DRIFT AND DIFFUSION IN PHASE SPACE

Luigi Chierchia[1], Giovanni Gallavotti[2]

Abstract: *The problem of stability of the action variables (i.e. of the adiabatic invariants) in perturbations of completely integrable (real analytic) hamiltonian systems with more than two degrees of freedom is considered. Extending the analysis of* [A]*, we work out a general quantitative theory, from the point of view of* dimensional analysis*, for* a priori unstable *systems (i.e. systems for which the unperturbed integrable part possesses separatrices), proving, in general, the existence of the so–called Arnold's diffusion and establishing upper bounds on the time needed for the perturbed action variables to* drift *by an amount of* $O(1)$*.*
The above theory can be extended so as to cover cases of a priori stable systems *(i.e. systems for which separatrices are generated near the resonances by the perturbation). As an example we consider the "D'Alembert precession problem in Celestial Mechanics" (a planet modelled by a rigid rotational ellipsoid with small "flatness"* η*, revolving on a given Keplerian orbit of eccentricity* $e = \eta^c$*,* $c > 1$*, around a fixed star and subject only to Newtonian gravitational forces) proving in such a case the existence of Arnold's drift and diffusion; this means that there exist initial data for which, for any* $\eta \neq 0$ *small enough, the planet changes, in due (*η*–dependent) time, the inclination of the precession cone by an amount of* $O(1)$*. The homo/heteroclinic angles (introduced in general and discussed in detail together with homoclinic splittings and scatterings) in the D'Alembert problem are not exponentially small with* η *(in spite of first order predictions based upon Melnikov type integrals).*

Key words: *perturbed hamiltonian systems, stability theory, Arnold's diffusion, homoclinic splitting, heteroclinic trajectories, KAM theory, whiskered tori, dimensional estimates, Celestial Mechanics, D'Alembert Equinox Precession problem.*

[1] Dip. di Matematica, II^a Università di Roma, "Tor Vergata", via della Ricerca Scientifica, 00133 Roma, Italia
[2] Dip. di Fisica, Università di Roma, "La Sapienza", P. Moro 5, 00185 Roma, Italia

CONTENTS

Appendices:

§1 Introduction and description of the results

A typical question about diffusion in phase space is the following: could the Earth axis tilt? To put the question in mathematical form we consider a model for the Earth precession, well known since D' Alembert, [L].

Let a planet $\mathcal{E}$ be a homogenous rigid body with rotational symmetry about its N-S axis and with polar and equatorial inertia moments J_3, J_1: hence with *mechanical* polar *flattening* $\eta = (J_3 - J_1)/J_3$, which is supposed to be small. Let the planet move on a keplerian orbit $t \to \vec{r}_T(t)$, with eccentricity e, about a fixed heavenly body $\mathcal{S}$ with mass $m_{\mathcal{S}}$; also e will supposed to be small and in fact we shall assume that *the eccentricity and the flattening coefficient are related by a power law*: $e = \eta^c$ for some positive constant c. Wishing to be closer to reality one could also assume that $\mathcal{E}$ had a satellite $\mathcal{M}$: what follows could be adapted to this stranger situation (in the case of the $\mathcal{E}$arth this is particularly relevant as the $\mathcal{M}$oon accounts for 2/3 of the lunisolar precession). But here, as it will be far too clear below, we are addressing a purely conceptual question and we have no pretension that our results apply directly to the solar system or subsystems thereof.

Regarding the flattening η (and hence the eccentricity e) as a (non vanishing) parameter, we consider initial conditions close to those in which the planet is rotating around its symmetry axis, at a daily angular velocity ω_D, and precessing around the normal to the orbit, at an angular velocity denoted $\omega_p \equiv -\eta\omega$, on a cone with inclination i. And we ask whether, *no matter how small the flattening coefficient η may be* (below some η_0), there is an initial condition such that, after due time, one can find the planet precessing on a cone with inclination $i' \neq i$, with i, i' fixed *a priori, independent on η*. Such a phenomenon will be called *drift* in phase space.

We have not worried, above, about fine points like the distinction between the symmetry axis of $\mathcal{E}$ and the angular momentum or the angular velocity axes: such a distinction is not a minor one and is of course relevant to a rigorous analysis of the problem which we defer to §12.

Closely related to the drift in phase space is the *diffusion*: we shall see that the same mechanism that we discuss to show the existence of drift also shows the existence of orbits along which the inclination does not increase monotonically (in average) from i to i' but rather it evolves, on a suitably large scale of time, so as to either increase or decrease the inclination by an amount $O(\varepsilon)$ according to a prefixed pattern at least for a number of time steps of order $\gg O(\varepsilon^{-1})$, for some ε small compared to η. If one chooses the initial datum randomly and with equal distribution among the initial data of the above orbits, one will see the inclination change as a brownian motion, at least as long as it takes to reach the target value i' (or its symmetric value with respect to i).

This work is a generalization of the well known example given by Arnold, [A]. The basic feature of Arnold's example was that the drift took place around invariant tori of dimension $l-1$ if l is the number of degrees of freedom of the system and that the system considered had a very special form: the tori around which the diffusion took place were explicit exact solutions of the equations of motion. This is a property which does not hold in general and a fraction of the work in this paper is devoted to a detailed construction of the tori and of the flow around them (an analysis started in [M]). Furthermore the instability of the tori is also explicit in the model in [A]. The general system, however, will be such that most of the tori will have dimension l and the unstable tori arise near resonances.

Some details of the mechanism generating unstable tori of dimension $l-1$ along which diffusion takes place may be quite involved, in general.

The point of view of this work has been to see if, starting with the ideas in the well known example of Arnold, one could develop the theory to a point to make it applicable to the above celestial problem (for which the invariant tori arise only near resonances). We felt that such a precise goal, if pursued without further simplifying hypotheses, would provide a natural selection of possible assumptions (which could, otherwise, appear as *ad hoc* to the reader).

To achieve such a goal several intermediate problems had to be solved.

1) In §2 we define precisely a class of systems that we study: it is a system of $l-1$ rotators coupled to a pendulum. Arnold's example is in this class, but not so the D' Alembert model for the Earth precession. The simplifying aspect of the systems in such a class is that it is obvious from their definition that they are unstable (the instability simply occurs near the pendulum separatix): thus we call them *a priori unstable*. A detailed theory of such systems is necessary to attack the far harder *a priori stable* systems (defined below).

2) In §3 we point out the main (easy) properties of the uncoupled (*free*) systems of a pendulum and several rotators.

3) In §4 we introduce the key notion of *diffusion path*: it is a curve in the rotator action space, along which the free rotators angular velocities form a vector with suitable diophantine properties. It will play the role of marking the projection in action space of a drifting or diffusing motion.

4) In §5 we prove that the points of the diffusion curves can be interpreted as $l-1$ dimensional invariant tori: most of them persist after the perturbation (*i.e.* the coupling between the pendulum and the rotators) is switched on. The stability of low dimensional tori has been studied in the literature by various authors: we present it from scratch because we need very detailed bounds and analyticity properties of the perturbed tori equations and a simple *normal form* for the motion of a large class ($l+1$ dimensional) of nearby points. The bounds must be general and at the same time simple enough to be applicable to the harder cases that we analyze later (like the D' Alembert model). Hence we need results stated in terms of the few really important features of the hamiltonian. We therefore proceed by identifying the relevant parameters (basically ratios of the independent time scales that govern the motions) and produce a proof in which the only ingredient is the use of the Cauchy theorem to bound the derivative of a holomorphic function by the ratio between the maximum modulus, in the considered analyticity domain, and the distance to the boundary of the analyticity domain. We call, for obvious reasons, such bounds *di-*

mensional bounds, see lemmata 1,1' of §5. The normal coordinates that we describe are a generalization of the celebrated Jacobi coordinates near the unstable equilibrium point of the pendulum (see lemma 0 of §5, and appendix 9 for a description of the classical Jacobi map).

5) In §6 we develop the perturbation theory of the asymptotic manifolds of the invariant tori constructed in §5; following Arnold, we call such manifolds *whiskers*. The theory is discussed to arbitrary order of perturbation theory: such a generality is necessary only if one has in sight applications to *a priori* stable systems (such as the celestial one of D' Alembert). Such analysis requires establishing, for the purpose of a consistency check, some remarkable *homoclinic identities*, established in appendix A12.
For the models in the class of the *a priori* unstable systems the theory to first order is sufficient and we deduce that the homoclinic angles (*i.e.* the angles between tangent vectors to the stable and to the unstable whiskers) are, no wonder, described by a tensor (that we call the *intersection tensor*) related to the Melnikov integrals, reproducing results of Melnikov which are well known, [Me].

6) In §7 we show that, given a diffusion path, if the perturbation has suitable properties (expressed in terms of some explicit condition of absence of low order resonant harmonics in the Fourier development of the perturbation at path points) then the set of points along the path representing invariant tori (for the full hamiltonian) is so dense that one can find a sequence of them spaced by an amount far smaller than the size of the homoclinic angles.

7) In §8, using the normal form described in §5 in a very essential way, we show that in the assumptions of §7 the diffusion path is *open for diffusion* and show the existence of initial conditions which evolve in time so that the projection of the motion in action space follows the diffusion path. We also find an *explicit* estimate of the time needed by the drifting motions to reach the other extreme of the diffusion path. The path is independent on the size μ of the perturbation and it is non trivial (*i.e.* not a single point) if $l > 3$ (no diffusion or drift are possible if $l = 2$ by the KAM stability).

The time it takes is of the order $O(\exp -k\mu^{-2})$: Arnold's example is covered by the theorem, but our result is less general than Arnold's one as it can be applied to diffusion paths which are segments of length of $O(1)$ but not arbitrarily placed on the action axis: this is the price that we have to pay to get concrete bounds on the drift time (and not only a finiteness result). We do not know if this restriction would also be present by using Arnold's method (*i.e.* whether Arnold's method could give, in his example, actual constructive upper bounds on the diffusion time).

8) In §9 we begin to worry about the fact that the above analysis does not cover *a priori* unstable systems in which the pendulum Lyapunov exponent (*i.e.* in physical terms the gravity acceleration), that we call here η, is not fixed but it is linked to the perturbation size (usually much smaller) that we call μ. The reason is that in such cases the first order of perturbation theory is "degenerate" in the sense that it predicts homoclinic splitting with some angles of size $O(\mu \exp -k\eta^{-1/2})$, for some $k > 0$. This leads essentially to a situation in which the first order perturbation theory is not sufficient, even to establish

the existence of the homoclinic splitting, not to speak of the existence of drift: it is well known that there are examples in which the situation does not improve by going to higher order (see *e.g.* [La2]).

In fact the problem is already quite hard in the case of a forced pendulum (*i.e.* $l = 2$) and with the rotator being a clock model, perfectly isochronous; this means that the rotator action B appears in the form of an additive term in the hamiltonian equal to ωB and the rest of the hamiltonian depends only on the pendulum coordinates I, φ and on the conjugate angle λ, "position of the clock arm". If the perturbation size is supposed $\mu = O(\eta^d)$ for some $d > 0$ the problem is non trivial (a case reducible to the ones treated in §6,§7,§8 would be if $\mu = O(\exp - c\eta^{-b})$ with $b > 1/2$: but this is, unfortunately, a case of little interest in view of the expected size of μ in the applications).

If $l > 2$ the angles are in general rather hard to describe: we find some rather implicit expressions for them, in general, but we can make use of them in the one case with $l = 3$ which motivated our work (*i.e.* the D' Alembert equinox precession model). Actually we point out an ambiguity about what one defines to be the homoclinic angles of splitting as there are at least two different interesting sets of coordinates that can be considered. To relate them we introduce the concept of homoclinic phase shift (a quite remarkable notion in itself: see item 13) below for a qualitative description of it).

In general, in the cases with an exponentially small splitting to first order, we *do not* discuss a proof of the existence of a homoclinic point: although the results that we have developed are probably sufficient for constructing a proof. The reason is not only to cut a little shorter this paper but mainly because the theory is, nevertheless, not empty: in fact we can apply it to a special but wide class of models for which the homoclinic point problem is (well known to be) exactly soluble (in the sense that one can show the existence, and locate exactly the position, of the homoclinic point). We call such class the *even models*: as the property is based on a symmetry of such hamiltonians. Many models of forced pendula fall in this class that we introduce and treat, for completeness, in §9.

9) In §10 we discuss in more detail the notion of homoclinic phase shift particularly in the case of even models with $l = 3$, in which one of the rotators is a clock and the other is "slow", *i.e.* its free angular velocity is of order η while also the pendulum gravity constant is of order η. The introduced formalism allows us to show that the phenomenon of *large homoclinic splitting* takes place even in presence of fast rotations, *as long as there is at least one slow among them*: this property holds only in systems with $l \geq 3$ (and generically it does happen, as we show) and in spite of first order (Melnikov type) computations (which predict exponentially small splittings). Some detailed calculations are performed in appendix A13 and they are interesting by themselves.

The existence of one fast rotation and other slow ones looks very special but we show in §12 that the D' Alembert precession model, which is *a priori* stable, is reducible to such a case: this is due to the extra degeneracies present in all celestial problems.

10) The actual application of the theory to even models with $l = 3$, relevant for the precession problem, requires some extra work performed in §11 and the technique is also

an illustration of a rigorous application of the usually qualitative *averaging methods*.

11) In §12 we finally study the *a priori* stable D' Alembert precession model. The original D' Alembert model took the planet orbit to be circular: in this case the model has $l = 2$ and diffusion is not possible. Therefore we take the orbit to be keplerian with eccentricity $e > 0$; this leads to a large class of models obtained by truncating the eccentricity series to order k; we study for simplicity only the case $k = 2$: the general case (k arbitrary), does not seem to offer more difficulties, except notational ones. The work having been organized in order to treat this case, the discussion is rather simple.

We choose in our example as diffusion path a line which has the physical interpretation of a $1 : 2$ resonance between the "day" period and the "year" period, and is such that a motion along it has the interpretation of changing the size of the angle between the ecliptic and the angular momentum of the planet ("inclination"). We just have to check that the model can be reduced, by a suitable change of coordinates, to a $l = 3$ system of a pendulum with small gravity of order η forced by a fast clock and by a slow anisochronous rotator; the perturbation parameter is the eccentricity e of the orbit, which we have to take small with η, *e.g.* $e = \eta^c$, for some convenient $c > 0$. The model is even, in the sense of §9,§10, and the theory of §9,§11 fully applies at least to portions of $O(1)$ of the diffusion path: for many of them we thus get the existence of drift (and diffusion).

13) The notion of *homoclinic scattering* and *phase shifts* arises naturally as a byproduct of the analysis performed to describe the phase shifts occurring on the homoclinic motion and near it. Calling $\vec{\alpha}$ the rotators angular coordinates and φ the pendulum angle suppose that at some arbitrarily fixed reference angle $\varphi = \bar{\varphi}$ there is a homoclinic point at $\vec{\alpha} = \vec{\alpha}_0$. Two points starting at $t = 0, \varphi = \bar{\varphi}$, one on the stable whisker and one on the unstable whisker of some invariant torus with the same position coordinates $\vec{\alpha}$, will evolve towards the invariant torus (respectively forward and backward in time) so that their asymptotic motion gives two points which move quasi periodically keeping a time independent *phase* with respect to the homoclinic motion. It will be a function of the distance of the initial points to the homoclinic point, *i.e.* of $\vec{\alpha}$. The difference $\vec{\sigma}[\vec{\alpha}]$ between such phases evaluated at $t = \pm\infty$ will be the *phase shift*. The "scattering" will be the family of derivatives of $\vec{\sigma}[\vec{\alpha}]$ at $\vec{\alpha} = \vec{\alpha}_0$. In other words we use the homoclinic point as a gauge to fix the origin of the angles on the standard torus on which the quasi periodic motion is linear and we look at the trajectory starting on the unstable whisker at $t = -\infty$ infinitesimally close to the invariant torus and evolving into a point with $\varphi = \bar{\varphi}$ and some $\vec{\alpha}$ at $t = 0$, "jump" on the stable whisker (keeping the values of $\vec{\alpha}, \varphi$), and evolve towards the invariant torus again. The trajectory will be asymptotically lagging behind the homoclinic trajectory by an amount $\vec{\psi}$, say, at $t = -\infty$ and by an amount $\vec{\psi} + \vec{\sigma}[\vec{\alpha}]$ at $t = +\infty$. The notion of $\vec{\sigma}[\vec{\alpha}]$ is intrinsic as the coordinates on which the motion on the torus appears as linear and which are "close" to the corresponding unperturbed ones are uniquely defined.

In presence of perturbations the phase shift is a non trivial function of the distance to the homoclinic point. We define analytically the phase shifts in §10 and briefly discuss them in §10 and, appendix A11, how they are related to the homoclinic splitting.

We present all details in a self contained way. Some of the details are, however, exposed in a series of appendices. Some of the appendices also contain classical results not so easy to find in the literature in the form in which we need them. Some, (very few), of them are not really necessary but they are reported because they clarify conceptual and historical aspects of the problem (namely the statement of Nekhorossev theorem (§A1), the D' Alembert precession theory for the Earth (§A6,§A7), the Jacobi map (§A9)), the bounds on the homoclinic scattering (§A11) and they occupy a negligible amount of space.

Acknowledgments: We are indebted to CNR-GNFM and to the research funds of MPI for partial support. L.C. also acknowledges support for a visit to Rutgers University under grant # DMR 89–18903. G.G. is grateful to J. Lebowitz for his encouragement and for his support to this work which was partly developed at Rutgers University with the support of the grants # DMR 89–18903 during the years 1989 through 1991.

§2 A priori unstable systems. Regularity assumptions

Let $(\vec{A},\vec{\alpha})$, (I,φ) be canonical coordinates describing a mechanical system with l degrees of freedom. We suppose $\vec{A}\in V\subset R^{l-1},\vec{\alpha}\in T^{l-1}$, $I\in R^1$ and $\varphi\in T^1$, where V is the closure of some open bounded set and T^s is the s-dimensional torus. We shall regard T^s interchangeably as $[-\pi,\pi]^s$ with opposite sides identified or we regard it as C_1^s={product of s unit circles in the s-dimensional complex space C^s} via the identification $\vec{\varphi}=(\varphi_1,\ldots,\varphi_s)\in T^s\longleftrightarrow\vec{z}=(z_1,\ldots,z_s)\in C^s$ with $z_j\equiv e^{i\varphi_j}$, $(j=1,\ldots,s)$.

The *free* system will consist of $l-1$ rotators described by the angles $\vec{\alpha}$ and their conjugate momenta $\vec{A}$, and one pendulum described by the angle φ with conjugate momentum I.

The pendulum oscillates with energy:

$$P_0(I,\vec{A},\varphi)=\frac{1}{2}\frac{I^2}{J_0(\vec{A})}+g(\vec{A})^2J_0(\vec{A})(\cos\varphi-1)\tag{2.1}$$

where $J_0(\vec{A})$ is a suitable *inertia moment* and $2\pi g(\vec{A})^{-1}$ is the characteristic period of the small oscillations or, as well, $g(\vec{A})$ is the Lyapunov exponent of the unstable fixed point. We call (2.1) a *standard pendulum* hamiltonian.

The rotators will move without being affected by the pendulum oscillations. A complete example hamiltonian will be:

$$h_0=\frac{1}{2}\frac{\vec{A}^2}{R}+P_0(I,\vec{A},\varphi)\tag{2.2}$$

where R is another inertia moment.

More generally we shall consider $\vec{\alpha}$-independent hamiltonians like:

$$H_0(I, \vec{A}, \varphi, \mu) = h(\vec{A}, \mu) + P(I, \vec{A}, \varphi, \mu) \tag{2.3}$$

where P is a real analytic hamiltonian depending on a parameter μ and describing a pendulum in the sense discussed below, and $h(\vec{A}, \mu)$ will also be assumed real analytic.

To clarify what we mean by a *pendulum* hamiltonian P we recall the characteristics of the pendulum phase portrait. The isoenergy lines in (I, φ)-space with $P = E$ are closed continuous curves with topological properties that may change as E varies. The lines of separation between the regions covered by curves of the same type (*i.e.* curves which do not contain an equilibrium point and which can be deformed into each other without crossing an equilibrium point) are called separatrices and contain at least one equilibrium point, and at most finitely many (as we are only considering analytic hamiltonians).

In our case we want to allow an explicit $(\mu, \vec{A})$-dependence of P: hence the above picture is $\mu, \vec{A}$ dependent. We shall require that, for all values of $\vec{A}$ of interest, the pendulum P has a linearly unstable fixed point $I_\mu(\vec{A}), \varphi_\mu(\vec{A})$ which is the only such point on the corresponding separatrix and, furthermore, we require that $I_\mu(\vec{A}), \varphi_\mu(\vec{A})$, together with its Lyapunov exponent $g(\vec{A}, \mu)$ ($\neq 0$ by assumption), depend analytically on $\vec{A}, \mu$.

Clearly the above is a very mild restriction, only exceptionally false: it emerges from the analysis that all we really want is that in the whole range of the $\vec{A}$'s the unstable fixed point, which we select for our analysis, depends analytically on $\vec{A}, \mu$ and does not merge, as $\vec{A}, \mu$ vary, with other fixed points. We shall call the above equilibrium point a *selected unstable equilibrium point of* P.

In such a situation we shall say that (2.3) describes an *a priori* unstable free assembly of rotators witnessed in their rotations by a free pendulum with a selected unstable point of equilibrium.

It is not restrictive, under the above circumstances, to assume that the selected unstable point is the origin $I = 0, \varphi = 0$, and that its energy is $P = 0$. In fact one can always change coordinates by using the canonical map generated by the function: $(\varphi - \varphi_\mu(\vec{A}'))I' + I_\mu(\vec{A}')\sin(\varphi - \varphi_\mu(\vec{A}')) + \vec{\alpha} \cdot \vec{A}'$, *i.e.* :

$$\begin{aligned} I =& I' + I_\mu(\vec{A}')\cos(\varphi - \varphi_\mu(\vec{A}')), \quad \vec{A} = \vec{A}', \quad \varphi' = \varphi - \varphi_\mu(\vec{A}') \\ \vec{\alpha}' =& \vec{\alpha} - \big(I' + I_\mu(\vec{A}')\cos(\varphi - \varphi_\mu(\vec{A}'))\big)\partial_{\vec{A}}\varphi_\mu(\vec{A}') + \partial_{\vec{A}} I_\mu(\vec{A}')\sin(\varphi - \varphi_\mu(\vec{A}')) \end{aligned} \tag{2.4}$$

which is clearly well defined and which generates a new hamiltonian of type (2.3) which has $I = 0, \varphi = 0$ as selected unstable equilibrium point. Furthermore if $P(\vec{A}, \mu) \equiv P(0, \vec{A}, 0, \mu)$ we can always redefine P as $P - P(\vec{A}, \mu)$ by accordingly changing h: hence the requirement that also $P(0, \vec{A}, 0, \mu) = 0$ is not restrictive.

The aspects of the regularity properties that we use, motivated by the above descriptions, are as follows:

Assumption 1: *The unperturbed hamiltonian H_0 has the form (2.3) and the pendulum energy P has the origin $(I=0,\varphi=0)$ as a selected unstable equilibrium point where P takes the value 0 (for all $\vec A$ and μ in the domain of definition of H_0); the associated (non negative) Lyapunov exponent, $g(\vec A,\mu)$:*

$$g^2\equiv[\big(\partial^2_{I\varphi}P\big)^2-\partial^2_IP\partial^2_\varphi P]|_{(I,\varphi)=(0,0)} \tag{2.5}$$

is bounded away from zero as $(\vec A,\mu)$ vary in their domain of definition.

Assumption 2: *The functions h and P are real analytic in their arguments. Hence they are holomorphic in their variables in a complex domain $S_{\rho',\rho,\xi',\xi,\bar\mu}$, described by five parameters $\rho',\rho,\xi',\xi,\bar\mu>0$ as:*

$$\begin{aligned}S_{\rho',\rho,\xi',\xi,\bar\mu}=\Big\{I,\vec A,\zeta,\vec z,\mu\Big|\ &|I|\le\rho',\ \text{and there is }\vec a\in V,\ \text{for which}\\ &|A_i-a_i|\le\rho\text{ and }e^{-\xi'}<|\zeta|<e^{\xi'},\ e^{-\xi}<|z_j|<e^{\xi},\ \text{and }|\mu|\le\bar\mu\Big\}\end{aligned} \tag{2.6}$$

with $z_j\equiv e^{i\alpha_j},\zeta=e^{i\varphi}$.

Assumption 3: *The following non degeneracy conditions:*

$$\det(\partial^2_{\vec A}h)\ne0,\qquad \det(\partial^2_{(I,\varphi)}P|_{I=0,\varphi=0})\ne0,\qquad \partial_{\vec A}h\cdot(\partial^2_{\vec A}h)^{-1}\partial_{\vec A}h\ne0 \tag{2.7}$$

hold on $S_{\rho',\rho,\xi',\xi,\bar\mu}$.

Then we set:

Definition: *Hamiltonians verifying all the above assumptions 1÷3 will be briefly referred to as* regular *anisochronous* a priori *unstable free hamiltonians.*

They are called *a priori* unstable, because the instability assumption is clearly built in the free system definition.

Such hamiltonians are quite common in the theory of the resonances of anisochronous systems.

For instance consider an l degrees of freedom system with free hamiltonian h of the form $h(\vec A,B)$ in action angle coordinates $\vec A,B,\vec\alpha,\lambda$ such that the equation of the resonance is simply $\partial_Bh(\vec A,B)=0$. Suppose that $B=B(\vec A)$ is the consequent resonance surface. Then, if $\varepsilon f(\vec A,\vec\alpha,B,\lambda)$ is a perturbation, one can find canonical coordinates $(\vec A',\vec\alpha',I,\varphi)$ apt to describe the motions that take place near the resonance and in which the hamiltonian takes the form (2.3) (in square brackets in the following expression) plus a *small* correction:

$$\begin{aligned}&\Big[h_p(\vec A',I,\varepsilon)+\varepsilon G_p(I,\vec A',\varphi,\varepsilon)\Big]+\varepsilon^pf_p(I,\vec A',\vec\alpha',\varphi,\varepsilon)\\ &h_p=h(\vec A',I)+\varepsilon\overline f(\vec A',I)+O(\varepsilon^2),\qquad G_p=\tilde f(I,\vec A',\varphi)+O(\varepsilon)\end{aligned} \tag{2.8}$$

with $\overline f$ equal to the average of f over the angles $\vec\alpha$, λ and G_p equal to the average of the function $f-\overline f$ over the $\vec\alpha$ alone; here p can be fixed arbitrarily and ε is the strength of the perturbation. But, the larger p is, the harder it is to find the functions G_p, f_p and a coordinate system in which (2.8) holds and the smaller becomes the (tiny) region of phase space around the resonance surface where the new coordinates can be used to describe the motion, (this is essentially the Nekhorossev theorem, see [BG], and appendix A1).

We consider hamiltonians H which are perturbations of regular free *a priori* unstable hamiltonians H_0, defining the latter by the assumptions 1÷3 above:

$$H = H_0(I, \vec{A}, \varphi, \mu) + \mu f(I, \vec{A}, \varphi, \vec{\alpha}, \mu) \tag{2.9}$$

with f holomorphic in the domain $S_{\rho',\rho,\xi',\xi,\bar{\mu}}$, see (2.6). We shall often refer to the Fourier expansion of f in the $\vec{\alpha}$ variables, which we shall write as:

$$f(I, \vec{A}, \varphi, \vec{\alpha}, \mu) = \sum_{\vec{\nu}\in Z^{l-1}} f_{\vec{\nu}}(I, \vec{A}, \varphi, \mu)\, e^{i\vec{\nu}\cdot\vec{\alpha}} \tag{2.10}$$

The problem of phase space drift and diffusion will be posed as follows:

Diffusion problem: *Given $\vec{A}_1, \vec{A}_2$, with $H_0(0, \vec{A}_1, 0, 0) = H_0(0, \vec{A}_2, 0, 0)$ can one find for all μ small enough, but non zero, initial data close (as $\mu \to 0$) to $(0, \vec{A}_1, 0)$ in the $(I, \vec{A}, \varphi)$-variables which, in due time (μ-dependent, of course) evolve into data close to $(0, \vec{A}_2, 0)$? More bluntly can one realize a displacement of $O(1)$ in the $\vec{A}$ variables with a perturbation of order μ as small as we please?*

§3 The free system. Diffusion paths and whisker ladders

To formulate our results we need several concepts. The first is the notion of *diffusion path* on a energy level E, whose value will be kept fixed throughout this section, as well as the value $\mu = 0$.

Let $s \to \vec{A}_s$ be a curve $\mathcal{L}$, piecewise analytic in $s \in [s_1, s_2]$, joining $\vec{A}^1 = \vec{A}_{s_1}$ to $\vec{A}^2 = \vec{A}_{s_2}$, such that, using the notation in (2.3), one can find two constants τ, t for which:

1) $H_0(0, \vec{A}_s, 0, 0) = h(\vec{A}_s, 0) \equiv E$, for $s \in [s_1, s_2]$,
2) if $\vec{\omega}_s = \partial_{\vec{A}} h(\vec{A}_s, 0)$ and if we set:

$$\begin{aligned} C(s) =& \sup_{\vec{\nu}\neq\vec{0},\, \vec{\nu}\in Z^{l-1}} |\vec{\nu}|^{-\tau} |\vec{\nu}\cdot\vec{\omega}_s|^{-1} \qquad \textit{“non resonance constant”} \\ \Sigma(C) =& \{\text{set of the } s \in [s_1, s_2] \text{ such that } C(s) < C\} \end{aligned} \tag{3.1}$$

then there is a $\mathcal{L}$-dependent constant $\bar{K} > 0$ such that:

$$(s_2 - s_1)^{-1} \cdot (\text{measure of the set } \Sigma(C)) \geq (1 - \bar{K}(DC)^{-1/t}) \tag{3.2}$$

if D is the maximum of $|\partial_{\vec{A}} h(\vec{A},0)|$ in a neighborhood of the curve $\mathcal{L}$.

Definition: *if $\mathcal{L}$ is a curve with the properties 1),2) above we call it a diffusion path.*

Clearly under the genericity assumption (2.7), $\det \partial^2_{\vec{A}} h \neq 0$, a diffusion path consists of just one point if $l=2$ (because $h=E$ fixes A): no diffusion path exists between distinct points in action space, if $l=2$. For this trivial reason our results, which otherwise do not distinguish $l=2$ from $l>2$, will be occasionally uninteresting if $l=2$.

In appendix A2 we show that under the genericity assumption (2.7) the constants t,τ can be taken to be $t=l-1$ and $\tau=(l-1)^2$. But on special curves it could be possible to make better choices: for instance in §11 we discuss an application with $l=3$ in which $t=1$.

Note that the diffusion paths lie, by definition, in the space of the $\vec{A}$-variables which are the "rotators" velocities (or *fast action variables*, or *adiabatic invariants*: using the terminology borrowed from the theory of resonances mentioned in connection with (2.8),(2.9); see also appendix A1); it is a notion depending solely on the free system hamiltonian ($\mu=0$) evaluated when the *pendulum* (or *slow*, or *secular*) variables (*i.e.* (I,φ)) are set to the equilibrium position.

It is easy to see that if $l>2$ there are, under the non degeneracy conditions (2.7) many diffusion paths joining *any* two close enough points $\vec{A}^1,\vec{A}^2$ lying on a connected portion of the energy shell $h(\vec{A},0)=E$, see appendix A2. The argument is similar to the one usually invoked to prove the abundance of diophantine irrationals (see, for instance, [G]): the $l-1=2$ case is particularly easy and the condition is fulfilled by any curve with non vanishing curvature; in the case $l-1=3$ one has to consider a curve joining $\vec{A}^1,\vec{A}^2$ with nowhere vanishing curvature and torsion, *etc.*

To see the connection between the torsion and the above mentioned values of τ,t one should recall that a smooth curve $s\to\vec{a}(s)$ in d-dimensions is said to have all its $d-1$ torsion coefficients non vanishing if, for each fixed s, the first d derivatives of $\vec{a}(s)$ are linearly independent: the torsion coefficients are suitable orthogonal invariants associated with the derivatives of order higher than the first (hence their number is $d-1$).

The first non degeneracy condition of the second line of (2.7) permits us to conclude that any curve with all its $l-2$ torsion coefficients non zero verifies (3.2); the last non degeneracy condition in the second line of (2.7) implies that a curve which, in a local chart on the energy surface, has all (the $l-3$) torsion coefficients non vanishing will also have all the $(l-2)$ torsion coefficients non vanishing when it is regarded as lying on the $l-1$ dimensional action space.

The values of the exponents arise from the remark that if the curve has all torsions non zero then a codimension one plane cutting it in a point cannot have a contact of order higher than $l-1$ with the curve. Thus a layer of width δ does not contain, locally, an arc length exceeding $O(\delta^{1/(l-1)})$. Therefore the statement follows by choosing $\delta=1/(C|\vec{\nu}|^{\tau+1})$ with $(\tau+1)/(l-1)>l-1$, so that one can sum the arc lengths over $\vec{\nu}$ (as it is clearly necessary); *i.e.* the choice $\tau = (l-1)^2$, $t = l-1$ is sufficient, see appendix A2.

For every $\vec{A}$ one can define the $(l-1)$-dimensional torus invariant for the motion governed by H_0:

$$\mathcal{T}_0(\vec{A}) = \big\{\vec{A}', \vec{\alpha}', I', \varphi' \big| \, \vec{A}' = \vec{A},\, I' = 0,\, \varphi' = 0,\, \vec{\alpha}' \in T^{l-1}\big\} \tag{3.3}$$

Such tori represent data in which the $l-1$ rotators are mindlessly and freely rotating while the pendulum *stands up* in its selected unstable equilibrium position. The picture, hence the tori, is obviously unstable and in fact the tori posses stable and unstable manifolds, called *whiskers* by Arnold, [A], (for reasons that emerge as soon as one tries to make a symbolic drawing of the situation). The whiskers correspond to data in which the rotators continue to rotate freely witnessing the pendulum falling from or climbing to the equilibrium position (respectively describing the unstable or the stable whiskers) and performing one of the two *separatrix* swings. More mathematically:

$$\begin{aligned} W_{unstable}(\vec{A}) =& \big\{\vec{\alpha} \in T^{l-1},\, \mathrm{sign}\, I = \mathrm{sign}\, \varphi, P(\vec{A}, I, \varphi, 0) = 0\big\} \\ W_{stable}(\vec{A}) =& \big\{\vec{\alpha} \in T^{l-1},\, \mathrm{sign}\, I = -\mathrm{sign}\, \varphi, P(\vec{A}, I, \varphi, 0) = 0\big\} \end{aligned} \tag{3.4}$$

where, to fix the ideas, we have assumed that $I > 0$ means $\dot{\varphi} > 0$, while $I < 0$ means $\dot{\varphi} < 0$ and each separatrix swing takes place over the complete circle $\varphi \in [-\pi, \pi]$ (as in the standard pendulum case; in these cases we shall speak of "open separatrices"). Such properties may fail in some pendula (*e.g.* one of the separatrices could be contractible to a point): in these cases (3.4) has to be changed in an obvious way.

It is always true, however, that the set $P(I, \vec{A}, \varphi, 0) = 0$ will consist of two branches which will be called the *separatrix swings*: in the case of the standard pendulum they are the subsets of $W(\vec{A})$ with $I > 0$ or $I < 0$. Furthermore the following well known accident happens:

$$\begin{aligned} &W(\vec{A}) = W_{stable}(\vec{A}) \equiv W_{unstable}(\vec{A}); \\ &W(\vec{A}) \cap W(\vec{A}') = \emptyset \qquad \text{if } \vec{A} \neq \vec{A}' \end{aligned} \tag{3.5}$$

hence in the general case both sets in (3.4) will be equal and coinciding with the separatrix data. Given a diffusion path $\mathcal{L}$ we can associate to it, for $\mu = 0$, a one parameter family $s \to \mathcal{T}_0(s) \equiv \mathcal{T}_0(\vec{A}_s)$ of $(l-1)$-dimensional tori, invariant with respect to the free evolution.

The family $s \to (\mathcal{T}_0(s), W_{stable}(s), W_{unstable}(s))$ of the above tori and of their whiskers will be said to form a *whisker ladder*, *leaning* on $\mathcal{L}$; again try a drawing for the word motivation.

§4 Motion on the separatrices. Melnikov integrals

Suppose, for simplicity, a (*open*) separatrix encircling the circle, with a monotonic motion taking place on it (e.g. such that the sign of I and that of $\dot\varphi$ coincide). We shall write the parametric equations for the branch $I<0, \varphi>0$ of $W(\vec A)$ as:

$$I = i(\varphi,\vec A),\quad \varphi\in(0,2\pi),\quad \vec\alpha\in T^{l-1} \tag{4.1}$$

where i is the separatrix swing with $I<0$ (*i.e.* the branch with $I<0$ of the curve $P|_{\mu=0}=0$ through the selected unstable equilibrium point, (see §2)). In the general case (when the separatrix may be shorter than the full circle, "closed separatrix case") one cannot use φ to parameterize a full separatrix swing, *i.e.* a branch of $W(\vec A)$: one would have to use a different extra parameter to describe $W(\vec A)$ at the cost of conceptually uninteresting complications.

If $X(\vec A,\varphi,\vec\alpha) = (i(\varphi,\vec A),\vec A,\varphi,\vec\alpha)$ is the point (4.1), let us denote with the symbol $X^0(\vec A,\varphi,\vec\alpha,t)\equiv(I^0(t),\vec A,\varphi^0(t),\vec\alpha^0(t))$ the point into which $X(\vec A,\varphi,\vec\alpha)$ evolves at time t in the motion governed by the hamiltonian equations with hamiltonian (2.3) with $\mu=0$.

The (3.4),(2.3) and our choice of coordinates (in which $I=0,\varphi=0$ is the selected unstable point) imply:

$$I^0(t),\varphi^0(t) = O(e^{\mp gt})\xrightarrow[t\to\pm\infty]{}0 \tag{4.2}$$

where $g\equiv g(\vec A)\equiv g(\vec A,0)$ is the Lyapunov exponent of the selected equilibrium point, *i.e.* it is given by $g^2=-\det(\partial^2P(0,\vec A,0,0))$, where ∂^2P is the matrix of the second derivatives with respect to I,φ. Furthermore, denoting $\vec\omega(\vec A)=\partial_{\vec A}h(\vec A,0)$:

$$\begin{aligned}\vec\alpha^0(t) =&\vec\alpha+\vec\omega(\vec A)t+\int_0^t\partial_{\vec A}P(\vec A,I^0(\tau),\varphi^0(\tau),0)\,d\tau\equiv\\ \equiv&\vec\alpha+\vec\omega(\vec A)t+\vec\vartheta(t;\vec A,\varphi)\end{aligned} \tag{4.3}$$

where we have used that $P(0,\vec A,0,0)\equiv0$, by our assumptions 1÷3, §2, so that the integrand tends to zero by (4.2); the function $\vec\vartheta$ is defined by (4.3).

It is convenient to fix once and for all an origin on the separatrix corresponding to the action $\vec A$: we take it to be the point $\bar I,\bar\varphi$ with $\bar\varphi$ such that the solution $I(\vec A,\varphi)$ of the equation $P(I,\vec A,\varphi,0)=0$ for I, parameterized by $\vec A$, reaches its absolute maximum value as a function of φ. We call this point *the origin of the separatrix*. In the case in which P is a standard pendulum (*i.e.* it is given by (2.1)) the position $\bar\varphi$ is $\bar\varphi=\pi$, where the pendulum attains the maximum velocity.

Therefore we can define the asymptotic *phase shifts* $\vec\vartheta^{\pm}(\vec A)$ equal to the limits as $t\to\pm\infty$ of $\vec\vartheta(t;\vec A,\bar\varphi)$. They depend on the starting point, *i.e.* on $\bar\varphi$, which however we keep fixed as above, and on $\vec A$; their difference $\vec\vartheta(\vec A)$ is:

$$\vec\vartheta^+(\vec A)-\vec\vartheta^-(\vec A)\equiv\vec\vartheta(\vec A)=\int_{-\infty}^{\infty}\partial_{\vec A}P(I^0(\tau),\vec A,\varphi^0(\tau),0)d\tau \tag{4.4}$$

and $-\vec\vartheta(\vec A)/2$ has the geometric interpretation of the $\vec A$-gradient of the area enclosed between the considered branch ($I<0, \varphi>0$) of separatrix and the $I=0$ axis (for closed separatrices it is the $\vec A$ gradient of the area enclosed by the considered separatrix swing).

We set the following definition in terms of the above concepts:

Definition 1: *The free system rotators and pendulum are independent at $\vec A$ if $\vec\vartheta(\vec A)\equiv 0$.*

The obviously interesting case (2.1) with $\vec A$ independent $R(\vec A), g(\vec A)$ is clearly very special and it is an example of independence in the above sense. If, on the other hand, in (2.1), the functions $g(\vec A), R(\vec A)$ are not constant the *phase shifts* $\vec\vartheta(t,\vec A;\bar\varphi)$ are easily computed:

$$\vec\vartheta(t,\vec A;\bar\varphi)=\vec\vartheta_0\,\tanh gt, \qquad \vec\vartheta_0\equiv\,-4\partial_{\vec A}(gR) \tag{4.5}$$

We shall call $X^0(t)\equiv(I^0(t),\vec A,\bar\varphi(t),\vec\alpha^0(t))\equiv X^0(\vec A,\bar\varphi,\vec\alpha,t)$ the separatrix motion corresponding to the initial point with $\varphi=\bar\varphi$ and some initial $\vec A,\vec\alpha$ (cfr. paragraph after (4.1)).

Given a diffusion curve $\mathcal{L}$, $s\to\vec A_s$ we introduce the following notations: $\vec\omega_s\equiv\vec\omega(\vec A_s)$ $\equiv\ \partial_{\vec A}h(\vec A_s,0)$, $\vec\vartheta_s(t)=\vec\vartheta(t;\vec A_s,\bar\varphi)$, and define (see also (2.10)):

$$F(t;\vec\alpha,s)=-\sum_{\vec\nu\neq\vec 0}\frac{e^{i\vec\alpha\cdot\vec\nu}}{i\vec\omega_s\cdot\vec\nu}\partial_t\Big[f_{\vec\nu}(I^0(t),\vec A_s,\bar\varphi(t),0)e^{i\vec\vartheta_s(t)\cdot\vec\nu}\Big] \tag{4.6}$$

which makes sense for $s\in\Sigma(\infty)$, (which, in general, is a subset of full measure of $\mathcal{L}$) see (3.1).

Clearly the function $F(t;\vec\alpha,s)\xrightarrow[t\to\pm\infty]{}0$ exponentially fast (see (4.2)), and the following *Melnikov integral* is well defined, see (3.1), for $s\in\Sigma(\infty)\subseteq[s_1,s_2]$:

$$M_f(\vec\alpha,s)=\int_{-\infty}^{\infty}F(t;\vec\alpha+\vec\omega_st,s)dt \tag{4.7}$$

(similar quantities were considered by Poincarè in [P]; see also [A]). Note that in the special case of a degenerate phase shift, *i.e.* of independence of the rotators and the pendulum, the M_f are defined for all $s\in[s_1,s_2]$ because the part involving the small denominators in (4.6) disappears by integration by parts. In the latter case, in fact, it is:

$$M_f(\vec\alpha,s)=c(\vec A,s)+\int_{-\infty}^{\infty}f(t;\vec\alpha+\vec\omega_st,s)dt \tag{4.8}$$

where $c(\vec A,s)\equiv\int_{-\infty}^{\infty}[f_{\vec 0}(I^0,\vec A_s,\bar\varphi,0)-f_{\vec 0}(0,\vec A_s,0,0)]dt$ is a constant which shall play no role and $f(t;\vec\alpha,s)=f(I^0(t),\vec A_s,\bar\varphi(t),\vec\alpha,0)-f(0,\vec A_s,0,\vec\alpha,0)$.

Such a case with $f(0,\vec A_s,0,\vec\alpha,0)=0$ was considered by Arnold in [A].

For $s \in \Sigma(\infty)$, see (3.1), the equation:

$$\vec{\partial}_{\vec{\alpha}} M_f(\vec{\alpha}, s) = \vec{0} \tag{4.9}$$

admits necessarily at least two solutions (*e.g.* one is at $\vec{\alpha} = \vec{\alpha}_s$ when $\vec{\alpha}_s$ is a minimum for M_f and the other when $\vec{\alpha}_s$ is a maximum).

The following definition will be important:

Definition 2: *We say that the arc of diffusion path corresponding to $s \in [\bar{s}_1, \bar{s}_2] \subseteq [s_1, s_2]$ is directly open for diffusion under the perturbation f, see (2.9),(2.10), if:*

1) *no f-resonance occurs for $s \in [\bar{s}_1, \bar{s}_2]$, in the sense that $f_{\vec{\nu}}(0, 0, \vec{A}_s, 0)/\vec{\omega}_s \cdot \vec{\nu}$ is analytic in $s \in [\bar{s}_1, \bar{s}_2]$ for all $\vec{\nu}$.*
2) *the equation (4.9) admits a continuous solution $\vec{\alpha} \to \vec{\alpha}_s$ for all $s \in [\bar{s}_1, \bar{s}_2]$ and such that:*

$$\det\, \partial^2_{\vec{\alpha}} M_f(\vec{\alpha}_s, s) \neq 0 \qquad s \in [\bar{s}_1, \bar{s}_2] \tag{4.10}$$

More generally we say that an arc of a diffusion path is open for diffusion under the perturbation f if it can be covered by finitely many arcs directly open for diffusion.

Note that the non resonance condition is a very strong condition: except for very special f we can expect to find open diffusion paths only when f is a trigonometric polynomial. In the latter case, however, it is clear that, in general, there will be many open, possibly very long, such paths.

Consider a diffusion path and assume that property 1) of the above definition is verified because f is a trigonometric polynomial with no non vanishing coefficients $f_{\vec{\nu}}$ corresponding to $\vec{\nu}$'s for which $\vec{\omega}_s \cdot \vec{\nu} = 0$ for some s. Then given a point of parameter s on the path, it will be generically true that s is inside some arc of $\mathcal{L}$ directly open for diffusion under f: the genericity is with respect to the choices of the non zero coefficients of the trigonometric polynomial f. This is a consequence of the explicit formula (4.6) and of the remark that one can change rather arbitrarily the function $M_f(\vec{\alpha}, s)$ by changing f and the change is effectively computable.

Our main result in the above anisochronous, *a priori* unstable, case is the following.

Proposition: *Consider a hamiltonian like (2.9) with H_0 verifying the assumptions 1÷3 of §2 and f being a trigonometric polynomial of degree d.*
Given a diffusion path $\mathcal{L}$ directly open for diffusion, suppose that $\vec{\omega}(\vec{A}) \cdot \vec{\nu} \equiv \partial_{\vec{A}} h(\vec{A}, 0) \cdot \vec{\nu} \neq 0$ for $\vec{A}$ in $\mathcal{L}$ and for all $|\vec{\nu}| < c\,d$, for some constant $c > 0$.

If c is large enough then one can find, for all $\mu \neq 0$ small enough, initial data with "fast action variables" (i.e. $\vec{A}$ variables) close to one extreme of $\mathcal{L}$, $\vec{A}^1$, and "slow variables" (i.e. (I, φ)) close to the selected unstable equilibrium position, which evolve, drift, into data with the $\vec{A}$ variables close to the other extreme, $\vec{A}^2$, of $\mathcal{L}$. And close can be taken to mean within a distance $\delta_\mu \xrightarrow[\mu \to 0]{} 0$.

One can find constants $T_1, c_1 > 0$ such that:

$$T(\mu) \equiv T_1 e^{c_1 |\mu|^{-2}} \tag{4.11}$$

provides an upper bound to the minimum time necessary for the drift from $\vec{A}^1$ to $\vec{A}^2$.

If the path $\mathcal{L}$ is open for diffusion, but not directly open, one can show the same result with a function $T(\mu)$ whose expression will depend on the structure of $\mathcal{L}$: in particular, it will depend on the number of segments directly open for diffusion.

The above theorem does not convey all the information that we gather by proving it: the dimensional nature of our bounds makes them very flexible and we use them in the later sections of this paper to cover a variety of cases in which the non degeneracy conditions are not verified, and eventually lead us to the result on the *a priori* stable heavenly problem described in the introduction.

§5 Existence of ladders of whiskers.

In this section we consider a hamiltonian (2.9) verifying the assumptions 1÷3 of §2 and study the persistence of the unperturbed whiskered tori and their regularity properties (see, also, [M], [Gr], [Z]).

The basic technical facts concerning the existence of the $l-1$ dimensional invariant tori and the normal form of the flow in their vicinity are stated in the following lemmata 1,1' and in lemma 2 (formulated after the proofs).

Since the theorem presented in the lemmata is a local theorem in the vicinity of the unperturbed invariant tori, it is useful to introduce a system of coordinates in which it is most conveniently studied. Thus we introduce a new system of canonical coordinates $(I,\vec{A},\varphi,\vec{\alpha}) = \overline{\mathcal{R}}_\mu(p_0,q_0,\vec{A}_0,\vec{\alpha}_0)$ defined by a canonical transformation $\overline{\mathcal{R}}_\mu$ enjoying the properties explained in the following lemma.

Let $C_\xi \equiv \{z \,|\, e^{-\xi} < |z| < e^{\xi}\}$, and consider the sets of the points $I,\vec{A},\zeta,\vec{z},\mu \in U$ and, respectively, $p,q,\vec{A},\vec{z},\mu \in W$ with:

$$\begin{aligned} U(\rho',\rho,\xi',\xi,\bar{\mu},\vec{a}) &\equiv \big\{|I| \le \rho',\ |A_i - a_i| \le \rho,\ \zeta \in C_{\xi'},\ z_j \in C_\xi,\ |\mu| \le \bar{\mu}\big\} \\ W(\kappa,\rho,\xi,\bar{\mu},\vec{a}) &\equiv \big\{|p|,|q| < \kappa,\ |A_j - a_j| < \rho,\ z_j \in C_\xi,\ |\mu| < \bar{\mu}\big\} \end{aligned} \tag{5.1}$$

Recall the definition of V (beginning of §2) and that H in (2.9) is holomorphic in $U(\rho',\rho,\xi',\xi,\bar{\mu},\vec{a})$ (assumption 2, §2).

Lemma 0: *For all $\vec{a} \in V$ there exist positive constants $\bar{\kappa}_0,\bar{\rho}_0,\bar{\xi}_0$ and a canonical transformation $(I,\varphi,\vec{A},\vec{\alpha}) = \overline{\mathcal{R}}_\mu(p_0,q_0,\vec{A}_0,\vec{\alpha}_0)$ defined and holomorphic in $W(\bar{\kappa}_0,\bar{\rho}_0,\bar{\xi}_0,\bar{\mu},\vec{a})$ with values in a domain $U(\rho',\rho,\xi',\xi,\bar{\mu},\vec{a})$ of holomorphy of (2.9) and casting H in the form:*

$$\begin{aligned} &h_0(\vec{A}_0,p_0q_0,\mu) + f_0(\vec{A}_0,\vec{\alpha}_0,p_0,q_0,\mu), \\ &h_0(\vec{A}_0,J,\mu) = h(\vec{A}_0,\mu) + G(J,\vec{A},\mu), \qquad \partial_J G(0,\vec{A},\mu) \equiv g(\vec{A},\mu) \end{aligned} \tag{5.2}$$

where f_0 is divisible by μ and h_0, f_0 are analytic in $W(\bar\kappa_0,\bar\rho_0,\bar\xi_0,\bar\mu,\vec a)$. Expressions for possible values of $\bar\kappa_0,\bar\rho_0,\bar\xi_0$ in terms of $\rho',\rho,\xi',\xi,\bar\mu$ and of a few constants depending on h,f can be found in appendix A3, see ♣A3.39.

The **proof** is given in appendix A3.

The map $\overline{\mathcal{R}}_\mu$ will have the form:

$$\begin{aligned} I =&R(\vec A_0,p_0,q_0,\mu), & \varphi &= S(\vec A_0,p_0,q_0,\mu)\\ \vec\alpha =&\vec\alpha_0+\vec\delta(\vec A_0,p_0,q_0,\mu), & \vec A &= \vec A_0 \end{aligned} \tag{5.3}$$

with $R,S,\vec\delta$ real–analytic in $W(\bar\kappa_0,\bar\rho_0,\bar\xi_0,\bar\mu,\vec a)$ (often (5.3) will be supposed to act also on the variable μ, trivially changing μ into itself).

The result in lemma 0 is well known: it extends a celebrated theorem by Jacobi who proved the above lemma in a variety of cases, first of all for the standard pendulum. In the latter case the Jacobi map $\overline{\mathcal{R}}_\mu$ can be constructed quite explicitly by using the theory of the jacobian elliptic functions, see appendix A9.

Lemma 1 below gives us a *normal form* for the hamiltonian flow near the unperturbed whiskers. It tells us that *most* of the structure of unstable tori and of corresponding manifolds survives the onset of the perturbation. In particular the tori are obtained by setting suitable coordinates p,q equal to 0; and the whiskers, in the vicinity of the tori, are obtained by setting $p=0$ (unstable whisker) or $q=0$ (stable whisker). The whisker ladder still exists, with a *few rounds missing* (where $s\notin\Sigma_\mu$, see below).

Lemma 1*: Consider a hamiltonian (2.9), verifying the assumptions 1÷3 of §2. Let $\mathcal{L}$ be a diffusion path $s\to\vec A_s$ with energy E (see 1),2) of §3), and let $s\to\mathcal{T}_0(s)\equiv\mathcal{T}_0(\vec A_s)$ be the family of $(l-1)$-dimensional tori, see (3.3), associated with $\mathcal{L}$. Suppose that $\bigcup_{\vec a\in\mathcal{L}}U(\rho',\rho,\xi',\xi,\bar\mu,\vec a)$, which is contained in the holomorphy domain of (2.9), is a region where the map $\overline{\mathcal{R}}_\mu$ can be defined via lemma 0 above: i.e. $U(\rho',\rho,\xi',\xi,\bar\mu,\vec a)\supset\overline{\mathcal{R}}_\mu W(\bar\kappa_0,\bar\rho_0,\bar\xi_0,\ \bar\mu,\vec a)$.*

Fixed $n>0$ and μ real, there exists, on the energy level E of the perturbed system, a family $s\to\mathcal{T}_\mu(s)$ of $(l-1)$-dimensional "whiskered" tori, C^n-close to the line of tori $s\to\mathcal{T}_0(s)$ within $O(\mu)$ as $\mu\to0$, which for μ small enough verify the following properties:

1) *There exist positive constants $c,\bar c,K,k$ such that the tori $\mathcal{T}_\mu(s)$ are invariant for $s\in\Sigma_\mu\subset[s_1,s_2]$ where: $\Sigma_\mu\equiv\{s|\,C(s)<k|\mu|^{-1/\bar c}\}$ and:*

$$(s_2-s_1)^{-1}\int_{\Sigma_\mu}ds\ge(1-K|\mu|^{1/c}) \tag{5.4}$$

2) *The tori $\mathcal{T}_\mu(s)$ are part of a family of l–dimensional invariant surfaces having energy E and parameterized by $\vec\psi\in T^{l-1}$, $|p|,|q|<\bar\kappa$, for some $\bar\kappa>0$, as:*

$$\begin{aligned} \vec A =&\vec A'+\vec\Xi(\vec\psi,p,q,s,\mu) & \vec\alpha &= \vec\psi+\vec\Delta(\vec\psi,p,q,s,\mu)+\vec\delta(\vec A',p,q,\mu)\\ I =&R(\vec A',p,q,\mu)+\Lambda(\vec\psi,p,q,s,\mu) & \varphi &= S(\vec A',p,q,\mu)+\Theta(\vec\psi,p,q,s,\mu) \end{aligned} \tag{5.5}$$

where $\vec{A}' \equiv \vec{A}_s(pq,\mu)$ with $\vec{A}_s(J,\mu)$ analytic in J, C^n–smooth in J,s,μ and $\vec{A}_s(0,0)$ coincides with the diffusion curve $\vec{A}_s$; $\vec{\Xi},\vec{\Delta},\Lambda,\Theta$ are analytic in $\vec{\psi},p,q$, divisible by μ and C^n-smooth in $\vec{\psi},p,q,s,\mu$, and $R,S,\vec{\delta}$ are as in lemma 0 (hence depend on s,μ,p,q only and are analytic in their variables).

3) *There are functions $\gamma'(J,s,\mu),\gamma(J,s,\mu)$ analytic in J for $|J|<\bar{\kappa}^2$, C^n-smooth in J,s,μ and divisible by μ if $J=0$ (and by J if $\mu=0$), such that the motion on the invariant surfaces is simply:*

$$\vec{\psi}(t)=\vec{\psi}+(1+\gamma)\vec{\omega}_s t,\qquad p(t)=pe^{-g_s(1+\gamma')t},\qquad q(t)=qe^{+g_s(1+\gamma')t} \tag{5.6}$$

where $\gamma=\gamma(pq,s,\mu),\gamma'=\gamma'(pq,s,\mu),g_s\equiv g(\vec{A}_s,0)$, see (2.5), $\vec{\omega}_s\equiv\vec{\omega}(\vec{A}_s)$, see (3.1). Hence the tori $\mathcal{T}_\mu(s)$ and their stable/unstable whiskers $W_\mu(s)$ are obtained by setting in (5.5), respectively, $p=q=0$; $p\neq 0$, $q=0$; and $p=0$, $q\neq 0$.

4) *The smallness condition on μ and the constants $k,K,\bar{c},c,\bar{\kappa}$ can be given an explicit dimensional form in terms of a few parameters associated with h,f, (see ♣5.76, ♣5.90, ♣5.82, ♣5.67, ♣5.18 below); similarly one can construct explicit bounds on the smallness of $\vec{\Xi},\vec{\Delta},\Lambda,\Theta,\gamma,\gamma'$, (see lemma 2 and ♣5.89,♣5.79 ,♣5.88 below).*

Instead of fixing the energy E of the invariant tori and the frequency ratios of the corresponding quasi periodic motions one can fix the frequencies (*i.e.* γ in (5.6)) at the cost of leaving E free.

Calling $s\to\vec{A}_s$ the diffusion curve equation and defining the two functions $\vec{\omega}_s=\partial_{\vec{A}}h_0(\vec{A}_s,0,0)$, and $g_s=\partial_J h_0(\vec{A}_s,0,0)$, see (5.2), we introduce a real parameter u and consider the vectors:

$$\vec{\omega}_{su}=(1+u)\vec{\omega}_s,\quad u\text{ real },\qquad \Big(\,\vec{\omega}_s\equiv\partial_{\vec{A}}h_0(\vec{A}_s,0,0)\,\Big) \tag{5.7}$$

We define the *diffusion sheet* $\tilde{\mathcal{L}}$: $(s,u)\to\vec{A}_{su}$ by:

$$\partial_{\vec{A}}h_0(\vec{A}_{su},0,0)=\vec{\omega}_{su} \tag{5.8}$$

This is well defined, taking into account the non degeneracy conditions (2.7), by the implicit function theorem, if $|u|$ is small enough. We shall suppose that u varies in an interval $[-\bar{u},\bar{u}]$ so small that:

$$\text{setting }\partial_J h_0(\vec{A}_{su},0,0)=(1+u'_{su})g_s\equiv g_{su}\ \text{ it is }:\ \ |u|,|u'_{su}|<4^{-1} \tag{5.9}$$

More stringent requirements on $\bar{u}$ will be imposed later.

One then obtains results similar to those described in lemma 1 with the basic difference that all the main functions will be *analytic also in μ near $\mu=0$*, and the energy of the motions on the invariant surfaces will *no longer be fixed*. More precisely one obtains the following statement:

Lemma 1': *Consider, as in lemma 1, a hamiltonian (2.9), verifying the assumptions 1÷3 of §2. Let $\mathcal{L}$ be a diffusion path $s\to\vec{A}_s$ with energy E, and let $s,u\to\vec{A}_{su}$ be the* *diffusion sheet, defined in (5.7),(5.8), and let $s, u \to \mathcal{T}_0(s,u)$ be the family of $(l-1)$-dimensional tori (see (3.3) with $\vec{A} = \vec{A}_{su}$)) associated with $\mathcal{L}$. Suppose, as in lemma 1, that $\bigcup_{\vec{a}\in\mathcal{L}} U(\rho', \rho, \xi', \xi, \bar{\mu}, \vec{a})$, is a region where a map $\overline{\mathcal{R}}_\mu$ can be defined via lemma 0.*

Fix $n > 0$, let u be real and small, and μ complex. Then there exists a family $s, u \to \mathcal{T}_\mu(s,u)$ of $(l-1)$-dimensional "whiskered" tori, C^n-close to the sheet of tori $s, u \to \mathcal{T}_0(s,u)$ as $\mu \to 0$, which for μ small enough verify the following properties:

1) *The tori $\mathcal{T}_\mu(s,u)$ are invariant for $s \in \Sigma_\mu \subset [s_1, s_2]$ and for $u \in [-\bar{u}, \bar{u}]$ for a suitable $\bar{u} > 0$: here Σ_μ is the same set defined in 1) of lemma 1 and verifies the same bound (5.4) (same constants).*
2) *The tori $\mathcal{T}_\mu(s,u)$ are part of a family of invariant l–dimensional surfaces parameterized by $\vec{\psi} \in T^{l-1}$, $|p|, |q| < \bar{\kappa}$, ($\bar{\kappa}$ as in lemma 1), as:*

$$\begin{aligned} \vec{A} =& \vec{A}' + \vec{\Xi}(\vec{\psi}, p, q, s, u, \mu) & \vec{\alpha} =& \vec{\psi} + \vec{\delta}(\vec{A}', p, q, \mu) + \vec{\Delta}(\vec{\psi}, p, q, s, u, \mu) \\ I =& R(\vec{A}', p, q, \mu) + \Lambda(\vec{\psi}, p, q, s, u, \mu) & \varphi =& S(\vec{A}', p, q, \mu) + \Theta(\vec{\psi}, p, q, s, u, \mu) \end{aligned} \tag{5.10}$$

where $\vec{A}' \equiv \vec{A}_{su}(pq, \mu)$ with $\vec{A}_{su}(J, \mu)$ analytic in J, μ, C^n–smooth in J, μ, s, u and $\vec{A}_{su}(0, 0) = \vec{A}_{su}$ (see (5.8)); $\vec{\Xi}, \vec{\Delta}, \Lambda, \Theta$ are analytic in $\vec{\psi}, p, q, \mu$, divisible by μ, and C^n-smooth in all their arguments, and $R, S, \vec{\delta}$, which depend on s, μ, p, q only, are as in (5.3).

3) *There is a function $\gamma'(J, s, u, \mu)$ analytic in J, μ for $|J| < \bar{\kappa}^2$ and μ small enough, C^n-smooth in s, u, μ, J and divisible by μ if $J = 0$ (and by J if $\mu = 0$), such that the motion on the invariant surfaces is simply:*

$$\vec{\psi}(t) = \vec{\psi} + \vec{\omega}_{su} t, \qquad p(t) = p e^{-g_{su}(1+\gamma')t}, \qquad q(t) = q e^{+g_{su}(1+\gamma')t} \tag{5.11}$$

where $\gamma' = \gamma'(pq, s, u, \mu), g_{su} \equiv g(\vec{A}_{su}, 0), \vec{\omega}_{su} \equiv (1+u)\vec{\omega}_s$ and $\gamma \equiv u$ is now fixed a priori.

4) *The constants $k, K, \bar{c}, c, \bar{\kappa}$ are as in lemma 1 above; furthermore the smallness condition on $|\mu|$ and the (new) functions $\vec{\Xi}, \vec{\Delta}, \Lambda, \Theta, \gamma'$ satisfy the same bounds of the corresponding objects of lemma 1 (see point 4) of lemma 1).*

In fact the strategy of our analysis will be to prove lemma 1' first and deduce lemma 1 by showing that the parameter u can be determined so that the real part of the energy *maintains a prefixed value E.*

Proof: the first step is to change variables $(I, \varphi, \vec{A}, \vec{\alpha}) \to (p_0, q_0, \vec{A}_0, \vec{\alpha}_0)$ using the canonical change of coordinates of lemma 0 to put (2.9) in the form (5.2).

By our assumption this is possible and we call $\bar{\kappa}_0, \bar{\rho}_0, \bar{\xi}_0$ parameters such that $\overline{\mathcal{R}}_\mu W(\bar{\kappa}_0, \bar{\rho}_0, \bar{\xi}_0, \bar{\mu}, \vec{a})$ is, for all $\vec{a} \in \mathcal{L}$ contained in the set $\bigcup_{\vec{a}\in\mathcal{L}} U(\rho', \rho, \xi', \xi, \bar{\mu}, \vec{a})$ where the hamiltonian is defined.

In this way we define h_0, f_0 on $W = W(\bar{\kappa}_0, \bar{\rho}_0, \bar{\xi}_0, \bar{\mu}, \vec{a})$ for all $\vec{a} \in \mathcal{L}$. Let E_0, η_0, Γ_0 be the suprema, in W and $\vec{a} \in \mathcal{L}$, of the functions $||\partial h_0||$ and $||(\partial^2_{\vec{A}} h_0)^{-1}||$, $||(\partial_J h_0)^{-1}||$, respectively. The norm of a vector or matrix will be, for simplicity, the maximum of the components.

Consider the equation (5.8). By a simple implicit function analysis we see that if:

$$|u| < \tilde{u} \equiv \frac{1}{\bar{B}^2(E_0\eta_0\bar{\rho}_0^{-1})^2} \tag{5.12}$$

for $\bar{B}$ large enough, it admits a solution $\vec{A}_{su}$ such that, (see appendix A4, ♣A4.3):

$$|\vec{A}_{su} - \vec{A}_s| < \tilde{\rho} \equiv \frac{\bar{\rho}_0}{(\bar{B}E_0\eta_0\bar{\rho}_0^{-1})} < \bar{\rho}_0/4 \tag{5.13}$$

We also consider the equation:

$$\partial_{\vec{A}} h_0(\vec{A}, J, \mu) = \vec{\omega}_s(1+u) \equiv \vec{\omega}_{su} \tag{5.14}$$

and we see that if $|u|$ verifies (5.12) and $|J| < \tilde{\kappa}^2$, $|\mu| < \tilde{\mu}$ with:

$$\tilde{\kappa} \equiv \frac{\bar{\kappa}_0}{\bar{B}(E_0\eta_0\bar{\rho}_0^{-1})}, \qquad \tilde{\mu} \equiv \frac{\bar{\mu}}{\bar{B}^2(E_0\eta_0\bar{\rho}_0^{-1})^2} \tag{5.15}$$

then the equation has a solution $\vec{A}^0(s, u, J, \mu)$ close to $\vec{A}_{su}$ within $\bar{\rho}_0/4$, (see appendix A4, ♣A4.5) and, obviously, $\vec{A}^0(s, u, 0, 0) \equiv \vec{A}_{su}$.

Recalling that $g_{su} \equiv \partial_J h_0(\vec{A}_{su}, 0, 0)$ and setting:

$$(1 + u'(s, u, \vec{A}, J, \mu))g_{su} \equiv \partial_J h_0(\vec{A}, J, \mu), \qquad \lambda_0 \equiv \sup|u'| \tag{5.16}$$

we find that in a domain $|\vec{A} - \vec{A}^0(s, u, J, \mu)| < \rho_0$, $|J| < \kappa_0^2, |\mu| < \mu_0$ the following bound holds for a suitable constant $\hat{B}$:

$$\lambda_0 = \sup|u'| \le 2l\Gamma_0 E_0 \Big(\frac{\tilde{\kappa}^2}{\bar{\kappa}_0^2} + \frac{\tilde{\rho}}{\bar{\rho}_0} + \frac{\tilde{\mu}}{\bar{\mu}}\Big) \le \hat{B}(\Gamma_0 E_0)\ (E_0\eta_0\bar{\rho}_0^{-1})^2 \tag{5.17}$$

see appendix A4, ♣A4.6; in such bounds we have used "dimensional" (or "Cauchy") estimates: see below.

Therefore we can fix $\rho_0, \kappa_0^2, \bar{u}$ so that (5.12) holds (hence $|u| < 1/4$) and also $4\lambda_0\Gamma_0E_0 < 1$ (hence $|u'|$, $|u'_{su}| < 1/4$, (see (5.9) for the definition of u'_{su}) because $|u'_{su}|$ can also be bounded by the r.h.s. of (5.17) by a similar estimate; see (5.9) for the definition of u'_{su}). A possible choice is:

$$\begin{aligned}
&\bar{u} \equiv \frac{1}{\bar{B}^2(E_0\eta_0\bar{\rho}_0^{-1})^2}, \qquad && \mu_0 \equiv \min\{\frac{\bar{\mu}}{\bar{B}^2(E_0\eta_0\bar{\rho}_0^{-1})^2(E_0\Gamma_0)^2}\ ,\ 1\} \\
&\kappa_0 \equiv \frac{\bar{\kappa}_0}{\bar{B}\,(E_0\eta_0\bar{\rho}_0^{-1})}\frac{1}{(E_0\Gamma_0)}, \qquad && \rho_0 \equiv \min\{\frac{\bar{\rho}_0}{\bar{B}(E_0\eta_0\bar{\rho}_0^{-1})(E_0\Gamma_0)^2}\ ,\ \kappa_0^2\}\ , \qquad \xi_0 \equiv \bar{\xi}_0
\end{aligned} \tag{5.18}$$

where ρ_0, μ_0 are taken to be necessarily smaller than κ_0^2 and 1, respectively, for later convenience. The constant $\bar{B}$ can be taken to be the same in all the above formulae, possibly readjusting it (to avoid the introduction of too many symbols, *a procedure that we shall use very often below*).

The functions h_0, f_0 will be holomorphic in the new coordinates in a domain that we have, to some extent, tailored to our needs. They will, in fact, be regarded as holomorphic in a domain containing:

$$\begin{aligned} &W_0 \equiv W(\kappa_0, \rho_0, \xi_0, \mu_0) \equiv \\ &\equiv \bigcup_{s,u \in \tilde{\mathcal{I}}_0} \{|p_0|, |q_0| < \kappa_0,\ |A_{0j} - A_j^0(s, u, p_0 q_0, \mu)| < \rho_0\ , e^{-\xi_0} < |z_j| < e^{\xi_0},\ |\mu| < \mu_0\} \end{aligned} \tag{5.19}$$

where the sheet $(s, u) \to \vec{A}^0(s, u, J, \mu)$ is defined by (5.14) with:

$$(s, u) \in \tilde{\mathcal{I}}_0 \equiv \mathcal{I}_0 \times [-\bar{u}, \bar{u}]\ , \quad \mathcal{I}_0 \equiv \Sigma(C_0) \quad \text{for some } C_0 > \Gamma_0 \tag{5.20}$$

In the coming analysis the constant C_0 will be left as a *free parameter* and will be chosen at the end in order to check (5.4). Thus, using $|u| < 1/4$, in $\tilde{\mathcal{I}}_0$ it will be true that:

$$\begin{aligned} |\vec{\omega}_{su} \cdot \vec{\nu}|^{-1} &\le C_0 |\vec{\nu}|^{\tau}, \qquad \forall \vec{\nu} \in Z^{l-1},\ \vec{\nu} \ne \vec{0} \\ |g_{su}|^{-1} &\le \Gamma_0 \end{aligned} \tag{5.21}$$

where, see (3.1), τ is a diophantine constant. Note that the just introduced parameters $\kappa_0, \rho_0, \xi_0, \mu_0$ are not, in any sense, the maximal ones compatible with the analyticity properties of h_0, f_0.

All our arguments will have *dimensional nature* involving combinations of the *sizes* of various functions, hence it is convenient to define the size of a function $\vec{F}$, holomorphic in a domain W, as:

$$||\vec{F}|| \equiv ||\vec{F}||_W \equiv \sup_{j,W} |F_j(p, q, \vec{A}, \vec{z}, \mu)| \tag{5.22}$$

where, of course, the symbol $||\cdot||$ is incomplete and (therefore) it will be always accompanied by the specification of the domain W considered in evaluating (5.22), unless obvious from the context.

Let us collect here the positive parameters $E_0, \Gamma_0, \varepsilon_0, \eta_0, \vartheta_0$ that we use to measure the size of h_0, f_0 (compare (2.7)):

$$\begin{aligned} ||\partial h_0|| \le E_0, \quad ||(\partial_J h_0)^{-1}|| \le \Gamma_0, \qquad ||f_0|| \le \varepsilon_0, \\ ||(\partial_{\vec{A}}^2 h_0)^{-1}|| \le \eta_0, \quad ||[(\partial_{\vec{A}}^2 h_0)^{-1} \vec{\partial}_{\vec{A}} h_0 \cdot \vec{\partial}_{\vec{A}} h_0]^{-1}|| \le \vartheta_0 \end{aligned} \tag{5.23}$$

where $||\cdot||$ is considered in W_0, see (5.19). This is consistent with the previous meaning and usage of the previously defined values of E_0, η_0, Γ_0.

The holomorphy of h_0, f_0 imposes restrictions on the relative values of the above constants; namely there exists $B_0 > 0$ depending only on the number l of degrees of freedom and such that:

$$E_0 C_0 > E_0 \Gamma_0 \geq B_0, \quad \eta_0 E_0 \rho_0^{-1} \geq B_0, \quad \vartheta_0 E_0^2 \eta_0 \geq B_0 \tag{5.24}$$

which we will repeatedly use for the purpose of simplifying bounds, at the expense of their sharpness; (one can take $B_0 = l^{-1}$, see appendix A4).

The quantities in (5.24) have the physical interpretation of ratios of the various relevant time scales relevant for our problem.

Our basic tool (already used in obtaining (5.17)) for bounds on a function F, of one variable, holomorphic in a domain $\mathcal{D}$ will be to restrict it to a smaller domain $\mathcal{D}' \subset \mathcal{D}$ and to estimate the n-th derivative of F in $\mathcal{D}'$ by $n!$ times r^{-n}, with $r =$ distance between $\mathcal{D}'$ and $\partial\mathcal{D}$, times the supremum of F in $\mathcal{D}$. We call such a bound a *dimensional estimate*: it is a consequence of (one among) the Cauchy's theorem(s).

In performing dimensional bounds it is convenient to deal with *dimensionless combinations* of the main parameters (5.23). Thus all our bounds will naturally involve the following dimensionless combinations of the parameters $E_0, \eta_0, C_0, \Gamma_0, \vartheta_0, \varepsilon_0, \rho_0, \kappa_0, \xi_0, \lambda_0, \mu_0$ that we have associated with our hamiltonian (see (5.23), (5.18), (5.20), (5.17),(5.19)):

$$\begin{aligned}
&E_0 C_0,\ C_0 \Gamma_0^{-1},\ \eta_0 E_0 \rho_0^{-1},\ \vartheta_0 E_0^2 \eta_0,\ \kappa_0^2 \rho_0^{-1},\ \hat{\xi}_0^{-1} \equiv \xi_0^{-1}(1+\xi_0),\ \mu_0,\ \lambda_0 \\
&\varepsilon_0 E_0^{-1} \rho_0^{-1}
\end{aligned} \tag{5.25}$$

and we see, from (5.17),(5.24),(5.20), and from (5.18) and the comment following it, that all the elements of the first line are $\geq B_0 > 0$; we shall impose, without loss of generality, that the element of the second line is $\leq 1/2$.

To help reading the formulae we often close in parentheses the above dimensionless combinations of parameters, even tough they may not be necessary.

Given a function F holomorphic on W_0, see (5.19), we introduce the Fourier coefficients $F_{\vec{\nu}}(\vec{A}, p, q)$ and the Taylor coefficients $F_{,hk}(\vec{A}, \vec{z})$ of the expansions:

$$F(\vec{A}, \vec{z}, p, q) = \sum_{\vec{\nu} \in Z^{l-1}} F_{\vec{\nu}}(\vec{A}, p, q) \vec{z}^{\vec{\nu}} = \sum_{h,k=0}^{\infty} F_{,hk}(\vec{A}, \vec{z}) p^h q^k \tag{5.26}$$

where $z_j = \exp i\alpha_j$ and $\vec{z}^{\vec{\nu}} = \prod_j z_j^{\nu_j} = \prod_j e^{i\nu_j \alpha_j}$, (the latter two notations will be used interchangeably).

Thus we can introduce the following functions (*truncations of F*), for $N \geq 0, |\vec{\nu}| \equiv \sum_j |\nu_j|$:

$$\begin{aligned}
&F^{[\leq N]}(\vec{A}, \vec{z}, p, q) = \sum_{|\vec{\nu}| \leq N} F_{\vec{\nu}}(\vec{A}, p, q) e^{i\vec{\nu}\vec{\alpha}}, \qquad F^{[>N]} \equiv F - F^{[\leq N]} \\
&F^D(\vec{A}, \vec{z}, w) = \sum_{h \geq 0} F_{,hh}(\vec{A}, \vec{z}) w^h
\end{aligned} \tag{5.27}$$

We can now begin our sequence of estimates leading to the proof of lemma 1' and lemma 1.

The following dimensional estimates hold for various truncations of f_0; given N_0, δ_0:

$$\begin{aligned} ||f_0^{[\le N_0]}|| \le & B_1\varepsilon_0\hat{\xi}_0^{-\beta_1}\delta_0^{-\beta_1}, \qquad ||f_0^{[>N_0]}|| \le B_1\varepsilon_0\hat{\xi}_0^{-\beta_1}\delta_0^{-\beta_1}e^{-\xi_0\delta_0 N_0/2}, \\ ||f_0^{[\le N_0]D}|| \le & B_1\varepsilon_0\hat{\xi}_0^{-\beta_1}\delta_0^{-\beta_1}, \qquad ||f_0^{[\le N_0]} - f_0^{[\le N_0]D}|| \le B_1\varepsilon_0\hat{\xi}_0^{-\beta_1}\delta_0^{-\beta_1} \end{aligned} \tag{5.28}$$

where the $||\cdot||$ is evaluated from (5.22) on the domain $W(\kappa_0 e^{-\delta_0}, \rho_0, \xi_0 e^{-\delta_0}, \mu_0)$ and the inequalities express simple dimensional estimates in the sense defined above: the constants that arise have been adjusted so that only the two parameters $B_1, \beta_1 > 0$ are needed. Sharper bounds would require more constants; but we are not interested in sharpness of the estimates (in this paper).

We assign, *a priori* , a sequence $\delta_0 > \delta_1 \ldots$ of positive numbers such that $4\sum_{j=0}^{\infty}\delta_j < \log 2$ and such that δ_j does not approach zero too fast (*e.g.* $\delta_j = (1+j^2)^{-1}2^{-4}\log 2$): it will be a set of auxiliary parameters that we shall use in our inductive construction. Below we introduce sequences of other parameters $B_1, B_2, B_3, \ldots$ and $\beta_1, \beta_2, \beta_3, \ldots$, depending only on the number of degrees of freedom l (and on the diophantine constant τ, see (3.1)), and we shall suppose the B_j's and the β_j's increasing (there will be, however, only finitely many such constants).

Let N_0 be such that (5.28) implies $||f_0^{[>N_0]}|| \le O(\varepsilon_0^2)$; for instance, recalling that $\varepsilon_0 E_0^{-1}\rho_0^{-1} < 1/2$, by the remark following (5.25):

$$N_0 = -2\xi_0^{-1}\delta_0^{-1}\log(\varepsilon_0 E_0^{-1}\rho_0^{-1}) \quad \Rightarrow \quad ||f_0^{[>N_0]}|| \le B_1\hat{\xi}_0^{-\beta_1}\delta_0^{-\beta_1}\varepsilon_0^2 E_0^{-1}\rho_0^{-1} \tag{5.29}$$

Calling $(\vec{A}_0, \vec{\alpha}_0, p_0, q_0)$ the canonical coordinates in which we describe our initial hamiltonian as in (5.2), we consider the canonical map defined via a generating function which, denoting the new variables with a prime, is a function $\Phi(\vec{A}', p', \vec{\alpha}_0, q_0, \mu)$ given by:

$$\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\alpha}_0}\Phi + g_0\,[q_0\partial_{q_0}\Phi - p'\partial_{p'}\Phi] = -f_0^{[\le N_0]} + \overline{f}_0^{[\le N_0]^D} \tag{5.30}$$

where $g_0 \equiv g_0(\vec{A}', J, \mu) \equiv \partial_J h_0(\vec{A}', J, \mu)$, $\vec{\omega}_0 \equiv \vec{\omega}(\vec{A}', J, \mu) \equiv \partial_{\vec{A}'} h_0(\vec{A}', J, \mu)$ with $J \equiv p' q_0$ and the bar denotes average over the $\vec{\alpha}$-variables.

The function Φ can be written:

$$\Phi(\vec{A}', \vec{\alpha}_0, p', q_0, \mu) = \sum_{\substack{|h-k|+|\vec{\nu}|>0 \\ |\vec{\nu}|\le N_0}} \frac{f_{0\vec{\nu},hk}(\vec{A}', \mu)e^{i\vec{\nu}\cdot\vec{\alpha}_0}p'^h q_0^k}{-i\vec{\omega}_0(\vec{A}', p'q_0, \mu)\cdot\vec{\nu} - g_0(\vec{A}', p'q_0, \mu)(k-h)} \tag{5.31}$$

The function Φ is defined in a domain:

$$W(\kappa_0 e^{-\delta_0}, \tilde{\rho}_0, \xi_0 e^{-\delta_0}, \mu_0), \quad \tilde{\rho}_0 < \rho_0 \tag{5.32}$$

(hence smaller than the one where (5.28) hold), where $\tilde\rho_0$ is so chosen to control the denominators in (5.31). By dimensional bounds one checks easily that if:

$$\tilde\rho_0=\rho_0[4lE_0C_0N_0^{\tau+1}]^{-1},\qquad \lambda_0E_0\Gamma_0<4^{-1} \tag{5.33}$$

(cfr. also (5.18)) then, for $|A'_i-A^0_i(s,u,J,\mu)|<\tilde\rho_0$, $0<|\vec\nu|\le N_0$ and $|J|<\kappa_0^2$ one has:

$$|-i\vec\omega_0(\vec A',J,\mu)\cdot\vec\nu+g_0(\vec A',J,\mu)(h-k)|^{-1}\le 2C_0(|\vec\nu|^\tau+|h-k|) \tag{5.34}$$

see appendix A5.

The last inequality can be combined with dimensional bounds to imply:

$$||\Phi||\le B_2\hat\xi_0^{-\beta_2}\delta_0^{-\beta_2}\,\varepsilon_0C_0 \tag{5.35}$$

for suitably chosen $B_2,\beta_2>0$, and in the domain (5.32).

The canonical map associated with Φ is generated by the following standard relations (omitting the explicit μ–dependence):

$$\begin{aligned}\vec A_0&=\vec A'+\partial_{\vec\alpha_0}\Phi(\vec A',\vec\alpha_0,p',q_0),\quad & p_0&=p'+\partial_{q_0}\Phi(\vec A',\vec\alpha_0,p',q_0),\\ \vec\alpha'&=\vec\alpha_0+\partial_{\vec A'}\Phi(\vec A',\vec\alpha_0,p',q_0),\quad & q'&=q_0+\partial_{p'}\Phi(\vec A',\vec\alpha_0,p',q_0)\end{aligned} \tag{5.36}$$

which could be written in the more precise complex variables notation, (see comment after (5.26)), by replacing $\vec\alpha_0$ by $\vec z_0$ in the argument of Φ, writing $iz_{0j}\partial_{z_{0j}}$ for $\partial_{\alpha_{0j}}$ and replacing the third of (5.36) by:

$$z'_j=z_{0j}\exp[i\partial_{A'_j}\Phi(\vec A',\vec z_0,p',q_0)] \tag{5.37}$$

To obtain a map $\tilde{\mathcal C}$ from (5.36), one has to use the implicit functions theorem: in so doing the domain of definition of $\tilde{\mathcal C}$ has to be taken somewhat smaller than the domain, (5.32), of definition of Φ. If we want $\tilde{\mathcal C}(\vec A',\vec\alpha',p',q',\mu)$ to be defined on the domain:

$$\tilde W\equiv W(\kappa_0e^{-2\delta_0},\tilde\rho_0e^{-\delta_0},\hat\xi_0e^{-2\delta_0},\mu_0) \tag{5.38}$$

(*i.e.* "just giving up regularity" by an extra δ_0) we must impose a condition implying that it is $\tilde\rho_0^{-1}\hat\xi_0^{-1}\delta_0^{-2}||\Phi||\ll 1$, *i.e.* :

$$x\equiv B_3\hat\xi_0^{-\beta_3}\delta_0^{-\beta_3}(\varepsilon_0C_0\rho_0^{-1})(E_0C_0)N_0^{\tau+1}<1 \tag{5.39}$$

with B_3,β_3 conveniently large.

This follows from a trivial implicit function theorem. After a moment of thought one realizes that such a condition implies at the same time the injectivity of the map (5.36), the non vanishing of its jacobian and it also imposes that the image of the boundary of the domain $W(\kappa_0e^{-\delta_0},\tilde\rho_0,\xi_0e^{-\delta_0})$ where $||\Phi||$ is defined stays well away from the boundary of $\tilde W$: in appendix A4 we have called such an argument an *image of the boundary* lemma, (see, for instance, [G], §5.11). Here (5.18), *i.e.* $\rho_0 < \kappa_0^2$, has been used to eliminate κ_0 from the condition.

The map $\tilde{\mathcal{C}}$: $(\vec{A}', \vec{\alpha}', p', q', \mu) \in \tilde{W} \to (\vec{A}_0, \vec{\alpha}_0, p_0, q_0)$ will take the form:

$$\begin{aligned} \vec{A}_0 =& \vec{A}' + \vec{\Xi}_0(\vec{A}', \vec{\alpha}', p', q') \qquad\qquad & p_0 = p' + \Lambda_0(\vec{A}', \vec{\alpha}', p', q') \\ \vec{\alpha}_0 =& \vec{\alpha}' + \vec{\Delta}_0(\vec{A}', \vec{\alpha}', p', q') \qquad\qquad & q_0 = q' + \Theta_0(\vec{A}', \vec{\alpha}', p', q') \end{aligned} \tag{5.40}$$

and in the domain $\tilde{W}$ the bounds:

$$\begin{aligned} ||\vec{\Xi}_0|| <& x\tilde{\rho}_0\delta_0, \qquad\qquad & ||\Lambda_0|| < x\kappa_0\delta_0 \\ ||\vec{\Delta}_0|| <& x\delta_0 \qquad\qquad & ||\Theta_0|| < x\kappa_0\delta_0 \end{aligned} \tag{5.41}$$

are valid, with x defined by (5.39), and $\rho_0 < \kappa_0^2$ has been again used.

The map $\tilde{\mathcal{C}}$ will transform the Hamiltonian (5.2) into:

$$h_1(\vec{A}', p'q', \mu) + f_1(\vec{A}', \vec{\alpha}', p', q', \mu) \tag{5.42}$$

where:

$$\begin{aligned} h_1 =& h_0(\vec{A}', p'q', \mu) + \overline{f_0}^D(\vec{A}', p'q', \mu) \\ \overline{f_0}^D(\vec{A}', p'q', \mu) \equiv& \sum_{k=0}^{\infty} f_{0\vec{0},kk}(\vec{A}', \mu)(p'q')^k \equiv \int f_0^D(\vec{A}', \vec{\alpha}', p'q', \mu)\frac{d\vec{\alpha}'}{(2\pi)^{l-1}} \end{aligned} \tag{5.43}$$

The functions h_1, f_1 are easily controlled (by "just giving up a bit δ_0 of regularity" in each variable) in:

$$\overline{W} = W(\kappa_0 e^{-3\delta_0}, \tilde{\rho}_0 e^{-2\delta_0}, \hat{\xi}_0 e^{-3\delta_0}, \mu_0) \tag{5.44}$$

by using dimensional estimates, from (5.40),(5.41) and along a well known elementary scheme, see [G] §5.12; the result is:

$$\begin{aligned} ||\partial h_1||_{\overline{W}} \le& E_0(1 + B_4\delta_0^{-\beta_4}(\varepsilon_0\rho_0^{-1}E_0^{-1})) \\ ||f_1||_{\overline{W}} \le& B_4\hat{\xi}^{\beta_4}\delta_0^{-\beta_4}\varepsilon_0(\varepsilon_0 E_0^{-1}\rho_0^{-1})(E_0C_0)^2 N_0^{\tau+1} \end{aligned} \tag{5.45}$$

The next step is to study the equations for $\vec{a} \in C^{l-1}$ given by:

$$\partial_{\vec{A}'} h_1(\vec{A}^0 + \vec{a}, J, \mu) = \vec{\omega}_{su}, \tag{5.46}$$

with $(s, u) \in \tilde{\mathcal{I}}_0$, $\vec{A}^0 = \vec{A}^0(s, u, J, \mu)$, see (5.20).

By Taylor expansion this can be written, setting $M_0 = \partial_{\vec{A}\vec{A}} h_0(\vec{A}^0, J, \mu)$, as:

$$\vec{a} + M_0^{-1}\vec{n}(\vec{a}) = \vec{0} \tag{5.47}$$

and $M_0^{-1}\vec{n}(\vec{a}) \equiv -\vec{m}(\vec{a})$ can be bounded by:

$$|M_0^{-1}\vec{n}(\vec{a})| \le 4l^2\eta_0(\varepsilon_0\rho_0^{-1} + E_0\rho^2\rho_0^{-2}) \;, \qquad \text{if} \;\; |\vec{a}| < \rho < \tilde{\rho}_0/2 \tag{5.48}$$

The (5.47) can be studied by applying the implicit function theorem. The usual argument about the "image of the boundary" implies the existence of a unique solution to the equation $\vec{a} = \vec{m}(\vec{a})$ with $|\vec{a}| < \rho$ if $b||\vec{m}||_\rho \; \rho^{-1} < 1$ for a suitably large b: for instance [G, proposition 19, p.490] shows that $b = 2^8$ is sufficient (but $b = 2$ would also be sufficient).

Therefore we take:

$$\rho \equiv \frac{\tilde{\rho}_0(\varepsilon_0 E_0\rho_0^{-1})^\chi\delta_0}{4l^2 b(\eta_0 E_0\rho_0^{-1})} < \frac{\tilde{\rho}_0\delta_0(\varepsilon_0 E_0\rho_0^{-1})^\chi}{4} \tag{5.49}$$

where $\chi \in (0,1)$ is a free parameter that we eventually fix close to 0 (*e.g.* 1/4) and b is as above (*e.g.* $b = 2$).

We deduce that a sufficient condition for the existence of a solution to (5.47) with $|\vec{a}| < \rho$ is:

$$\begin{aligned} 1 >& 4l^2 \; b \; \eta_0(\varepsilon_0\rho_0^{-1} + E_0\rho^2\rho_0^{-2})\rho^{-1} \equiv 4l^2 b(\eta_0 E_0\rho_0^{-1})\Big(\varepsilon_0 E_0^{-1}\rho_0^{-1}\frac{\rho_0}{\rho} + \frac{\rho}{\rho_0}\Big) \Leftarrow \\ &\Leftarrow 4l^2 b(\varepsilon_0 E_0^{-1}\rho_0^{-1})^{1-\chi}\frac{4l^3 b(\eta_0 E_0\rho_0^{-1})^2}{\delta_0}\frac{\rho_0}{\tilde{\rho}_0} + \frac{1}{2} < 1 \end{aligned} \tag{5.50}$$

and the latter condition can be imposed by requiring:

$$B_5\hat{\xi}_0^{-\beta_5}\delta_0^{-\beta_5}(\varepsilon_0 E_0^{-1}\rho_0^{-1})^{1-\chi}(\eta_0 E_0\rho_0^{-1})^2(E_0 C_0)\big[-\log(\varepsilon_0 E_0^{-1}\rho_0^{-1})\big]^{(\tau+1)} < 1 \tag{5.51}$$

for suitably large B_5, β_5.

Setting $\vec{A}^1(s,u,J,\mu) = \vec{A}^0(s,u,J,\mu) + \vec{a}$ we get (see (5.48)):

$$|\vec{A}^1(s,u,J,\mu) - \vec{A}^0(s,u,J,\mu)| < \rho < \frac{1}{4}\tilde{\rho}_0\delta_0(\varepsilon_0 E_0^{-1}\rho_0^{-1})^\chi \tag{5.52}$$

The free constant χ could in fact be taken zero, at the price of having no ε_0 dependence in the r.h.s. of (5.52): a property that we do not want in later estimates.

Therefore (5.52) insures also that $(\vec{A}, \vec{\alpha}, p, q, \mu) \equiv (\vec{A}^1, \vec{\alpha}, 0, 0, 0)$ lies very well inside the domain, $\overline{W}$, of definition of $h_1 + f_1$, (*i.e.* of f_1).

Choosing suitably B_6 and β_6 one easily checks that the two conditions:

$$\begin{aligned} &B_6\hat{\xi}_0^{-\beta_6}\delta_0^{-\beta_6}(\varepsilon_0 E_0^{-1}\rho_0^{-1})^{1-\chi}(\eta_0 E_0\rho_0^{-1})^2(C_0 E_0)[\log(E_0\rho_0\varepsilon_0^{-1})]^{\tau+1} < 1 \\ &\lambda_0\Gamma_0 E_0 < 4^{-1} \end{aligned} \tag{5.53}$$

imply all the conditions imposed so far (*i.e.* imposed in (5.51),(5.33),(5.39)).

With the above defined $(s,u)\to \vec{A}^1(s,u,J,\mu)$ we can define, via (5.19) with $0\to 1$, the set:

$$W_1 \equiv W(\kappa_1,\rho_1,\xi_1,\mu_1) \tag{5.54}$$

where:

$$\begin{aligned}
&\kappa_1\equiv\kappa_0 e^{-4\delta_0},\quad \xi_1\equiv\hat{\xi}_0 e^{-4\delta_0},\quad \mu_1\equiv\mu_0,\\
&\rho_1\equiv\rho_0\big(B_7\hat{\xi}_0^{-\beta_7}\delta_0^{-\beta_7}(E_0C_0)(\log E_0\rho_0/\varepsilon_0)^{\tau+1}\big)^{-1}\\
&C_1\equiv C_0,\quad \Gamma_1\equiv\Gamma_0
\end{aligned}\tag{5.55}$$

Note that for B_7,β_7 large enough it follows that $\rho_1<\tilde{\rho}_0/2$ so that the domain W_1 is strictly contained in the domain $\overline{W}$, see (5.44), of definition of h_1,f_1 and the above definitions, via dimensional estimates, allows to control *all the derivatives* of h_1 and f_1 in W_1.

The new parameters measuring the size of h_1,f_1 (cfr. (5.23), (5.16)) can be taken, by (5.45), to be any parameters $E_1,\varepsilon_1,\eta_1,\lambda_1$ verifying the following inequalities:

$$\begin{aligned}
&E_1\ge E_0(1+B_8\delta_0^{-\beta_8}(\varepsilon_0\rho_0^{-1}E_0^{-1}))\\
&\varepsilon_1\rho_1^{-1}E_1^{-1}\ge B_8\hat{\xi}_0^{-\beta_8}\delta_0^{-\beta_8}(\varepsilon_0\rho_0^{-1}E_0^{-1})^2(E_0C_0)^3(-\log\varepsilon_0E_0^{-1}\rho_0^{-1})^{2(\tau+1)}\\
&\eta_1\ge\eta_0\big(1+B_8\delta_0^{-\beta_8}(\eta_0\rho_0^{-1}E_0)(\varepsilon_0E_0^{-1}\rho_0^{-1})\big)\\
&\lambda_1\ge\lambda_0+B_8(E_0\Gamma_0)(\varepsilon_0\rho_0^{-1}E_0^{-1})
\end{aligned}\tag{5.56}$$

provided the conditions in (5.53) hold.

Following the familiar pattern of KAM theory we are now going to iterate the above scheme, *i.e. we shall label by indices $j=0,1,2...$ the Hamiltonians h_j+f_j together with their size parameters $(\varepsilon_j,E_j,...)$ obtained by sequentially applying the above scheme.* This procedure makes sense *provided* the analogous of condition (5.53) are satisfied at each step of the construction.

We claim that one can find B,β depending only on l, τ and large enough so that:

$$B\hat{\xi}_0^{-\beta}(\varepsilon_0\rho_0^{-1}E_0^{-1})(E_0C_0)^6(\eta_0E_0\rho_0^{-1})^3<1 \tag{5.57}$$

implies that the above scheme can be carried out an infinite number of times.

To prove the claim we proceed by induction and to simplify the discussion we introduce the following *dimensionless* parameters:

$$\bar{\varepsilon}_j=\varepsilon_j\rho_j^{-1}E_j^{-1},\quad \bar{E}_j=E_jC_j,\quad \bar{\eta}_j=\eta_jE_j\rho_j^{-1} \tag{5.58}$$

and a number 1^- which is any prefixed number less than 1. Given 1^- we fix the so far free χ so that $\chi>1-1^-\equiv 0^+$ (one could already say that χ is any prefixed number close to 0 (*e.g.* 1/4) and 1^- is a free parameter to be eventually fixed slightly above 1/2: but we prefer to keep the parameters free as the inequalities look probably more transparent in this way).

Furthermore we impose the following conditions which permit simple bounds on the r.h.s. of (5.56) and (5.53):

$$
\begin{aligned}
&B_8\delta_j^{-\beta_8}\bar\varepsilon_j^{1^-}<1,\qquad B_8\xi_j^{-\beta_8}\delta_j^{-\beta_8}\bar\varepsilon_j^{1^-}(\log\bar\varepsilon_j^{-1})^{2(\tau+1)})\bar E_j^3<1,\quad B_8\delta_j^{-\beta_8}\bar\varepsilon_j^{1^-}\bar\eta_j<1\\
&B_6\xi_j^{-\beta_6}\delta_j^{-\beta_6}\bar\varepsilon_j^{1^-}\bar\eta_j^2\bar E_j(-\log\bar\varepsilon_j)^{\tau+1}<1
\end{aligned}
\tag{5.59}
$$

and in terms of this definition we fix the definition of the parameters verifying the analogous of (5.56) for general j as follows:

$$
\begin{aligned}
\bar E_{j+1}=&\bar E_j(1+\bar\varepsilon_j^{0^+}),\qquad \bar\varepsilon_{j+1}=\bar\varepsilon_j^{1^+},\qquad \xi_{j+1}=e^{-4\delta_j}\xi_j\\
\eta_{j+1}=&\eta_j(1+\bar\varepsilon_j^{0^+}),\qquad \lambda_{j+1}=\lambda_j+\bar\varepsilon_j^{0^+},\qquad \rho_{j+1}=\frac{\rho_j(\xi_j\delta_j)^{\beta_7}}{B_7\bar E_j\log\bar\varepsilon_j^{-1}}\\
C_{j+1}=&C_j,\qquad \Gamma_{j+1}=\Gamma_j
\end{aligned}
\tag{5.60}
$$

where $1^+=2-1^-$, $0^+=1-1^-$ (and for $j=0$, $\xi_{j=0}=\hat\xi_0$).

Hence if $\bar\varepsilon_0$ is small enough (depending on the value chosen for 1^-) we see that $\forall j$:

$$
E_j\le\sqrt2E_0,\qquad \eta_j\le\sqrt2\eta_0,\qquad \lambda_j<2\lambda_0,\qquad \kappa_j\ge\kappa_0/2,\qquad \xi_j\ge\hat\xi_0/2 \tag{5.61}
$$

so that, if $\bar\varepsilon_0$ is small enough compared to 1 (depending on the choice of the number denoted 1^-) and if $\lambda_0E_0\Gamma_0$ is small enough (*i.e.* $<8^{-1}$), it will be $\lambda_jE_j\Gamma_j<4^{-1}$ and:

$$
\begin{aligned}
\rho_{j+1}\ge&\rho_j\frac{\hat\xi_0^{\beta_7}}{2B_7(E_0C_0)}\frac{2^4(\log2)^{-1}}{(1+j^2)^{\beta_7}}\frac{1}{(1^+)^j\log\bar\varepsilon_0^{-1}}\ge\\
\ge&\rho_0\big(B_7'\hat\xi_0^{-\beta_7}(E_0C_0)\log\bar\varepsilon_0^{-1}\big)^{-j-1}(1+j)^2!^{-\beta_7}(1^+)^{-j(j+1)/2}\equiv\sigma_j^{-1}
\end{aligned}
\tag{5.62}
$$

$$
\bar\eta_j\le4\bar\eta_0\sigma_j
$$

for some B_7'.

Thus we see, by taking into account the rapidity of convergence to zero of $\bar\varepsilon_j$ and if β_-,B_- are suitably large, that the conditions in (5.59) are equivalent to:

$$
\begin{aligned}
&B_-\bar\varepsilon_0^{1^-}<1,\qquad B_-\hat\xi_0^{-\beta_-}\bar\varepsilon_0^{1^{--}}\bar E_0^3<1\\
&B_-\hat\xi_0^{-\beta_-}(\bar\varepsilon_0)^{(1+)^j\cdot1^{--}}\bar E_0^j\sigma_j^2\bar\eta_0^2\bar E_0<1,\qquad j\ge0
\end{aligned}
\tag{5.63}
$$

if 1^{--} is defined to be slightly smaller (by any prefixed amount) than the value fixed for 1^- appearing in (5.59).

Choosing $1^-,1^{--}$ slightly larger than $1/2$, and taking into account the expression in (5.62) for σ_j, it follows that all conditions are implied by the following:

$$
B_9\hat\xi_0^{-\beta_9}\bar\varepsilon_0\bar E_0^6\bar\eta_0^3<1 \tag{5.64}
$$

where B_9, β_9 are constants depending only on l.

The above discussion contains some "hidden" assumptions on the initial data, namely (5.20),(5.9), and $8\lambda_0 E_0 \Gamma_0 < 1$. They are verified automatically if the parameters $\bar\rho_0, \bar\kappa_0, \bar\mu_0$ are chosen as prescribed by (5.18),(5.20).

Hence we can say that (5.64) together with (5.9),(5.20) and $4\lambda_0 E_0 \Gamma_0 < 1$ are implied by:

$$B_9 \hat\xi_0^{-\beta_9} \bar\varepsilon_0 \bar E_0^6 \bar\eta_0^3 < 1 \tag{5.65}$$

possibly readjusting B_9, β_9 (recall, as well, that $\Gamma_0 < C_0$). As we shall see below this is the final condition under which lemma 1' holds.

This completes our check of the claim in (5.57).

Thus we can construct, for all $j \geq 0$, canonical transformations $\tilde{\mathcal{C}}_j$ ($\tilde{\mathcal{C}}_0 = \tilde{\mathcal{C}}$) mapping W_{j+1} into W_j (W_j is defined in (5.54) with $1 \to j$; recall that $\mu_j \equiv \mu_0$). Such maps are close to the identity within $\|\Phi_j\|\rho_{j+1}^{-1}$ in the $\vec A$–variables and within $\|\Phi_j\|\kappa_{j+1}^{-1}\delta_{j+1}^{-1}$ in the p, q–variables and within $||\Phi_j||\xi_{j+1}^{-1}\delta_{j+1}^{-1}$ in the $\vec\alpha$–variables.

Their derivatives of order k in $\vec A$'s, h in the $\vec\alpha$'s, z in the p, q are bounded by multiplying the above bounds by $\rho_{j+1}^{-k}\kappa_{j+1}^{-z}\delta_{j+1}^{-h-z}$. Since $\|\Phi_j\| \leq B_2 \hat\xi_j^{-\beta_2}\delta_j^{-\beta_2}\varepsilon_j C_0$, see (5.35), (5.60), we realize that the map $\tilde{\mathcal{C}}_j$ approaches the identity very quickly.

Taking into account the (5.52) it also follows that the sheets $\bar{\mathcal{L}}^j$ defined by $(s, u) \in \tilde{\mathcal{I}}_0 \to \vec A^j(s, u, 0, \mu)$ approach a limit sheet:

$$\bar{\mathcal{L}}^\infty \text{ defined by } \quad (s, u) \in \tilde{\mathcal{I}}_0 \to \vec A^\infty(s, u, 0, \mu) \ , \quad \vec A^\infty(s, u, 0, 0) \equiv \vec A_{su} \tag{5.66}$$

and control is kept on any prefixed number of derivatives of $\bar{\mathcal{L}}^\infty$: here we have used that, see (5.52), $\chi > 0$.

Furthermore the domains of holomorphy of the maps $\tilde{\mathcal{C}}_j$, hence of $\tilde{\mathcal{C}}_0\tilde{\mathcal{C}}_1 ... \tilde{\mathcal{C}}_j = \tilde{\mathcal{C}}^{(j)}$ do not shrink to zero in the $\vec\alpha, p, q, \mu$ variables.

If we call $\tilde\Phi_j(\vec A', \vec\alpha, p', q, \mu)$ the generating function of the composite map $\tilde{\mathcal{C}}^{(j)}$, the above remarks imply that $\tilde\Phi_j$ can be extended to a C^n function defined in the vicinity of the sets W_j: the extension, which we still denote $\tilde\Phi_j$, can be made in class C^n for any n so that $\tilde\Phi_j$ converges in the C^n-norm to a limit $\tilde\Phi_\infty$ (simply because the variations of the $\tilde\Phi_j$'s are basically bounded as the Φ_j, *i.e.* by $\tilde B_2 \hat\xi_0^{-\beta_2}\bar\varepsilon_0^{(3/2)^j}\bar E_0 \rho_j$ see (5.35), (5.58), (5.61),(5.62), in their analyticity domain; hence they have their derivatives very small and therefore can be extended remaining small), see [La], [Sv], [CG] and [Pö] for similar constructions.

The limit $\tilde\Phi_\infty$ will be uniquely defined on $\bar{\mathcal{L}}^\infty \times T^{l-1} \times \{|p'| < \kappa_\infty, \ |q| < \kappa_\infty\}$, with (cfr. 2) of lemma 1):

$$\kappa_\infty = \frac{\kappa_0}{2} \equiv \bar\kappa \tag{5.67}$$

and it will be real-analytic in the $\vec\alpha_0, p', q_0$ variables, C^n–smooth in $\vec A', \vec\alpha_0, p', q, \mu$ prefixed number n of derivatives if ε_0 is small enough.

Therefore $\tilde{\Phi}_\infty$ generates a canonical map, $\tilde{\mathcal{C}}_\infty$, which for $\vec{A}' \in \bar{\mathcal{L}}^\infty$ takes the form (cfr. (5.5)):

$$\begin{aligned} \vec{A}_0 =& \vec{A}' + \vec{\Xi}'(\vec{A}', \vec{\psi}, p, q, \mu) \qquad & \vec{\alpha}_0 = \vec{\psi} + \vec{\Delta}'(\vec{A}', \vec{\psi}, p, q, \mu) \\ p_0 =& p + \Lambda'(\vec{A}', \vec{\psi}, p, q, \mu) \qquad & q_0 = q + \Theta'(\vec{A}', \vec{\psi}, p, q, \mu) \end{aligned} \tag{5.68}$$

and for $\vec{A}' = \vec{A}^\infty(s, u, pq, \mu)$ the solutions of the motion equations take the form (5.11) with $\vec{\omega}_{su}, g_{su}$ defined in (5.7), (5.9) and with $\gamma' \equiv u'_\infty(s, u, J, \mu)$ defined by (cfr. (5.16)):

$$\partial_J h_\infty(\vec{A}^\infty(s, u, J, \mu), J, \mu) = (1 + u'_\infty) g_{su} \equiv (1 + \gamma') g_{su} \tag{5.69}$$

where $h_\infty \equiv \lim_{j \to \infty} h_j$. Note that if we denote by $H_\infty(\vec{A}', \vec{\psi}, p, q, \mu)$ the original hamiltonian (5.2) computed in the new variables defined by $\tilde{\Phi}_\infty$, it is:

$$h_\infty(\vec{A}^\infty(s, u, pq, \mu), pq, \mu) = H_\infty(\vec{A}^\infty(s, u, pq, \mu), \vec{\psi}, pq, \mu) \tag{5.70}$$

Furthermore $u'_\infty(s, u, pq, \mu)$ is analytic in p, q, μ, if (s, u) are fixed in $\tilde{\mathcal{I}}_0$. The parametric equations of the whiskers (5.10) are now immediately obtained in terms of (5.68) and of the transformation (5.3) of lemma 0. Setting:

$$\begin{aligned} & z \equiv (\vec{A}^\infty(s, u, pq, \mu), \vec{\psi}, p, q, \mu) \quad \hat{z} \equiv (\vec{A}^\infty(s, u, pq, \mu), p, q, \mu) \\ & \hat{\zeta} \equiv \hat{z} + \Big(\vec{\Xi}'(z), \Lambda'(z), \Theta'(z), 0\Big) \end{aligned} \tag{5.71}$$

we find (cfr. (5.10)):

$$\begin{aligned} & \vec{A}_{su}(J, \mu) \equiv \vec{A}^\infty(s, u, J, \mu) \ , \qquad (\Rightarrow \ \vec{A}_{su}(0, 0) \equiv \vec{A}_{su}) \\ & \vec{\Xi}(\vec{\psi}, p, q, s, u, \mu) \equiv \vec{\Xi}'(z) \ , \quad \vec{\Delta}(\vec{\psi}, p, q, s, u, \mu) \equiv \vec{\Delta}'(z) + \vec{\delta}(\hat{\zeta}) - \vec{\delta}(\hat{z}) \\ & \Lambda(\vec{\psi}, p, q, s, u, \mu) \equiv R(\hat{\zeta}) - R(\hat{z}) \ , \quad \Theta(\vec{\psi}, p, q, s, u, \mu) \equiv S(\hat{\zeta}) - S(\hat{z}) \end{aligned} \tag{5.72}$$

The linearity of the flow on the surfaces (5.10) follows because f_j tends to zero very fast with all its derivatives, including the $\vec{A}$ derivatives in spite of the fact that the $\vec{A}$-domain shrinks: in fact the derivatives are bounded, for real $\vec{A}, \vec{\alpha}, p, q$, by ε_j times some inverse power of ρ_j and $\varepsilon_j \rho_j^{-k} \to 0$ for all $k \geq 0$, by the inequality (5.62).

Note that the H_∞, by our construction, has derivatives with respect to $\vec{\psi}$ vanishing if $\vec{A}' = \vec{A}^\infty(s, u, pq, \mu)$, $(s, u) \in \tilde{\mathcal{I}}_0$, see (5.70); points it depends non trivially on $\vec{\psi}$. Hence for $p = q = 0$, $\vec{A}' = \vec{A}^\infty(s, u, 0, \mu)$ and $(s, u) \in \tilde{\mathcal{I}}_0$ the (5.68) describe invariant tori $\mathcal{T}_\mu(s, u)$, and their whiskers are obtained by considering $p = 0$, $q \neq 0$ or $q = 0$ and $p \neq 0$.

We express (5.65) in terms of the more fundamental parameters $\bar{\kappa}_0, \bar{\rho}_0, \bar{\xi}_0$ of lemma 0; see (5.18), (5.25), (5.58). If we assume for simplicity that for a suitable constant $\bar{B}_0 \geq 1$ one has:

$$\bar{B}_0 \bar{\rho}_0 \leq \bar{\kappa}_0^2 \tag{5.73}$$

so that (see (5.18)):

$$\rho_0 = \frac{\bar\rho_0}{\bar B(E_0\eta_0\bar\rho_0^{-1})(E_0\Gamma_0)^2} \tag{5.74}$$

then (5.65) becomes:

$$B_{10}(\varepsilon_0 E_0^{-1}\bar\rho_0^{-1})(E_0C_0)^6(E_0\Gamma_0)^8(E_0\eta_0\bar\rho_0^{-1})^7\hat\xi^{-\beta_{10}} < 1 \tag{5.75}$$

Finally, we see that, in the case of interest to us, ε_0 *is of the form* $\mu\widetilde\varepsilon_0$ *for some* $\widetilde\varepsilon_0$, so that the condition of μ small takes the form (see (5.75)):

$$|\mu| < \mu_0 \equiv \big[B_{11}\hat\xi_0^{-\beta_{11}}(\widetilde\varepsilon_0\bar\rho_0^{-1}E_0^{-1})(E_0C_0)^6(E_0\Gamma_0)^8(\eta_0E_0\bar\rho_0^{-1})^7\big]^{-1} \tag{5.76}$$

provided $2\Gamma_0 < C_0$ and for suitable constants B_{11},β_{11}.

This is still not completely explicit as the values of $\bar\rho_0,\bar\kappa_0,\bar\xi_0$ are not the analyticity parameters of the original hamiltonian. In fact they can be deduced from the latter via the application of lemma 0.

Lemma 0 allows us to take (see appendix A3, ♣A3.39 and (5.73)):

$$\begin{aligned}\bar\kappa_0 \ &= \frac{1}{2}\frac{\kappa}{Bm^7}\min\Big\{\frac{\rho'}{\kappa^2}\ ,\ \xi'\ ,\ \frac{1}{(\mathcal{E}\Gamma\kappa^{-2})\sigma_2^2\sigma_3}\ ,\ \frac{\rho\xi}{\rho'\hat\sigma}\Big\}\\ \bar\rho_0 &\equiv\ \min\{\frac{\rho}{2}\ ,\ \frac{\bar\kappa_0^2}{\bar B_0}\}\ , \qquad \bar\xi_0\equiv\frac{\xi}{2}\end{aligned} \tag{5.77}$$

if ρ',ξ',ρ,ξ are the original hamiltonian regularity parameters (see §2), and $m,\Gamma,\mathcal{E},B,\kappa$, σ_2, $\sigma_3,\hat\sigma$ are introduced in appendix A3, see ♣A3.3,♣A3.47, ♣A3.49 and ♣A3.53,♣A3.39.

We can also deduce, from the analyticity in μ, a simple bound on the size of the variation $|\vec A^\infty(s,u,J,\mu)-\vec A^0(s,u,J,\mu)|$ and of the variation of the *whisker graphs*, *i.e.* of the functions in the r.h.s. of (5.68) and, by dimensional estimates, consequent bounds on their derivatives. We see from the above analysis that the bounds (5.41),(5.52) must hold, with different constants replacing β_3,B_3 for the corresponding functions in (5.68). Hence for suitable constants $G_A,G_{\mathcal{C}}$:

$$\begin{aligned}&\bar\rho_0^{-1}(||\vec A^\infty(s,u,\cdot)-\vec A^0(s,u,\cdot)||)\le G_A|\mu|\mu_0^{-1}\\ &\bar\rho_0^{-1}||\vec\Xi'||+\bar\xi_0^{-1}||\vec\Delta'||+\rho'^{-1}||\Lambda'||+\xi'^{-1}||\Theta'||<G_{\mathcal{C}}|\mu|\mu_0^{-1}\end{aligned} \tag{5.78}$$

where the norms are evaluated by fixing $(s,u)\in\bar{\mathcal{L}}$; here we have just bounded the value at z of a function holomorphic in a disk of radius z_0 and vanishing at the center $z=0$ by its supremum times $|z|/z_0$: we take $z=\mu$ and use the holomorphy in μ.

And, using (5.52) and (5.35),(5.41) the constants $G_A,G_{\mathcal{C}}$ can be easily expressed in terms of our dimensionless constants:

$$\begin{aligned}G_A =&B_{12}\big[(\widetilde\varepsilon_0E_0^{-1}\bar\rho_0^{-1})(\eta_0E_0\bar\rho_0^{-1})\big]^\chi\\ G_{\mathcal{C}} =&B_{12}(E_0\Gamma_0)\hat\xi_0^{-\beta_{12}}\ \big[(\widetilde\varepsilon_0E_0^{-1}\bar\rho_0^{-1})(E_0C_0)(\eta_0E_0\bar\rho_0^{-1})\big]^>\end{aligned} \tag{5.79}$$

for suitably chosen constants B_{12}, β_{12} and having denoted $[x]^>$ the function $x \log x^{-1}$ for $x > 1$.

The function γ' in (5.11) is the value u'_∞ in (5.69): it is analytic in $|J| < \kappa_0^2/2$ (see (5.61)) and $|\mu| < |\mu_0|$ and it is bounded there, for all $s, u \in \tilde{\mathcal{I}}_0$ (see (5.20), (5.17), (5.61)), by:

$$|\gamma'(J, s, u, \mu)| \le 2\hat{B}(\Gamma_0 E_0)(E_0 \eta_0 \bar{\rho}_0^{-1}) \Big(\frac{|J|}{\kappa_0^2} + \frac{|\mu|}{\mu_0} \Big) \tag{5.80}$$

To check (5.4) we simply use that the above proof has a free parameter C_0. The set $\Sigma(C_0)$, see (3.1), has measure at least $(s_2 - s_1)[1 - (\bar{K}/(DC_0)^{1/t})]$ by the assumption that $\mathcal{L}$ is a diffusion path, see (3.2). Therefore we choose, taking into account that the constant C_0 appears to the power 6 in the basic condition (5.75):

$$C_0 \equiv \Gamma_0 |\mu|^{-1/7} \qquad \Rightarrow \ \Sigma_\mu = \Sigma(C_0) \tag{5.81}$$

Then we see that the constants $k, K, \bar{c}$ and c of lemma 1' can be taken:

$$k = \Gamma_0, \quad K = \bar{K}(E_0 \Gamma_0)^{-1/t}\ , \quad \bar{c} \equiv 7\ , \quad c \equiv 7\ t \tag{5.82}$$

(where we have replaced D by E_0: see 2) §3) and, what is more important, the smallness condition on μ can still be met.

This finishes the proof of lemma 1'. Note that the smallness condition on $|\mu|$ (*i.e.* (5.75) with $C_0 \equiv \Gamma_0 |\mu|^{-1/7}$) does not involve ϑ_0 (defined in (5.23)): such a quantity will appear in fixing the energy in order to get lemma 1 as a corollary of lemma 1'.

We now let p, q be such that $|p|, |q| < \kappa_0/2$, $J \equiv pq$, $\vec{\psi} = \vec{0}$, $\vec{A}^\infty \equiv \vec{A}^\infty(s, u, J, \mu)$ and, fixing $s \in \mathcal{I}_0$, we try to find $u \equiv u(s, J, \mu)$ so that the real part of the energy $E(s, u, J, \mu)$ associated to the initial data $(\vec{A}^\infty, \vec{0}, p, q, \mu)$ coincide with the *prefixed value* $E \equiv h_0(\vec{A}_s, 0, 0)$ (see 1) of §3). In view of the above construction, the energy $E(s, u, J, \mu)$ is given by (compare with (5.69), (5.70)):

$$E(s, u, J, \mu) = h_\infty(\vec{A}^\infty(s, u, J, \mu), J, \mu) \tag{5.83}$$

and by Taylor expansion at $\mu = 0$, $u = 0$ (see (5.7),(5.8), (5.66)):

$$\begin{aligned} \mathrm{Re}\, E(s, u, J, \mu) &= E + u \Big(\partial_{\vec{A}} h_0(\vec{A}_s, 0, 0) \cdot [\partial_u \vec{A}_{su}]_{\mu=0} \Big) + \beta \tilde{G} \Big(\frac{u^2}{\bar{u}^2} + \frac{|\mu|}{\mu_0} + \frac{|J|}{\kappa_0^2} \Big) = \\ &= E + u\, [\vec{\omega}_s \cdot (\partial_{\vec{A}}^2 h_0)^{-1} \vec{\omega}_s] + \beta \tilde{G} \Big(\frac{u^2}{\bar{u}^2} + \frac{|\mu|}{\mu_0} + \frac{|J|}{\kappa_0^2} \Big) \end{aligned} \tag{5.84}$$

where the derivative $\partial_u \vec{A}_{su}$ is computed by differentiating (5.8) and $\partial_{\vec{A}}^2 h_0 \equiv \partial_{\vec{A}}^2 h_0(\vec{A}_s, 0, 0)$; β is some C^∞ function (at s fixed) with $|\beta| \le 1$ and the constant $\tilde{G}$ can be taken to be proportional, via a constant depending only on l, to $E_0(E_0 \eta_0 \bar{\rho}_0^{-1})^2$ (recall that the constant $\bar{u}$, see (5.9), can be taken to be a numerical constant times $(\eta_0 E_0 \bar{\rho}_0^{-1})^{-2}$, see (5.18)).

The first two derivatives of β with respect to u, μ, J can be bounded by our dimensionless constants. Hence, recalling the definition in (5.23) of ϑ_0, we see, by the implicit function theorem, that under the further condition:

$$B_{13}(|\mu|\mu_0^{-1})(\vartheta_0 E_0 \bar\rho_0)(\eta_0 E_0 \bar\rho_0^{-1})^2 \equiv B_{13}(|\mu|\mu_0^{-1})(\vartheta_0 E_0^2 \eta_0)(E_0 \eta_0 \bar\rho_0^{-1}) < 1 \tag{5.85}$$

we can find $u = u(s, J, \mu)$ as desired *i.e.* so that (see (5.83)):

$$E(s, u(s, J, \mu), J, \mu) \equiv E \tag{5.86}$$

Therefore condition (5.85) together with (5.75) with $C_0 \equiv \Gamma_0|\mu|^{-1/7}$ are sufficient to yield lemma 1' *and* lemma 1.

The functions $\gamma', \vec\Xi, \vec\Delta, \Lambda, \Theta$ of lemma 1 are obviously related to the corresponding (but different) functions of lemma 1': just set $u \equiv u(s, J, \mu)$ in (5.72), (5.71) and in the definition of γ' (cfr. (5.69)); *e.g.* :

$$\gamma'(J, s, \mu) \equiv \gamma'(J, s, u(J, s, \mu), \mu) \ , \quad \vec\Xi(\vec\psi, p, q, s, \mu) \equiv \vec\Xi(\vec\psi, p, q, s, u(J, s, \mu), \mu) \ , \ etc \tag{5.87}$$

and the function $\gamma(J, s, \mu)$ is just $\gamma \equiv u(J, s, \mu)$. The functions γ, γ' are easily seen to satisfy the bound (cfr. (5.69),(5.17),(5.18)):

$$\sup_{\substack{s\in\mathcal{I}_0, \mu\in[-\mu_0\cdot\mu_0] \\ |J|<\kappa_0^2/2}} |\gamma|, |\gamma'| \le \bar u \equiv \bar B^2 (E_0 \eta_0 \bar\rho_0^{-1})^2 \tag{5.88}$$

Bounds on $|\vec\Xi|, |\vec\Delta|, |\Lambda|, |\Theta|$ are easily obtained by recalling their definitions, (5.71), (5.72) (for the functions of lemma 1') and (5.87) (for the functions of lemma 1), the bounds (5.78), (5.79), and the bounds on $|R|, |S|, |\vec\delta|$ (see ♣A3.54):

$$\bar\rho_0^{-1}||\vec\Xi|| + \bar\xi_0^{-1}||\vec\Delta|| + {\rho'}^{-1}||\Lambda|| + {\xi'}^{-1}||\Theta|| < G_{\mathcal{C}}|\mu|\mu_0^{-1} \tag{5.89}$$

where $G_{\mathcal{C}}$ ia as in (5.79) (actually increased by a factor 2) and the norms are taken at $s \in \mathcal{I}_0$ fixed for the functions of lemma 1, or at $(s, u) \in \bar{\mathcal{L}}$ fixed for the functions of lemma 1'.

Finally, we remark that all the requirements ((5.75),(5.76),(5.81),(5.85)) we needed to prove lemma 1',1 can be enforced by requiring the single condition:

$$|\mu| < \mu^* \equiv \Big[\big[B\hat\xi_0^{-\beta} (\eta_0 E_0 \bar\rho_0^{-1})^7 (\Gamma_0 E_0)^{14} (\widetilde\varepsilon_0 \bar\rho_0^{-1} E_0^{-1}) \big]^7 (\vartheta_0 E_0^2 \eta_0) \Big]^{-1} \tag{5.90}$$

where $B, \beta > 0$ are suitable constants depending only on l and τ (see (5.23), (5.25), lemma 0 and (5.76) to refresh the memory about the various parameters involved).

This completes the proof of lemma 1' and lemma 1. In fact we have proved:

Lemma 2: *There exists a canonical map $\mathcal{C}(p',q',\vec{A}',\vec{\alpha}')=(I,\vec{A},\varphi,\vec{\alpha})$ of class C^n and a line $\mathcal{L}_\mu : s\to\vec{A}_\mu(s)$, contained in the energy surface of energy E for the perturbed hamiltonian (2.9), of class C^n with the properties:*

1) *$\mathcal{C}$ is C^n-close to the identity as $\mu\to0$, $\mathcal{L}_\mu$ is C^n-close to $\mathcal{L}$ as $\mu\to0$ and the domain of $\mathcal{C}$ is a set of the form $V\times T^{l-1}\times S^2$ where V is a neighborhood of $\mathcal{L}$ containing $\mathcal{L}_\mu$ and S^2 is a neighborhood of the origin in R^2.*
2) *for s in a set of measure $\ge(1-K|\mu|^{1/c}), K,c>0$, the set $\mathcal{C}(\vec{A}_\mu(s)\times T^{l-1}\times S^2)$ is invariant for the flow generated by the perturbed hamiltonian (2.9);*
3) *the derivatives in $\vec{A}'$ of the hamiltonian (2.9) regarded as a function of the new coordinates $(p',q',\vec{A}',\vec{\alpha}')$ as well as those in p',q' at constant $p'q'$ vanish on the above set, so that the flow is linear in the $\vec{\psi}\in T^{l-1}$ variables and hyperbolic in the p,q variables.*
4) *explicit bounds on the parametric equations of the invariant tori, on their whiskers and on the main dimensionless parameters involved in the construction are provided by the bounds found in the course of the above proof.*

This lemma is a quick if a little mysterious, way of summarizing the analysis of this section.

Another important *corollary of the above lemmata is that they can be shown to cover the case of a forced system*:

$$H\equiv\omega B+H_0+\mu f\equiv\omega B+h(\tilde{A},\mu)+P(I,\tilde{A},\varphi,\mu)+\mu f(I,\tilde{A},\varphi,\tilde{\alpha},\varphi,\lambda,\mu)\tag{5.91}$$

where B,λ are a pair of conjugate action angle variables and $(\tilde{A},\tilde{a})\in R^{l-2}\times T^{l-2}$ are other action angle coordinates which will be supposed anisochronous, *i.e.* such that $||(\partial^2_{\tilde{A}}h)^{-1}||=\eta_0<+\infty$; to compare with the previous notations one should set $\vec{A}\equiv(B,\tilde{A})$ and $\vec{\alpha}\equiv(\lambda,\tilde{\alpha})$.

In this case the notion of diffusion path has to be suitably adapted. We consider a curve in $\tilde{A}$ space, $\mathcal{L}=\{s\to\tilde{A}_s\}$ and define $\Sigma(C)$ exactly as in (3.1) with $\vec{A}$ replaced by $\tilde{A}$ and we say that $\mathcal{L}$ is a diffusion path if (3.2) holds. In other words in forced systems *the action of the "forcing reservoir" B does not enter into the definition.*

The following technique, invented by Poincaré, applies remarkably well to this case, see [P] p. 118, tome I, ch. III. Note that, if $H_0(I,\tilde{A},\varphi)\equiv h(\tilde{A},\mu)+P(I,\tilde{A},\varphi,\mu)$ and H is as in (5.91), the hamiltonian:

$$\begin{aligned}&H_2(\vec{A},\vec{\alpha},\mu)\equiv\frac{H^2}{2E}\equiv h_2+f_2\ ,\qquad \vec{A}\equiv(B,\tilde{A})\ ,\ \vec{\alpha}\equiv(\lambda,\tilde{\alpha})\\&h_2(I,B,\tilde{A},\varphi)\equiv\frac{1}{2E}\big(\omega^2B^2+H_0^2+2\omega BH_0\big)\end{aligned}\tag{5.92}$$

where $E\neq0$ is fixed arbitrarily, is such that h_2 has the property that [see 1) of appendix A9]:

$$\det(\partial^2_{\vec{A}}h_2)\equiv\det(\partial^2_{B\tilde{A}}h_2)=\frac{\omega^2}{E}\Big(\frac{h_2}{E}\Big)^{l-2}\det(\partial^2_{\tilde{A}}h)\tag{5.93}$$

where in (5.93) we evaluate the derivatives at $I=0$, $\varphi=0$ (hence $P=0$). Thus h_2 is non degenerate and, furthermore, at $I=0$, $\varphi=0$:

$$\partial_{\vec A} h_2 \equiv (\partial_B h_2, \partial_{\tilde A} h_2) \equiv (\omega,\tilde\omega) \qquad \text{if} \qquad h_2|_{(I,\varphi)=(0,0)} \equiv \omega B + h = E \tag{5.94}$$

and we see that the line $\mathcal{L}_2$ obtained form $\mathcal{L}$ by adding to each of its points a coordinate B_s computed from the equation $\omega B + h(\tilde A_s,0) = E$ is a diffusion path for h_2 in the sense of §2.

It is immediate to check that if $z(t)\equiv(\vec A(t),\vec\alpha(t))$ is a motion for (5.92), then $\bar z(t) \equiv z(t/\sigma)$, with $\sigma\equiv(H_2(\vec A(0),\vec\alpha(0),\mu)/E)$, is a motion for (5.91). We can thus construct, by using the above lemmata, whiskered tori for h_2+f_2 (and hence for H).

For a proper usage of the bounds involved in the above lemmata, one has to estimate the basic dimensional quantities. Fixing the arbitrary parameter $E\equiv 4\bar\rho_0\max\{||\partial_{\tilde A}h||,|\omega|\}$ we see that:

$$||\frac{\omega B + h(\tilde A)}{E}-1|| \equiv \sup_{\substack{|B-B_s|\le\bar\rho_0\\ ||\tilde A-\tilde A_s||\le\bar\rho_0}} |\frac{\omega B+h(\tilde A)}{E}-1| \le \frac{1}{2}\,, \qquad B_s\equiv\frac{E-h(\tilde A_s)}{\omega} \tag{5.95}$$

and this allows to bound the norms, $||\partial h_2||$, $||(\partial_J h_2)^{-1}||$, $||f_2||$, associated to H_2 in (5.92) *in terms of constant times the corresponding quantities for (5.91)*, while we can take $E_0\equiv\tilde\beta(||\partial_{\tilde A}h||+|\omega|)$ and $\vartheta_0\equiv\tilde\beta(E_0\bar\rho_0)^{-1}$ [see also 1) of appendix A9] for a suitable constant $\tilde\beta$; of course the norms referring to (5.92) are taken over the action domain $|B-B_s|\le\bar\rho_0$, $||\tilde A-\tilde A_s||\le\bar\rho_0$ (see (5.95)). Thus we see that the statements of lemma 1' and lemma 1 just carry over to the present case under a condition like (5.90) with:

$$\mu^* = \Big[\big[B\hat\xi_0^{-\beta}(\eta_0E_0\rho_0^{-1})^7(\Gamma_0E_0)^{14}(\widetilde\varepsilon_0\rho_0^{-1}E_0^{-1})\big]^7(\eta_0E_0\rho_0^{-1})\Big]^{-1} \tag{5.96}$$

and with the same quantitative bounds established in the proofs, provided we interpret the notion of diffusion path in the way described above.

Finally, we remark that the whiskered tori that we obtain for (5.91) via lemma 1' applied to (5.92) and via the rescaling described after (5.94), have, *for all* $s\in\Sigma_\mu$ *and* $u\in[-\bar u,\bar u]$, the λ–frequency equal to ω (as it should as the clock velocity ω cannot change, just because it is a clock).

§6 Large whiskers. Homoclinic points and angles

In §5 we have constructed invariant tori surviving the onset of a perturbation as well as the parts of their whiskers in their immediate vicinity. We now derive the equations of the whiskers away from the invariant tori with the purpose of finding whether they contain homoclinic intersections.

The whiskers can be continued to form a full invariant manifold by evolving them with the solution map $(I, \vec{A}, \varphi, \vec{\alpha}) \to S_t^\mu(I, \vec{A}, \varphi, \vec{\alpha})$ associated with the perturbed Hamilton equations generated by (2.9). We regard the map S_t^μ as defined in the original coordinates, which are *globally* describing our system and we shall call *local* the part of the whiskers constructed so far, via lemma 1', denoting it by $W^{loc}(s, u)$.

The full stable whisker will be:

$$W_{stable}(s, u) = \bigcup_{t \leq 0} S_t^\mu \left\{ (I, \vec{A}, \varphi, \vec{\alpha}) \in W_{stable}^{loc}(s, u) \right\} \tag{6.1}$$

for values of u small and $s \in \mathcal{I}_0 \equiv \Sigma(C_0)$, see lemma 1' §5 and (5.20). Lemmata 0,1' imply that this set can be described, for $|\varphi| < \tilde{\varphi}$, and $|\mu|$ small enough, by parametric equations:

$$I = I(\varphi, \vec{\alpha}, \mu) \ , \qquad \vec{A}(\varphi, \vec{\alpha}, \mu), \qquad |\varphi| < \tilde{\varphi}, \ \vec{\alpha} \in T^{l-1} \tag{6.2}$$

where $\tilde{\varphi}$ is *a priori* fixed in the following discussion (to be $< 2\pi$ in the case of an *open* (see §4) separatrix while for *closed* separatrices, $|\varphi| < \tilde{\varphi}$ should be replaced by $\varphi_{\min} + \delta < \varphi < \varphi_{max} - \delta$ for some $\delta > 0$). To fix ideas and simplify notations we shall discuss here mainly the *stable case*, the unstable one being completely analogous (see also below); however, when needed, we shall attach to the above functions (6.2) superscripts to distinguish among the two different cases (such superscripts should not be confused with the parameters s, u in (6.1) and elsewhere).

Fixing $s \in \mathcal{I}_0$ and u small, the functions $\vec{A}(\varphi, \vec{\alpha}, \mu), I(\varphi, \vec{\alpha}, \mu)$ have to be such that for any $|\varphi'| < \tilde{\varphi}, \vec{\alpha}' \in T^{l-1}$ there are $\varphi, \vec{\alpha}$ such that:

$$S_t^\mu(I(\varphi', \vec{\alpha}', \mu), \varphi', \vec{A}(\varphi', \vec{\alpha}', \mu), \vec{\alpha}') = (I(\varphi, \vec{\alpha}, \mu), \varphi, \vec{A}(\varphi, \vec{\alpha}, \mu), \vec{\alpha}) \tag{6.3}$$

and we know, from lemma 1', that for $|\varphi|, |\mu|$ small enough $I(\varphi, \vec{\alpha}, \mu), \vec{A}(\varphi, \vec{\alpha}, \mu)$ are analytic functions in the perturbation parameter μ.

We shall fix $(\varphi, \vec{\alpha})$ and try to determine the functions $I(\varphi, \vec{\alpha}), \vec{A}(\varphi, \vec{\alpha})$.

We begin by noting that S_t^μ is close to $S_t^\mu|_{\mu=0} \equiv S_t^0$ and depends analytically on $t, \vec{\alpha}, \varphi, \vec{A}, I, \mu$. Also S_t^0 expands any $\varphi \neq 0$ to a value larger (in absolute value) than $\tilde{\varphi}$ in a finite (positive or negative) time.

Hence it is clear that, fixed $\tilde{\varphi}$ and given $W_{stable}^{loc}(s, u)$, we can use (6.3) with $|\varphi| < \delta$, $\vec{\alpha} \in T^{l-1}$ and $|t| < t_\delta$ to define $I(\varphi, \vec{\alpha}), \vec{A}(\varphi, \vec{\alpha})$ for $|\varphi| < \tilde{\varphi}$, $\vec{\alpha} \in T^{l-1}$ and δ can be taken to be any prefixed small positive number and t_δ a suitably long (but finite) time.

And the remarked analyticity of S_t^μ together with the analyticity of $W_{stable}^{loc}(s, u)$, see lemma 1,1',2 of §5, imply the analyticity of $I(.), \vec{A}(.)$ in their arguments, at fixed s, u.

¿From now on, in order to avoid confusion with upper *indices* $\sigma = s, u$, indicating stable/unstable, we drop the dependence upon the *parameters* s, u, which in this section will be kept fixed.

Given a hamiltonian $H = H_0 + \mu f$ as in (2.9) we write the equations of motion for the vector $(I, \vec{A}, \varphi, \vec{\alpha}) = X$ as:

$$\dot{X} = G_0(X) + \mu G(X) \tag{6.4}$$

and we remark that if $\varphi, \vec{\alpha} \to (I^s(\varphi, \vec{\alpha}), \vec{A}^s(\varphi, \vec{\alpha}), \varphi, \vec{\alpha})$, for $|\varphi| < \tilde{\varphi}$ and for $\vec{\alpha} \in T^{l-1}$, are the equations of the stable whisker, then:

$$I^s(\varphi, \vec{\alpha}) = I^0(\varphi) + \sum_{k=1}^{\infty} \mu^k v^{ks}(\varphi, \vec{\alpha}) \ , \qquad \vec{A}^s(\varphi, \vec{\alpha}) = \vec{A}^0 + \sum_{k=1}^{\infty} \mu^k \vec{h}^{ks}(\varphi, \vec{\alpha}) \tag{6.5}$$

Since in this section we shall mainly discuss the stable whiskers, we shall also drop the suffix s (for stable) when this does not lead to confusion.

It will be useful to consider also the slightly more general case in which the variable $\vec{\alpha}$ is a function of μ:

$$\vec{\alpha} \equiv \vec{\alpha}_\mu \equiv \sum_{k \geq 0} \vec{\alpha}^k \mu^k \tag{6.6}$$

while the variable φ will be fixed once for all to be the value $\bar{\varphi}$ corresponding to the point where $|I|$ is maximal for the unperturbed hamiltonian (cfr. §4; $\bar{\varphi} = \pi$ for the pendulum (2.1)).

Furthermore if $X^s(t)$ is the solution of the Hamilton equations with initial data $X^s(0) = (I^s(\varphi, \vec{\alpha}), \vec{A}^s(\varphi, \vec{\alpha}), \varphi, \vec{\alpha})$ then, for large enough t, $X^s(t)$ is inside the vicinity of the unperturbed torus $\vec{A} = \vec{A}^0$, $I = \varphi = 0$ where we can use the coordinates, described in §5, $(p, q, \vec{\psi})$.

Actually, by using the analyticity of the flow S_t^μ and the analyticity properties in $p, q, \vec{\psi}$ discussed in §5, one can analytically continue the functions $R(\vec{A}', p, q, \mu), \Lambda(\vec{\psi}, p, q, \mu), \ldots$ in (5.10) to a domain around the real p, q such that $|pq| < \bar{\kappa}^2$ and around the real $\vec{\psi}$, so large to cover a vicinity of the points $p = \bar{p}, q = \bar{q}, \vec{\psi} \in T^{l-1}$ corresponding to $\bar{I}, \bar{\varphi}, \vec{\alpha} \in T^{l-1}$, where $\bar{I}, \bar{\varphi}$ is the separatrix point chosen as the origin. Therefore, after the analytic continuation, we can write for $\mathrm{Re}\, t \geq 0, |\,\mathrm{Im}\, t| \leq \hat{\xi}$ and $\hat{\xi}$ small enough (cfr. (5.10), (5.11) and recall that we are dropping the parameters s, u from the notation):

$$\begin{aligned}
I^s(t) =& R(\vec{A}', pe^{-(1+\gamma')gt}, 0, \mu) + \Lambda(\vec{\psi} + \vec{\omega}t, pe^{-(1+\gamma')gt}, 0, \mu)\\
\vec{A}^s(t) =& \vec{A} + \vec{\Xi}(\vec{\psi} + \vec{\omega}t, pe^{-(1+\gamma')gt}, 0, \mu)\\
\varphi^s(t) =& S(\vec{A}', pe^{-(1+\gamma')gt}, 0, \mu) + \Theta(\vec{\psi} + \vec{\omega}t, pe^{-(1+\gamma')gt}, 0, \mu)\\
\vec{\alpha}^s(t) =& \vec{\psi} + \vec{\omega}t + \vec{\Delta}(\vec{\psi} + \vec{\omega}t, pe^{-(1+\gamma')gt}, 0, \mu) + \vec{\delta}(\vec{A}, pe^{-(1+\gamma')gt}, 0, \mu)
\end{aligned} \tag{6.7}$$

and the expressions of $\vec{\alpha}, \varphi$ in terms of $\vec{\psi}, p$ are deduced from the above relations with $t = 0$. The constant $\hat{\xi}$ is, of course, small and cannot exceed the width of the holomorphy domain of $X^0(t)$. The same holds for the unstable whisker with the obvious changes (*i.e.* exchanging the roles of of p and q and considering $\mathrm{Re}\, t \leq 0$).

If:

$$X^s(t,\vec\alpha)=X^0(t,\vec\alpha)+\sum_{k=1}^{\infty}\mu^k X^{ks}(t,\vec\alpha)\equiv X^0(t,\vec\alpha)+\bar X^s(t,\vec\alpha) \tag{6.8}$$

denotes the evolution of the initial point in (6.5),(6.6), on the stable whisker, it follows from §5 that $X^s(t,\vec\alpha)$ has the form $X^s(\vec\omega t,t;\vec\alpha)$ for a suitable analytic function $X^s(\vec\psi,t;\vec\alpha)$ periodic in $\vec\psi,\vec\alpha$ converging at an exponential rate as $t\to\infty$ to $X^s(\vec\psi,\infty;\vec\alpha)$.

Analogously, if $X^{ks}(t,\vec\alpha_\mu)$ denotes the $k^{\rm th}$ Taylor coefficient in the μ expansion of $X^s(t,\vec\alpha_\mu)$, then $X^{ks}(t,\vec\alpha_\mu)=X^{ks}(\vec\omega t,t;\vec\alpha_\mu)$.

Note that X^{ks} depends only on the first $k+1$ coefficients of $\vec\alpha_\mu$; to stress this fact we shall sometimes write $X^{ks}(\vec\psi,t;\vec\alpha_\mu^{[\le k]})$. Note also that $X^s(\vec\psi,t,\vec\alpha)$ is holomorphic in a domain $|\,{\rm Im}\,\psi_j|<\bar\xi,|\mu|<\bar\mu_0$, ${\rm Re}\,t\ge T$ and ${\rm Im}\,t$ arbitrary if $|\mu|<\bar\mu_0$ with $\bar\mu_0,\bar\xi$ small enough and if T is large enough (for instance so that $(\bar p+\bar\kappa)e^{-gT/2}<\kappa$, see lemma 1').

Recursive expressions for $v^k,\vec h^k$ and the $X^{ks}(\vec\psi,t;\vec\alpha_\mu)$ could be deduced from §5, however it is more convenient to derive them directly.

To do this we put (6.8) into (6.4) and introduce the following notations.

If $G(I,\vec A,\varphi,\vec\alpha,\mu)$ is a function and if p, $\vec m=(m_1,\ldots,m_{2l})$ and k^i_j are integers, we denote:

$$\begin{aligned}&(G)^p_{\vec m}(\cdot)\equiv\Big(\frac{\partial^p_\mu\partial^{m_1}_I\partial^{m_2}_{A_1}\ldots\partial^{m_l}_{A_{l-1}}\partial^{m_{l+1}}_\varphi\partial^{m_{l+2}}_{\alpha_1}\ldots\partial^{m_{2l}}_{\alpha_{l-1}}G}{p!\,m_1!\,m_2!\,m_3!\,\ldots m_{l+2}!\,\ldots m_{2l}!}\Big)(\cdot)\\&(k^i_j)_{\vec m,p}\equiv(k^1_1,\ldots,k^1_{m_1},k^2_1,\ldots,k^2_{m_2},\ldots,k^{2l}_1,\ldots,k^{2l}_{m_{2l}})\qquad\text{s.t.}\ \sum k^i_j=p\end{aligned} \tag{6.9}$$

where $k^i_j\ge1$ if $m_i>0$. Then, given $\vec\alpha_\mu$, the (6.4) can be translated into a hierarchy of equations for the Taylor coefficients X^{ks} of $X(t,\vec\alpha_\mu)\equiv X(t)$; it becomes (omitting the stable index s):

$$\begin{aligned}\dot X^k_r=&\sum_j(\partial_jG_{0r})(X^0(t))X^k_j+\sum_{|\vec m|+p>1}(G_{0r})^p_{\vec m}(X^0(t))\sum_{(k^i_j)_{\vec m,k-p}}\prod_{i=1}^{2l}\prod_{j=1}^{m_i}X_i^{k^i_j}+\\&+\sum_{|\vec m|+p>0}(G_r)^p_{\vec m}(X^0(t))\sum_{(k^i_j)_{\vec m,k-p-1}}\prod_{i=1}^{2l}\prod_{j=1}^{m_i}X_i^{k^i_j}\end{aligned} \tag{6.10}$$

where $X^0(t)\equiv X^0(t,\vec\alpha_\mu)$ and the first term in the r.h.s. is separated from the others being the only one involving X^h with $h=k$. We write (6.10) as:

$$\dot X^k=LX^k+F^k \tag{6.11}$$

where $\partial_jG_{0r}\equiv L_{jr}$ and $F^k=F^{ks}$ is implicitly defined by (6.10) and (6.11), so that:

$$L\equiv L(t)=\begin{pmatrix}-\partial^2_{I\varphi}H_0&-\partial^2_{\vec A\varphi}H_0&-\partial^2_{\varphi\varphi}H_0&0\\0&0&0&0\\\partial^2_{II}H_0&\partial^2_{\vec AI}H_0&\partial^2_{I\varphi}H_0&0\\\partial^2_{\vec AI}H_0&\partial^2_{\vec A\vec A}H_0&\partial^2_{\vec A\varphi}H_0&0\end{pmatrix},\quad F^1(t)=\begin{pmatrix}-\partial_\varphi f-\partial^2_{\varphi\mu}H_0\\-\partial_{\vec\alpha}f\\\partial_If+\partial^2_{I\mu}H_0\\\partial_{\vec A}f+\partial_{\vec A\mu}H_0\end{pmatrix} \tag{6.12}$$

where the zeroes in the matrix L appear because we make use of the form (2.3) of H_0 and all the derivatives in (6.12) are evaluated at the point $X^0(t) = (I^0(t), \vec{A}^0, \varphi^0(t), \vec{\alpha}_0 + \vec{\omega}t + \vec{\vartheta}(t))$ and at $\mu = 0$ ($\varphi(0) = \bar{\varphi}$ and $\vec{\vartheta}$ is the phase shift introduced in §4). And more generally $F^k(t)$ are defined by solving recursively (6.11), using (6.10).

In fact we will check directly that (6.10) can be solved for every $k \geq 1$ and that, for each $\vec{\alpha}^k$, the initial data:

$$X^k(0) = \begin{pmatrix} v^k \\ \vec{h}^k \\ 0 \\ \vec{\alpha}^k \end{pmatrix} \tag{6.13}$$

can be fixed (*i.e.* v^k, $\vec{h}^k$ can be fixed) for each $k \geq 1$ so that $X^k(t)$ has the asymptotic properties dictated by lemma 1' (in particular that $X^k(t)$ is bounded).

The check can be done by studying with some care the wronskian of (6.11), *i.e.* the solution to the equation:

$$\dot{W} = LW, \qquad W(0) = 1 \tag{6.14}$$

In fact the solution to (6.11) can then be written:

$$X(t) = W(t)\Big(X(0) + \int_0^t W(\tau)^{-1} F(\tau) d\tau\Big) \tag{6.15}$$

where we drop the index k on X and F to simplify the notation while performing the k independent algebra that follows.

As mentioned above, we also suppose, for convenience, that the fixed $\bar{\varphi}$ is so chosen that the solution $i = i(\varphi, \vec{A}^0)$ of the separatrix equation $P(\vec{A}^0, i, \varphi, 0) = 0$ is maximal for $\varphi = \bar{\varphi}$, see (4.1).

The general properties of the wronskian matrix $W(t)$, solution of (6.14), can be found easily from lemma 0, §5. Let us write (cfr. (5.3), (5.2)):

$$\begin{pmatrix} I \\ \vec{A} \\ \varphi \\ \vec{\alpha} \end{pmatrix} = \begin{pmatrix} R(p, \vec{a}, q) \\ \vec{a} \\ S(p, \vec{a}, q) \\ \vec{\psi} + \vec{\delta}(p, \vec{a}, q) \end{pmatrix} \equiv V^0(p, q, \vec{\psi}, \vec{a}) \qquad H_0 \equiv h_0(\vec{a}, pq, 0) \equiv h(\vec{a}) + E(pq, \vec{a}) \tag{6.16}$$

the canonical map of lemma 0, for $\mu = 0$ (and dropped from the notation), reducing to normal form the free part of the hamiltonian: $E(pq, \vec{a}) = P(\vec{a}, I, \varphi)$ (if the free pendulum is an ordinary pendulum like (2.1), the (6.16) is the well known Jacobi map, and R, S are suitable Jacobian elliptic functions, see appendix 9; note that here we use a different order of the variables from that used in §5).

If we replace:

$$\begin{aligned} &p \to pe^{-gt}, \quad q \to qe^{+gt}, \quad \vec{\psi} \to \vec{\psi} + \vec{\Omega}t \\ &g \equiv \partial_J E(J, \vec{a}), \qquad J \equiv pq \\ &\vec{\Omega} \equiv \partial_{\vec{A}} h(\vec{a}) + \partial_{\vec{a}} E(J, \vec{a}) \end{aligned} \tag{6.17}$$

the map is still canonical (as (6.17) is the solution of the Hamilton equations in normal form for the free hamiltonian). Hence its jacobian is a canonical matrix that can be written:

$$U(t)=\begin{pmatrix}\partial_p R & \partial_{\vec a} R & \partial_q R & 0\\ 0 & 1 & 0 & 0\\ \partial_p S & \partial_{\vec a} S & \partial_q S & 0\\ \partial_p\vec\delta & \partial_{\vec a}\vec\delta & \partial_q\vec\delta & 1\end{pmatrix}\cdot\begin{pmatrix}e^{-gt} & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & e^{gt} & 0\\ 0 & t\,\partial_{\vec a}\vec\Omega & 0 & 1\end{pmatrix}\equiv U_0(t)\bar U(t) \tag{6.18}$$

where the derivatives of $R,S,\vec\delta,\vec\Omega$ are evaluated at $(pe^{-gt},qe^{gt},\vec a)$.

If $(\bar p,0)$ and $(0,\bar q)$ denote the points corresponding to $\bar\varphi$, we denote by $U^s_\cdot(t),U^u_\cdot(t)$ the above matrices evaluated, respectively, at the points $(\bar pe^{-gt},0)$, $(0,\bar qe^{gt})$; in both cases $\partial_{\vec a}\vec\Omega\equiv\bar H\equiv\partial^2_{\vec A}h$, $g=\partial_J E(0,\vec a)$. Note also that the entries involving the $\vec a-$ derivatives and all the $\partial\vec\delta$ vanish as $t\to\pm\infty$ (+ for the stable case and $-$ for the unstable one).

The representation (6.18) is symbolic as the 1's are in fact $(l-1)\times(l-1)$ identity matrices, the $\partial_{\vec a}R$, $\partial_{\vec a}S$ are row vectors (or $1\times(l-1)$ matrices), the $\partial\vec\delta$ are column vectors (or $(l-1)\times1$ matrices) while $\bar H,\partial\delta_{\vec a}$ are $(l-1)\times(l-1)$ matrices. Since, however, the notation is (after a moment of thought) self evident we shall use it also in the following without describing the obvious meaning of the matrix elements of $U(t)$ and of the corresponding ones of $U(t)^{-1}$.

The inverse of the matrix $U(t)$ is immediately computed (because $U(t)$ is canonical):

$$U(t)^{-1}=\bar U(-t)U_0(t)^{-1}=\bar U(-t)\begin{pmatrix}\partial_q S & \partial_q\vec\delta & -\partial_q R & 0\\ 0 & 1 & 0 & 0\\ -\partial_p S & -\partial_p\vec\delta & \partial_p R & 0\\ -\partial_{\vec a}S & -(\partial_{\vec a}\vec\delta)^T & \partial_{\vec a}R & 1\end{pmatrix} \tag{6.19}$$

The $U(t)$ is the jacobian of a family of solutions of the equations of motion, hence it verifies (6.14) except for the initial condition; so that:

$$W(t)\equiv U(t)U(0)^{-1} \tag{6.20}$$

We proceed to investigate the functions $X^k(t)$ with the aim of finding explicit conditions which can be used to determine the initial data $v^k,\vec h^k$, from the boundedness at $+\infty$ of the X^k and, more in general, to determine recursive equations for the $X^k(\vec\psi,t;\vec\alpha)$.

Let us write (5.10) as:

$$\begin{aligned}I&=R(p,\vec a,q)+V_+(p,q,\vec\psi,\mu) &\qquad \varphi&=S(p,\vec a,q)+V_-(p,q,\vec\psi,\mu)\\ \vec A&=\vec a+\vec V_\uparrow(p,q,\vec\psi,\mu) &\qquad \vec\alpha&=\vec\psi+\vec\delta(p,\vec a,q)+\vec V_\downarrow(p,q,\vec\psi,\mu)\end{aligned} \tag{6.21}$$

where $\vec a\in\Sigma_\mu$ (defined in 1) of lemma 1, §5) will be fixed throughout the the analysis and it will be dropped from the notations. As discussed in §5, the functions V are analytic in $p,q,\vec\psi,\mu$ for pq, $\mathrm{Im}\,p$, $\mathrm{Im}\,q$, $\mathrm{Im}\,\psi_j$ small, say $|pq|<\bar\kappa^2$, $|\mathrm{Im}\,p|,|\mathrm{Im}\,q|<\bar\kappa$, $|\mathrm{Im}\,\psi_j|<\bar\xi$

and $|\mu| < \mu_0$, for some $\bar\xi, \bar\kappa, \mu_0$ suitably chosen as functions of the hamiltonian parameters; furthermore the R, S are analytic in the same domain and the V are divisible by μ.

If we define $p_0(\vec\alpha), \vec\psi^s(\vec\alpha)$ as the solution of the equations:

$$\begin{aligned}\bar\varphi =& S(p_0,0) + V_-(p_0, 0, \vec\psi^s, \mu)\\ \vec\alpha =& \vec\psi^s + \vec\delta(p_0, 0) + \vec V_\downarrow(p_0, 0, \vec\psi^s, \mu)\end{aligned} \tag{6.22}$$

we can remark that the above equations can be solved at $\mu = 0$ with a non zero jacobian $\partial_p S(p, 0)$ (because at $(\bar p, 0)$ it is $\varphi = \bar\varphi$ and therefore $\dot\varphi$ is, by the above definition of $\bar\varphi$, maximal so that (since $q = 0$) $S_p(\bar p, 0) = -(g\bar p)^{-1}\dot\varphi \neq 0$ and $R_p(\bar p, 0) = 0$). The functions $p_0, \vec\psi^s$ will be analytic in $\vec\alpha$ for $|\,\mathrm{Im}\,\vec\alpha_j| < \bar\xi$, and $|\mu| < \mu_0$ imagining to redefine $\bar\xi, \mu_0$ so that they are the same here and in the analyticity domain of V, to avoid introducing too many parameters.

We shall also use the following notations:

$$\begin{aligned}&Z(p, q, \vec\psi, \mu) \equiv V^0(p, q, \vec\psi) + V(p, q, \vec\psi, \mu)\\ &X^s(\vec\psi, t; \vec\alpha, \mu) = Z(p_0(\vec\alpha)e^{-g(1+\gamma')t}, 0, \vec\psi + \vec\psi^s(\vec\alpha), \mu)\\ &X^u(\vec\psi, t; \vec\alpha, \mu) = Z(0, q_0(\vec\alpha)e^{g(1+\gamma')t}, \vec\psi + \vec\psi^u(\vec\alpha), \mu)\end{aligned} \tag{6.23}$$

where the functions $q_0, \vec\psi^u$ are defined as in (6.22) exchanging the roles of p and q.

Thus we can define:

$$\begin{aligned}&X^0(\vec\psi, t; \vec\alpha) = V^0(\bar p e^{-gt}, 0, \vec\psi + \vec\psi^{0s})\\ &\bar X^s(\vec\psi, t; \vec\alpha) = V(p_0 e^{-g(1+\gamma')t}, 0, \vec\psi + \vec\psi^s, \mu) - V^0(p_0 e^{-gt}, 0, \vec\psi + \vec\psi^s, \mu)\end{aligned} \tag{6.24}$$

γ' being the correction to the Lyapunov exponent in lemma 1' so that $X(t) = X(\vec\omega t, t; \vec\alpha) = X^0(\vec\omega t, t; \vec\alpha) + \bar X(\vec\omega t, t; \vec\alpha)$; $\vec\psi^{0s}$ is simply $\vec\alpha - \vec\delta(\bar p, 0, \vec a)$ while $\vec\psi^{0u}$ is $\vec\alpha - \vec\delta(0, \bar q, \vec a)$. Furthermore, all the functions $t \to X(\vec\psi + \vec\omega t, t; \vec\alpha)$ are orbits on the stable whiskers.

The analyticity in p implies that all the functions $X(\vec\psi, t; \vec\alpha)$ converge as $t \to \infty$ at exponential rate. The functions $X(\vec\psi, t; \vec\alpha)$ are analytic in $t, \vec\alpha, \vec\psi, \mu$ in a domain:

$$\bar{\mathcal D} = \{|\,\mathrm{Im}\,\alpha_j| < \bar\xi, |\,\mathrm{Im}\,\psi_j| < \bar\xi,\ |\,\mathrm{Im}\,t| < \bar\xi g^{-1}/2 \text{ or } |\,\mathrm{Re}\,t| > K g^{-1}, |\mu| < \mu_0\} \tag{6.25}$$

for a suitable K (so that the point $(p_0 + \bar\kappa)e^{-gt}$ is inside the analyticity domain for the X functions), again by lemma 1' and having once more redefined $\bar\kappa, \bar\xi$ to avoid introducing too many symbols.

If ρ', ρ, ξ', ξ are the analyticity parameters of the original hamiltonian, we shall use, to measure the size of the vectors $X = (X_+, \vec X_\uparrow, X_-, \vec X_\downarrow)$ the dimensionless norm:

$$|X| = {\rho'}^{-1}|X_+| + \rho^{-1}|\vec X_\uparrow| + {\xi'}^{-1}|X_-| + \xi^{-1}|\vec X_\downarrow| \tag{6.26}$$

and the above statements can be summarized by the first of:

$$||\bar X||_{\bar{\mathcal D}} \equiv \sup_{\bar{\mathcal D}} |\bar X| \le \bar v, \qquad |\gamma'| < \gamma_0 \tag{6.27}$$

where $\bar v, \gamma_0$ are suitable constants proportional to $|\mu|\mu_0^{-1}$ (see §5); the second bound provides a further property and also comes from lemma 1'; of course, all the above constants can be derived explicitly (if one wishes) from the dimensional bounds in the proof of lemma 1' (*i.e.* the above statement is "constructive"); the second inequality is a quantitative bound on the Lyapunov exponent, also part of lemma 1'.

The above remarks can be used to bound $\bar X^k_{\vec\nu}$, the Fourier transform with respect to the $\vec\psi$ variables of the $k^{\rm th}$– Taylor coefficient (in μ) of $\bar X$, defined in (6.24); in fact one immediately gets:

$$\begin{aligned} |\bar X^k_{\vec\nu}(t;\vec\alpha)| &< \bar v\mu_0^{-k} e^{-\bar\xi|\vec\nu|} \\ |\bar X^k_{\vec\nu}(t;\vec\alpha) - \bar X^k_{\vec\nu}(+\infty;\vec\alpha)| &< \bar v\mu_0^{-k} e^{-\bar\xi|\vec\nu|} e^{-g(1-\gamma_0)\,{\rm Re}\, t} \end{aligned} \qquad \text{in } \mathcal D \tag{6.28}$$

in a domain $\mathcal D$ defined as in (6.25) by omitting the condition on the $\vec\psi$ variables.

A further consequence of the above remarks is that if we define:

$$F(\vec\psi,t;\vec\alpha,\mu) = E\,[\mu\partial f(X^0+\bar X) + \partial H_0(X^0+\bar X) - \Big(\partial H_0(X^0) + \partial^2 H_0(X^0)\bar X\Big)] \tag{6.29}$$

where E is the standard $2l\times 2l$ matrix which in block form looks like $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $X^0, \bar X$ denote as above $V^0(\bar p e^{-gt}, 0, \vec\psi + \vec\psi^{0s})$, $\bar X(\vec\psi, t; \vec\alpha)$, then:

$$||F||_{\bar{\mathcal D}} < \bar f \tag{6.30}$$

for a suitable $\bar f$; again here we possibly redefine the constants $\mu_0, \bar\xi, \bar\kappa$. In the notations of (6.8) $\div$ (6.12), it is $F^k(t) = F(\vec\omega t, t; \vec\alpha, \mu)$.

The (6.30) implies the bounds:

$$\begin{aligned} |F^k_{\vec\nu}(t;\vec\alpha)| &< \bar f\mu_0^{-k} e^{-\bar\xi|\vec\nu|} \\ |F^k_{\vec\nu}(t;\vec\alpha) - F^k_{\vec\nu}(+\infty;\vec\alpha)| &< \bar f\mu_0^{-k} e^{-\bar\xi|\vec\nu|} e^{-g(1-\gamma_0)\,{\rm Re}\, t} \end{aligned} \tag{6.31}$$

in $\mathcal D$.

If $\vec\alpha$ is replaced by $\vec\alpha_\mu$, all the above estimates (6.31), (6.28) hold provided $|\,{\rm Im}\,\vec\alpha_\mu| < \bar\xi$ for $|\mu| < \mu_0$ and since F^k depends in such a case only on the first $k-1$ coefficients of $\vec\alpha_\mu$, the proper notation will be $F^k(\vec\psi, t; \vec\alpha_\mu^{[<k]})$.

Remark: in the case in which $P_0(I, \vec A, \varphi)$ is an ordinary pendulum hamiltonian the matrix elements of $U_0(t), W(t)$ can be computed essentially explicitly (see appendix 9): they have some remarkable analyticity and symmetry properties in t; namely the matrix elements of $U_0(t)$ are holomorphic in the domain $|\,\mathrm{Im}\,t| < (1-\varepsilon)\pi/(2g_0)$, for $\varepsilon > 0$, and are bounded there by $\bar u\varepsilon^{-2}$ for some $\bar u$, see ♣A9.8; furthermore the $W(t)$ matrix elements of the block $(+,\uparrow)\times(+,\uparrow)$ and those of the block $(-,\downarrow)\times(-,\downarrow)$ are even in t while those of the other two blocks are odd.

In general the "rows" of the matrix $U_0(t)^{-1}$, (6.19), will be denoted $\xi_+,\xi^0_\uparrow,\xi_-,\xi_\downarrow+\xi^0_\downarrow$, where $\xi^0_{\uparrow j}=(0,\vec e_j,0,0)$, $\xi^0_{\downarrow j}=(0,0,0,\vec e_j)$ with $\vec e_j$ being the unit $(l-1)$-vector with the $j-th$ component equal to 1. The splitting $\xi_\downarrow+\xi^0_\downarrow$ is performed so that all the matrices ξ_j have the last $l-1$ components zero.

With the above notations for $\bar X, F, U(t)^{-1}$ we deduce immediately an explicit expression for X^k in terms of F^k; it is $X^k(t)=U_0(t)Y^k(t)$ with:

$$Y^k(t)=\bar U(t)\left(U_0(0)^{-1}X^k(0)+\int_0^t\bar U(-\tau)U_0(\tau)^{-1}F^k(\tau)d\tau\right) \tag{6.32}$$

so that one finds:

$$\begin{aligned}
Y^k_+(t)=&e^{-gt}\Big(\xi_+(0)X^k(0)+\int_0^t e^{g\tau}\xi_+(\tau)F^k\,d\tau\Big)\\
Y^k_\uparrow(t)=&\vec h^k+\int_0^t\vec F^k_\uparrow\,d\tau\\
Y^k_-(t)=&e^{gt}\Big(\xi_-(0)X^k(0)+\int_0^t e^{-g\tau}\xi_-(\tau)F^k\,d\tau\Big)\\
Y^k_\downarrow(t)=&t\,\bar H\Big(\vec h^k+\int_0^t\vec F^k_\uparrow\,d\tau\Big)+\xi_\downarrow(0)X^k(0)+X^k_\downarrow(0)+\\
&+\int_0^t[-\tau\bar H F^k_\uparrow+\xi_\downarrow(\tau)F^k+F^k_\downarrow]\,d\tau
\end{aligned} \tag{6.33}$$

where the arguments of F^k are $(\vec\omega\tau,\tau;\vec\alpha^{[<k]}_\mu)$; when needed we regard m vectors as $1\times m$ or $m\times1$ matrices and use the standard rules for matrix multiplication (*e.g.* in the first of (6.33) ξ_+ is a $(1\times2l)$ matrix while X^k is a $(2l\times1)$ matrix and their product is a scalar); in the following formulae we shall drop the explicit dependence on $\vec\alpha$ as such dependence plays here no role. Hence, since the boundedness as $t\to+\infty$ of $X^k(t)$ is equivalent to the boundedness of $Y^k(t)$, a straightforward asymptotic analysis based on the bounds (6.28),(6.31) shows that the latter corresponds to the following conditions on the initial data $X^k(0)$ for $k\ge1$:

$$\begin{aligned}
&\langle\vec F^k_\uparrow(\cdot,\infty)\rangle=0\\
&\xi_-(0)X^k(0)+\int_0^\infty e^{-g\tau}\xi_-(\tau)F^k(\vec\omega\tau,\tau)\,d\tau=0\\
&\langle\vec F^k_\downarrow(\cdot,\infty)\rangle+\bar H\Big(\vec h^k+\int_0^\infty[\vec F^k_\uparrow(\vec\omega\tau,\tau)-\vec F_\uparrow(\vec\omega\tau,\infty)]\,d\tau-(\vec\omega\cdot\vec\partial)^{-1}\vec F^k_\uparrow(\vec 0,\infty)\Big)=\vec 0
\end{aligned} \tag{6.34}$$

where to derive the third equality we have used the first identity and the fact that $\xi_\downarrow(\infty) = 0$, see the comment following (6.18); $\langle\cdot\rangle$ denotes average over $\vec{\psi}$, $(\vec{\omega}\cdot\partial)^{-1}$ acts by dividing the $\vec{\nu}$–Fourier coefficient (with respect to $\vec{\psi}$) by $i\vec{\omega}\cdot\vec{\nu}$: notice that it is possible to apply $(\vec{\omega}\cdot\partial)^{-1}$ to $\vec{F}_\uparrow(\cdot,\infty)$ because of the first identity. The first condition must be an identity, as it does not involve the initial conditions which we suppose to have already determined for $X^h(0)$, $h = 1,\ldots,k-1$, (otherwise lemma 1' could not possibly hold). If one feels uneasy with such a boldly indirect proof one can check the statement directly (see appendix 12).

The third condition fixes $\vec{h}^k(0)$, while $\vec{X}^k_\downarrow(0)$ has to be $\vec{\alpha}^k$, by definition, and finally the second condition fixes $v^k = X^k_+(0)$, because $X^k_-(0) \equiv 0$ (by definition) and :

$$\begin{aligned} \xi_-(0)X^k(0) = -S_p(\bar{p},0)X^k_+ - (\partial_{\vec{a}}\delta)^T\vec{h}^k \equiv \\ \equiv (g\bar{p})^{-1}\big(X^k_+(0)\partial_I P + X^k_\uparrow(0)\cdot\partial_{\vec{A}}P\big) \end{aligned} \tag{6.35}$$

and more generally:

$$e^{-gt}\xi_-(t) \equiv (g\bar{p})^{-1}w(t), \quad w(t) \equiv (\partial_I P, \partial_{\vec{A}}P, \partial_\varphi P, 0) \tag{6.36}$$

where the derivatives are evaluated at $X^0(t)$. Notice that $w(t)$ in (6.36) is such that $Ew = \dot{X}^0(t)$, if E denotes the standard symplectic matrix. Furthermore:

$$e^{-gt}\xi_-(t) = (g\bar{p})^{-1}w(t) \equiv e^{gt}\xi_+(t)\ , \qquad \text{for all} \quad -\infty < t < \infty \tag{6.37}$$

and $w(t)$ is holomorphic in t, for all real values of t: the above statements reflect the degeneracy of the unperturbed whiskers.

It is also useful to rewrite the vector $Y^k(t)$ after the above (6.34) are taken into account, leading to a few cancellations which make clear the asymptotic boundedness as $t\to\infty$ of Y^k (imposed, indeed, by (6.34)). We find:

$$\begin{aligned} Y^k_+(t) =& e^{-gt}\rho^k + e^{-gt}\int_0^t e^{g\tau}\xi_+(\tau)F^k(\vec{\omega}\tau,\tau)\,d\tau \\ \vec{Y}^k_\uparrow(t) =& \vec{h}^k + \int_0^t \vec{F}^k_\uparrow(\vec{\omega}\tau,\tau)\,d\tau \\ Y^k_-(t) =& -e^{gt}\int_t^\infty e^{-g\tau}\xi_-(\tau)F^k(\vec{\omega}\tau,\tau)\,d\tau \\ \vec{Y}^k_\downarrow(t) =& \vec{\chi}^k - \int_t^\infty \Big[\xi_\downarrow F^k(\vec{\omega}\tau,\tau) + [\vec{F}^k_\downarrow(\vec{\omega}\tau,\tau) - \vec{F}^k_\downarrow(\vec{\omega}\tau,\infty)]\Big]\,d\tau + \\ & + (\vec{\omega}\cdot\vec{\partial})^{-1}\mathcal{P}\vec{F}^k_\downarrow(\vec{\omega}t,\infty) - \bar{H}\int_t^\infty (t-\tau)[\vec{F}^k_\uparrow(\vec{\omega}t,\tau) - \vec{F}^k_\uparrow(\vec{\omega}t,\infty)]\,d\tau + \\ & + \bar{H}(\vec{\omega}\cdot\partial)^{-2}\vec{F}^k_\uparrow(\vec{\omega}t,\infty) \end{aligned} \tag{6.38}$$

where $(\vec{\omega}\cdot\vec{\partial})$, $\mathcal{P}$, $\langle\cdot\rangle$ are operators acting on the $\vec{\psi}$ dependence of the F^k- functions: they multiply the Fourier transforms of F^k by $i\vec{\omega}\cdot\vec{\nu}$, $\delta_{|\vec{\nu}|\neq 0}$ and $\delta_{\vec{\nu}=\vec{0}}$ respectively (hence $\langle\cdot\rangle$ is the average over $\vec{\psi}$), $\delta_\cdot$ being the Kronecker δ. The constants $\rho^k, \vec{\chi}^k$ are:

$$
\begin{aligned}
\rho^k =&(\partial_q S)_0 v^k + (\partial_q\vec{\delta})_0\cdot\vec{h}^k\\
\vec{\chi}^k =&\vec{\alpha}_k - (\partial_{\vec{a}}S)_0 v^k - (\partial_{\vec{a}}\vec{\delta})_0^T\vec{h}^k + \int_0^\infty \xi_\downarrow(\tau)F^k(\vec{\omega}\tau,\tau)\,d\tau +\\
&+\int_0^\infty [\vec{F}_\downarrow^k(\vec{\omega}\tau,\tau) - \vec{F}_\downarrow^k(\vec{\omega}\tau,\infty)]\,d\tau - (\vec{\omega}\cdot\vec{\partial})^{-1}\mathcal{P}\vec{F}_\downarrow^k(\vec{0},\infty)\\
&-(\vec{\omega}\cdot\vec{\partial})^{-2}\bar{H}\vec{F}_\uparrow^k(\vec{0},\infty) - \bar{H}\int_0^\infty \tau[\vec{F}_\uparrow^k(\vec{\omega}\tau,\tau) - \vec{F}_\uparrow^k(\vec{\omega}\tau,\infty)]\Big]\,d\tau
\end{aligned}
\tag{6.39}
$$

where $(\cdot)_0$ means "at $t=0$"; in the cases in which there is no coupling between the pendulum and the rotators in the free system (see §4) the ρ^k and also the coefficients of v^k and $\vec{h}^k$ in $\vec{\chi}^k$ vanish.

Note also that, in the case $\vec{\alpha}=\vec{\alpha}_\mu$, the dependence on the coefficients of $\vec{\alpha}_\mu$ (as it follows by inspection from (6.15),(6.20),(6.29)) is:

$$
\begin{aligned}
X^k(\vec{\psi},t;\vec{\alpha}_\mu^{[\le k]}) =&\begin{pmatrix}0\\0\\0\\\vec{\alpha}^k\end{pmatrix} + X^k(\vec{\psi},t;\vec{\alpha}^{[<k]})\\
F^k(\vec{\psi},t;\vec{\alpha}_\mu^{[\le k-1]}) =&E\partial^2 f(X^0)\begin{pmatrix}0\\0\\0\\\vec{\alpha}^{k-1}\end{pmatrix} + F^k(\vec{\psi},t;\vec{\alpha}^{[<k-1]})
\end{aligned}
\tag{6.40}
$$

The theory of the unstable whisker is identical, with the obvious changes which essentially consist in taking $\mathrm{Re}\,t$ negative, replacing ∞ with $-\infty$ and exchanging the role of p and q and of the components $+,-$ of Y^k. Note that, in general, the functions F^k, X^k for the unstable whisker *are not* the analytic continuation of those of the stable whisker with the same initial conditions (which is a fact that would mean that all stable whisker orbits are also orbits on the unstable whisker, so that the homoclinic points would be completely degenerate).

As remarked above, see (6.37), $W(t)$ (but not $U(t)$) is analytic for all t because of the degeneracy of the free system. Hence in what follows we should append a superscript s or u (or equivalently $+$ or $-$) to F, X, Y, U, U_0 depending on whether we consider them for a stable or an unstable whisker. When, as usual, we do not follow this convention *we shall assume that the functions being considered are associated with the stable whisker if their time argument has* $\mathrm{Re}\,t>0$ *and with the unstable if* $\mathrm{Re}\,t<0$. Of course we label the initial conditions $X^\sigma(0)$ with $\sigma=s$ or $\sigma=u$, depending on the whisker to which they refer.

We now look at the *homoclinic conditions.* The homoclinic points that we study will be *a priori* supposed to have the form:

$$\varphi = \bar{\varphi}, \qquad \alpha_\mu = \vec{\alpha}^0 + \mu\,\vec{\alpha}^1 + \mu^2\vec{\alpha}^2 + \dots \tag{6.41}$$

with the series in (6.41) being at least asymptotic as $\mu \to 0$. Such points correspond to initial conditions $X^{k\sigma}(0) = (v^{k\sigma}, \vec{h}^{k\sigma}, 0, \vec{\alpha}^k)$.

Remarking that the definition of $X(\vec{\psi}, t; \vec{\alpha}, \mu)$, (6.23), implies that $X^s(\vec{\psi}, \infty; \vec{\alpha}, \mu) = X^u(\vec{\psi} + \vec{\sigma}(\vec{\alpha}, \mu), -\infty; \vec{\alpha}, \mu)$ with $\vec{\sigma} = \vec{\psi}^s(\vec{\alpha}) - \vec{\psi}^u(\vec{\alpha})$ and also $F^s(\vec{\psi}, \infty; \vec{\alpha}, \mu) = F^u(\vec{\psi} + \vec{\sigma}, -\infty; \vec{\alpha}, \mu)$ so that for each k:

$$\langle X^{ks}(\cdot, +\infty; \vec{\alpha}_\mu)\rangle = \langle X^{ku}(\cdot, -\infty; \vec{\alpha}_\mu)\rangle\ , \qquad \langle F^{ks}(\cdot, +\infty; \vec{\alpha}_\mu)\rangle = \langle F^{ku}(\cdot, -\infty; \vec{\alpha}_\mu)\rangle \tag{6.42}$$

we see that, in general, the homoclinic conditions $\vec{h}^{ks} = \vec{h}^{ku}$ become, (if $\sigma_t = \mathrm{sign}\,(t)$):

$$-\int_{-\infty}^{\infty} \left[\vec{F}^k_\uparrow(\vec{\omega}\tau, \tau; \vec{\alpha}^{[<k]}_\mu) - \vec{F}^k_\uparrow(\vec{\omega}\tau, \sigma_\tau\infty; \vec{\alpha}^{[<k]}_\mu)\right] d\tau + \sum_{\substack{\vec{\nu}\neq 0\\ \sigma=\pm}} \sigma \frac{\vec{F}^k_{\uparrow\vec{\nu}}(\sigma\infty; \vec{\alpha}^{[<k]}_\mu)}{i\vec{\omega}\cdot\vec{\nu}} = 0 \tag{6.43}$$

which we call "the conditions associated with the $\vec{A}$ variables". Note that (6.43) is immediately derived from (6.34) in the anisochronous case (*i.e.* in the case $\bar{H}$ is invertible); however it is possible to deduce it directly also in the important case of partially isochronous systems (*i.e.* periodically time–dependent systems) associated to hamiltonians of the form:

$$\omega B + h(\tilde{A}) + P(I, \tilde{A}, \varphi) + \mu f(I, \tilde{A}, \varphi, \lambda, \tilde{\alpha})\ , \qquad (B, \tilde{A}) \equiv \vec{A}\ , (\lambda, \tilde{\alpha}) \equiv \vec{\alpha} \tag{6.44}$$

with $\det \partial^2_{\tilde{A}} h \neq 0$. To check (6.43) in such a case, consider the second of (6.38) observing that $\vec{Y}^k_\uparrow \equiv \vec{X}^k_\uparrow$; add and subtract to the l.h.s. $\vec{X}^k_\uparrow(\vec{\omega}t, \sigma_t\infty)$ and to the r.h.s. $\int_0^t \vec{F}^k_\uparrow(\vec{\omega}\tau, \sigma_\tau\infty) d\tau$; take the quasi–periodic average $\lim_{T\to\sigma\infty}(1/T)\int_0^T \cdot$ using (6.28), (6.31) and the first of (6.34); finally use (6.42). Hence we conclude that a *unified treatment of anisochronous systems and of forced systems is possible.*

To derive the other homoclinic condition, $v^{ks} = v^{ku}$, we use the (6.35), (6.36) to find an expression of $v^{ks} - v^{ku}$ which is symmetric for the two whiskers contributions, and, by the second of (6.34) we find:

$$\left[w_\uparrow(0)(\vec{h}^{ks} - \vec{h}^{ku}) + \int_{-\infty}^{\infty} w(\tau) F^k(\vec{\omega}\tau, \tau; \vec{\alpha}^{[<k]}_\mu) d\tau\right] = 0 \tag{6.45}$$

($w_\uparrow \equiv \partial_{\vec{A}} P$ and notice that $w \to 0$ exponentially fast as $t \to \pm\infty$). We call the above equation "the homoclinic condition corresponding to the I variable".

Hence (6.43),(6.45) is the complete set of homoclinic conditions. It is clear that the the l equations in (6.43),(6.45) cannot be independent: and in fact one can check that either the (6.43) or the (6.45) together with any $l-2$ of the (6.43) imply the remaining one: this expresses the fact that the energy of the two whiskers is the same (because they are asymptotic to the same torus).

The (6.43) is not yet very explicit, as a recursive equation for $\vec{\alpha}^k$. But it is easy to make it clearer.

We begin by considering it for $k=1$. The function F of order 1, $F^1(\vec{\psi},t;\vec{\alpha}^{[<1]})$, is in this case:

$$F^1(\vec{\psi},t;\vec{\alpha}^0)=\begin{pmatrix}-\partial_\varphi f(I^0(t),\vec{a},\varphi^0(t),\vec{\psi}+\vec{\alpha}^0+\vec{\vartheta}(t),0)\\ -\partial_{\vec{\alpha}} f(I^0(t),\vec{a},\varphi^0(t),\vec{\psi}+\vec{\alpha}^0+\vec{\vartheta}(t),0)\\ \partial_I f(I^0(t),\vec{a},\varphi^0(t),\vec{\psi}+\vec{\alpha}^0+\vec{\vartheta}(t),0)\\ \partial_{\vec{A}} f(I^0(t),\vec{a},\varphi^0(t),\vec{\psi}+\vec{\alpha}^0+\vec{\vartheta}(t),0)\end{pmatrix} \tag{6.46}$$

where $\vec{\vartheta}(t)$ is the phase shift introduced in §4. And the homoclinic conditions for the $\vec{A}$ variables becomes (cfr. (4.6)÷(4.8))

$$\begin{aligned}&\int_{-\infty}^{\infty}[\partial_{\vec{\alpha}} f|_t-\partial_{\vec{\alpha}} f|_{\sigma_t\infty}]\,d\tau-\sum_{\substack{\vec{\nu}\neq\vec{0}\\ \sigma=\pm}}\frac{\sigma}{i\vec{\omega}\cdot\vec{\nu}}\partial_{\vec{\alpha}} f_{\vec{\nu}}(0,\vec{a},0,0)e^{i(\vec{\omega}\cdot\vec{\alpha}^0+\vec{\nu}\cdot\vartheta(\sigma\infty))}=\\ &=\int_{-\infty}^{\infty}\sum_{\vec{\nu}\neq\vec{0}}i\vec{\nu}e^{i(\vec{\omega}\cdot\vec{\nu}\tau+\vec{\nu}\cdot\vec{\alpha}^0)}\partial_\tau\Big(e^{i\vec{\nu}\cdot\vec{\vartheta}(\tau)}f_{\vec{\nu}}(I^0(\tau),\vec{a},\varphi^0(\tau),0)\Big)\,d\tau=\vec{0}\end{aligned} \tag{6.47}$$

where $f|_t\equiv f$ evaluated at $(I^0(t),\vec{a},\varphi^0(t),\vec{\alpha}^0+\vec{\omega}t+\vec{\vartheta}(t),0)$ while $f|_{\sigma_t\infty}$ denotes f evaluated at $(0,\vec{a},0,\vec{\alpha}^0+\vec{\omega}t+\vec{\vartheta}(\sigma_t\infty),0)$ if $\sigma_t=\operatorname{sign} t$. And we see that the (6.47) always has at least two solutions, namely the critical points of the periodic function:

$$M_f(\vec{\alpha})=\int_{-\infty}^{\infty}\sum_{\vec{\nu}\neq\vec{0}}e^{i(\vec{\omega}\cdot\vec{\nu}\tau+\vec{\nu}\cdot\vec{\alpha})}\partial_\tau\big(e^{i\vec{\nu}\cdot\vec{\vartheta}(\tau)}f_{\vec{\nu}}(I^0(\tau),\vec{a},\varphi^0(\tau),0)\big)\,d\tau \tag{6.48}$$

The equality of the whiskers energies then implies that the homoclinic equation relative to I is also satisfied.

Therefore the equations for $k=1$ determine the zeroth order approximation $\vec{\alpha}^0$ to the homoclinic point. Let M_0 be the jacobian matrix (with respect to $\vec{\alpha}$) of the $l-1$ functions in (6.43) evaluated at a solution point $\vec{\alpha}^0$ (equivalently: $M_0\equiv\partial^2_{\vec{\alpha}}M_f(\vec{\alpha}_0)$) and *assume that M_0 is non degenerate*, see (4.10).

To determine the higher orders $\vec{\alpha}_k$, $k\geq 1$, we remark that (6.40) shows that the dependence of F^k on $\vec{\alpha}_{k-1}$ is rather simple. If we call $\vec{D}_k$ the l.h.s. of (6.43) evaluated by replacing $\vec{\alpha}_\mu^{[<k]}$ with $\vec{\alpha}_\mu^{[<k-1]}$, we see that, by (6.40), the (6.43) take the form:

$$M_0\vec{\alpha}_{k-1}=\vec{D}_k \tag{6.49}$$

Therefore the non degeneracy of M_0, the hessian of (6.48), at the solution point $\vec{\alpha}^0$ is all one needs to perform the perturbation theory to arbitrary order.

But the first order calculation is usually sufficient to prove the existence of the homoclinic point. Its exact equation is in fact:

$$\vec{H}(\vec{\alpha}) \equiv \sum_{k=0}^{\infty} \mu^{k-1}(\vec{H}^{k+}(\vec{\alpha}) - \vec{H}^{k-}(\vec{\alpha})) = 0 \tag{6.50}$$

where $\vec{H}^{k\sigma}(\vec{\alpha})$ is just the l.h.s. of (6.43).

We see that the equation (6.50) admits, for $\vec{a}$ on a diffusion path open for diffusion (in the sense of §4, see (4.10), a non degenerate solution for $\mu = 0$, by our assumption on $\vec{\alpha}^0$: furthermore the function in the l.h.s. of (6.50) is analytic in μ for $|\mu| < \mu_0$ by lemma 1': therefore for all the invariant tori associated with the diffusion path, by lemma 1', the equation has an analytic solution $\vec{\alpha}_\mu$ whose Taylor expansion coefficients have already been determined in (6.49).

We also need some informations about the non degeneracy of the intersection between the two whiskers at $\vec{\alpha}_\mu$. They are provided by the Taylor coefficients of the expansion in $\vec{\alpha} - \vec{\alpha}_\mu$ of the functions $\vec{H}(\vec{\alpha})$ around $\vec{\alpha}_\mu$. In particular the first order coefficients will define a $(l-1) \times (l-1)$ matrix that we call *the homoclinic intersection tensor* or the *homoclinic angles matrix*: it is clear that such matrix is just M_0: hence we see the physical meaning of the matrix M_0. The name is slightly improper as the matrix M_0 has the dimension of an action and (the trigonometric tangent of) the physical angles between the tangent vectors to the two whiskers will not be the eigenvalues of M_0 but proportional to them via a constant bearing the dimension of an action.

We conclude that, *at a homoclinic point, there is a uniform lower bound to the angles between pairs of vectors tangent to the whiskers associated with the invariant tori of a diffusion path open for diffusion.*

§7 Whisker ladders and rounds density

Given a diffusion path $\mathcal{L}$, we see from lemma 1, §5, that we can expect that there are gaps on it, in which whiskered tori are missing (*i.e.* ladder segments with no *rounds*), which are quite large and, in fact they can be as large as $K\mu^{1/c}$; see (5.4) and the comment after (5.90): so that $c = 7t$.

However the rule is that *the gaps are much narrower if the path $\mathcal{L}$ is not too badly placed in the action space*: this depends of course on the perturbation f. The situation is clearest if f is a trigonometric polynomial. We give the following, general, definition:

Definition: Given a free hamiltonian h as in (2.3), a region $V \subset R^{l-1}$ of the action variables $\vec{A}$ is free of resonances to order N if it is:

$$\vec{\omega}(\vec{A}) \cdot \vec{\nu} \neq 0 \qquad \forall |\vec{\nu}| \leq N \tag{7.1}$$

having set, as usual, $\vec\omega(\vec A) \equiv \partial_{\vec A} h(\vec A, 0, 0)$.

If f *is a trigonometric polynomial of degree* N_f*:*

$$f(I, \vec A, \varphi, \vec\alpha, \mu) = \sum_{|\vec\nu| \le N_f} f_{\vec\nu}(I, \vec A, \varphi, \mu) e^{i\vec\nu\cdot\vec\alpha} \tag{7.2}$$

we say that h *and* f *do not resonate to degree* p *in the region* V *in action space if the region* V *is free of resonances to order* $N = pN_f$.

Recalling (3.1),(3.2),(5.4), (5.82) we can prove the following lemma:

Lemma 3: *Let* f *be a trigonometric polynomial which does not resonate to degree* p *with the free hamiltonian* h *in the phase space domain* V*. And let* $\mathcal{L}$ *be a diffusion path contained in* V*. Then the set* Σ_μ *in (5.4) can be taken to verify:*

$$(s_2 - s_1)^{-1} \int_{\Sigma_\mu} ds \ge (1 - K|\mu|^{(p+1)/c}) \tag{7.3}$$

and all the remaining statements of lemma 1 and 1' stay unchanged.

So in particular, if $p > 2c - 1$ any interval of length $O(\mu^2)$ on $\mathcal{L}$ will necessarily contain rounds of the whisker ladder, if μ is small enough. The analysis in §5 allows us to say that c can be taken, for instance, $c = 7(l-1)$ hence for $p = 14(l-1)$ we have gaps of relative size of $O(\mu^2)$.

Proof: The assumptions are just what one needs to perform with no troubles perturbation theory to order p. We shall take for simplicity f, h to be μ independent.

Let Φ be a generating function for a canonical map: it will depend on variables $(\vec A', \vec\alpha_0, p', q_0) \in V \times T^{l-1} \times S^2_{\bar\kappa_0}$, assuming to have already done the first change of coordinates considered in the proof of lemma 1, §5, to change I, φ into the more natural (but local) coordinates p_0, q_0 of the pendulum (and $S_\kappa = \{p : \ |p| < \kappa\}$), see lemma 0, §5.

We take Φ to be a polynomial of order p in μ:

$$\Phi = \mu\Phi^{(1)} + \mu^2\Phi^{(2)} + \ldots + \mu^p\Phi^{(p)} \tag{7.4}$$

which we impose to be a solution to order $O(\mu^{p+1})$ of the Hamilton-Jacobi equation:

$$\begin{aligned} &h_0(\vec A' + \partial_{\vec\alpha_0}\Phi, (p' + \partial_q\Phi)q_0) + \mu f(\vec A' + \partial_{\vec\alpha_0}\Phi, \vec\alpha_0, p' + \partial_q\Phi, q_0) = \\ &\qquad = \tilde h(\vec A', p'(q_0 + \partial_{p'}\Phi), \mu) + O(\mu^{p+1}) \\ &\tilde h(\vec A', J) = h_0(\vec A', J) + \mu h^{(1)}(\vec A', J) + \ldots + \mu^p h^{(0)}(\vec A', J) \end{aligned} \tag{7.5}$$

where $\Phi, h^{(1)}, \ldots, h^{(p)}$ are regarded as unknown.

The equations for $\Phi^{(1)}, \ldots, \Phi^{(p)}, h^{(1)}, \ldots, h^{(p)}$ generated by (7.5) can be recursively solved if the non resonance property (7.1) holds for $N = pN_f$ in a domain around the diffusion curve. The domain of definition of the solution can be given the form $V \times T^{l-1} \times S^2_{\bar\kappa}$, with $\tilde{\kappa}$ small enough. In fact, assuming to have solved the equations for $i = 1, \ldots, k$ and to have determined $\Phi^{(i)}, h^{(i)}$, $i \le k$, as analytic functions on $V \times T^{l-1} \times S^2_{\tilde{\kappa}}$, with $\Phi^{(i)}$ being a trigonometric polynomial of degree iN_f, we see that (7.5) says:

$$\begin{aligned}&\vec{\omega}_0(\vec{A}', p'q_0) \cdot \partial_{\vec{\alpha}} \Phi^{(k+1)} + g_0(\vec{A}', p'q_0)\big(q_0 \partial_{q_0} \Phi^{(k+1)} - p' \partial_{p'} \Phi^{(k+1)}\big) + \\ &+ P^{(k+1)}(\vec{A}', \vec{\alpha}_0, p', q_0) = h^{(k+1)}(\vec{A}', p'q_0)\end{aligned} \tag{7.6}$$

where $\vec{\omega}_0(\vec{A}, J) \equiv \partial_{\vec{A}'} h_0(\vec{A}', J)$, $g_0(\vec{A}', J) \equiv \partial_J h_0(\vec{A}', J)$ and $P^{(k+1)}$ is a polynomial in the derivatives of f of order $\le k$ and in $\partial\Phi^{(1)}, \ldots, \partial\Phi^{(k)}$, $\partial h^{(1)}, \ldots, \partial h^{(k)}$ and its monomials:

$$(\partial\Phi^{(1)})^{n_1} \ldots (\partial\Phi^{(k)})^{n_k} (\partial h^{(1)})^{m_1} \ldots (\partial h^{(k)})^{m_k} \tag{7.7}$$

must be such that their order verifies:

$$\sum_{j=1}^{k} j(n_j + m_j) \le k + 1 \tag{7.8}$$

Hence we see that $P^{(k+1)}$ is a trigonometric polynomial of degree $N_{k+1} \le (k+1)N_f$, and (7.7) can be solved, using the notations introduced in (5.27), by:

$$\Phi^{(k+1)}_{\vec{\nu},hh'} = P^{(k+1)}_{\vec{\nu},hh'}(\vec{A}')\big[-i\vec{\omega}(\vec{A}', p'q_0) \cdot \vec{\nu} - (h' - h) g_0(\vec{A}', p'q_0)\big]^{-1} \tag{7.9}$$

for $|h' - h| + |\vec{\nu}| \ne 0$, $|\vec{\nu}| \le (k+1)N_f$, and:

$$h^{(k+1)}(\vec{A}', p'q_0) = \sum_{h=0}^{\infty} P^{(k+1)}_{\vec{0},hh}(\vec{A}')(p'q_0)^h \tag{7.10}$$

The induction construction will work until $k \le p$, as it is clear that $\Phi^{(k+1)}, h^{(k+1)}$ have the same analyticity domains as $\Phi^{(1)}, \ldots, \Phi^{(k)}, h^{(1)}, \ldots, h^{(k)}$.

The canonical map generated by Φ will have a somewhat smaller domain $V' \times T^{l-1} \times S^2_{\kappa'}$, with V' differing from V and κ' differing from $\tilde{\kappa}$ only by an amount of order $O(\mu)$ (which means that the boundaries of V and V' are close within a distance $O(\mu)$).

This completes the proof of lemma 3 because, in the new coordinates, the hamiltonian takes the form:

$$h'(\vec{A}', p'q', \mu) + \mu^{p+1} f'(\vec{A}', \vec{\alpha}', p', q', \mu) \tag{7.11}$$

where the difference $h' - h$ is divisible by μ and h', f' are analytic in $V' \times T^{l-1} \times S^2_{\kappa'}$.

Under the above circumstances lemma 1 applies: but this time μ is replaced by μ^{p+1}, and therefore $1/c$, in (5.4), is replaced by $(p+1)/c$.

Lemma 3 above has some rather obvious extensions to the cases in which f is not a trigonometric polynomial. In such cases, if $\bar{N}$ is the maximum value such that no resonance of h occurs for $|\vec{\nu}| \le \bar{N}$ we can write $f = f^{[\le \bar{N}/c]} + f^{[> \bar{N}/c]}$ and if $\mu||\bar{f}_0||e^{-\xi \bar{N}/c} \ll \mu_*$ and, at the same time, $|\mu| \ll \mu_*$ of lemma 1, see (5.90). Hence we can hope to reach the same conclusion of lemma 3 in various particular cases. But we see that, unless f is a trigonometric polynomial, the improvement of the density estimate of the whisker is a delicate matter. Since the ideas are quite clear, we refrain from formulating precise results on the cases when f is not a trigonometric polynomial.

§8 Heteroclinic intersections. Drift and diffusion along directly open paths

We now want to study intersections between stable whiskers of tori corresponding to some value of s, and unstable whiskers of tori corresponding to a different s value: such intersections are called *heteroclinic* intersections. Then, by means of *chains* of heteroclinic intersections (cfr. [A]; see also [D]) corresponding to tori related to directly open diffusion paths, we will show how one can construct orbits *shadowing* all the heteroclinic orbits of the chain. In particular, we will estimate the time needed by the diffusion orbit to drift from one end of the chain to the other end.

We consider a line of tori (and corresponding whiskers) with a given energy E associated with a diffusion line $\mathcal{L}$ *directly open for diffusion*; see definition 2 of §4. Given $\mathcal{L}$ the existence of such a line of tori is consequence of lemma 1 of §5. We consider also the associated diffusion sheet $\tilde{\mathcal{L}}$ (see (5.8)) and the relative whiskered tori constructed by lemma 1'.

The equations for the whiskers associated to $\tilde{\mathcal{L}}$, deduced in §6, take the form, at the point corresponding to the value s, u of the diffusion sheet parameters:

$$\vec{A}^{\pm}(\varphi, \vec{\alpha}; s, u) = \vec{A}_{su} + \mu\{\partial_{\vec{\alpha}} M_f^{\pm}(\varphi, \vec{\alpha}; s, u) + \vec{\xi}_{su}^{\infty}\} + O(\mu^2) \tag{8.1}$$

where the function $M_f^{\pm}$ is defined in terms of the Melnikov function F, see (4.6), (6.48) and (see (6.42)):

$$\vec{\xi}_{su}^{\infty} = \langle X^{1+}(\cdot, +\infty; \vec{\alpha})\rangle = \langle X^{1-}(\cdot, -\infty; \vec{\alpha})\rangle \tag{8.2}$$

In this section "$+/-$" will mean "stable/unstable" as we reserve the labels u, s to have the meaning they have in §5: so that s is a parameter describing points on $\mathcal{L}$ and u describes the variation of the quasi periodic motions frequencies.

The results of §5 show that we can imagine that the functions $X, \vec{\xi}$ are defined for *all* $(s, u) \in [s_1, s_2] \times [-\bar{u}, \bar{u}]$ and for $|\mu| \le \mu^*$, $\vec{\alpha} \in T^{l-1}$ and $|\varphi| < \tilde{\varphi}$ and there they are of class C^p where p is any prefixed integer. Below we imagine to have fixed such an extension with $p = 2$. Of course this is just a convenient way of expressing the regularity properties of functions that are defined on sets with a lot of holes: the values of the functions in points with $s \notin \Sigma_\mu$ (cfr. 1) of lemma 1, §5) are not interesting and all that is being said is that the interesting values of our functions can be smoothly interpolated. At a point with a prefixed φ coordinate, say $\varphi = \bar{\varphi}$ (as in §4 and elsewhere) the condition for heteroclinic intersection between $W_{stable}(s)$ and $W_{unstable}(s')$, in the prefixed energy surface, becomes to first order in μ:

$$\begin{aligned} &\mu\left\{\partial_{\vec{\alpha}}[M_f^+(\bar{\varphi},\vec{\alpha};s,u)-M_f^-(\bar{\varphi},\vec{\alpha};s',u')]+\vec{\xi}^{\infty}_{su}-\vec{\xi}^{\infty}_{s'u'}\right\}+\vec{A}_{su}-\vec{A}_{s'u'}+\mu^2R=\vec{0}\\ &E^+(s,u,\mu)\equiv E\equiv E^-(s',u',\mu)\end{aligned}\tag{8.3}$$

where $E^\pm$ denotes the value of the hamiltonian H on the whiskers $W^\pm$, $E\equiv H_0(0,\vec{A}_s,0,0)$ (see §3) and the remainder $R\equiv R(\vec{\alpha},s,s',u,u',\mu)$ is a C^2 function of its arguments.

We regard (8.3) as an implicit function equation determining $\vec{\alpha},u,u'$ at fixed $\bar{\varphi}$ in terms of the parameters s,s',μ.

Note that finding solutions of (8.3) means that the whiskers contain a point with equal $\vec{\alpha},\varphi,\vec{A}$ coordinates: hence the whiskers have a point in common as their energy is the same and hence also the I coordinate has to be the same (notice, in fact, that the derivative of the hamiltonian with respect to I cannot vanish, at $\mu=0$ and at the φ value that we have chosen).

We first determine $u^+(s,\mu)$ and $u^-(s',\mu)$ so that the energy constraint is satisfied: that this is possible follows from the non degeneracy hypothesis (2.7) (see also the last of (5.23)) and from the condition (5.85). The precise argument is basically a repetition of the argument used to deduce lemma 1 from lemma 1', see (5.83) $\div$ (5.85) and we shall not repeat it here. We then substitute $u=u^+$ and $u'=u^-$ in the first of (8.3) and observe that $u^\pm(\cdot,0)=0$.

Hence we are interested in the jacobian matrix of the first of (8.3) at the non-degenerate solution point $s'=s$, $u=u'=0$. Here, if the diffusion curve $\mathcal{L}$ is open for direct diffusion (as we suppose throughout this section), the above (8.3) has the solution $\vec{\alpha}'_s=\vec{\alpha}_s+O(\mu)$, $(\varphi=\bar{\varphi})$, described in the definition following (4.9). In fact by (4.9) at $\vec{\alpha}=\vec{\alpha}_s$, $u=u'=0$:

$$\partial_{\vec{\alpha}}(M_f^+-M_f^-)\equiv\partial_{\vec{\alpha}}M_f(\vec{\alpha};s,0)=\vec{0}\tag{8.4}$$

and by (4.10) the jacobian of (8.3) at $\vec{\alpha}_s$, $s=s'$, $u=u'=0$, is:

$$\mu\,\partial_{\vec{\alpha}\vec{\alpha}}(M_f^+-M_f^-)+O(\mu^2)\equiv\mu\,\partial_{\vec{\alpha}\vec{\alpha}}M_f(\vec{\alpha}_s;s,0)+O(\mu^2)\tag{8.5}$$

so that by the implicit function theorem, for μ small and $\varphi=\bar{\varphi}$, we have the solution $\vec{\alpha}'_s$ of (8.3) when $s=s'$.

Again, the implicit function theorem implies, giving up control of the values of the constants (for simplicity rather than by necessity), that there is a constant $G_1>0$ such that for:

$$|s-s'|<G_1\mu^2\tag{8.6}$$

and for μ small enough, the equation (8.3) admits a non degenerate solution $\vec{\alpha}(s,s')=\vec{\alpha}'_s+O(\mu)$. Here we do not find how the constant G_1 depends on dimensionless quantities but, of course, this could be done if necessary.

Hence if $s, s' \in \Sigma_\mu$ and verify (8.6) there is a heteroclinic intersection $H_{ss'}$ between the whiskers $W_{stable}(s)$ and $W_{unstable}(s')$ at energy level E.

As an application we consider the case in which f is a trigonometric polynomial of degree N_f and we suppose that the path $\mathcal{L}$ is contained in a region free of resonances (see definition following (7.1)) to order $p = 2c$ for the perturbing function f, where c is the constant appearing in lemma 1, §5, see (5.4) and the comment after lemma 3 of §7.

Then by (7.3) we can find a sequence $s_0 = \sigma_1 < \ldots < \sigma_N = s_1$ with:

$$|\sigma_{i+1} - \sigma_i| < G_1 |\mu|^2, \qquad i = 1, \ldots, N-1 \qquad \text{and } \sigma_i \in \Sigma_\mu, \tag{8.7}$$

provided $|\mu|$ is small enough, *i.e.* provided $|\mu| < \mu^{*'}$ so that the results of §5,§6,§7 can be used to infer the existence of heteroclinic intersections: it appears that $N < G_2 |\mu|^{-2}$, with $G_2 > 0$ conveniently chosen.

Therefore for each $i = 1, \ldots, N-1$ we have an invariant torus $\mathcal{T}_i \equiv \mathcal{T}(\sigma_i)$ with two whiskers W^i_{stable} and $W^i_{unstable}$ and for $i = 1, \ldots, N-1$ we consider the heteroclinic intersections of $W^i_{unstable}$ with W^{i+1}_{stable}. Such intersection contains a curve: it is an orbit "spiraling" onto $\mathcal{T}_i$ as $t \to -\infty$ and onto $\mathcal{T}_{i+1}$ as $t \to +\infty$. On each of such orbits we fix a unique point:

$$H_i \in W^i_{unstable} \cap W^{i+1}_{stable} \tag{8.8}$$

with a φ coordinate prefixed and equal to $\bar\varphi$ (independently of i) as discussed in the previous sections (in particular $\bar\varphi$ is distinct from the equilibrium positions on the separatrix and for the standard pendulum (2.2) it is $\bar\varphi = \pi$).

The motion on the tori is quasi periodic with $l-1$ frequencies and, by lemmata 1,3, we can suppose that the frequencies of such quasi periodic motions have a non resonance constant $C(\sigma_i)$ verifying (5.81):

$$C(\sigma_i) < G_3 |\mu|^{-G_4} \equiv C_\mu \tag{8.9}$$

(one can take $G_4 = (p+1)/7$ by the discussion of §7,§5).

We imagine to draw around each of the tori $\mathcal{T}_i$ a small vicinity U_i of radius $\hat r$, i independent but so small that inside it the tori and their whiskers can be described by parametric equations, analytic in some standard coordinates $(p, \vec A', q, \vec\psi)$ with: $|\vec A' - \vec A'_i(0)| < \hat r$, $|p| < \hat r$, $|q| < \hat r$, $\vec\psi \in T^{l-1}$, where:

$$\vec A'_i(J) \equiv \vec A^\infty(s_i, u(s_i, J, \mu), J, \mu) \tag{8.10}$$

where (consistently with the use of the symbol in §5 we call $\vec A^\infty(s, u, J, \mu)$ the solution to (5.14) with h_0 replaced by h_∞: see also the comment following (5.69) and ♣8.16 below).

In this section we shall refer to such coordinates as to "normal coordinates". We suppose $\hat r$ so small that the heteroclinic points H_j are outside the sets U_i. We shall proceed in a asymmetric fashion, treating differently the stable and the unstable directions (but we could, everywhere below, interchange their roles). The following analysis can be followed rather easily if on tries to draw a picture (which after some thought becomes not too hard) of the various geometrical concepts that we need and introduce below.

For $i = 1, \ldots, N-1$ we fix in U_i a point $E_i^u \in W^i_{unstable}$ which is on the heteroclinic orbit of H_i and with normal coordinates given by:

$$E_i^u \in W^i_{unstable}, \qquad E_i^u = (0, \vec{A}'_i(0), \hat{r}/2, \vec{\psi}_i), \quad i = 1, \ldots, N-1 \tag{8.11}$$

Here $\vec{\psi}_i$ is uniquely determined: just evolve with the solution of the equations of motion, $t \to S_t H_i$, the datum H_i backwards in time until its q-coordinate becomes (meaningful and) equal to $\hat{r}/2$:

$$E_i^u = S_{-T_i^u} H_i \tag{8.12}$$

The times T_i^u are all bounded by some T' which is i-independent, which can be estimated at small enough μ only in terms of $\hat{r}$ and of the free part of the hamiltonian: $T_i^u \le T' \equiv G_5$.

We now define the surface element $\tilde{\Delta}_i^s \equiv \big(W^{i+1}_{stable} \cap M_{\bar{r}}(H_i)\big)^*$ where $M_{\bar{r}}(H_i)$ is the ball of radius $\bar{r}$ around H_i and the $*$ signifies that:

1) we consider the connected component of the intersection containing the center of $M_{\bar{r}}$ ($\bar{r}$, at the moment, is rather arbitrary, for instance take it so that $M_{\bar{r}}(H_i) \cap U_i = \emptyset$ and $\bar{r} < \hat{r}$);
2) inside the connected component we select the points with φ coordinate equal to $\bar{\varphi}$ (*i.e.* equal to the φ coordinate of the heteroclinic point).

The set $\tilde{\Delta}_i^s$ is a $(l-1)-$ dimensional regular submanifold of the $2(l-1)-$ dimensional manifold $\tilde{M}_i$ obtained by intersecting $M_{\bar{r}}$ with the energy level E (E being the energy of the diffusion curve and, therefore, of the whiskers) and with the points with $\varphi = \bar{\varphi}$:

$$\tilde{M}_i = M_{\bar{r}} \cap \{H = E\} \cap \{\varphi = \bar{\varphi}\}. \tag{8.13}$$

The same is true for the analogous surface element $\tilde{\Delta}_i^u \equiv (W^i_{unstable} \cap M_{\bar{r}})^*$. Moreover, the nondegeneracy condition (4.10), *i.e.* [see (8.1), (8.4), (8.5)]:

$$\det \partial_{\vec{\alpha}} (\vec{A}^i_{unstable} - \vec{A}^{i+1}_{stable})|_{(\vec{\alpha},\varphi)=(\vec{\alpha}_i,\bar{\varphi})} \neq 0, \tag{8.14}$$

(where $\vec{\alpha}_i \equiv \vec{\alpha}_{\sigma_i,\sigma_{i+1}}(\bar{\varphi}, \mu)$ is the locally unique nondegenerate solution of $\vec{A}^i_{unstable} = \vec{A}^{i+1}_{stable}$ whose existence has been proved in §7), implies that $\tilde{\Delta}_i^s$ and $\tilde{\Delta}_i^u$ *intersect transversally* at H_i. This means that any pair of tangents to the two surfaces form an angle bounded away from 0.

Note also that the determinants in (8.14) are $O(\mu)$ so that also the angle between $\tilde{\Delta}_i^s$ and $\tilde{\Delta}_i^u$ (*i.e.* the smallest angle between corresponding tangent vectors) is bounded below by a quantity of the same order.

We transport the above two surface elements inside the region U_i by using the hamiltonian flow S_t, solving the Hamilton equations, as follows.

Consider a $2(l-1)$ neighborhood, M_i, of E_i^u, obtained by considering points with $q-$coordinate equal to $\hat{r}/2$ and having energy equal to E. By taking $\bar{r}$ small enough we can construct a (smooth) diffeomorphism, F_i, of $\tilde{M}_i$ into M_i as follows. For each $x \in \tilde{M}_i$ we can find a (smooth) real function $\tau = \tau(x)$ so that $S_{-(T_i^u+\tau)}(x)$ has $q-$coordinate exactly equal to $\hat{r}/2$ (in particular: $\tau(H_i) = 0$) and then we set $F_i(x) = S_{-(T_i^u+\tau)}(x)$.

Hence we can define $\Delta_i^s \equiv F_i\tilde{\Delta}_i^s$, $\Delta_i^u \equiv F_i\tilde{\Delta}_i^u$ and observe that, since transversality is preserved under diffeomorphisms, Δ_i^s and Δ_i^u intersect transversally at E_i^u (with angle of order μ).

Furthermore Δ_i^s and Δ_i^u can be represented, in normal coordinates, as regular *graphs* over the angles $\vec{\psi}$'s for $\vec{\psi}$ varying in some open $(l-1)$-dimensional set, D_i, whose size is independent of μ. For simplicity (and without loss of generality) we let D_i be a $(l-1)$-ball around $\vec{\psi}_i$: $D_i \equiv \{\vec{\psi} \in T^{l-1} : |\vec{\psi}-\vec{\psi}_i| < \delta\}$ for some positive δ. Then the above construction and lemma 1 of §5 imply that Δ_i^u is simply $\{(p, \vec{A}', q, \vec{\psi}) = (0, \vec{A}'(0), \hat{r}/2, \vec{\psi}_i)|\vec{\psi} \in D_i\}$, while Δ_i^s will have the form:

$$\{(\vec{A}', \vec{\psi}, p, q) = (\vec{\bar{A}}_i(\vec{\psi}), \vec{\psi}, p_i(\vec{\psi}), \hat{r}/2)|\,\vec{\psi} \in D_i\} \tag{8.15}$$

for suitable (smooth) functions $\vec{\bar{A}}_i$ and p_i $\delta-$close, respectively, to $\vec{A}_i'(0)$ and 0.

Note that, since the energy E is fixed, the $p_i(\vec{\psi})$ is actually computable in terms of $\vec{\psi}$ and $\bar{A}_i$; in fact the hamiltonian H, in normal coordinates, takes the form:

$$H_\infty \equiv h_\infty(\vec{A}', pq, \mu) + f_\infty(\vec{A}', \vec{\psi}, p, q, \mu) \tag{8.16}$$

with $h_\infty - h_0$ and f_∞ of $O(\mu)$ and with f_∞ vanishing, together with all its derivatives, when $\vec{A}' = \vec{A}_i'(pq)$. This is in fact the content of the results of §5, see lemma 2; therefore (see (2.7) and use $J = pq$, $r = \hat{r}/2$):

$$\partial_p H_\infty|_{(\vec{A}_i'(0),\vec{\psi}_i,0,\hat{r}/2)} = \frac{\hat{r}}{2}\partial_J h_0 + O(\mu) > 0, \tag{8.17}$$

which, by the implicit function theorem, allows us to express p in terms of $\vec{A} = \vec{\bar{A}}(\vec{\psi})$, provided δ is small enough.

In fact, recalling that on Δ_1^s the energy is fixed, we realize that by taking the $\vec{\psi}$-gradient of the relation $H_\infty(\vec{A}_1(\vec{\psi}), \vec{\psi}, p_i(\vec{\psi}), \hat{r}/2) \equiv E$ and evaluating the result at the center $\vec{\psi}_i$ (see (8.16) and the comments after it) one obtains: $\partial_{\vec{\psi}} p_i(\vec{\psi}_1) = 2(\hat{r}\partial_J h_\infty)^{-1}\partial_{\vec{\psi}}\vec{\bar{A}}_1 \cdot \vec{\omega}_1$ with $\vec{\omega}_1 = \vec{\omega}_s(1+u) \neq \vec{0}$, $(\partial_J h_\infty)^{-1} = (1+u')g_{su} \neq \infty$, see (5.6),(5.21); thus the vector $\partial_{\vec{\psi}} p_i(\vec{\psi}_1) \neq 0$ and its length is of size $O(\mu)$, see ♣8.18.

In the normal coordinates, transversality of Δ_i^s and Δ_i^u reads simply:

$$\det \partial_{\vec{\psi}}\vec{\bar{A}}_i|_{\vec{\psi}_i} \neq 0 \tag{8.18}$$

and since such a determinant is of $O(\mu)$ one sees that there exists a constant $G_6 > 0$ such that the set $\{p =$ value determined by the energy conservation; $\vec{\bar{A}}_i(\vec{\psi})$, $q = 0$, $\vec{\psi} \in D_i\}$ contains a $(l-1)$- ball of radius which is not too small, namely $G_6|\mu|$, around $\vec{\bar{A}}_i(\vec{\psi}_i) = \vec{A}_i'(0)$. This property is important in the following construction.

Let $\varepsilon_{\|}^i > \varepsilon_{\perp}^i > 0$ and let $\xi_i \in \Delta_i^s$ be a point with $\vec{\psi}$-coordinate equal to some $\vec{\chi}_i$, $|\vec{\chi}_i - \vec{\psi}_i| < \delta/2$ and p-coordinate $p_i = p(\vec{\chi}_i) > 2\hat{r}\varepsilon_{\perp}^i$. Let B_i be the set:

$$B_i = \Big\{|\vec{\psi} - \vec{\chi}_i| < \varepsilon_{\|}^i,\ |q - \frac{\hat{r}}{2}| < \varepsilon_{\perp}^i \frac{\hat{r}}{2},\ \ |p - p_i| < \varepsilon_{\perp}^i \hat{r},\ |\vec{A} - \vec{\bar{A}}_i(\vec{\chi}_i)| < \hat{r}^2 \varepsilon_{\perp}^i\Big\} \tag{8.19}$$

It is clear that if $\varepsilon_{\|}^i \le \delta/C_1$, for C_1 large enough, and $\varepsilon_{\perp}^i$ is small enough $B_i \subset U_i \cap S_{-T_i^u} M_{\bar{r}}(H_i)$ and in a suitable time, denoted $T_i^u + T_i^s$ and bounded uniformly in i, *it evolves into a set containing the point* $\xi_i' \equiv (\vec{A}_{i+1}'(0), \vec{\chi}_i', \hat{r}/2, 0) \in (W_s^{i+1} \cap U_{i+1})^{cc}$ (where $^{cc} \equiv$ connected component) *as well as a set*:

$$B_i' = \Big\{|\vec{\psi} - \vec{\chi}_i'| < \frac{\varepsilon_{\|}^i}{C_2},\ |q| < \varepsilon_{\perp}^i \hat{r}/2C_2,\ |p - \frac{\hat{r}}{2}| < \frac{\varepsilon_{\perp}^i \hat{r}}{C_2},\ |\vec{A} - \vec{A}_{i+1}'(0)| < \frac{\hat{r}^2 \varepsilon_{\perp}^i}{C_2}\Big\} \tag{8.20}$$

where $C_2 > 1$ is a suitable constant of $O(1)$ (more precisely of order $O(\exp g_i(T_i^u + T_i^s))$, uniformly bounded in i).

All that the latter statement is saying is that the flow S_t takes a finite time to carry a point on the heteroclinic orbit (or close to it) and at distance $\sim \hat{r}$ from the torus $\mathcal{T}_i$ to a point close within the same distance to the torus $\mathcal{T}_{i+1}$. During such finite time nothing bad can really happen; all expansions and contractions in phase space being bounded by a suitably large constant (basically depending only on the size of $\hat{r}$).

We take $\varepsilon_{\|}^1 = \delta/C_1$, $\varepsilon_{\perp}^1 \le G_7\mu$ and suppose μ small enough: this guarantees that not only B_1 has the above mentioned inclusion property but also that as $\vec{\psi}$ varies around $\vec{\psi}_1$ then $p_1(\vec{\psi})$ becomes different from 0 and spans an interval of $O(\mu)$ by (8.17),(8.18) (thus implying the existence of the point $\vec{\chi}_1$ with the above properties).

Having constructed B_i and hence B_i' we consider the evolution of the set B_i': we shall see that it evolves in time and crosses Δ_{i+1}^s at a time T_i that can be chosen so that its image still contains a set B_{i+1} around some point $\xi_{i+1} \in \Delta_{i+1}^s$ described by (8.19) with a suitable $\vec{\chi}_{i+1}$ and with $\varepsilon_{\perp}^{i+1}$, $\varepsilon_{\|}^i$ such that:

$$\varepsilon_{\perp}^{i+1} = (C_3^{-1}\varepsilon_{\perp}^i)^a \quad \varepsilon_{\|}^{i+1} = C_3^{-1}\varepsilon_{\|}^i, \quad T_i \le 2g^{-1}\log(C_3/2\varepsilon_{\perp}^i) \tag{8.21}$$

for suitable constants $a, C_3 > 1$ (determined below).

To simplify the analysis we take $\varepsilon_{\perp}^1$ exponentially small with respect to μ otherwise the above expression for T_i would be more involved (see ♣8.29). Hence we define $\varepsilon_{\perp}$ by:

$$\varepsilon_{\|}^1 \equiv \frac{\delta}{C_1}, \qquad \varepsilon_{\perp}^1 \equiv \frac{C_2}{2}\exp\left[-\bar{g}\vartheta\Big(\frac{C_2}{|\mu|\varepsilon_{\|}^1}\Big)^{G_{11}}\right] \tag{8.22}$$

where C_2 is the constant fixed above (cfr. (8.20)), $\bar{g}$ is a ($i-$independent) upper bound on g_i and ϑ and G_{11} are suitable constants related to diophantine properties of the motion on the invariant tori (see below).

Assuming (8.21),(8.22) it will follow that the time of drift can be bounded by $T^* \le \sum_{i=1}^N (T_i + 2T_0)$ if T_0 is a i-independent bound on $T_i^u + T_i^s$. Recalling that $N \le G_2\mu^{-2}$, we see that (8.21) implies (given the choice of $\varepsilon_{\|}^1$ and $\varepsilon_\perp^1$ in (8.22)):

$$T^* \le C_4 \vartheta e^{C_5/\mu^2} \tag{8.23}$$

for μ small enough and suitable C_4, C_5.

It remains to check the above recursion in (8.21). Let $\vec{\chi}_i$ be the $\vec{\psi}$-coordinate of ξ_i'. In B_i' we consider, for $|q| < \varepsilon_\perp^i \hat{r}/2C_2$ and μ small enough the points $z(q,\vec{\psi}) = (\hat{r}/2, \vec{A}'_{i+1}(q\hat{r}/2), q, \vec{\psi})$ with $|\vec{\psi} - \vec{\chi}_i'| < \varepsilon_{\|}^i / C_2$: such points are indeed in B_i' because the function $J \to \vec{A}'_{i+1}(J)$ is differentiable with derivative of order μ, and we see that, see (8.20):

$$|\vec{A}'_{i+1}(q\hat{r}/2) - \vec{A}'_{i+1}(0)| < G_8|\mu| q\hat{r}/2 < \hat{r}^2 \varepsilon_\perp^i / C_2 \tag{8.24}$$

if $C_2 G_8 q |\mu| \hat{r}^{-1} < \varepsilon_\perp^i$, *i.e.* if $|\mu|$ is small enough.

The evolution of z, which for all $q, \vec{\psi}$ has energy E, is simply (by lemma 1, §5):

$$S_T\, z(q,\vec{\psi}) = (q e^{g_i T},\ \vec{A}'_{i+1}(q\hat{r}/2),\ \hat{r} e^{-g_i T}/2,\ \vec{\psi} + \vec{\omega}_i T) \tag{8.25}$$

with ω_i, g_i depending on μ and on the product $q\hat{r}/2$, see (5.6). Therefore we can define $T(q)$ by:

$$q e^{g_i T(q)} = \hat{r}/2 \tag{8.26}$$

and $T(q) \xrightarrow[q\to 0]{} \infty$.

Let T_ε be the time necessary in order that a trajectory on the torus T^{l-1} fills, running quasi periodically with velocity $\vec{\omega}_i$, the torus within a distance $\varepsilon/2$, *i.e.* such that no point of T^{l-1} has distance $\ge \varepsilon/2$ from the set $\{\vec{\omega}_i t |\, 0 \le t \le T_\varepsilon\}$. Let $\eta(\varepsilon) \equiv e^{-\bar{g} T_\varepsilon}$ if $\bar{g}, g$ are such that $\bar{g} \ge g_i \ge g > 0$ for all $i = 1, \ldots, N-1$.

Call $q_{\max} \equiv \hat{r}\varepsilon_\perp^i / C_2$ and $q_{\min} \equiv q_{\max} \exp(-g_i T_{\varepsilon_{\|}^i/C_2})$ so that $q_{\min} \ge q_{\max} \eta(\varepsilon_{\|}^i / C_2)$ and $T(q_{\min}) - T(q_{\max}) = T_{\varepsilon_{\|}^i/C_2}$.

Then it is clear that as q varies between $q_{\max}$ and $q_{\min}$ there is a value $\bar{q}$ such that $T(\bar{q})$ verifies:

$$|\vec{\chi}_i' + \vec{\omega}_i T(\bar{q}) - \vec{\psi}_{i+1}| < \frac{\varepsilon_{\|}^i}{2C_2}, \qquad T(\bar{q}) \le T(q_{\max}) + T_{\varepsilon_{\|}^i/C_2} \tag{8.27}$$

¿From the theory of quasi periodic motions it follows that T_ε is bounded above in terms of the non resonance constants C_0, τ of the frequencies $\vec{\omega}_i$, see (3.1), by $G_9' C_0 \varepsilon^{-G_9}$ for some $G_9', G_9 > 0$ (one can take $G_9 = \tau + 2$). And in our application it is, see (8.9), $C_0 \le G_3 |\mu|^{-G_4}$, so that:

$$T_\varepsilon \le \vartheta\, (\varepsilon|\mu|)^{-G_{11}} \equiv \bar{T}(\varepsilon) \tag{8.28}$$

for suitable $\vartheta, G_{11} > 0$.

Hence (8.27) implies that:

$$T(\bar{q}) \le \vartheta\,[C_2(\mu\varepsilon^i_\parallel)^{-1}]^{G_{11}} + g^{-1}\log[C_2(2\varepsilon^i_\perp)^{-1}] \equiv T_i \tag{8.29}$$

At this time the trajectory of $z(\bar{q},\vec{\chi}'_i)$ is in a point $\bar{\xi}_{i+1}$ which has q-coordinate equal to $\hat{r}/2$ and $\vec{A}$ coordinate still equal to the original value, while the p coordinate is:

$$p = \frac{\hat{r}}{2}e^{-g_i T(\bar{q})} \ge \frac{\hat{r}}{2}e^{-\bar{g}\hat{T}_i} > 0 \tag{8.30}$$

The set image of B'_i has dimensions constant in the $\vec{\psi}$ variables (because the quasi periodic motion is rigid), while it contracts by at most $e^{-\bar{g}\hat{T}_i}G_{12}$ in the $\vec{A}$ variables (recall (8.29) and that $\bar{g}$ is an upper bound for the expansion rates), for some G_{12} that we suppose > 1, and also by at most $G_{12}e^{-\bar{g}\hat{T}_i}$ in the p-variables.

The $\vec{A}_{i+1}$-coordinate, being equal to its initial value, has distance from the point $\vec{\bar{A}}_{i+1}(\vec{\psi}_{i+1}) \equiv \vec{A}'_{i+1}(0)$ bounded by (8.24), *i.e.* by $G_8|\mu|\hat{r}\bar{q}/2 \le G_8|\mu|\hat{r}^2\varepsilon^i_\perp/2C_2$. This is *far less than* $|\mu|\varepsilon^i_\parallel/2C_2$ if, as we shall suppose (see below), $\varepsilon^i_\parallel \gg G_8\hat{r}^2\varepsilon^i_\perp$. Hence we can find a point $\vec{\bar{\psi}}$ such that:

$$\vec{\bar{A}}_{i+1}(\vec{\bar{\psi}}) = \vec{A}'_{i+1}(\bar{q}\hat{r}/2), \qquad |\vec{\bar{\psi}} - \vec{\psi}_{i+1}| < \varepsilon^i_\parallel/4C_2 \tag{8.31}$$

In fact such $\vec{\psi}$ can be estimated by the implicit function theorem by:

$$|\vec{\bar{\psi}} - \vec{\psi}_{i+1}| < |\vec{A}'_{i+1}(q\hat{r}/2) - \vec{A}'_{i+1}(0)|O(|\mu|^{-1}) \le \frac{G_8|\mu|\hat{r}^2\varepsilon^i_\perp}{2C_2G_{13}|\mu|} \tag{8.32}$$

provided:

$$G_8\frac{\hat{r}^2\varepsilon^i_\perp}{2G_{13}} < \frac{\varepsilon^i_\parallel}{4C_2} \tag{8.33}$$

By energy conservation it must be that $\bar{p}_{i+1}(\vec{\bar{\psi}}) \equiv \exp[-g_iT(\bar{q})]\hat{r}/2$ is such that the point $(\vec{\bar{A}}_{i+1}(\vec{\bar{\psi}}), \vec{\bar{\psi}}, \bar{p}_{i+1}(\vec{\bar{\psi}}), \hat{r}/2)$ will be inside W^u_{i+1} and it is such that the set B_{i+1} defined by (8.19) with i replaced by $i+1$, $\vec{\chi}_{i+1} \equiv \vec{\bar{\psi}}$, $p_{1\,i+1} \equiv \bar{p}_{i+1}(\vec{\bar{\psi}})$ and parameters:

$$\varepsilon^{i+1}_\perp \le \frac{\varepsilon^i_\perp\, e^{-\bar{g}\hat{T}_i}}{2G_{12}}, \qquad \varepsilon^{i+1}_\parallel \le \frac{\varepsilon^i_\parallel}{4C_2} \tag{8.34}$$

is contained into M_{i+1} and $S_{-T(\bar{q})}(B_{i+1}) \subset B'_i$.

It is now easy to check that letting $C_3 \equiv C_2^2G_{12}/2$, $a \equiv C_3^{G_{11}}$ (> 3) and defining the $\varepsilon^{i+1}_\cdot$'s as in (8.21),(8.22),(8.19) we see (inductively) that for all i's the second term in (8.29) dominates over the first and (8.34) holds for all i's together with $T_i \le \hat{T}_i \le$

$2g^{-1}\log(C_3/2\varepsilon^i_\perp)$ for $|\mu|$ small enough; and, finally, also (8.33) will be readily verified for small $|\mu|$.

Hence we conclude that drift takes place on a time scale T^* *bounded above* by (8.23).

Clearly instead of trying to go systematically forward along the ladder of whiskers we could have chosen an arbitrary up/down pattern and found the existence of an initial datum which would have followed the prescribed pattern (taking essentially the same time). In this way we can construct a collection of 2^N, N being of the order of $|\mu|^{-2}$, of sets of (very small but positive) measure of initial data collected according to the up/down pattern that they follow in their evolution along the ladder of whiskers. If we give equal probability to data corresponding to each pattern and choose one of them at random we shall see that it climbs brownianly the ladder: *i.e.* the existence of what we have called drift and of diffusion are essentially the same phenomenon.

Finally, we remark that the analysis of this section *extends to the case of forced systems* (5.91). Recall from the discussion at the end of §5, that one can construct whiskered tori $W^\pm(s,u)$ for (5.91) (with the right λ–frequency ω) for all $s\in\Sigma_\mu$ and $u\in[-\bar u,\bar u]$. The second of (8.3) is trivially solved (as (5.91) is linear in the clock action B) and the analysis of the first of (8.3) is then carried out as discussed above (of course various notions, such as the diffusion path, have now to be properly reinterpreted: see end of §5).

§9 A class of exactly soluble homoclines

After reduction to normal form, the motion of a quasi integrable hamiltonian system near a resonance is described by a hamiltonian with two perturbation parameters [N,BG]:

$$H=\vec\omega\cdot\vec A+\frac{\vec A^2}{2J}+\frac{I^2}{2J_0}+v(\varphi)+f'(I,\vec A,\varphi,\vec\alpha)\tag{9.1}$$

with:

$$\vec\omega=\vec\omega_0\eta^{-1/2},\qquad f'=\mu f(\varphi,\vec\alpha)\tag{9.2}$$

where $\eta>0,\mu>0$, the function J is analytic in $\eta^{1/2}\vec A$, η; and the functions J_0,v,f_j are regular analytic in the variables $\eta^{1/2}I,\eta^{1/2}\vec A,\eta$ while μ is generally much smaller than η *but related* to it: *e.g.* $\mu=\eta^Q$ with Q large (in fact as large as wished, at the expense of the complexity of the f's). All functions are analytic in the angles on which they depend.

If the model comes from the perturbation theory of a degenerate system, like a celestial mechanics system or a forced system with clock variables $(B,\lambda)\equiv(A_1,\alpha_1)$, the (9.1) may contain some additional features and some variations, see §11, §12.

In this section we look at a more special class of systems which we shall call the *even class*:

$$H_{even}=\vec\omega\cdot\vec A+h(\vec A)+\frac{I^2}{2J_0}+J_0g_0^2(\cos\varphi-1)+\mu\sum_\nu f_\nu\cos(\vec\nu\cdot\vec\alpha+n\varphi)\tag{9.3}$$

where J_0, g_0^2, f depend, in general, on $I, \vec{A}$ and $\nu \equiv (n, \vec{\nu})$, $\vec{\nu} \neq \vec{0}$.

The hamiltonian (9.3) is remarkable because the homoclinic points at $\varphi = \pi$ can be, often, computed exactly (namely they are at $\vec{\alpha} = \vec{0}$, modulo some convergence questions).

An important class of examples, with $l = 2$, is:

$$H_p = \omega B + \frac{I^2}{2J_0} + J_0 g_0^2(\cos\varphi - 1) + \mu f(\eta^{1/2} I, \varphi, \lambda) \tag{9.4}$$

with J_0, g_0 positive constants, $\omega = \omega_1/\eta^{1/2}$. And for $l = 3$:

$$H_c = \omega B + \eta\frac{A^2}{2J} + \frac{I^2}{2J_0} + J_0 g_0^2(\cos\varphi - 1) + \mu f(\eta^{1/2} I, \eta^{1/2} A, \varphi, \vec{\alpha}) \tag{9.5}$$

with $J = J(\eta^{1/2} A, \eta)$, J_0, g_0 positive constants, $\omega = \omega_1/\eta^{1/2}$, and, typically, $A = O(\eta^{-1/2})$. The functions f may also depend on η, μ and they will be supposed analytic in η, μ near 0 and in $\vec{\alpha}$ for $|\operatorname{Im}\alpha_j| < \xi$, $\eta e^{|\operatorname{Im}\varphi|} < \xi$, and averageless and uniformly bounded in this domain (as $\eta \to 0$). We make the identification $\alpha_1 \equiv \lambda$, $A_1 \equiv B$, where λ, B are the clock angle and the clock action.

The model (9.4) is a classical forced pendulum (compare [La2, HMS, ACKR, DS, Ge]) and the model (9.5) is a system arising in some celestial mechanics problems, see §12.

The determination of the location of the homoclinic point is based on a very simple *symmetry argument*. Therefore we present it in the simple case of (9.3) with $J, J_0, f_{\vec{\nu},m}$ action independent (*i.e.* constants) and all the phases are isochronous (quite simple, of course, but useful for illustration purposes).

We shall show that for all the models in the even class the homoclinic equations at $\varphi = \pi$ ((6.47)) can be solved to all orders of perturbation theory and have $\vec{\alpha} = \vec{0}$ as solution. Since the difference between the two whiskers at $\varphi = \pi$ and at any $\vec{\alpha}$ is analytic in $\vec{\alpha}, \mu$, by lemma 1', we see that $\vec{\alpha} = \vec{0}$ is an actual solution of the equation, as long as μ is small enough.

We begin by remarking that from the explicit form of $W(t)$, see appendix 9, the wronskian can be thought as formed with four $l \times l$ blocks (l is arbitrary): the action action and the phase phase blocks are even functions of t, while the other two blocks are odd functions of t. If p denotes an even function (of t) and d denotes an odd function, we can write symbolically $W(t) = \begin{pmatrix} p & d \\ d & p \end{pmatrix}$. Similarly the vectors F^k and X^k in (6.8) can be thought as column vectors with two l dimensional columns.

Set right away $\vec{\alpha} = \vec{0}$: then the blocks of F^1 are of parity $\begin{pmatrix} d \\ p \end{pmatrix}$, thus we see that, regarding the initial data for X^1 as (constant) even or odd functions:

$$X^1(t) = \begin{pmatrix} p & d \\ d & p \end{pmatrix} \left[\begin{pmatrix} p \\ 0 \end{pmatrix} + \int_0^t \begin{pmatrix} \tilde{p} & \tilde{d} \\ \tilde{d} & \tilde{p} \end{pmatrix} \begin{pmatrix} d' \\ p' \end{pmatrix} d\tau \right] = \begin{pmatrix} p'' \\ d'' \end{pmatrix} \tag{9.6}$$

and we see that the homoclinic conditions have the form, see (6.43):

$$(p_\infty - p_{-\infty}) + \int_{-\infty}^{\infty} (pd + dp) = 0 \tag{9.7}$$

because one can see that also the contribution from $t = \pm\infty$ vanish as F^1 is even in the phase components. Thus (9.7) is automatically satisfied. In this case one sees by direct calculation that $p_\infty - p_{-\infty} = 0$.

However in general functions $p_\infty - p_{-\infty}$ vanish, if constructed in the way they are in the homoclinic equation. In fact, see (6.43), they are generated by an odd function $F(\vec\omega t, t)$ which as $t \to \pm\infty$ converges exponentially to $F(\vec\omega t, \pm\infty)$ in the sense of our functions, so that we can proceed to computing the harmonics of F via the formula:

$$\begin{aligned} F_{\vec\nu}(+\infty) &= \lim_{T\to+\infty} T^{-1} \int_0^T F(\vec\omega t, t) e^{-i\vec\omega\cdot\vec\nu t} dt = \\ &= \lim_{T\to+\infty} -T^{-1} \int_0^T F(-\vec\omega t, -t) e^{-i\vec\omega\cdot\vec\nu t} dt = -F_{-\vec\nu}(-\infty) \end{aligned} \tag{9.8}$$

which immediately implies that the expression outside the integral in the homoclinic equations for the $\vec A$ variables, *i.e.* $\sum_{\vec\nu} (F_{\vec\nu}(+\infty) - F_{\vec\nu}(-\infty))(\vec\omega\cdot\vec\nu)^{-1}$, denoted symbolically $p_\infty - p_{-\infty}$ in (9.7) does vanish.

The above formula implies also immediately that $F(\vec\omega t, \infty) = -F(-\vec\omega t, -\infty)$ so that the integrand in (6.43) is odd and therefore the homoclinic condition holds.

We now assume inductively that the block structure of F^k and X^k is respectively d, p and p, d: if so, the homoclinic conditions will be satisfied to all orders, for the same reason they were satisfied at $k = 1$. Clearly if the assumption is verified for $h = 1, \ldots, k-1$ and if we prove that as a consequence the F^k has the correct parity property, also X^k will have the correct parity property.

The parity of F^k is immediately checked by inspection of (6.29) *i.e.* :

$$F^k(\vec\omega t, t) = \frac{\partial_\mu^k}{k!} E\,[\mu\partial f(X_0 + \bar X^k) + \partial H_0(X_0 + \bar X^k) - \Big(\partial H_0(X_0) + \partial^2 H_0(X_0)\bar X\Big)]|_{\mu=0} \tag{9.9}$$

where $\bar X^k$ is the order k truncation of $\bar X$: just observe that the matrix E inverts parities, that $\partial b(X)$ is of type $\begin{pmatrix} p \\ d \end{pmatrix}$ for any $X = \begin{pmatrix} p \\ d \end{pmatrix}$ and any function b even in $\vec\alpha$ and φ, and finally that $\partial^2 H_0(X^0) = \begin{pmatrix} p & d \\ d & p \end{pmatrix}$ so that $\partial^2 H_0 X = \begin{pmatrix} p \\ d \end{pmatrix}$.

One easily checks that the argument is neither affected by action dependence of the various coefficients nor by lack of isochrony: in fact the unperturbed motion is such that the action dependence on t is even and this is all one really needs: the only change is that not all the matrix elements of F^k relative to the angle block (*i.e.* lower) are zero (but they are still even).

§10 Homoclinic scattering. Large separatrix splitting

The concept of *homoclinic scattering* arises naturally if one compares the homoclinic splitting as seen in the original $\vec{\alpha}$ coordinates or in the *intrinsic* coordinate $\vec{\psi}$ associated with the whiskers normal forms of §5. The first part of this section is devoted to it.

The second part will deal with the theory of the homoclinic splitting in systems with more than 2 degrees of freedom: the main point in the analysis will be that when the unperturbed frequences depend on a paramenter η and one of them becomes large (we shall say *fast*) as $\eta \to 0$ *it is not, in general, true that the homoclinic splitting is smaller than any power in* η, *i.e.* the determinant of the intersection matrix is not exponentially small with some inverse power of η as $\eta \to 0$ (see below for a formal definition). *Unless* all the frequences are fast.

Suppose that there is a homoclinic point at $\vec{\alpha} = \vec{\alpha}_0, \varphi = \pi$. We can regard such point either as a point on the stable whisker, or as a point on the unstable whisker. In this way the point receives, from the parametrization in lemma 1', the coordinates $p = p_0, q = 0$, $\vec{\psi} = \vec{\psi}_0^s$ or $p = 0, q = q_0$, $\vec{\psi} = \vec{\psi}_0^u$, so that:

$$\begin{aligned} Z_-(p_0, 0, \vec{\psi}_0^s) =& \pi = Z_-(0, q_0, \vec{\psi}_0^u) \\ \vec{Z}_\downarrow(p_0, 0, \vec{\psi}_0^s) =& \vec{\alpha}_0 = \vec{Z}_\downarrow(0, q_0, \vec{\psi}_0^u) \end{aligned} \tag{10.1}$$

with the notations of (6.23), *i.e.* Z is the r.h.s. of (6.21); of course all the above functions $(\cdot)_0$ are analytic functions of μ.

We can consider the point on the stable manifold with coordinates $p = p_{\vec{\psi}}, q = 0$ with $p_{\vec{\psi}}$ such that:

$$Z_-(p_{\vec{\psi}}, 0, \vec{\psi}_0^s + \vec{\psi}) = \pi \tag{10.2}$$

We can also consider the point on the unstable manifold with coordinates $q_{\vec{\psi}\,'}, \vec{\psi}_0^u + \vec{\psi}\,'$ such that:

$$Z_-(0, q_{\vec{\psi}\,'}, \vec{\psi}_0^u + \vec{\psi}\,') = \pi \tag{10.3}$$

and lemma 1' guarantees that $p_{\vec{\psi}}, q_{\vec{\psi}}$ are analytic in $\vec{\psi}, \mu$ for μ small; note that $p_{\vec{0}} = p_0$, $q_{\vec{0}} = q_0$.

We shall establish a correspondence between $\vec{\psi}$ and $\vec{\psi}\,'$, if they describe the same $\vec{\alpha}$, so that:

$$\vec{Z}_\downarrow(p_{\vec{\psi}}, 0, \vec{\psi}_0^s + \vec{\psi}) = \vec{Z}_\downarrow(0, q_{\vec{\psi}\,'}, \vec{\psi}_0^u + \vec{\psi}\,') \equiv \vec{\alpha} \tag{10.4}$$

and we denote it as:

$$\vec{\psi}\,' = \vec{\psi} + \vec{\sigma}(\vec{\psi}) \tag{10.5}$$

calling the function $\vec{\sigma}$ the *scattering phase shift*; here we think of ψ, ψ' as functions of the independent variable $\vec{\alpha}$. We also introduce the same function regarded as a function of the common value $\vec{\alpha}$ of the two sides of (10.4):

$$\vec{\sigma}[\vec{\alpha}] \equiv \vec{\sigma}(\vec{\psi}\,) \tag{10.6}$$

if $\vec{\alpha}, \vec{\psi}$ are coordinates of the same points. The functions $\vec{\sigma}$ are, by lemma 1', analytic on T^{l-1} and in μ for small μ, and $\vec{\sigma}(\vec{0}) = \vec{0}, \vec{\sigma}[\vec{\alpha}_0] = \vec{0}$, by definition.

The scattering measures the degree of interaction between the pendulum and the rotators. If the rotators are not isochronous the scattering is an interesting homoclinic property: of course if the j-th rotator is a clock (*i.e.* it is isochronous) then $\sigma_j \equiv 0$ (as it should because it is a clock).

The equations for the homoclinic point can be written:

$$\begin{aligned} Q_+(\vec{\alpha}) =& Z_+(p_{\vec{\psi}}, 0, \vec{\psi}^s_0 + \vec{\psi}) - Z_+(0, q_{\vec{\psi}\,+\vec{\sigma}(\vec{\psi}\,)}, \vec{\psi}^u_0 + \vec{\psi} + \vec{\sigma}(\vec{\psi})) = 0 \\ \vec{Q}_\uparrow(\vec{\alpha}) =& Z_\uparrow(p_{\vec{\psi}}, 0, \vec{\psi}^s_0 + \vec{\psi}) - Z_\uparrow(0, q_{\vec{\psi}+\vec{\sigma}(\vec{\psi})}, \vec{\psi}^u_0 + \vec{\psi} + \vec{\sigma}(\vec{\psi})) = \vec{0} \end{aligned} \tag{10.7}$$

or also as:

$$\begin{aligned} Q^0_+(\vec{\psi}\,) =& Z_+(p_{\vec{\psi}}, 0, \vec{\psi}^s_0 + \vec{\psi}) - Z_+(0, q_{\vec{\psi}}\,, \vec{\psi}^u_0 + \vec{\psi}) = 0 \\ \vec{Q}^0_\uparrow(\vec{\psi}\,) =& \vec{Z}_\uparrow(p_{\vec{\psi}}, 0, \vec{\psi}^s_0 + \vec{\psi}) - \vec{Z}_\uparrow(0, q_{\vec{\psi}}, \vec{\psi}^u_0 + \vec{\psi}) = \vec{0} \end{aligned} \tag{10.8}$$

Hence the *homoclinic splitting* can be measured as a function of $\vec{\alpha}$ by $\vec{Q}(\vec{\alpha})$ or as a function of $\vec{\psi}$ by $\vec{Q}^0(\vec{\psi}\,)$. Note that if $X^\sigma(t;\vec{\alpha})$ denote the evolutions of the stable/unstable motions with initial angle coordinates $(\pi, \vec{\alpha})$, it is $X^s(0;\vec{\alpha}) - X^u(0;\vec{\alpha}) \equiv Q(\vec{\alpha})$.

Therefore there are two interesting sets of homoclinic angles. One set is described by the $\vec{\psi}$ derivatives of $\vec{Q}^0_\uparrow$ defined by (10.8) which we could call the *intrinsic* homoclinic angles. The other set is described by the $\vec{\alpha}$ derivatives of the $\vec{Q}_\uparrow$ at $\vec{\psi}\, = \vec{0}$ or, respectively, at $\vec{\alpha} = \vec{\alpha}_0 \equiv \vec{\alpha}_{hom}$ (the latter derivatives are proportional, via the matrix $\partial_{\vec{\alpha}}\vec{\psi}$ (close to the identity), to the $\vec{\psi}$ derivatives of $\vec{Q}_\uparrow$ in (10.7), phase shift included. The latter will be called the *natural* homoclinic angles. Note that Q^0_+, Q_+ do not appear in the definitions of the intersection matrix as they can be computed from the $Q^0_\uparrow, Q_\uparrow$ by using the fact that the energy of the whiskers can be supposed fixed (and equal for both).

The higher $\vec{\alpha}$ or $\vec{\psi}$ derivatives of $\vec{Q}_\uparrow, \vec{Q}^0_\uparrow$ in $\vec{\alpha}$ or $\vec{\psi}\,$, respectively, will define the *intersection tensors*. The *homoclinic angles* will be the eigenvalues of the the *intersection matrix*, *i.e.* of the matrix of the first derivatives of the $\vec{Q}_\uparrow$, or $\vec{Q}^0_\uparrow$.

More precisely the trigonometric tangent of the homoclinic angles is proportional to the mentioned eigenvalues. The latter have the dimension of an action and, therefore, a normalization constant with the dimension of an inverse of an action has to be introduced to really define the tangents of the angles. A natural normalization could be $(J_0|\vec{\omega}|)^{-1}$, see (9.1).

Definition: *In the case of hamiltonians depending on a parameter η, e.g. (9.4),(9.5), we shall say, that the homoclinic splitting is "smaller than any power" in the parameter η if there exists $c > 0$ such that when the perturbation constant μ is $\mu = \eta^c$ the determinant of the intersection matrix tends to 0 as $\eta \to 0$ faster than any power in η. Likewise in the same situation we say that the homoclinic splitting is "exponentially small" if the determinant of the intersection matrix is asymptotically equal to the determinant of its first order approximation (in μ) and the latter tends to 0 as an exponential of some inverse power of the parameter η.*

If $l = 2$ the results of [Nei] imply, for the model (9.4), that the scattering phase shifts are smaller than any power, together with their derivatives at the homoclinic point. The same holds if $l > 2$ and all the angles rotate at fast speed (*i.e.* with $\vec{\omega}$ as in (9.2)). The intersection matrix and all its derivatives are, under the same conditions, smaller than any power.

The relatioship between the two notions of homoclinic angles and their connection with the scattering phase shifts is outlined in Appendix A11.

If $l > 2$ with mixed fast and slow rotators, like (9.4),(9.5), more analysis is necessary to understand such cases.

In the remaining part of this section we study some more detailed results concerning:

1) the relation between the intersection tensors in the two systems of coordinates ($\vec{\alpha}$ or $\vec{\psi}$). In fact we shall prove that for even hamiltonians all the odd derivatives of $\vec{\sigma}[\vec{\alpha}]$ vanish at the homoclinic point $\vec{\alpha} = \vec{\alpha}_0 \equiv \vec{0}$. Hence *the homoclinic angles are the same* in the natural and in the intrinsic coordinates (in general even models).
The even derivatives do not vanish, in general, but they are bounded in a useful way only if $l = 2$ (or if $l > 2$ and $\vec{\omega}$ has non resonant and large components, see (9.2) and take η small). In such cases, as mentioned above, in fact the homoclinic splitting is smaller than any power in η and therefore (by the discussion in appendix A11) on can see that generically the scattering phase shifts turn out to be also smaller than any power.
2) We consider even hamiltonians, depending on a parameter $\eta > 0$, of the type (9.3) or (9.4),(9.5) and the vector $\vec{\omega}$ will be supposed to have one of the two forms:
$$\vec{\omega} = \vec{\omega}_0 \eta^{-1/2}, \quad \vec{\omega} = (\eta^{-1/2}\omega_1, \eta^{1/2}\omega_2, \ldots, \eta^{1/2}\omega_{l-1}) \tag{10.9}$$
and to verify a diophantine condition $|\vec{\omega} \cdot \vec{\nu}|^{-1} \le \eta^{-b} C_0 |\vec{\nu}|^\tau$ for some $b, C_0, \tau > 0$: in the first case take $b = 1/2$ and $\vec{\omega}_0$ diophantine with constants C_0, τ; in the second case let $(\omega_2, ..., \omega_{l-1})$ verify a diophantine condition with constants C_0 and $\tau > l - 2$, then given $a > 1/2$, it is easy to see that there exists a set $\Omega_1 \subset [\overline{\omega}, \infty)$ with:
$$\text{meas.}\{[\overline{\omega}, \infty) \backslash \Omega_1\} \le \frac{K}{C_0} \eta^a \Big(\sum_{i=2}^{l-1} |\omega_i / \overline{\omega}| \Big)^{\tau - l + 2} \tag{10.10}$$

such that if $\omega_1/\eta \in \Omega_1$ then the above estimate on $|\vec{\omega}\cdot\vec{\nu}|$ holds with $b \equiv a - 1/2$.

We shall study the whiskers of an invariant torus run quasi periodically with rotation spectrum $\vec{\omega}$.

The first of (10.9) will be called the *fast rotation case* and the second will be called the *mixed fast-slow case*. The angles whose rotation velocity is $O(\eta^{-1/2})$ will be called *fast angles* or *fast modes*; the others *slow*. Thus in the first of the cases in (10.9) all the angles (or modes) are fast while in the second case the first angle α_1 is fast and the others are slow.

We shall usually add the hypothesis that f is a trigonometric polynomial of degree N in the $\vec{\alpha}$'s. The results a),b) below will be derived under the additional assumption that g_0, J_0, J, f depend on the parameter η and are unformly bounded and holomorphic in:

$$\mathcal{D} = \{|\eta^{1/2}\vec{A}|,\ |\eta^{1/2}I| < r,\ |\operatorname{Im}\alpha_j| < \xi_0, |\eta(\cos\varphi - 1)| < \xi_0\} \tag{10.11}$$

while the results c),d) require the same properties for g_0, J_0, J but put on f only the requirement of boundedness and holomorphy in $\vec{A}, I$ as above and for $|\operatorname{Im}\alpha_j|, |\operatorname{Im}\varphi| < \xi_0$.

The above η dependence will be recorded by appending a subscript η to the hamiltonian as in H_η.

The hamiltonians we are considering are *even* in the sense of §9. Therefore if $\vec{\omega}$ is as above, the invariant tori constructed by using lemma 1' of §5 will have, for μ small enough (depending on η, in the η dependent cases), whiskers homoclinic at $\varphi = \pi, \vec{\alpha} = \vec{0}$.

The following theorem summarizes our main results about the *homoclinic splitting for even models*:

Theorem 3: *a) The odd derivatives of the scattering phase shifts vanish at the homoclinic point. Hence the homoclinic angles will be the same in both systems of coordinates (hence smaller than any power in the fast rotation cases, i.e. if $l = 2$ or if $l > 2$ and $\vec{\omega}$ is given by the first of (10.9)). This shows that the difference between the two notions of splitting of the whiskers (at the homoclinic point) is a higher order effect.*

b) In the fast rotation cases, i.e. if $l = 2$ or if $l > 2$ and $\vec{\omega}$ is given by the first of (10.9) all the even derivatives of the phase shifts and the odd derivatives of the homoclinic splitting are smaller than any power.

c) The jacobian determinant of the derivatives of the scattering phase shifts are not, in general, smaller than any power in the mixed cases, (i.e. if $l > 2$ and $\vec{\omega}$ is given by the second of (10.9)). The same can be said of the jacobian determinant of the derivatives with respect to $\vec{\alpha}$ or to $\vec{\psi}$ of the homoclinic intersection tensors.

d) If $l > 2$ the second order value of the determinant of the intersection matrix is not smaller than any power as $\eta \to 0$, in general, for the mixed rotation cases.

Remark: Hence one should not be led erroneously to believe that the homoclinic splitting is, as a rule, exponentially small when there is one or more rapidly rotating angle (unless all of them do rotate at fast speed). This is particularly striking in the case d).

Here we prove a); the part of the statement b) concerning the connection between the homoclinic splitting and the homoclinic scattering is briefly discussed in appendix A11 while the statement about the size smaller than any power is not analyzed here as it is well known ([N], [Nei]) and we do not really need it; the proof of c),d) is an explicit check and, to set an example, the calculation is performed in appendix A13 for the statement d).

Proof: We shall take $J, J_0, g_0 > 0$ constants and f depending only on $\vec{\alpha}, \varphi$. Most arguments being based on symmetry properties, the general case is identical. We make the above simplifying assumptions only to have a lighter notation and to exhibit the essence of the argument (as we did in the analogous situation in §9).

The wronskian matrix for such a case is simply related to ♣A9.8:

$$W(t) = \begin{pmatrix} w_{11}(t) & 0 & w_{12}(t) & 0 \\ 0 & 1 & 0 & 0 \\ w_{21}(t) & 0 & w_{22}(t) & 0 \\ 0 & J^{-1}t & 0 & 1 \end{pmatrix} \quad , \quad W(t)^{-1} = \begin{pmatrix} w_{22}(t) & 0 & -w_{12}(t) & 0 \\ 0 & 1 & 0 & 0 \\ -w_{21}(t) & 0 & w_{11}(t) & 0 \\ 0 & -J^{-1}t & 0 & 1 \end{pmatrix} \tag{10.12}$$

where w_{ij}, $i, j = 1, 2$, is the matrix in ♣A9.8. And we shall write, if $g \equiv g_0$:

$$\begin{aligned} w_{11}(t) =& c'(x^{-1} + x)/2 + C^{11}(x) + g\,\sigma t\,\bar{C}^{11}(x) \\ w_{21}(t) =& c\sigma(x^{-1} - x)/2 + \sigma\,g\,\sigma t\,\bar{C}^{21}(x) \equiv c\,\sigma\,(x^{-1} - x)/2 + w^0_{21}(t) \\ w_{12}(t) =& \sigma\,C^{12}(x), \qquad w_{22}(t) = C^{22}(x) \end{aligned} \tag{10.13}$$

where $\sigma \equiv \sigma_t \equiv \mathrm{sign}\,(\,\mathrm{Re}\,t)$, $x = e^{-g\sigma t}$, and c, c' are constants and the C functions are analytic in x at $x = 0$. The radius of convergence of the series defining $C^{ij}, \bar{C}^{ij}$ is 1, but all the above functions of x can be perfectly continued beyond, as their singularities are poles at $x = \pm i$.

As function of t the $\sigma^{i+j}C^{ij}, \sigma^{i+j}\bar{C}^{ij}$ are holomorphic in t with poles, at most double, at $i(2n+1)\pi/2g_0$, where n is an integer.

We shall consider functions of t which can be represented as:

$$M(t) = \sum_{j=0}^{s} \frac{(\sigma t g)^j}{j!} M^\sigma_j(x, \vec{\omega}t) \ , \quad x \equiv e^{-\sigma g t} \ , \quad \sigma \equiv \mathrm{sign}\,(\,\mathrm{Re}\,t) \tag{10.14}$$

with $s < \infty$, $M^\sigma_j(x, \vec{\psi}\,)$ holomorphic, *at σ fixed equal to + or −*, in the x-plane in a strip $|\,\mathrm{Im}\,x| < 1$ except, possibly, for a polar singularity at $x = 0$. We restrict also M_j to be trigonometric polynomials in the $\vec{\psi}$ variables. We call $\mathcal{M}$ such class of functions. We call $\mathcal{M}_0$ the class obtained by requiring that no $M^\sigma_j(x, \vec{\psi}\,)$ in (10.14) is $\vec{\psi}$ *and* x independent: *i.e.* we "quotient" $\mathcal{M}$ with respect to polynomials in t.

Note that a function $M(t)$ can admit at most one representation like (10.14) with the above mentioned analyticity properties: *i.e.* given $M(t)$ one can compute $M_j(x, \vec{\psi}\,, \sigma)$. In fact, assuming for simplicity that $x \to M_j(x,\vec{\psi})$ are analytic at $x=0$, then:

$$\lim_{T\to\infty} \frac{s!}{(gT)^s}\int_0^T M^\sigma(t)e^{-i\vec{\omega}\cdot\vec{\nu}t}d(gt) = M^\sigma_{s,o\vec{\nu}} \tag{10.15}$$

where we write $M^\sigma_j \equiv \sum_{\vec{\nu},k\ge0} M^\sigma_{j,k\vec{\nu}}x^k e^{i\vec{\psi}\cdot\vec{\nu}}$; to compute $M^\sigma_{s,1\vec{\nu}}$ substitute in the above equation $M^\sigma(t)$ with $(M_\sigma(t)-\sum_{\vec{\nu}} M_{s,o\vec{\nu}}e^{i\vec{\omega}\cdot\vec{\nu}t})e^{\sigma gt}$; and so on; having computed $M^\sigma_s(x,\vec{\psi})$ repeat the procedure with s replaced by $(s-1)$ to $M_{s-1}(t) \equiv\ M^\sigma(t) - \frac{(gt)^s}{s!}M^\sigma_s$.

We define a linear operation $\mathcal{I}$ on the functions $M\in\mathcal{M}_0$ by defining its action on the monomials:

$$M(t) = \frac{(g\sigma t)^h}{h!}x^k\sigma^\vartheta e^{i\rho\vec{\omega}\cdot\vec{\nu}t} \tag{10.16}$$

with h,k integers, $\vartheta=0,1$, $\rho=\pm1$ and $gk\pm i\vec{\omega}\cdot\vec{\nu}\ne0$:

$$\mathcal{I}M(t) = -g^{-1}\sigma^{\vartheta+1}x^k e^{i\rho\vec{\omega}\cdot\vec{\nu}t}\sum_{p=0}^h \frac{(g\sigma t)^{h-p}}{(h-p)!}\frac{1}{(k-i\rho\sigma g^{-1}\vec{\omega}\cdot\vec{\nu})^{p+1}} \tag{10.17}$$

Note that the $\mathcal{I}$ is *not defined* on the polynomials of t,σ, *i.e.* if $k=0$ and $\vec{\omega}\cdot\vec{\nu}=0$ (so that no exponentials are present in the monomial defining M).

The operation $\mathcal{I}$ yields, *at fixed* σ a special primitive of M, in fact:

$$\partial_t\mathcal{I}M \equiv M \tag{10.18}$$

A few further features of $\mathcal{I}$ are the following:

1) if M is odd in t then $\mathcal{I}M$ is even; if M is even then $\mathcal{I}M$ is odd.
2) if M is analytic in t and odd then $\mathcal{I}M$ is analytic and even.
3) if M is analytic in t and even then $\mathcal{I}M(t)$ can be continued analytically from $t>0$ (or $t<0$) to a function $\mathcal{I}^+M(t)$ (or, respectively, to $\mathcal{I}^-M(t)$) defined for all t's and $\mathcal{I}^+M(t)-\mathcal{I}^+M(0)$ is odd (or, respectively, $\mathcal{I}^-M(t)-\mathcal{I}^-M(0)$ is odd). In general $\mathcal{I}^+M(t)\ne\mathcal{I}^-M(t)$ unless $\mathcal{I}^+M(0)=0$ (or $\mathcal{I}^-M(0)=0$). In the latter case $\mathcal{I}^+M(0)=\mathcal{I}^-M(0)=0$.
4) the function $\mathcal{I}_RM(t)\equiv\int_{\sigma\infty}^t e^{-Rg\sigma\tau}M(\tau)d\tau$ is defined for $\mathrm{Re}\,R$ large enough and it admits an analytic continuation to $\mathrm{Re}\,R<0$ and:

$$\mathcal{I}M(t)\equiv\mathcal{I}_0M(t) \tag{10.19}$$

5) if M is such that $M(t)\equiv M(\vec{\omega}t,\sigma)$ for some $M(\vec{\psi},\sigma)$ defined on the torus, then:

$$\mathcal{I}M(t) = (\vec{\omega}\cdot\vec{\partial}_{\vec{\psi}})^{-1}M(\vec{\omega}t,\sigma) \tag{10.20}$$

and $\mathcal{I}M$ is analytic if M is analytic in t and the functions $\mathcal{I}M$ and M have the opposite parity, if M has well defined parity in t.

7) if M depends on other $l-1$ dimensional angles $\vec{\alpha}$ as a linear combination of monomials:

$$\frac{(g\sigma t)^h}{h!} x^k \sigma^\vartheta \cos_{\vartheta'}(\vec{\omega}\cdot\vec{\nu}t + \vec{\alpha}\cdot\vec{\mu}) \equiv \frac{(g\sigma t)^h}{h!} x^k \sigma^\vartheta \frac{(-1)^{[\vartheta'/2]}}{2i^{\vartheta'/2}} \sum_{\rho=\pm 1} \rho^{\vartheta'} e^{i\rho(\vec{\omega}\cdot\vec{\nu}t+\vec{\mu}\cdot\vec{\alpha})} \tag{10.21}$$

with $\vartheta, \vartheta' = 0, 1$ and $\cos_{\vartheta'} y = \cos y$ if $\vartheta' = 0$ and $\cos_{\vartheta'} y = \sin y$ if $\vartheta' = 1$, the $\mathcal{I}M$ has the same form. We shall say that M is *time-angle even* if $\vartheta + \vartheta' =$ even for all monomials of M. If, instead, $\vartheta + \vartheta' =$ odd for all monomials we say that M is *time-angle odd.* It then follows that the time angle parities of M and $\mathcal{I}M$ are opposite (when either is well defined).

8) $\mathcal{I}$ does not change the trigonometric degree of M: *i.e.* if $M_j(t, \vec{\psi}, \sigma)$ had a maximum trigonometric degree N in the $\vec{\psi}$ variables also the functions representing $\mathcal{I}M$ will have trigonometric degree $\le N$. And the operator $\mathcal{I}$ does not increase the degree in t.

9) we extend the operation $\mathcal{I}$ to $\mathcal{M}$ by setting $\mathcal{I}t^n = t^{n+1}/(n+1)$: the above parity properties remain valid. Property 4) holds for $\mathcal{I}_R F(t) - \mathcal{I}_R F(0)$ in general. Property 8) changes as the degree in t of the "non exponential" monomials of the form $(g\sigma t)^h \sigma^\vartheta$ is increased by 1. *However, such a rather arbitrary extension will play no role in the following, as in our context $\mathcal{I}$ is applied only to functions in $\mathcal{M}_0$.*

After the above remarks we make the inductive assumption that $F^{h\sigma}(t, \vec{\alpha})$ has action components $(+, \uparrow)$, denoted symbolically d, of odd time-angle parity in the above sense (different from the one used in §9) and angle components $(-, \downarrow)$, denoted p, of even time angle parity. Opposite parity assumptions will be made for $X^{h\sigma}(t, \vec{\alpha})$. We shall write:

$$F^h = \begin{pmatrix} d \\ p \end{pmatrix}, \quad X^h = \begin{pmatrix} p \\ d \end{pmatrix} \tag{10.22}$$

dropping the label σ from F and X. In fact the main goal of the above formalism is to treat simoultaneously the stable and the unstable whiskers: for $t > 0$ it is $\sigma = 1$ and F^h, X^h represent F^{h+}, X^{h+} while for $t < 0$, $\sigma = -1$ and F^h, X^h represent F^{h-}, X^{h-}. Hence we can symbolically write:

$$\begin{aligned} F^h &= \sum \delta\, x^k\, (g\sigma t)^{k'} \sigma^\vartheta \cos_{\vartheta'}(\vec{\omega}\cdot\vec{\nu}t + \vec{\alpha}\cdot\vec{\mu}) \\ X^h &= \sum \xi\, x^k\, (g\sigma t)^{k'} \sigma^\vartheta \cos_{\vartheta'}(\vec{\omega}\cdot\vec{\nu}t + \vec{\alpha}\cdot\vec{\mu}) \end{aligned} \tag{10.23}$$

with suitable σ-independent coefficients δ, ξ and $\vartheta + \vartheta' =$ even for the $(+, \uparrow)$ components and odd for the $(-, \downarrow)$ components in the case of X, and with reversed parities in the case of F.

Remark: It will be useful to use also complex notation: in this case the $(t, \vec{\alpha})$ parities reflect into $(\sigma, \vec{\nu}, \vec{\mu})$ parities. More precisely if γ is either δ or ξ and if $\lambda \equiv (\vec{\nu}, \vec{\mu})$ we can rewrite the r.h.s. of (10.23) as:

$$\sum i^{\vartheta'} \hat{\gamma}^{\vartheta'}_{\lambda,k,k'} x^k (\sigma g t)^{k'} \sigma^\vartheta e^{i\vec{\nu}\cdot\vec{\omega}t} e^{i\vec{\alpha}\cdot\vec{\mu}}\ , \quad \lambda \equiv (\vec{\nu}, \vec{\mu}) \tag{10.24}$$

where $\hat{\gamma} \in R$ and $\hat{\gamma}^{\vartheta'}_{-\lambda} = (-1)^{\vartheta'} \hat{\gamma}^{\vartheta'}_{\lambda}$; here, as usual, we use the convention the the sum over λ runs over vectors with non–negative first component (so as not to repeat identical terms). ¿From (10.24) it follow immediately the usual parity rules: $p \cdot d = d$ and $p \cdot p = d \cdot d = p$; $p \equiv$ time/angle even, $d \equiv$ time/anlge odd.

Having set the above definitions we deduce immediately from the wronskians structure (10.13), from (6.15) and from the above property of $\mathcal{I}$ that X^h will have the *opposite* structure to F^h (*i.e.* if $F^h = \begin{pmatrix} d \\ p \end{pmatrix}$ then $X^h = \begin{pmatrix} p \\ d \end{pmatrix}$). Explicit expressions of X in terms of the $\mathcal{I}$ operators can be found in appendix A13, see ♣A13.3÷♣A13.5, for the cases J, J_0, g_0, f_ν constant.

The above remark and (6.10) imply that if $X^{h'}$ has the structure $\begin{pmatrix} d \\ p \end{pmatrix}$ for $h' < h$ then F^h has $\begin{pmatrix} d \\ p \end{pmatrix}$ structure. And since it is obvious that F^1 has $\begin{pmatrix} d \\ p \end{pmatrix}$ structure, the (10.22) follows by induction.

To establish a connection between the above remarks and the scattering theory (and to prove theorem 3) we consider the functions $X^\sigma(t, \vec{\alpha}) \equiv X^\sigma(\vec{0}, \vec{\alpha}, t)$, see (6.24), defined in §6 and describing the σ whiskers orbits with initial data at angles $(\pi, \vec{\alpha})$. We can write:

$$\begin{aligned} X^s(t, \vec{\alpha}) =& Z(p(\vec{\alpha}) e^{-gt}, 0, \vec{\psi}^{\,s}_{\vec{\alpha}} + \vec{\omega} t) \\ X^u(t, \vec{\alpha}) =& Z(0, q(\vec{\alpha}) e^{gt}, \vec{\psi}^{\,u}_{\vec{\alpha}} + \vec{\omega} t) \end{aligned} \tag{10.25}$$

for suitable $p(\vec{\alpha}), q(\vec{\alpha})$. We write (10.25), at $t = 0$, for the $\downarrow$ components as:

$$\vec{\alpha} = \vec{\psi}^{\,s}_{\vec{\alpha}} + \vec{\Delta}^s_{\vec{\alpha}}, \qquad \vec{\alpha} = \vec{\psi}^{\,u}_{\vec{\alpha}} + \vec{\Delta}^u_{\vec{\alpha}} \tag{10.26}$$

where $\vec{\psi}^{\,\sigma}_{\vec{\alpha}}$ are the $\vec{\psi}$ coordinates, see lemma 1', of the point on the σ-whisker with angle coordinates $(\pi, \vec{\alpha})$. Then the lemma 1' statement that :

$$\vec{Z}_\downarrow(p, q, \vec{\psi}\,) \equiv \vec{\psi} + \vec{\zeta}_\downarrow(p, q, \vec{\psi}\,) \tag{10.27}$$

with $\vec{\zeta}_\downarrow$ analytic in its arguments (for p, q small) and (10.26) imply that:

$$\vec{\Delta}^s(\vec{\alpha}) = \vec{\zeta}_\downarrow(p(\vec{\alpha}), 0, \vec{\psi}^{\,s}_{\vec{\alpha}}), \qquad \Delta^u(\vec{\alpha}) = \vec{\zeta}_\downarrow(0, q(\vec{\alpha}), \vec{\psi}^{\,u}_{\vec{\alpha}}) \tag{10.28}$$

Hence:

$$\begin{aligned} \vec{X}_\downarrow(t, \vec{\alpha}) =& \vec{\psi}_{\vec{\alpha}} + \vec{\omega} t + \vec{\zeta}_\downarrow(p(\vec{\alpha}) e^{-gt}, 0, \vec{\psi}_{\vec{\alpha}} + \vec{\omega} t) = \\ =& \vec{\alpha} + \vec{\omega} t + \vec{\zeta}_\downarrow(p(\vec{\alpha}) e^{-gt}, 0, \vec{\alpha} + \vec{\Delta}^s(\vec{\alpha}) + \vec{\omega} t) - \vec{\zeta}_\downarrow(p(\vec{\alpha}), 0, \vec{\psi}_{\vec{\alpha}}) \end{aligned} \tag{10.29}$$

and therefore we conclude that:

$$T^{-1} \int_0^T (X^s_\downarrow(t, \vec{\alpha}) - \vec{\alpha} - \vec{\omega} t) dt \xrightarrow[T \to \infty]{} - \vec{\zeta}_\downarrow(p(\vec{\alpha}), 0, \vec{\psi}_{\vec{\alpha}}) + const \tag{10.30}$$

The scattering phase shifts $\vec{\sigma}[\vec{\alpha}] \equiv \vec{\psi}^{\,u}_{\vec{\alpha}} - \vec{\psi}^{\,s}_{\vec{\alpha}} = \vec{\Delta}^s_{\vec{\alpha}} - \vec{\Delta}^u_{\vec{\alpha}}$, by (10.26),(10.28) will be:

$$\vec{\sigma}[\vec{\alpha}] = \vec{\zeta}_{\downarrow}(p(\vec{\alpha}), 0, \vec{\psi}_{\vec{\alpha}}) - \vec{\zeta}_{\downarrow}(0, q(\vec{\alpha}), \vec{\psi}_{\vec{\alpha}}) \tag{10.31}$$

and "all it remains to do" is to find expressions for $\vec{\zeta}_{\downarrow}$ via (10.30). Note that (10.30) and (10.31) are quite general and could be used also for non even hamiltonians. But we keep concentrating on the even case, for simplicity.

To find more concrete expressions we use (10.23): if we set $t = 0, x = 1$ we find in fact $\vec{\zeta}_{\downarrow}(p(\vec{\alpha}), 0, \vec{\psi}^{\,s}_{\vec{\alpha}})$ when $\sigma = +$ and $\vec{\zeta}_{\downarrow}(0, q(\vec{\alpha}), \vec{\psi}^{\,u}_{\vec{\alpha}})$ when $\sigma = -$ from the $\downarrow$ components of X.

Hence we see that only the terms with $k' = 0, \vartheta = 1, \vartheta' = 0$ can contribute to $\vec{\sigma}[\vec{\alpha}]$ as the time-angle parity must be odd (and as the monomials in the expansion of X have no discontinuity at $t = 0$ if $\vartheta = 0$). Therefore:

$$\vec{\sigma}^h[\vec{\alpha}] = 2 \sum \xi^h \left(\cos \vec{\alpha} \cdot \vec{\mu} - 1\right) \tag{10.32}$$

where the -1 has been introduced recalling the convention that $\vec{\sigma}[\vec{0}] = \vec{0}$.

Hence in the even models the odd $\vec{\alpha}$-derivatives of the scattering phase shifts vanish, at $\vec{\alpha} = \vec{0}$ (*i.e.* at the symmetric homoclinic point), More general expansions for the scattering phase shifts and for the splitting are derived in appendix A13.

Although we have always referred to hamiltonians like (9.3) with $J, J_0, g_0, f_{\vec{\nu} m}$ constants, we have only used parity properties which remain unchanged if J is allowed to depend on $\vec{A}$ and if $J_0, g_0, f_{\vec{\nu} m}$ are allowed to depend on $I, \vec{A}, 1 - \cos \varphi$. The only difference will be a more complicated wronskian, still with the even time parity properties for its action-action or angle-angle matrix elements and with odd parity for the action-angle and angle-action elements.. The matrix elements will still have the property of being expressible as power series in $x = e^{-g\sigma t}$ and $\sigma g t$ with σ independent coefficients up to some parity fixing factors σ as in (10.13), and C^h will be expressed in terms of F^h via the $\mathcal{I}$ operations through suitable extensions of the formulae in appendix A13.

This proves a) of the theorem 3 showing the coincidence of the homoclinic angles in the intrinsic coordinates $\vec{\psi}$ and in the natural $\vec{\alpha}$ coordinates. Part b) is a simple corollary of the results in §9, (we allude to its check in Appendix A11).

The proof of part c),d) simply consists in exhibiting an explicit example as we "just" have to show that *in general* the homoclinic splitting is large if $l \geq 3$.

It emerges, from the example, that for the $l = 2$ systems it is possible to think that the homoclinic intersection tensors are all exponentially small because in such cases a special property holds. Namely that an expression like $\sum_i \vec{\nu}_i \cdot \vec{\omega}$ is either 0 or it is necessarily $\geq |\vec{\omega}|$.

This property is no longer true if $l > 2$ and we can obtain slow non zero velocity $\vec{\omega} \cdot \sum_i \vec{\nu}_i$, much smaller than ω_1, even by combining modes $\vec{\nu}_i$ which have fast velocity $\vec{\omega} \cdot \vec{\nu}_i$ (unless of course all the angles rotate at fast velocity).

It will be sufficient to show that, in a model, the second order contribution to the first derivatives of the homoclinic splitting $\vec{Q}_{\uparrow}(\vec{\alpha})$ (with respect to $\vec{\alpha}$) define a matrix (called above the *intersection matrix*) with determinant which is not exponentially small. This means that it is not bounded by an exponential of an inverse power of the parameter η in (10.9), as $\eta \to 0$ at second order (it is easy to see that if α_1 is the fast angle then, to first order, one still has exponentially small splitting at least in the α_1 direction).

The analogous analysis for the scattering phase shifts is essentially identical and in appendix A13 we only derive the expression of the second order phase shift without actually computing it. In fact we shall not really need, in the application analyzed in §12, the homoclinic angles in the intrinsic coordinates and, therefore, we shall not really need the part of theorem 3 concerning the phase shifts.

As a final comment we point out that the inductive check of (10.23) yields a somewhat stronger result if one examines it more carefully. In fact one can check, inductively, that the terms with $k = 0$ have $k' = 0$ as well; and furthermore the terms with $k \geq 1$ have $k' \leq h - 1$ for F^h and $k' \leq h$ for X^h:

$$\begin{aligned}
F^{h\sigma} =& F^{h(\infty)}(\vec{\omega}t,\sigma) + \sum_{p=0}^{h-1} \frac{(g\,\sigma t)^p}{p!} e^{-g\,\sigma t} F^{h(p)}(e^{-g\,\sigma t},\vec{\omega}t,\sigma) \\
X^{h\sigma} =& X^{h(\infty)}(\vec{\omega}t,\sigma) + \sum_{p=0}^{h} \frac{(g\,\sigma t)^p}{p!} e^{-g\,\sigma t} X^{h(p)}(e^{-g\,\sigma t},\vec{\omega}t,\sigma)
\end{aligned} \tag{10.33}$$

and, of course, $\vec{\psi} \to X^{h(\infty)}(\vec{\psi},\pm)$ yield (different) parametrizations of the invariant torus.

§11 Variable coefficients. Fast mode averaging

Let $l = 3$ and consider a hamiltonian $H_\mu \equiv H_\mu(I,\vec{A},\varphi,\vec{\alpha};\eta)$ dependent on a small parameter η having the form:

$$\eta^{-1/2}\omega B + h(\eta^{1/2}A) + \frac{I^2}{2J_0} + J_0 g_0^2(\cos\varphi - 1) + \mu \sum_\nu f_\nu \cos(\vec{\nu}\cdot\vec{\alpha} + n\varphi) \tag{11.1}$$

where $\vec{A} \equiv (B,A)$, $\vec{\alpha} \equiv (\lambda,\alpha)$ are canonically conjugated variables; $\nu \equiv (\vec{\nu},n)$ is an integer vector; and where $J \equiv (h'')^{-1}$, J_0, g_0, f_ν are not constants. In fact we shall allow J_0, g_0, f_ν to be functions of $\eta^{1/2}A, \eta^{1/2}I, \eta(\cos\varphi - 1)$. We shall call $\bar{J}, \bar{J}_0, \bar{g}_0, \bar{f}_\nu$ the values of such functions at zero arguments, assuming them to be $\neq 0$. This implies that when J_0, J, g_0 are replaced by their values at 0 in (11.1) (*i.e.* they are given the barred values) then the theory of §9 applies. It is not restrictive to suppose that J is I independent and that J_0 is φ independent and we shall do so. *Here, the choice of the origin $\eta^{1/2}A = 0$ is arbitrary and in fact, later, we shall replace the center of the rescaled action variables as* *an arbitrary point on the diffusion path* (of course $\eta^{1/2}I = 0$ is instead fixed, being related to the unstable equilibrium of the pendulum).

The φ-dependence of $J_0^{-1}I^2$ can be put together with the $J_0 g_0(\cos\varphi - 1)$ part of the pendulum hamiltonian; and the I dependence of J can be removed by shifting the origin of the I variables by a suitably chosen, $\vec{A}$ dependent, quantity G: this can be achieved (up to corrections in $(1-\cos\varphi)$ that can be included in the g_0) by a canonical transformation generated by:

$$I'\varphi + A'\alpha - G(A'\eta^{1/2})\sin\varphi \tag{11.2}$$

and we can take $G(a) \equiv -\frac{1}{2}\eta^{1/2}a^2 J_0(a,0)\partial_i J^{-1}(a,i)|_{i=0}$. Thus, we shall consider (11.1) with J, J_0 functions of $(\eta^{1/2}A, \eta^{1/2}I)$ and g_0, f_ν functions of $(\eta^{1/2}A, \eta^{1/2}I, \eta(\cos\varphi - 1))$. Such functions will be supposed analytic in their arguments and admitting holomorphic extensions "by ρ_0" in the variables $\eta^{1/2}A, \eta^{1/2}I$ (for I near $I = 0$ and for A near a real interval $\Delta = \eta^{-1/2}[a_1, a_2]$) and "by ξ_0" in the angles. The functions will be supposed to admit upper bounds uniform in η as $\eta \to 0$ (this is only slightly more general than the assumption that they are in fact fixed functions of $\eta^{1/2}I, \eta^{1/2}A$, which is what we really need). In fact we could, in most of what follows, permit a η dependence on the bounds of f proportional to η^{-b} for some $b \ge 0$: but we require boundedness to simplify the formulation of the results, occasionally commenting on extensions of the latter type. Uniformly and positive lower bounds on $|J|$, $|J_0|$, $|g_0|$, $|\partial_A h|$ will be supposed to hold, as well.

The f will be supposed a trigonometric polynomial in the $\vec{\alpha}$ angles with degree $\le N$ for some $N > 0$. In fact if we want I-independence of J we see from (11.2) that f cannot, without loss of generality, be considered a trigonometric polynomial in the φ variables even if the original f in (11.1) was such.

The above hamiltonian (11.1) is taken as an example because of its relevance for the applications of §12: and essentially all the results of §10 extend to the cases of variable coefficients (in the above sense).

The dependence on the action variables through their values scaled by a small parameter $\eta^{1/2}$ is natural. At least if one thinks that (11.1) arises from the change of variables $a = \eta^{1/2}A, i = \eta^{1/2}I, b = \eta^{1/2}B$ accompanied by a multiplication of the hamiltonian by a factor $\eta^{-1/2}$ (which is a transformation leaving the Hamilton equations invariant) followed by a rescaling $\bar{t} = \eta^{-1/2}t$ of the time (which divides the hamiltonian by another factor $\eta^{1/2}$: here t is the time for (11.1) while $\bar{t}$ is the unscaled time for ♣11.3 below)), and starting from a hamiltonian $\bar{H}_\mu$:

$$\bar{\omega}_1 b + \eta h(a) + \frac{i^2}{2J_0} + \eta J_0 g_0^2(\cos\varphi - 1) + \mu\eta\sum_{\vec{\nu},n} f_{\vec{\nu},n}\cos(\vec{\alpha}\cdot\vec{\nu} + n\varphi), \quad \bar{\omega}_1 \equiv \omega \tag{11.3}$$

with h analytic in a (and $h'' \neq 0$); $J_0 \neq 0$ analytic in a, i; $g_0^2 \neq 0$ analytic in $a, i, \eta(\cos\varphi - 1)$; and $f_{\vec{\nu},n}$ analytic in $a, i, \eta(\cos\varphi - 1)$ and vanishing if $|\vec{\nu}| + |n| > N$ for some $N > 0$.

The (11.3) is a natural form in which the hamiltonian appears after the first basic approximations, in many Celestial Mechanics problems, as the three body example discussed in §12.

We call $\vec{\omega}$ an *admissible* velocity vector if $\omega_1 = \eta^{-1/2}\omega$ and $\omega_2 \in \eta^{1/2}[\bar{\omega}, \tilde{\omega}]$ varies in an interval covering the values taken by $\partial_A h(\eta^{1/2}A)$ as A varies in the interval $\Delta = \eta^{-1/2}(a_1, a_2)$ around which H is defined, see above. The set $\Delta(C)$ of the A's such that $|\vec{\omega}\cdot\vec{\nu}|^{-1} < C|\vec{\nu}|^2$ has relative measure $> (1 - K\eta^{-1/2}C^{-1})$, for some suitable $K > 0$ and η small; (such a straightforward bound, obtained, as usual, by summing up over all $\vec{\nu} \neq 0$ the resonant intervals of length $\sim (C\eta|\vec{\nu}|^2\nu_1)^{-1}$, could be improved by taking into account that the centers of the above resonant intervals have to be in Δ: this observation leads to a relative measure $> (1 - K'\eta^{1/2}C^{-1})$ with any $C > \bar{C}\eta^{-1/2}$ for a suitable $\bar{C}$).

We fix $\vec{\omega} = (\omega_1, \omega_2)$ admissible and verifying a diophantine condition, $|\vec{\omega}\cdot\vec{\nu}|^{-1} < C|\vec{\nu}|^2$, for some $C > 0$ and we consider the invariant torus constructed by lemma 1', §5, with rotation velocity $\vec{\omega}$, (when existing).

A precise description of what we have in mind by saying "constructed via lemma 1'" is as follows.

We are in a situation considered already in the corollary to lemma 1' described in §5, see (5.91) and following. The parameters in (5.96) are, of course, η dependent in the present case. And we easily see that ξ_0 (hence $\hat{\xi}_0$), can be taken η independent, while $E_0, \eta_0, \Gamma_0, \rho_0, \tilde{\varepsilon}_0$ can be taken proportional to $\eta^{-1/2}, \eta^{-1}, 1, \eta^{-1/2}, 1$ respectively.

Then (5.96) shows that we shall be able to construct a family of invariant tori with rotation velocities $\vec{\omega}$ and Lyapunov exponent g' with $\omega_1 \equiv \eta^{-1/2}\omega$, with the second angular velocity given by $\omega_2 = \partial_A h(A)$ for A in $\Delta(\mu^{-1/7})$, *i.e.* for a set of a's with relative measure $(1 - K\eta^{-1/2}\mu^{1/7})$ if:

$$|\mu| < \mu^* \equiv B^*\Big[[(\eta^{-1}\eta^{-1/2}\eta^{1/2})^7(\eta^{-1/2})^{14}\eta]^7(\eta^{-1})\Big]^{-1} < B^*\eta^{92} \tag{11.4}$$

where 92 is what comes out of a blind application of the general results of §5 (see (5.96). This constant can be greatly improved by taking into account the special properties of our particular case. Already by using more carefully the estimates of §5, *i.e.* using the weaker (but sufficient) conditions (5.76) and (5.85) with $C = \eta^{-(\delta+1/2)}$ one would obtain a condition like $|\mu| < \bar{B}^*\eta^{-(16+6\delta)}$ with the set $\Delta(\eta^{-(\delta+1/2)})$ having relative measure $> (1 - K\eta^{\delta})$.

However, for simplicity, we shall use (11.4). If f is supposed to be bounded proportionally to η^{-b}, for some $b \geq 0$, instead of being uniformly bounded, the results of the theorem change by suitably increasing 92 to a (linearly b–dependent) new constant.

The above invariant tori for the hamiltonian (11.1) are run quasi periodically with angular velocity $\vec{\omega}$ and Lyapunov exponent g which have the form:

$$\omega_1 = \bar{\omega}_1\eta^{-1/2}, \quad \omega_2 = \bar{\omega}_2\eta^{1/2}, \quad g = g_0(1 + \gamma'(\eta)) \tag{11.5}$$

with $\gamma' \to 0$ as $\eta \to 0$, $\bar{\omega}_2 \in [\bar{\omega}, \tilde{\omega}]$.

Given any admissible $\vec{\omega}$ verifying a diophantine condition $|\vec{\omega} \cdot \vec{\nu}|^{-1} < C|\vec{\nu}|^2$ with $C < |\mu|^{-1/7}$), there will be an invariant torus run quasi periodically with angular velocities $\vec{\omega}$, if η is small enough and if μ verifies (11.4).

In what follows we suppose that η is small enough, that μ verifies (11.4) and *study one of the above invariant tori*, with prefixed angular velocities and Lyapunov exponents given by $\vec{\omega}, g$ like in (11.5). And we want to estimate the whiskers splitting at the symmetric homoclinic point $\varphi = \pi, \vec{\alpha} = \vec{0}$.

Remark: It is convenient to fix $a \in \overline{\Delta} \equiv \eta^{1/2}\Delta$ as the origin of the unscaled variables so that $\omega_2 \equiv \partial_a h(a)$ (obviously the condition on μ will not be affected by such a choice and we shall assume that the associated invariant torus is persistent). *Hence, the values of $\bar{J}, \bar{J}_0, \bar{g}_0, \bar{f}_\nu$ are now the values of the corresponding functions evaluated at $\eta^{1/2}A = a$, $\eta^{1/2}I = 0$.*

Consider first the case $\mu = 0$. In this case one can perform an elementary discussion of the separatrix quadratures. For instance if $\mu = 0$ and the pendulum and rotators are independent in the sense of §4 (*i.e.* J_0, g_0 are A independent) one finds:

$$\dot{\varphi} = \frac{I}{J_0} - \frac{I^2}{2J_0}\frac{\partial_I J_0}{J_0} - 2\partial_I(J_0 g_0^2)\sin^2\frac{\varphi}{2}, \qquad I = \pm 2 J_0 g_0 \sin\frac{\varphi}{2} \tag{11.6}$$

One can check that the pendulum wronskian matrix elements verify, for η small:

$$|w_{ij}| \le \bar{u}\eta^{-1}, \qquad \text{if} \quad |\operatorname{Im} t| \le \frac{\pi}{2\bar{g}_0}(1 - \kappa\eta^{1/2}), \quad |\operatorname{Re} t| \le \bar{g}_0^{-1} \tag{11.7}$$

Here the constants $\kappa, \bar{u}$ have to be taken large enough, depending on the functions J, J_0, g_0. Furthermore the functions $w_{ij}(t)$ admit expansions (10.13): this is now true also for the components that were zero in the cases considered in §9,§10. In appendix A9, part 4, we have studied the full wronskian in a rather general pendulum system. Also the parity properties in t of the full wronskian are the same (*i.e.* even in the action-action or angle-angle blocks and odd in the other, mixed, blocks). The above "large" domain bounds (11.7), are useful in the fast rotation cases (*i.e.* when both frequencies have size of $O(\eta^{-1/2})$, or $\bar{\omega}_2 = O(\eta^{-1})$), discussed in §10 but they will not be really necessary in what follows (boundedness in a finite strip being sufficient).

The parity properties and the analyticity together with the bounds (11.7) are the only ingredients necessary to perform the analysis of §9,§10 as we have repeatedly claimed and as it is easy to check.

Therefore the same conclusions about the homoclinic angles at the symmetric homocline $\vec{\alpha} = \vec{0}$ hold. In particular we consider the intersection matrix to the lowest non trivial order (*i.e.* to the lowest order that makes its determinant not exponentially small as $\eta \to 0$, (namely the second)). By the final result of appendix A13, see ♣A13.22,♣A13.23:

$$M = \begin{pmatrix} 0 & \delta \\ \delta & \bar{\gamma}\mu \end{pmatrix}, \qquad \delta \equiv \mu^2 \sum_{\substack{\vec{\mu}\,fast \\ \vec{\nu}\,slow}} nm\mu_1\nu_2 K_n (-1)^m \frac{\bar{f}_{\vec{\nu},n}\bar{f}_{\vec{\mu},m}}{\bar{J}_0\bar{g}_0^2}\frac{\vec{\omega}\cdot\vec{\nu}}{(\vec{\omega}\cdot\vec{\mu})^2} \tag{11.8}$$

is the leading part of the intersection matrix $\partial_{\vec\alpha}\vec Q_\uparrow(\vec 0)$ (as $\eta\to 0$ and at second order in μ); see §10, (10.7), for the definition of Q; δ denotes here the $\delta_{12}=\delta_{21}$ of appendix A13; $\vec\omega$ is defined in (11.5), $\bar\gamma$ is a constant at fixed $\bar\omega_1,\bar\omega_2$ and the first matrix element is exponentially small (to second order); K_n,Γ_{mn} are defined by the integrals in ♣A13.21,♣A13.23; and, finally, see the remark after (11.5) for the values of $\bar f_\nu,\bar J_0,\bar g_0$.

The (11.8), therefore, shows that generically $\delta=O(\mu^2\eta^{3/2})$, and hence if $\mu<\eta^{1/2}$ (consequence of (11.5)) this leads over the terms of order μ^3 and higher and the splitting is not exponentially small as $\eta\to 0$, but of the order of det $M=-\delta^2$.

The δ in (11.8) has to be multiplied by $\eta^{1/2}$ if one wants to regard it as the intersection tensor for (11.3), (the reason being that with our definitions the "homoclinic angles" have the dimensions of an action and scale as such upon coordinates rescalings).

A more formal statement of the above conclusions is:

Lemma 4: *Consider the hamiltonian (11.1) near the segment Δ where $\partial_A h$ varies, as $A\in\Delta$, in an interval $\eta^{1/2}[\bar\omega,\tilde\omega]$ with $\bar\omega>0$. There is $c>0$ such that if $\mu=\eta^c$ then (11.1) admits invariant tori which, if η is small enough, have whiskers with a "homoclinic splitting" ($\equiv$ determinant of the above intersection matrix $\partial_{\vec\alpha}\vec Q_\uparrow(\vec 0)$) of $O((\eta^{3/2}\mu^2)^2)$ as $\eta\to 0$, provided the sum in (11.8) does not vanish accidentally.*

As a second important extension of the results of §9,§10 we consider a situation also met in some applications, (*e.g.* see §12). Namely a hamiltonian obtained by adding to H_μ in (11.1) a further perturbation:

$$F=\sum_{0<\nu<N}F_\nu\cos\nu\,\vec\nu_0\cdot\vec\alpha \tag{11.9}$$

where F is analytic in $\eta^{1/2}A,\eta^{1/2}I$ and in $(\cos\varphi-1)$ *but it is not small.*

We suppose that F, which by assumption contains only harmonics that are *multiples of a given mode* $\vec\nu_0$, depends on the fast variable $\lambda\equiv\alpha_1$: *i.e.* $\nu_{01}\neq 0$. We say that F is *unimodal* on a fast mode, with mode $\vec\nu_0$.

Hence the hamiltonian that we consider is:

$$H\equiv H_{\mu,F}=H_\mu+F,\qquad\text{or}\qquad \bar H\equiv\bar H_{\mu,F}=\bar H_\mu+\eta\bar F \tag{11.10}$$

depending on whether we regard it as a function of A,B,I or of $a=\eta^{1/2}A, b=\eta^{1/2}B, i=\eta^{1/2}I$.

We shall refer to the above two equivalent representations of the same mechanical problem as the *scaled representation* (H) and as the *unscaled* or *natural* representation ($\bar H$).

In the case (11.10) we cannot apply directly the results of §9,§10. But one can remark that the angle $\vec\alpha\cdot\vec\nu_0$ is a "fast angle", *i.e.* it rotates (if unperturbed) at speed $\eta^{-1/2}\omega_1$ compared to the speed $\omega_2\eta^{1/2}$ of the "slow mode" $\alpha=\alpha_2$.

The idea of the *averaging method* is just the remark that *quickly oscillating* perturbing forces of order 1 can, in fact, for many purposes, be treated as small. However the method does not consist in the brutal setting of $F=0$, so familiar in heuristic treatments, but rather consists in treating F as a perturbation by putting a formal parameter β in front of the total perturbing terms $\beta(F+\mu f)$, and by taking as many orders in β as it might be necessary to match the precision required, when $\beta=1$ (and eventually setting $\beta=1$).

We consider the (11.1) perturbed by (11.9) *i.e.* we consider (11.10):

$$H=\eta^{-1/2}\omega B+\eta^{1/2}\bar\omega_2 A+\eta\frac{A^2}{2J}+\frac{I^2}{2J_0}+J_0 g_0^2(\cos\varphi-1)+\beta(F+\mu f) \tag{11.11}$$

where we have expanded $h(A)=\eta^{1/2}\bar\omega_2 A+\eta\frac{A^2}{2J}$ for a J analytic in $\eta^{1/2}A$ ($J(0)\neq 0$) and with the functions J, J_0, g_0, F, f having the analyticity and boundedness properties described above around a line:

$$\mathcal{L}=\{\eta^{-1/2}[a_1,a_2]\}\times\{I=0\}\equiv\Delta\times\{I=0\} \tag{11.12}$$

In addition F, f will be supposed to be trigonometric polynomials of degree $\le N$ in (α,φ), allowing also an analytic dependence of the Fourier coefficients on the variable $z\equiv\eta(1-\cos\varphi)$.

In the present case we show that the method allows one to establish the existence of the invariant tori and to compute the leading order expressions of the homoclinic angles, as $\eta\to 0$, $|\mu|=\eta^c$, with $c>0$ large enough.

Lemma 5: *Fix $x>0$ and $0<\sigma<1/2$. There exist constants η_0, $B^*>0$ such that if $|\mu|<\eta^c$, $c>10$, and $|\beta|<B^*\,\eta^{-\sigma}$ then one can construct a holomorphic canonical map casting the hamiltonian H (11.11), for all $0<\eta<\eta_0$, in a form:*

$$\frac{\omega B}{\eta^{1/2}}+\eta^{1/2}\bar\omega_2 A+\frac{\eta A^2}{2\hat J(\eta^{1/2}A)}+g(\eta^{1/2}A,pq)pq+\eta^x\hat f(p,q,A,\lambda,\alpha) \tag{11.13}$$

with $\hat J, g, \hat f$ (depending also on β) bounded and holomorphic in the complex domain:

$$\big\{|p|,|q|<\kappa,\ |A|<\rho\eta^{-1/2},\ |\operatorname{Im}\alpha_j|<\xi,\ |\mu|<\eta^c\big\}\ \times\ \{|\beta|<B^*\eta^{-\sigma}\} \tag{11.14}$$

for suitable η–independent $\kappa,\rho,\xi>0$. The smallness condition on η_0 can be taken to be $\eta_0^{1/2}\log\eta_0^{-1}<Dx$ for a suitable $D>0$.

The reason for the validity of (11.13) is simply that in $H_{0,F}$, see (11.10), no strong resonances occur with $|\vec\nu|\le O(\eta^{-1})$, all denominators being bounded below by $O(\eta^{1/2})$. Hence we are in an essentially better situation compared to that in §7, as we can proceed to perturbation theory of much larger order, essentially $O(\eta^{-1/2})$, after taking advantage in the first step of *large* denominators (of order $O(\eta^{-1/2})$) to reduce the size of F.

The method used to deduce (11.13) is the usual method developed in the Nekhorossev resonance theory. Note, however, that in (11.13) the angle $\vec\alpha$ are in the remainder term (while, perhaps, one would expect them to remain of order μ): the mechanism for this is essentially the same as the one used in §7.

The estimates leading to (11.13) are carried out in detail in appendix A10, using the scaled variables form $H_{\mu,F}$ of the hamiltonian.

Once the hamiltonian has been put in the above form we are in a situation in which the theory of §5 becomes applicable, at least if c and x are chosen large enough. Assuming that the value $c = 92$ (*i.e.* the value dictated by the "blind" bound discussed above: see (11.4)) is also large enough for Lemma 4 to hold, and taking $|\mu| = \eta^c$ we see that the methods of §2 ÷ §8 are applicable.

In fact we see that for a set of $A \in \Sigma_\eta \subset \Delta$ of relative measure $\le K\eta^{-1/2}\eta^{x/7}$, with K being a suitable constant, it is (if $\omega_2 = \partial_A h(A) \;\equiv\; \eta^{1/2}\bar\omega_2 + \partial_A[\eta A^2/(2J(A)]$; see also the definition of "admissible" after (11.3)):

$$|\eta^{-1/2}\bar\omega_1\nu_1 + \omega_2\nu_2|^{-1} \le \bar g_0^{-1}|\vec\nu|^2\eta^{-x/7} \tag{11.15}$$

and for each such A there is an invariant whiskered torus run quasi periodically with angular velocities $\vec\omega = (\eta^{-1/2}\bar\omega_1, \omega_2)$. Thus if Lemma 4 holds and if $x > 2597$ (so that the round spacing $\sim O(\eta^{x/7})$ is larger than the "homoclinic splitting" $\sim O((\eta^{3/2}\mu^2)^2))$ we see that drift and diffusion take place along $\mathcal{L}$.

Remark 1): Thus we see that along the line $\mathcal{L}$ there is a whiskers ladder with very small rounds spacing, as $\eta \to 0$: *i.e.* of order $O(\eta^{x/7})$. We see that this is so *in spite* of the presence of F (which is of $O(1)$) and of μf (which is smaller than F but still very large, of $O(\eta^c) = O(\mu)$, compared to the spacing in the ladder).

Remark 2): The holomorphy and uniform boundedness in β is very important: it allows us to conclude that the tori equations as well as those of their whiskers can be computed as power series in β. And since $\beta = 1$ is inside the radius of convergence $(B^*\eta^{-\sigma})$ we get immediately that the various orders in β, (*note the distinction between orders in β and orders in μ or in η*), give contributions to the whiskers parametric equations or to the size of the homoclinic angles whose size decreases with the order k in β as $\eta^{\sigma k}$ at least. Hence if to some order some contribution has a size of some power of η *it becomes a matter of a calculation to finite order to check if it is the dominant contribution to the quantity being calculated.* See remark 5) below.

Remark 3): All the above invariant tori will have whiskers homoclinic at $\varphi = \pi, \vec\alpha = \vec 0$ because all the above hamiltonians are even in the sense of §9 (it is easy to see, although not necessary, that all the canonical changes of coordinates that we use (in appendix A10) to perform the perturbation theory construction of $\hat f$ do not change the even nature of the hamiltonians). Hence it makes sense to ask about their homoclinic angles or tensors.

Remark 4): The above analysis shows that if we introduce an artificial parameter β that we put in front of both ηF and $\eta\mu f$, then then for η small, we can compute the whiskers for

$\beta = 1$ in power series of β. In other words we can apply perturbation theory to compute the intersections tensors. We have to push perturbation theory up to an order n (in β) such that the exactly computed terms are larger than the remainder (which is of order $\eta^{-n/2}$). In concrete cases this might mean just the second order (never the first as we have seen that to first order the intersection tensors are degenerate).

Remark 5): And drift or diffusion will follow along the line $\mathcal{L}$ by the theory of §8, for most choices of β around $\beta = 1$ (and possibly β *exactly* equal to 1), provided there is an order at which one sees that the homoclinic intersection tensor is not exponentially small. Because in this case the rounds spacing in the ladder of whiskers is, by the averaging phenomenon, always faster than any power (being $O(\eta^{x/7})$ for a prefixed x, if η is small enough) and the splitting cannot be exponentially small unless there is a cancellation between finitely many orders in β on a segment of order $O(1)$ of $\mathcal{L}$. By the analyticity in β, this can only be for exceptional values of β. Of course in a given problem one has to exclude that $\beta = 1$ is not an exceptional value (unless β happens to be a natural parameter in the problem and one is just interested in showing existence of drift or diffusion for some values of β). The check of the latter property is reduced in general, by the above analysis, to a finite order calculation which, in concrete cases, could be conceivably performed with the help of an electronic device.

The analysis is thus concluded and one can try to apply it to some concrete problem. This is better than trying to continue proceeding in general because in this way we can avoid formulating too abstract results, and apparently unphysical hypotheses on the perturbations.

§12 Planetary precession. Existence of drift and diffusion

Imagine a planet $\mathcal{E}$ as a homogeneous rigid body with cylindrical symmetry. The body surface will be described in polar coordinates by $\rho = Rh(\cos\vartheta)$ for some R and some h, $R > 0$, $0 < h \le 1$, *e.g.* for a rotation ellipsoid with equatorial radius R and polar radius $R/(1+2\eta)^{1/2}$ it is $h(z) = (1+2\eta z^2)^{-1/2}$.

We suppose the planet center T to revolve on a keplerian orbit $t \to \vec{r}_T(t)$: the orbit plane will be called the *ecliptic* plane and $\vec{\vec{k}}$ will denote its unit normal vector which sees the planet rotating counterclockwise.

The longitude λ_T of $\vec{r}_T$ on the ecliptic will be reckoned from the major semiaxis of the ellipse; hence $\lambda_T = 0$ is the *aphelion* position *i.e.* when $r_T \equiv |\vec{r}_T|$ is maximal: $r_T(0) = a(1+e)$, a being the major semiaxis of the Keplerian ellipse and e its eccentricity.

With these conventions, r_T and λ_T are related by the *focal equation* (see, e.g. [G] p.304):

$$r_T \equiv |\vec{r}_T| = \frac{p}{1 - e\cos\lambda_T}, \quad p \equiv a(1-e^2). \tag{12.1}$$

In this section we shall always denote by e the eccentricity of the orbit and to avoid confusion with the Neper constant we denote the exponential of a number α by $\exp\alpha$*, while* e^α *will denote everywhere the* α*-th power of the eccentricity* e*.*

Kepler's law, $\dot\lambda_T r_T^2 = const$, and (12.1) imply that if λ is the keplerian *average anomaly*:

$$\lambda \equiv (1-e^2)^{3/2}\int_0^{\lambda_T}\frac{d\beta}{(1-e\cos\beta)^2} = \lambda_T + 2e\sin\lambda_T + \frac{3}{4}e^2\sin 2\lambda_T + \ldots, \tag{12.2}$$

then:

$$\lambda_T = \lambda - 2e\sin\lambda + (5/4)e^2\sin 2\lambda + \ldots, \qquad \frac{a}{r_T} = 1 - e\cos\lambda + e^2\cos 2\lambda + \ldots \tag{12.3}$$

and the motion is $\lambda \to \lambda + \omega_T t$, where $2\pi/\omega_T = 2\pi a^{3/2} g_N^{-1/2}$ is the year of the planet, $g_N \equiv k(m_S + m_T)$ if k is Newton's constant and m_T, m_S are the masses of the planet and of its star.

The unit vector $\vec{\bar{\imath}}$ pointing from the focus towards the aphelion will be used together with $\vec{\bar{k}}$ and a third vector $\vec{\bar{\jmath}}$ to form an orthonormal triad $(\vec{\bar{\imath}}, \vec{\bar{\jmath}}, \vec{\bar{k}})$ of fixed directions in space.

A comoving frame $(T;\vec\imath_1,\vec\imath_2,\vec\imath_3)$ will be attached to the planet with $\vec\imath_3$ axis coinciding with the symmetry axis (*polar axis*) of the planet and $\vec\imath_1$ is arbitrarily chosen on the *equatorial plane*, (*i.e.* the plane orthogonal to $\vec\imath_3$).

The position of $(T;\vec\imath_1,\vec\imath_2,\vec\imath_3)$ referred to $(T;\vec{\bar{\imath}},\vec{\bar{\jmath}},\vec{\bar{k}})$ will be determined by the three Euler angles $\bar\vartheta,\bar\varphi,\bar\psi$ with $\bar\vartheta$ being the angle between $\vec{\bar{k}}$ and $\vec\imath_3$, $\bar\varphi$ being the angle on the ecliptic between $\vec{\bar{\imath}}$ and the ecliptic – equator node $\vec{\bar{n}}$, while $\bar\psi$ is the angle on the equator between $\vec{\bar{n}}$ and $\vec\imath_1$, (drawings with the above and the following notations can be found in [G, p.318 $\div$ 321]).

In the coordinates $(\bar\vartheta,\bar\varphi,\bar\psi)$ the motion of the planet $\mathcal{E}$ is described by the Euler-Lagrange equation associated to the lagrangian:

$$\mathcal{L} \equiv \frac{1}{2}J_3(\dot{\bar\varphi}\cos\bar\vartheta + \dot{\bar\psi})^2 + \frac{1}{2}J_1(\dot{\bar\vartheta}^2 + \dot{\bar\varphi}^2\sin^2\bar\vartheta) + \int_{\mathcal{E}}\frac{k m_T m_S}{|\vec r_T + \vec x|}\frac{d\vec x}{|\mathcal{E}|} \tag{12.4}$$

where $J_3, J_1 \equiv J_2$ are the inertia moments of $\mathcal{E}$, m_T its mass, $|\mathcal{E}|$ its volume, m_S is the mass of the heavenly body keeping the planet $\mathcal{E}$ on its celestial path, $t \to \vec r_T(t)$, and k is Newton's constant.

Very remarkable is a theorem by Andoyer-Deprit, see [G, p.318 $\div$ 321], which produces canonically conjugate variables casting the Hamiltonian corresponding to $\mathcal{L}$ in a simple form. To describe such variables we consider the unit vector $\vec k$ parallel to the angular momentum $\vec K_T \equiv M\vec k$, $M = |\vec K_T|$ and call *angular momentum plane* the plane orthogonal to $\vec k$. We define the angle δ and ϑ between $\vec{\bar{k}}$ and $\vec k$ and, respectively, $\vec k$ and $\vec\imath_3$, so that the components of the angular momentum on $\vec{\bar{k}}$ and on $\vec\imath_3$ will be, respectively:

$$K = M\cos\delta, \qquad L = M\cos\vartheta \tag{12.5}$$

We also associate with $\vec{K}_T$ two more remarkable angles: in fact the angular momentum plane has a node $\vec{m}$ on the ecliptic plane and one $\vec{n}$ on the equator plane. We call γ the angle on the ecliptic between $\vec{m}$ and $\vec{\imath}$ and φ the angle on the angular momentum plane between the node $\vec{m}$ and the node $\vec{n}$. Finally we let ψ denote the angle between $\vec{\bar{n}}$ and $\vec{\imath}_1$.

Deprit's theorem states that the variables $(K,\gamma),(M,\varphi),(L,\psi)$ are canonically conjugate for the hamiltonian H associated to $\mathcal{L}$, and that H in such variables takes the form:

$$H=\frac{M^2}{2J_3}+\frac{J_3-J_1}{2J_1J_3}(M^2-L^2)+\omega_T B+V \tag{12.6}$$

where V is the integral in (12.4) changed in sign, and (B,λ) is a fourth pair of canonical coordinates with λ being the average anomaly of the planet in its revolution about the ellipse focus, see (12.2). The pair (B,λ) has been introduced in order to eliminate the explicit time dependence from the hamiltonian.

It is convenient to bear in mind that $\omega_T B$ has a simple physical interpretation: it is the energy stored in the device providing the external force that keeps the heavenly body $\mathcal{E}$ on its keplerian celestial path, $t\to\vec{r}_T(t)$.

By symmetry considerations it is clear that V is a function of the angle λ (or of λ_T), and of the angle α between the position vector $\vec{r}_T$ and the axis $\vec{\imath}_3$ of the planet. In fact, it is easy to find out an expression for $V=V(\alpha,\lambda)$. Recalling the relation between the Legendre polynomials $P_l(z)$ and their generating function $(1+x^2-2xz)^{-1/2}$, one finds:

$$\begin{aligned}V=&\frac{-km_Sm_T}{|\vec{r}_T|}\int\frac{d\vec{x}}{|\mathcal{E}|}\left(1+(\frac{|\vec{x}|}{r_T})^2+2\frac{|\vec{x}|}{r_T}(\frac{\vec{x}}{|\vec{x}|}\cdot\frac{\vec{r}_T}{r_T})\right)^{-1/2}=\\ =&\frac{-km_Sm_T}{|\vec{r}_T|}\sum_{l=0}^{\infty}\int_{\mathcal{E}}\frac{d\vec{x}}{|\mathcal{E}|}(\frac{-|\vec{x}|}{r_T})^lP_l(\frac{\vec{x}}{|\vec{x}|}\cdot\frac{\vec{r}_T}{r_T})\end{aligned} \tag{12.7}$$

and the above expression can be used to compute the series expansion of the potential energy in the eccentricity.

If we perform the calculation neglecting the terms in (12.7) which come from $l\geq 4$ (which roughly means neglecting $O(R/a)^2$ compared to 1, with R being the planet radius and a being the major semiaxis of its orbit, because the odd orders in l vanish by symmetry), it is well known, (see [L]), that the only properties of the rigid body that matter are the inertia moments. It is also clear that the hamiltonian must be expressible in terms of the physical quantities that establish the orders of magnitude of the problem. Thus we expect the hamiltonian to be a function depending, besides on the angles and their conjugate moments, on the daily rotation of the planet ω_D, on the yearly rotation ω_T and on the inertia moments J_i. The physical periods are introduced into the problem through the initial data, which we denote $\bar{K}$, $\bar{M}$, $\bar{L}$, around which we want to set up a perturbation theory. Denoting $\eta\equiv(J_3-J_1)/J_3$, $\eta'\equiv(J_3-J_1)/J_1$, $\omega_D\equiv\bar{M}/J_1$, $\cos i_0\equiv\bar{K}\bar{M}^{-1}$ (the cosine of the planet inclination i_0 over the ecliptic), and:

$$\omega=\omega_T^2\omega_D^{-1}\cos i_0 \tag{12.8}$$

it is a classical calculation to check that the exact form of the hamiltonian is, to order k in the eccentricity e and denoting $[\cdot]^{[\le k]}$ the truncation to power k of a series in e:

$$H=\frac{M^2}{2J_1}-\eta'\frac{L^2}{2J_3}+\omega_T B+\eta\omega\frac{\bar M^2}{\bar K}\big[\frac{(1-e\cos\lambda_T)^3}{(1-e^2)^3}\cos^2\alpha\big]^{[\le k]} \tag{12.9}$$

The model thus obtained will be called the *D' Alembert precession-nutation* model. The reason for the above definition [especially (12.8)] is that $-\eta\omega$ has the physical meaning of the average angular velocity of *precession of the equinoxes*, as it appears also from the following analysis: for more details see Appendices 6,7 where we discuss this celebrated result of D'Alembert using canonical formalism.

Concerning the approximations involved in passing from (12.6) to (12.9) we note that the terms of $O(\eta(R/a)^4)$ are believed to be really negligible for all practical purposes in many astronomy problems while, for the truncation approximation, D' Alembert did not have data on the Moon mass accurate enough to wish to consider orders $k>0$ in his theory of lunisolar precession. Here we consider only the case $k=2$: but it is clear that what follows does not really require neither the truncation nor neglecting the higher orders in $\eta(R/a)^2$. Considering such more general problems would only lead to some (minor) modifications, except in the case $k=0$, where the result is simply false (*i.e.* no drift or diffusion can take place) and the case $k=1$ which cannot be decided by a "lowest order" perturbation theory as, instead, the cases $k\ge 2$ are, (at least if the initial data are chosen as we are going to do).

To compute the D' Alembert hamiltonian (12.9) we have, of course, to find how $\cos^2\alpha$ depends on the canonically conjugated variables $(K,M,L,B,\gamma,\varphi,\psi,\lambda)$.

Simple spherical trigonometry arguments, see appendix A8, lead to:

$$\begin{aligned}\cos\alpha=&\sin(\lambda_T-\gamma)\,\big(\cos\varphi\sin\vartheta\cos\delta+\sin\delta\cos\vartheta\big)-\cos(\lambda_T-\gamma)\sin\vartheta\sin\varphi=\\ =&\sin(\lambda_T-\gamma)\,\Big((K/M)\,\big(1-(L/M)^2\big)^{1/2}\cos\varphi+(L/M)\,\big(1-(K/M)^2\big)^{1/2}\Big)-\\ &-\big(1-(L/M)^2\big)^{1/2}\sin\varphi\cos(\lambda_T-\gamma)\equiv s(\kappa\nu c_\varphi+\mu\sigma)-\nu s_\varphi c\end{aligned} \tag{12.10}$$

where:

$$\begin{array}{llll}\mu\equiv L/M, & \nu^2\equiv 1-\mu^2, & s\equiv\sin(\lambda_T-\gamma), & s_\varphi\equiv\sin\varphi,\\ \kappa\equiv K/M, & \sigma^2\equiv 1-\kappa^2, & c\equiv\cos(\lambda_T-\gamma), & c_\varphi\equiv\cos\varphi.\end{array} \tag{12.11}$$

Hence we see that (12.9), as well as the full (12.7), does not contain ψ. Therefore L is a constant of motion and it will be regarded as a parameter. It has the physical interpretation that $\bar\nu=(1-L^2/\bar M^2)^{1/2}$ is the angle between the spin axis and the symmetry axis and in the theory of nutation it is called the *eulerian nutation constant*, at the initial *epoch*, *i.e.* at a prefixed reference time, when $\bar M,\bar L,\bar K,\bar\varphi,\bar\psi,\bar\gamma$ are the values of the canonical variables.

Therefore setting:

$$V\equiv\big[\frac{(1-e\cos\lambda_T)^3}{(1-e^2)^3}\cos^2\alpha\big]^{[\le 2]}\equiv V_0+eV_1+e^2V_2 \tag{12.12}$$

and $\gamma_0 \equiv \gamma - \lambda$, $\chi_0 \equiv \varphi, \lambda_0 \equiv \lambda$ and using:

$$\frac{(1 - e\cos\lambda_T)^3}{(1-e^2)^3} = 1 + \frac{3}{2}e^2 - 3e\cos\lambda + \frac{9}{2}e^2\cos 2\lambda + \dots \tag{12.13}$$

one finds that:

$$V = \sum_{h=0}^{2} e^h \sum_{\substack{r,p,j \\ r,\, p+h=even}} \bar{B}^h_{rpj} \cos(r\gamma_0 + p\lambda_0 + j\chi_0) \tag{12.14}$$

where $\bar{B}^h_{rpj}$ are suitable coefficients depending on M, K, listed in appendix A14. For instance:

$$\bar{B}^0_{000} \equiv c_0 \equiv \frac{1}{4}[2\sigma^2\mu^2 + (1+\kappa^2)\nu^2] \;, \qquad \bar{B}^0_{201} \equiv d_1 \equiv \; -\frac{(1+\kappa)\mu\sigma}{2}\nu \tag{12.15}$$

Thus, setting $\bar{E} \equiv \omega\frac{\bar{M}^2}{\bar{K}}$ and dropping from (12.9) the additive constant $\eta' L^2/(2J_3)$, the full ("order 2") D' Alembert hamiltonian, in the canonical variables $(\gamma_0, K_0) \equiv (\gamma-\lambda, K)$, $(\chi_0, M_0) \equiv (\varphi, M)$, $(\lambda_0, B_0) \equiv (\lambda, K + B_T)$, takes the form:

$$\omega_T B_0 + h_0(K_0, M_0; \eta) + \eta f_0(K_0, M_0, \gamma_0, \chi_0, \lambda_0; e) \;, \tag{12.16}$$

where:

$$\begin{aligned} h_0 &\equiv \; -\omega_T K_0 + \frac{M_0^2}{2J_1} + \eta\bar{E}\; c_0(K_0, M_0) \\ f_0 &\equiv \bar{E}\; [V_0 - c_0 + eV_1 + e^2 V_2], \qquad \langle f_0\rangle = \frac{3}{2}e^2 c_0 \;, \end{aligned} \tag{12.17}$$

with $V_i = V_i(K_0, M_0, 2\gamma_0, \chi_0, \lambda_0)$, $c_0 = c_0(K_0, M_0)$, and $\langle\cdot\rangle$ denotes average over the angles.

Note that, taking into account the coefficients calculated in appendix A14 and neglecting terms of $O(\nu^2)$, *the integrable part of the hamiltonian becomes:* $c_0 = \sigma^2/2 = (1 - K_0^2/M_0^2)/2$ *leading to the standard "D' Alembert equinox precession"* $\dot{\gamma} = -\eta\omega$, see (12.8) and appendix A6,A7.

To analyze the motions of the D' Alembert hamiltonian we shall consider particular regions of phase space that we single out also for convenience and for the sake of concreteness. Fix $\bar{M}, \bar{K} > 0$, such that $\bar{K}/\bar{M} \le 1/8$; the hamiltonian (12.16) will be studied in the vicinity of the set:

$$\Upsilon_0 = \{M_0 = \bar{M},\; |K_0 - \bar{K}| \le \frac{\bar{M}}{8}\; \} \tag{12.18}$$

where $\bar{M}, \bar{K}$ are values around which drift or diffusion will take place; the role of B_0 is trivial: see, however, end of §5 [especially (5.95)] and the final remark of §8. ¿From now on $(\mu, \nu, \kappa, \sigma)$ *denote the functions in (12.11) evaluated at* $(K, M) = (K_0, M_0)$. The condition $|\bar{K}/\bar{M}| \le 1/8$ implies that, on Υ_0, $|\kappa| \le 1/4$, so that σ is well defined and $15/16 \le \sigma^2 \le 1$.

The physical meaning of a drift along Υ_0 is a variation of K at M, L fixed: hence it represents a change in the inclination of the spin axis, see §1.

Furthermore, the hamiltonian (12.16) is holomorphic in a complex neighborhood of Υ_0. To be more precise, let $\bar{L} \equiv L$ be such that $15/16 < \bar{\mu} < 1$, $(\bar{\mu} \equiv \bar{L}/\bar{M})$, let $\rho_0 > 0$ and define:

$$\Upsilon_{0,\rho_0} \equiv \bigcup_{(K_0', M_0') \in \Upsilon_0} \{(K_0, M_0) \in C^2 \ : \ |K_0 - K_0'| < \rho_0 \ , \ |M_0 - M_0'| < \rho_0\} \ . \qquad (12.19)$$

Then, we can choose $\eta_0, e_0 < 1/4$, $\rho_0 > 0$ and an arbitrary ξ_0 such that (12.16) is holomorphic on:

$$\{B_0 \in C\} \times \Upsilon_{0,\rho_0} \times \{(\gamma_0, \chi_0, \lambda_0) \in S^3_{\xi_0}\} \times \{|\eta| < \eta_0\} \times \{|e| < e_0\} \ , \qquad (12.20)$$

with $S^n_\xi \equiv \{\vec{\alpha} \in C^n : \exp(-\xi) < |\exp i\alpha_j| < \exp \xi\}$ being the standard complex neighborhood of T^n, and so that, for $(K_0, M_0) \in \Upsilon_{0,\rho_0}$, one has:

$$|\frac{M_0}{\bar{M}} - 1| < \frac{1}{4} \ , \quad |\nu - \bar{\nu}| < \frac{\bar{\nu}}{4} \ , \quad |\mu - \bar{\mu}| < \frac{1}{4} \ , \quad |\kappa| < \frac{1}{2} \ , \quad \frac{1}{2} \leq |\sigma| \leq \frac{3}{2} \ . \qquad (12.21)$$

If we suppose that $\eta = 0$ and that $p_0 \bar{M}/J_1 \equiv p_0 \omega_D = q_0 \omega_T$ for some integers p_0, q_0, the set Υ_0 is clearly a resonance for our hamiltonian if $\eta = 0$.

We shall fix, in the example that we treat here, $p_0 = 1, q_0 = 2$; thus we set $\bar{M} \equiv 2\omega_T J_1$, so that $\bar{E} \equiv J_1 \omega_T^2$.

The condition $\omega_D = 2\omega_T$, *i.e.* a day/year simple resonance $2:1$, is a condition far from the ones relevant for the Earth nutation, but it might be more realistic for other situations (*e.g.* for Mercury there is a similar simple resonance which is relevant, namely the $3:2$). In Celestial Mechanics there is however a rather general feature in the data: usually the bodies are almost spherical and the symmetry axis and the spin axis are very close. In fact the angle between such axes, measured by $\bar{\nu} = (1 - L^2/\bar{M}^2)^{1/2}$ is usually much smaller than the parameter η (in the Earth case $\bar{\nu} < \eta^2$), as it has to be according to various models of planet formation by accretion. But, unfortunately, we must require also that the initial value of L, which is a constant of the motion, verifies $\bar{\nu} \geq \bar{\nu}_0 > 0$ for some $\bar{\nu}_0$, no matter how small η is. This is a feature that makes our model somewhat unrealistic: it is a necessary requirement to guarantee that the function $\nu = (1 - L^2/M_0^2)^{1/2}$ does not become singular in the domain in which we consider it (note that the M_0 derivative causes problems even in writing down the equations for $\dot{\chi}_0$, if we do not impose that ν stays away from zero). With our choices above, we can take $\bar{\nu}_0 \equiv \bar{\nu}\ 3/4$.

We shall perform a few (trivial) changes of coordinates and rescalings to put the hamiltonian in a standard form to which the theory of §11 can be easily applied.

Because of our selection of the resonance, the harmonic $(2\gamma_0 + \chi_0)$ will be the angle of the pendulum–part of the hamiltonian. Therefore, we perform the following linear (canonical) change of variables,

$$(K_0, \gamma_0), \ (M_0, \chi_0), \ (B_0, \lambda_0) \ \to (I_0, \varphi_0), \ (A_0, \alpha_0), \ (B_0', \lambda_0) \qquad (12.22)$$

defined by $B_0 = B'_0 - (A_0 - a_0)$, with $a_0 \equiv \bar K - 2\bar M$ and:

$$\begin{array}{ll}\gamma_0 = -(\alpha_0 + \lambda_0 + \pi/2) & K_0 = 2I_0 - (A_0 - a_0) + 4\omega_T J_1 \equiv 2I_0 - A_0 + \bar K \\ \chi_0 = 2(\alpha_0 + \lambda_0) + \varphi_0 & M_0 = I_0 + 2\omega_T J_1 \equiv I_0 + \bar M\end{array} \tag{12.23}$$

where the shifts have been introduced so that the unstable point of the pendulum is η–close to $(I_0, \varphi_0) = (0,0)$ and so that the initial datum $(\bar K, \bar M)$ corresponds to $(\bar I_0, \bar A_0) \equiv (0,0)$. The hamiltonian (12.16), in the canonical variables (I_0, φ_0), (A_0, α_0), (B'_0, λ_0), takes the form (up to a neglected constant):

$$\big(\omega_T B'_0 + \frac{I_0^2}{2J_1} + \eta J_1 g^2\,(\cos\varphi_0 - 1) + \eta h\big) + \eta \bar E\Big[V_0^{(1)} + eV_1^{(1)} + e^2 V_2^{(1)}\Big] \tag{12.24}$$

where $h = h(A_0, I_0)$, $g = g(A_0, I_0)$ and the functions $V_j^{(1)}$ are simply defined in terms of the $\bar B^h_{rpj}$ functions above evaluated in the new coordinates with $V_0^{(1)}$ being derived from $V_0 - c_0$ by extracting from it the $J_1 g^2 \cos\varphi_0$ term which we call the "pendulum term". Using the appendix A14 one sees that:

$$h \equiv \bar E\Big(\frac{\sigma_0^2\mu_0^2}{2} + \frac{(1+\kappa_0)\mu_0\sigma_0}{2}\nu_0 + \frac{1+\kappa_0^2}{4}\nu_0^2\Big)\,, \quad g^2 \equiv \omega_T^2 \frac{(1+\kappa_0)\mu_0\sigma_0}{2}\nu_0 \tag{12.25}$$

where:

$$(\mu_0, \nu_0, \kappa_0, \sigma_0) \equiv (\mu, \nu, \kappa, \sigma)|_{\{(K_0, M_0) = (2I_0 - A_0 + \bar K, I_0 + \bar M)\}} \tag{12.26}$$

Analogously, one finds:

$$V_h^{(1)} = \sum_{\substack{rpj\ :\ r,p+h=even\\ |r|\le 6, |p|\le 6+h, |j|\le 2}} B^h_{rpj} \cos(r\alpha_0 + p\lambda_0 + j\varphi_0) \tag{12.27}$$

and a simple calculation yields the values of the coefficients $B^h \equiv B^h(A_0, I_0)$ in terms of the functions in (12.11) evaluated at $K_0 = 2I_0 - A_0 + \bar K$ and $M_0 = I_0 + \bar M$. The results are in appendix A14, where all the coefficients are derived. Here we just remark that all the non–trivial modes are *fast i.e.* $r \ne 0$ (see §11 and below) with the *only exception* of the two *slow* modes:

$$B^2_{-20-1} \equiv \frac{9}{2} c^0_{-1} \equiv \frac{\kappa_0\sigma_0\mu_0}{2}\nu_0\,, \qquad B^2_{202} \equiv -\frac{17}{2} d^0_2 \equiv \frac{17}{2}\frac{(1+\kappa_0)^2}{8}\nu_0^2 \tag{12.28}$$

This is the reason for having kept the second order in the e–expansion.

The initial datum $(\bar K, \bar M)$ becomes, in the new variables, $(\bar I_0, \bar A_0) \equiv (0,0)$ so that the resonance Υ_0 gets mapped into:

$$\Upsilon_1 \equiv \{I_0 = 0\ ,\ |A_0| \le \frac{\bar M}{8}\}\,, \tag{12.29}$$

and we can easily find a ρ_1, ξ_1 such that (12.23) is analytic on $\Upsilon_{1,\rho_1} \times S^3_{\xi_1}$ and the image of such domain under the canonical transformation (12.23) is contained in $\Upsilon_{0,\rho_0} \times S^3_{\xi_0}$.

To see that (12.24) can be put in the form (11.3) so that the theory of §11 can be applied, we perform a change of variables setting the unstable equilibrium of the pendulum *exactly* in the origin. Let:

$$h(A_0, I_0) \equiv h_0(A_0) + h_1(A_0, I_0)I_0 \ , \qquad h_0(A_0) \equiv h(A_0, 0) \tag{12.30}$$

with h_1 analytic and define (via the implicit function theorem) $G(A', I'; \eta) = -J_1 h_1(A', I')$ $+O(\eta)$ as the solution of:

$$\frac{(I' + \eta G)^2}{2J_1} + \eta h(A', I' + \eta G) = \frac{{I'}^2}{2J_1} + \eta h_0(A') \tag{12.31}$$

Then it is easy to check that the canonical transformation (A_0, α_0), $(I_0, \varphi_0) \to (A', \alpha')$, (I', φ') and $(B'_0, \lambda_0) \equiv (B', \lambda')$, generated by:

$$B'\lambda_0 + I'\varphi_0 + A'\alpha_0 + \eta G(A', I'; \eta) \sin\varphi_0 \tag{12.32}$$

transforms the hamiltonian (12.24) into a hamiltonian like (11.3), *i.e.* into:

$$\omega_T B' + \eta h_0(A') + \frac{{I'}^2}{2J_1} + \eta J_1 g_0^2 \ (\cos\varphi' - 1) + \eta \bar{E}[v_0 + e v_1 + e^2 v_2] \tag{12.33}$$

where (cfr. (12.25) $\div$ (12.27)):

$$h_0(a) \equiv h(a, 0) \ , \qquad v_h \equiv \sum_{\substack{rpj \ : \ r,p+h=even \\ |r|\le 6, |p|\le 6+h, |j|\le 2}} b^h_{rpj} \cos(r\alpha' + p\lambda' + j\varphi') \tag{12.34}$$

with $g_0 \equiv g_0(A', I', \eta(\cos\varphi - 1); \eta)$ and $b^h_{rpj} \equiv b^h_{rpj}(A', I', \eta(\cos\varphi - 1); \eta)$ being analytic functions of their variables A', I', $z \equiv \eta(\cos\varphi - 1)$, η and:

$$g_0^2(A', 0, 0; 0) = g^2(A', 0) \ , \qquad b^h_{rpj}(A', 0, 0; 0) \equiv B^h_{rpj}(A', 0) \tag{12.35}$$

Furthermore along the resonance Υ_1 (see (12.29) and replace (I_0, A_0) with (I', A')) one has:

$$\partial_a h_0(a) = \omega_T \frac{\hat\kappa(a)}{2}[1 + O(\bar\nu)] \ , \qquad \left(\partial_a^2 h_0(a)\right)^{-1} = -4J_1[1 + O(\bar\nu)] \tag{12.36}$$

where (see (12.26)):

$$\hat\kappa(a) \equiv \kappa|_{(K_0, M_0) = (\bar K - a, \bar M)} \ , \quad \bar\nu \equiv (1 - \bar\mu^2)^{1/2} \ , \quad \bar\mu \equiv \frac{\bar L}{\bar M} \tag{12.37}$$

To conform with the analysis of §11 we use also the scaled form of (12.33); setting:

$$\alpha' \equiv \alpha \;, \quad \lambda' \equiv \lambda \;, \quad \varphi' \equiv \varphi \;, \quad A' \equiv \eta^{1/2} A \;, \quad B' \equiv \eta^{1/2} B \;, \quad I' \equiv \eta^{1/2} I \tag{12.38}$$

and multiplying the hamiltonian by a factor η^{-1} (see remark before (11.3)) we get from (12.33), *introducing also the auxiliary parameter* β (eventually to be set equal to 1):

$$\eta^{-1/2}\omega_T B + h_0(\eta^{1/2}A) + \frac{I^2}{2J_1} + J_1 g_0^2\,(\cos\varphi - 1) + \beta\bar E[v_0 + ev_1 + e^2 v_2] \tag{12.39}$$

where g_0 and v_h are now evaluated at $\alpha' \equiv \alpha,\ldots,\ I' \equiv \eta^{1/2} I$: see (12.38).

Note that the "fast" term F in (11.9), (11.10) corresponds here to v_0 with $N = 6$ and $\vec\nu_0 = 2(1,1)$. Given the final form (12.39), we fix a diffusion curve η–close to the resonance Υ_1 (cfr. (11.12)):

$$\mathcal{L} \equiv \{|A| < \eta^{-1/2}\frac{\bar M}{8}\} \times \{I = 0\} \equiv \Delta \times \{I = 0\} \tag{12.40}$$

We take $e = \eta^c$, with c large enough to apply the theory of §11 (*e.g.* $c > 92$). If we check that the determinant of the intersection matrix given by the generalization of (11.8) is not exponentially small as $\eta \to 0$, then it will follow that the homoclinic angles are not exponentially small but have the size of a power of η, while their spacing has exponentially small size (by the averaging properties discussed in §11: see remarks 1 and 5). Recall that on a portion $\Sigma_\eta \times \{I = 0\}$ of the diffusion curve, with Σ_η of relative measure $\le K\eta^{-1/2}$ $\eta^{x/7}$ (see the discussion around (11.2)), the diophantine property (11.15) holds (with $\bar\omega_1$ $= \omega_T$ and $\omega_2 = \partial_A h_0(\eta^{1/2}A)$).

We write the leading terms which can arise only, in the second order of perturbation theory in the auxiliary parameter β, from the "interference" between v_0 and v_2. Recall that such a perturbation theory has a radius of convergence $B^*\eta^{-s}$ with a s arbitrarily close to $1/2$ (here s corresponds to the constant σ of lemma 5 of §11), so that its term sizes decrease with the order p at least as η^{2s}. With the notations of (11.8) (but note that here $\vec\nu \longleftrightarrow (p,r)$) we see that $\vec\mu$ fast corresponds to $p \ne 0$ and that $\vec\nu$ slow corresponds to $p = 0$, $r \ne 0$; thus the definitions in §11 (cfr. (11.8): $-\delta^2 \equiv$ leading term of the determinant of the intersection matrix $\partial_{\vec\alpha}\vec Q_\uparrow(\vec 0)$) and the computations in appendix A14 yield *to leading terms* (as $\eta \to 0$ with $e = \eta^c$):

$$\delta = \eta^{3/2} e^2\ \frac{4\bar\omega_2}{\omega_T^2 J_1 \bar g_0^2}\ \bar E^2\ \Big(\sum_{r\ne 0, j\ne 0} (-1)^j \frac{j}{r} B^0_{2r,2r,j}\Big)\ \big(2K_2 B^2_{202} + K_1 B^2_{-20-1}\big) \tag{12.41}$$

where (see also (12.36), (12.35), (12.25)) $\bar\omega_2 \equiv \partial_a h_0(a)$, $\bar g_0^2 \equiv g^2(a,0)$ and the coefficients B^h are computed on the image of the diffusion curve $\mathcal{L}$:

$$\begin{aligned} &B^h_{rpj} \equiv B^h_{rpj}(a,0)\;, \qquad |a| < \frac{\bar M}{8}\;; \qquad K_j = \int_{-\infty}^{+\infty} \frac{u \sin j\tilde\varphi(u)}{2\cosh u} du \\ &\tilde\varphi(u) \equiv \varphi_0(u/g) = 4\text{arctg}\, e^{-u}\;, \quad K_1 = 2\;, \quad K_2 = \frac{10\pi}{3} \end{aligned} \tag{12.42}$$

On the basis of the results of §11 we know that the order p contribution to the intersection matrix (in the expansion in the auxiliary parameter β) will be bounded at least by $O(\eta^{2s})$: but they might be much smaller. In fact if $p = 1$ they are essentially exponentially small in $\eta^{-1/2}$ as $\eta \to 0$. If $p = 2$ we see from (12.41) that they are of order $\eta^{3/2}e^2$.

To order p we can get non exponentially small contributions only from terms like $v_0^{p-1}v_2$ or $v_1^2 v_0^{p-2}$ or by terms involving at least v_1^3 or $v_1 v_2$ which contribute corrections of size e^3 at least (hence negligible with respect to (12.41)): this can be seen directly by inspection of the Fourier transform structure of the v_0, v_1, v_2 in appendix A14. But the terms $v_0^{p-1}v_2$ or $v_1^2 v_0^{p-2}$ contribute $\eta^{sp}e^2\eta^{1/2}$ where the last $\eta^{1/2}$ arises because if $\bar\omega_2 = 0$ we would have corrections smaller than any power coming even from the above terms (note that the order p is *analytic* in $\bar\omega_2$ *and* in e^2).

Hence we can set $\beta = 1$ and the leading term is actually (12.41), if it does not vanish: and the determinant of the intersection matrix is of order $\eta^3 e^4 = \eta^{3+4c}$.

It is easy to check that the sum in (12.41) does not vanish on the diffusion curve $\mathcal{L}$: in fact from (12.41),♣A14.10, ♣A14.6, (12.37), (12.36), (12.35), (12.25) it follows immediately that on $\mathcal{L}$ it is, to leading order in η and up to a factor $[1 + O(\bar\nu)]$:

$$\begin{aligned} &\delta = \frac{9}{2}\eta^{3/2}e^2\ (J_1\omega_T)\ \hat\delta_a \qquad \text{with :}\\ &\hat\delta_a \equiv \frac{1-\hat\kappa(a)}{1+\hat\kappa(a)}\ \hat\kappa(a)^2\ \hat\sigma(a)\ \bar\nu \equiv \tan^2\left(\frac{i_a}{2}\right)\cos^2 i_a\ \sin i_a\ \sin\vartheta \end{aligned} \tag{12.43}$$

where $|a| < \bar M/8$; $\hat\sigma(a)$ is defined analogously to $\hat\kappa$ (see (12.37); ϑ is defined in (12.5) and i_a is defined here and represents the planet inclination over the ecliptic (cfr. (12.8)). Equation (12.43) holds, of course, for values of the coordinate a in Σ_η (*i.e.* for all points outside a family of gaps of size smaller than any prefixed power in η: see comment after (12.40)) where the appropriate diophantine inequality holds. One could also check that (12.43) holds uniformly on paths $\mathcal{L}$ with Δ (see (12.40)) replaced by any closed interval such that $\hat\kappa(a) \neq \pm 1$ or equivalently $i_a \neq 0, \pi$ (which means that the spin axis is not parallel to the normal to the ecliptic).

This implies that on the diffusion path $\mathcal{L}$ the homoclinic angles are much larger than the whiskered tori spacing, so that we shall have (by the analysis of §8) heteroclinic ladders along which Arnold's drift and diffusion will take place on a time scale proportional to $\exp(b\eta^{-d})$ for suitable positive constants b, d.

Appendix A1 Resonances: Nekhorossev theorem

Let $h(\vec A)$ be an anisochronous hamiltonian (*i.e.* $\det \partial^2_{\vec A} h \neq 0$), analytic on $V_R \times T^l$, where $V_R = \{\vec{A} = (A_1, \ldots, A_l),\ |A_j| \le R\}$ and holomorphic in:

$$W(V_R;\rho,\xi) = \{\vec{A},\vec{z}\ \ |(\vec{A},\vec{z}) \in C^{2l};\ |A_j - A_{0j}| < \rho \\ \text{for some } \vec{A}_0 \in V_R,\ e^{-\xi} < |z_j| < e^{\xi}\} \tag{A1.1}$$

see [BG], p. 296: it is convenient to regard h as defined on W even though it is independent on the angles φ_j, $(z_j \equiv e^{i\varphi_j})$.

Let $\vec{\nu}_1 \in Z^l$ and let $\mathcal{M}$ be the line parallel to $\vec{\nu}_1$; define the *resonance surface* $\Sigma_{\mathcal{M}}$ as: $\Sigma_{\mathcal{M}} = \{\vec{A}|\vec{\omega}(\vec{A}) \cdot \vec{\nu}_1 = 0\}$, with $\vec{\omega}(\vec{A}) \equiv \partial_{\vec{A}} h(\vec{A})$. There is no loss of generality if one takes $\vec{\nu}_1 = (0,0,\ldots,0,1)$, (see [BG] proposition 3 and p.303, for the obvious change of coordinates).

So we suppose that $\Sigma_{\mathcal{M}}$ is defined by:

$$\partial_{A_l} h(\vec{A}) = 0 \tag{A1.2}$$

and rename $\vec{F} = (A_1, \ldots, A_{l-1})$, $S = A_l$, and $\vec{\varphi}, \sigma$ the conjugate angles; the notation is motivated by the fact that σ is a *slow angle* (indeed on the resonance $\sigma(t)$ does not move at all) in opposition to φ which is *fast* as it evolves on a time scale of order 1.

Let $f(S,\vec{F},\sigma,\vec{\varphi})$ be a perturbation and consider the hamiltonian:

$$H_\varepsilon(S,\vec{F},\sigma,\vec{\varphi}) = h(\vec{F},S) + \varepsilon f(S,\vec{F},\sigma,\vec{\varphi}) \tag{A1.3}$$

We shall assume that $\partial_S^2 h(\vec{F},S) \ne 0$ on $\Sigma_{\mathcal{M}}$ and let $\vec{F} \to s(\vec{F})$ be the equation for $\Sigma_{\mathcal{M}}$, ($i.e.$ $\partial_S\ h(\vec{F},\ s(\vec{F})) \equiv 0$).

Let $\mathcal{U}_{\mathcal{M}}$ be a resonant region of order 1 with parameters $b, \sigma_1, \sigma_2, \lambda_1^0, \lambda_2^0$, defined by:

$$|\vec{\omega}(\vec{F},S) \cdot \vec{\nu}_1| < \lambda_1^0 \varepsilon^{\sigma_1} \\ |\vec{\omega}(\vec{F},S) \cdot \vec{\nu}| > \lambda_2^0 \varepsilon^{\sigma_2}, \qquad \forall\ |\vec{\nu}| < \varepsilon^{-b}, \quad \vec{\nu} \text{ not parallel to } \vec{\nu}_1 \tag{A1.4}$$

where, see [BG], (3.1)%(3.6):

$$\begin{aligned} b &= (8l(l+1))^{-1}, \qquad\qquad \sigma_1 = 8^{-1} \\ \sigma_2 &= (1 - 1/l(l+1))8^{-1}, \qquad \lambda_i^0 = E2^{i-1}\big(m/8lM\big)^{-l+i} \\ M &= \{\text{maximum of the absolute values of the eigenvaluesof } \partial^2_{\vec{A}\vec{A}} h\} \\ m &= \{\text{minimum of the absolute values of the eigenvaluesof } \partial^2_{\vec{A}\vec{A}} h\} \\ E &= \{\text{maximum of } |\partial_{\vec{A}} h|\} \end{aligned} \tag{A1.5}$$

where the max and min are considered in the holomorphy domain, (A1.1).

Let $W_\varepsilon \equiv W(\mathcal{U}_{\mathcal{M}};\rho',\xi')$ be a vicinity of $\mathcal{U}_{\mathcal{M}}$, with:

$$\rho' = \varepsilon^{1/4} \lambda_1^0 (8M)^{-1}, \qquad \xi' = \xi/8 \tag{A1.6}$$

Then, see [BG] proposition 2, ii), if ε is small enough, (*e.g.* $|\varepsilon| < \varepsilon_c$ with ε_c defined in (3.6) of [BG])), for all $1 \le p \le \varepsilon^{-b/3}$ one can find a change of coordinates changing the hamiltonian into:

$$h_p(\vec{F}, S, \varepsilon) + \varepsilon G_p(\vec{F}, S, \sigma, \varepsilon) + \varepsilon^p f_p(\vec{F}, S, \vec{\varphi}, \sigma, \varepsilon) \tag{A1.7}$$

where the new $(\vec{F}, S, \vec{\varphi}, \sigma)$ coordinates vary in W_ε and describe at least all the points which are in $W_{\varepsilon/4}$ in terms of the original coordinates.

Furthermore the change of coordinates is analytic and in the whole domain W_ε:

$$\begin{aligned} h_p(\vec{F}, S, \varepsilon) \equiv & h(\vec{F}, S) + \varepsilon \bar{f}(\vec{F}, S) + O(\varepsilon^2) \\ G_p(\vec{F}, S, \sigma, \varepsilon) \equiv & \tilde{f}(\vec{F}, S, \sigma) + O(\varepsilon), \qquad\qquad |f_p(S, \vec{F}, \sigma, \varepsilon)| \le O(1) \end{aligned} \tag{A1.8}$$

where $\bar{f}$ is the average of f over both $\vec{\varphi}$ and σ and $\tilde{f}(\vec{F}, S, \sigma, \varepsilon)$ is the average of $f - \bar{f}$ over $\vec{\varphi}$ alone.

Finally if one is interested in a fixed, ε-independent, value of p then one can fix ρ' in (A1.6) to be ε-independent, see also the following appendices A10, A11.

Appendix A2 Diffusion paths and diophantine conditions

Here we prove the claims in §3. Fix E and $\vec{A}_0$ such that $h(\vec{A}_0, 0) = E$; see (2.3) and assumptions 1÷3 of §2. We simplify the notation in this appendix by replacing $h(\vec{A}, 0)$ with $h(\vec{A})$. We consider first the case $h(\vec{A}) = \vec{A}^2/2$, so that $\vec{\omega}(\vec{A}) \equiv \partial_{\vec{A}} h(\vec{A}) \equiv \vec{A}$. We consider a small vicinity U_ε of $\vec{A}_0$ with diameter ε. Given $\vec{A}^1, \vec{A}^2$ in U_ε we define the curve $\mathcal{L}_0$ as:

$$s \to \vec{A}^1 + \sum_{j=1}^{l-1} s^j (A^2_j - A^1_j)_j \vec{u}_j = \vec{A}(s) \equiv \vec{A}_s \qquad s \in [0, 1] \tag{A2.1}$$

where $\vec{u}_1, \ldots, \vec{u}_{l-1}$ are the natural basis in R^{l-1}. Without loss of generality we suppose that $\vec{A}^1$ and $\vec{A}^2$ have different corresponding coordinates: $A^1_i \ne A^2_i$. Let $\vec{\omega}_s = \vec{\omega}(\vec{A}_(s))$.

At every point of $\mathcal{L}_0$ the derivatives of $\vec{A}_s$ of orders $1, \ldots, k \le l-1$ span a k dimensional space, *i.e.* the curve has "full torsion".

No codimension 1 plane can have a contact with $\mathcal{L}_0$ of order higher than $l-1$. Therefore, given $\vec{\nu} \in Z^l$, $\vec{\nu} \ne \vec{0}$, the set of the values of s for which $|\vec{\omega}_s \cdot \vec{\nu}|/|\vec{\nu}| \le \eta_{\vec{\nu}}$ has a measure that can be bounded by *const* $|\eta_{\vec{\nu}}|^{1/(l-1)}$. It follows that the measure of the set of values of s for which $|\vec{\omega}_s \cdot \vec{\nu}| < \eta_{\vec{\nu}} |\vec{\nu}|$ will be bounded by *const* $\sum |\eta_{\vec{\nu}}|^{1/(l-1)}$.

Thus, if we choose $|\eta_{\vec{\nu}}| = 1/C|\vec{\nu}|^{(l-1)^2+1}$ it is:

$$|\vec{\omega}_s \cdot \vec{\nu}|\, |\vec{\nu}|^{-1} > C^{-1} |\vec{\nu}|^{-(l-1)^2-1} \tag{A2.2}$$

on a set Σ of values of s with measure of order:

$$1 - const \sum_{\vec{\nu}} (C|\vec{\nu}|^{(l-1)^2+1})^{-1/(l-1)} \geq 1 - const\, C^{-1/(l-1)} \tag{A2.3}$$

Thus the curve $s \to \vec{A}(s)$ has the property (3.1),(3.2), but it does not necessarily verify $h(\vec{A}(s)) = E$.

Therefore we modify $\mathcal{L}_0$ into $\mathcal{L}$ defined by: $s \to \vec{A}(s)+\vec{\delta}(s)$. We determine the correction $\vec{\delta}(s)$ together with an auxiliary parameter $\gamma(s)$ by the equations:

$$\begin{aligned} \vec{\partial} h(\vec{A}+\vec{\delta}) =& \vec{\omega}(\vec{A})(1+\gamma) \\ h(\vec{A}+\vec{\delta}) =& E \end{aligned} \tag{A2.4}$$

with $\vec{A} = \vec{A}(s)$.

The latter equations, in linearized form, look like:

$$\begin{aligned} \partial^2_{\vec{A}} h(\vec{A})\vec{\delta} =& \gamma\vec{\omega}(\vec{A}) + \dots & \vec{\delta}(\vec{A}) =& \gamma(\partial^2_{\vec{A}} h)^{-1}(\vec{A})\vec{\omega}(\vec{A}) + \dots \\ \vec{\omega}(\vec{A}) \cdot \vec{\delta} =& E - h(\vec{A}) + \dots & \gamma(\vec{A}) =& \frac{E - h(\vec{A})}{\vec{\omega}(\vec{A}) \cdot (\partial^2 h(\vec{A}))^{-1}\vec{\omega}(\vec{A})} + \dots \end{aligned} \tag{A2.5}$$

which show that (A2.3) can be solved expressing $\vec{\delta}, \gamma$ as analytic functions of $\vec{A}, h - E$ at least if the size ε of U_ε is small enough. Hence also the first property of the diffusion paths holds for $\mathcal{L}$ which is therefore an example of a (short) diffusion path.

The general case in which $h(\vec{A})$ is not $\vec{A}^2/2$, but still $\det \partial_{\vec{A}\vec{A}} h \neq 0$, can be reduced to the above by changing variables $\vec{A} \longleftrightarrow \vec{\omega} = \vec{\omega}(\vec{A}) = \partial_{\vec{A}} h(\vec{A})$ and by drawing the curves in $\vec{\omega}$ coordinates.

If one gives up full constructivity one can produce a somewhat different class of examples: they are even better, as far as the exponents in (3.1), (3.2) are concerned. But they are constructed with the help of measure theoretic lemmata and, therefore, are not *really* constructive examples.

Appendix A3: normal hyperbolic coordinates for a pendulum

Here we prove lemma 0 of §5. Although the following proof is elementary, we report it here in detail to establish the values of several constants needed in the main text. The proof is based on a iteration method in the style of §5, but it is clear that softer methods could also be used.

Let $P(I,\varphi) \equiv P(I,\vec{A},\varphi,\mu)$ be a pendulum hamiltonian holomorphic for $|I| < \rho', e^{-\xi'} < |e^{i\varphi}| < e^{\xi'}$, see (2.3) and assumptions 1,2 of §2. We fix $\vec{a} \in V$ and we regard $\vec{A}, \mu$ in:

$$(\vec{A},\mu) \in \mathcal{D} \equiv \{\vec{A} \in C^{l-1}, \mu \in C :\ |\vec{A} - \vec{a}| \leq \rho\ ,\ |\mu| \leq \bar{\mu}\} \tag{A3.1}$$

as parameters, which often will be omitted from the notation. The following analysis is local (near $(I,\varphi)=(0,0)$) therefore we shall consider only $|\varphi|<\xi'$.

By assumption 1, §2, $(I,\varphi)=(0,0)$ is an unstable equilibrium point; hence $\partial_{(I,\varphi)}P(0,0) = 0$ and the matrix $M\equiv\partial^2P(0,0)$ can be put in a off diagonal form via a canonical transformation $\mathcal{R}\equiv\mathcal{R}(\vec{A},\mu)$:

$$\mathcal{R}^T M\mathcal{R} = g\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \text{with } g^2\equiv(\partial^2_{I\varphi}P)^2-\partial^2_{II}P\partial^2_{\varphi\varphi}P \tag{A3.2}$$

with the derivatives evaluated at $(I,\vec{A},\varphi,\mu)=(0,\vec{A},0,\mu)$: in fact $\mathcal{R}$ is one of the (∞^1-many) matrices with determinant 1 (and hence symplectic, since we are in dimension two) that diagonalize EM where $E\equiv\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ (given one of such matrices one obtains the others by right multiplication by $\begin{pmatrix}\sigma^{-1} & 0\\ 0 & \sigma\end{pmatrix}$, for any $\sigma\neq 0$). We select one among the above canonical maps as follows. Let:

$$\kappa^2\equiv(16)^{-1}\cdot\ [\text{ area enclosed by the two separatrices swings }] \tag{A3.3}$$

this is a natural unit of measure for the pendulum action I (*e.g.* , for the standard pendulum (2.1) it is $\kappa^2=J_0g$); define the "dimensionless energy" $P_0(x_1,x_2)\equiv P(x_1\kappa^2,x_2)/(g\kappa^2)$ and $M_0\equiv\partial^2P_0|_{(x_1,x_2)=(0,0)}$, so that:

$$M=g\,D\,M_0\,D,\quad D\equiv\begin{pmatrix}\kappa^{-1} & 0\\ 0 & \kappa\end{pmatrix} \tag{A3.4}$$

Now let $\pm\lambda^{\pm1}$, $\lambda>0$, be the eigenvalues of M_0 (recall that $\det M=-g^2$ so that $\det M_0 = -1$), and let U,V be the unitary matrices with determinant 1 that diagonalize, respectively, M_0 and $\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$:

$$U^TM_0U=\begin{pmatrix}\lambda & 0\\ 0 & -\lambda^{-1}\end{pmatrix},\quad V^T\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}V=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix},\quad V\equiv\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix} \tag{A3.5}$$

then it is immediate to check that (see (A3.4)):

$$\mathcal{R}=D^{-1}U\Lambda V^T,\qquad \Lambda\equiv\begin{pmatrix}\lambda^{-1/2} & 0\\ 0 & \lambda^{1/2}\end{pmatrix} \tag{A3.6}$$

We also set:

$$\mathcal{R}_0\equiv D\mathcal{R},\quad \mathcal{R}_0=\begin{pmatrix}a & b\\ c & d\end{pmatrix},\quad m^2\equiv\sup_{\mathcal{D}}[|a|^2+|b|^2+|c|^2+|d|^2]=\sup_{\mathcal{D}}(\lambda+\lambda^{-1}) \tag{A3.7}$$

(to check last equality, note that UV^T is unitary so that the sum of the squared absolute values of the entries is equal to 2).

We define a first canonical map via:

$$\begin{pmatrix} I \\ \varphi \end{pmatrix} = \begin{pmatrix} \mathcal{R}_{11} & \mathcal{R}_{12} \\ \mathcal{R}_{21} & \mathcal{R}_{22} \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \tag{A3.8}$$

For instance in the case of the standard pendulum in (2.1) it is $\kappa^2 = J_0 g$ and:

$$\mathcal{R} = \frac{1}{\sqrt{2}} \begin{pmatrix} \kappa & \kappa \\ -\kappa^{-1} & \kappa^{-1} \end{pmatrix}, \qquad \mathcal{R}^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} \kappa^{-1} & -\kappa \\ \kappa^{-1} & \kappa \end{pmatrix} \tag{A3.9}$$

For later convenience it is useful to write explicitly the generating function of (A3.8):

$$F(p,\varphi) \equiv F(p,\varphi;\vec{A},\mu) \equiv \mathcal{R}_{22}^{-1}(\mathcal{R}_{12}\frac{\varphi^2}{2} + p\varphi - \mathcal{R}_{21}\frac{p^2}{2}) \ , \ I = F_\varphi \ , \ q = F_p \tag{A3.10}$$

provided:

$$\sup_{\mathcal{D}} |\mathcal{R}_{22}|^{-1} \equiv \hat{\kappa} < \infty \tag{A3.11}$$

(clearly (A3.11) can be achieved, by taking $\rho, \bar{\mu}$ small enough, if $\mathcal{R}_{22}(\vec{a},0) \neq 0$, otherwise we would have to choose different independent "mixed variables"; see [G] for general informations on canonical transformations).

In the new coordinates p, q it is clear that $P(I,\varphi)$ becomes:

$$P(I,\varphi) = G_0(pq) + Q_0(p,q) \equiv K_0(p,q) \tag{A3.12}$$

where Q_0, G_0 are holomorphic for $|p|,|q| < \tilde{\kappa}_0$, if $\tilde{\kappa}_0 > 0$ is suitably small: to guarantee that the image under (A3.8) of the complex domain $\{|p|,|q| < \tilde{\kappa}_0\}$ is contained in $|I| < \rho'$, $|\varphi| < \xi'$, we can take (see (A3.7)):

$$\tilde{\kappa}_0 \le \frac{\kappa}{2m} \min\{\rho'\kappa^{-2}, \xi'\} \tag{A3.13}$$

The Taylor series of Q_0 starts at third order:

$$Q_0(p,q) = \sum_{\substack{h+k\ge 3 \\ h\ne k}} Q^0_{hk} p^h q^k, \qquad (p,q) \in W_{\tilde{\kappa}_0} \equiv \{|p|,|q| < \tilde{\kappa}_0\} \tag{A3.14}$$

and $G_0(J)$ is holomorphic for $|J| < \tilde{\kappa}_0^2$:

$$G_0(J) = \sum_{h=1}^{\infty} g_h J^h, \quad g_1 \equiv g \ , \quad J \in \tilde{W}_{\tilde{\kappa}_0} \equiv \{|J| < \tilde{\kappa}_0^2\} \tag{A3.15}$$

Suppose that:

$$||G_0||_{\tilde{\kappa}_0} < E_0\tilde{\kappa}_0^2, \qquad ||(\partial_J G_0)^{-1}||_{\tilde{\kappa}_0} < \gamma_0, \qquad ||Q_0||_{\tilde{\kappa}_0} < \varepsilon_0\tilde{\kappa}_0^2 \tag{A3.16}$$

where $||\cdot||_{\tilde{\kappa}} \equiv ||\cdot||_{\tilde{\kappa},\rho,\bar{\mu}}$ denotes the maximum in $\tilde{W}_{\tilde{\kappa}}$ (or $W_{\tilde{\kappa}}$, whichever makes sense) times $\mathcal{D}$, see (A3.1). Let $\delta_j = \delta_0/2^j$, $\delta_0 = (1/8)\log 2$, be a convenient sequence of positive numbers. We define a canonical map $(p,q) = \mathcal{C}_0(p',q')$ via a generating function:

$$\Phi_0(p',q) = \sum_{\substack{h+k\geq 3\\ h\neq k}} \frac{{p'}^h q^k Q^0_{hk}}{-(k-h)\partial_J G_0(p'q)} \tag{A3.17}$$

The function Φ_0 can be bounded, together with its derivatives, in $W_{\tilde{\kappa}_0 e^{-\delta_0/2}}$ by:

$$\tilde{\kappa}_0||\partial\Phi_0||_{\tilde{\kappa}_0 e^{-\delta_0/2}}, ||\Phi_0||_{\tilde{\kappa}_0 e^{-\delta_0/2}} \leq B_1\gamma_0\varepsilon_0\kappa_0^2\delta_0^{-1} \tag{A3.18}$$

if B_1 is a suitable constant ($e.g.$ $B_1 = \sup_\delta \delta\sum_{0\neq h\neq k}|h-k|^{-1}e^{-\delta(h+k)/2}$). To estimate the domain of definition of the map generated by Φ_0, $i.e.$ defined by the relations:

$$p = p' + \partial_q\Phi_0(p',q) \qquad q' = q + \partial_{p'}\Phi_0(p',q) \tag{A3.19}$$

we use the implicit function theorem. The latter (see for instance, [G, p.490]) will guarantee that $\mathcal{C}_0$ and $\mathcal{C}_0^{-1}$ have a domain containing $W_{\tilde{\kappa}_0 e^{-\delta_0}}$ with images contained in $W_{\tilde{\kappa}_0 e^{-\delta_0/2}}$ provided:

$$B_2\gamma_0\varepsilon_0\delta_0^{-3} < 1 \tag{A3.20}$$

for B_2 large enough (we simply use that the image of the boundary of a set under a holomorphic map is the boundary of the image; this gives, for instance, $B_2 = 16B_1$). Notice that $\mathcal{C}_0, \mathcal{C}_0^{-1}$ have the form: identity + second order polynomial (in the p,q variables). Assuming (A3.20) valid we can write the hamiltonian (A3.12) in the new coordinates, writing it as:

$$K_1(p',q') = G_1(p'q') + Q_1(p',q') \tag{A3.21}$$

where Q_1 is defined in terms of Q':

$$Q'(p',q') = G_0(pq) - G_0(p'q') + Q_0(p,q) \equiv \sum_{h+k\geq 3} Q'_{hk}{p'}^h{q'}^k \tag{A3.22}$$

by setting:

$$Q_1(p',q') \equiv \sum_{\substack{h+k\geq 3\\ h\neq k}} Q'_{hk}{p'}^h{q'}^k, \quad G_1(J) \equiv G_0(J) + \sum_{h\geq 2} Q'_{hh}J^h \equiv G_0 + \Delta_0 \tag{A3.23}$$

The estimate of the size of Q_1 can be performed by taking into account that Φ_0 has been chosen so as to verify the *first order Hamilton-Jacobi equation*:

$$\partial_J G_0(p'q)\big(q\partial_q\Phi_0(p',q) - p'\partial_{p'}\Phi_0(p',q)\big) + Q_0(p',q) = 0 \tag{A3.24}$$

so that, using (A3.16) and $\gamma_0 E_0 \geq 1$, we find, for a suitable B_3:

$$||Q'||_{\tilde{\kappa}_0 e^{-2\delta_0}} \leq B_3\, E_0 \tilde{\kappa}_0^2\, \gamma_0^2 \varepsilon_0^2\, \delta_0^{-6} \tag{A3.25}$$

Therefore for a suitable B_4:

$$\begin{aligned} ||Q_1||_{\tilde{\kappa}_0 e^{-3\delta_0}} \leq & B_4\, E_0 \tilde{\kappa}_0^2 (\gamma_0\, \varepsilon_0)^2\, \delta_0^{-7} \\ ||\Delta_0||_{\tilde{\kappa}_0 e^{-3\delta_0}} \leq & B_4\, E_0 \tilde{\kappa}_0^2\, (\gamma_0\, \varepsilon_0)^2\, \delta_0^{-7} \end{aligned} \tag{A3.26}$$

This, in turn implies, for a suitable B_5:

$$\begin{aligned} ||G_1||_{\tilde{\kappa}_0 e^{-4\delta_0}} \leq & E_0 \tilde{\kappa}_0^2 + B_5\, E_0 \tilde{\kappa}_0^2 \gamma_0^2 \varepsilon_0^2\, \delta_0^{-7} \\ ||(\partial_J G_1)^{-1}||_{\tilde{\kappa}_0 e^{-4\delta_0}} \leq & \gamma_0 (1 + B_5\, E_0 \gamma_0 \varepsilon_0^2 \gamma_0^2\, \delta_0^{-8}) \qquad \text{if } B_5\, E_0 \gamma_0 \varepsilon_0^2 \gamma_0^2 \delta_0^{-8} < 1 \end{aligned} \tag{A3.27}$$

Hence if we suppose (see (A3.20), (A3.27)), that for some B_6 large enough it is:

$$B_6\, E_0 \gamma_0 (\varepsilon_0 \gamma_0)^{1/2} \delta_0^{-8} < 1 \tag{A3.28}$$

we see that both conditions in (A3.20), (A3.27), are satisfied; and the hamiltonian K_1 is defined in terms of the functions G_1, Q_1 which can be bounded as in (A3.16) with constants $\varepsilon_1, \gamma_1, E_1, \tilde{\kappa}_1$ which can be taken as given by:

$$\begin{aligned} \tilde{\kappa}_1 = & \tilde{\kappa}_0 e^{-4\delta_0}, \quad \gamma_1 = \gamma_0 (1 + (\varepsilon_0 \gamma_0)^{1/2}), \\ E_1 = & E_0 (1 + (\varepsilon_0 \gamma_0)^{1/2}), \quad \gamma_1 \varepsilon_1 = (\gamma_0 \varepsilon_0)^{3/2} \end{aligned} \tag{A3.29}$$

Hence if we disregard (A3.28) and define $(\gamma_j, E_j, \tilde{\kappa}_j, \varepsilon_j)$ by iterating (A3.29) we see that, if $\gamma_0 \varepsilon_0$ is small enough, the sequence verifies:

$$\tilde{\kappa}_j \geq \frac{2}{3} \tilde{\kappa}_0, \qquad E_j \leq 2E_0, \qquad \gamma_j \leq 2\gamma_0, \qquad \gamma_j \varepsilon_j \leq (\gamma_0 \varepsilon_0)^{(3/2)^j} \tag{A3.30}$$

This allows us to infer that (A3.28) with j replacing 0 will be automatically verified for all but a finite number of values of j if $\varepsilon_0 \gamma_0$ is sufficiently smaller than 1. Therefore, under a condition like:

$$B_7 (E_0 \gamma_0)^2 (\varepsilon_0 \gamma_0) < 1 \tag{A3.31}$$

with B_7 suitably chosen, we see that (A3.28), with j instead of 0, is verified for all $j \geq 0$. Thus we conclude that $\mathcal{C}_j \to$ identity very fast with all its derivatives, in the slightly smaller domain $W_{\kappa'}$ with $\kappa' = \tilde{\kappa}_0 \exp - \sum_j^\infty 4\delta_j \equiv \tilde{\kappa}_0/2$. The composition:

$$(p, q) = \mathcal{C}(p_0, q_0) \equiv \mathcal{C}(p_0, q_0; \vec{A}_0, \mu) \equiv \lim_{n \to \infty} \mathcal{C}_n \mathcal{C}_{n-1} \ldots \mathcal{C}_0 (p_0, q_0)\ , \quad (p_0, q_0) \in W_{\tilde{\kappa}_0/2} \tag{A3.32}$$

is clearly a canonical map casting K_0 in the form $K_\infty(p_0q_0)$, for a suitable function K_∞, and defining *normal hyperbolic coordinates.* Finally we remark that in our case, since Q_0 has a third order zero it is not restrictive to suppose that:

$$\varepsilon_0 = \overline{Q}\ \tilde\kappa_0 \quad \text{for some constant} \quad \overline{Q} > 0 \tag{A3.33}$$

hence (A3.31) can be fulfilled for $\tilde\kappa_0$ small enough, *i.e.* :

$$\tilde\kappa_0 < \kappa[B_7(E_0\gamma_0)^2(\overline{Q}\kappa\gamma_0)]^{-1} \tag{A3.34}$$

It is clear that this conclusion is what we need to establish the claims from which this appendix is called. In fact the map $\overline{\mathcal{R}}_\mu$ we are after (see (5.3)) is given by:

$$\begin{gathered}\begin{pmatrix} R(\vec A_0, p_0, q_0, \mu) \\ S(\vec A_0, p_0, q_0, \mu)\end{pmatrix} \equiv \mathcal{R}(\vec A_0, \mu)\mathcal{C}(p_0, q_0; \vec A_0, \mu) \\ \vec\alpha = \vec\alpha_0 - \partial_{\vec A}F(p, S; \vec A_0, \mu) - \partial_{\vec A_0}\Phi(p_0, q; \vec A_0, \mu) \equiv \vec\alpha_0 + \vec\delta(\vec A_0, p_0, q_0, \mu)\end{gathered} \tag{A3.35}$$

where (p, q) is as in (A3.32) and $p_0q + \Phi(p_0, q; \vec A_0, \mu)$ denotes the generating function associated to $\mathcal{C}$.

The above condition (A3.31) can be easily used to infer explicit values $\bar\rho_0, \bar\kappa_0, \bar\xi_0$ which are needed in the first step of the proof of lemmata 1,1'. From the definition of F, (A3.10), and the fact that (see (A3.7):

$$\sup_{W_{\frac{\tilde\kappa_0}{2}} \times \mathcal{D}} |S| \le m\frac{\tilde\kappa_0}{\kappa} \tag{A3.36}$$

it follows for a suitable constant $B_8 > 1$:

$$\sup_{W_{\frac{\tilde\kappa_0}{2}} \times \mathcal{D}} |F(p, S, \vec A, \mu)| \le B_8\ m^3\tilde\kappa_0^2\frac{\hat\kappa}{\kappa} \tag{A3.37}$$

where we have used that $m > 1$ and (A3.11). Thus if $\mathcal{D}'$ is defined as in (A3.1) with ρ replaced by $\rho/2$ (so as to be able to perform dimensional bounds) we obtain easily:

$$\sup_{W_{\frac{\tilde\kappa_0}{2}} \times \mathcal{D}'} |\vec\delta| \le B_9 m^3\big(1 + \frac{\hat\kappa}{\kappa}\big)\frac{\tilde\kappa_0^2}{\rho}\ , \quad \mathcal{D}' \equiv \{|\vec A - \vec a| \le \frac{\rho}{2}, |\mu| \le \bar\mu\} \tag{A3.38}$$

for a suitable constant $B_9 > 1$ (we have also used $\varepsilon_0\gamma_0 < 1$ in view of (A3.31)). Thus we see that we can take:

$$\bar\kappa_0 \equiv \frac{\tilde\kappa_0}{2}\ , \quad \bar\rho_0 \equiv \frac{\rho}{2}\ , \quad \bar\xi_0 \equiv \frac{\xi}{2} \tag{A3.39}$$

provided $\tilde\kappa_0$ satisfies (A3.13), (A3.34) (or (A3.31)) and:

$$\tilde\kappa_0^2 < \xi\rho\ [2B_9\ m^3\ (1 + \hat\kappa/\kappa)]^{-1} \tag{A3.40}$$

a condition that guarantees that if $|\operatorname{Im}\alpha_{0j}| < \bar\xi_0$ then $|\operatorname{Im}\alpha_j| < \xi$ (see (A3.38)). As already pointed out, the choice of $\bar\rho_0$ allowed us to perform dimensional bounds and get (A3.38).

As an example consider the case of a standard pendulum hamiltonian, see (2.1), with g, J_0 being $\vec A$ independent. Then, $\kappa = \sqrt{J_0 g}$, $m = \sqrt{2}$, $\hat\kappa = (2J_0 g)^{-1}$ and, by (A3.9):

$$P = P\Big((p+q)(J_0 g/2)^{1/2}, \frac{q-p}{(2J_0 g)^{1/2}}\Big) \tag{A3.41}$$

and we see that:

$$\begin{aligned} G_0 =& gpq + J_0 g^2 \sum_{k=2}^{\infty} \binom{2k}{k} \big(\frac{pq}{2J_0 g}\big)^k \frac{1}{(2k)!} \\ Q_0 =& \sum_{k=2}^{\infty} \frac{(-1)^k}{(2k)!} \Big(\frac{p-q}{(2J_0 g)^{1/2}}\Big)^{2k}_{\neq} \end{aligned} \tag{A3.42}$$

where the subscript $\neq$ means that the expansion of the binomial has to be carried out omitting the terms with $p^k q^k$. So that setting:

$$\tilde\kappa_0 = x\sqrt{J_0 g}, \qquad x < 1/4 \tag{A3.43}$$

where the restriction $x < 1/4$ is imposed to simplify the analysis, we easily find that we can take:

$$E_0 = b_1 g, \qquad \varepsilon_0 = b_2 g x^2, \qquad \gamma_0 = b_3 g^{-1} \tag{A3.44}$$

(*e.g.* $b_1 = \cosh\sqrt{2}$, $b_2 = 4b_1$, $b_3 = 2$). The condition (A3.31) becomes simply: $x < b_4^{-1}$ for a suitable $b_4 > 1$ and to match also (A3.13), (A3.34) we see that we can take $\rho_0 = \rho/2$, $\xi_0 = \xi/2$ and:

$$\tilde\kappa_0 \equiv b_4^{-1}\sqrt{J_0 g}\,\min\big(\xi', \frac{\rho'}{J_0 g}, \xi'\frac{\rho}{J_0 g}\big) \tag{A3.45}$$

(Note that in the present case $\vec\delta$ is actually identically zero and in (A3.45) one can drop the third argument in the minimum).

In the $\vec A, \mu$ dependent case, if $\kappa_{\min} = \min\sqrt{J_0 g}$ and $\kappa_{\max} = \max\sqrt{J_0 g}$ with the extrema evaluated as $\vec A, \mu$ vary in $\mathcal{D}$, the (A3.45) is replaced by:

$$\tilde\kappa_0 \equiv b_4^{-1}\kappa_{\min}\min\big(\xi', \frac{\rho'}{\kappa_{\max}}, \xi'\frac{\rho}{\kappa_{\max}}\big) \tag{A3.46}$$

and we can still take $\bar\rho_0 = \rho/2, \bar\xi_0 = \xi/2$.

In general it is possible to express more explicitly the conditions (A3.13), (A3.31) and (A3.40). Note that the domain of definition of the dimensionless energy $P_0(x_1, x_2)$, defined before (A3.4), is $|x_1| < \rho'/\kappa^2, |\varphi| < \xi'$. Recall the definitions of $m, \hat\kappa$ ((A3.7), (A3.11));

define the parameters $\mathcal{E}, \Gamma, m$ by the following suprema in the latter domain of definition times $\mathcal{D}$ (see (A3.1)):

$$\mathcal{E} = \sup |P_0|, \quad \Gamma = \sup |g^{-1}|, \quad m^2 = \sup(\lambda + \lambda^{-1}) \tag{A3.47}$$

and observe that from the construction of the map $\mathcal{C}$ it follows (see (A3.35), (A3.2)):

$$P(I, \varphi) \equiv \frac{1}{2} M \begin{pmatrix} I \\ \varphi \end{pmatrix} \cdot \begin{pmatrix} I \\ \varphi \end{pmatrix} + \sum_{h+k \geq 3} P_{hk} I^h \varphi^k = g\ pq + \sum_{h+k \geq 3} P_{hk} I^h \varphi^k \tag{A3.48}$$

Set $\beta \equiv \kappa^2/\rho'$ and introduce the following parameters:

$$\hat{\sigma} \equiv \max\{1, \frac{\hat{\kappa}}{\kappa}\}\ , \quad \sigma_2 \equiv \max\{1, \beta^2\ , \frac{\beta}{\xi'}\ , \frac{1}{\xi'^2}\}\ , \quad \sigma_3 \equiv \max\{1, \beta^3\ , \beta^2 \frac{1}{\xi'}\ , \frac{\beta}{\xi'^2}\ , \frac{1}{\xi'^3}\} \tag{A3.49}$$

and observe that (see (A3.7)):

$$\sup |I| \leq m\kappa\tilde{\kappa}_0\ , \quad \sup |\varphi| \leq m \frac{\tilde{\kappa}}{\kappa}\ , \quad |P_{hk}| \leq \frac{\mathcal{E}}{{\rho'}^h {\xi'}^k}\ , \quad \sup |g| \leq 4 \frac{\mathcal{E}}{\rho' \xi'} \tag{A3.50}$$

(where $(I, \varphi) = \mathcal{C}(p, q)$ and the suprema are taken over the usual domain $W_{\tilde{\kappa}_0/2} \times \mathcal{D}$). Then one check easily that, for a suitable constant $B_{10} > 1$ the parameters in (A3.16) can be taken to be (cfr. (A3.47), (A3.33)):

$$E_0 \equiv B_{10}(\mathcal{E}\kappa^{-2}) m^2 \sigma_2\ , \quad \gamma_0 \equiv 2\Gamma\ , \quad \varepsilon_0 \equiv B_{10}(\mathcal{E}\kappa^{-2}) m^3 \sigma_3 \frac{\tilde{\kappa}_0}{\kappa} \equiv \overline{Q}\tilde{\kappa}_0 \tag{A3.51}$$

provided:

$$B_{10}(\Gamma\mathcal{E}\kappa^{-2}) \frac{\tilde{\kappa}_0}{\kappa} \sigma_3 < 1 \tag{A3.52}$$

a condition which is needed in bounding $(\partial_J G_0)^{-1}$ in terms of Γ. Finally we see that all the smallness requirements on $\tilde{\kappa}_0$ (*i.e.* (A3.13), (A3.34), (A3.40), (A3.52)) are enforced by taking, for a suitable $B > 1$:

$$\tilde{\kappa}_0 < \frac{\kappa}{Bm^7} \min\{\frac{\rho'}{\kappa^2}\ , \xi'\ , \frac{1}{(\mathcal{E}\Gamma\kappa^{-2})\sigma_2^2\sigma_3}\ , \frac{\rho\xi}{\rho'\hat{\sigma}}\} \tag{A3.53}$$

determining the range of $\tilde{\kappa}_0$ in terms of the analyticity radii, ρ', ρ, ξ', ξ and the parameters $\mathcal{E}, \Gamma, m, \hat{\kappa}$ defined in (A3.47),(A3.7), (A3.11).

This means that the map $\overline{\mathcal{R}}_\mu$ (A3.35) can be defined in a domain $\overline{W} \equiv W(\bar{\kappa}_0, \bar{\rho}_0, \bar{\xi}_0, \bar{\mu}, \vec{a})$ with $\bar{\kappa}_0, \bar{\rho}_0, \bar{\xi}_0$ as in (A3.39) and a $\tilde{\kappa}_0$ verifying (A3.53); moreover such a map satisfy on $\overline{W}$ the bounds (see (A3.50), (A3.39), (A3.38), (A3.49), (A3.53)):

$$\sup_{\overline{W}} |R| < \frac{\rho'}{4}\ , \quad \sup_{\overline{W}} |S| < \frac{\xi'}{4}\ , \quad \sup_{\overline{W}} |\vec{\delta}| < \frac{\xi}{4} \tag{A3.54}$$

Appendix A4 Diffusion sheets. Relative size of the time scales

The following 1),2),3) are the estimates needed to construct the diffusion sheet discussed in the proof of lemma 1'.The 4),5),6) describe the relative size of the various time scales involved in lemma 1'.

1) We write (5.8), if $\vec A\equiv\vec A_s+\vec a$ and $M_s\equiv\partial^2_{\vec A}h_0(\vec A_s,0,0)$ as:

$$M_s\vec a+\Big[\partial_{\vec A}h_0(\vec A_s+\vec a,0,0)-\partial_{\vec A}h_0(\vec A_s,0,0)-M_s\vec a-u\vec\omega_s\Big]=\vec 0 \tag{A4.1}$$

or $\vec a=\vec m(\vec a)$, with $\vec m(\vec a)$ defined as M_s^{-1} applied to the term in square brackets.

Hence, using the holomorphy of h_0 and Cauchy's theorem, we see that if $\tilde\rho\le\bar\rho_0/2$, we can bound the $\vec m$ as:

$$||\vec m||\le|u|\eta_0E_0+4\eta_0E_0\frac{\tilde\rho^2}{\bar\rho_0^2} \tag{A4.2}$$

Fixing $\bar B$ large (*e.g.* $\bar B>20l^2$) and:

$$\tilde u=\frac{1}{\bar B^2(E_0\eta_0\bar\rho_0^{-1})^2},\qquad \tilde\rho=\frac{\bar\rho_0}{\bar B(E_0\eta_0\bar\rho_0^{-1})} \tag{A4.3}$$

we see that $\tilde u<1/4,\tilde\rho<\bar\rho_0/4$ and $4||\vec m||\tilde\rho^{-1}<1$ for $|\vec a|<\tilde\rho$, $|u|<\tilde u$. The constant $\bar B$ can be taken depending only on the dimension l because $E_0\eta_0\bar\rho_0^{-1}\ge l^{-1}$, see item 5) below).

Thus the equation $\vec a=\vec m(\vec a)$ has a solution with $|\vec a|<\tilde\rho<\bar\rho_0/4$ because $4||\vec m||\tilde\rho^{-1}<1$. The implicit functions theorem used here is the essentially obvious argument "on the image of the boundary": it is the same used in appendix A3, see (A3.20).

2) The (5.14) can be similarly written, setting $\vec A=\vec A_{su}+\vec a$ as $\vec a=\vec n(\vec a)$ with $\vec n$ verifying:

$$||\vec n||\le 2l^2E_0\eta_0\Big(\frac{\tilde\rho^2}{\bar\rho_0^2}+\frac{\tilde\kappa^2}{\bar\kappa_0^2}+\frac{\tilde\mu}{\bar\mu_0}\Big),\qquad |\vec a|<\tilde\rho,\ |J|<\tilde\kappa^2,\ |\mu|<\tilde\mu \tag{A4.4}$$

Hence if:

$$\tilde\kappa=\frac{\bar\kappa_0}{\bar B(E_0\eta_0\bar\rho_0^{-1})},\qquad \tilde\mu=\frac{\bar\mu_0}{\bar B^2(E_0\eta_0\bar\rho_0^{-1})^2} \tag{A4.5}$$

we see that, possibly readjusting $\bar B$, it is $4\tilde\rho^{-1}||\vec n||<1$. Therefore, there is a solution with $|\vec a|<\tilde\rho<\bar\rho_0/4$.

3) Considering (5.16) and recalling that $g_{su}\equiv\partial_Jh_0(\vec A_{su},0,0)$ (defined after (5.9)) we see that:

$$\begin{aligned}|u'_{su}|&\equiv|g_{su}^{-1}(g_{su}-g_s)|<2\Gamma_0lE_0\tilde\rho\bar\rho_0^{-1}\\ |u'|&\le\Gamma_0\Big|\partial_Jh_0(\vec A_{su}+\vec a,J,\mu)-\partial_Jh_0(\vec A_{su},0,0)\Big|\le2l\Gamma_0E_0\Big(\frac{\tilde\kappa^2}{\bar\kappa_0^2}+\frac{\tilde\rho}{\bar\rho_0}+\frac{\tilde\mu}{\bar\mu}\Big)\end{aligned} \tag{A4.6}$$

for $|J|<\kappa_0<\tilde\kappa$, $|\mu|<\bar\mu<\tilde\mu$, $|\vec a|<\rho_0<\tilde\rho$.

4) From $|\vec{\omega}_0\cdot\vec{\nu}|\ge C_0^{-1}|\vec{\nu}|^{(l-1)^2}$ and choosing $\vec{\nu}=(1,\ldots,0)$ one finds:

$$C_0|\vec{\omega}_{01}|\ge 1\Rightarrow C_0E_0\ge 1 \tag{A4.7}$$

and a similar argument yields also $\Gamma_0E_0\ge 1$. The relation $E_0C_0>E_0\Gamma_0$ is simply our assumption (5.33).

5) From $[(\partial^2_{\vec{A}}h_0)^{-1}(\partial^2_{\vec{A}}h_0)]_{11}=1$ one finds:

$$1\le\sum_j|[(\partial^2_{\vec{A}}h_0)^{-1}]_{1j}[(\partial^2_{\vec{A}}h_0)]_{j1}|\le E_0\eta_0\rho_0^{-1}l \tag{A4.8}$$

6) From $[(\partial^2_{\vec{A}}h_0)^{-1}\vec{\omega}_0\cdot\vec{\omega}_0]^{-1}[(\partial^2_{\vec{A}}h_0)^{-1}\vec{\omega}_0\cdot\vec{\omega}_0]=1$ one deduces $\vartheta_0\eta_0E_0^2\ge 1$

Hence we can take $B_0=l^{-1}$.

Appendix A5: Divisor bounds

If $k-h=0$ the denominator in (5.31) becomes, using also (5.21),(5.23) and a dimensional bound (recall also that $\vec{\omega}_{su}\equiv\partial_{\vec{A}}h_0(\vec{A}^0(s,u,0,\mu)$ by the definition (5.14) of $\vec{A}^0(s,u,J,\mu)$):

$$\begin{aligned}|\vec{\omega}_0(\vec{A},J,\mu)\cdot\vec{\nu}|=&|\vec{\omega}_{su}\cdot\vec{\nu}+(\vec{\omega}_0(\vec{A},J,\mu)-\vec{\omega}_{su})\cdot\vec{\nu}|\ge\\ \ge&|\vec{\omega}_{su}\cdot\vec{\nu}|(1-|\vec{\omega}_{su}\cdot\vec{\nu}|^{-1}|(\vec{\omega}_0(\vec{A},J,\mu)-\vec{\omega}_{su})\cdot\vec{\nu}|)\ge\\ \ge&C_0^{-1}|\vec{\nu}|^{-\tau}(1-lN_0^{\tau+1}C_0E_0\tilde{\rho}_0/\rho_0)\ge(2C_0)^{-1}|\vec{\nu}|^{\tau}\end{aligned} \tag{A5.1}$$

valid if $|A_i-A^0(s,u,J,\mu)_i|\le\tilde{\rho}_0$ and $\tilde{\rho}_0$ verifies:

$$\tilde{\rho}_0<\rho_0/(2lE_0C_0N_0^{\tau+1}) \tag{A5.2}$$

If $k-h=p=$ is a non zero integer, then, assuming $g_s>0$ for definiteness, and recalling that $\vec{\omega}_{su}$ is real and $g_0(\vec{A},J,\mu)\equiv(1+u')g_{su}$, see (5.16),(5.30):

$$\begin{aligned}&|i\vec{\omega}_0(\vec{A},J,\mu)\cdot\vec{\nu}+g_0(\vec{A},J,\mu)p|\ge|Re\,[i(\vec{\omega}_0(\vec{A},J,\mu)-\vec{\omega}_{su})\cdot\vec{\nu}+g_0(\vec{A},J,\mu)p]|\ge\\ &\ge|p|(g_{su}-E_0\lambda_0-\frac{N_0E_0\tilde{\rho}_0}{\rho_0})\ge|p|\Gamma_0^{-1}(1-\Gamma_0E_0\lambda_0-\frac{N_0E_0\Gamma_0\tilde{\rho}_0}{\rho_0})\ge\frac{|p|}{2\Gamma_0}\end{aligned} \tag{A5.3}$$

as, recalling (5.20):

$$\tilde{\rho}_0<\rho_0(4lE_0C_0N_0^{\tau+1})^{-1},\qquad C_0>\Gamma_0,\qquad 4\lambda_0E_0\Gamma_0<1 \tag{A5.4}$$

so that (5.34) follows (for $|J|<\kappa_0$, *i.e.* if the above quantities make sense).

Note that (A5.3) holds trivially in the case $\vec{\nu}=\vec{0}$; hence.:

$$|i\vec{\omega}_0\cdot\vec{\nu}+g_0(h-k)|^{-1}<2\big(C_0|\vec{\nu}|^\tau+\Gamma_0|h-k|\big) \tag{A5.5}$$

for all $|\vec{\nu}|+|h-k|>0$.

Appendix A6 The equinox precession

Consider the D'Alembert Lagrangian (12.4) and the associated hamiltonian (12.6) $\div$ (12.15). Suppose that the eccentricity of the planet orbit is neglected (*i.e.* that the orbit is taken circular with radius a equal to the major semiaxis of the keplerian ellipse), then the average of the hamiltonian H over the angles γ,φ,ψ and over $\lambda\equiv\omega_T t$ is:

$$\begin{aligned}\bar{H}_p\equiv\frac{M^2}{2J_3}+\eta\Big(&\frac{M^2-L^2}{2J_1}+\frac{3km_Tm_S}{5a}(\frac{R}{a})^2\\&[\frac{K^2}{M^2}(1-\frac{L^2}{M^2})\frac{1}{4}+\frac{1}{2}(1-\frac{K^2}{M^2})\frac{L^2}{M^2}+\frac{1}{4}(1-\frac{L^2}{M^2})]\Big)\end{aligned} \tag{A6.1}$$

with an error of order $O(\eta e)$.

Suppose also that $M=L$ (*i.e.* neglect the non alignment between the planet axis and the angular momentum), so that $\gamma=\bar{\varphi}$. And, furthermore, assume that the hamiltonian H can be replaced by $\bar{H}_p$ for the purpose of evaluating the average motion over many periods of revolutions (see Ch. 5, of [G] §10 $\div$ 12 for a more rigorous treatment). Then the precession angular velocity would be $\partial_K\bar{H}_p$:

$$\dot{\gamma}\equiv\lambda_p^S\equiv-\eta\frac{3km_Tm_S}{5a}(\frac{R}{a})^2\frac{K}{M^2}+O(\eta e^2) \tag{A6.2}$$

Appendix 7 Application to the Earth precession

¿From (A6.2) and neglecting the small variations of the average inclination, i_0, of the planet axis and denoting ω_D is the angular velocity of the daily rotation, and T the period of revolution, the solar precession rate is:

$$\lambda_p^S=-\frac{3}{2}\eta\frac{g_N}{a}(\frac{R}{a})^2\frac{\cos i_0}{R^2\omega_D}=-\frac{6\pi^2\eta}{\omega_DT^2}\cos i_0=-\frac{3}{2}\eta\frac{\omega_T^2}{\omega_D}\cos i_0 \tag{A7.1}$$

having used the third Kepler law to eliminate the gravitational constant (*i.e.* having used that $T=\pi(2a)^{3/2}(2km_S)^{-1/2}$); the fact that the precession is negative is often referred as a *retrograde* precession. This shows also that the period of precession, is $T_p^S = -2\pi/\lambda_p^S = T(\omega_D T\cos i_0)/3\pi\eta$, or since $T = 1.\,\text{year} = 375581495\, s$ and $\eta \simeq 0.003$, $T_p^S \sim 7.94\,10^4$years.

A rough analysis of the lunar precession can be made assuming that the Moon is on the ecliptic and that its orbit is circular. One easily checks that the solar precession analysis can the be applied to the Moon influence and that the lunar precession would be, if m_L, a_L denote respectively the Moon mass and the radius of its orbit:

$$\dot\gamma = \lambda_p^L = -\eta\frac{3km_T m_L}{5a_L}(\frac{R}{a_L})^2\frac{K}{A^2} + O(\eta e^2) = \lambda_p^S(\frac{a}{a_L})^3\frac{m_L}{m_S} \tag{A7.2}$$

so that, taking also into account that the Moon orbit forms an angle i_L with the ecliptic and that the orbit eccentricity e_L is quite large, the total luni-solar precession would be:

$$\lambda_p = \lambda_p^S + \lambda_p^L = \lambda_p^S\left((1+\frac{3}{2}e_T^2) + (\frac{a}{a_L})^3\frac{m_L}{m_S}(1+\frac{3}{2}e_L^2)(1-\frac{3}{2}\sin^2 i_L)\right) \sim 3\lambda_p^S \tag{A7.3}$$

where the eccentricity corrections are obtained by remarking that the above theory with $e = 0$ has taken $(a/r_T)^2 \equiv 1$: but $1+(3/2)e^2$ is the actual average of a^2/r_T^2 over the period with a time evolution based on the Kepler laws; in a similar way one takes into account the inclination of the Moon orbit to found the second correcting factor. Of course one could do also the latter corrections in a less empirical way by using the canonical formalism, but we do not reproduce the details.

Using the data:

$$\begin{aligned} a =&1.496\,10^8\,Km, \qquad a_L = 3.844\,10^5\,Km, \qquad i_L = 5^o\,1' \\ e_L =&0.0549 \\ m_L =&81.3\,m_T, \qquad m_S = 1.99\,10^{30}\,Kg_N, \qquad m_T = 5.98\,10^{24}\,Kg_N \end{aligned} \tag{A7.4}$$

the total rate of lunisolar precession in the above approximation gives, after a small correction for the Moon inclination over the ecliptic is taken into account, $T_p \sim 2.51\,10^4$ years, or a yearly precession of the equinoxes of $\sim 51''$ per (sidereal) year. So that only 1/3 of the luni-solar precession is due to the Sun.

Even assuming that Jupiter gravitated around the Earth on a circular orbit its contribution to the precession would be much smaller (as, with obvious notations, it would be a fraction of the order of $(a/a_J)^3 m_J/m_S$, *i.e.* $O(10^{-5})$ of the solar precession).

A more fundamental formula is obtained if the Earth is not supposed an homogeneous ellipsoid, but is supposed only to be rigid. In this case one finds that (A7.3) remains the same if η is defined in terms of the inertia moments as $\eta = (2J_3 - J_1 - J_2)/2J_3$; the analysis is unchanged and the constant η thus defined is called the *mechanical flattening* and it is independent on the Earth shape and mass distribution, as long as it can be supposed rigid: hence it is this quantity that can be really deduced from the observed rate of the precession of the equinoxes, and it is $\eta = 1/304 = 0.0329$, (while the observed polar radius of the

Earth is by 0.0035 shorter than the equatorial radius, showing that the ellipsoidal model is, to some extent, not satisfactory).

The above calculation, due to D' Alembert (who did not use the canonical formalism) (see [L]: vol II, book V, §6, fourth formula to the last, where $l = \lambda_p$, $m = \omega_T$, $n = \omega_D$, $h = i_0$, $\lambda = (a/a_L)^3(m_L/m_S)$, and e_T, e_L, i_L are neglected) was in fact *used* to determine η from the known precession rate, in terms of the masses of the Sun and of the Moon.

Appendix 8 Trigonometry of the Andoyer–Deprit angles

We refer here to fig. 4.11, 4.12, 4.10 of [G, p. 321÷323] and to the well known spherical trigonometry identities:

$$\begin{aligned}\frac{\sin A}{\sin\alpha} =& \frac{\sin B}{\sin\beta} = \frac{\sin C}{\sin\gamma}\\ \cos A =& \cos B\cos C + \sin B\sin C\cos\alpha\\ \sin C\cos\beta =& \cos B\sin A - \sin B\cos A\cos\gamma\\ \cos A\cos\gamma =& \sin A\cot B - \sin\gamma\cot\beta\end{aligned} \tag{A8.1}$$

the inversion can be actually performed via the relations:

$$\begin{aligned}\cos\delta =& \frac{K_z}{A}, \qquad \cos\vartheta = \frac{L}{A}\\ \cot(\bar\varphi-\gamma) =& (\cos\varphi\cos\delta + \sin\delta\cot\vartheta)/\sin\varphi\\ \cot(\bar\psi-\psi) =& (-\cos\varphi\cos\gamma + \sin\varphi\cot\delta)/\sin\vartheta\\ \sin\bar\vartheta =& \sin\vartheta\frac{\sin\varphi}{\sin(\bar\varphi-\gamma)}\end{aligned} \tag{A8.2}$$

which follow immediately from the definitions, see [G, p. 323], and the result is, after some algebra, (12.10).

Appendix 9: Determinants, wronskians, Jacobi's map

1) Consider (see (5.92) with $(I,\varphi) = (0,0)$):

$$h_2(\vec A) \equiv \frac{[\omega B + h(\tilde A)]^2}{2E}\,, \qquad \vec A \equiv (B,\tilde A) \in R\times R^{l-2}\,, \qquad E \neq 0 \tag{A9.1}$$

Then, setting $\tilde\omega \equiv \partial_{\tilde A} h$ and assuming $\tilde\omega_i \ne 0$ for $i = 1, ..., l-2$, one has:

$$\det \partial^2_{\vec A} h_2 \equiv \det \frac{1}{E}\begin{pmatrix} \omega^2 & \omega\tilde\omega \\ \omega\tilde\omega^T & (h_2/E)\partial^2_{\tilde A} h_2 + \tilde\omega^T \times \tilde\omega \end{pmatrix} = \frac{\omega^2}{E}\Big(\frac{h_2}{E}\Big)^{l-2} \det \partial^2_{\tilde A} h_2 \qquad (A9.2)$$

as it follows by multiplying, for $i = 1, ..., l-2$, the first row by ω_i/ω and subtracting it to the i-th following row: this proves (5.93). Furthermore the following general identity is valid for any $n \times n$ matrix H and row vector $\vec\omega$:

$$\det \begin{pmatrix} 0 & \vec\omega \\ \vec\omega^T & H \end{pmatrix} = -(\vec\omega \cdot H^{-1}\vec\omega) \det H \qquad (A9.3)$$

where if H is not invertible the right hand side has to be interpreted as $-(\vec\omega \cdot \tilde H \vec\omega)$, with $\tilde H_{ij} \equiv (i,j)^{th}-$cofactor of $H \equiv (-1)^{i+j} \times$ the determinant of the matrix obtained by deleting the i^{th} row and the j^{th} column.

2) The standard pendulum: $P_0 = I^2/2J_0 + g_0^2 J_0(\cos\varphi - 1)$ has a separatrix motion $t \to \varphi^0(t)$ which is easily computable. One finds, starting at $\varphi = \pi$ at $t = 0$ on finds:

$$\begin{aligned} \sin\varphi^0(t)/2 &= 1/\cosh g_0 t, \qquad & \sin\varphi^0(t) &= 2\sinh g_0 t\,(\cosh g_0 t)^{-2} \\ \cos\varphi^0(t)/2 &= \tanh g_0 t, \qquad & \cos\varphi^0(t) &= 1 - 2\,(\cosh g_0 t)^{-2} \end{aligned} \qquad (A9.4)$$

A further elementary discussion of the pendulum quadratures near $E = 0$, allows us to find the E derivatives of the separatrix motion and leads to:

$$\begin{aligned} I^0 &= \frac{-2g_0 J_0}{\cosh g_0 t} = -2g_0 J_0 \sin\frac{\varphi^0}{2} \qquad & \partial_E I^0 &= J_0 (I^0)^{-1}\Big(1 + J_0 g_0^2 (\partial_E \varphi^0)\sin\varphi^0\Big) \\ \varphi^0 &= 4\,\mathrm{arctg}\, e^{-g_0 t} \qquad & \partial_E \varphi^0 &= \frac{-1}{8g_0^2 J_0}\big(2g_0 t + \sinh 2g_0 t\big)\sin\frac{\varphi^0}{2} \end{aligned} \qquad (A9.5)$$

exhibiting the analyticity properties in the complex t plane that are useful in discussing the size of the homoclinic angles. The (A9.5) allows us to compute the wronskian matrix of the above separatrices, *i.e.* the solution of the pendulum equation, linearized on the separatices:

$$\dot{\overline W} = \overline L(t)\overline W, \quad \overline W(0) = 1, \quad \overline L(t) = \begin{pmatrix} 0 & J_0 g_0^2 \cos\varphi^0(t) \\ J_0^{-1} & 0 \end{pmatrix} \qquad (A9.6)$$

and we get:

$$\overline W(t) = \begin{pmatrix} \partial_E I^0/c_1 & \dot I^0/c_2 \\ \partial_E \varphi^0/c_1 & \dot\varphi^0/c_2 \end{pmatrix}, \qquad \begin{aligned} c_1 &= \partial_E I^0(0) \\ c_2 &= \dot\varphi^0(0) \end{aligned} \qquad (A9.7)$$

where the E derivative is computed by imagining motions close to the separatrix (which has energy $E=0$) and with the same initial $\varphi=\pi$. This becomes:

$$\overline{W}(t)=\begin{pmatrix}(1-\frac{\mathcal{F}}{4}\frac{\sinh g_0t}{\cosh^2 g_0t})\cosh g_0t & -J_0g_0\frac{\sinh g_0t}{\cosh^2 g_0t}\\ \frac{\mathcal{F}}{4J_0g_0} & \frac{1}{\cosh g_0t}\end{pmatrix},\qquad \mathcal{F}\equiv\frac{2g_0t+\sinh 2g_0t}{\cosh g_0t} \tag{A9.8}$$

We are also interested in the matrices $U^s(t),U^u(t)$ of §6. If we write $\varphi=S(p_0,0)$ and if $\bar p$ is the value of p_0 such that $S(\bar p,0)=\pi$, we see that:

$$\begin{aligned}\varphi^0(t)=&S(\bar pe^{-g_0t},0)=4\,\mathrm{arctg}\,e^{-g_0t}\\ I^0(t)=&J_0(\bar pe^{-g_0t},0)=J_0\dot\varphi^0(t)=-J_0\bar pg_0e^{-g_0t}\partial_pS(\bar pe^{-g_0t},0)\end{aligned} \tag{A9.9}$$

and it can be seen that $\bar p=(32J_0g_0)^{1/2}$; so that, noting that at $q_0=0,E=0$ it is $\partial_q=p_0g_0\partial_E$ at $p_0=const$:

$$U^s(0)=\begin{pmatrix}\partial_pI & \partial_qI\\ \partial_p\varphi & \partial_q\varphi\end{pmatrix}=\begin{pmatrix}0 & -\bar p/2\\ 2/\bar p & 0\end{pmatrix} \tag{A9.10}$$

and, similarly:

$$U^u(0)=\begin{pmatrix}\bar p/2 & 0\\ 0 & 2/\bar p\end{pmatrix} \tag{A9.11}$$

3) the theory of the jacobian elliptic functions shows how to perform a complete calculation of the functions R,S, see [GR] (8.198),(8.153), (8.146), (8.128), (8.197). The result (a celebrated theorem by Jacobi, and a strongly instructive exercise in Mechanics) is reported here for completeness and is discussed in terms of the pendulum energy:

$$\frac{J_0\dot\varphi^2}{2}+J_0g_0^2(1-\cos\varphi)=E \tag{A9.12}$$

where the origin in φ is set at the stable equilibrium, to adhere to the notations in the theory of elliptic functions.

Setting $u=t(E/2J_0)^{1/2}\equiv\varepsilon^{1/2}g_0t$, $k^2=2J_0g_0^2/E=\varepsilon^{-1}$ where ε is the *dimensionless* energy so that $\varepsilon=1$ is the separatrix, and:

$$\mathbf{K}(k)=\int_0^{\pi/2}\frac{d\alpha}{(1-k^2\sin^2\alpha)^{1/2}} \tag{A9.13}$$

One sets, using whenever possible, the standard notations for the jacobian elliptic integrals except for $x(.)$ which is usually denoted $q(.)$ but which we would confuse with the canonical variable q of lemma 0:

$$\begin{aligned}k'=&(1-k^2)^{1/2},\qquad g_J=g_0\frac{\pi}{2k\mathbf{K}(k')},\qquad \lambda\equiv\frac{1}{2}\frac{1-k^{1/2}}{1+k^{1/2}}\\ x(k')=&e^{-\pi\mathbf{K}(k)/\mathbf{K}(k')}=\lambda+2\lambda^5+15\lambda^9+150\lambda^{13}+1707\lambda^{17}+\dots\end{aligned} \tag{A9.14}$$

In terms of the above conventions we have, directly from the definitions (*i.e.* from the equations of motion):

$$\begin{aligned} I(t) =& J_0\dot\varphi = -2J_0 g_0\varepsilon^{1/2}\,\mathrm{dn}\,(u,k)\\ \varphi(t) =& 2\,\mathrm{am}\,(tg_0\varepsilon^{1/2}) \end{aligned} \tag{A9.15}$$

which yield, changing at this point the origin for φ to the unstable point to conform with our notations (*i.e.* obtaining $\varphi(t) = 2(\,\mathrm{am}\,(tg_0\varepsilon^{1/2}) + \pi/2)$):

$$R = -2J_0 g_0\varepsilon^{1/2}\frac{\mathrm{dn}\,(iu,k')}{\mathrm{cn}\,(iu,k')},\qquad \sin\frac{S}{2} = \frac{1}{\mathrm{cn}\,(iu,k')},\qquad \cos\frac{S}{2} = i\frac{\mathrm{sn}\,(iu,k')}{\mathrm{cn}\,(iu,k')} \tag{A9.16}$$

which, using also $R(p,q) = g_J J_0(-p\partial_p + q\partial_q)S(p,q)$ to evaluate S from R, imply immediately the Jacobi map:

$$\begin{aligned} R(p,q) =& -2J_0 g_J\Big[\frac{p}{1+p^2} + \frac{q}{1+q^2} - \sum_{n=1}^{\infty}(-1)^n\frac{1+x^{2n-1}}{1-x^{2n-1}}(p^{2n-1}+q^{2n-1})\Big]\\ S(p,q) =& 2\left[\,\mathrm{arctg}\,p - \,\mathrm{arctg}\,q - \sum_{n=1}^{\infty}(-1)^n\frac{1+x^{2n-1}}{1-x^{2n-1}}\frac{(p^{2n-1}-q^{2n-1})}{2n-1}\right]\\ \sin\frac{S(p,q)}{2} =& \frac{\pi}{2k\mathbf{K}(k')}\Big[\frac{p}{1+p^2} - \frac{q}{1+q^2} - \sum_{n=1}^{\infty}(-1)^n\frac{1-x^{2n-1}}{1+x^{2n-1}}(p^{2n-1}-q^{2n-1})\Big]\\ \cos\frac{S(p,q)}{2} =& \frac{-\pi}{4k\mathbf{K}(k')}\Big[\frac{1-p^2}{1+p^2} + \frac{1-q^2}{1+q^2} + 2\sum_{n=1}^{\infty}(-1)^n\frac{1-x^{2n}}{1+x^{2n}}(p^{2n}+q^{2n})\Big] \end{aligned} \tag{A9.17}$$

with $x \equiv pq$.

Note that g_J depends on k, and so do k', k: hence the coefficients of the first and of the last two of (A9.17) are also functions of $x = pq$.

Furthermore the (dimensionless) energy becomes a function of $x = pq$ defined by inverting the map:

$$\varepsilon \to x(k') \equiv x((1-\varepsilon^{-1})^{1/2}) \tag{A9.18}$$

and the point corresponding to $\varphi = \pi$ and to a dimensionless energy ε, has coordinates:

$$p \equiv 1,\quad q \equiv x(k') \tag{A9.19}$$

(a rearrangement of the above series showing the convergence for $p = 1$ and $|x| < 1$ is exhibited below).

However the derivative of the energy with respect to pq is not proportional to g, defined above: this would mean that the map (A9.17) between I,φ and p,q would be a canonical map. The above implies, nevertheless, that the jacobian determinant $D(x) = \det(\partial(p,q)/\partial(p_J,q_J))$ with respect to the (yet unknown) canonical coordinates of lemma

0 is a function of the x variable identical to the determinant of the jacobian matrix $\partial(p,q)/\partial(I,\varphi) \equiv D(x)$ (which has to be a function of x and can be computed from (A9.17)).

To be precise, (p_J, q_J) denote here *some* canonical variables which transform the pendulum hamiltonian into a function of the product $(p_J q_J)$ (but clearly there is a large ambiguity in the construction of such variables and therefore (p_J, q_J) need not coincide with the variables constructed in lemma 0). It is easy to check that with the above definitions it is:

$$g_J = 2J_0 g_0^2 \frac{d\varepsilon}{dx} D \tag{A9.20}$$

which (using (8.197) of [GR]) yields at once $D(0) = (32 J_0 g_0)^{-1}$.

Setting $p = p_J$, $q = f(x) q_J$ we see that the jacobian determinant $\partial(p,q)/\partial(p_J,q_J) \equiv f + (1 - xf'/f)^{-1} x f'$. Hence, if f is the solution of the differential equation:

$$f + \frac{xf'}{1 - xf'/f} = D(x), \qquad f(0) = 1/(32 J_0 g_0)^{1/2} \tag{A9.21}$$

regular at $x = 0$, then the map $(I,\varphi) \to (p_J, q_J)$ is canonical; in fact, one finds: $f(x) = x(\int_0^x D^{-1}(y)dy)^{-1}$.

Therefore a canonical Jacobi map is, in terms of the Jacobi map (A9.17), simply obtained by substituting p, q with p_J, $q_J F(p_J q_J)$, where $x_J \equiv p_J q_J$ and $F(x_J)$ is implicitly defined by $x = x_J f(x) = x_J F(x_J)$, *i.e.* $f(x) = F(x_J)$.

In lemma 0 we did not use dimensionless p, q coordinates: it is easy to check that with the conventions of §5, and appendix 3 the p, q coordinates above are related to canonical (dimensional) ones, which we henceforth denote p_0, q_0, coinciding with the coordinates constructed by lemma 0 up to first order in p_0, q_0, by:

$$p = \frac{-p_0}{(32 J_0 g_0)^{1/2}}, \qquad q = \frac{-q_0}{(32 J_0 g_0)^{1/2}} \tag{A9.22}$$

so that p_0, q_0 have the dimension of the square root of an action.

The above (A9.17) is written in the form in which it is easily recognized in the elliptic functions tables. However, once derived it, it can be rewritten in the following form:

$$\begin{aligned}
R(p,q) =& -4J_0 g\left[\sum_{m=0}^{\infty}\left(\tfrac{x^m p}{1+x^{2m}p^2} + \tfrac{x^m q}{1+x^{2m}q^2}\right)\right]\\
S(p,q) =& 4\left[\sum_{m=0}^{\infty}\left(\operatorname{arctg} x^m p - \operatorname{arctg} x^m q\right)\right]\\
\sin\tfrac{S(p,q)}{2} =& \tfrac{\pi}{k\mathbf{K}(k')}\left[\sum_{m=0}^{\infty}(-1)^m\left(\tfrac{x^m p}{1+x^{2m}p^2} - \tfrac{x^m q}{1+x^{2m}q^2}\right)\right]\\
\cos\tfrac{S(p,q)}{2} =& \tfrac{-\pi}{2k\mathbf{K}(k')}\left[1 - 2\sum_{m=0}^{\infty}(-1)^m\left(\tfrac{x^{2m}p^2}{1+x^{2m}p^2} + \tfrac{x^{2m}q^2}{1+x^{2m}q^2}\right)\right]
\end{aligned} \tag{A9.23}$$

exhibiting some of the properties of the Jacobi map in a better way.

5) In general the wronskian W in (6.14) [and hence the solution of (6.15)] can be computed quite explicitly. It is however convenient, for computational purposes, to rearrange its rows and columns by writing them in the order $(I,\varphi,\vec A,\vec\alpha)$ instead of $(I,\vec A,\varphi,\vec\alpha)$ used in §6.

Consider the equation for I' $P(\vec A,I',\varphi',0)=e$ and let $I'=i(\vec A,e,\varphi')$ be a solution. Then define the functions $\Phi(t,e)$ and $I(t,e)=i(\vec A,e,\Phi(t,e))$, with $\vec A,e$ regarded as parameters, solutions of:

$$\dot\Phi=\partial_I P(\vec A,i(\vec A,e,\Phi,0),\Phi,0),\qquad \Phi(0,e)=\varphi,\qquad I(0,e)=i(\vec A,e,\varphi)\tag{A9.24}$$

The above functions $t\to(\Phi(t,e),I(t,e))$ will be a family of motions of the pendulum with energy e close to the separatrix motion, $(e=0)$. The functions:

$$t\to S_1(t)=\begin{pmatrix}\partial_e I(t)\\ \partial_e\Phi(t)\end{pmatrix}_{e=0}\frac{1}{\partial_e I(0)_{e=0}},\qquad t\to S_2(t)=\begin{pmatrix}\dot I(t)\\ \dot\Phi(t)\end{pmatrix}_{e=0}\frac{1}{\dot\Phi(0)_{e=0}}\tag{A9.25}$$

verify the equations of motion linearized around the separatrix motion:

$$\dot{\vec S}_j=\overline L(t)\vec S_j,\qquad \overline L(t)=\begin{pmatrix}-\partial^2_{I\varphi}H_0 & -\partial^2_{\varphi\varphi}H_0\\ \partial^2_{II}H_0 & \partial^2_{I\varphi}H_0\end{pmatrix}\tag{A9.26}$$

here $H_0=h(\vec A,0)+P(\vec A,I,\varphi,0)$ and all the derivatives are evaluated at the point $X^0_{su}(t)$ and at $\mu=0$, see also (6.12); note that $\partial_e I(0)_{e=0}=\partial_e i(\vec A,0,\varphi)=1/\partial_I P(i(\vec A,0,\varphi))>0$.

Furthermore $S_{12}(0)=S_{21}(0)=0$: in fact $\Phi(0)\equiv\varphi$ so that $S_{12}(0)\equiv0$, and $\dot\Phi(0)=0$ (because $\dot I(0,0)=\partial_\varphi i(\vec A,0,\varphi)=0$ by our choice of φ as the point where i has a maximum), hence $S_{21}(0)=0$.

Hence the matrix $\overline W(t)=(S_1(t),S_2(t))$ verifies the equation:

$$\dot{\overline W}=\overline L(t)\overline W,\qquad \overline W(0)=1\tag{A9.27}$$

and we realize that $\overline W$ is the wronskian for the separatrix motion of the pendulum $h(\vec A)+P(\vec A,i,\varphi,0)$, with $\vec A=\vec A_{su}$. For the standard pendulum it is given by (A9.8).

The wronskian $W(t)$ can be expressed in terms of $\overline W$ as follows:

$$W(t)=\begin{pmatrix}\overline W(t) & x(t) & 0\\ 0 & 1 & 0\\ y(t)^T & H(t)+R(t) & 1\end{pmatrix}\tag{A9.28}$$

where x,y are $2\times l-1$ matrices and R,H are $(l-1)\times(l-1)$ matrices: we shall think x,y as rows of 2-vectors $x=(\vec x_1,\ldots,\vec x_{l-1})$, $y=(\vec y_1,\ldots,\vec y_{l-1})$, with $\vec x_i,\vec y_i$ being column 2-vectors; or, alternatively, a one column of two $l-1$ vectors: $x=\begin{pmatrix}\vec x^1\\ \vec x^2\end{pmatrix}$, $y=\begin{pmatrix}\vec y^1\\ \vec y^2\end{pmatrix}$; by the matrix multiplication rules we have:

$$(y^Tx)_{ij}=\vec y_i\cdot\vec x_j,\quad(i,j=1,...,l-1),\quad(yx^T)_{ij}=\vec y^{(i)}\cdot\vec x^{(j)},\quad(i,j=1,2)\tag{A9.29}$$

The conditions that (A9.28) verifies $\dot W = LW$, $W(0) = 1$ are:

$$\begin{aligned}
\dot{\vec x}_i =&\overline{L}\vec x_i + \vec\xi_i && \vec x_I(0) = 0\\
\dot{\vec y}_i =&\sigma\overline{W}^{-1}\vec\xi_i && \vec y_I(0) = 0\\
\dot R =&(\sigma\xi)^T x && R(0) = 0\\
\dot H =&(\partial^2_{\vec A\vec A}H_0)(\vec A, I(t), \varphi(t), 0) \equiv M && H(0) = 0\\
\vec\xi_i(t) =& \begin{pmatrix} -\partial_{A_i\varphi}H_0(\vec A, I(t), \varphi(t), 0)\\ \partial_{A_iI}H_0(\vec A, I(t), \varphi(t), 0)\end{pmatrix} && \sigma \equiv \begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}
\end{aligned} \tag{A9.30}$$

because L can be written:

$$L = \begin{pmatrix} \overline{L} & \xi & 0\\ 0 & 0 & 0\\ (\sigma\xi)^T & M & 0\end{pmatrix} \tag{A9.31}$$

and $\sigma\overline{W}^T\sigma = -\overline{W}^{-1}$, as $\det\overline{W} \equiv 1$, and $\sigma^2 = -1$. Hence:

$$\begin{aligned}
\vec x_i(t) =&\overline{W}(t)\int_0^t \overline{W}(\tau)^{-1}\vec\xi_i(\tau)d\tau \equiv \overline{W}(t)\vec z_i(t)\\
\vec y_i(t) =&\sigma\int_0^t \overline{W}(\tau)^{-1}\vec\xi_i(\tau)d\tau \equiv \sigma\vec z_i(t)\\
R_{ij}(t) =& \int_0^t (\sigma\vec\xi_i(\tau))^T\cdot\vec x_j(\tau)d\tau, \qquad H(t) = \int_0^t M(\tau)d\tau
\end{aligned} \tag{A9.32}$$

where $\vec z_i(t)$ is defined by the second equality in the first line of (A9.32).

It is important to find the asymptotic expansion of $\overline{W}, x, y, z$ as $t \to +\infty$. It can be derived from the expansion of $S_1(t)$ and from:

$$\xi_i(t) = \overline{\xi}^1_i e^{-gt} + \overline{\xi}^2_i e^{-2gt} + \ldots, \qquad S_i(t) = \overline{S}^1_i e^{-gt} + \overline{S}^2_i e^{-2gt} + \ldots \tag{A9.33}$$

where g is the Lyapunov exponent of the selected unstable equilibrium point of the pendulum (to be consistent with §3 we should replace everywhere below g with κg). The expansions of S_1 and S_2 are deduced from the corresponding quadratures; for instance that of S_1 is derived from the quadrature:

$$t = \int_\varphi^{\Phi(t,e)} (\partial_I P)(\vec A, i(\vec A, e, \psi, 0), \psi, 0)^{-1}d\psi \tag{A9.34}$$

by differentiating with respect to e and setting $e = 0$.

For suitably chosen constants $\gamma, \gamma_0, \gamma', \gamma'_0$ one easily finds:

$$\overline{W}(t) = \begin{pmatrix} \gamma e^{gt} + \sigma_{11}(t) & S_{21}(t)\\ \gamma' e^{gt} + \sigma_{12}(t) & S_{22}(t)\end{pmatrix} \tag{A9.35}$$

where σ_{11}, σ_{12} converge to γ_0, γ_0' at speed $O(te^{-gt})$.

Therefore, see (A9.32):

$$\begin{aligned}
\overline{W}(t)^{-1} &= \begin{pmatrix} S_{22}(t) & -S_{21}(t) \\ -\gamma' e^{gt} - \sigma_{12}(t) & \gamma e^{gt} + \sigma_{11}(t) \end{pmatrix} \\
\vec{z}_i(t) &= \int_0^t \overline{W}(\tau)^{-1} \vec{\xi}_i(\tau) d\tau = \overline{\zeta}_i t \binom{0}{1} + \begin{pmatrix} \zeta_i^1 \\ \zeta_i^2 \end{pmatrix} + \begin{pmatrix} \zeta_i^{r1}(t) \\ \zeta_i^{r2}(t) \end{pmatrix}
\end{aligned} \tag{A9.36}$$

where, setting $\hat{\xi}_i = \bar{\xi}_i^2 e^{-gt} + \bar{\xi}_i^3 e^{-2gt} + \ldots$ and $\Delta(t) = -\gamma' \hat{\xi}_{i1}(t) + \gamma \hat{\xi}_{i2}(t)$:

$$\begin{aligned}
&\overline{\zeta}_i = -\gamma' \, \overline{\xi}_{i1} + \gamma \overline{\xi}_{i2} \\
&\vec{\zeta}_i = \begin{pmatrix} \zeta_i^1 \\ \zeta_i^2 \end{pmatrix} = \int_0^{+\infty} d\tau \begin{pmatrix} \xi_{i1}(\tau) S_{22}(\tau) - \xi_{i2}(\tau) S_{21}(\tau) \\ \Delta(\tau) - \xi_{i1} \sigma_{12}(\tau) + \xi_{i2}(\tau) \sigma_{22}(\tau) \end{pmatrix} \\
&\vec{\zeta}^{\,r}_i(t) = \int_t^{\infty} d\tau \begin{pmatrix} \xi_{i1}(\tau) S_{22}(\tau) - \xi_{i2}(\tau) S_{21}(\tau) \\ \Delta(\tau) - \xi_{i1}(\tau) \sigma_{12}(\tau) + \xi_{i2}(\tau) \sigma_{22}(\tau) \end{pmatrix} = \begin{pmatrix} O(e^{-2gt}) \\ O(e^{-gt}) \end{pmatrix}
\end{aligned} \tag{A9.37}$$

and we shall denote $\vec{\bar{\zeta}}$ the $(l-1)$-vector with components $\bar{\zeta}_i$, $i = 1, \ldots, l-1$, thinking also, see (A9.29):

$$\vec{\zeta} = \begin{pmatrix} \vec{\zeta}^{\,1} \\ \vec{\zeta}^{\,2} \end{pmatrix}, \quad \vec{\xi} = \begin{pmatrix} \vec{\xi}^1 \\ \vec{\xi}^2 \end{pmatrix}, \quad etc \tag{A9.38}$$

Therefore the complete wronskians $W(t)$ and $W(t)^{-1}$ are:

$$\begin{pmatrix} \overline{W}(t) & \overline{W}(t)\vec{z}(t) & 0 \\ 0 & 1 & 0 \\ (\sigma \vec{z}(t))^T & H+R & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} \overline{W}(t)^{-1} & -\vec{z}(t) & 0 \\ 0 & 1 & 0 \\ -(\sigma \vec{z})^T \overline{W}^{-1} & -H - R + (\sigma \vec{z})^T \vec{z} & 1 \end{pmatrix} \tag{A9.39}$$

respectively, the inverse matrix being computed by the general formula:

$$W = \begin{pmatrix} \overline{W} & B & 0 \\ 0 & 1 & 0 \\ C^T & K & 1 \end{pmatrix} \longleftrightarrow W^{-1} = \begin{pmatrix} \overline{W}^{-1} & -\overline{W}^{-1} B & 0 \\ 0 & 1 & 0 \\ -C^T \overline{W}^{-1} & -K + C^T \overline{W}^{-1} B & 1 \end{pmatrix} \tag{A9.40)}$$

and one finds for $W(t)W(\tau)^{-1}$, shortening $(H(t) + R(t))$ into $(H+R)_t$:

$$\begin{pmatrix} \overline{W}(t)\overline{W}(\tau)^{-1} & \overline{W}(t)(\vec{z}(t) - \vec{z}(\tau)) & 0 \\ 0 & 1 & 0 \\ (\sigma(\vec{z}(t) - \vec{z}(\tau))^T \overline{W}^{-1}(\tau) & (H+R)_t - (H+R)_\tau - (\sigma(\vec{z}(t) - \vec{z}(\tau))^T \vec{z}(\tau)) & 1 \end{pmatrix} \tag{A9.41}$$

Appendix 10 High order perturbation theory and averaging

We consider the hamiltonian, see (11.11):

$$H = \eta^{-1/2}\omega B + h(\eta^{1/2}A) + \frac{I^2}{2J_0} + J_0 g_0^2(\cos\varphi - 1) + \beta(F + \mu f) \tag{A10.1}$$

where $h \equiv$ constant $+\eta^{1/2}\bar\omega_2 A + \eta(A^2/2J(A))$; J_0 depends on A and I while g_0 depends on $A, I, z \equiv (\cos\varphi - 1)$; possibly such functions depend also on η, μ. *The parameter β is an auxiliary complex parameter that will be eventually set equal to* 1.

We assume that each function in (A10.1) is holomorphic for $|\eta| < \bar\eta_0, |\mu| < \bar\mu_0$ and in a domain obtained by complexifying, by an amount $\bar\rho_0 = \eta^{-1/2}\bar\rho$ ($\bar\rho$ being some given positive action) the actions and by an amount $\bar\xi_0$ the angles, around the real domain:

$$\Upsilon = \big\{I = 0,\ A \in \eta^{-1/2}[\bar A, \tilde A]\big\} \tag{A10.2}$$

for some $\bar A, \tilde A,\ 0 < \bar A < \tilde A$, and we call the latter domain $\Upsilon(\bar\rho_0, \bar\xi_0)$. As usual we suppose for simplicity that all $\xi_.$ variables are < 1 (no loss of generality). The functions will be supposed to verify uniform bounds (with respect to η) in the above holomorphy domain.

Note that the "large" size, $O(\eta^{-1/2})$ of the analyticity domains in the A, I variables simply reflects the assumptions in §10 that the dependence of J, J_0, F, f on the A, I is via $\eta^{1/2}A, \eta^{1/2}I$. This assumption implies, together with the boundedness assumption, that the rotation vector $\vec\omega = (\eta^{-1/2}\omega, \partial_A h(\eta^{1/2}A))$ is a vector $\vec\omega = (\omega_1, \omega_2)$ with ω_2 varying between $\eta^{1/2}(\bar\omega, \tilde\omega)$ as A varies in the interval $[\bar A, \tilde A]$: which we assume to exclude the origin (this takes into account the fact that (A10.1) has to come from (11.3) with $\bar\omega_2 > 0$).

In the following discussion the assumption that the functions F, f depend on ηz, *i.e.* that they have a large analyticity domain in the z variable is not necessary (although it is part of the assumptions of lemma 5, to simplify the formulation).

We also assume that the functions f, F, whose Fourier transforms will be denoted by affixing a label $\vec\nu, n$, are trigonometric polynomials in the $\vec\alpha \equiv (\alpha_1, \alpha_2) \equiv (\lambda, \alpha)$ variables: *i.e.* their Fourier transforms vanish if $|\vec\nu| > N$, for some $N > 0$: this hypothesis , as we shall see, is not really necessary and is done only for simplicity.

Finally we suppose that F has zero average in the $\vec\alpha$ variables and that it contains only harmonics multiples of a fixed $\vec\nu_0$ (*e.g.* $\vec\nu_0 = (1,1)$) which is a "fast mode", (*i.e.* $\nu_{01} \neq 0$, see (11.9)).

We shall show that, if c is a large enough constant, *for any $x > 0$* and any $0 < \sigma < 1/2$ there exist constants $\kappa, \rho, \xi > 0$ such that, for all $\eta > 0$ small enough, in the domain:

$$\begin{aligned}\Omega(\kappa, \rho, \xi, \bar\mu) = \big(\{&|p|, |q| < \kappa, |\operatorname{Im} A| < \rho,\\ &\operatorname{Re} A \in \eta^{-1/2}[\bar A, \tilde A],\ |\operatorname{Im}\psi_j| < \xi, |\mu| < \bar\mu\big\}\end{aligned} \qquad \bar\mu = \eta^c \tag{A10.3}$$

there exist functions $\delta, R, S, \Lambda, \Theta, \Xi, \Delta$ holomorphic in (A10.3) defining a map:

$$\begin{aligned} I &= R(p,a,q) + \Lambda(p,a,q,\vec\psi), \qquad A = a + \Xi(p,a,q,\vec\psi), \qquad \lambda = \psi_1\\ \varphi &= S(p,a,q) + \Theta(p,a,q,\vec\psi), \qquad \alpha = \psi_2 + \delta(p,a,q,\vec\psi) + \Delta(p,a,q,\vec\psi)\end{aligned} \tag{A10.4}$$

with the μ, η, β dependence of the above functions not explicitly shown, and with $\Delta, \Xi, \Lambda, \Theta$ of order $\beta\sqrt{\eta}$. The functions are uniformly bounded in $|\beta| < B^*\eta^{-\sigma}$. And the map is canonical and changes the hamiltonian (A10.1) (up to a trivial constant) into:

$$H = \eta^{-1/2}\omega b + \eta^{1/2}\bar{\omega}_2 a + \frac{\eta a^2}{2\bar{J}(a)} + \bar{g}(a, pq) + \beta^{2x}\eta^x \bar{f}(p, a, q, \vec{\psi}) \tag{A10.5}$$

with $\bar{J}, \bar{g}, \bar{f}$ depending on η, μ, β. The size of ρ, ξ, κ and how small should η be depend on x can be easily deduced from the proof below. The constant c can be taken a suitably large number (*e.g.* $c = 10$ is proposed in the proof).

Proof: We begin by performing the (generalization of the) Jacobi map of lemma 0, §5. This gives a canonical map defined on $W(\tilde{\kappa}_0, \tilde{\rho}_0, \tilde{\xi}_0, \tilde{\mu}_0)$ for some $\tilde{\kappa}_0, \tilde{\rho}_0 \equiv \eta^{-1/2}\tilde{\rho}, \tilde{\xi}_0, \tilde{\mu}_0 > 0$ (where we use the notations of §5) like (A10.4) with $\Delta, \Xi, \Theta, \Lambda = 0$, and we take for simplicity $\tilde{\kappa}_0^2 \equiv \tilde{\rho}$. The parameters $\tilde{\kappa}_0, \tilde{\rho}_0, \ldots$ are larger than the corresponding barred ones in (A10.3).

The map transforms the original hamiltonian into:

$$\eta^{-1/2}\omega B + h(\eta^{1/2}A) + \tilde{g}(A, pq) + \beta(\tilde{F} + \mu\tilde{f}) \tag{A10.6}$$

where $\tilde{F}$, $\tilde{f}$ are evaluated at $(p, q, A, \vec{\alpha})$. The various functions depend also on η, μ and are analytic in the above domain and in $\eta^{1/2}$, for η small enough.

The variables in (A10.6) should be primed as they are different from the original ones, but we do not do so.

The main property of $\tilde{F}$ is that in the new variables it is still a trigonometric polynomial with the same modes in $\vec{\alpha}$ as F (*i.e.* only finitely many multiples of a given fast mode), hence with zero average. This happens because the variation of $\vec{\alpha}$ in the transformation of lemma 0, §5, is $\vec{\alpha}$-independent, see (5.3).

As a preliminary step we apply perturbation theory to remove the $\tilde{F}$ by performing a perturbation expansion to first order in the auxiliary variable β. This leads to the Hamilton Jacobi equation ($J \equiv pq$):

$$\eta^{-1/2}\omega\partial_\lambda\Phi + \eta^{1/2}\omega_2\,\partial_\alpha\Phi + \partial_J\tilde{g}\,(q_0\partial_{q_0}\Phi - p\partial_p\Phi) + \tilde{F} = 0 \tag{A10.7}$$

where $\eta^{1/2}\omega_2 = \partial_A h + \partial_A\tilde{g}(A, J) = O(\eta^{1/2})$, see also (5.31).

The hypothesis that F, hence $\tilde{F}$, has only one fast mode is easily seen to imply that Φ exists, is holomorphic in a suitable domain and it generates a canonical transformation, close to the identity within $O(\beta\eta^{1/2})$, on a domain $W(\bar{\kappa}_0, \bar{\rho}_0, \bar{\xi}_0, \tilde{\mu}_0)$ for some $\bar{\kappa}_0, \bar{\rho}_0 \equiv \eta^{-1/2}\bar{\rho}, \bar{\xi}_0, \tilde{\mu}_0 > 0$, $\bar{\rho} > 0$. Of course this holds if $\beta\eta^{1/2}$ is small enough

The finite modality of F is not really necessary: it gives easily the property that the divisors found in solving (A10.7) cannot vanish even for complex phase space points (which could in general happen as discussed in Appendix A5). But such divisors can arise only if

$w \equiv \nu i\vec{\omega}\cdot\vec{\nu}_0 + gp$ (with $g \equiv \partial_J \tilde{g}$) vanishes for some p, ν different from zero: this is impossible for η small as one can easily check that $|w| >$ *const* $\eta^{-1/2}$ if ν and p do not vanish.

The canonical map transforms (A10.5) into:

$$\eta^{-1/2}\omega B_0 + \hat{h}(\eta 1/2) + \hat{g}(A, p_0 q_0) + \beta\mu\hat{f} + \beta^2\eta^{1/2}\hat{F} \tag{A10.8}$$

as the largest "second order term" comes from $\tilde{F}$ itself and therefore it has size $O(\beta\partial\Phi) = O(\beta^2\eta^{1/2})$.

Note that the fact that Φ has size of order $O(\beta\eta^{1/2})$ is a consequence of the fast mode assumption on F, forbidding the appearance of divisors of order $\eta^{1/2}$ or even of order 1, in the solution of (A10.7), see also (5.31). But what said so far would hold rather generally if we only assumed that F contained just fast modes, not necessarily parallel to a fixed one $\vec{\nu}_0$.

We proceed by remarking that the assumption that F contains only modes parallel to $\vec{\nu}_0$ has the simple consequence that also $\tilde{F}$ *and* $\hat{F}$ have the same property. The $\hat{F}$ might have non zero average over the angles $\vec{\alpha}$: by a further canonical change of coordinates of the type of lemma 0, §5, we can transform (A10.8) back into a hamiltonian of the same form with $\hat{F}$ with zero average (and new $\hat{h}$, $\hat{g}$). In this way we see that we can assume that the canonical map transforms (A10.6) into ♣A10.10 with $\hat{F}$ unimodal, fast, and with zero average over the $\vec{\alpha}$ angles.

The latter properties would fail if F had not been assumed unimodal (but just fast).

Hence we can repeat the above argument and further reduce, in suitable new coordinates the size of F to $\beta^{2^n}\eta^{(2^n-1)/2}$, after n steps. No small divisors problems can arise, again by our unimodality assumption (which makes the system, for the purposes of the present discussion, effectively one dimensional). If $\mu = \eta^c$ we continue until $2^n - 1 > 2c$. At each step we must put a requirement on the size of $\beta\eta^{1/2}$ in order to solve the implicit functions problems that arise at each step in passing from the generating functions Φ to the actual map. However, see the remark 1) after lemma 1', §5, we could continue indefinitely and build a canonical transformation casting (A10.6) into the same form with $\tilde{F} \equiv 0$. The quadratic decrease of the size of F is such that the successive conditions on the size of $\beta\eta^{1/2}$ or the decrease in the analyticity domains (*i.e.* of the constants measuring their size) become essentially negligible: but it is sufficient to stop when the size of the new F has become of order η^c for all $|\beta|\eta^{1/2}$ small enough. This happens if one considers an order n_0 such that $n_0 \sim \log\big((\log\eta^{c+1/2})(\log\beta\eta^{1/2})^{-1}\big)$. However, we shall prefer, in order to have *analyticity* in β, to consider a smaller domain, say, $|\beta| < B^*\eta^{-\sigma}$ with $0 < \sigma < 1/2$, in which case it will be sufficient to take $n_0 \ge (\log 2)^{-1}\log[(2c+1)/(1-2\sigma)]$.

At this point we have put the original hamiltonian in the form (A10.6) with $\tilde{F} = 0$ and a new $\tilde{f}$, as $\mu \le \eta^c$ is supposed to hold. However the new $\tilde{f}$ will in general have all harmonics (*i.e.* it will no longer be a trigonometric polynomial in the $\vec{\alpha}$).

Thus we see the "averaging" phenomenon: the problem of casting the (A10.1) into the form (A10.4) is equivalent (if $\beta\eta^{1/2}$ is small enough) to the same problem with $F = 0$, $\beta =$

1, $\mu=\eta^c$ and (another) f with the same analyticity properties and with the (minor as we shall see) difference that it is not a trigonometric polynomial but "only" $\bar\xi_0$ analytic in the $\vec\alpha$ variables with some $\bar\xi_0$.

Hence we look at the same problem with $F=0,\beta=1$, and at first with f being a trigonometric polynomial of degree N: and we perform the Jacobi map to put the hamiltonian in the form (A10.6); this time with $\tilde F$ and $\beta=1$.

We denote $A_0,p_0,q_0,\vec\alpha_0$ the canonical coordinated describing our problem after the Jacobi map and in a domain $W(\kappa_0,\rho_0,\xi_0,\bar\mu)$, with $\bar\mu=\eta^c$. By the assumption that f is a trigonometric polynomial *we can fix* ξ_0 *arbitrarily*, (≤ 1); and ρ_0 is of order $\eta^{-1/2}$.

The hamiltonian will be written $H\equiv H_0+\varepsilon x$ with $x\equiv\mu\tilde f$ and with ε being a formal parameter to be set eventually equal to 1.

The function $x\equiv\mu f$ can be treated perturbatively up to order s (to be fixed later), at least formally, in the sense that one can define:

$$\Phi=\varepsilon\Phi^1+\varepsilon^2\Phi^2+\ldots+\varepsilon^s\Phi^s \tag{A10.9}$$

recursively so that, if $H=H_0+\varepsilon x$, it verifies in the sense of formal series in ε:

$$\begin{aligned}&H_0(B+\partial_{\lambda_0}\Phi,A+\partial_{\alpha_0}\Phi,p+\partial_{q_0}\Phi,q_0)+\varepsilon x(A+\partial_\alpha\Phi,p+\partial_{q_0}\Phi,q_0,\lambda_0,\alpha_0)=\\&=H_0(B,A,p\cdot(q_0+\partial_p\Phi),q_0)+\sum_{k=1}^s\varepsilon^kH_k(A,p\cdot(q_0+\partial_p\Phi))+O(\varepsilon^{s+1})\end{aligned} \tag{A10.10}$$

with suitable functions H_h, up to order s in powers of ε.

The (A10.10) gives the following set of recursive equations ($\vec\alpha_0\equiv(\lambda_0,\alpha_0)$):

$$\vec\omega\cdot\partial_{\vec\alpha_0}\Phi^k+\bar g\,(q_0\partial_{q_0}\Phi^k-p\partial_p\Phi^k)+x^k(p,q_0,A,\vec\alpha_0)-\langle x^k\rangle^D(p\,q_0,A)=0 \tag{A10.11}$$

for $k=1,\ldots,s$; here $\vec\omega=\vec\omega(A,pq_0)=(\eta^{-1/2}\omega,\partial_A(\,\eta A^2/(2\bar J)+\bar g),\ \bar g=\bar g(A,pq_0)$ and the D superscript denotes the "diagonal cut" operation, defined for any function $f(p,q)=\sum_{r,s=0}^\infty f_{rs}p^rq^s$ as the map:

$$f(p,q)\to f^D(p\cdot q)=\sum_{r=0}^\infty(pq)^rf_{rr} \tag{A10.12}$$

By assumption (see comment following (A10.2)) $\vec\omega$ has the form $(\eta^{-1/2}\omega,\eta^{1/2}\omega')$ with $0<\bar\omega<\omega'<\tilde\omega$ in the real part of the definition domains.

If we label with $i=1,2,3,4,5$ the five variables conjugated to $(A,p,q_0,\alpha_0,\lambda_0)$ (*i.e.* α_0, q_0, p,A,B) it is, with the notations of (6.10):

$$x^k(p,A,q_0,\vec\alpha)=\sum_{\substack{\vec m\\ m_3=0,|\vec m|>1}}H_0^{\vec m}(A,p,q_0)\sum_{(k^i_j)_{\vec m,k}}\prod_{i=1}^5\prod_{j=1}^{m_i}\partial_i\Phi^{k^i_j}+$$

$$
\begin{aligned}
&+\sum_{\substack{\vec m\\ m_3,m_4=0,|\vec m|\ge 1}} \mu f^{\vec m}(A,p,q_0,\vec\alpha_0)\sum_{(k^i_j)_{\vec m,k-1}}\prod_{i=1}^{5}\prod_{j=1}^{m_i}\partial_i\Phi^{k^i_j}+\\
&-\widetilde{\sum_{\substack{m_j\,\equiv\,0,j\ne 3\\ r=0,\ldots,k-1}}} H^{\vec m}_k(A,pq_0)\sum_{(k^3_j)_{m_3,k-r}}\prod_{j=1}^{m_3}\partial_3\Phi^{k^3_j} \qquad (A10.13)\\
H_k(A,pq) =&\langle x^k\rangle^D\equiv x^{kD}
\end{aligned}
$$

where $\widetilde\Sigma$ means that $m_3>1$ if $r=0$ and $m_3\ge 1$ for $r\ge 1$.

The key remark is that the above equation can be solved recursively, producing Φ^k, x^k *which are trigonometric polynomials of degree* $\le kN$ *for all* $k\le s_0$*, provided* s_0 *is such that* $|\vec\omega\cdot\vec\nu|>\eta^{1/2}\omega/2$ *for* $0<|\nu_i|\le s_0N$. Since $\vec\omega=(\eta^{-1/2}\omega,\eta^{1/2}\omega_2)$, with $\bar\omega<\omega_2<\tilde\omega$ by assumption, this holds if $s_0=\omega/(2N\eta\tilde\omega)\equiv b/N\eta$.

We need, however, also some bounds on Φ^k,H_k,x^k, for $k\le s_0$. We set, for $\delta>0$ to be chosen later:

$$
\rho_h=\rho_0(1-h\delta),\qquad \xi_h=\xi_0(1-h\delta),\quad \kappa_h=\kappa_0(1-h\delta) \qquad (A10.14)
$$

and $\rho'_h=\rho_{h-1/2}$, $\kappa'_h=\kappa_{h-1/2}$, $\xi'_h=\xi'_{h-1/2}$, so that $\rho_h>\rho'_h>\rho_{h-1}$, *etc.*

To simplify the analysis we shall often replace, in the following, ρ_0 with κ_0^2, in spite of the fact that $\rho_0=O(\eta^{-1/2})$ while $\kappa_0=O(1)$.

Note that, for $k=1$ and suitable D_X,D_H,D_Φ it is ($x^1=\mu f$):

$$
|x^1|_{\rho_1,\xi_1,\kappa_1,\mu_0}<D_X,\quad |H_1|_{\rho_1,\xi_1,\kappa_1,\mu_0}<D_H,\quad |\Phi^1|_{\rho_1,\xi_1,\kappa_1,\mu_0}<D_\Phi \qquad (A10.15)
$$

Possibly reducing by a factor 2 the size of the original analyticity domains, (an operation which we may and shall assume as unnecessary, possibly by redefining the analyticity parameters κ_0,ξ_0), we see immediately that in our case we can take, for some $\gamma'>\gamma>1$, $D_X=\gamma\mu$, $D_H=\gamma'\mu$.

We can also take $D_\Phi=D_X\eta^{-1/2}\gamma''$ for some $\gamma''>\gamma$. This is because, in general for $k\le s_0$, we can bound $\partial\Phi$ in terms of x by $D_XK\,(\xi\delta)^{-4}\eta^{-1/2}$ if $K>1$ is a suitable constant and if δ denotes the analyticity loss in the domain of Φ with respect to that of x (*i.e.* if κ,ρ,ξ are the analyticity parameters of x and $\kappa(1-\delta)$, $\rho(1-\delta)$,$\xi(1-\delta)$ are those of Φ).

The point being that for $k\le s_0$ the smallest divisors are bounded below by $O(\eta^{1/2})$ and the sum giving Φ runs over four integer indices, see (5.31), so that $\partial_\alpha\Phi$ requires a bound "of dimension" 4 (in fact there are only two angle variables and we could get a better bound of the order $\xi^{-3}\delta^{-4}$; but here and below we do not do so, for simplicity).

Suppose also that for $1\le h\le k-1<s_0$ it is:

$$
\begin{aligned}
|x^h|_{\rho_h,\xi_h,\kappa_h,\mu_0}<&D_XB^{h-1},\qquad |H_h|_{\rho_h,\xi_h,\kappa_h,\mu_0}<D_HB^{h-1},\\
|\Phi^h|_{\rho_h,\xi_h,\kappa_h,\mu_0}<&D_\Phi B^{h-1}
\end{aligned} \qquad (A10.16)
$$

for a suitably chosen B. This holds for $h=1$ (with *any* B) by the above comments.

Then we see that (A10.13) can be used to find a bound on x^k in $W_{\rho'_k,\xi'_k,\kappa'_k,\mu_0}$

$$
\begin{aligned}
|x^k|_{\rho'_k,\xi'_k,\kappa'_k,\mu_0} \le & \sum_{m\ge 2} E_0 B^k \big(\kappa_0^{-2} d\delta^{-1}\xi_0^{-1} D_\Phi B^{-1}\big)^m \binom{k-1}{m-1} + \\
& + \sum_{m\ge 1} D_X B^{k-1} \big(\kappa_0^{-2} d\delta^{-1}\xi_0^{-1} D_\Phi B^{-1}\big)^m \binom{k-1}{m-1} + \\
& + \sum_{r=1}^{k-1}\sum_{m=1}^{k-r} D_H B^{k-r} \big(\kappa_0^{-2} d\delta^{-1}\xi_0^{-1} D_\Phi B^{-1}\big)^m \binom{k-r-1}{m-1}
\end{aligned}
\tag{A10.17}
$$

where the factor d is a numerical constant arising from various bounds (for instance from bounding from below $\delta/2$ and $(1-e^{-\xi_0\delta/4})$ by a constant times δ and $\xi_0\delta$, respectively, in the dimensional bounds leading to (A10.17)); the term with $r=0$ in the third line of (A10.17) is bounded here by the first line (having absorbed numerical constants in the definition of d).

This implies that $|x^k|_{\rho'_k,\xi'_k,\kappa'_k,\mu_0}$ is bounded above by:

$$
B^{k-1}\Big(1+\frac{dD_\Phi}{\kappa_0^2\xi_0 B\delta}\Big)^k \Big[E_0 B(\frac{dD_\Phi}{\kappa_0^2\xi_0 B\delta})^2 + D_X(\frac{dD_\Phi}{\kappa_0^2\delta\xi_0 B}) + D_H(\frac{dD_\Phi}{\kappa_0^2\xi_0 B\delta})\Big] \tag{A10.18}
$$

Let $\bar\gamma>1$ be fixed large enough so that the diagonal cut of the function x can be bounded by $\bar\gamma\delta^{-2}$ times (A10.18) on the smaller domain $W(\rho_k,\xi_k,\kappa_k,\mu_0)$. Note that the function Φ_k can be bounded dimensionally in the domain $W(\rho_k,\xi_k,\kappa_k,\mu_0)$ by $2^4K(\xi_0\delta)^{-4}\eta^{-1/2}$ times the bound (A10.18) of x_k in the domain $W(\rho'_k,\xi'_k,\kappa'_k,\mu_0)$. We shall proceed by choosing B so large that $(1+dD_\Phi/(\kappa_0^2\xi_0 B\delta))^k<e$ and so that each of the three term in square brackets are bounded by $D_X(\bar\gamma^{-1}\xi_0^2\delta^2)^2/(3e)$.

This is achieved if we suppose that $\delta=1/4k$, so that $\xi_h>\xi_0/2$, for $h\le 4$, that $\bar\gamma>2^4K$ and if we impose:

$$
\delta=\frac{1}{4k},\qquad \bar\gamma^{-1}\xi_h^4\delta^2<1\quad\text{for } h\le k \tag{A10.19}
$$

and:

$$
\begin{aligned}
B > & \frac{dkD_\Phi}{\kappa_0^2\xi_0\delta}, \qquad B>\frac{\bar\gamma^2}{\delta^4\xi_0^4}3eE_0\frac{D_\Phi^2}{D_X}\frac{d^2}{(\kappa_0^2\xi_0\delta)^2},\\
B > & \frac{\bar\gamma^2}{\delta^4\xi_0^4}3e\frac{dD_\Phi}{\kappa_0^2\xi_0\delta}, \qquad B>\frac{\bar\gamma^2}{\xi_0^4\delta^4}3e\frac{D_H D_\Phi}{D_X}\frac{d}{\kappa_0^2\xi_0\delta}
\end{aligned}
\tag{A10.20}
$$

We see that:

$$
|x^k|_{\rho_k,\xi_k,\kappa_k,\mu_0}\le D_X B^{k-1}(\bar\gamma^{-1}\xi_0^2\delta^2)^2 \tag{A10.21}
$$

Giving up explicit control of the constants (for simplicity of notation) we see that (A10.20) can be implied by the stronger conditions, having chosen $\delta=1/4k$ for an arbitrarily fixed k, and replacing D_Φ by $D_X\gamma''\eta^{-1/2}$, D_H by $\gamma'\mu$ and D_X by $\gamma\mu$ (see the comment after (A10.15):

$$
\begin{aligned}
B >&c_0\mu\eta^{-1/2}k^2, \qquad B > c_0\mu k^6\eta^{-1}\\
B >&c_0\mu\eta^{-1/2}k^5, \qquad B > \mu k^5\eta^{-1/2}
\end{aligned}
\tag{A10.22}
$$

for some $c_0 > 0$.

Therefore, if $h \le k$, we see that B can be taken $B = B_0 k^6\mu\eta^{-1}$ for some constant $B_0 > 1$. And also, recalling that γ', γ'' have been chosen larger than γ, that $\bar\gamma^2 > 2^4 K$ and (A10.19) we see that :

$$
\begin{aligned}
&|\Phi_k|_{\rho_k,\xi_k,\kappa_k,\mu_0} \le D_X B^{k-1}(\bar\gamma^{-1}\xi_h^2\delta^2)^2\frac{2^4K}{(\xi_0\delta)^4\eta^{1/2}} \le B^{k-1}D_\Phi\\
&|x^{kD}|_{\rho_k,\xi_k,\kappa_k,\mu_0} \le B^{k-1}D_X \le B^{k-1}D_H
\end{aligned}
\tag{A10.23}
$$

Hence the inductive proof works and we get:

$$
\begin{aligned}
|x^k|_{\rho_k,\xi_k,\kappa_k,\mu_0} <&c_1\mu(\mu\eta^{-1}B_0)^{k-1}(k-1)!^6,\\
|H_k|_{\rho_k,\xi_k,\kappa_k,\mu_0} <&c_1\mu(\mu\eta^{-1}B_0)^{k-1}(k-1)!^6\\
|\Phi^k|_{\rho_k,\xi_k,\kappa_k,\mu_0} <&c_1\mu\eta^{-1/2}(\mu\eta^{-1}B_0)^{k-1}(k-1)!^6
\end{aligned}
\tag{A10.24}
$$

for some B_0, c_1 being constants depending on the maximum of the coefficients in F, f, J, J_0, g_0 in their analyticity domains as well as on the sizes $\bar\rho_0, \bar\xi_0$ of the domains.

We see that the above results (A10.4),(A10.5) follow immediately, *under the present assumption that f is a trigonometric polynomial*: and in fact we get a better bound as the remainder will be of order $O(e^{-b/(\eta N)})$ (recall the definition of b in the comment after (A10.13)) *i.e.* much smaller than what declared in (A10.10). This is obtained by pushing the perturbation analysis up to an order $s_0 = b/N\eta$. The remainders are estimated via the analyticity. They are of order $e^{-b/N\eta}$, if $\mu = \eta^c$ and $\eta^{c-7}B_0(b/N)^6 < e^{-1}$.

But we still have to relax the trigonometric polynomial assumption. We follow the usual cut off technique to exploit the fast decay as $|\vec\nu| \to \infty$ of the Fourier transform, which allows us to regard f "almost" as a trigonometric polynomial.

More precisely let N_0 be a cut off parameter so that if f is $\bar\rho_0, \bar\xi_0$ analytic then:

$$
||f^{[>N_0]}||_{\frac{\bar\rho_0}{2}\frac{\bar\xi_0}{2}} < c_3\bar\xi_0^{-1}||f||_{\bar\rho_0\bar\xi_0}e^{-\bar\xi_0 N_0/2}, \qquad ||f^{[\le N_0]}||_{\frac{\bar\rho_0}{2}\frac{\bar\xi_0}{2}} < c_3\bar\xi_0^{-1}||f||_{\bar\rho_0\bar\xi_0}
\tag{A10.25}
$$

for a suitable constant c_3.

We fix, therefore, $N_0 = \eta^{-1/2}$ and apply the above argument to the hamiltonian with $f^{[\le N_0]}$ replacing f. Then we can perform perturbation theory up to $s_0 = b/N_0\eta = b\eta^{-1/2}$. We construct in this way a canonical transformation casting the hamiltonian $H_0 + \mu f^{[\le N_0]}$ exactly in the form (A10.5), provided $\eta^{c-10}B_0 b^6 < e^{-1}$ and $e^{-b/\sqrt{\eta}} < \eta^x$.

The same transformation will cast the total hamiltonian (*i.e.* with f rather than $f^{[\le N_0]}$) in the form (A10.5) with a remainder which will be $O(e^{-\bar\xi_0 N_0/2}) = O(e^{-c'/\sqrt{\eta}})$ by (A10.25). This yields (A10.5).

Therefore we see that the above analysis can be carried out if $c > 10$ and if $|\beta| < B^*\eta^{-\sigma}$, $0 < \sigma < 1/2$ (which was necessary in the first part of our discussion).

Thus, by the analyticity in β, the invariant tori and their whiskers, constructed for (A10.1) via lemma 1', can be expanded in powers of β and, if $\beta = 1$, the power series terms are bounded, at order k, proportionally to $(\bar b\eta^{-\sigma})^k$, for some $\bar b > 0$. The round spacing in the whiskers ladders will be smaller than η^y for some $y > 0$ provided x is large enough: and by taking x large enough and η small enough we can make y larger than any prefixed amount.

In other words the F in (A10.1) can be regarded as formally of order $\eta^{1/2}$ (actually η^σ with any $\sigma < 1/2$) and along the line $A \in \eta^{-1/2}[\bar A, \tilde A]$ the whiskers form a ladder with round spacing that is smaller than any power of η as $\eta \to 0$.

Appendix 11: Scattering phases shifts and intrinsic angles

In this section we show that if the homoclinic splitting is exponentially small, also the scattering phase shifts are such. A fact checked, for some even models, by explicit estimates in §11.

To express the homoclinic angles in the intrinsic coordinates, we consider the derivatives of Q, see (10.7), with respect to $\vec\alpha$ at $\vec\alpha = \vec 0$. This means that we consider:

$$\begin{aligned}\partial_{\alpha_j} Q \equiv{}& \partial_{\vec\psi} Q \cdot \partial_{\alpha_j}\vec\psi = \\ &= \Big[\Big(\partial_p Z(p_{\vec 0}, 0, \vec\psi^{\,s}_0)\,\partial_{\vec\psi}\, p_{\vec\psi} - \partial_q Z(0, q_{\vec 0}, \vec\psi^{\,u}_0)\,\partial_{\vec\psi}\, q_{\vec\psi}\;(1 + \partial_{\vec\psi}\vec\sigma)\Big) + \\ &+ \sum_{\vec\nu} i\vec\nu\Big(X^s_{\vec\nu} - X^u_{\vec\nu}(1 + \partial_{\vec\psi}\vec\sigma)\Big)\Big] \cdot \partial_{\alpha_j}\vec\psi\end{aligned} \tag{A11.1}$$

where $X^s = X^{sk} = X^{sk}(p_0, 0, \vec\psi^{\,s}_0)$, $X^u = X^{uk} = X^{uk}(0, q_0, \vec\psi^{\,u}_0)$, and all the derivatives are evaluated at $\vec\psi = \vec 0$.

The above (A11.1) is written symbolically: various indices are omitted as the contraction rules are obvious.

For instance the derivatives with respect to $\vec\psi$ of the Q^0 functions, which give what we have just called the intrinsic angles, are given by (A11.1) with $\vec\sigma \equiv \vec 0$. The second line in (A11.1), with $\vec\sigma \equiv \vec 0$ can be studied by remarking that (10.2),(10.3) imply:

$$\begin{aligned}&\partial_p Z_-(p_{\vec 0}, 0, \vec\psi^{\,s}_0)\,\partial_{\vec\psi}\, p_{\vec\psi} + \partial_{\vec\psi} Z_-(p_{\vec 0}, 0, \vec\psi^{\,s}_0) = 0 \\ &\partial_q Z_-(0, q_{\vec 0}, \vec\psi^{\,u}_0)\,\partial_{\vec\psi}\, q_{\vec\psi} + \partial_{\vec\psi} Z_-(0, q_{\vec 0}, \vec\psi^{\,s}_0) = 0\end{aligned} \tag{A11.2}$$

Furthermore considering:

$$\zeta_s(t) = Z(p_{\vec 0} e^{-gt}, 0, \vec\psi^{\,s}_0 + \vec\omega t), \qquad \zeta_u(t) = Z(0, q_{\vec 0} e^{gt}, \vec\psi^{\,u}_0 + \vec\omega t) \tag{A11.3}$$

we get:

$$\dot\zeta_s(t) = -e^{-gt} g p_{\vec 0} \partial_p Z + \vec\omega \cdot \vec\partial_{\vec\psi} Z, \qquad \dot\zeta_u(t) = e^{gt} g q_{\vec 0} \partial_q Z + \vec\omega \cdot \vec\partial_{\vec\psi} Z \tag{A11.4}$$

Denoting $\partial.Z|_p$ the derivatives of Z at $p = p_0, q = 0, \vec\psi = \vec 0$ and $\partial.Z|_q$ the derivative at $q = q_0, p = 0, \vec\psi = \vec 0$ we add and subtract terms and use (A11.4) to transform the part of the second line in (A11.1) not containing the terms proportional to $\partial_{\vec\psi}\vec\sigma$ into:

$$\begin{aligned}
&\tfrac{\partial_p Z}{\partial_p Z_-}\,\partial_{\vec\psi} Z_-|_p - \tfrac{\partial_q Z}{\partial_q Z_-}\,\partial_{\vec\psi} Z_-|_q \equiv\\
&\equiv \Big[\tfrac{\partial_p Z}{\partial_p Z_-} - \tfrac{\partial_q Z}{\partial_q Z_-}\Big]\,\partial_{\vec\psi} Z_-|_p + \tfrac{\partial_q Z}{\partial_q Z_-}\Big[\partial_{\vec\psi} Z_-|_p - \partial_{\vec\psi} Z_-|_q\Big] \equiv\\
&\equiv -\Big[\tfrac{\dot\zeta_s - \vec\omega\cdot\vec\partial_{\vec\psi} Z|_p}{\dot\zeta_{s-} - \vec\omega\cdot\vec\partial_{\vec\psi} Z_-|_p} - \tfrac{\dot\zeta_u - \vec\omega\cdot\vec\partial_{\vec\psi} Z|_q}{\dot\zeta_{u-} - \vec\omega\cdot\vec\partial_{\vec\psi} Z_-|_q}\Big]\,\partial_{\vec\psi} Z_-|_p + \tfrac{\partial_q Z}{\partial_q Z_-}\Big[\partial_{\vec\psi} Z_-|_p - \partial_{\vec\psi} Z_-|_q\Big]
\end{aligned} \tag{A11.5}$$

which simply are the derivatives of the above introduced functions Q^0 instead of the Q,

We can remark that $\dot\zeta_s = \dot\zeta_u$, at $t = 0$, because we are at a homoclinic point. Furthermore the denominators in (A11.5) are $(-2g)$ for $\mu = 0$, hence they are bounded away from zero if μ is small enough. And the derivatives $\partial_{\vec\psi} Z_-|_p$ and $\partial_q Z$ are small of order μ as $\mu \to 0$ (as they vanish if $\mu = 0$). Hence we see that the derivatives of the homoclinic equation in the $\vec\alpha$ coordinates are related to the derivatives in the $\vec\psi$ coordinates by terms proportional either to the $\vec\psi$ derivatives of the scattering phase shifts or to the $\vec\psi$ derivatives of the splitting itself multiplied by $O(\mu)$.

We see that this implies that if $\partial_{\psi_i}\sigma_j(\vec 0) = \vec 0$, as in fact we show in §10 for even models, and if the intrinsic intersection tensor is exponentially small then also the natural intersection tensor is exponentially small. One can, likewise check that also the converse holds: in general, for even models and to leading order in μ, the two notions of angles coincide.

We can also deduce, in general, from the knowledge that the splitting is exponentially small for all $\vec\alpha$, as it is the case in even models if $l = 2$ or $l \geq 3$ and all rotations fast, as shown in §11, that the scattering phase shifts must be exponentially small (as we can "compute" them by difference from (A11.1), and as we can infer, by dimensional bound that exponentially small in $\vec\alpha$ implies exponentially small in $\vec\psi$ (because the $\vec\alpha$ are analytic in the $\vec\psi$ and viceversa).

Appendix 12: Compatibility. Homoclinic identities

1) Here we want to check directly that $\vec{F}^k_{\uparrow\infty}$ has vanishing mean value, which is a crucial fact in the derivation of the main equations of §6.

More precisely, *assume* that, for $1 \le j \le k-1$, $\vec{X}^j(t)$ in (6.11) has the form $\vec{X}^j(t) = \vec{X}^j(\vec{\omega}_{su}t, t)$, with $\vec{X}^j(\vec{\psi}, t)$ periodic in $\vec{\psi}$ and $\partial_t \vec{X}^j(\vec{\psi}, t)$ converging to 0 exponentially fast as $t \to \infty$ and recall that, if $\vec{\alpha}$ is the point over which we construct the whisker, then $\vec{X}^0 = (I^0(t), \varphi^0(t), \vec{A}, \vec{\alpha} + \vec{\vartheta}(t))$ so that:

$$\vec{X}^0(\psi, \infty) \equiv \vec{X}^0_\infty(\vec{\psi}) = (0, 0, \vec{A}, \vec{\psi} + \vec{\vartheta}(\infty)) \tag{A12.1}$$

(the limit $\vec{\vartheta}(\infty)$ being reached at an exponential rate). Then, for $1 \le j \le k$, $\vec{F}^j(t)$, which is defined in terms of the $\vec{X}^i$, $0 \le i \le j-1$ [see (6.10)], has also the form $\vec{F}^j(t) = \vec{F}^j(\vec{\psi} + \vec{\omega}_{su}t, t)$ (with $\partial_t \vec{F}^j(\vec{\psi}, t) \to 0$ as $t \to \infty$ exponentially fast) and the limits $\vec{X}^j(\vec{\psi}, \infty) \equiv \vec{X}^j_\infty(\vec{\psi})$ $(j \le k-1)$ and $\vec{F}^j(\vec{\psi}, \infty) \equiv \vec{F}^j_\infty(\vec{\psi})$ $(1 \le j \le k)$are well defined.

We then show that *from the above assumptions it follows* that:

$$\int \vec{F}^k_{\infty\uparrow}(\vec{\psi}) d\vec{\psi} = \vec{0} \tag{A12.2}$$

(recall that the subscripts $(\cdot)_+, (\cdot)_-, (\cdot)_\uparrow, (\cdot)_\downarrow$ denote projections onto, respectively, I,φ, $\vec{A}$,$\vec{\alpha}$-coordinates). The argument is adapted from the similar argument in [CZ].

¿From the definitions of §6 it is clear that the function $\vec{F}^k_\infty(\vec{\alpha})$ is given by:

$$\vec{F}^k_{\infty\uparrow} \equiv -[\vec{\partial} f(\sum_{j=0}^{k-1} \mu^j \vec{X}^j_\infty)]_{k-1} = [\vec{\partial} H(\sum_{j=0}^{k-1} \mu^j \vec{X}^j_\infty)]_k \tag{A12.3}$$

where $[\cdot]_k$ denotes the k-th order coefficient of a power series in μ; $\vec{\partial} \equiv \partial_{\vec{\alpha}}$; $H = H_0 + \mu f$ is the hamiltonian and the sums are the arguments of the functions $\vec{\partial} f, \vec{\partial} H$.

On the other hand the function $Y(\vec{\psi}) = \sum_{j=0}^{k-1} \mu^j X^j_\infty(\vec{\psi})$ is such that $Y(\vec{\omega}_{su}t)$ verifies the Hamilton equations up to order $k-1$ in μ, *i.e.* if D is the operator $D = \vec{\omega}_{su} \cdot \vec{\partial}$ (where, of course, $\vec{\partial}$ stands now for $\partial_{\vec{\psi}}$):

$$\dot{Y} \equiv DY = E\partial H(Y) \qquad \text{up to order } k-1 \tag{A12.4}$$

where $\partial = (\partial_I, \partial_\varphi, \partial_{\vec{A}}, \vec{\partial})$ and E is the obvious symplectic matrix (so that $E^2 = -1$): $\partial H(Y)$ is ∂H evaluated at $Y(\vec{\psi})$ and DY as well as Y are evaluated at $\vec{\psi}$, with $\vec{\psi} = \vec{\omega}t$.

We rewrite (A12.4) as:

$$E\,DY + \partial H(Y) = 0 \qquad \text{up to order } k-1 \tag{A12.5}$$

and we note the identities:

$$\int \partial Y(\vec\psi)\cdot EDY(\vec\psi)\,d\vec\psi\equiv 0,\qquad \int \partial Y(\vec\psi)\cdot \partial H(Y(\alpha))\,d\vec\psi=\vec 0 \tag{A12.6}$$

valid for *any* periodic function $Y(\vec\psi)$ of $\vec\psi$; hence in our case they are valid to *all* orders in μ. They are trivial consequences of the periodicity in $\vec\psi$ of $Y(\vec\psi)$ and of the peculiar structure of E or, in the case of the second, of the remark that the integrand is just the $\vec\partial$ of $H(Y(\vec\psi))$.

Therefore we have:

$$\int \partial Y\cdot(EDY+\partial H(Y))\,d\vec\psi=0 \tag{A12.7}$$

to all orders in μ. We shall write it explicitly to order k: remarking that Y has order $k-1$ and that (A12.5) holds up to order $k-1$, this gives:

$$\int [\partial Y]_0\cdot[\partial H(Y)]_k=0 \tag{A12.8}$$

but $[Y(\vec\psi)]_0=(0,0,\vec A,\vec\psi+\vec\vartheta_\infty)$ so that (A12.8) becomes simply (A12.2).

2) In fact, the above method can be used to prove other interesting *homoclinic identities*, namely:

$$\begin{aligned}-2\int_0^\infty F^k_{\uparrow\vec 0i}(t)dt=&\sum_{j=1}^{k-1}\langle X^j_+(\cdot,\infty)\partial_{\psi_i}X^{k-j}_-(\cdot,\infty)+\vec X^j_\uparrow(\cdot,\infty)\cdot\partial_{\psi_i}X^{k-j}_\downarrow(\cdot,\infty)\rangle+\\ &-\sum_{j=1}^{k-1}\langle X^j_+(\cdot,0)\partial_{\psi_i}X^{k-j}_-(\cdot,0)+\vec X^j_\uparrow(\cdot,0)\cdot\partial_{\psi_i}X^{k-j}_\downarrow(\cdot,0)\rangle\\ 2\int_0^{+\infty}w(\tau)F^k_{\vec 0}(\tau)=&\,w_\uparrow(0)\cdot X^k_{\uparrow\vec 0}(0)\end{aligned}\tag{A12.9}$$

where $\vec F^k_{\uparrow\vec 0}\equiv(F^k_{\uparrow\vec 01},\ldots,F^k_{\uparrow\vec 0l-1})$, and $w(t)$ is defined in (6.36).

To check this identity we use that $X(\vec\psi+\vec\omega t;t,\vec\alpha)\equiv X(\vec\psi+\vec\omega t,t)$ describes for all $\vec\psi$ a motion on the stable whisker and in particular satisfies the Hamilton equation:

$$\frac{d}{dt}X^j(\vec\psi+\vec\omega t,t)=L(t)X^j(\vec\psi+\vec\omega t,t)+F^j(\vec\psi+\vec\omega t,t) \tag{A12.10}$$

which, performing the t-derivative and using the arbitrariness of $\vec\psi$, can be rewritten as:

$$E\,DX^j(\vec\psi,t)+E\,\partial_tX^j(\vec\psi,t)+\Big[\partial H\Big(\sum_{i=0}^j\mu^iX^i(\vec\psi,t)\Big)\Big]_j=0\,,\qquad(0\le j\le k-1)\tag{A12.11}$$

Similarly to above, let $Y(\vec{\psi},t) \equiv \sum_{j=0}^{k-1} \mu^j X^j(\vec{\psi},t)$, then by (A12.6) we see that:

$$\int \partial_{\psi_i} Y \cdot \big(E\ DY + E\ \partial_t Y + \partial H(Y)\big) d\vec{\psi} = \int \partial_{\psi_i} Y \cdot E\ \partial_t Y d\vec{\psi} \tag{A12.12}$$

to all orders in μ; and taking the order k in μ, using (A12.10), (A12.11), one recognizes:

$$-F^k_{\uparrow\vec{0}i}(t) = \sum_{j=1}^{k-1} \langle \partial_{\psi_i} X^j \cdot E\ \partial_t X^{k-j} \rangle \tag{A12.13}$$

Finally one integrates the latter identity between 0 and ∞: the integral can be performed by parts first in t and then in ψ_i and, after changing j to $k-j$ and using $E^T = -E$, the (A12.9) is easily obtained.

The second identity is obtained in the same way by multiplying by $\dot{Y}(t)$ the expression $E\dot{Y}(t) + \partial H(Y(t))$, which vanishes up to order $k-1$, and by integrating from 0 to t. On obtains the variation $H(Y(t) - H(Y(0))$. Writing the above identity to order k and using that the energy $H(Y)$ is conserved up to order $k-1$ we get:

$$\int_0^t \dot{X}^0(\tau)\big[\partial H(Y(\tau))\big]_k d\tau = \Big[\partial H_0(X^0(\tau))\, X^k(\tau)\Big]\Big|_0^t \tag{A12.14}$$

We can write the Hamilton equations in the form $\dot{X} = E\partial H$ so that the above relation is (using $\dot{X}^0 = Ew + E\partial h$, and (6.36) and $\big[E\partial H(Y)\big]_k \equiv F^k$ and $\partial H_0 X^k \equiv w X^k + \vec{\omega}\cdot X^k_\uparrow$):

$$-\int_0^t w(\tau) F^k_{\vec{0}}(\tau) d\tau = \langle w(t)X^k(t) - w(0)X^k(0)\rangle + \vec{\omega}\cdot\langle X^k_\uparrow(t) - X^k_\uparrow(0)\rangle \tag{A12.15}$$

(which would hold even without the averages over $\vec{\psi}$). Hence, using $\dot{X}^k_\uparrow = F^k_\uparrow$ and the second of (6.34) and (6.36) [which tell us that $w(0)X^k(0) = -\int_0^\infty w(\tau)F^k(\tau)$] we get (A12.9). In the special case considered in §9 it is $w_\uparrow(\tau) \equiv \vec{0}$, $w_\downarrow(\tau) \equiv \vec{0}$, $w_+(\tau) \equiv \dot{\varphi}_0(\tau) = \dot{\varphi}_0(0) w_-(\tau)$ and we get:

$$2\int_0^\infty w_-(\tau) F^h_{+\vec{0}}(\tau) d\tau = -2\dot{\varphi}(0)^{-1} \int_0^\infty \vec{\omega}\cdot F^k_{\uparrow\vec{0}}(\tau) \tag{A12.16}$$

Appendix 13 Second (and third) order whiskers and phase shifts.

1) Here we prove $c)$ of theorem 3, §10.

Consider the hamiltonian:

$$H = \vec{\omega}\cdot\vec{A} + \frac{1}{2}J^{-1}\vec{A}\cdot\vec{A} + \frac{I^2}{2J_0} + g_0^2 J_0(\cos\varphi - 1) + \mu\sum_\nu f_\nu \cos(\vec{\alpha}\cdot\vec{\nu} + n\varphi) \tag{A13.1}$$

with $\nu \equiv (n, \vec{\nu})$, $\vec{\nu} \neq \vec{0}$, J_0, g_0, f_ν constants and J^{-1} being a constant diagonal matrix.

Then $F_-^h, F_\downarrow^h$ vanish identically and:

$$\begin{aligned}
F_+^1 =& -\partial_\varphi f, \qquad F_\uparrow^1 = -\partial_{\vec{\alpha}} f\\
F_+^2 =& -\frac{J_0 g_0^2}{2}\sin\varphi\,(X_-^1)^2 - \partial_{\varphi^2} f X_-^1 - \partial_{\varphi\vec{\alpha}} f X_\downarrow^1\\
F_\uparrow^2 =& -\partial_{\vec{\alpha}\varphi} f X_-^1 - \partial_{\vec{\alpha}^2} f X_\downarrow^1\\
F_+^3 =& -\frac{J_0 g_0^2}{2}\,2X_-^1 X_-^2 \sin\varphi - \frac{J_0 g_0^2}{3!}\cos\varphi (X_-^1)^3 - \partial_{\varphi^2} f\, X_-^2 +\\
&- \partial_{\vec{\alpha}\varphi} f X_\downarrow^2 - \frac{1}{2}\partial_{\varphi\vec{\alpha}^2} f X_\downarrow^1 X_\downarrow^1 - \partial_{\varphi^2\vec{\alpha}} f X_\downarrow^1 X_-^1 - \frac{1}{2}\partial_{\varphi^3} f\,(X_-^1)^2\\
F_\uparrow^3 =& -\partial_{\vec{\alpha}\varphi} f X_-^2 - \partial_{\vec{\alpha}^2} f X_\downarrow^2 - \frac{1}{2}\partial_{\vec{\alpha}^3} f X_\downarrow^1 X_\downarrow^1 - \partial_{\varphi\vec{\alpha}^2} f\, X_\downarrow^1 X_-^1 - \frac{1}{2}\partial_{\vec{\alpha}\varphi^2} f X_-^1 X_-^1
\end{aligned} \tag{A13.2}$$

where φ in $\cos\varphi$, $\sin\varphi$ is $\varphi \equiv \varphi(t) \equiv \varphi_0(t)$, see (A9.5), *i.e.* it is the unperturbed separatrix motion.

To compute the X^h we define the operator $\mathcal{I}$ as in §10 and we introduce the following operators:

$$\begin{aligned}
\overline{\mathcal{I}}^2 F(t) =& \mathcal{I}^2 F(t) - \mathcal{I}^2 F(0^\sigma)\\
\mathcal{O}F(t) =& w_{21}(t)\big(\mathcal{I}\, w_{22}(\tau)F(\tau)\big)(t) - w_{22}(t)\big(\mathcal{I} w_{21}(\tau)F(\tau)\big)|_{0^\sigma}^t\\
\mathcal{O}_+F(t) =& w_{11}(t)\big(\mathcal{I}\, w_{22}(\tau)F(\tau)\big)(t) - w_{12}(t)\big(\mathcal{I} w_{21}(\tau)F(\tau)\big)|_{0^\sigma}^t
\end{aligned} \tag{A13.3}$$

where the notation of (10.13) is used, and $\sigma = \mathrm{sign}\,(t)$. Then:

$$\begin{aligned}
X_-^1(t) =& w_{21}(t)X_+^1(0) + w_{21}(t)\int_0^t w_{22}(\tau)F_+^1(\tau)d\tau - w_{22}(t)\int_0^t w_{21}(\tau)F_+^1(\tau)\,d\tau =\\
=& w_{21}(t)\int_{\sigma\infty}^t w_{22}(\tau)F_+^1(\tau)d\tau - w_{22}(t)\int_0^t w_{21}(\tau)F_+^1(\tau)d\tau = \mathcal{O}F_+^1(t)\\
X_+^1 =& w_{11}(t)\int_{\sigma\infty}^t w_{22}(\tau)F_+^1(\tau)d\tau - w_{12}(t)\int_0^t w_{21}(\tau)F_+^1(\tau)d\tau = \mathcal{O}_+F_+^1(t) \qquad (A13.4)\\
X_\downarrow^1(t) =& J^{-1}\overline{\mathcal{I}}^2 F_\uparrow^1(t), \qquad X_\uparrow^1(t) = \mathcal{I}F_\uparrow^1(t)
\end{aligned}$$

where we used the boundedness criterion of §6 to eliminate the exponentially, or linearly (in the case of $X_\downarrow$), divergent terms and to find the initial conditions.

In fact (A13.4) *can be immediately generalized to arbitrary order*:

$$
\begin{aligned}
X^h_-(t) =&\mathcal{O}F^h_+(t), &\qquad X^h_+(t) = \mathcal{O}_+F^h_+\\
X^h_\downarrow(t) =&J^{-1}\overline{\mathcal{I}}^2F^h_\uparrow(t), &\qquad X^h_\uparrow(t) = \mathcal{I}F^h_\uparrow(t)
\end{aligned}
\tag{A13.5}
$$

More explicitly:

$$
\begin{aligned}
X^1_- =&\mathcal{O}(-\partial_\varphi f), \qquad X^1_\downarrow = J^{-1}\overline{\mathcal{I}}^2(-\partial_{\vec\alpha} f)\\
X^2_- =&\mathcal{O}\Big(-\frac{J_0g_0^2}{2}\sin\varphi\,(X^1_-)^2\Big)+\mathcal{O}\Big(-\partial_{\varphi^2}fX^1_- - \partial_{\vec\alpha\varphi}f\cdot X^1_\downarrow\Big)\\
X^2_\downarrow =&J^{-1}\overline{\mathcal{I}}^2\Big(-\partial_{\vec\alpha^2}fX^1_\downarrow - \partial_{\varphi\vec\alpha}fX^1_-\Big)
\end{aligned}
\tag{A13.6}
$$

leading to:

$$
\begin{aligned}
X^1_\uparrow(t) =& -\mathcal{I}(\partial_{\vec\alpha}f)\\
X^2_\uparrow(t) =& -\mathcal{I}\Big(\partial_{\vec\alpha\varphi}f\,\mathcal{O}(-\partial_\varphi f)\Big)-\mathcal{I}\Big(\partial_{\vec\alpha^2}fJ^{-1}\overline{\mathcal{I}}^2(-\partial_{\vec\alpha}f)\Big)\\
X^3_\uparrow(t) =& -\mathcal{I}\Big(\partial_{\vec\alpha\varphi}f\mathcal{O}\big(-\frac{J_0g_0^2}{2}\sin\varphi\mathcal{O}(\partial_\varphi f)\mathcal{O}(\partial_\varphi f)\big)\Big)+\\
&-\mathcal{I}\Big(\partial_{\vec\alpha\varphi}f\mathcal{O}\big(\partial_{\varphi^2}f\mathcal{O}(\partial_\varphi f)\big)\Big)-\mathcal{I}\Big(\partial_{\vec\alpha\varphi}f\mathcal{O}\big(\partial_{\vec\alpha\varphi}f\,J^{-1}\overline{\mathcal{I}}^2(\partial_{\vec\alpha}f)\big)\Big)+\\
&-\mathcal{I}\Big(\partial_{\vec\alpha^2}fJ^{-1}\overline{\mathcal{I}}^2\big(\partial_{\vec\alpha^2}fJ^{-1}\overline{\mathcal{I}}^2(\partial_{\vec\alpha}f)\big)\Big)-\mathcal{I}\Big(\partial_{\vec\alpha^2}f\,J^{-1}\overline{\mathcal{I}}^2\big(\partial_{\vec\alpha\varphi}f\mathcal{O}(\partial_\varphi f)\big)\Big)+\\
&-\frac{1}{2}\mathcal{I}\Big(\partial_{\vec\alpha^3}\,fJ^{-1}\overline{\mathcal{I}}^2(\partial_{\vec\alpha}f)J^{-1}\overline{\mathcal{I}}^2(\partial_{\vec\alpha}f)\Big)-\mathcal{I}\Big(\partial_{\varphi\vec\alpha^2}fJ^{-1}\overline{\mathcal{I}}^2(\partial_{\vec\alpha}f)\mathcal{O}(\partial_\varphi f)\Big)+\\
&-\frac{1}{2}\mathcal{I}\Big(\partial_{\vec\alpha\varphi^2}f\mathcal{O}(\partial_\varphi f)\mathcal{O}(\partial_\varphi f)\Big)
\end{aligned}
\tag{A13.7}
$$

We shall see that, setting $E^\lambda_\nu(t)\equiv e^{i\lambda\vec\nu\cdot\vec\omega\,t}e^{i\lambda n\varphi(t)}$, with $\lambda=\pm1$, the splitting and the phase shifts can be expressed in terms of the matrices:

$$
\begin{aligned}
H^1_{\lambda\nu}(t) =& \frac{1}{2i\lambda}\mathcal{I}(E^\lambda_\nu)(t)\\
H^2_{\lambda\nu\rho\mu} =& \frac{1}{2^2i\rho}\big(nm\mathcal{I}E^\lambda_\nu\mathcal{O}(E^\rho_\mu)+\vec\mu\cdot J^{-1}\vec\nu\,\mathcal{I}(E^\lambda_\nu\overline{\mathcal{I}}^2E^\rho_\mu)\big)\\
H^3_{\lambda_1\nu_1\lambda_2\nu_2\lambda_3\nu_3} =& \frac{\vec\nu_1\cdot J^{-1}\vec\nu_2\,\vec\nu_2\cdot J^{-1}\vec\nu_3}{2^3i\lambda_3}\mathcal{I}(E^{\lambda_1}_{\nu_1}\overline{\mathcal{I}}^2(E^{\lambda_2}_{\nu_2}\overline{\mathcal{I}}^2E^{\lambda_3}_{\nu_3}))+\\
+\frac{1}{2}&\frac{\vec\nu_1\cdot J^{-1}\vec\nu_2\,\vec\nu_1\cdot J^{-1}\vec\nu_3}{2^3i\lambda_1\lambda_2\lambda_3}\mathcal{I}\Big(E^{\lambda_1}_{\nu_1}(\overline{\mathcal{I}}^2E^{\lambda_2}_{\nu_2})\,(\overline{\mathcal{I}}^2E^{\lambda_3}_{\nu_3})\Big)+
\end{aligned}
\tag{A13.8}
$$

$$+\frac{n_2 n_3 \vec{\nu}_1\cdot J^{-1}\vec{\nu}_2}{2^3 i\lambda_3}\mathcal{I}\Big(E^{\lambda_1}_{\nu_1}\overline{\mathcal{I}}^2(E^{\lambda_2}_{\nu_2}\mathcal{O}E^{\lambda_3}_{\nu_3})\Big)+\frac{n_1 n_2 \vec{\nu}_2\cdot J^{-1}\vec{\nu}_3}{2^3 i\lambda_3}\mathcal{I}\Big(E^{\lambda_1}_{\nu_1}\mathcal{O}(E^{\lambda_2}_{\nu_2}\overline{\mathcal{I}}^2 E^{\lambda_3}_{\nu_3})\Big)+$$

$$+\frac{n_1 n_3 \vec{\nu}_1\cdot J^{-1}\vec{\nu}_2}{2^3 i\lambda_1\lambda_2\lambda_3}\mathcal{I}\Big(E^{\lambda_1}_{\nu_1}\overline{\mathcal{I}}^2(E^{\lambda_2}_{\nu_2}\mathcal{O}(E^{\lambda_3}_{\nu_3}))\Big)+$$

$$-\frac{n_1 n_2 n_3}{2^3\lambda_2\lambda_3}\mathcal{I}\Big(E^{\lambda_1}_{\nu_1}\mathcal{O}(-\frac{J_0 g_0^2}{2}\sin\varphi\,\mathcal{O}(E^{\lambda_2}_{\nu_2})\mathcal{O}(E^{\lambda_3}_{\nu_3}))\Big)+$$

$$+\frac{n_1 n_2^2 n_3}{2^3 i\lambda_3}\mathcal{I}\Big(E^{\lambda_1}_{\nu_1}\mathcal{O}(E^{\lambda_2}_{\nu_2}\mathcal{O}(E^{\lambda_3}_{\nu_3}))\Big)+\frac{n_1^2 n_2 n_3}{2}\frac{1}{2^3 i\lambda_1\lambda_2\lambda_3}\mathcal{I}\Big(E^{\lambda_1}_{\nu_1}\mathcal{O}(E^{\lambda_2}_{\nu_2})\mathcal{O}(E^{\lambda_3}_{\nu_3})\Big)$$

where $\mu\equiv(m,\vec{\mu})$ each of the addends above arises from a corresponding one in (A13.7): if we label 12345678 the addends in (A13.7) then they generate the r.h.s of (A13.8) in the order 46537128.

In fact one finds:

$$\begin{aligned}\vec{Q}^1_\uparrow(\vec{\alpha})=&\sum_\sigma\sigma\sum_{\lambda\nu}\vec{\nu}f_\nu e^{i\lambda\vec{\nu}\cdot\vec{\alpha}}H^1_{\lambda\nu}(0^\sigma)\\ \vec{Q}^2_\uparrow(\vec{\alpha})=&\sum_\sigma\sigma\sum_{\lambda\nu\rho\mu}\vec{\nu}f_\nu f_\mu e^{i(\lambda\vec{\nu}+\rho\vec{\mu})\cdot\vec{\alpha}}H^2_{\lambda\nu\rho\mu}(0^\sigma)\\ \vec{Q}^3_\uparrow(\vec{\alpha})=&\sum_\sigma\sigma\sum_{\lambda_i\nu_i}\vec{\nu}_1(\prod f_{\vec{\nu}_i})e^{i\sum\lambda_i\vec{\nu}_i\cdot\vec{\alpha}}H^3_{\lambda_1\nu_1\lambda_2\nu_2\lambda_3\nu_3}(0^\sigma)\end{aligned}\tag{A13.9}$$

Note that the $\vec{Q}_\uparrow$ functions are naturally expressed in terms of the $\vec{\alpha}$ variables in (A13.9), and the $\vec{\psi}$ variables of the definition (10.7) do not appear explicitly.

The scattering phase shifts are related to the time average of $-\vec{X}^h_\downarrow$, see (10.30),(10.31), so that $\vec{\sigma}^h[\vec{\alpha}]\equiv\,-J^{-1}\sum_\sigma(\mathcal{I}^2F_\uparrow)(0^\sigma)\sigma$:

$$\begin{aligned}\vec{\sigma}^1[\vec{\alpha}]=&-J^{-1}\sum_\sigma\sigma\sum_{\lambda\nu}\vec{\nu}f_\nu e^{i\lambda\vec{\nu}\cdot\vec{\alpha}}\mathcal{I}(H^1_{\lambda\nu})(0^\sigma)\\ \vec{\sigma}^2[\vec{\alpha}]=&-J^{-1}\sum_\sigma\sigma\sum_{\lambda\nu\rho\mu}\vec{\nu}f_\nu f_\mu e^{i(\lambda\vec{\nu}+\rho\vec{\mu})\cdot\vec{\alpha}}\mathcal{I}(H^2_{\lambda\nu\rho\mu})(0^\sigma)\\ \vec{\sigma}^3[\vec{\alpha}]=&-J^{-1}\sum_\sigma\sigma\sum_{\lambda_i\nu_i}\vec{\nu}_1(\prod f_{\nu_i})e^{i\sum\lambda_i\vec{\nu}_i\cdot\vec{\alpha}}\mathcal{I}(H^3_{\lambda_1\nu_1\lambda_2\nu_2\lambda_3\nu_3})(0^\sigma)\end{aligned}\tag{A13.10}$$

To study the H matrices we introduce:

$$\begin{aligned}\varepsilon^{n\lambda}(t)=&e^{i\lambda\varphi(t)n}-1,\qquad\qquad \sin n\varphi(t)\,\sinh gt=2n+\gamma_n(t)\\ \bar{\varepsilon}^n(t)=&\varepsilon^{n\lambda}(t)-i\lambda\sin n\varphi(t)\equiv\,\cos n\varphi(t)-1\end{aligned}\tag{A13.11}$$

with $\bar{\varepsilon}^n,\gamma_n\xrightarrow[t\to\pm\infty]{}0$ faster than $e^{-g\,\sigma t}$ by $\simeq e^{-g\,\sigma t}$. And we see that:

$$\sum_\sigma H^1_{\lambda\nu}(0^\sigma)\sigma=\sum_\sigma\frac{1}{2i\lambda}\mathcal{I}\big(e^{i\lambda\vec{\nu}\cdot\vec{\omega}\,\tau}(1+\varepsilon^{n\lambda}(\tau))\big)(0^\sigma)=\frac{1}{2i\lambda}\int_{+\infty}^{-\infty}\varepsilon^{n\lambda}(\tau)e^{i\lambda\vec{\nu}\cdot\vec{\omega}\,\tau}d\tau\tag{A13.12}$$

Also $4i\rho\sum_\sigma \sigma H^2_{\lambda\nu\rho\mu}$ is given by:

$$
\begin{aligned}
&(\vec\nu J^{-1}\cdot\vec\mu)\int_{+\infty}^{-\infty}\Big(\frac{\varepsilon^{n\lambda}(t)}{(i\rho\vec\omega\cdot\vec\mu)^2}+\frac{\varepsilon^{m\rho}(t)}{(i\lambda\vec\omega\cdot\vec\nu)^2}\Big)e^{i(\lambda\vec\nu+\rho\vec\mu)\cdot\vec\omega t}dt+\\
&-(\vec\nu J^{-1}\cdot\vec\mu)\int_{+\infty}^{-\infty}\Big(\frac{\varepsilon^{n\lambda}(t)e^{i\lambda\vec\nu\cdot\vec\omega t}}{(i\rho\vec\omega\cdot\vec\mu)^2}+\frac{\varepsilon^{m\rho}(t)e^{i\rho\vec\mu\cdot\vec\omega t}}{(i\lambda\vec\omega\cdot\vec\nu)^2}\Big)dt+\\
&+(\vec\nu J^{-1}\cdot\vec\mu)(\int_{+\infty}^{-\infty}\varepsilon^{n\lambda}(t)e^{i\lambda\vec\nu\cdot\vec\omega\, t}dt)(\int_{+\infty}^{0}\varepsilon^{m\rho}(t)e^{i\rho\vec\mu\cdot\vec\omega\, t}t\,dt)+\\
&+(\vec\nu J^{-1}\cdot\vec\mu)(\int_{+\infty}^{-\infty}\varepsilon^{m\rho}(t)e^{i\rho\vec\mu\cdot\vec\omega\, t}dt)(\int_{-\infty}^{0}\varepsilon^{n\lambda}(t)e^{i\lambda\vec\nu\cdot\vec\omega\, t}t\,dt)+\\
&+(\vec\nu J^{-1}\cdot\vec\mu)\int_{+\infty}^{-\infty}\varepsilon^{n\lambda}(t)e^{i\lambda\vec\nu\cdot\vec\omega\, t}\int_{+\infty}^{t}e^{i\rho\vec\mu\cdot\vec\omega\,\tau}(t-\tau)\varepsilon^{m\rho}(\tau)\,d\tau+4i\rho\vec\Delta^2
\end{aligned}
\tag{A13.13}
$$

where $\vec\Delta^2$ is the same expression evaluated at $J^{-1}=0$.

If $\tilde w^n_\lambda(t)\equiv w^0_{21}(t)e^{in\lambda\varphi(t)}+ic\lambda\gamma_n(t)+\bar\varepsilon^n(t)c\sinh gt$ (with $w^0_{12}\equiv cgt(\cosh gt)^{-1}$ and $c=(2J_0g_0)^{-1}$, see (10.13) and (A9.8)), $\vec\Delta^2$ is:

$$
\begin{aligned}
&\frac{4i\rho\vec\Delta^2}{nm}=\int_{+\infty}^{-\infty}\int_{+\infty}^{t}\big(e^{i\lambda\vec\nu\cdot\vec\omega\, t}\tilde w^n_\lambda(t)w_{22}(\tau)E^\rho_\mu(\tau)-e^{i\rho\vec\mu\cdot\vec\omega\,\tau}\tilde w^m_\rho(\tau)w_{22}(t)E^\lambda_\nu(t)\big)\,d\tau dt+\\
&-(\int_{+\infty}^{-\infty}w_{22}(\tau)E^\rho_\mu(\tau)d\tau)(\int_0^{-\infty}e^{i\lambda\vec\nu\cdot\vec\omega\, t}\tilde w^n_\lambda(t)dt)+\\
&-(\int_{+\infty}^{-\infty}w_{22}(\tau)E^\lambda_\nu(\tau)d\tau)(\int_0^{+\infty}e^{i\rho\vec\mu\cdot\vec\omega\, t}\tilde w^m_\rho(t)dt)+\\
&-\frac{c}{2}\int_{+\infty}^{-\infty}\Big[\big(\frac{e^{i\lambda\vec\nu\cdot\vec\omega\, t}e^{gt}-1}{i\lambda\vec\omega\cdot\vec\nu+g}-\frac{e^{i\lambda\vec\nu\cdot\vec\omega\, t}e^{-gt}-1}{i\lambda\vec\omega\cdot\vec\nu-g}\big)w_{22}(t)E^\rho_\mu(t)\Big]\,dt+\\
&-\frac{c}{2}\int_{+\infty}^{-\infty}\Big[\big(\frac{e^{i\rho\vec\mu\cdot\vec\omega\, t}e^{gt}-1}{i\rho\vec\omega\cdot\vec\mu+g}-\frac{e^{i\rho\vec\mu\cdot\vec\omega\, t}e^{-gt}-1}{i\rho\vec\omega\cdot\vec\mu-g}\big)w_{22}(t)E^\lambda_\nu(t)\Big]\,dt+\\
&-\int_{+\infty}^{-\infty}w_{22}(t)\Big(\frac{2ni\lambda c}{i\lambda\vec\nu\cdot\vec\omega}E^\rho_\mu(t)(e^{i\lambda\vec\nu\cdot\vec\omega\, t}-1)+\frac{2mi\rho c}{i\rho\vec\mu\cdot\vec\omega}E^\lambda_\nu(t)(e^{i\rho\vec\mu\cdot\vec\omega\, t}-1)\Big)\,dt
\end{aligned}
\tag{A13.14}
$$

It is easy to see that the above expressions (A13.13),(A13.14) *are symmetric in the exchange* $\lambda\,\nu\longleftrightarrow\rho\,\mu$. Furthermore the terms with $e^{\pm gt}$ are slightly improperly written as the $\mathcal{I}$ operation is not exactly an ordinary integral as such terms contain quantities which oscillate at ∞. To write them correctly we introduce $e^{in\lambda\varphi(t)}\tanh gt\equiv\vartheta^\lambda_n(t)+\tanh gt$. Recalling that $w_{22}=1/\cosh gt$ and setting $\Omega\equiv(\lambda\vec\nu+\rho\vec\mu)\cdot\vec\omega$, we see that the (four) just mentioned terms can be written as:

$$
\begin{aligned}
&-c\int_{+\infty}^{-\infty}e^{i\Omega t}\Big(-\frac{i\rho\vec\mu\cdot\vec\omega}{\vec\omega\cdot\vec\mu^2+g^2}\vartheta^\lambda_n(t)+\frac{g(e^{i\varphi(t)n\lambda}-1)}{\vec\omega\cdot\vec\mu^2+g^2}+(\lambda\nu\longleftrightarrow\rho\mu)\Big)+\\
&+c\sum_\sigma\sigma\mathcal{I}\Big(e^{i\Omega t}\Big(\frac{i\rho\vec\mu\cdot\vec\omega}{\vec\omega\cdot\vec\mu^2+g^2}+\frac{i\lambda\vec\nu\cdot\vec\omega}{\vec\omega\cdot\vec\nu^2+g^2}\Big)\tanh gt\Big)(0^\sigma)
\end{aligned}
\tag{A13.15}
$$

and the last row can be explicitly computed in terms of the coefficients of the series: $(1-x^2)(1+x^2)^{-1} \equiv \sum_{k=0}^{\infty} t_k x^{2k}$:

$$-2c\Big(\frac{i\rho\vec{\mu}\cdot\vec{\omega}}{\vec{\omega}\cdot\vec{\mu}^2+g^2}+\frac{i\lambda\vec{\nu}\cdot\vec{\omega}}{\vec{\omega}\cdot\vec{\nu}^2+g^2}\Big)\sum_{k=1}^{\infty} t_k \frac{2i\Omega}{\Omega^2+4g^2k^2} \tag{A13.16}$$

The above formulae show that, if $l=2$, $\partial^p_{\vec{\alpha}}\vec{Q}^2_{\uparrow}(\vec{0})$ is exponentially small if $\omega\to\infty$. This, together with the results of §9, implies immediately that also $\partial^p\vec{\sigma}^2|_{\alpha=0}$ are exponentially small, see also appendix A11.

If $l=2$ the homoclinic splitting has first derivative exponentially small to all orders and higher derivatives exponentially small at least to second order; the same holds for $l>2$ if $\vec{\omega}$ is given by the first of (10.9). In fact, by §9 this holds to all orders.

Before proceeding to a higher order calculation we examine a case with $l>2$ and $\vec{\omega}$ given by the second of (10.9): *i.e.* a mixed case in which ω_1 is fast and ω_j, $j>1$ are slow.

We fix our attention on the simple model:

$$f(\vec{\alpha},\varphi) = f_1\cos(\nu_1\alpha_1+n_1\varphi)+f_2\cos(\nu_2\alpha_2+n_2\varphi) \tag{A13.17}$$

which has only two *modes* $\vec{\nu}$: $\vec{\nu}=\vec{\vartheta}_1=\nu_1(1,0)$ and $\vec{\nu}=\vec{\vartheta}_2=\nu_2(0,1)$.

In the latter case, assuming $\omega_1=\bar{\omega}_1\eta^{-1/2}$, $\omega_2=\bar{\omega}_2\eta^{1/2}, g=O(1)$, we study the first derivatives of $\vec{Q}_{\uparrow}(\vec{\alpha})$ to second order.

If we study the α_1 derivatives of $\vec{Q}_{\uparrow}(\vec{\alpha})$ we see that we must consider in (A13.9) only modes $\vec{\mu},\vec{\nu}$ such that $\lambda\nu_1+\rho\mu_1\neq 0$. Since the case $\vec{\nu}=\vec{\mu}=\vec{\vartheta}_1$ has to be discarded because it gives exponentially small contributions to the (A13.14) we see that the only terms that are not obviously contributing exponentially small quantities are pairs $\vec{\nu},\vec{\mu}$ with $\lambda\nu_1+\rho\mu_1\neq 0$, $\vec{\nu}\cdot\vec{\mu}=0$. So that only the part $\vec{\Delta}^2$ can contribute to the α_1 derivatives of $\vec{Q}^2_{\uparrow}(\vec{\alpha})$ in (A13.14), at $\vec{\alpha}=\vec{0}$.

If we consider (A13.14),(A13.15) we see that many terms are exponentially small as $\omega_1\to\infty$ if $\lambda\nu_1+\rho\mu_1\neq 0$. The part of $\tilde{\Delta}^2$ of $\vec{\Delta}^2$ which is not obviously exponentially small as $\omega_1\to\infty$ corresponds to:

$$\begin{aligned}\frac{4i\rho\tilde{\Delta}^2_{\lambda\vec{\nu}\rho\vec{\mu}}}{nm} =& -\Big(\int_{+\infty}^{-\infty} w_{22}(t)E^{\rho}_{\mu}(t)dt\Big)\Big(\int_0^{-\infty} e^{i\lambda\vec{\nu}\cdot\vec{\omega}\,\tau}\tilde{w}^n_{\lambda}(\tau)d\tau\Big)+\\ & -\Big(\int_{+\infty}^{-\infty} w_{22}(t)E^{\lambda}_{\nu}(t)dt\Big)\Big(\int_0^{+\infty} e^{i\rho\vec{\mu}\cdot\vec{\omega}\,\tau}\tilde{w}^m_{\rho}(\tau)d\tau\Big)+\\ & +cg\int_{+\infty}^{-\infty} w_{22}(t)\Big(\frac{E^{\rho}_{\mu}(t)}{\vec{\omega}\cdot\vec{\nu}^2+g^2}+\frac{E^{\lambda}_{\nu}(t)}{\vec{\omega}\cdot\vec{\mu}^2+g^2}\Big)\,d\tau+\\ & +2c\int_{+\infty}^{-\infty} w_{22}(t)\Big(\frac{nE^{\rho}_{\mu}(t)}{\vec{\omega}\cdot\vec{\nu}}+\frac{mE^{\lambda}_{\nu}(t)}{\vec{\omega}\cdot\vec{\mu}}\Big)\,d\tau+\\ & -2c\Big(\frac{i\rho\vec{\mu}\cdot\vec{\omega}}{\vec{\omega}\cdot\vec{\mu}^2+g^2}+\frac{i\lambda\vec{\nu}\cdot\vec{\omega}}{\vec{\omega}\cdot\vec{\nu}^2+g^2}\Big)\sum_{k=1}^{\infty} t_k\frac{2i\Omega}{\Omega^2+4g^2k^2}\end{aligned} \tag{A13.18}$$

We look for $\partial_{\alpha_h} Q^2_{\uparrow k}(\vec\alpha)|_{\vec\alpha=\vec 0} \equiv \delta_{hk}$ and we see that (A13.18) contributes, taking into account the parity properties of the integrals, $\sum 2^{-2}\nu_k\mu_h 4i\rho\tilde\Delta^2_{\lambda\vec\nu\rho\vec\mu}$. Hence for $h=1$ it must be $\vec\mu=\vec\vartheta_1$ and $\vec\nu=\vec\vartheta_2$, otherwise we se that (A13.18) contributes exponentially small quantities. Therefore up to exponentially small terms δ_{12} is given by:

$$
\begin{aligned}
\delta_{12} = & -\nu_1\nu_2\varphi_1 f_2\Big(n_2\int_{+\infty}^{-\infty} w_{22}\cos(\nu_2\omega_2 t+n_2\varphi(t))\Big)\cdot \qquad (A13.19)\\
& n_1\Big[\int_0^{+\infty} 2^{-1}\sum_\rho e^{i\omega_1\nu_1\rho t}\tilde w_\rho^{n_1}(t)\,dt - c\frac{g}{\nu_1^2\omega_1^2+g^2}-\frac{2n_1 c}{\nu_1\omega_1}d\tau\Big] = \\
= & -\nu_1\nu_2 f_1 f_2 c\Big(\int_{+\infty}^{-\infty} n_2 w_{22}\cos(\nu_2\omega_2 t+n_2\varphi(t))dt\Big)\cdot n_1\Big[-\frac{g}{\nu_1^2\omega_1^2+g^2}-\frac{2n_1}{\nu_1\omega_1}- \\
& -\int_{+\infty}^{0}\big(\frac{gt}{\cosh gt}\cos(\nu_1\omega_1 t+n_1\varphi)-\gamma_{n_1}(t)\sin\nu_1\omega_1 t+\bar\varepsilon^{n_1}(t)\sinh gt\cos\nu_1\omega_1 t\big)dt\Big] = \\
= & \delta_{21} = 2(-1)^{n_1}\, c f_1 f_2\,\frac{\nu_2\omega_2}{g\nu_1^2\omega_1^2}\nu_1 n_1\nu_2 n_2 K_{n_2}
\end{aligned}
$$

where $c\equiv 1/(2g_0J_0)$ and the equalities holds up to exponentially small terms as $\omega_1\to\infty$ except the last which holds up to a factor $1+O(g/\omega_1+\omega_2/g)$, and:

$$
K_n \equiv \int_{-\infty}^{\infty} t\frac{\sin n\varphi_0(t)}{\cosh t}dt,\qquad \varphi_0(t)\equiv 4\,\mathrm{arctg}\, e^{-t} \qquad (A13.20)
$$

The symmetry $\delta_{hk}=\delta_{kh}$ is a direct consequence of the symmetry remarked after (A13.14) above; the asymptotic analysis is made easier if one remarks that:

$$
\begin{aligned}
&\gamma_n(0) = -2n, \qquad \int_{-\infty}^{+\infty}(\cos n\varphi)(\cosh gt)^{-1}dt = 0 \\
&\int_{-\infty}^{+\infty} t\frac{\sin\varphi_0(t)}{\cosh t}dt = 2, \qquad \int_{-\infty}^{+\infty} t\frac{\sin 2\varphi_0(t)}{\cosh t}dt = \frac{10\pi}{3}
\end{aligned}
\qquad (A13.21)
$$

where the last two formulae are useful in the applications of §13.

The above formulae show that, if $l=2$, $\partial^p_{\vec\alpha}\vec Q^2_\uparrow(\vec 0)$ is exponentially small if $\omega\to\infty$. This, together with the results of §9, implies immediately that also $\partial^p\vec\sigma^2|_{\alpha=0}$ are exponentially small, see also appendix A11.

If $l=2$ the homoclinic splitting has first derivative exponentially small to all orders and higher derivatives exponentially small at least to second order; the same holds for $l>2$ if $\vec\omega$ is given by the first of (10.9). In fact, by §9 this holds to all orders.

Before proceeding to a higher order calculation we examine a case with $l=3$ and $\vec\omega$ given by the second of (10.9): *i.e.* a mixed case in which ω_1 is fast and ω_2 is slow, but we do not suppose the orthogonality between the slow and fast modes. More precisely we assume that $\omega_1 = \bar\omega_1\eta^{-1/2}$, $\omega_2 = \bar\omega_2\eta^{1/2}$, $g = O(1)$; we shall also assume here, for simplicity, that f is a trigonometric polynomial *i.e.* $f_\nu = 0$ if $|\nu| > N$ for a suitable $N > 0$.

Calling $\mathcal{F} = \{\mu : \vec\mu\cdot\vec\omega = O(\eta^{-1/2})\}$ = "fast modes" and $\mathcal{S} = \{\mu : \vec\mu\cdot\vec\omega = O(\eta^{1/2})\}$ = "slow modes", the above formulae yield easily that the terms in (A13.13)(A13.14) with μ,ν both fast give an exponentially small contribution and that the leading contribution to δ_{12} is given by:

$$\delta_{12} = \delta_{21} = \sum_{\mu\in\mathcal{F},\vec\nu\in\mathcal{S}} \frac{\mu_1\nu_2(-1)^m}{g\vec\mu\cdot\vec\omega^2}[-(\vec\nu\cdot J^{-1}\vec\mu)\hat K_n + 2nmcK_n\vec\nu\cdot\vec\omega] \tag{A13.22}$$

where $c \equiv (2g_0J_0)^{-1}$, and:

$$\hat K_n \equiv \int_{-\infty}^{\infty}(\cos n\tilde\varphi - 1)du\ ,\quad K_n \equiv \int_{-\infty}^{\infty}\frac{u\sin n\tilde\varphi}{\cosh u}du\ ,\quad \tilde\varphi(u) \equiv 4\arctan e^{-u} \tag{A13.23}$$

Note also that $\delta_{11} = 0$ up to exponentially small terms (in the general case). This proves c) for what concerns the homoclinic splitting. The check concerning the phase shifts is a similar calculation (and it works for the above considered example, (A13.17)), but we omit the details.

Since the first order yields exponentially small contributions to the α_1-derivatives of both components of $\vec Q_\uparrow$ as well as to $\partial_{\alpha_2}Q_1$, we see that (A13.22) gives that the intersection matrix determinant is, to leading order, equal to $-\delta_{12}\delta_{21}$ and this proves d).

2) We study, now, the third order for the simple model:

$$f(\alpha,\varphi) = f\cos(\alpha+\varphi) \tag{A13.24}$$

with the purpose of performing an instructive calculation showing quite clearly one among several cancellation mechanisms behind the smallness of the homoclinic angles when $l = 2$.

We shall study the contributions to $\Delta^3(\vec\alpha)$ of order J^{-2} as $J\to 0$. Note that $\Delta^3(\vec\alpha)$ is a polynomial of degree 2 in J^{-1}.

We see from (A13.8) that in this case we must consider:

$$\sum_{\lambda_1\lambda_2\lambda_3}\Big[\frac{2^{-3}}{i\lambda_3}\mathcal{I}\Big(E^{\lambda_1}\overline{\mathcal{I}}^2\big(E^{\lambda_2}\overline{\mathcal{I}}^2E^{\lambda_3}\big)\Big) + \frac{2^{-3}}{2i\lambda_1\lambda_2\lambda_3}\mathcal{I}\Big(E^{\lambda_1}\big(\overline{\mathcal{I}}^2E^{\lambda_2}\big)\big(\overline{\mathcal{I}}^2E^{\lambda_3}\big)\Big)\Big] \tag{A13.25}$$

Each E^λ is split as $e^{i\lambda\omega t}$ plus $e^{i\lambda\omega t}\varepsilon^\lambda$ and each addend in (A13.25) generates eight terms. To avoid considering improper integrals we shall consider only the terms obtained by operating the first three second choices.

The same mechanism, once understood applies equally well to the other seven terms (and actually it is more convenient not to separate them and to operate with the improper integrals). We hope to examine the general theory elsewhere.

Setting $\vec{\omega}\cdot\vec{\nu}_j \equiv \omega_j$, and omitting writing the summation symbol over $\lambda_1,\lambda_2,\lambda_3$, we get a contribution denoted C_3^σ given by:

$$
\begin{aligned}
&\frac{2^{-3}}{i\lambda_3}\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\int_{\sigma\infty}^{t} d\tau (t-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}\int_{\sigma\infty}^{\tau} d\vartheta\,(\tau-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}+ \qquad (A13.26)\\
&+\frac{2^{-3}}{2i\lambda_1\lambda_2\lambda_3}\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\Big(\int_{\sigma\infty}^{t} d\tau (t-\tau)\varepsilon^{\lambda_2}e^{i\lambda_2\omega_2\tau}\Big)\big(\int_{\sigma\infty}^{t} d\vartheta\,(t-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}\Big)-\\
&-\frac{2^{-3}}{i\lambda_3}\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\int_{\sigma\infty}^{t} d\tau (t-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}\int_{\sigma\infty}^{0} d\vartheta\,(-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}+\\
&-\frac{2^{-3}}{2i\lambda_1\lambda_2\lambda_3}\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\Big(\int_{\sigma\infty}^{t} d\tau (t-\tau)\varepsilon^{\lambda_2}e^{i\lambda_2\omega_2\tau}\Big)\big(\int_{\sigma\infty}^{0} d\vartheta\,(-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}\Big)-\\
&-\frac{2^{-3}}{i\lambda_3}\Big(\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\Big)\Big(\int_{\sigma\infty}^{0} d\tau\,(-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}\int_{\sigma\infty}^{\tau} d\vartheta\,(\tau-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}\Big)+\\
&-\frac{2^{-3}}{2i\lambda_1\lambda_2\lambda_3}\Big(\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\Big(\int_{\sigma\infty}^{0} d\tau(-\tau)\varepsilon^{\lambda_2}e^{i\lambda_2\omega_2\tau}\Big)\int_{\sigma\infty}^{t} d\vartheta\,(t-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}\Big)-\\
&+\frac{2^{-3}}{i\lambda_3}\big(\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\big)\big(\int_{\sigma\infty}^{0} d\tau(-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}\big)\big(\int_{\sigma\infty}^{0} d\vartheta\,(-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}\big)+\\
&+\frac{2^{-3}}{2i\lambda_1\lambda_2\lambda_3}\Big(\int_{\sigma\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\Big)\Big(\int_{\sigma\infty}^{0} d\tau(-\tau)\varepsilon^{\lambda_2}e^{i\lambda_2\omega_2\tau}\Big)\Big(\int_{\sigma\infty}^{0} d\vartheta\,(-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}\Big)
\end{aligned}
$$

and, computing $\sum\sigma C_3^\sigma \equiv C_3$ up to terms giving, obviously, exponentially small contributions as $\eta\to 0$, we see that C_3 is given by:

$$
\begin{aligned}
&-\frac{2^{-3}}{i\lambda_3}\int_{-\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\int_{-\infty}^{+\infty} d\tau (t-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}\int_{+\infty}^{\tau} d\vartheta\,(\tau-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}+\\
&-\frac{2^{-3}}{i\lambda_3}\int_{+\infty}^{-\infty} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\int_{+\infty}^{t} d\tau (t-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}\int_{+\infty}^{0} d\vartheta\,(-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}+ \qquad (A13.27)\\
&-\frac{2^{-3}}{2i\lambda_1\lambda_2\lambda_3}\int_{+\infty}^{-\infty} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\int_{+\infty}^{t} d\tau (t-\tau)\varepsilon^{\lambda_2}e^{i\lambda_2\omega_2\tau}\int_{+\infty}^{0} d\vartheta\,(-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}-\\
&+\frac{2^{-3}}{i\lambda_3}\int_{-\infty}^{0} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\int_{-\infty}^{+\infty} d\tau(-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}\int_{+\infty}^{\tau} d\vartheta\,(\tau-\vartheta)e^{i\lambda_3\omega_3\vartheta}\varepsilon^{\lambda_3}+\\
&-\frac{2^{-3}}{2i\lambda_1\lambda_2\lambda_3}\int_{+\infty}^{-\infty} dt e^{i\lambda_1\omega_1 t}\varepsilon^{\lambda_1}\int_{+\infty}^{t} d\tau (t-\vartheta)\varepsilon^{\lambda_3}e^{i\lambda_3\omega_3\vartheta}\int_{+\infty}^{0} d\tau\,(-\tau)e^{i\lambda_2\omega_2\tau}\varepsilon^{\lambda_2}
\end{aligned}
$$

and we easily see that only the cases $\lambda_1+\lambda_2+\lambda_3=\pm1$ can give non exponentially small contributions to C_3. But in such cases the sum of the second, third and fifth terms vanishes identically (exhibiting the mentioned cancellation mechanism); while the sum over the λ's of the first and fourth terms is also exponentially small.

Appendix A14: Development of the perturbatrix

We study here the perturbation V in the D'Alembert model of §12. We set, for a general angle x, $s_x \equiv \sin x$, $c_x \equiv \cos x$ and:

$$\begin{aligned}
&a \equiv \kappa\nu,\ b \equiv \mu\sigma,\ d \equiv -\nu, \qquad \gamma_0 = \gamma - \lambda\\
&\bar s = \sin(\lambda_T - \gamma), \qquad \bar c = \cos(\lambda_T - \gamma)\\
&c_{\chi_0} \equiv \cos\chi_0, \quad s_{\chi_0} \equiv \sin\chi_0, \quad s_{\gamma_0} = \sin\gamma_0, \quad c_{\gamma_0} = \cos\gamma_0\\
&\cos\alpha \equiv c_\alpha = a\bar s c_{\chi_0} + b\bar s + d\bar c s_{\chi_0}
\end{aligned} \tag{A14.1}$$

and, dropping everywhere below the subscript $\cdot_0$ from the angles, we have, see (12.3):

$$\begin{aligned}
\bar s =& - s_\gamma(1-2e^2 {s_\lambda}^2) + c_\gamma\,(-2es_\lambda + \frac{5}{4}e^2 s_{2\lambda})\\
\bar c =& c_\gamma(1-2e^2 {s_\lambda}^2) + s_\gamma\,(-2es_\lambda + \frac{5}{4}e^2 s_{2\lambda})
\end{aligned} \tag{A14.2}$$

up to $O(e^3)$. It follows immediately that if:

$$A \equiv a s_\gamma c_\chi + b s_\gamma - d c_\gamma s_\chi, \qquad B \equiv a c_\gamma c_\chi + b c_\gamma + d s_\gamma s_\chi \tag{A14.3}$$

it is (again up to $O(e^3)$):

$$c_\alpha^2 = A^2 + 4es_\lambda AB + 4e^2\big((B^2 - A^2){s_\lambda}^2 - \frac{5}{8}AB s_{2\lambda}\big) \tag{A14.4}$$

More explicitly we find:

$$\begin{aligned}
A^2 =& \frac14(a^2+2b^2+d^2) + \frac14(a^2-d^2)c_{2\chi} + \frac14(-a^2-2b^2+d^2)c_{2\gamma}+\\
&-\frac14(a^2+d^2)c_{2\gamma}c_{2\chi} + abc_\chi - abc_{2\gamma}c_\chi+\\
&+\frac14 ad(c_{2\gamma+2\chi} - c_{2\gamma-2\chi}) + \frac12 bd(c_{2\gamma+\chi} - c_{2\gamma+\chi})\\
B^2 =& \frac14(a^2+2b^2+d^2) + \frac14(a^2-d^2)c_{2\chi} - \frac14(-a^2-2b^2+d^2)c_{2\gamma}+\\
&+\frac14(a^2+d^2)c_{2\gamma}c_{2\chi} + abc_\chi + abc_{2\gamma}c_\chi+\\
&-\frac14 ad(c_{2\gamma+2\chi} - c_{2\gamma-2\chi}) - \frac12 bd(c_{2\gamma+\chi} - c_{2\gamma+\chi})
\end{aligned} \tag{A14.5}$$

$$
\begin{aligned}
B^2 - A^2 =& \frac{1}{2}(a^2+2b^2-d^2)c_{2\chi} + \frac{1}{4}(a^2+d^2)(c_{2\gamma+2\chi}+c_{2\gamma-2\chi}) + \\
& + ab(c_{2\gamma+\chi}+c_{2\gamma-\chi}) - \frac{1}{2}ad(c_{2\gamma+2\chi}-c_{2\gamma-2\chi}) - bd(c_{2\gamma+\chi}-c_{2\gamma-\chi}) \\
4ABs_\lambda =& -\frac{1}{2}(a^2+2b^2-d^2)(c_{2\gamma+\lambda}-c_{2\gamma-\lambda}) - \frac{1}{2}(a^2+d^2)c_{2\chi}(c_{2\gamma+\lambda}-c_{2\gamma-\lambda}) + \\
& - 2abc_\chi(c_{2\gamma+\lambda}-c_{2\gamma-\lambda}) + adc_{2\gamma}(c_{\lambda+2\chi}-c_{\lambda-2\chi}) + 2bdc_{2\gamma}(c_{\lambda+\chi}-c_{\lambda-\chi})
\end{aligned}
$$

Defining $c_j \equiv c_j(K_0, M-0),\ d_j(K_0, M_0)$ by:

$$
\begin{aligned}
c_0 &\equiv \frac{1}{4}(a^2+2b^2+d^2) = \frac{1}{4}[2\sigma^2\mu^2 + (1+\kappa^2)\nu^2] \\
d_0 &\equiv \frac{1}{4}(-a^2-2b^2+d^2) = -\frac{\sigma^2}{4}(2\mu^2-\nu^2) \\
c_{\pm1} &\equiv \frac{1}{2}ab = \frac{\kappa\sigma\mu}{2}\nu\ , \quad c_{\pm2} = \frac{1}{8}(a^2-d^2) = -\frac{\sigma^2}{8}\nu^2 \\
d_{\pm1} &\equiv\ -\frac{1}{2}ab \pm \frac{1}{2}bd = \mp\frac{(1\pm\kappa)\mu\sigma}{2}\nu \\
d_{\pm2} &= -\frac{1}{8}(a^2+d^2) \pm \frac{1}{4}ad = -\frac{(1\pm\kappa)^2}{8}\nu^2
\end{aligned}
\tag{A14.6}
$$

we see from (A14.4) that:

$$
\begin{aligned}
& c_\alpha^2 \equiv \bar{V}_0 + e\bar{V}_1 + e^2\bar{V}_2 \\
& \bar{V}_0 \equiv A^2 = \sum_{j=-2}^{2} c_j \cos j\chi + d_j \cos(j\chi+2\gamma) \\
& \bar{V}_1 \equiv 4s_\lambda AB = \sum_{j=-2}^{2} 2d_j \cos(2\gamma+\lambda+j\chi) - 2d_j\cos(2\gamma-\lambda+j\chi) \\
& \bar{V}_2 \equiv\ -\frac{5}{8}4s_{2\lambda}AB + 4{s_\lambda}^2(B^2-A^2) = \qquad (A14.7) \\
& \quad = \sum_{j=-2}^{2} d_j\Big[-\frac{5}{4}\cos(2\gamma+2\lambda+j\chi) + \frac{5}{4}\cos(2\gamma-2\lambda+j\chi) + \\
& \qquad -4\cos(2\gamma+j\chi) + 2\cos(2\gamma+2\lambda+j\chi) + 2\cos(2\gamma-2\lambda+j\chi)\Big] = \\
& \quad = \sum_{j=-2}^{2} d_j\Big[\frac{3}{4}\cos(2\gamma+2\lambda+j\chi) + \frac{13}{4}\cos(2\gamma-2\lambda+j\chi) - 4\cos(2\gamma+j\chi)\Big]
\end{aligned}
$$

so that (up to order $O(e^2)$):

$$
\begin{aligned}
& (\bar{V}_0 + e\bar{V}_1 + e^2\bar{V}_2)(1 - 3ec_\lambda + \frac{3}{2}e^2 + \frac{9}{2}e^2c_{2\lambda}) = \\
& \bar{V}_0 + e(\bar{V}_1 - 3c_\lambda\bar{V}_0) + e^2(\bar{V}_2 - 3c_\lambda\bar{V}_1 + \frac{3}{2}\bar{V}_0 + \frac{9}{2}c_{2\lambda}\bar{V}_0)
\end{aligned}
\tag{A14.8}
$$

and, reinserting the lower index 0:

$$
\begin{aligned}
V_0 &= \sum_{j=-2}^{2} c_j \cos j\chi_0 + d_j \cos(2\gamma_0 + j\chi_0) \\
V_1 &= \sum_{j=-2}^{2} \Big(-3c_j \cos(\lambda_0 + j\chi_0) + \\
&\quad + \frac{1}{2} d_j \big(\cos(2\gamma_0 + \lambda_0 + j\chi_0) - 7\cos(2\gamma_0 - \lambda_0 + j\chi_0) \big) \Big) \\
V_2 &= \sum_{j=-2}^{2} \Big[c_j \Big(\frac{3}{2} \cos j\chi_0 + \frac{9}{2} \cos(2\lambda_0 + j\chi_0) \Big) + \\
&\quad + d_j \Big(\frac{17}{2} \cos(2\gamma_0 - 2\lambda_0 + j\chi_0) - \frac{5}{2} \cos(2\gamma_0 + j\chi_0) \Big) \Big]
\end{aligned}
\tag{A14.9}
$$

and the coefficients $\bar{B}^h_{rpj}$ (cfr. (12.14)) vanish unless they belong to the following list where $|j| \leq 2$:

$$
\begin{aligned}
&\bar{B}^0_{00j} \equiv c_j \ , \quad \bar{B}^0_{20j} \equiv d_j \\
&\bar{B}^1_{01j} \equiv -3c_j \ , \quad \bar{B}^1_{211} \equiv \frac{d_j}{2} \ , \quad \bar{B}^1_{2-1j} \equiv -\frac{7}{2} d_j \\
&\bar{B}^2_{00j} \equiv \frac{3}{2} c_j \ , \quad \bar{B}^2_{02j} \equiv \frac{9}{2} c_j \ , \bar{B}^2_{20j} \equiv -\frac{5}{2} d_j \ , \bar{B}^2_{2-2j} \equiv \frac{17}{2} d_j
\end{aligned}
\tag{A14.10}
$$

Next, after the linear change of variables (12.22), (12.23) one gets (12.24) with the B^h_{rpj} vanishing unless they belong to the following list where $|j| \leq 2$ and where c^0_j, d^0_j are the functions in (A14.6) evaluated at $K_0 = 2I_0 - A_0 + \overline{K}$, $M_0 = I_0 + 2\omega_T J_3$:

$$
\begin{aligned}
&B^0_{2j,2j,j} \equiv c^0_j \quad (j \neq 0) \ , \quad B^0_{2(j-1),2(j-1),j} \equiv -d^0_j \quad (j \neq 1) \\
&B^1_{2j,2j+1,j} \equiv -3c^0_j \ , \quad B^1_{2(j-1),2j-1,j} \equiv -\frac{d^0_j}{2} \ , \quad B^1_{2(j-1),2j-3,j} \equiv \frac{7}{2} d^0_j \\
&B^2_{2j,2j,j} \equiv \frac{3}{2} c^0_j \ , \quad B^2_{2j,2(j+1),j} \equiv \frac{9}{2} c^0_j \\
&B^2_{2(j-1),2(j-2),j} \equiv -\frac{17}{2} d^0_j \ , \quad B^2_{2(j-1),2(j-1),j} \equiv \frac{5}{2} d^0_j
\end{aligned}
\tag{A14.11}
$$

References

[A] Arnold, V.: *Instability of dynamical systems with several degrees of freedom*, Sov. Mathematical Dokl., 5, 581-585, 1966.

[ACKR] Amick C., Ching E.S.C., Kadanoff L.P., Rom–Kedar V.: *Beyond All Orders: Singular Perturbations in a Mapping* J. Nonlinear Sci. 2, 9–67, 1992.

[BG] Benettin, G., Gallavotti, G.: *Stability of motions near resonances in quasi-integrable hamiltonian systems*, J. Statistical Physics, 44, 293-338, 1986.

[CG] Chierchia, L., Gallavotti, G.: *Smooth prime integrals for quasi-integrable Hamiltonian systems* Il Nuovo Cimento, 67 B, 277-295, 1982.

[CZ] Chierchia, L., Zehnder, E.: *Asymptotic expansions of quasi-periodic motions*, Annali della Scuola Normale Superiore di Pisa, Serie IV Vol XVI Fasc.2 (1989).

[D] Douady, R.: *Stabilité ou instabilité des points fixes elliptiques*, Annales Scientifiques de l' Ecole Normale Superieure, 21, 1-46, 1988.

[DS] Delshams, A., Seara, M.T.: *An asymptotic expression for the splitting of separatrices of rapidly forced pendulum*, preprint 1991.

[G] Gallavotti, G.: *The elements of Mechanics*, Springer, 1983.

[Ge] Gelfreich, V. G.: *Separatrices splitting for the rapidly forced pendulum*, Preprint, 1992.

[GLT] Gelfreich, V. G., Lazutkin, V.F., Tabanov, M.B.:*Exponentially small splitting in Hamiltonian systems*, Chaos, 1 (2), 1991.

[Gr] Graff, S. M.: *On the conservation for hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations 15, 1-69, 1974.

[HMS] Holmes, P., Marsden, J., Scheurle,J: *Exponentially Small Splittings of Separatrices in KAM Theory and Degenerate Bifurcations*, Preprint, 1989.

[L] de la Place, S.: *Mécanique Céleste*, tome II, book 5, ch. I, 1799, english translation by Bodwitch, E., reprinted by Chelsea, 1966.

[La1] Lazutkin, V. F.: *The existence of caustics for a billiard problem in a convex domain*, Izv. Akad. Nauk. SSSR, 37 (1), 1973

[La2] Lazutkin, V.F.: *Separatrices splitting for standard and semistandard mappings*, Preprint, 1989.

[LW] de la Llave, R., Wayne, E.: *Whiskered Tori*, preprint 1990

[M] Moser, J.: *Convergent series expansions for quasi-periodic motions*, Matematische Annalen, 169, 136-176, 1967.

[Me] Melnikov, V.K.: *On the stability of the center for time periodic perturbations*, Trans. Moscow Math Math. Soc., 12, 1-57, 1963.

[N] Nekhorossev, N.: An exponential estimate of the time of stability of nearly integrable hamiltonian systems, Russian Mathematical Surveys, 32, 1-65, 1975.

[Nei] Neihstad, A.I.: *The separation of motions in systems with rapidly rotating phase*, PMM U.S.S.R. 48, 133–139, 1984.

[P] Poincarè, H.: *Les Méthodes nouvelles de la mécanique céleste*, 1892, reprinted by Blanchard, Paris, 1987.

[Pö] Pöschel, J.: *Integrability of Hamiltonian systems on Cantor sets*, Communications Pure Appl. Math, 35, 653-696, 1982.

[Sv] Svanidze, N.V.: *Small perturbations of an integrable dynamical system with an integral invariant*, Proceed. Steklov Institute of Math., 2, 1981

[Z] Zehnder, E.: *Generalized implicit function theorems with applications to some small divisor problems* I,II, Communications Pure Applied Mathematics, 28, 91-140, 1975.