0$. This implies that $\alpha_\mu = O(\mu/r)$ and from equation (15.11) we get (15.9)i).

We can now prove (15.9) – ii). Let $i_0 \in \{1, \dots, k-1\}$ be such that $\forall 1 \leq i \leq k-1$, $|\tilde{w}_{i_0}| \geq |\tilde{w}_i|$. For $j \in \{1, \dots, k-1\}$, $j \neq i_0$ we can write $\tilde{w}_j = \tilde{w}_{i_0} + s_j$ with $s_j = \sum_{i=i_0}^{j-1} (\tilde{w}_{i+1} - \tilde{w}_i)$ and hence, by (15.9)i)

$$|s_j| \leq \sum_{i=i_0}^{j-1} |\tilde{w}_{i+1} - \tilde{w}_i| \leq C\mu |j - i_0| \quad (15.15)$$

for some constant $C > 0$. Hence

$$\tilde{c}_j - \tilde{c}_{i_0} = \sum_{i=i_0}^{j-1} \tilde{w}_i (\tilde{\theta}_{i+1} - \tilde{\theta}_i) = \tilde{w}_{i_0} (\tilde{\theta}_j - \tilde{\theta}_{i_0}) + \sum_{i=i_0}^{j-1} s_i (\tilde{\theta}_{i+1} - \tilde{\theta}_i) \quad (15.16)$$

and then by (15.15)

$$|\tilde{c}_j - \tilde{c}_{i_0}| \geq |\tilde{w}_{i_0}| |\tilde{\theta}_j - \tilde{\theta}_{i_0}| - C\mu |j - i_0| |\tilde{\theta}_j - \tilde{\theta}_{i_0}| = (|\tilde{w}_{i_0}| - C\mu |j - i_0|) |\tilde{\theta}_j - \tilde{\theta}_{i_0}|. \quad (15.17)$$

Since $|\tilde{\theta}_{i+1} - \tilde{\theta}_i| > C_1 |\ln \mu| + O(1)$ (by (12.4)), $\forall i = 1, \dots, k-1$, $|\tilde{\theta}_j - \tilde{\theta}_{i_0}| > C_1 |j - i_0| \cdot |\ln \mu|$. Take $\bar{j} \in \{1, \dots, k-1\}$ such that $|\bar{j} - i_0| = [(\sqrt{\mu} \sqrt{|\ln \mu|})^{-1}] + 1$ (such a \bar{j} certainly exists since, by (14.7), $k \approx 1/\mu$ for μ small). Then we obtain, using that $|\tilde{c}_i| \leq r$ for all $i = 1, \dots, k$,

$$2r \geq |\tilde{c}_{\bar{j}} - \tilde{c}_{i_0}| \geq (|\tilde{w}_{i_0}| - C \frac{\sqrt{\mu}}{\sqrt{|\ln \mu|}} - C\mu) C_1 \frac{\sqrt{|\ln \mu|}}{\sqrt{\mu}},$$

i.e. $|\tilde{w}_{i_0}| \leq \frac{(2r + CC_1)\sqrt{\mu}}{C_1 \sqrt{|\ln \mu|}} + C\mu$. We have thus proved the important property (15.9)–ii). \blacksquare

Remark 15.1 By (15.5), $(\tilde{w}_i - \bar{w}_i) = \tilde{w}_i + O(\mu/|\ln \mu|)$, so that, by (14.8), (15.9) implies

$$|\tilde{w}_i - \bar{w}_i| = O\left(\frac{\sqrt{\mu}}{\sqrt{|\ln \mu|}}\right), \quad |\tilde{w}_{i+1} - \tilde{w}_i| = O(\mu). \quad (15.18)$$

Note that, from (12.8), we would just obtain $|\tilde{w}_i - \bar{w}_i| = O(1/|\ln \mu|)$. (15.18) can be seen as an a-priori estimate satisfied by the minimum point $(\tilde{\theta}, \tilde{\varphi})$.

The following lemma proves the existence of a local minimum of the reduced action functional in the interior of W and hence of a true diffusion orbit.

Lemma 15.3 Let (\tilde{b}, \tilde{c}) be a minimum point of \mathcal{F}_μ over \bar{W} . Then $(\tilde{b}, \tilde{c}) \in W$, namely

$$|\tilde{c}_i| < r \quad \text{for all } i \in \{1, \dots, k\} \quad (15.19)$$

and

$$|\tilde{b}_i| < \bar{b} \quad \text{for all } i \in \{1, \dots, k\}. \quad (15.20)$$

PROOF. By (15.9) we have $||\tilde{w}_{i+1}|^2 - |\tilde{w}_i|^2| \leq |\tilde{w}_{i+1} - \tilde{w}_i| \cdot (|\tilde{w}_{i+1}| + |\tilde{w}_i|) = O(\mu^{3/2})$, and hence, from (15.7) we derive

$$\partial_{b_i} \mathcal{F}_\mu(\tilde{b}, \tilde{c}) = \mu \partial_b \tilde{\Gamma}(\bar{w}_i, \eta_i + \tilde{b}_i, \xi_i + \tilde{c}_i) + O(\mu^{3/2}) + S_i. \quad (15.21)$$

Let us first assume by contradiction that $\exists i$ such that $|\tilde{c}_i| = r$ and $|\tilde{b}_i| < \bar{b}$. In this case we claim that

$$\partial_c \tilde{\Gamma}(\bar{w}_i, \eta_i + \tilde{b}_i, \xi_i + \tilde{c}_i) \cdot \tilde{c}_i \leq \nu_2/2 \quad \text{and} \quad |\partial_b \tilde{\Gamma}(\bar{w}_i, \eta_i + \tilde{b}_i, \xi_i + \tilde{c}_i)| \leq \nu_1/2 \quad (15.22)$$

contradicting (15.1), since $|\eta_i|, |\xi_i| \leq \delta$. Let us prove (15.22). Since (\tilde{b}, \tilde{c}) is a minimum point

$$\partial_{c_i} \mathcal{F}_\mu(\tilde{b}, \tilde{c}) \cdot \tilde{c}_i = (\tilde{w}_{i-1} - \tilde{w}_i) \cdot \tilde{c}_i + \mu \partial_c \tilde{\Gamma}(\bar{w}_i, \eta_i + \tilde{b}_i, \xi_i + \tilde{c}_i) \cdot \tilde{c}_i + R_i \cdot \tilde{c}_i = \alpha_\mu \tilde{c}_i \cdot \tilde{c}_i = \alpha_\mu r^2 \leq 0.$$

By (15.14) and (15.8) it follows that $\partial_c \tilde{\Gamma}(\bar{\omega}_i, \eta_i + \tilde{b}_i, \xi_i + \tilde{c}_i) \cdot \tilde{c}_i \leq \nu_2/2$. Moreover since $|\tilde{b}_i| < \bar{b}$ we have $\partial_{b_i} \mathcal{F}_\mu(\tilde{b}, \tilde{c}) = 0$, and by (15.21), (15.8) it follows that $|\partial_b \tilde{\Gamma}(\bar{\omega}_i, \eta_i + \tilde{b}_i, \xi_i + \tilde{c}_i)| \leq \nu_1/2$ (provided μ is small enough). Estimate (15.22) is then proved. As a result, if (15.20) holds, so does (15.19).

Let us finally prove (15.20). If by contradiction $\exists i$ with $|\tilde{b}_i| = \bar{b}$, by (15.21), (15.8) and since (\tilde{b}, \tilde{c}) is a minimum point, arguing as before, we deduce that $\partial_b \tilde{\Gamma}(\bar{\omega}_i, \eta_i + \tilde{b}_i, \xi_i + \tilde{c}_i) \text{sign}(\tilde{b}_i) \leq \nu_1/2$. This contradicts (15.2) since $|\eta_i|, |\xi_i| \leq \delta$. The lemma is proved. ■

PROOF OF THEOREM 10.1. Lemmas 15.3 and 11.3 imply the existence of a diffusion orbit $z_\mu(t) := (\varphi_\mu(t), q_\mu(t), I_\mu(t), p_\mu(t))$ with $\dot{\varphi}_\mu(\tilde{\theta}_1) = \omega_I + O(\mu)$ and $\dot{\varphi}_\mu(\tilde{\theta}_k) = \omega_I + O(\mu)$ ($z_\mu(\cdot)$ connects a $O(\mu)$ -neighborhood of \mathcal{T}_{ω_I} to a $O(\mu)$ -neighborhood of \mathcal{T}_{ω_F} in the time-interval (τ_1, τ_2) where $\tau_1 := (\tilde{\theta}_1 + \tilde{\theta}_2)/2$, $\tau_2 := (\tilde{\theta}_{k-1} + \tilde{\theta}_k)/2$). The estimate on the diffusion time is a straightforward consequence of (14.13) and the fact that $\tilde{\theta}_{1,k} = \bar{\theta}_{1,k} + O(1)$. That $\text{dist}(I_\mu(t), \gamma([0, L])) < \eta$ for all t , provided μ is small enough, results from (15.18) and the estimates of lemma 11.1.

Finally we observe that, if the perturbation is $\mu(f + \mu \tilde{f})$, then lemma 11.1 still applies with the same estimates. Moreover in the development of the reduced functional the term containing $\mu^2 \tilde{f}$ gives, in time intervals $\bar{\theta}_{i+1} - \bar{\theta}_i \leq \text{const.} |\ln \mu| / \sqrt{\mu}$, negligible contributions $o(\mu)$. Therefore the same variational proof applies. ■

PROOF OF THEOREM 10.2. If the perturbation is of the form $f(\varphi, q, t) = (1 - \cos q) f(\varphi, t)$, by remark 11.1-2, we can prove that the development (12.22) holds along any path γ of the action space (without any condition as (12.21)). Therefore the previous variational argument applies. ■

For $\beta > 0$ small let \mathcal{D}_N^β be the set of frequencies “ β -non-resonant with the perturbation” $\mathcal{D}_N^\beta := \{\omega \in \mathbb{R}^d \mid |\omega \cdot n + l| > \beta, \forall 0 < |(n, l)| \leq N\}$. If β becomes small with μ our estimate on the diffusion time required to approach to the boundaries of $\mathcal{C} \cap \mathcal{D}_N^\beta$ slightly deteriorates. In the same hypotheses as in Theorem 10.1 we have the following result.

Theorem 15.1 $\forall R > 0, \forall 0 \leq a < 1/4$, there exists $\mu_8 > 0$ such that $\forall 0 < \mu \leq \mu_8, \forall \omega_I, \omega_F \in \mathcal{C} \cap \mathcal{D}_N^{\mu^a} \cap B_R(0)$ there exist a diffusion orbit $(\varphi_\mu(t), q_\mu(t), I_\mu(t), p_\mu(t))$ of (\mathcal{S}_μ) and two instants $\tau_1 < \tau_2$ with $I_\mu(\tau_1) = \omega_I + O(\mu)$, $I_\mu(\tau_2) = \omega_F + O(\mu)$ and

$$|\tau_2 - \tau_1| = O(1/\mu^{1+a}). \quad (15.23)$$

PROOF. For simplicity we consider the case in which $\beta(\omega_I) = O(\mu^a)$ and $\beta(\omega_F) = O(1)$. With respect to Theorem 10.1 we only need to prove the existence of a diffusion orbit connecting ω_I to some fixed ω^* lying in the same connected component of $\mathcal{D}_N^\beta \cap B_R(0)$ containing ω_I . In order to construct an orbit connecting ω_I to ω^* we can define $\bar{\omega}_i := \omega_I + i(\omega^* - \omega_I)/k$, for $0 \leq i \leq k$ and $k := \lceil |\omega^* - \omega_I|/\rho\mu \rceil + 1$. We obtain that $\beta_j = \beta(\bar{\omega}_j) \geq C(\mu^a + j\rho\mu)$ for some $C > 0$ and we choose $\bar{\theta}_{j+1} - \bar{\theta}_j \geq \text{const.} \beta_j^{-2}$ verifying in this way the hypotheses of lemma 12.5. If ω_I belongs to some Q_M the transition times $|\ln \mu|/\sqrt{\mu}$ needed to cross Q_M (see lemma 14.2) still satisfy (12.21). We finally obtain a diffusion time $\bar{\theta}_k - \bar{\theta}_1 = \sum_{j=1}^{k-1} (\bar{\theta}_{j+1} - \bar{\theta}_j) = O(1/\mu^{1+a})$. ■

16 The stability result and the optimal time

In this section we will prove, via classical perturbation theory, stability results for the action variables, implying, in particular, Theorem 10.3. We shall use the following notations: for $l \in \mathbb{N}$, $A \subset \mathbb{C}^l$ and $r > 0$, we define $A_r := \{z \in \mathbb{C}^l \mid \text{dist}(z, A) \leq r\}$ and $\mathbb{T}_s^l := \{z \in \mathbb{C}^l \mid |\text{Im } z_j| < s, \forall 1 \leq j \leq l\}$ (thought of as a complex neighborhood of \mathbb{T}^l). Given two bounded open sets $B \subset \mathbb{C}^2$, $D \subset \mathbb{C}^l$ and $f(I, \varphi, p, q)$, real analytic function with holomorphic extension on $D_\sigma \times \mathbb{T}_{s+\sigma}^l \times B_\sigma$ for some $\sigma > 0$, we define the following norm $\|f\|_{B,D,s} = \sum_{k \in \mathbb{Z}^l} \sup_{\substack{(p,q) \in B \\ I \in D}} |\hat{f}_k(I, p, q)| e^{|k|s}$ where $\hat{f}_k(I, p, q)$ denotes the k -Fourier coefficient of the periodic function $\varphi \rightarrow f(I, \varphi, p, q)$.

Let us consider Hamiltonian \mathcal{H}_μ defined in (10.1) and assume that $f(I, \varphi, p, q, t)$, defined in (10.2), is a real analytic function, possessing, for some $r, \bar{r}, \tilde{r}, s > 0$, complex analytic extension on $\{I \in \mathbb{R}^d \mid |I| \leq \bar{r}\}_r \times \mathbb{T}_s^d \times \{p \in \mathbb{R} \mid |p| \leq \tilde{r}\}_r \times \mathbb{T}_s \times \mathbb{T}_s$.

It is convenient to write Hamiltonian \mathcal{H}_μ in autonomous form. For this purpose let us introduce the new action-angle variables (I_0, φ_0) with $t = \varphi_0$, that will still be denoted by $I := (I_0, I_1, \dots, I_n)$ and $\varphi := (\varphi_0, \varphi_1, \dots, \varphi_n)$. Defining $h(I) := I_0 + |I|^2/2$ and $E := E(p, q) := p^2/2 + (\cos q - 1)$, \mathcal{H}_μ is then equivalent to the autonomous Hamiltonian

$$H := H(I, \varphi, p, q) := h(I) + E(p, q) + \mu f(I, \varphi, p, q). \quad (16.1)$$

Clearly, Hamiltonian H is a real analytic function, with complex analytic extension on

$$\{I \in \mathbb{R}^{d+1} \mid |I| \leq \bar{r}\}_r \times \mathbb{T}_s^{d+1} \times \{p \in \mathbb{R} \mid |p| \leq \tilde{r}\}_r \times \mathbb{T}_s.$$

In the sequel we will denote by $z(t) := (I(t), \varphi(t), p(t), q(t))$ the solution of the Hamilton equations associated to Hamiltonian (16.1) with initial condition $z(0) = (I(0), \varphi(0), p(0), q(0))$.

The proof of the stability of the action variables is divided in two steps:

- (i) **(Stability far from the separatrices of the pendulum:)** prove stability in the region

$$\begin{aligned} \mathcal{E}_1 &:= \mathcal{E}_1^+ \cup \mathcal{E}_1^- := \left\{ (I, \varphi, p, q) \mid E(p, q) \geq \mu^{c_d} \right\} \\ &\cup \left\{ (I, \varphi, p, q) \mid -2 + \mu^{c_d} \leq E(p, q) \leq -\mu^{c_d} \right\} \end{aligned}$$

in which we can apply the Nekhoroshev Theorem obtaining actually stability for exponentially long times,

- (ii) **(Stability close to the separatrices of the pendulum and to the elliptic equilibrium point:)** prove stability in the region

$$\begin{aligned} \mathcal{E}_2 &:= \mathcal{E}_2^+ \cup \mathcal{E}_2^- := \left\{ (I, \varphi, p, q) \mid -2\mu^{c_d} \leq E(p, q) \leq 2\mu^{c_d} \right\} \\ &\cup \left\{ (I, \varphi, p, q) \mid -2 \leq E(p, q) \leq -2 + 2\mu^{c_d} \right\} \end{aligned}$$

in which we use some *ad hoc* arguments,

where $0 < c_d < 1$ is a positive constant that will be chosen later on, see (16.12).

We first prove (i). In the regions⁹⁸ $\tilde{\mathcal{E}}_1^\pm := \Pi_{q,p}\mathcal{E}_1^\pm$ we first write the pendulum Hamiltonian $E(p, q)$ in action-angle variables. In the region⁹⁹ $\tilde{\mathcal{E}}_1^+ \cup \{p > 0\}$ the new action variable P is defined by the formula

$$P := P^+(E) := \frac{\sqrt{2}}{\pi} \int_0^\pi \sqrt{E + (1 + \cos \psi)} d\psi.$$

while in the region $\tilde{\mathcal{E}}_1^-$ the new action variable is

$$P := P^-(E) = \frac{2\sqrt{2}}{\pi} \int_0^{\psi_0(E)} \sqrt{E + (1 + \cos \psi)} d\psi$$

where $\psi_0(E)$ is the first positive number such that $E + (1 + \cos \psi_0(E)) = 0$. We will use the following lemma, proved in [33], regarding the analyticity radii of these action-angle variables close to the separatrices of the pendulum.

Lemma 16.1 *There exist intervals $D^\pm \subset \mathbb{R}$, symplectic transformations $\phi^\pm = \phi^\pm(P, Q)$ real analytic on $D^\pm \times \mathbb{T}$ with holomorphic extension on $D_{r_0}^\pm \times \mathbb{T}_{s_0}$ and functions E^\pm real analytic on D^\pm with holomorphic extension on $D_{r_0}^\pm$ such that $\phi^\pm(D^\pm \times \mathbb{T}) = \tilde{\mathcal{E}}_1^\pm$ and*

$$E(\phi^\pm(P, Q)) = E^\pm(P),$$

with $r_0 = \text{const} \mu^{c_d}$ and $s_0 = \text{const}/|\ln \mu|$. Moreover, for E bounded, the following estimates on the derivatives hold¹⁰⁰

$$\frac{dE^\pm}{dP}(P^\pm(E)) \approx \ln^{-1}\left(1 + \frac{1}{\sqrt{|E|}}\right) \quad (16.2)$$

$$\pm \frac{d^2 E^\pm}{dP^2}(P^\pm(E)) \approx \frac{1}{|E|} \ln^{-3}\left(1 + \frac{1}{\sqrt{|E|}}\right). \quad (16.3)$$

After this change of variables Hamiltonian H becomes

$$H^\pm := H^\pm(I, \varphi, P, Q) := h^\pm(I, P) + \mu f^\pm(I, \varphi, P, Q) := h(I) + E^\pm(P) + \mu f^\pm(I, \varphi, P, Q)$$

where $f^\pm(I, \varphi, P, Q) := f(I, \varphi, \phi^\pm(P, Q))$.

Stability in the region \mathcal{E}_1^+ . In the region \mathcal{E}_1^+ , the proof of the stability of the actions variables follows by a straightforward application of the Nekhoroshev Theorem as proved in Theorem 1 of [93]. In order to apply such Theorem we need some definitions.

⁹⁸ $\Pi_{p,q}$ denotes the projection onto the (p, q) variables.

⁹⁹The case with $p < 0$ is completely analogous.

¹⁰⁰If $f(x), g(x)$ are positive function, with the symbol $f \approx g$ we mean that $\exists c_1, c_2 > 0$ such that $c_1 g(x) \leq f(x) \leq c_2 g(x), \forall x$.

For $l, m > 0$, a function $h := h(J)$ is said to be l, m -quasi-convex on $A \subset \mathbb{R}^{d+1}$, if at every point $J \in A$ at least one of the inequalities

$$|\langle h'(J), \xi \rangle| > l|\xi|, \quad \langle h''(J)\xi, \xi \rangle \geq m|\xi|^2$$

holds for each $\xi \in \mathbb{R}^{d+1}$. Using the previous lemma it is possible to prove that, for every $\bar{r} > 0$, the Hamiltonian h^+ is l, m -quasi-convex in the set $S := D_{r_0}^+ \times \{I \in \mathbb{R}^{d+1} \mid |I| \leq \bar{r}\}_{r_0}$ with $l, m = O(1)$. In the previous set also holds

$$\|(h^+)''\| =: M = O(\mu^{-c_d} \ln^{-3}(1/\mu)), \quad \|(h^+)'\| =: \Omega_0 = O(1).$$

Putting

$$\begin{aligned} \varepsilon &:= \mu \|f^+\|_{S, s_0} = O(\mu), \quad \alpha := (1 - 2c_d(d+3))/2(d+2), \\ \varepsilon_0 &:= 2^{-10} r_0^2 m (m/11M)^{2(d+2)} = O(\mu^{2c_d(d+3)} \ln^{6(d+2)}(1/\mu)), \end{aligned}$$

we obtain that, if the initial data $(I(0), \varphi(0), p(0), q(0)) \in \mathcal{E}_1^+$, that is $P(0) \in D^+$, then

$$|I(t) - I(0)| \leq \text{const.} \mu^\alpha \ln^{-3}(1/\mu), \quad \text{for } |t| \leq \text{const.} \exp(\text{const.} \mu^{-\alpha} \ln^2(1/\mu)). \quad (16.4)$$

If $c_d < 1/2(d+3)$ then $\alpha > 0$ and we obtain stability for exponentially long times.

Stability in the region \mathcal{E}_1^- . In the region \mathcal{E}_1^- we cannot use the Nekhoroshev Theorem as proved in [93], because E^- is concave and so h^- is not quasi-convex. However we can still apply the Nekhoroshev Theorem in its original and more general form as proved in [90] (see also [91]); in fact the function h^- proves to be *steep* (see Definition 1.7.C. pag. 6 of [90]).

For simplicity we prove the *steepness* of the function h^- in the case $d = 1$ only. In this case $h^- = h^-(I_0, I_1, P) = I_0 + I_1^2/2 + E^-(P)$. We need more informations on the function E^- . In the following, in order to simplify the notation, we will forget the apex $-$ writing, for example, $E = E^-$ and $P = P^-$.

By (1.11) of [90], since $\nabla h^- \neq 0$, a sufficient condition for h^- to be steep is that the system

$$\begin{aligned} \eta_1 + I\eta_2 + E'(P)\eta_3 &:= 0 \\ \eta_2^2 + E''(P)\eta_3^2 &:= 0 \\ E'''(P)\eta_3^3 &:= 0 \end{aligned} \quad (16.5)$$

has no real solution apart from the trivial one $\eta_1 = \eta_2 = \eta_3 = 0$.

Making the change of variable $\psi = \arccos(1 - \tilde{E} + \xi \tilde{E})$, where $\tilde{E} = E + 2$, we get¹⁰¹

$$\dot{P}(E) = \int_0^1 F_1(\xi; E) d\xi, \quad \ddot{P}(E) = 3^{-1/2} \int_0^1 F_2(\xi; E) d\xi, \quad \ddot{\ddot{P}}(E) = \int_0^1 F_3(\xi; E) d\xi, \quad (16.6)$$

¹⁰¹We will denote with “ $\ddot{\cdot}$ ” the derivative with respect to E , and with “ $\dot{\cdot}$ ” the derivative with respect to P .

where

$$\begin{aligned}
F_1(\xi; E) &:= \frac{\sqrt{2}}{\pi\sqrt{\xi}\sqrt{1-\xi}\sqrt{\tilde{E}\xi - E}} \\
F_2(\xi; E) &:= \frac{\sqrt{6}\sqrt{1-\xi}}{2\pi\sqrt{\xi}(\tilde{E}\xi - E)^{3/2}} \\
F_3(\xi; E) &:= \frac{3\sqrt{2}(1-\xi)^{3/2}}{4\pi\sqrt{\xi}(\tilde{E}\xi - E)^{5/2}}.
\end{aligned} \tag{16.7}$$

From the equation $E(P(E)) = E$, deriving with respect to E , we obtain that

$$E'''(P(E)) = -(\dot{P}(E))^{-5}[\dot{P}(E)\ddot{P}(E) - 3(\ddot{P}(E))^2].$$

We want to prove that

$$E'''(P(E)) < 0. \tag{16.8}$$

for every E with $-2 < E < 0$. This is equivalent to prove that $\dot{P}(E)\ddot{P}(E) > 3(\ddot{P}(E))^2$. Using (16.7) we see that $F_1F_3 = F_2^2$ and hence, noting that $F_3(\xi; E)$ is not proportional to $F_1(\xi; E)$ for every E fixed, we conclude that $\int F_1 \int F_3 > (\int F_2)^2$ by a straightforward application of Cauchy-Schwarz inequality and (16.8) follows from (16.6).

By (16.8) the unique solution of the system (16.5) is the trivial one $\eta_1 = \eta_2 = \eta_3 = 0$, hence the function h^- is steep. It is simple to prove that the so called *steepness coefficients* and *steepness indices* (see again Definition 1.7.C. pag. 6 of [90]) can be taken uniformly for $-2 + \mu^{c_d} \leq E \leq -\mu^{c_d}$: that is they do not depend on μ .

Now we are ready to apply the Nekhoroshev Theorem in the formulation given in Theorem 4.4 of [90]. In order to use the notations of [90] we need the following substitutions¹⁰²:

$$(I, P) \rightarrow I, \quad (\varphi, Q) \rightarrow \varphi, \quad H^- \rightarrow H, \quad h^- \rightarrow H_0, \quad \mu f^- \rightarrow H_1, \quad r_0 \rightarrow \rho,$$

$$\{I \in \mathbb{R}^{d+1} \mid |I| \leq \bar{r}\} \times D^- \rightarrow G, \quad \{I \in \mathbb{R}^{d+1} \mid |I| \leq \bar{r}\}_{r_0} \times \mathbb{T}_{s_0}^{d+1} \times D_{r_0}^+ \times \mathbb{T}_{s_0} \rightarrow F.$$

Defining $m := \sup_F \|\frac{\partial^2 H_0}{\partial I^2}\|$ and remembering (16.3) and the definition of r_0 , we have

$$m \leq \text{const.} \mu^{-c_d} \ln^{-3}(1/\mu), \quad \rho = \text{const.} \mu^{c_d}. \tag{16.9}$$

In order to apply the Theorem we have only to verify the following condition

$$M := \sup_F |H_1| < M_0 \tag{16.10}$$

where M_0 depends only on the steepness coefficients and steepness indices (which are independent of μ) and on m and ρ (which depend on μ). Moreover we use the fact that the dependence of M_0 on m and ρ is, “polynomial” (although it is quite cumbersome):

¹⁰²We observe that we do not need to introduce the (p, q) variables so in our case $C = +\infty$.

that is there exist constant $\tilde{c}_d, \bar{c}_d > 0$ such that $M_0(m, \rho) \geq \text{const.} m^{-\tilde{c}_d} \rho^{\bar{c}_d}$ (see §6.8 of [91]). So condition (16.10) becomes, using (16.9),

$$\mu \leq \text{const.} \mu^{c_d(\tilde{c}_d + \bar{c}_d)} \ln^{3\tilde{c}_d}(1/\mu),$$

which is verified choosing $c_d < (\tilde{c}_d + \bar{c}_d)^{-1}$.

Now we can apply the Nekhoroshev Theorem as formulated in Theorem 4.4 of [90], obtaining that if $(I(0), \varphi(0), p(0), q(0)) \in \mathcal{E}_1^-$ then

$$|I(t) - I(0)| \leq d/2 := M^b/2 = O(\mu^b) \quad \forall |t| \leq T := \frac{1}{M} \exp\left(\frac{1}{M}\right)^a = O\left(\frac{1}{\mu} \exp\left(\frac{1}{\mu}\right)^a\right) \quad (16.11)$$

where $a, b > 0$ are some constants depending only on the steepness properties of H_0 . Finally, choosing

$$c_d < \min\{(2d+6)^{-1}, (\tilde{c}_d + \bar{c}_d)^{-1}\}, \quad (16.12)$$

we have proved the exponential stability in the region \mathcal{E}_1 .

Stability in the region \mathcal{E}_2^+ . In the following we will denote $I^* := (I_1, \dots, I_d)$ the projection on the last d coordinates. We shall prove the following lemma

Lemma 16.2 $\forall \kappa > 0, \exists \kappa_0, \mu_8 > 0$ such that $\forall 0 < \mu \leq \mu_8$, if $(I(t), \varphi(t), p(t), q(t)) \in \mathcal{E}_2^+$ for $0 < t \leq \bar{T}$, then

$$|I^*(t) - I^*(0)| \leq \frac{\kappa}{2} \quad \forall t \leq \min\left\{\frac{\kappa_0}{\mu} \ln \frac{1}{\mu}, \bar{T}\right\}.$$

It is quite obvious that for initial conditions $(I(0), \varphi(0), p(0), q(0)) \in \mathcal{E}_2^+$, Theorem 10.3 follows from lemma 16.2 and the exponential stability in the region \mathcal{E}_1 .

In order to prove lemma 16.2 let us define, for some fixed $0 < \delta < \pi/4$, the following two regions in the phase space : $U := \{(I, \varphi, p, q) \mid |q| \leq \delta \bmod 2\pi, |E(p, q)| \leq 2\mu^{c_d}\}$ and $V := \{(I, \varphi, p, q) \mid |q| > \delta \bmod 2\pi, |E(p, q)| \leq 2\mu^{c_d}\}$. We first note that¹⁰³

$$\begin{aligned} z(t) \in V \quad \forall t_1 < t < t_2, \quad |q(t_1)|, |q(t_2)| = \delta \bmod 2\pi \\ \Downarrow \\ t_2 - t_1 < c_1, \quad |I(t_2) - I(t_1)| \leq c_2(t_2 - t_1)\mu. \end{aligned} \quad (16.13)$$

Indeed in this case $\forall t_1 < t < t_2, c_3 \leq |\dot{q}(t)| \leq c_4$. This implies that $t_2 - t_1 \leq c_1$ and then, integrating the equation of motion $\dot{I} = -\mu \partial_\varphi f$ in (t_1, t_2) , we immediately get (16.13). We also claim that

$$\forall t_1 < t < t_2, z(t) \in U \text{ and } |q(t_1)|, |q(t_2)| = \delta \bmod 2\pi \Rightarrow t_2 - t_1 \geq c_5 |\ln \mu|. \quad (16.14)$$

We denote with t_U^i (resp. t_V^i) the i -th time for which the orbits enters in (resp. goes out from) U , so that $t_U^i < t_V^i < t_U^{i+1} < t_V^{i+1}$ for $0 \leq i \leq i_0$. From (16.14) it follows that

¹⁰³In the following we will use c_i to denote some positive constant independent on μ .

$i_0 \leq c_6 \kappa_0 / \mu$ and, from (16.13), that the time T_V spent by the orbit in the region V is bounded by $c_7 \kappa_0 / \mu$.

In order to prove (16.14) we use the following normal form result for the pendulum Hamiltonian $E(p, q)$ in a neighborhood of its hyperbolic equilibrium point (see e.g. [49])

Lemma 16.3 *There exist $R, \tilde{\delta} > 0$, an analytic function g , with $g'(0) = -1$ and an analytic canonical transformation $\Phi : B \rightarrow \{|p| \leq \tilde{\delta}\} \times \{|q| \leq \delta \bmod 2\pi\}$ where $B := \{|P|, |Q| \leq R\}$, such that $E(\Phi(P, Q)) = g(PQ)$.*

In the coordinates (Q, P) the local stable and unstable manifolds are resp. $W_{loc}^s = \{P = 0\}$ and $W_{loc}^u = \{Q = 0\}$ and Hamiltonian (16.1) writes as

$$\tilde{H} := \tilde{H}(I, \varphi, P, Q) := h(I) + g(PQ) + \mu \tilde{f}(I, \varphi, P, Q)$$

where $\tilde{f}(I, \varphi, P, Q) := f(I, \varphi, \Phi(P, Q))$.

We are now able to prove (16.14). Certainly there exists an instant $t_1^* \in [t_1, t_2]$ for which $(p(t_1^*), q(t_1^*)) \in \Phi(B)$ but, $\forall t_1 < t < t_1^*$, $(p(t), q(t)) \notin \Phi(B)$. It follows that, if we take the representant $q(t_1) \in [-\delta, \delta]$, then $p(t_1^*)q(t_1^*) < 0$. We will denote with $Z(t) := (I(t), \varphi(t), P(t), Q(t)) = (I(t), \varphi(t), \Phi^{-1}(p(t), q(t)))$ the corresponding solution of the Hamiltonian system associated to \tilde{H} . From the fact that $|q(t_1^*)| = \delta$ or $(p(t_1^*), q(t_1^*)) \in \partial\Phi(B)$ and that $|g(PQ)| \leq \mu^{c_d}$, $p(t_1^*)q(t_1^*) < 0$, it follows that $|P(t_1^*)| \leq c_8 \mu^{c_d}$ and $|Q(t_1^*)| \geq c_9$.

In the same way there exists an instant t_2^* with $t_1 < t_1^* < t_2^* < t_2$ for which $(P(t_2^*), Q(t_2^*)) \in B$ but, $\forall t > t_2^*$ $(P(t), Q(t)) \notin B$; in particular it results $|P(t_2^*)| \geq c_{10}$. We claim that $t_2^* - t_1^* \geq c_{11} \ln(1/\mu)$. Indeed $P(t)$ satisfies the Hamilton's equation $\dot{P}(t) = -g'(P(t)Q(t))P(t) - \mu \partial_Q \tilde{f}(I(t), \varphi(t), P(t), Q(t))$ with initial condition $|P(t_1^*)| \leq c_8 \mu^{c_d}$. Since $|P(t_2^*)| \geq c_{10}$, we can derive from Gronwall's lemma that $t_2^* - t_1^* \geq c_{11} \ln(1/\mu)$, which implies (16.14).

By the following normal-form lemma there exists a close to the identity symplectic change of coordinates removing the non-resonant angles φ in the perturbation up to $O(\mu^2)$. It can be proved by standard perturbation theory (see for similar lemmas section §5 of [49]).

Lemma 16.4 *Let $\beta > 0$. There exist $R, \rho > 0$ so small that, defining $\lambda := \min_{|\xi| \leq R^2} |g'(\xi)|$, $S := \max_{|\xi| \leq R^2} |g''(\xi)|$, then $\lambda \geq 2SR^2$ and $\rho \leq \min\{\lambda/4N, R^2/8S, \beta/2N, r\}$. Let Λ be a sublattice of \mathbb{Z}^{d+1} . Let $\mathcal{D} \subset \mathbb{R}^{d+1}$ be bounded and β -non-resonant mod Λ , i.e. $\forall I \in \mathcal{D}$, $h \in \mathbb{Z}^{d+1} \setminus \Lambda$, $|h| \leq N$ it results $|(1, I^*) \cdot h| \geq \beta$. Suppose that*

$$\varepsilon := \mu \|\tilde{f}\|_{B, D, s} \leq 2^{-11} \beta_* \rho s, \quad (16.15)$$

where¹⁰⁴ $D := \mathcal{D}_\rho$, $\beta_* := \min\{\beta, \lambda/2\}$. Then there exists an analytic canonical transformation

$$\begin{aligned} \Psi : \overline{D} \times \mathbb{T}_{s/4}^{d+1} \times \overline{B} &\longrightarrow D \times \mathbb{T}_s^{d+1} \times B \\ (\overline{I}, \overline{\varphi}, \overline{P}, \overline{Q}) &\longmapsto (I, \varphi, P, Q) \end{aligned} \quad (16.16)$$

¹⁰⁴ B and D are thought as complex domains, as in the sequel \overline{B} and \overline{D} .

with $\overline{B} := \{|\overline{P}|, |\overline{Q}| \leq R/8\}$, $\overline{D} := \mathcal{D}_{\rho/4}$, such that

$$\overline{H} := \overline{H}(\overline{I}, \overline{\varphi}, \overline{P}, \overline{Q}) := \widetilde{H} \circ \Psi = h(\overline{I}) + \overline{g}(\overline{I}, \overline{\varphi}, \overline{P}, \overline{Q}) + \overline{f}(\overline{I}, \overline{\varphi}, \overline{P}, \overline{Q})$$

with $\overline{g}(\overline{I}, \overline{\varphi}, \xi) := g(\xi) + f^*(\overline{I}, \overline{\varphi}, \xi)$, $f^*(\overline{I}, \overline{\varphi}, \xi) = \sum_{h \in \Lambda, |h| \leq N} f_h^*(\overline{I}, \xi) e^{ih \cdot \overline{\varphi}}$ and $\|f^*\|_{\overline{B}, \overline{D}, s/4} \leq \varepsilon$. Moreover the following estimates hold

$$|\overline{I} - I| \leq \frac{2^4 \varepsilon}{\beta_* s}, \quad |\overline{P} - P|, |\overline{Q} - Q| \leq \frac{2^5 \varepsilon}{R \beta_*}, \quad \|\overline{f}\|_{\overline{B}, \overline{D}, s/4} \leq \frac{2^9 \varepsilon^2}{\beta_* \rho s}. \quad (16.17)$$

Let \mathcal{L} be the (finite) set of the maximal sublattices $\Lambda = \langle h_1, \dots, h_s \rangle \subset \mathbb{Z}^{d+1}$ for some independent $h_i \in \mathbb{R}^{d+1}$ with $|h_i| \leq N$ for $i = 1, \dots, s \leq d$. For $\Lambda \in \mathcal{L}$ we define the Λ -resonant frequencies $R^\Lambda := \{I^* \in \mathbb{R}^d \mid (1, I^*) \cdot h = 0, \forall h \in \Lambda\}$ and the set of the s -order resonant frequencies $Z^s := \cup_{\dim \Lambda = s} R^\Lambda$.

Setting $h_i = (l_i, n_i)$ with $l_i \in \mathbb{R}$, $n_i \in \mathbb{R}^d$, we remark that if $R^\Lambda \neq \emptyset$ then n_1, \dots, n_s are independent. We also define the $(d-s)$ -dimensional linear subspace (associated with the affine subspace R^Λ) $L^\Lambda := \cap_{i=1}^s n_i^\perp \subset \mathbb{R}^d$ and we denote by Π^Λ the orthogonal projection from \mathbb{R}^d onto L^Λ .

Since \mathcal{L} is a finite set, $\alpha := \min_{\Lambda \in \mathcal{L}} \min_{n \in \mathbb{Z}^d, |n| \leq N, \Pi^\Lambda n \neq 0} |\Pi^\Lambda n|$ is strictly positive.

We now perform a suitable version of the standard “covering lemma” in which the whole frequency space is covered by non-resonant zones. The fundamental blocks used to construct this covering will be r -neighborhoods of any R^Λ i.e. $R_r^\Lambda := \{I^* \in \mathbb{R}^d \mid \text{dist}(I^*, R^\Lambda) \leq r\}$ for suitable $r > 0$ depending on $\dim \Lambda$. Let $r_d > 0$ be such that $(d+1)r_d < c_{12}\kappa$, for some c_{12} sufficiently small to be determined. For $1 \leq s \leq d-1$ we can define recursively numbers r_s sufficiently small such that $0 < r_s < \alpha r_{s+1}/2N$, verifying¹⁰⁵

$$\dim \Lambda = \dim \Lambda' = s, \quad R^\Lambda \neq R^{\Lambda'} \implies R_{(s+1)r_s}^\Lambda \cap R_{(s+1)r_s}^{\Lambda'} \subset \cup_{i=s+1}^d Z_{r_i}^i. \quad (16.18)$$

We also define, for $1 \leq s \leq d-1$, $S^0 := \mathbb{R}^d \setminus (\cup_{i=1}^d Z_{2r_i}^i)$ and $S^s := Z_{(s+1)r_s}^s \setminus (\cup_{i=s+1}^d Z_{(s+2)r_i}^i)$, i.e. the s -order resonances minus the higher-order ones. We claim that $\mathbb{R}^d = S^0 \cup \dots \cup S^{d-1} \cup Z_{(d+1)r_d}^d$ is the covering that we need. We also define $S^0 \subset S_*^0 := \mathbb{R}^d \setminus (\cup_{i=1}^d Z_{r_i}^i)$ and $S^s \subset S_*^s := Z_{(s+1)r_s}^s \setminus (\cup_{i=s+1}^d Z_{(s+1)r_i}^i)$.

If the orbit lies near a certain R^Λ (but far away from higher order resonances) then the following lemma says that the drift of the actions I^* in the direction which is parallel to R^Λ is small.

Lemma 16.5 *Suppose that $I^*(0) \in S^s$, $I^*(t) \in S_*^s$ and $|I^*(t)| \leq \bar{r} + r/2$, $\forall 0 \leq t \leq T^*$ for some $T^* \leq \kappa_0 |\ln \mu|/\mu$ and $0 \leq s \leq d-1$. Then, if $s \geq 1$, there exists a sublattice $\Lambda \subset \mathbb{Z}^{d+1}$, $\dim \Lambda = s$ such that $I^*(t) \in R_{(s+1)r_s}^\Lambda \setminus (\cup_{i=s+1}^d Z_{(s+1)r_i}^i)$, $\forall 0 \leq t \leq T^*$. Moreover if κ_0 is sufficiently small¹⁰⁶*

$$|\Pi^\Lambda(I^*(t) - I^*(0))| \leq r_1/2 \quad \forall 0 \leq t \leq T^* \quad (16.19)$$

¹⁰⁵ Assumption (16.18) means that, in order to go from a neighborhood of a $(d-s)$ -order resonance to a different one, we have to pass through an higher order dimensional one.

¹⁰⁶ In the case $s = 0$ Π^Λ is simply the identity on \mathbb{R}^d .

and hence, for $s \geq 1$, $|I^*(t) - I^*(0)| \leq 2(s+1)r_s + r_1/2$. In particular for $I^*(0) \in S^0$ we have that $|I^*(t) - I^*(0)| \leq r_1/2$, $\forall 0 \leq t \leq T^*$.

PROOF. In the case $s = 0$ we take $\Lambda = \{0\}$. The existence of Λ is trivial because $I^*(0) \in S^s$ and hence $I^*(0) \in R_{(s+1)r_s}^\Lambda$ for some $\Lambda \in \mathcal{L}$ with $\dim \Lambda = s$. The fact that $I^*(t) \in R_{(s+1)r_s}^\Lambda \setminus (\cup_{i=s+1}^d Z_{(s+1)r_i}^i)$, $\forall 0 \leq t \leq T^*$, follows from $I^*(t) \in S^s$, $\forall 0 \leq t \leq T^*$ and (16.18). Now we want to apply lemma 16.4 with $\beta := \alpha r_1/2$ and $\mathcal{D} := R_{(s+1)r_s}^\Lambda \setminus (\cup_{i=s+1}^d Z_{(s+1)r_i}^i)$. We have to verify that \mathcal{D} is β -non-resonant mod Λ . Fix $|h_0| \leq N$, $h_0 = (l_0, n_0) \notin \Lambda$ (resp. $\neq 0$ for $s = 0$). We first estimate $|l_0 + n_0 \cdot I_0^*|$ for all $I_0^* \in \mathcal{D}_0 := R^\Lambda \setminus (\cup_{i=s+1}^d Z_{(s+1)r_i}^i)$. If $\Lambda' := \Lambda \oplus \langle h_0 \rangle$ and $n_0^* := \Pi^\Lambda n_0$ we have two cases: $n_0^* \neq 0$ or $n_0^* = 0$. In $n_0^* \neq 0$ we can perform the following decomposition: $I_0^* = I_1^* + v$ with $I_1^* \in R^{\Lambda'}$, $v \in L^\Lambda$ and moreover¹⁰⁷ $v = \pm |v|n_0^*/|n_0^*|$. Since $I_0^* \notin (\cup_{i=s+1}^d Z_{(s+1)r_i}^i)$ then $I_0^* \notin Z_{(s+1)r_{s+1}}^{\Lambda'}$ and, hence $|v| \geq (s+1)r_{s+1}$. Using the previous estimate, the fact that $I_1^* \in \Lambda'$ and $|n_0^*| \geq \alpha$, we conclude that

$$|l_0 + n_0 \cdot I_0^*| = |(l_0 + n_0 \cdot I_1^*) + n_0 \cdot v| = |n_0 \cdot v| = |n_0^* \cdot v| = |v||n_0^*| \geq \alpha(s+1)r_{s+1}. \quad (16.20)$$

Now we consider the case in which $n_0^* = 0$. In this case it is simple to see that $h_0 = (l', 0) + h$ where $h \in \Lambda$ and $l' \in \mathbb{Z} \setminus \{0\}$. So $|l_0 + n_0 \cdot I_0^*| = |l'| \geq 1$. Now we can prove that $|l_0 + n_0 \cdot I^*| \geq \beta$ for all $I^* \in \mathcal{D}$. In fact $I^* = I_0^* + u$ with $I_0^* \in \mathcal{D}_0$ and $|u| \leq (s+1)r_s$. Using (16.20) and $r_s < \alpha r_{s+1}/2N$, we have

$$|l_0 + n_0 \cdot I^*| \geq |l_0 + n_0 \cdot I_0^*| - |n_0 \cdot u| \geq \alpha(s+1)r_{s+1} - N(s+1)r_s \geq \alpha(s+1)r_{s+1}/2 \geq \beta,$$

proving that \mathcal{D} is β -non-resonant mod Λ . Finally we can verify (16.15) if μ_8 is sufficiently small. Now we are ready to apply lemma 16.4 in order to prove (16.19). Using (16.13), the fact that f^* contains only the Λ -resonant Fourier coefficients, (16.17) and Hamilton's equation for \bar{H} we have

$$\begin{aligned} |\Pi^\Lambda(I^*(t) - I^*(0))| &\leq c_2 T_V \mu + c_{13} \mu^2 (\kappa_0 |\ln \mu| / \mu) + c_{14} i_0 \mu \\ &\leq c_2 c_7 \kappa_0 + c_{13} \mu \kappa_0 |\ln \mu| + c_{14} c_6 \kappa_0 \leq r_1/2 \end{aligned}$$

if κ_0 and μ_8 are sufficiently small. ■

PROOF OF LEMMA 16.2. Suppose first that $|I^*(t)| \leq \bar{r} + r/2 \forall 0 \leq t \leq \kappa_0 |\ln \mu| / \mu$. If $I^*(0) \in Z_{(d+1)r_d}^d$ and $I^*(t) \in Z_{(d+1)r_d}^d \forall 0 \leq t \leq \kappa_0 |\ln \mu| / \mu$ then $|I^*(t) - I^*(0)| \leq 2(d+1)r_d$ and the lemma is proved if $c_{12} < 1/4$. Otherwise we can suppose that $I^*(0) \in S^s$ for some $0 \leq s \leq d-1$. If $I^*(t) \in S_*^s \forall 0 \leq t \leq \kappa_0 |\ln \mu| / \mu$ then we can apply the lemma 16.5 proving the lemma for c_{12} small enough. Suppose that $\exists 0 < T^* < \kappa_0 |\ln \mu| / \mu$ such that $I^*(t) \in S_*^s \forall 0 \leq t < T^*$ but $I^*(T^*) \notin S_*^s$. We will prove that

$$I^*(T^*) \in S^0 \cup \dots \cup S^{s-1} \quad (16.21)$$

that means that the orbit can only enter in zones that are “less” resonant. In fact by lemma 16.5 we see that $I^*(T^*) \notin \cup_{i=s+1}^d Z_{(s+1)r_i}^i$, moreover, since $I^*(T^*) \notin S_*^s$, we have

¹⁰⁷We observe that $\text{dist}(I_0^*, R^{\Lambda'}) = |v|$.

that $I^*(T^*) \notin Z_{(s+1)r_s}^s$ and hence $I^*(T^*) \notin \cup_{i=s}^d Z_{(s+1)r_i}^i$. If $I^*(T^*) \in S^0$ we have finished. If $I^*(T^*) \notin S^0$ then $I^*(T^*) \in \cup_{i=1}^{s-1} Z_{2r_i}^i \subseteq \cup_{i=1}^{s-1} Z_{(i+1)r_i}^i$. If $I^*(T^*) \in S^1$ we have finished. If $I^*(T^*) \notin S^1$ then $I^*(T^*) \notin Z_{2r_1}^1 \setminus \cup_{i=2}^d Z_{3r_i}^i$ and hence $I^*(T^*) \in \cup_{i=2}^{s-1} Z_{(i+1)r_i}^i$. Iterating this procedure we prove (16.21).

The conclusion is that if the order of resonance changes along the orbit, it can decrease only so that the orbit may eventually arrive in the completely non resonant zone S^0 where there is stability. Considering the “worst” case i.e. when $I^*(0) \in Z_{(d+1)r_d}^d$ and the orbit arrives in S^0 , summing all the contributions from lemma 16.5, we have that, if c_{12} is sufficiently small,

$$|I^*(t) - I^*(0)| \leq 2(d+1)r_d + \sum_{s=1}^{d-1} (2(s+1)r_s + r_1/2) + r_1/2 = \sum_{s=1}^d 2(s+1)r_s + dr_1/2 \leq \kappa/2. \quad (16.22)$$

In order to conclude the proof of the lemma we have only to prove that if $|I^*(0)| \leq \bar{r}$ then $|I^*(t)| \leq \bar{r} + r/2 \forall 0 \leq t \leq \kappa_0 |\ln \mu|/\mu$. This is an immediate consequence of (16.22) and of the fact that $\kappa \leq r$. ■

Stability in the region \mathcal{E}_2^- . If, for all $t \geq 0$ $(p(t), q(t)) \in \mathcal{E}_2^-$, then it follows easily that $|p(t)|, |q(t) - \pi| = O(\mu^{c_d/2})$. Then, defining $f_1(I, \varphi) := f(I, \varphi, 0, \pi)$ and $f_2(I, \varphi, t) := \mu^{-c_d/2}[f(I, \varphi, p(t), q(t)) - f_1(I, \varphi)]$, it results that $|\partial_I f_2(I, \varphi; t)|, |\partial_\varphi f_2(I, \varphi; t)| \leq \text{const}$. Clearly if $(I(t), \varphi(t), q(t), p(t))$ is a solution of (16.1) then $(I(t), \varphi(t))$ is solution of Hamiltonian

$$H_1 := H_1(I, \varphi; t) := h(I) + \mu f_1(I, \varphi) + \mu^{1+(c_d/2)} f_2(I, \varphi; t).$$

Now¹⁰⁸ one can construct, in the standard way, an analytic symplectic map $\Phi : (\bar{I}, \bar{\varphi}) \rightarrow (I, \varphi)$ with $|\bar{I} - I| = O(\mu/\beta)$, and two analytic functions \bar{h}, \bar{f} such that $[h + \mu f_1] \circ \Phi(\bar{I}, \bar{\varphi}) = \bar{h}(\bar{I}) + \bar{f}(\bar{I}, \bar{\varphi})$ with $\|\bar{f}\| = O(\mu^2)$. Defining $f_3 := f_3(\bar{I}, \bar{\varphi}; t) := f_2(\Phi(\bar{I}, \bar{\varphi}); t)$ we also get that $|\partial_{\bar{I}} f_3(\bar{I}, \bar{\varphi}; t)|, |\partial_{\bar{\varphi}} f_3(\bar{I}, \bar{\varphi}; t)| \leq \text{const.}/\beta$. The solutions of Hamiltonian H_1 are symplectically conjugated, via Φ^{-1} , to the solutions of the Hamiltonian

$$H_2 := H_2(\bar{I}, \bar{\varphi}; t) := \bar{h}(\bar{I}) + \bar{f}(\bar{I}, \bar{\varphi}) + \mu^{1+(c_d/2)} f_3(\bar{I}, \bar{\varphi}; t)$$

for which we obtain, directly from Hamilton's equations, the estimates

$$|\bar{I}(t) - \bar{I}(0)| \leq \text{const} \cdot \mu^{c_d/4}, \quad \forall |t| \leq \text{const} \cdot \mu^{-1-c_d/4}.$$

It follows that, if $(I(0), \varphi(0), p(0), q(0)) \in \mathcal{E}_2^-$, then $\forall |t| \leq \text{const} \cdot \mu^{-1-c_d/4}$,

$$|I(t) - I(0)| \leq |I(t) - \bar{I}(t)| + |\bar{I}(t) - \bar{I}(0)| + |\bar{I}(0) - I(0)| \leq \text{const} \cdot \mu^{c_d/4},$$

(if at some instant t the solution $z(t)$ escapes outside \mathcal{E}_2^- it is exponentially stable in time).

Finally, from the previous steps, we can conclude that there exists $\mu_1 > 0$ such that $0 < \mu \leq \mu_1$ Theorem 10.3 holds.

¹⁰⁸For brevity we prove only the case in which $I(0)$ is in a non-resonant zone. The resonant case can be treated as in \mathcal{E}_2^+ .

Part V

Appendix

A Proof of lemma 11.1.

We shall use the following lemma:

Lemma A.1 *There exists $T_0 > 0$ such that, $\forall T \geq T_0$, for all continuous $f : [-1, T+1] \rightarrow \mathbb{R}$, there exists a unique solution h of*

$$-\ddot{h} + \cos Q_T(t)h = f, \quad h(0) = h(T) = 0. \quad (\text{A.1})$$

The Green operator $\mathcal{G} : C^0([-1, T+1]) \rightarrow C^2([-1, T+1])$ defined by $\mathcal{G}(f) := h$, satisfies

$$\max_{t \in [-1, T+1]} |h(t)| + |\dot{h}(t)| \leq C \max_{t \in [-1, T+1]} |f(t)| \quad (\text{A.2})$$

for some positive constant C independent of T .

PROOF. We first note that the homogeneous problem (A.1) (i.e. $f = 0$) admits only the trivial solution $h = 0$. This immediately implies the uniqueness of the solution of (A.1). The existence result follows by the standard theory of linear second order differential equations. We now prove that any solution h of (A.1) satisfies (A.2). It is enough to show that $\max_{t \in [-1, T+1]} |h(t)| \leq C' \max_{t \in [-1, T+1]} |f(t)|$. Indeed we obtain by (A.1) that $\max_{t \in [-1, T+1]} |h(t)| + |\dot{h}(t)| \leq (2C' + 1) \max_{t \in [-1, T+1]} |f(t)|$ and, by elementary analysis, this implies (A.2) for an appropriate constant C .

Arguing by contradiction, we assume that there exist sequences $(T_n) \rightarrow \infty$, (f_n) , (h_n) such that

$$-\ddot{h}_n + \cos Q_{T_n}(t)h_n = f_n, \quad h_n(0) = h_n(T_n) = 0, \quad |h_n|_n := \max_{t \in [-1, T_n+1]} |h_n(t)| = 1, \quad |f_n|_n \rightarrow 0.$$

By the Ascoli-Arzelà Theorem there exists $h \in C^2([-1, \infty), \mathbb{R})$ such that, up to a subsequence, $h_n \rightarrow h$ in the topology of C^2 uniform convergence in $[-1, M]$ for all $M > 0$. Since $Q_{T_n} \rightarrow q_0 - 2\pi$ uniformly in all bounded intervals of $[-1, \infty)$, we obtain that

$$-\ddot{h} + \cos q_0(t)h = 0, \quad h(0) = 0, \quad \sup_{t \in [-1, \infty)} |h(t)| \leq 1. \quad (\text{A.3})$$

Now the solutions of the linear differential equation in (A.3) have the form $h = K_1\xi + K_2\psi$, where $(K_1, K_2) \in \mathbb{R}^2$, $\xi(t) = \dot{q}_0(t) = \frac{2}{\cosh t}$ and $\psi(t) = \frac{1}{4}(\sinh t + \frac{t}{\cosh t})$ satisfies $\psi\xi - \xi\psi = 1$. The bound on h implies that $K_2 = 0$ and $h(0) = 0$ implies that $K_1 = 0$. Hence $h = 0$. In the same way we can prove that $h_n(\cdot - T_n) \rightarrow 0$ uniformly in every bounded subinterval of $(-\infty, 1]$.

Now let us fix \bar{t} such that for all n large enough, for all $t \in [\bar{t}, T_n - \bar{t}]$, $\cos Q_{T_n}(t) \geq 1/2$ (\bar{t} does exist because of (11.4)). By the previous step, for n large enough, there exists a maximum point $t_n \in (\bar{t}, T_n - \bar{t})$ of $h_n^2(t)$, i.e. $h_n^2(t_n) = |h_n|_n^2 = 1$. Then $(h_n^2)(t_n) = 2h_n(t_n)\dot{h}_n(t_n) = 0$ and $(\ddot{h}_n^2)(t_n) = 2\ddot{h}_n(t_n)h_n(t_n) + 2\dot{h}_n^2(t_n) \leq 0$. By the differential equation satisfied by h_n , we can derive from the latter inequality that $\cos Q_{T_n}(t_n)h_n^2(t_n) \leq f_n(t_n)h_n(t_n)$, i.e. $\cos Q_{T_n}(t_n) \leq f_n(t_n)$, which, for n large enough, contradicts the property of \bar{t} and the fact that $|f_n|_n \rightarrow 0$. ■

Now we can deal with the existence result of lemma 11.1. Let $T := (\theta^- - \theta^+)$, $\omega = (\varphi^- - \varphi^+)/T$, $\bar{\varphi}(t) := \omega(t - \theta^+) + \varphi^+$. In the following we call c_i constants depending only on f . We are searching for solutions (φ, q) of (11.1) with $\varphi(\theta^\pm) = \varphi^\pm$, $q(\theta^\pm) = \mp\pi$, in the following form

$$\begin{cases} \varphi(t) = \omega(t - \theta^+) + \varphi^+ + v(t - \theta^+) \\ q(t) = Q_T(t - \theta^+) + w(t - \theta^+). \end{cases}$$

Hence we need to find a solution, in the time interval $I := [-1, T + 1]$, of the following two equations

$$\begin{cases} \ddot{v}(t) = -\mu[F_\varphi(v, w)](t), & v(0) = v(T) = 0, \\ [L(w)](t) = [G(v, w)](t) := -[S(w)](t) + \mu[F_q(v, w)](t), & w(0) = w(T) = 0, \end{cases} \quad (\text{A.4})$$

where

$$\begin{aligned} [F_\varphi(v, w; \lambda, \mu)](t) &:= \partial_\varphi f(\omega t + \varphi^+ + v(t), Q_T(t) + w(t), t + \theta^+), \\ [F_q(v, w; \lambda, \mu)](t) &:= \partial_q f(\omega t + \varphi^+ + v(t), Q_T(t) + w(t), t + \theta^+), \\ [S(w)](t) &:= \sin(Q_T(t) + w(t)) - \sin(Q_T(t)) - \cos(Q_T(t))w(t), \\ [L(w)](t) &:= -\ddot{w}(t) + \cos Q_T(t)w(t). \end{aligned}$$

We want to solve (A.4) as a fixed point problem. By lemma A.1, the second equation of (A.4) can be written $w = K := \mathcal{G}(-S + \mu F_q)$. Moreover the first equation (A.4) can be written

$$v(t) = J(t) := [J(v, w; \lambda, \mu)](t) := \bar{J}(t) - \frac{\bar{J}(0)(T - t) + \bar{J}(T)t}{T}, \quad (\text{A.5})$$

where, setting $F_\varphi(s) = F_\varphi(v(s), w(s))$,

$$[\bar{J}(v, w; \lambda, \mu)](t) := -\mu \int_{T/2}^t \int_{T/2}^x F_\varphi(s) \, ds \, dx.$$

Let us consider the Banach space $Z = V \times W := \mathcal{C}^1(I; \mathbb{R}^d) \times \mathcal{C}^1(I; \mathbb{R})$, endowed with the norm $\|z\| = \|(v, w)\| := \max\{\|v\|_V, \|w\|_W\}$, defined by

$$\|v\|_V := \sup_{t \in I} [|v(t)|(1 + c_1 \mu T^2)^{-1} \beta^2 + |\dot{v}(t)|\beta], \quad \|w\|_W := \sup_{t \in I} [|w(t)| + |\dot{w}(t)|]. \quad (\text{A.6})$$

A fixed point of the operator $\Phi : Z \rightarrow Z$ defined $\forall z \in Z$ as $\Phi(z) := \Phi(z; \lambda, \mu) := (J(z), K(z))$ is a solution of (A.4). We shall prove in the sequel that Φ is a contraction in the ball¹⁰⁹ $D := B_{\bar{c}\mu}(Z)$ for an appropriate choice of \bar{c}, c_1, C_0 , provided μ is small enough.

We have $|[S(w)](t)| \leq w^2(t)$, so that $\forall t$, $|[G(v, w)](t)| \leq \bar{c}^2 \mu^2 + c_4 \mu$. Now, choosing first \bar{c} sufficiently large and then μ sufficiently small, we can conclude using (A.2) that, if $z \in D$, $\|K(z)\|_W \leq \bar{c}\mu/4$. Now we study the behaviour of J . Let us first consider \bar{J} . We define

$$f_{nl}(t) := f_{nl}(Q_T(t) + w(t)), \quad g_{nl}(t) := f'_{nl}(Q_T(t) + w(t)),$$

¹⁰⁹If X is a Banach space and $r > 0$ we define $B_r(X) := \{x \in X; \|x\| \leq r\}$.

$$\alpha_{nl} := n \cdot \varphi^+ + l\theta^+, \quad \beta_{nl} := n \cdot \omega + l.$$

For $t \in [-1, T+1]$, $z \in D$, we want to estimate

$$\dot{\bar{J}}(t) = -\mu \int_{T/2}^t F_\varphi = -\mu \sum_{|(n,l)| \leq N} i n e^{i\alpha_{nl}} \int_{T/2}^t f_{nl}(s) e^{in \cdot v(s)} e^{i\beta_{nl}s} ds.$$

Integrating by parts, we obtain

$$-i\beta_{nl} \int_{T/2}^t f_{nl}(s) e^{in \cdot v(s)} e^{i\beta_{nl}s} ds = f_{nl}(T/2) e^{in \cdot v(T/2)} e^{i\beta_{nl}T/2} - f_{nl}(t) e^{in \cdot v(t)} e^{i\beta_{nl}t} \quad (\text{A.7})$$

$$+ \int_{T/2}^t g_{nl}(s) \dot{Q}_T(s) e^{in \cdot v(s)} e^{i\beta_{nl}s} ds \quad (\text{A.8})$$

$$+ \int_{T/2}^t (g_{nl}(s) \dot{v}(s) + f_{nl}(s) i n \cdot \dot{v}(s)) e^{in \cdot v(s)} e^{i\beta_{nl}s} ds \quad (\text{A.9})$$

By (11.4), the term (A.8) is bounded by $c_5 \max\{e^{-K_2 t}, e^{-K_2(T-t)}\}$. Hence, for $z \in D$,

$$\int_{T/2}^t F_\varphi = u(t) - u(T/2) + R(t), \quad \text{with} \quad |R(t)| \leq \frac{c_6}{\beta} \left[\max\{e^{-K_2 t}, e^{-K_2(T-t)}\} + \bar{c}(\mu + \frac{\mu}{\beta})T \right], \quad (\text{A.10})$$

where $u(t) = \sum (n/\beta_{nl}) e^{i\alpha_{nl}} f_{nl}(t) e^{in \cdot v(t)} e^{i\beta_{nl}t}$.

So we can write $\bar{J}(t) = j(t) + \mu(t - T/2)u(T/2)$, where

$$j(t) = \int_{T/2}^t -\mu u(s) ds + \int_{T/2}^t -\mu R(s) ds.$$

By the bound of $R(t)$ given in (A.10), the second integral can be bounded by $c_7(\mu/\beta)[1 + \bar{c}T^2\mu/\beta]$. Integrating once again by parts as above, we find that the first integral is bounded by $c_8(\mu/\beta^2)[1 + \bar{c}(\mu T/\beta)]$, hence, by the condition imposed on μT , it can be bounded by $\mu\bar{c}/8\beta^2$, provided that C_0 has been chosen small enough and \bar{c} is large enough. Hence

$$|j(t)| \leq \frac{\mu\bar{c}}{\beta^2} \left[\frac{c_7}{\bar{c}} + c_7\mu T^2 + \frac{1}{8} \right].$$

In addition

$$\left| \frac{d}{dt} j(t) \right| = \mu |u(t) + R(t)| \leq c_{10} \frac{\mu\bar{c}}{\beta} \left(\frac{1}{\bar{c}} + \frac{\mu T}{\beta} \right).$$

As a result $\|j\|_V \leq \mu\bar{c}/4$, provided \bar{c} and c_1 have been chosen large enough, C_0 small enough.

Now $\bar{J}(t) = j(t) + at + b$, where $a, b \in \mathbb{R}$, so that we may replace \bar{J} with j in (A.5). Since $|J(t)| \leq |j(t)| + \max\{|j(0)|, |j(T)|\}(T+2)/T$ and $|\dot{J}(t)| \leq |dj(t)/dt| + (1/T) \int_1^{T+1} |dj(s)/ds| ds$, we obtain $\|J\|_V \leq 3\|j\|_V \leq \mu 3\bar{c}/4$. We have finally proved that Φ maps D into itself (in fact into $B_{3\bar{c}\mu/4}$).

Now we must prove that Φ is a contraction. Φ is differentiable and for $z = (v, w) \in D$, $(D\Phi(z)[h, g])(t) = (r(t), s(t))$, r and $s : [-1, T+1] \rightarrow \mathbb{R}$ being defined by

$$\begin{aligned} \ddot{r}(t) &= a_1(t).h(t) + b_1(t)g(t), \quad r(0) = r(T) = 0, \\ L(s)(t) &= a_2(t).h(t) + b_2(t)g(t), \quad s(0) = s(T) = 0, \end{aligned} \quad (\text{A.11})$$

where

$$\begin{aligned} a_1(t) &= -\mu \partial_{\varphi\varphi} f(\omega t + \varphi^+ + v(t), Q_T(t) + w(t), t + \theta^+), \\ b_1(t) &= -\mu \partial_{\varphi q} f(\omega t + \varphi^+ + v(t), Q_T(t) + w(t), t + \theta^+), \quad a_2(t) = -b_1(t), \\ b_2(t) &= \cos(Q_T(t) + w(t)) - \cos Q_T(t) + \mu \partial_{qq} f(\omega t + \varphi^+ + v(t), Q_T(t) + w(t), t + \theta^+). \end{aligned}$$

By the same arguments as above $(A, B) \in V_1 \times V$ (where $V_1 := C^1(I, \mathbb{R}^{d^2})$) defined by

$$\ddot{A}(t) = a_1(t), \quad A(0) = A(T) = 0, \quad \ddot{B}(t) = b_1(t), \quad B(0) = B(T) = 0$$

satisfy $\|A\|_{V_1} + \|B\|_V \leq c_{11}\bar{c}\mu$ ($\|\cdot\|_{V_1}$ being defined in the same way as $\|\cdot\|_V$).

Using an integration by parts, we can derive from (A.11) and the bound on $\|A\|_{V_1} + \|B\|_V$ that

$$|\dot{r}(t)| \leq c_{12}\bar{c}\frac{\mu}{\beta} \left[\left(\frac{1 + c_1\mu T^2}{\beta^2} \|h\|_V + \|g\|_W \right) + T \left(\frac{\|h\|_V}{\beta} + \|g\|_W \right) \right]. \quad (\text{A.12})$$

Therefore, for C_0 small enough, $|\beta\dot{r}(t)| \leq 1/8 \max\{\|h\|_V, \|g\|_W\}$. We derive also from (A.12) that

$$|r(t)| \leq c_{13}\bar{c} \left[\frac{\mu T}{\beta^3} + \frac{c_1\mu^2 T^3}{\beta^3} + \frac{\mu T^2}{\beta^2} \right] \max\{\|h\|_V, \|g\|_W\},$$

which yields

$$\beta^2(1 + c_1\mu T^2)^{-1}|r(t)| \leq c_{14}\bar{c}(\mu T/\beta + (1/c_1)) \max\{\|h\|_V, \|g\|_W\} \leq \max\{\|h\|_V, \|g\|_W\}/8,$$

provided C_0 is small enough and c_1/\bar{c} is large enough. Finally we obtain that $\|r\|_V \leq \max\{\|h\|_V, \|g\|_W\}/4$.

Using the properties of L and the fact that

$$|a_2(t).h(t) + b_2(t)g(t)| \leq c_{15}\mu(1 + c_1\mu T^2)/\beta^2 \|h\|_V + c_{15}(|w(t)| + \mu)\|g\|_W$$

we easily derive $\|s\|_W \leq \max\{\|h\|_V, \|g\|_W\}/4$ (again provided that C_0 , more precisely $C_0 c_1$ is small enough). We have proved that for a good choice of \bar{c}, c_1, C_0 , $\|D\Phi(z)[h, g]\| \leq \|(h, g)\|/2$ for $z \in D$. Hence Φ is a contraction. As a result, it has a unique fixed point z_λ in D (which in fact belongs to $B_{3\bar{c}\mu/4}$). This proves existence.

Now there remains to prove that $\varphi_{\mu,\lambda}(t), q_{\mu,\lambda}(t)$ are C^1 functions of (λ, t) . Let (θ_0^+, θ_0^-) be fixed with $T_0 := \theta_0^- - \theta_0^+$ and let $\Lambda = \{\lambda \mid |\theta^+ - \theta_0^+| \leq 1/4, \quad |\theta^- - \theta_0^-| \leq 1/4\}$. For $\lambda \in \Lambda$ $I_0 := [-1/2, T_0 + 1/2] \subset [-1, \theta^- - \theta^+ + 1]$, hence the restrictions v_λ^0 and w_λ^0 of v_λ and w_λ to I_0 are well defined.

Let $V_0 \times W_0 := C^1(I_0, \mathbb{R}^n) \times C^1(I_0, \mathbb{R})$ be endowed with the norm $\|\cdot\|_0$ as defined in (A.6). Define $\Psi : \Lambda \rightarrow V_0 \times W_0$ by $\Psi(\lambda) = z_\lambda^0$. We shall justify briefly that Ψ is differentiable and that $\|D\Psi\| \leq c_{16}\mu$. z_λ^0 is the unique solution in $B_{\bar{c}\mu}$ of (A.4) (with $T = \theta^- - \theta^+$), which is equivalent to $(v_\lambda, w_\lambda) = \Phi(z_\lambda; \theta^+, \theta^-, \varphi^+, \varphi^-, \mu)$, where $\Phi : B_{\bar{c}\mu} \times \Lambda \times (0, \mu_2) \rightarrow V_0 \times W_0$ is smooth. Now, by the previous step, $\|D_z \Phi\| \leq 1/2$ everywhere, so that $I - D_z \Phi$ is invertible. Therefore, by the Implicit Function Theorem, Ψ is C^1 . This proves that $(\lambda, t) \mapsto \varphi_{\mu, \lambda}(t)$ (resp. $(\lambda, t) \mapsto q_{\mu, \lambda}(t)$) and $(\lambda, t) \mapsto \dot{\varphi}_{\mu, \lambda}(t)$ (resp. $(\lambda, t) \mapsto \dot{q}_{\mu, \lambda}(t)$) have continuous partial derivatives w.r.t. λ in the set $\{(\lambda, t) \mid -1/2 + \theta^+ < t < 1/2 + \theta^-\}$, and by the standard theory of differential equations, these partial derivatives have continuous extensions on $\{(\lambda, t) \mid -1 + \theta^+ < t < 1 + \theta^-\}$. Finally, by (11.1), $\ddot{\varphi}_{\mu, \lambda}$ and $\ddot{q}_{\mu, \lambda}$ depend continuously on (λ, t) .

B Proof of Theorem 13.2.

In order to prove Theorem 13.2 we need a preliminary lemma. Observe that Λ_R^* is a finite set which is symmetric with respect to the origin. Hence, if it is not empty there exists $p \in \Lambda_R^*$ such that $p \cdot \Omega = \alpha(\Lambda, \Omega, R)$.

Lemma B.1 *Assume that $\Lambda_R^* \neq \emptyset$ and let $p \in \Lambda_R^*$ be such that $p \cdot \Omega = \alpha := \alpha(\Lambda, \Omega, R)$. Assume moreover that $\alpha > 0$ and define $E := [p]^\perp$. Then $\Lambda_0 := \Lambda \cap E$ is a lattice of E . In addition :*

(i) $\frac{\alpha}{\beta|p|} \leq \frac{2}{R}$, where $\beta = \inf\{|q \cdot \Omega|; q \in (\Lambda_0)^*_{\sqrt{3}R/2}\}$, $(\Lambda_0)^* = \{q \in E \mid \forall x \in \Lambda_0 \ q \cdot x \in \mathbb{Z}\}$.

In particular $\alpha \leq 2\beta$.

(ii) $\alpha(\Lambda, \Omega, \sqrt{7}R/2) \leq \beta$.

PROOF. Since Λ is a lattice, it is not contained in E . Hence $p \cdot \Lambda$ is a non trivial subgroup of \mathbb{Z} , $p \cdot \Lambda = m\mathbb{Z}$ for some integer $m \geq 1$, which implies that $p/m \in \Lambda^*$. But $p/m \cdot \Omega = \alpha/m$ and $|p/m| \leq R$, hence by the definition and the positivity of α , $m = 1$. As a result there exists $\bar{x} \in \Lambda$ such that $p \cdot \bar{x} = 1$. Obviously $\Lambda_0 + \mathbb{Z}\bar{x} \subseteq \Lambda$. On the other hand all $x \in \Lambda$ can be written as $x = (x \cdot p)\bar{x} + y$, where $y \in \Lambda$, $y \cdot p = 0$, i.e. $y \in \Lambda_0$. So the reverse inclusion holds and we may write $\Lambda = \Lambda_0 + \mathbb{Z}\bar{x}$. As a consequence Λ_0 is a lattice of E and

$$\Lambda^* = \{r \in \mathbb{R}^l \mid r \cdot \Lambda_0 \subset \mathbb{Z} \text{ and } r \cdot \bar{x} \in \mathbb{Z}\} = \{q + ap : q \in \Lambda_0^*, a \in \mathbb{Z} - q \cdot \bar{x}\},$$

$$\Lambda_R^* = \{q + ap : q \in \Lambda_0^*, a \in \mathbb{Z} - q \cdot \bar{x}, 0 < |q|^2 + a^2|p|^2 \leq R^2\}.$$

If $\beta = +\infty$ there is nothing more to prove. If $\beta < +\infty$, let $q \in (\Lambda_0)^*_{\sqrt{3}R/2}$ be such that $q \cdot \Omega = \beta$. Let

$$S = \{a \in \mathbb{R} : q + ap \in \Lambda_R^*\} = \{a \in \mathbb{R} : a \in \mathbb{Z} - q \cdot \bar{x}, |a| \leq (R^2 - |q|^2)^{1/2}/|p|\}.$$

Since $|q|^2 \leq 3R^2/4$, $S \supseteq S' := (\mathbb{Z} - q \cdot \bar{x}) \cap [-R/2|p|, R/2|p|]$. Hence by the definition of α , for all $a \in S'$, $|(q + ap) \cdot \Omega| = |\beta + a\alpha| \geq \alpha$, i.e. $\beta/\alpha \notin (-1 - a, 1 - a)$.

As $|p| \leq R$, the interval $[-R/2|p|, R/2|p|]$ has length ≥ 1 and must intersect $(\mathbb{Z} - q \cdot \bar{x})$. Therefore $S' \neq \emptyset$, more precisely $S' = \{u, u + 1, \dots, u + K\}$, for some integer $K \geq 0$, where $u = \inf S'$. As a result,

$$\beta/\alpha \notin \bigcup_{k=0}^K (-1 - u - k, 1 - u - k) = (-1 - u - K, 1 - u).$$

Now $S' \cap [-1/2, 1/2] \neq \emptyset$, hence $u + K \geq -1/2$ and $-1 - u - K < 0$. As a consequence $\beta/\alpha \geq 1 - u$. Since $[-R/2|p|, -R/2|p| + 1] \subseteq [-R/2|p|, R/2|p|]$ intersects $\mathbb{Z} - q \cdot \bar{x}$, $u \leq -R/2|p| + 1$. Therefore $\beta/\alpha \geq R/2|p|$, which is (i). In particular, since $|p| \leq R$, $\alpha \leq 2\beta$.

Finally there exists $a \in [-1, 0) \cap (\mathbb{Z} - q \cdot \bar{x})$; $q + ap \in \Lambda^*$, and $|q + ap|^2 = |q|^2 + a^2|p|^2 \leq 3R^2/4 + R^2 = 7R^2/4$. Hence $q + ap \in \Lambda_{\sqrt{7}R/2}^*$. We have $|(q + ap) \cdot \Omega| = |\beta + a\alpha| \leq \beta$, because $-1 \leq a \leq 0$ and $\alpha \leq 2\beta$. This proves (ii). ■

Now we turn to the proof of Theorem 13.2 We first prove that the statement is true for $l = 1$, with $a_1 = 1/2$. Here $\Lambda = \lambda_0 \mathbb{Z}$ for some $\lambda_0 > 0$, and $\Lambda^* = (\lambda_0)^{-1} \mathbb{Z}$. We can assume without loss of generality that $\Omega > 0$. If $\lambda_0 < 2\delta$, then for all $x \in \mathbb{R}$, $d(x, \Lambda) < \delta$. Hence $T(\Lambda, \Omega, \delta) = 0$.

If $\lambda_0 \geq 2\delta$, then it is easy to see that $T(\Lambda, \Omega, \delta) = (\lambda_0 - 2\delta)/\Omega \leq \lambda_0/\Omega$. On the other hand, $1/\lambda_0 \in \Lambda_{1/2\delta}^*$ and $\alpha(\Lambda, \Omega, 1/(2\delta)) = \Omega/\lambda_0$. The result follows.

Now we assume that the statement holds true up to dimension $l - 1$ ($l \geq 2$). We shall prove it in dimension l .

Fix $R > 0$ and define $\delta_R = (4a_{l-1}^2/3 + 4)^{1/2}/R$. We claim that:

(a) If $\Lambda_R^* = \emptyset$ then $T(\Lambda, \Omega, \delta_R) = 0$.

(b) If $\Lambda_R^* \neq \emptyset$, let $p \in \Lambda_R^*$ be such that $p \cdot \Omega = \alpha := \alpha(\Lambda, \Omega, R)$, and define β as in lemma B.1. Then

$$T(\Lambda, \Omega, \delta_R) \leq \max\{\alpha^{-1}, \beta^{-1}\}.$$

Postponing the proof of (a) and (b), we show how to define a_l . In the case (b), by lemma B.1 (ii), $T(\Lambda, \Omega, \delta_R) \leq \alpha(\Lambda, \Omega, \sqrt{7}R/2)^{-1}$. This estimate obviously holds in the case (a) too. Hence for all $R > 0$,

$$T(\Lambda, \Omega, (4a_{l-1}^2/3 + 4)^{1/2}/R) \leq \alpha(\Lambda, \Omega, \sqrt{7}R/2)^{-1}.$$

As a consequence, the statement of Theorem 13.2 holds with $a_l = (\sqrt{7}(4a_{l-1}^2/3 + 4)^{1/2}/2)$.

There remains to prove (a) and (b). First assume that $\Lambda_R^* = \emptyset$. Let $p \in \Lambda^* \setminus \{0\}$ be such that for all $p' \in \Lambda^* \setminus \{0\}$, $|p| \leq |p'|$. Then $|p| > R$. Let E, Λ_0 be defined from p as in lemma B.1.

Arguing by contradiction, we assume that $(\Lambda_0)^*_{\sqrt{3}R/2} \neq \emptyset$. By the same arguments as previously there exist $q \in (\Lambda_0)^*_{\sqrt{3}R/2}$ and $a \in [-1/2, 1/2]$ such that $q + ap \in \Lambda^*$. But $|q + ap|^2 = |q|^2 + a^2|p|^2 \leq (3/4)R^2 + |p|^2/4 < |p|^2$ and this contradicts the definition of p . Hence $(\Lambda_0)^*_{\sqrt{3}R/2} = \emptyset$ and by the iterative hypothesis, all point of E lies at a distance from Λ_0 less than $2a_{l-1}/\sqrt{3}R$.

From the proof of lemma B.1, there exists $\bar{x} \in \Lambda$ such that $p \cdot \bar{x} = 1$ and $\Lambda = \Lambda_0 + \mathbb{Z}\bar{x}$. Therefore for all $x \in \mathbb{R}^l$, there is $x' \in x + \Lambda$ such that $|x' \cdot p| \leq 1/2$. This implies that $d(x', E) \leq 1/(2|p|) \leq 1/(2R)$ and hence that $d(x', \Lambda_0) \leq (4a_{l-1}^2/3 + 1/4)^{1/2}/R \leq \delta_R$. Hence the distance from any point of \mathbb{R}^l to Λ is not greater than δ_R . This completes the proof of (a).

Next assume that $\Lambda_R^* \neq \emptyset$ and let p be as in lemma B.1. Define α and β in the same way as in lemma B.1. Let $x \in \mathbb{R}^l$. Again $\Lambda = \Lambda_0 + \mathbb{Z}\bar{x}$ for some $\bar{x} \in \Lambda$ such that $p \cdot \bar{x} = 1$, hence there exists $x' \in x + \Lambda$ such that $p \cdot x' \in [0, 1)$. We have

$$x' = y + \frac{w}{|p|^2}p, \quad \Omega = U + \frac{\alpha}{|p|^2}p,$$

with $y, U \in E = [p]^\perp$, $w = p \cdot x' \in [0, 1)$. We shall assume that $\alpha > 0$ (if $\alpha = 0$, there is nothing to prove). Let $\bar{t} = w/\alpha$, and consider the time interval defined by

$$J = [0, 1/\beta] \text{ if } \bar{t} < 1/\beta, \quad J = [\bar{t} - 1/\beta, \bar{t}] \text{ if } \bar{t} \geq 1/\beta.$$

$J \subset [0, \max\{1/\beta, 1/\alpha\}]$, and it is enough to prove that there exists $t \in J$ such that $d(x', t\Omega + \Lambda_0) \leq \delta_R$. The length of J is not less than $1/\beta$. Hence by the iterative hypothesis, there exists $t \in J$ such that $d(y, tU + \Lambda_0) \leq 2a_{l-1}/(\sqrt{3}R)$ (notice that for all $q \in \Lambda_0^*$, $q \cdot U = q \cdot \Omega$, so that the linear flow (tU) creates a $2a_{l-1}/(\sqrt{3}R)$ -net of E/Λ_0 in time β^{-1}). We have

$$d(x', t\Omega + \Lambda_0)^2 = \left(\frac{(t - \bar{t})\alpha}{|p|}\right)^2 + d(y, tU + \Lambda_0)^2 \leq \left(\frac{\alpha}{\beta|p|}\right)^2 + \frac{4a_{l-1}^2}{3R^2}.$$

Hence, by lemma B.1 (i), $d(x', t\Omega + \Lambda_0) \leq (4a_{l-1}^2/3 + 4)^{1/2}/R$. This completes the proof of (b).

References

- [1] A. Ambrosetti, M. Badiale: *Homoclinics: Poincaré-Melnikov type results via a variational approach*, Annales I. H. P. - Analyse nonlin., vol. 15, n.2, 1998, pp. 233-252.
- [2] A. Ambrosetti, M. Berti: *Homoclinics and complex dynamics in slowly oscillating systems*, D.C.D.S. vol. 4, n. 3.,1998
- [3] A. Ambrosetti, V. Coti Zelati: *Multiple homoclinic orbits for a class of conservative systems*, Rend. Sem. Mat. Univ. Padova, vol. 89, (1993), p. 177-194
- [4] A. Ambrosetti, V. Coti Zelati, I. Ekeland: *Symmetry breaking in Hamiltonian systems*, J. Diff. Equat. 67, (1987), p. 165-184
- [5] A. Ambrosetti, A. Malchiodi: *On the symmetric scalar curvature problem on S^n* , J. Differential Equations 170 (2001), no. 1, 228–245.
- [6] A. Ambrosetti, A. Malchiodi, S. Secchi: *Multiplicity results for some nonlinear Schrödinger equations with potentials*, Arch. Ration. Mech. Anal. 159 (2001), no. 3, 253–271.
- [7] S. Angenent: *A variational interpretation of Melnikov's function and exponentially small separatrix splitting*, Lecture notes of the London Math. Soc., Symplectic Geometry, ed. D. Salamon
- [8] V.I. Arnold (editor): *Encyclopedia of Mathematical Sciences, Dynamical Systems III*, Springer-Verlag 3 (1988)
- [9] V.I. Arnold: *Small divisor problems in classical and Celestial Mechanics*, Russian Math. Survey 18, No. 6, 85–192 (1963)
- [10] V.I. Arnold: *Instability of dynamical systems with several degrees of freedom*, Sov. Math. Dokl. 6, 1964, pp. 581-585.
- [11] G. Benettin, A.M. Cherubini, F. Fassò: *Regular and chaotic motions of the fast rotating rigid body: a numerical study*. In print in Discete and Continuous Dynamical Systems B.
- [12] G. Benettin, F.Fassò: *Fast rotations of the symmetric rigid body: a general study by Hamiltonian perturbation theory. Part I*. Nonlinearity 9, 137–186 (1996).
- [13] G. Benettin, F. Fassò, M. Guzzo: *Fast rotations of the symmetric rigid body: a study by Hamiltonian perturbation theory. Part II, Gyroscopic rotations*. Nonlinearity 10, 1695–1717 (1997).
- [14] G. Benettin, F. Fassò, M. Guzzo: *Nekhoroshev-stability of L_4 and L_5 in the spatial restricted three-body problem*, Regul. Chaotic Dyn. 3, no. 3, 56–72 (1998)

- [15] G. Benettin, G. Gallavotti: *Stability of motions near resonances in quasi-integrable Hamiltonian systems*, J. Stat. Phys. **44**, 293–338 (1986)
- [16] P. Bernard: *Perturbation d'un hamiltonien partiellement hyperbolique*, C. R. Acad. Sci. Paris Sér. I Math. **323** (1996)
- [17] D. Bernstein, A. Katok: *Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians*, Invent. Math. **88**, (1987), p. 225–241
- [18] M. Berti, L. Biasco, P. Bolle: *Optimal stability and instability results of a class of nearly integrable Hamiltonian systems*, to appear on Rend. Mat. Acc. Naz. Lincei.
- [19] M. Berti, L. Biasco, P. Bolle: *Drift in phase space: a new variational mechanism with optimal diffusion time*, J. Math. Pures Appl., submitted (2002)
- [20] M. Berti, P. Bolle: *Variational construction of homoclinics and chaos in presence of a saddle-saddle equilibrium*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **27** (1998), no. 2, 331–377
- [21] M. Berti, P. Bolle: *Homoclinics and chaotic behaviour for perturbed second order systems*, Ann. Mat. Pura Appl. (4) **176** (1999), 323–378
- [22] M. Berti, P. Bolle: *Diffusion time and splitting of separatrices for nearly integrable isochronous Hamiltonian systems*, Rend. Mat. Acc. Naz. Lincei, s. 9, vol. 11, fasc. 4, 2000, pp. 235–243.
- [23] M. Berti, P. Bolle: *A functional analysis approach to Arnold Diffusion*, Annales de l'I.H.Poincaré, analyse non-lineaire, vol. 19, (2002).
- [24] M. Berti, P. Bolle: *Fast Arnold's diffusion in systems with three time scales*, Discrete and Continuous Dynamical Systems, series A, Vol. 8, n.3, 2002, pp.795–811.
- [25] M. Berti, A. Malchiodi: *Non-compactness and multiplicity results for the Yamabe problem on S^n* , J. Funct. Anal. **180** (2001), no. 1, 210–241.
- [26] U. Bessi: *An approach to Arnold diffusion through the calculus of variations*, Non-linear Analysis T. M. A., **26**, 1996, pp. 1115–1135.
- [27] U. Bessi: *Arnold's example with three rotators*, Nonlinearity, **10**, pp. 763–781, 1997.
- [28] U. Bessi: *Arnold's diffusion with two resonances*, J. Diff. Equ. **137**, (1997), 211–239
- [29] U. Bessi: *An analytic counterexample to KAM Theorem*, Erg. Theo. Dyn. Syst. **20**, (2000), 317–333
- [30] U. Bessi, L. Chierchia, E. Valdinoci: *Upper Bounds on Arnold Diffusion Time via Mather theory*, J. Math. Pures Appl. vol. **80**, 1, 2001, pp. 105–129.

- [31] U. Bessi, V. Semijopuva: *Minimal orbits close to periodic frequencies*, Comm. Math. Helv. 73, N. 4, (1998), 516-547
- [32] L. Biasco, L. Chierchia: *Nekhoroshev stability for the D'Alembert problem of Celestial Mechanics*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur., to appear
- [33] L. Biasco, L. Chierchia: *On the stability of some properly-degenerate Hamiltonian systems*, Discrete and Continuous Dynamical Systems, series A, to appear
- [34] L. Biasco, L. Chierchia: *Effective Hamiltonian for the D'Alembert Planetary model near a spin/orbit resonance*, Celestial Mech. Dynam. Astronom., to appear
- [35] L. Biasco, L. Chierchia: *Exponential stability for the resonant D'Alembert model of Celestial Mechanics*, Comm. Math. Phys., submitted (2002)
- [36] L. Biasco, L. Chierchia, D. Treschev: *Total stability of properly-degenerate Hamiltonian systems with two degrees-of-freedom*, preprint 2002
- [37] S. V. Bolotin: *The existence of homoclinic motions*, Moscow Univ. Math. Boll. 38, 1983, p. 117-123
- [38] S. V. Bolotin: *Motions doubly asymptotic to invariant tori in the theory of perturbed Hamiltonian systems*, P.M.M. U.S.S.R., Vol 54, N° 3, (1990) p. 412-417
- [39] S. V. Bolotin: *Variational methods of constructing chaotic motions in rigid-body dynamics*, J. Appl. Math. Mech. 56, 2, 1992, p. 198-205
- [40] S. V. Bolotin: *Homoclinic orbits to invariant tori of Hamiltonian systems*, Amer. Math. Soc. Transl. (2), Vol 168, (1995) p. 21-90
- [41] S. V. Bolotin: *Homoclinic trajectories of invariant sets of Hamiltonian systems*, 4, NoDEA, (1997), 359-389
- [42] S. V. Bolotin: *Infinite number of homoclinic orbits to hyperbolic invariant tori of Hamiltonian systems*, Regular and Chaotic Dyn. (2000)
- [43] S. V. Bolotin, P. Negrini: *Variational criteria for nonintegrability*, Russian J. of Math. Phy., 5, 415-439, 1996
- [44] S. V. Bolotin, P. Rabinowitz: *A variational construction of chaotic trajectories for a Hamiltonian system on a torus*, Boll. U.M.I. Sez. B, Artic. Ric. Mat. (8)1, N. 3, (1998), 541-570
- [45] S. V. Bolotin, D. V. Treshev: *Unbounded growth of energy in nonautonomous Hamiltonian systems*, 12, Nonlinearity, (1999), 365-387
- [46] S. V. Bolotin, D. V. Treshev: *Remarks on the definition of hyperbolic tori of Hamiltonian systems*, Regular and Chaotic Dyn. 5, N. 2, (2000), 139-156

- [47] J. Bourgain, F. Golse, B. Wennberg: *On the distribution of free path lengths for periodic Lorentz gas*, Comm. Math. Phys., vol. 190, 1998, pp. 491-508.
- [48] L. Chierchia: *Arnold diffusion for nearly-integrable analytic Hamiltonian systems*, Proceedings of the workshop "Variational and local methods in the study of Hamiltonian systems", A. Ambrosetti, G.F. Dell'Antonio eds., Trieste, 1994.
- [49] L. Chierchia, G. Gallavotti: *Drift and diffusion in phase space*, Ann. Inst. Henri Poincaré, Phys. Théor., **60**, 1-144 (1994). *Erratum*, Ann. Inst. Henri Poincaré, Phys. Théor., **68**, no. 1, 135 (1998)
- [50] L. Chierchia, E. Valdinoci: *A note on the construction of Hamiltonian trajectories along heteroclinic chains*, Forum Math. 12 (2000), no. 2, 247-255.
- [51] B.V. Chirikov, Zaslavskij: *Stochastic instability of non-linear oscillations*, Usp. Fiz. Nauk **105**, no. 1,2-40 (1971).
- [52] V. Coti Zelati, I. Ekeland, E. Séré: *A variational approach to homoclinic orbits in Hamiltonian systems*, Math. Annalen **288** (1990), 133-160
- [53] V. Coti Zelati, P. Rabinowitz: *Homoclinic orbits for second order Hamiltonian systems possessing super-quadratic potentials*, J. Amer. Math. Soc. 4, (1991), p. 693-727
- [54] V. Coti Zelati, P. Rabinowitz: *Multibump periodic solutions of a family of Hamiltonian systems*, Topol. Meth. Nonlin. Anal. 4, (1995), p. 31-57
- [55] J. Cresson: *Conjecture de Chirikov et optimalité des exposants de stabilité du théorème de Nekhoroshev*, preprint univ. Besançon.
- [56] J. Cresson, C. Guillet: *Periodic orbits and Arnold diffusion*, to appear on DCDS
- [57] J. Cresson: *The transfer lemma for Graff tori and Arnold diffusion time*, Discrete Contin. Dynam. Systems 7 (2001), no. 4, 787-800
- [58] A. Delshams, P. Gutierrez: *Effective stability and KAM theory* J. Differential Equations 128, no. 2, 415-490 (1996)
- [59] A. Delshams, P. Gutierrez: *Splitting and Melnikov potentials in Hamiltonian systems*, Hamiltonian systems and celestial mechanics HAMSYS-98, Proceedings of the 3rd international symposium, Pátzcuaro, Michoacán, Mexico, December 1998 Delgado (ed)
- [60] A. Delshams, V. G. Gelfreich, A. Jorba, T. M. Seara: *Exponentially small splitting under fast quasi-periodic forcing*, Comm. Math. Phys. 189, (1997), 35-71
- [61] A. Delshams, V. G. Gelfreich, A. Jorba, T. M. Seara: *Splitting of separatrices for (fast) quasi-periodic forcing*, Hamiltonian systems with three or more degrees of freedom, 533 NATO ASI Ser. C: Math. Phys. Sci. (1999) 367-371 Kluwer

- [62] A. Delshams, R. de la Llave, T. M. Seara: *A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by a potential of generic flows of \mathbb{T}^2* , Commun. Math. Phys. 209, (2000), 353-392
- [63] A. Delshams, T. M. Seara: *An asymptotic expression for the splitting of separatrices of the rapidly forced pendulum*, Comm. Math. Phys. 150 (1992), no. 3, 433-463
- [64] G. Forni, J. Mather: *Action minimizing orbits in Hamiltonian systems*, Transition to Chaos in Classical and Quantum Mechanics, CIME, Montecatini Terme, Italy, July 1991, S. Graffi (ed.), Springer-Verlag
- [65] G. Gallavotti: *The elements of mechanics*. Springer, 1983.
- [66] G. Gallavotti: *Arnold's diffusion in Isochronous Systems*, Mathematical Physics, Analysis and Geometry 1, 1999, pp. 295-312.
- [67] G. Gallavotti, G. Gentile, V. Mastropietro: *Separatrix splitting for systems with three time scales*, Commun. Math. Phys. 202, 197-236 (1999)
- [68] G. Gallavotti, G. Gentile, V. Mastropietro: *Melnikov approximation dominance. Some examples*, Rev. Math. Phys., 11, 451-461 (1999)
- [69] G. Gallavotti, G. Gentile, V. Mastropietro: *Hamilton-Jacobi equation and existence of heteroclinic chains in three time scale systems*, Nonlinearity, 13, 323-340 (2000)
- [70] V. Gelfreich: *A proof of the exponentially small transversality of the separatrices for the standard map*, Comm. Math. Phys., 201, (1999), 155-216.
- [71] V. Gelfreich: *Separatrix splitting for a high-frequency perturbation of the pendulum*, Russ. J. Math. Phys. vol. 7, n. 1, (2000) 48-71
- [72] M. Guzzo: *Nekhoroshev stability of quasi-integrable degenerate Hamiltonian Systems*. Regular and Chaotic Dynamics 4, 78-102 (1999).
- [73] P. S. de LaPlace: *Mécanique Céleste*, tome II, livre 5, ch. I, 1799, English Translation by N. Bowditch, reprinted by AMS Chelsea Publishing, 1966
- [74] V. F. Lazutkin: *Splitting of separatrices for the Chirikov's standard map*, preprint VINITI 6372/84, 1984
- [75] P. Lochack: *Arnold diffusion; a compendium of remarks and questions*, Proc. NATO's Adv. Stu. Inst. 3-D Ham. Syst. Simó (Ed.), (1996), Plenum Press
- [76] P. Lochack: *Effective speed of Arnold diffusion and small denominators*, Phys. Lett. A 143, (1990), p. 39-42
- [77] P. Lochack: *Canonical perturbation theory via simultaneous approximation*, Russ. Math. Surv. 47, (1992), 57-133

- [78] P. Lochack: *Hamiltonian perturbation theory: periodic orbits, resonances and intermittency*, Nonlinearity, 6, (1993), 885-904
- [79] P. Lochack, J. P. Marco, D Sauzin: *On the splitting of invariant manifolds in multi-dimensional near-integrable Hamiltonian Systems*, preprint 1999
- [80] J. P. Marco: *Transitions le long des chaînes de tores invariants pour les systèmes hamiltoniens analytiques*, Annales I. H. P., vol. 64, 1995, pp. 205-252
- [81] J. Mather: *Minimal action measure for positive-definite Lagrangian systems*, Math. Phys. 9th Int. Congr., Swansea/UK 1988, 466-468 (1989)
- [82] J. Mather: *Differentiability of the minimal average action as a function of the rotation number*, Bol. Soc. Bras. Mat., Nova Ser. 21, N. 1, (1990), 59-70
- [83] J. Mather: *Variational construction of orbits of twist diffeomorphism*, J. Am. Math. Soc. 4, N. 2, (1991), 207-263 -
- [84] J. Mather: *Action Minimizing Invariant Measures for Positive Definite Lagrangian Systems*, Math. Z., **207** (1991), 169
- [85] J. Mather: *Variational Construction of Connecting Orbits*, Ann. Ist. Fournier, **43** (1993), 1349-1386
- [86] J. Mather: Seminar on "Arnold diffusion" given in the conference of "Dynamishe Systeme" in Oberwolfach, July 2001
- [87] J. Mather: Seminar on "Arnold diffusion" given in Scuola Normale Superiore di Pisa, during the "Research Trimester on Dynamical Systems" February-April 2002
- [88] J. Moser: *Stable and random motions in dynamical systems, with special emphasis on Celestial Mechanics*, Ann. Math. Studies, **77** Princeton University Press (1973)
- [89] A. I. Nejshtadt: *Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions* Prikl. Mat. Mekh. **45**, No. 6, 1016-1025 (1981) (Russian); English trans.: J. Appl. Math. Mech. **45**, No. 6, 766-772 (1982)
- [90] N. N. Nekhoroshev: *An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems*, I, Usp. Mat. Nauk. **32** 5-66 (1977); Russ. Math. Surv. **32**, 1-65 (1977)
- [91] N. N. Nekhoroshev: *An exponential estimate of the time of stability of nearly-integrable Hamiltonian systems*, II, Tr. Semin. Petrovsk. **5**, pp. 5-50, 1979; In: Oleinik, O.A. (ed) Topics in Modern Mathematics, Petrovskii Semin., no. 5. New York: Consultant Bureau 1985.
- [92] H. Poincaré: *Les Methodes Nouvelles de la Mechanique Celeste*, Gauthier Villars, Paris (1892)

- [93] J. Pöschel: *Nekhoroshev estimates for quasi-convex Hamiltonian Systems* Math. Zeitschrift **213**, 187–216 (1993)
- [94] M. Procesi: *Estimates on Hamiltonian splittings. Tree techniques in the theory of homoclinic splitting and Arnold diffusion for a-priori stable systems*, PhD. Thesis, Università degli Studi “La Sapienza”, Roma (Italy), (2001)
- [95] M. Procesi: *Arnold diffusion for systems with one fast frequency*, preprint 2002.
- [96] D. Sauzin: *A new method for measuring the splitting of invariant manifolds*, Ann. Sci. Ecole Norm. Sup. (4) **34** (2001), no. 2, 159–221
- [97] E. Séré: *Existence of infinitely many homoclinic orbits in Hamiltonian systems*, Math. Zeit. **209** (1992), 27–42
- [98] E. Séré: *Looking for the Bernoulli shift*, Ann. I.H.P., Anal. Nonlin. **10** (1992), 27–42
- [99] D. V. Treshev: *The mechanism of destruction of resonance tori of Hamiltonian systems*, Math. USSR Sbornik **68**, (1991) 181–203
- [100] D. V. Treshev: *Hyperbolic tori and asymptotic surfaces in Hamiltonian systems*, Russ. J. Math. Phys. **2**(1), (1994), 93–110
- [101] D. V. Treshev: *Separatrix splitting from the point of view of symplectic geometry*, Math. Notes **61**, (1997), 744–757
- [102] D. V. Treshev: *Multidimensional symplectic separatrix maps*, preprint
- [103] D. V. Treshev: *Trajectories in a neighborhood of asymptotic surfaces of a priori unstable Hamiltonian systems*, preprint
- [104] E. Valdinoci: *Estimates for non-resonant normal forms in Hamiltonian perturbation theory* J. Statist. Phys. **101** (2000), no. 3–4, 905–919.
- [105] E. Valdinoci: *Families of whiskered tori for a-priori stable/unstable Hamiltonian systems and construction of unstable orbits* Math. Phys. Electron. J. **6** (2000), Paper 2, 31
- [106] E. Valdinoci: *Times of diffusion in Hamiltonian systems* Proceedings of the 2000 Topology and Dynamics Conference (San Antonio, TX). Topology Proc. **25** (2000), Spring, 309–320.

