Pendulum: separatrix splitting.

G. Gallavotti, G. Gentile, V. Mastropietro

Università di Roma 1,2,3
1 settembre 1997

Abstract: *An exact expression for the determinant of the splitting matrix is derived: it allows us to analyze the asymptotic behaviour needed to amend the large angles theorem proposed in* Ann. Inst. H. Poincaré, B-60, 1, 1994*. The asymtotic validity of Melnikov's formulae is proved for the class of models considered, which include polynomial perturbations.*

§1. Introduction.

Recently V. Gelfreich noted that a "theorem" in [CG] contains an error. The theorem gave a lower bound on the splitting angles in a three degrees of freedom system and it was needed to show the existence of heteroclinic chains in a class of hamiltonian systems with the aim of an application to a celestial mechanics problem. We correct it here by providing the correct lower bound and, at the same time, exposing again and in a more meditated form some ideas of [CG]. In the present paper we do not discuss the existence of Arnold's diffusion, [A2], in our systems. We do not discuss the celestial mechanics application of [CG] either, as two parts of it (see below) relied on the erroneous statement, and more work is needed. The present paper is, therefore, *not* a correction of the implications of the error in [CG] but only of the error itself. In order to derive the same implications further work is necessary as the erroneous result was used several times in the last three sections of [CG]. Each use has therefore to be treated separately.

Our analysis deals mainly with systems with *three* different time scales with the ratio between the smallest to the largest being $\ll 1$: *e.g.* the largest is $\eta^{-1/2}$, the intermediate is 1 and the smallest is η^a, $a \ge 0$ (the $a = 0$ case being a limiting two scales problem). The result will be called the *large angles theorem*, (theorem 2 in §5, proved in §6,§7 for the systems defined in (2.1),(7.1) below), and describes a property of a pendulum subject to both a slow periodic force and to a rapid periodic force (incommensurate to the former) that we imagine generated by a pair of rotators, whose positions are given by two angles α, λ while the pendulum position is given by an angle φ. The angles α, λ will be respectively called *slow* and *fast*.

When the system is perturbed some of the quasi periodic motions performed by rotators (or clocks) while the pendulum is in its unstable equilibrium persist. They become motions in which the pendulum is slightly moving quasi periodically around its unstable position ("without ever falling down"): one has therefore an invariant 2–dimensional torus in phase space. For each position (α, λ) of the rotators there is well defined value of the pendulum variables (I, φ) as well as of the rotators actions (A, B). And (α, λ) evolve quasi–periodically (*i.e.* (α, λ) can be written as $\alpha = \psi_1 + \Delta_1(\psi_1, \psi_2)$, $\lambda = \psi_2 + \Delta_2(\psi_1, \psi_2)$ and the time evolution is simply linear $\underline{\psi} \to \underline{\psi} + \underline{\omega}\, t$ for a suitable $\underline{\omega}$).

Obviously the above quasi periodic motion is unstable and perturbations will make the pendulum "fall". The 2–dimensional invariant tori have therefore stable and unstable manifolds: which coincide in absence of perturbations (because they correpond to the separatrix motions and the pendulum separatrix is degenerate). However under perturbations they "split" (*i.e.* become distinct 3 dimensional surfaces). Often, simply by symmetry, there is one trajectory that lies on both manifolds. Physically it corresponds to a motion that consists of just one swing of the pendulum from nearby the unstable position back to it.

If we observe the swinging trajectory at the moment it passes through the stable equilibrium position, $\varphi = \pi$, we see just a point that can be identified by the value $\underline{\alpha}_0$, at that moment, of the $\underline{\alpha}$ coordinates (where the two manifolds meet). If we move away from that point the $\underline{A}$ coordinates on the two manifolds become different, at the same $\underline{\alpha}$, and their difference is a "splitting" vector $\underline{Q}(\underline{\alpha})$. The 2×2 matrix $D = \partial_{\underline{\alpha}} \underline{Q}(\underline{\alpha}_0)$ is called the *intersection matrix.*

The homoclinic angles can be defined as the angles whose tangents are the eigenvalues of the intersection matrix. The *splitting* is usually defined as $\det D$. When the perturbing frequencies are held fixed and the perturbation is sent to 0 there is a well known asymptotic expression for the splitting, called *Melnikov's formula* and coinciding with the "first order perturbation theory result".

The splitting problem is to find under which conditions *Melnikov's formula* holds when *both* the perturbation and the shortest forcing period are sent to 0.

Informally our result is (see theorem 2 in §5 for a formal statement):

Suppose that the forcing frequency is very large, say $\eta^{-1/2}$ times the pendulum frequency, with η very small. Suppose also that the second characteristic frequency is relatively very small, say η^a times that of the pendulum with $a \ge 0$. Then there are perturbations of size $\varepsilon = O(\eta^c)$ (for some $c > 0$) such that the separatrix splitting has size given asymptotically by Melnikov's formula provided the frequencies of the motion verify a suitable diophantine condition. In fact this property holds generically, under the same conditions, for perturbations of size $\varepsilon = O(\eta^c)$ which are trigonometric polynomials.

There are many examples of systems for which the above property does not hold, see [S], [DJGS], [RW]. But the forcing frequencies relations are different.

The correct lower bound estimate on the splitting makes an analysis *to all orders* necessary: the analysis, performed here, becomes "marginal" but the final result on the existence of homoclinic splitting (*i.e.* a lower bound on it) remains valid. Therefore the present paper corrects the error in §10 of [CG] as far as its implications on the size of the splitting are concerned.

The techniques we use here *to bypass perturbation theory* were started in the Appendix A13 of [CG]: they were not pushed too far because the error made any developments unnecessary for the purposes of [CG].

The techniques were subsequently developed in [G3] (which *does not* repeat the computational error). The methods of [G3] were not developed to treat three time scales problems: the aim there being to prove the smallness of the splitting in *two time scales* problems (*i.e.* in systems not considered here with all rotators with comparably large velocity, $a = -\frac{1}{2}$). But they can be *easily* extended to three time scales problems and even lead to a remarkable *nonperturbative and exact* computation of the leading order (exponentially small) of the intersection matrix, see (6.12),(7.19): V. Gelfreich stressed necessity of a nonperturbative analysis by a simple cogent argument.

The reason why we study three time scales systems is simply that they arise in a celestial mechanics problem of interest to us, see §12 of [CG]. But most of our results hold also if the slow time scale is the same as that of the pendulum, *i.e.* the parameter a above (see also (2.1)) is $a = 0$, see also comments in §9: this is also an interesting case and it has been considered, to some extent, in [RW].

The recent work of [RW] considerably overlaps with ours: even when $a = 0$ (a two time scales problem) it cannot be used to achieve all our results because our assumptions can violate eq. (15) of [RW] (*i.e.* the bound on the constant called b in [RW], see (2.2), below, for the isochronous case and (7.3) for the anisochrounus). We also require, as an essential assumption, that the perturbations be trigonometric polynomials while the examples of [RW] require, as an essential assumption, that the perturbation has infinitely many non vanishing modes (see eq. (19) of [RW]). Another essential difference with respect to [G3] and [RW] is that we *only* study the splitting *at the homoclinic point* while they study it everywhere: hence our work is much more limited in scope. On the other hand we recover some of the results of [RW] because our proofs also apply nontrivially to two time scale cases with 3 degrees of freedom (the ones in [RW]). Nevertheless we feel that the main difference between our work and [RW] lies in the techniques: here we show that the techniques of [G3] do apply immediately to the problem. See concluding remarks for a more technical comparison.

§2. Isochronous clock–pendulum system.

Our main result concerns anisochronous systems, (7.1). Isochronous models will be considered only to illustrate the simplest cases: the cancellations that we find in the anisochronous case would look miraculous otherwise. Calling $(I,\varphi),(A,\alpha),(B,\lambda)$ pairs of canonical coordinates suppose them action–angle variables: (φ,α,λ) is a triple of angles (varying on the 3–dimensional torus $T^3=[0,2\pi]^3$) and $(I,A,B)\in R^3$.
The hamiltonian will be:

2.1 $$H=\eta^a A+\frac{I^2}{2}+\eta^{-1/2}B+\,g^2\,(\cos\varphi-1)+\mu\,\eta^c\,f(\varphi,\alpha,\lambda),\ \varepsilon\stackrel{def}{=}\mu\eta^c \tag{2.1}$$

where $\eta,\mu>0$ and f is an *even trigonometric polynomials* in the angles (φ,α,λ), *i.e.* for instance one could take $f(\varphi,\alpha,\lambda)=\big(\cos(\alpha+\varphi)+\cos(\lambda+\varphi)\big)$. The parameter η sets the ratios between time scales and is a free parameter which we take, eventually, to be close to zero in order to study asymptotic properties as $\eta\to0$.
If $\mu=0$ the 2–dimensional torus: $A=B=I=0$, $\varphi=0$ and $\underline{\alpha}=(\alpha,\lambda)$ arbitrary, is invariant and run quasi-periodically with rotation velocity $\underline{\omega}=(\eta^a,\eta^{-1/2})$: it will be called the *unperturbed torus*.
In the above case, *i.e. in the isochronous case only*, we shall suppose (not "for simplicity": but as an essential hypothesis) that the vector $\underline{\omega}=(\eta^a,\eta^{-1/2})$ is a *diophantine vector*:

2.2 $$|\,\underline{\omega}\cdot\underline{\nu}\,|>C^{-1}\eta^d|\,\underline{\nu}\,|^{-\tau} \tag{2.2}$$

for some *diophantine constant* C,τ and some $d>0$. This restricts the values of η that we can consider. For $a\in[0,\frac12]$ it still allows sequences η_k, such that (2.2) holds with prefixed $\tau>2, d=a$ and C large enough, with $\eta=\eta_k=\eta_1k^{-1}$, for some $\eta_1\in[\frac12,1]$ and all integers k large enough.

Remarks: (1) A related H with $a=\frac12$ arises in a celestial mechanics problem near a *double resonance*, responsible for the time scales differences (see (12.39) in [CG], after scaling away the factors ω_T to put H in dimensionless form and with several factors replaced here by constants, for simplicity). A first simplification of (2.1) compared to the "realistic" model in §12 of [CG] is the absence of an additive isochrony breaking term $\eta\frac{A^2}{2}$: taking it into account does not essentially change the analysis of the splitting results (even in its quantitative aspects on the asymptotics as $\eta\to0$, see §7). A second simplification is the absence in (2.1) of a further perturbation βf_0 which is not small (*i.e.* f_0,β are η,μ–independent) but depends only on the "fast" angle λ and on φ: $f_0=f_0(\varphi,\lambda)$. Taking it into account is a problem not discussed here. However it does not change the qualitative aspects but only the quantitative ones *as long as the system is isochronous*, see the discussion at the end of §8.
(2) g^2 in the above hamiltonian is *fixed* (*i.e.* it is η,μ independent): eventually we take $g\equiv1$ for simplicity. The parameters μ,η are free and we shall be interested in them having a "small value": note that if $\eta\to0$ the rotation vector $\underline{\omega}$ of the unperturbed torus ($I=\varphi=0$, $B=A=0$) has a size tending to ∞. The even parity assumptions on the "interactions" f simplifies, possibly in an essential way, the analysis.

In the precession problem of [CG] η is the deviation from spherical shape of a planet precessing around its baricenter, which moves on an ellipse of (fixed) eccentricity $\varepsilon = \mu\eta^c$.
(3) A general physical interpretation of (2.1) is that of a system consisting in (i) a forcing clock with angular velocity η^a (*i.e.* a point moving on a circle with angular velocity η^a, position α and action variable A); (ii) a pendulum (*i.e.* a point moving on a vertically placed circle with angular momentum I and position φ counted by taking $\varphi = 0$ as the unstable equilibrium position of the pendulum); (iii) a second forcing clock with angular velocity $\eta^{-1/2}$ and action variable B. Equivalently one can delete the $A\eta^a, B\eta^{-1/2}$ terms and replace α, λ by $\underline{\omega}\, t = (\eta^a t, \eta^{-1/2} t)$ thus regarding the system as a time dependent one, consisting of a pendulum subject to a quasi periodic force with periods $2\pi\eta^{-a}, 2\pi\eta^{1/2}$. The "characteristic time" of the pendulum system is $T_0 = g^{-1}$.
(4) Finally the "coupling constant" in (2.1) is written as $\mu\eta^c$ and not just μ, because the convergence radius of the expansions around the unperturbed torus is expected to be of the order of some power of η: it would be nice to know the best value of c (thus replacing the constant c by its optimal value) but it seems not known, see [HMS], [ACKR] (it seems that $c > \frac{1}{2}$ might be the right condition, but below we make no attempt at getting even close to such a small value).

§3. Separatrices and non degeneracy.

Supposing $\mu = 0$ in (2.1) we look at the unstable quasi periodic motion with $A = 0, B = 0, I = 0, \varphi = 0$ where $\alpha = \alpha_0 + \eta^a t$ and $\lambda = \lambda_0 + \eta^{-1/2} t$. This is a family of motions whose initial data are parameterized by $\underline{\alpha} \stackrel{def}{=} (\alpha_0, \lambda_0)$ and therefore form an invariant torus: we shall call it a *hyperbolic torus*. The energy of such motions is $H = 0$. If we fix A, B at other values we find a continuum of invariant hypebolic tori, with rotation vector $\underline{\omega} \stackrel{def}{=} (\eta^a, \eta^{-1/2})$. The energy of such motions is $\eta^{-1/2} B + \eta^a A$.
The torus is unstable and its unstable manifold W^- is 3–dimensional and contains the set parameterized by $(\varphi, \alpha, \lambda)$ via the equations:

3.1
$$A = 0, \qquad I = I(\varphi), \qquad (\alpha, \lambda) \in [0, 2\pi]^2,\ \varphi \in (0, 2\pi) \tag{3.1}$$

where $I(\varphi) = I^0(\varphi) \stackrel{def}{=} g\sqrt{2(1 - \cos\varphi)}$ is the pendulum *separatrix*. But also the set with $I = -I^0(\varphi)$ is part of the torus unstable manifold because of the meaning of φ as an angle. The stable manifold W^+ has equations $I(\varphi) = \pm I^0(\varphi)$: this means that $W^+ \equiv W^-$: a well known degeneracy of the pendulum motion.
A nice way of seeing the separatrix motions is by representing them in the (canonical) *Jacobi's coordinates*: these are coordinates (p, q) in terms of which the neighborhood of the (two) lines $I = \pm I^0(\varphi)$ in (3.1) is represented, near $I = \varphi = 0$, as:

3.2
$$I = R_0(p, q), \qquad \varphi = S_0(p, q) \tag{3.2}$$

where the functions R_0, S_0 are defined near the origin, where they vanish, and the pendulum motion becomes, in such coordinates, simply $p(t) = p_0 e^{-gt}, q(t) = q_0 e^{gt}$. Therefore the (p, q) coordinates describe globally, as $t \to +\infty$, the evolution of initial data $(p_0, 0)$ and globally, as $t \to -\infty$, those with data $(0, q_0)$. In such coordinates the equation of the stable manifold is simply $p = 0$ while that of the unstable manifold is $q = 0$.
The functions S_0, R_0 have well known holomorphy properties: the latter imply that the singularities of $R_0(pe^{gt}, 0), R_0(0, qe^{-gt})$, at fixed p or q, occur at $t = \pm i\frac{\pi}{2} g^{-1}$, and the same holds for S_0. This can be seen from the explicit solution of the pendulum equation in terms of elliptic integrals: an elementary analysis is in Appendix A9 of [CG], or in Appendix A1 of [G5].
The first result that we use to set up a general picture but strictly speaking not even necessary as shown in [Ge1], is the (well known, see [Gr]) stability of such unperturbed hyperbolic torus and of its stable and unstable manifolds. This will be stated as:

Theorem 1: *Suppose c in (2.1) large enough, and suppose that η is such that the rotation vector $\underline{\omega} \stackrel{def}{=} (\eta^a, \eta^{-1/2})$ verifies the diophantine property (2.2) and $\mu_0 > 0$ is small enough. One can define functions $\Gamma, \Xi, \Lambda, \Theta$ divisible by μ and analytic in the variables $(\psi_1, \psi_2, p, q, \mu)$, varying respectively on the torus $[0, 2\pi]^2$, on a neighborhood of $p = q = 0$, and in $|\mu| \le \mu_0$ with μ_0 small enough, such that setting, as $\underline{\psi}$ varies on $T^2 = [0, 2\pi]^2$:*

3.3
$$\begin{aligned} A =& A_0 + \Xi(\underline{\psi}, p, q), \quad B = B_0 + \Gamma(\underline{\psi}, p, q), \quad \alpha = \psi_1, \quad \lambda = \psi_2 \\ I =& R_0(p, q) + \Lambda(\underline{\psi}, p, q), \quad \varphi = S_0(p, q) + \Theta(\underline{\psi}, p, q) \end{aligned} \tag{3.3}$$

one defines, for all A_0, B_0, an invariant set on which the motion described by (2.1) takes place following $t \to (\underline{\psi} + \underline{\omega}\, t, p e^{-\overline{g}t}, q e^{+\overline{g}t})$ (as long as the (p, q) stay in the domain of definition of $\Xi, \Gamma, \Lambda, \Theta$), with $\overline{g} = (1 + \gamma(\mu))g$ and γ analytic in μ, near $\mu = 0$ and divisible by μ. The functions Γ, Ξ evaluated at $p = q = 0$ have zero average with respect to $\underline{\psi}$; also the time average of H vanishes on the above motions and (therefore) $H = 0$ for all of them.

The radius of convergence μ_0 is uniform in $\eta < 1$, as long as $\underline{\omega}$ verifies the diophantine condition.

Remarks: (1) The theorem implies that if $p = q = 0$ then the parametric equations $\varphi = \Theta(\underline{\psi}, 0, 0)$, $\alpha = \psi_1$, $\lambda = \psi_2$, $I = \Lambda(\underline{\psi}, 0, 0)$, $A = \Xi(\underline{\psi}, 0, 0)$ and $B = \Gamma(\underline{\psi}, 0, 0)$ describe, as $\underline{\psi}$ varies on the 2–dimensional torus $[0, 2\pi]^2$, an invariant torus. The quasi periodic rotation $\underline{\psi} \to \underline{\psi} + \underline{\omega}\, t$ with $\underline{\omega} = (\eta^a, \eta^{-1/2})$ gives, for all $\underline{\psi}$, a solution to the equations of motion. This means that the invariant torus of dimension 2 that we considered in §2 (*i.e.* $A = B = I = 0$, $\varphi = 0$) *survives the onset of perturbation* and persists, *slightly deformed*, with the same rotation vector $\underline{\omega}$ and a slightly varied pair of Lyapunov exponents (*i.e.* $\pm\overline{g} = \pm(1 + \gamma)\, g$ rather than $\pm g$). The zero average property for Γ, Ξ means that the perturbed torus is in the average (over $\underline{\psi}$) located at the same position as the unperturbed one from which it "emanates": this is useful as it allows us to parameterize the invariant tori by their average position in action space.

(2) Setting $p = 0, q \neq 0$ one finds a surface of dimension 3 which is *a part* of the *unstable* manifold W^- of the torus, while setting $q = 0, p \neq 0$ one gets a part of the stable manifod W^+. Such manifolds are colorfully called local *whiskers*. They are in fact "local" parts of larger "global" manifolds (see item (4) below) and can be called, for this reason, *local stable and local unstable manifolds.*

(3) The motion of A, B is computed from that of the other coordinates by quadrature (*e.g.* $\dot{B} = -\varepsilon\, \partial_\lambda f(\varphi, \alpha, \lambda)$, while B itself does not occurr in the equations of the other coordinates, and similarly for A). The symmetry of f does not imply that Λ, Θ have zero average over $\underline{\psi}$, if $\varphi = \pi$, [CG], and also the Lyapunov exponent $\overline{g}$ in general changes by a quantity γ of order $O(\mu)$ with respect to the unperturbed value $g(pq)$. The motions energy also changes, in general, by an amount analytic in μ and divisible by μ (*i.e.* of $O(\mu)$).

(4) Once the *local* parts of the torus whiskers have been defined as above we can extend them to *global* objects by simply applying time evolution to their points.

(5) The case $\mu = 0$ (one of the most classical results in Mechanics) is a non trivial exercise: it is developed in Appendix A9 of [CG] or Appendix A1 of [G5].

(6) The stable and unstable manifolds do not coincide, in general, for $\mu \neq 0$.

(7) in the anisochronous cases of §7 a very similar result holds with the difference that the relation between $\underline{\alpha}$ and $\underline{\psi}$ is also non trivial and described by a function $\underline{\Delta}(\underline{\psi}, p, q) = (\Delta(\underline{\psi}, p, q), 0)$ with zero average over $\underline{\psi}$ and divisible by μ and with the same domain of definition of the other functions in (3.3). In this case all the functions in (3.3), and $\underline{\Delta}$ as well, depend also on A_0 which must be restricted so that $\underline{\omega} = (\eta^a + \eta A_0, \eta^{-1/2})$ verifies a diophantine property with suitable diophantine constants $C(\eta), \tau$: the size of

the radius of convergence and the value of c will depend on the selected $C(\eta)$ (like (2.2) or (7.3)).

(8) The above theorem is well known; it is explicitly proved (in a much more general case) in the above form in §5 of [CG], see p. 38. Its proof is a rather straightforward adaptation of Arnold's method of proof of the KAM theorem, [A1], as exposed for instance in [G1]: the only element of "novelty" is perhaps the *normal form* of the motion in the coordinates $\underline{\psi}\,, p, q$, as long as the latter two remain in their domain of definition. A more modern proof, based on Eliasson's method, [E], for proving KAM (as exposed in [G2],[GG]) and extending it to the problem of tori of one dimension less than the maximal, can be derived from [Ge1] (where only the cases $p = 0$ or $q = 0$ are studied). Also unusual is the absence of the twist condition (present, and necessary, in the more general proof in [CG]): it can be eliminated because of the special structure of the hamiltonian (2.1). But it is a long and uninteresting proof for us here. We shall not need, however, the normal form and, as it will be clear, we only need the classical results in the weaker form discussed in [Gr],[Ge1].

§4. Splitting angles. A recursive determination.

The symmetry of f implies an intersection at $\varphi = \pi$ and $\underline{\alpha} = \underline{0}$ (see below or, for instance, p. 363 of [G3]) between stable and unstable manifolds of the torus into which the unperturbed torus (*i.e.* $A = B = 0, I = 0, f = 0$) is deformed by the perturbation. Therefore we set up an algorithm to study such intersection. For this purpose it is convenient to work in the original canonical coordinates and write the stable and unstable whiskers $W_\mu^\pm$ as:

4.1
$$W_\mu^\pm = \{(\varphi, \underline{\alpha}\,, I, \underline{A}\,) = \big(\varphi, \underline{\alpha}\,, I_\mu^\pm(\varphi, \underline{\alpha}\,), \underline{A}\,_\mu^\pm(\varphi, \underline{\alpha}\,)\big)\} \tag{4.1}$$

with $\underline{\alpha} \in T^2$, $\varepsilon < |\varphi| < 2\pi - \varepsilon$ where $\varepsilon > 0$ can be fixed *a priori* as small as we please provided we diminish the value of the analyticity radius μ_0 in theorem 1.

In other words the whiskers deformation is of $O(\mu)$ in every closed subinterval of $(-2\pi, 2\pi)$: therefore they remain parametrizable by $\varphi, \underline{\alpha}$ for φ in any closed subinterval if μ is small enough (just note that for $\mu = 0$ they *are* parametrizable). We say that, in such region of $(\varphi, \underline{\alpha}\,)$, they are *graphs* over the angle variables. We only need that $\varphi = \pi$ is allowed (hence $\varepsilon = \frac{\pi}{2}$ will do).

We define the *splitting vector* $\underline{Q}\,(\underline{\alpha}\,)$, the *splitting* or *intersection matrix* and the *splitting* between W_μ^+ and W_μ^- at $\varphi = \pi$ and $\underline{\alpha} = \underline{0}\,$, respectively, as:

4.2
$$\underline{Q}\,(\underline{\alpha}\,) \overset{def}{=} \underline{A}\,_\mu^+(\pi, \underline{\alpha}\,) - \underline{A}\,_\mu^-(\pi, \underline{\alpha}\,),\; D \overset{def}{=} \partial_{\underline{\alpha}}\, \underline{Q}\,(\underline{\alpha}\,)\big|_{\underline{\alpha}=\underline{0}}\,, \text{ and } \det D \tag{4.2}$$

A relevant remark (Lochak and Sauzin, private communication) is that the whiskers are *lagrangian manifolds* so that, for a suitable generating function $S^\pm(\varphi, \underline{\alpha}\,)$, the whiskers have equation: $\underline{A}\,^\pm = \partial_{\underline{\alpha}}\, S(\varphi, \underline{\alpha}\,), I^\pm = \partial_\varphi S^\pm(\varphi, \underline{\alpha}\,)$ around every point $(\varphi_0, \underline{\alpha}\,_0)$ where they can be locally regarded as graphs over the angles $(\varphi, \underline{\alpha}\,)$, *e.g.* at $\varphi = \pi$ and for μ small. This implies that D is *symmetric*, as we indeed find in (6.10), (7.10).

We now derive recursive formulae for $I_\mu^\pm$, $\underline{A}\,_\mu^\pm$ in (4.1) and their time evolution, keeping in mind that for $\mu = 0$ it is $I_0^+ = I_0^- = I^0(\varphi) = g\sqrt{2(1-\cos\varphi)}$, $\underline{A}\,_0^\pm = \underline{0}\,$. The unperturbed motion is simply: $X^0(t) \equiv (\varphi^0(t), \underline{\alpha} + \underline{\omega}\,t, I^0(\varphi^0(t), \underline{0}\,)$, where $(\varphi^0(t), I^0(\varphi^0(t)))$ is the free (*i.e.* with $\mu = 0$) separatrix motion, generated by the pendulum in (2.1) setting the origin of time when the pendulum swings through $\varphi = \pi$. The following §4,§5 are presented here only for completeness: although selfcontained *they are not* meant as a substitute of the work done in [G3] but serve the purpose of guiding the reader to dig out of that paper what he may want to see in more detail.

Let $X_\mu^\sigma(t; \underline{\alpha}\,)$, $\sigma = \pm$, be the evolution of the point on W_μ^σ whose initial coordinates are given by $(\pi, \underline{\alpha}\,, I_\mu^\sigma(\pi, \underline{\alpha}\,), \underline{A}\,_\mu^\sigma(\pi, \underline{\alpha}\,))$; from now on we shall fix initial data with $\varphi = \pi$ (which amounts to studying the whiskers at the "Poincaré's section" $\{\varphi = \pi\}$).

The analyticity in μ implied by the above theorem 1 and the analyticity properties of the Jacobi functions R_0, S_0 allow us to consider the *convergent* (if c in (2.1) theorem 1 is large) Taylor series expansions, in μ, of the whiskers equations. Let:

4.3 $$X^\sigma_\mu(t) \equiv X^\sigma_\mu(t;\underline{\alpha}) \stackrel{def}{=} \sum_{k\ge0} X^{k\sigma}(t;\underline{\alpha})\varepsilon^k \equiv \sum_{k\ge0} X^{k\sigma}(t)\varepsilon^k, \qquad \sigma=\pm \tag{4.3}$$

be the power series in $\varepsilon=\mu\eta^c$ of X^σ_μ, (convergent for μ small); note that $X^{0\sigma}\equiv X^0$ is the degenerate unperturbed whisker. We shall often omit writing explicitly the $\underline{\alpha}$ variable among the arguments of various $\underline{\alpha}$ dependent functions, to simplify the notations.

Theorem 1, §3, tells us that the t–dependence of $X^\sigma_\mu(t)$ has the form:

4.4 $$X^\sigma_\mu(t) = X^\sigma_\mu(\underline{\omega}\, t, t;\underline{\alpha}) \tag{4.4}$$

where $X^\sigma_\mu(\underline{\psi}, t;\underline{\alpha})$ is a real analytic function, of *all* its arguments (μ included), which is periodic in $\underline{\psi}$ and $\underline{\alpha}$. And in fact in our *isochronous* case the dependence on $\underline{\psi}$ and $\underline{\alpha}$ is via $\underline{\alpha}+\underline{\psi}$ ("no phase shift" in the sense of [CG]). The two functions $X^{k\sigma}(t)$ will be regarded as forming a single function $X^k(t)$:

4.5 $$X^k(t) = \begin{cases} X^{k+}(t) & \text{if } \sigma = \operatorname{sign} t = + \\ X^{k-}(t) & \text{if } \sigma = \operatorname{sign} t = - \end{cases} \tag{4.5}$$

We label the 6 *components of* X *with an index* j, $j=0,1,\ldots,5$, *and write them (*notation used in [G3]*), with the convention:*

4.6 $$X_0 \stackrel{def}{=} X_-, \quad (X_j)_{j=1,2} \stackrel{def}{=} \underline{X}_\downarrow, \quad X_3 \stackrel{def}{=} X_+, \quad (X_j)_{j=4,5} \stackrel{def}{=} \underline{X}_\uparrow \tag{4.6}$$

and we write first the angle variables $((\varphi,\alpha,\lambda) = (X_-, X_\downarrow))$*, then the action variables* $((I,A,B) = (X_+, \underline{X}_\uparrow))$*; first the pendulum, then the rotator and then the clock variables.*

Therefore at order 0 in μ:

4.7 $$\begin{aligned} X_0(t) =&\varphi^0(t), \quad X_1(t) = \alpha_0 + \eta^a t, \quad X_2(t) = \lambda_0 + \eta^{-1/2} t \\ X_3(t) =&I^0(\varphi^0(t)), \quad X_4 = 0 = X_5, \end{aligned} \tag{4.7}$$

where $\varphi^0(t)$ is the free separatrix motion (*i.e.* $\varphi^0(t) = 4\,\mathrm{arctg}\, e^{-gt}$, see Appendix A1).

Inserting (4.3) in Hamilton's equations for (2.1) and comparing the various orders in μ, the coefficients $X^{k\sigma}(t) \equiv X^{k\sigma}(\underline{\omega}\, t, t;\underline{\alpha})$ are seen to satisfy the hierarchy of equations:

4.8 $$\frac{d}{dt} X^{k\sigma} \equiv \dot{X}^{k\sigma} = L X^{k\sigma} + F^{k\sigma} \tag{4.8}$$

where L is a 6×6 matrix with only two non vanishing elements $L_{03}=1$ and $L_{30} = g^2\cos\varphi^0(t)$. Expanding X^σ in powers of ε and imposing that the equations of motion are verified, we find recursively the expressions for F. For instance:

4.9 $$\begin{aligned} F^1_+ =& -\partial_\varphi f, \qquad \underline{F}^1_\uparrow = -\partial_{\underline{\alpha}} f \\ F^2_+ =& -\frac{g^2}{2}\sin\varphi\,(X^1_-)^2 - \partial_{\varphi^2} f X^1_- - \partial_{\varphi\underline{\alpha}} f\, \underline{X}^1_\downarrow \\ \underline{F}^2_\uparrow =& -\partial_{\underline{\alpha}\varphi} f X^1_- - \partial_{\underline{\alpha}^2} f\, \underline{X}^1_\downarrow \\ F^3_+ =& -\frac{g^2}{2} 2X^1_- X^2_- \sin\varphi - \frac{g^2}{3!}\cos\varphi (X^1_-)^3 - \partial_{\varphi^2} f\, X^2_- + \\ & -\partial_{\underline{\alpha}\varphi} f\, \underline{X}^2_\downarrow - \frac12 \partial_{\varphi\underline{\alpha}^2} f\, \underline{X}^1_\downarrow \underline{X}^1_\downarrow - \partial_{\varphi^2\underline{\alpha}} f\, \underline{X}^1_\downarrow X^1_- - \frac12 \partial_{\varphi^3} f\,(X^1_-)^2 \\ \underline{F}^3_\uparrow =& -\partial_{\underline{\alpha}\varphi} f X^2_- - \partial_{\underline{\alpha}^2} f\, \underline{X}^2_\downarrow - \frac12 \partial_{\underline{\alpha}^3} f\, \underline{X}^1_\downarrow \underline{X}^1_\downarrow - \partial_{\varphi\underline{\alpha}^2} f\, \underline{X}^1_\downarrow X^1_- - \frac12 \partial_{\underline{\alpha}\varphi^2} f X^1_- X^1_- \end{aligned} \tag{4.9}$$

where the functions are evaluated at $\varphi(t)=\varphi^0(t)$, $\underline{\alpha}\,(t)\stackrel{def}{=}\underline{\alpha}+\underline{\omega}\,t$. More generally, F^k depend upon $X^0,...,X^{k-1}$ *but not on* X^k. It is $F^k_-=\underline{F}^{\,k}_{\downarrow}\equiv 0$ for all $k\ge 1$.

Expressing the solution of a linear inhomogeneous equation like the one in (4.8) can be conveniently done in terms of the *wronskian matrix.* We recall therefore the notion of *wronskian matrix* $W(t)$ of a solution $t\to x(t)$ of a differential equation $\dot x=f(x)$ in R^n.

It is a $n\times n$ matrix whose columns are formed by n linearly independent solutions of the linear differential equation obtained by linearizing f around the solution x and assuming $W(0)=$ identity. In other words $W(t)$ is, in our case, the solution of the differential equation $\dot W(t)=L(t)\,W(t)$, $W(0)=1$.

The solubility by elementary quadrature of the free pendulum equations leads, on the separatrix, to the following formulae that have importance because the wronskian of the free separatrix motion can be expressed in terms of them. If $\varphi^0(t)=4\arctan e^{-gt}$:

4.10

$$\begin{aligned}\cos\varphi^0=&1-\frac{2}{(\cosh gt)^2}=1-8\frac{x^2}{(1+x^2)^2}\\ \sin\varphi^0=&2\frac{\sinh gt}{(\cosh gt)^2}=4\sigma x\frac{1-x^2}{(1+x^2)^2}\end{aligned}\tag{4.10}$$

for $x=e^{gt}$ *or, as well, for* $x=e^{-gt}$. Eq. (4.10) leads, see (A1.1) in Appendix A1, to the following expression for the wronskian $W(t)$ of the *separatrix motion* for the pendulum appearing in (2.1), with initial data at $t=0$ given by $\varphi=\pi, I=2g$:

4.11

$$W(t)=\begin{pmatrix}\frac{1}{\cosh gt} & \frac{\overline{w}}{4}\\ -\frac{\sinh gt}{\cosh^2 gt} & (1-\frac{\overline{w}}{4}\frac{\sinh gt}{\cosh^2 gt})\cosh gt\end{pmatrix},\qquad \overline{w}\equiv\frac{2gt+\sinh 2gt}{\cosh gt}\tag{4.11}$$

Notationally we follow here [G3] (in [CG] I,φ are exchanged). The evolution of the $X_\pm$ (*i.e.* I,φ) components can be determined by using the above wronskian:

4.12

$$\begin{pmatrix}X^{k\sigma}_-\\ X^{k\sigma}_+\end{pmatrix}=W(t)\begin{pmatrix}0\\ X^{k\sigma}_+(0)\end{pmatrix}+W(t)\int_0^t W^{-1}(\tau)\begin{pmatrix}0\\ F^{k\sigma}_+(\tau)\end{pmatrix}\,d\tau\tag{4.12}$$

Thus, denoting by w_{ij} $(i,j=0,3)$ the entries of W, we see immediately that:

4.13

$$\begin{aligned}X^{k\sigma}_-(t)=&w_{03}(t)X^{k\sigma}_+(0)+w_{03}(t)\textstyle\int_0^t w_{00}(\tau)F^{k\sigma}_+(\tau)d\tau-w_{00}(t)\int_0^t w_{03}(\tau)F^{k\sigma}_+(\tau)\,d\tau\\ X^{k\sigma}_+(t)=&w_{33}(t)X^{k\sigma}_+(0)+w_{33}(t)\textstyle\int_0^t w_{00}(\tau)F^{k\sigma}_+(\tau)d\tau-w_{30}(t)\int_0^t w_{03}(\tau)F^{k\sigma}_+(\tau)\,d\tau\end{aligned}\tag{4.13}$$

Integrating on the separatrix (4.8) for the $\uparrow,\downarrow$ components is “easier”: one can find it directly or, more systematically, by writing the full 6×6–wronskian matrix of the equation (4.8), which is trivially related to $W(t)$ above; we shall not write it here: see eq. (4.21) in [G3] for an explicit expression. The result is:

4.14

$$\underline{X}^{\,k\sigma}_{\downarrow}(t)=\underline{0}\,,\qquad \underline{X}^{\,k\sigma}_{\uparrow}(t)=\underline{X}^{\,k\sigma}_{\uparrow}(0)+\int_0^t\underline{F}^{\,k\sigma}_{\uparrow}(\tau)d\tau\tag{4.14}$$

having used that the $\underline{X}^{\,k\sigma}_{\downarrow}(0)\equiv\underline{0}$, because the initial datum is fixed and μ–independent.

If (2.1) is modified by adding an isochrony breaking term $\frac{\eta}{2}A^2$ the first component of the first of (4.14) becomes:

4.15

$$X^{k\sigma}_1(t)=\eta\,\big(tX^{k\sigma}_1(0)+\int_0^t(t-\tau)F^{k\sigma}_1(\tau)\,d\tau\big)\tag{4.15}$$

where $X^{k\sigma}_1$ is the first compnent of $\underline{X}^{\,k\sigma}_{\uparrow}$, see eq. (2.18) in [G3].

Equations (4.13), (4.14) can be used to find a reasonably simple algorithm to represent whiskers to all orders $k \geq 1$ of the perturbation expansion. It is very important to keep in mind that the initial data in (4.13), (4.14) *are not constants*: according to the convention following (4.3) they can depend on the $\underline{\alpha}$ variables of the initial data. This means that the functions X depend separately on $\underline{\alpha}$ and $\underline{\omega}\,t$. *Except* when, as in (2.1), the hamiltonian is linear in the A, B variables. In the latter case the dependence on $\underline{\alpha}$ and $\underline{\omega}\,t$ of the r.h.s. of (4.4) (where the notation is complete and all variables are indicated explicitly) *must be* through $\underline{\alpha} + \underline{\omega}\,t$, since the $\underline{\alpha}$ angles vary as $\dot{\underline{\alpha}} = \underline{\omega}$.

Note that the case (2.1) is *non trivial and, in fact, very interesting: it is equivalent to a problem on a non linear quasi periodic Schrödinger equation*, see [G2], [BGGM]. The extension to anisochronus cases (*i.e.* with a quadratic term in A added to (2.1)) is worked out in [CG] up to third order: a general analysis can be found in [G3].

The initial data (still unknown) in (4.13),(4.14) are determined by imposing the correct behavior at $\pm\infty$, and the correct dependence on $\underline{\alpha}$ and $\underline{\omega}\,t$ (*i.e.* a dependence on these two vectors through their sum). These conditions arise from the fact that the motion must be asymptotic to the quasi periodic motion on the invariant torus whose whiskers are described by X^+ or X^-. The scheme to do so is the following, see [G3].

Note that w_{03}, w_{33} in (4.11) behave, as $t \to \sigma\infty$ with $\sigma = \pm 1$, asymptotically as $\sigma e^{gt\sigma}/4$ and $-e^{gt\sigma}/4$, while the other two matrix elements become exponentially small. Hence we see that the tems in (4.13) proportional to $w_{33}(t)$ or $w_{30}(t)$ diverge, in general, as $t \to \sigma\infty$ exponentially fast (supposing the integrals convergent, *i.e.* supposing $F_+^{k\sigma}(\tau)$ not growing faster than a polynomial in t, as it will turn out to be). But they are multiplied by: $\big(X_+^{k\sigma}(0) + \int_0^t w_{00}(\tau) F_+^{k\sigma}(\tau)\, d\tau\big)$ so that the condition to determine $X_+^{k\sigma}(0)$ is:

4.16

$$X_+^{k\sigma}(0) + \int_0^{\sigma\infty} w_{00}(\tau) F_+^{k\sigma}(\tau)\, d\tau = 0 \tag{4.16}$$

Likewise, in the isochronous case, from the second of (4.14) we determine $\underline{X}_\uparrow^{k\sigma}(0)$ by imposing that $\underline{X}_\uparrow^{k\sigma}(t)$, *i.e.* the momentum component corresponding to the isochronous angles α, λ, depends asymptotically on t only via $\underline{\alpha} + \underline{\omega}\,t$: this will determine $\underline{X}_\uparrow^{k\sigma}(0)$ up to a constant.

And the average over $\underline{\alpha}$ of $\underline{F}_\uparrow^{k\sigma}(t)$ must tend to 0 as $t \to \sigma\infty$ (otherwise the second of (4.14) could not possibly be bounded as $t \to \sigma\infty$: but it has to be such because $\underline{X}_\uparrow^{k\sigma}(t)$ has to be bounded, by theorem 1). This means that the constant is not a function of $\underline{\alpha}$ and can be fixed arbitrarily: however we want that the averages of the Ξ, Γ in (3.3) are 0 so that $X^{k\sigma}(0)$ is completely determined. One finds, for *both components*:

4.17

$$\underline{X}_\uparrow^{k\sigma}(0) + \int_0^{\sigma\infty} \underline{F}_\uparrow^{k\sigma}(\tau) d\tau = \underline{0} \tag{4.17}$$

where the integral is usually improper: see (5.1) below for a proper definition (derived by simply looking at the meaning of the conditions imposed to determine $\underline{X}_\uparrow^{k\sigma}(0)$). Eq. (4.17) *remains the same* even in the anisochronous cases (see eq. (4.5) in [G3]).

The key to concrete calculations is that, f in (2.1) being a trigonometric polynomial, the function F^1 (see (4.9)) belongs to the class $\mathcal{M}$ of linear combinations of terms like:

4.18

$$M(t) = \frac{(g\sigma t)^h}{h!} x^k \sigma^\vartheta e^{i\rho\,\underline{\omega}\cdot\underline{\nu}\,t} \tag{4.18}$$

with $h \geq 0, k$ integers, $\vartheta = 0, 1$, $\rho = \pm 1$ and $gk \pm i\,\underline{\omega} \cdot \underline{\nu} \neq 0$ and we set $x = e^{-gt\sigma}$ with $\sigma = \operatorname{sign} t = \pm 1$. In fact F^1 is a *finite linear combination* of harmonics $\underline{\nu}$. By (4.10) we see that (4.11) is an analytic function of the variable $x = e^{gt}$ and of t, or of $x = e^{-gt}$ and of t: the explicit t–dependence arises because of the term $\overline{w}$ in (4.11).

Expressing F^1 as a sum of monomials like (4.18) requires a (convergent) infinite sum over k, h. In the case of F^1, in fact, there are *no monomials* with a power of t higher than 0 (*i.e.* with $h > 0$ in (4.18)). By induction *all* the F^k have the property of being expressible as sums of monomials like (4.18). However, starting with the second order, one sees that powers of t do appear (this reflects, see theorem 1, that the function γ, describing the change of the Lyapunov exponents to $\pm(1+\gamma)g$ with γ analytic in μ, is not identically zero; see, however, [Ge2]). A full description of the induction can be found in [G3].

For the analyticity properties of the series introduced above we refer to §2 of [G3] and we proceed to a quick discussion of the determination of the the initial constants.

§5. Improper integrals and the operators $\mathcal{I}, \mathcal{O}, \mathcal{O}_0$.

The integrations in (4.13),(4.14) can be expressed in terms of an operator $\mathcal{I}$ acting linearly on finite linear combinations of monomials like (4.18) with $k^2 + (\underline{\omega} \cdot \underline{\nu})^2 > 0$:

5.1
$$
\begin{aligned}
&\mathcal{I}M(t) \stackrel{def}{=} \int_{\sigma\infty}^{t} M(\tau)d\tau, \qquad \text{with :}\\
&\mathcal{I}M(t) = -\sigma^{\vartheta+1} x^k e^{i\rho\,\underline{\omega}\cdot\underline{\nu}\,t} \sum_{p=0}^{h} \frac{(\sigma t)^{h-p}}{(h-p)!} \frac{1}{(k - i\rho\sigma\,\underline{\omega}\cdot\underline{\nu}\,)^{p+1}}
\end{aligned}
\tag{5.1}
$$

where the first row is a formal definition whose mathematical meaning is given by the second row (note that if $k \le 0$ the first line is an improper integral), and we set $g = 1$.

Note that the $\mathcal{I}$ is *not defined* on the polynomials of t, σ, *i.e.* if $k = 0$ and $\underline{\omega} \cdot \underline{\nu} = 0$ (so that no exponentials are present in the monomial defining M). It can be naturally extended, for $j \ge 0$, to the polynomials by setting $\mathcal{I}t^j\sigma^\vartheta = \frac{\sigma^\vartheta}{j+1} t^{j+1}$, see (3.7) in [G3].

The $\mathcal{I}$ is an integration with respect to t with special initial data: in fact *at fixed* σ:

5.2
$$
\partial_t \mathcal{I}M \equiv M \tag{5.2}
$$

If M is such that $M(t) \equiv M(\underline{\omega}\,t, \sigma)$ for some $M(\underline{\psi}\,, \sigma)$ defined on the torus, then:

5.3
$$
\mathcal{I}M(t) = (\underline{\omega} \cdot \underline{\partial}_{\,\underline{\psi}})^{-1} M(\underline{\omega}\,t, \sigma) \tag{5.3}
$$

The integrals in (4.13),(4.14) can be expressed in terms of the operators:

5.4
$$
\begin{aligned}
\mathcal{O}F(t) =& w_{03}(t)\big(\mathcal{I}(\,w_{00}F)\big)(t) - w_{00}(t)\big(\mathcal{I}(w_{03}F)\big)(\tau)|_{0^\sigma}^{t}\\
\mathcal{O}_+F(t) =& w_{33}(t)\big(\mathcal{I}(\,w_{00}F)\big)(t) - w_{30}(t)\big(\mathcal{I}w_{03})F\big)(\tau)|_{0^\sigma}^{t}\\
\overline{\mathcal{I}}^2 F(t) =& \mathcal{I}^2F(t) - \mathcal{I}^2F(0^\sigma)
\end{aligned}
\tag{5.4}
$$

where $\sigma = \operatorname{sign} t$. Then one finds, in the general anisochronous case:

5.5
$$
X_-^h(t) = \mathcal{O}F_+^h(t), \quad X_+^h(t) = \mathcal{O}_+F_+^h, \quad \underline{X}_\downarrow^{\,h}(t) = \eta E\overline{\mathcal{I}}^2 F_\uparrow^h(t), \quad \underline{X}_\uparrow^{\,h}(t) = \mathcal{I}\,\underline{F}_\uparrow^{\,h}(t) \tag{5.5}
$$

where E is the projection over the first component, see (4.15), and F^h have to be expressed in terms of the $X^{h'}$ with $h' < h$.

Since, in this section, we consider an isochronous model the above formulae are slightly simpler as $\underline{X}_\downarrow^{\,h}(t) \equiv \underline{0}$, *i.e.* must take $E = 0$ in the third of (5.5). In §7 we shall need, however, (5.5).

There is no difficulty in setting up a general recursive scheme for the computation. We just give the result using the convenient notations ($k_j^i \ge 1, m_i > 0$):

5.6
$$
(k_j^i)_{\underline{m}\,,p} \equiv (k_1^0, \ldots, k_{m_0}^0, k_1^1, \ldots, k_{m_1}^1, \ldots, k_1^3, \ldots, k_{m_3}^3) \qquad \text{s.t.} \sum k_j^i = p \tag{5.6}
$$

referring to §2 of [G3] for more details (if needed). If $f_1 \stackrel{def}{=} f, f_0 \stackrel{def}{=} g^2\cos\varphi$, we find:

5.7

$$\begin{aligned}
&F^{k\sigma}_- \equiv 0\ ,\ \ \underline{F}^{k\sigma}_{\downarrow} \equiv \underline{0}\ ,\ \ \underline{F}^{k\sigma}_{\uparrow} = -\sum_{|\underline{m}|\ge 0,\delta=0,1} \frac{1}{\underline{m}!}(\partial^{\underline{m}}\partial_{\underline{\alpha}} f_\delta) \sum_{(k^i_j)_{\underline{m},k-1}} \prod_{i=0}^{l-1}\prod_{j=1}^{m_i} X_i^{k^i_j\sigma}\\
&F^{k\sigma}_+ \equiv -\sum^{*}_{|\underline{m}|\ge 2,\delta=0,1} \frac{1}{\underline{m}!}(\partial^{\underline{m}}\partial_{\varphi} f_\delta) \sum_{|\underline{m}|\ge 0} (\partial_0 f_\delta)_{\underline{m}} \sum_{(k^i_j)_{\underline{m},k-1}} \prod_{i=0}^{l-1}\prod_{j=1}^{m_i} X_i^{k^i_j\sigma}
\end{aligned}\tag{5.7}$$

where $(k^i_j)_{\underline{m},k}, (k^i_j)_{\underline{m},k-1}$ are defined in (5.6). The $*$ means that if $\delta = 0$ only vectors $\underline{m}$ with $|\underline{m}| \ge 2$ have to be considered in the sums. Note that if $\delta = 0$ the sum in the expression for F^h_+ can only involve vectors $\underline{m}$ with $m_j = 0$ if $j \ge 1$, because the function $f_0 = g^2\cos\varphi$ depends only on φ and not on $\underline{\alpha}$, (hence also $k^i_j = 0$ if $i > 0$). The functions are evaluated at $\varphi(t) = \varphi^0(t)$, $\underline{\alpha}(t) \stackrel{def}{=} \underline{\alpha} + \underline{\omega}\,t$. The indices in (5.7) are mutually contracted with a natural rule that we leave to the reader to work out.
The relations $\underline{F}^h_{\downarrow} = \underline{0}$ and $F^h_- \equiv 0$ are general (as the equations for $\underline{\dot X}_{\downarrow}, \dot X_-$ are linear) and in the *isochronous cases*:

5.8

$$\begin{aligned}
&\underline{X}^1_{\uparrow}(t) = -\mathcal{I}(\partial_{\underline{\alpha}} f),\quad \underline{X}^2_{\uparrow}(t) = -\mathcal{I}\Big(\partial_{\underline{\alpha}\varphi} f\,\mathcal{O}(-\partial_\varphi f)\Big)\\
&\underline{X}^3_{\uparrow}(t) = -\mathcal{I}\Big(\partial_{\underline{\alpha}\varphi} f\mathcal{O}\big(-\frac{1}{2}\sin\varphi\mathcal{O}(\partial_\varphi f)\mathcal{O}(\partial_\varphi f)\big)\Big) +\\
&\qquad -\mathcal{I}\Big(\partial_{\underline{\alpha}\varphi} f\mathcal{O}\big(\partial_{\varphi^2} f\mathcal{O}(\partial_\varphi f)\big)\Big) - \frac{1}{2}\mathcal{I}\Big(\partial_{\underline{\alpha}\varphi^2} f\mathcal{O}(\partial_\varphi f)\mathcal{O}(\partial_\varphi f)\Big)
\end{aligned}\tag{5.8}$$

(see the examples in (4.9)). We fix our attention on the models with f a trigonometric polynomial ("trigonometric perturbation") of degree N:

5.9

$$\begin{aligned}
&f(\underline{\alpha},\varphi) = f_S(\varphi,\alpha,\lambda) + f_F(\varphi,\alpha,\lambda)\\
&f_j(\varphi,\alpha,\lambda) = \sum_{n,\underline{\nu}}^{N} f_{j,(n,\underline{\nu})}\,\cos(\nu_1\alpha + \nu_2\lambda + n\varphi),\quad j = S, F
\end{aligned}\tag{5.9}$$

with $|n|, |\underline{\nu}| \le N$ and $f_{S,(n,\underline{\nu})} = 0$ unless $\nu_2 = 0$, *i.e.* $\underline{\nu}$ is a *slow* mode, while $f_{F,(n,\underline{\nu})} = 0$ unless $\underline{\nu}_2 \ne 0$, *i.e.* $\underline{\nu}$ is a *fast* mode. We also say that f_S depends only on slowly rotating angles and f_F depends on fastly rotating angles. A nontrivial example can be:

5.10

$$f(\alpha,\lambda) = \big(\cos(\alpha+\varphi) + \cos(\lambda+\varphi)\big)\tag{5.10}$$

The intersection matrix to order h, D^h_{ij}, can be expressed, to order $h = 1, 2, 3$ as:

5.11

$$\begin{aligned}
&D^1_{ij} = \int_{-\infty}^{\infty} dt\,\partial_{ij} f,\quad D^2_{ij} = -\int_{-\infty}^{\infty} dt\,\big[\partial_{ij0} f\,\mathcal{O}(\partial_0 f) + \partial_{j0} f\mathcal{O}(\partial_{i0} f)\big]\\
&D^3_{22} = \int_{-\infty}^{\infty} dt\,\Big[w_{30}\,\mathcal{O}(\partial_{220} f)\,\mathcal{O}(\partial_0 f)^2 + 2w_{30}\,\mathcal{O}(\partial_{20} f)^2\,\mathcal{O}(\partial_0 f) +\\
&\quad + \partial_{200} f\,\mathcal{O}(\partial_{20} f)\,\mathcal{O}(\partial_0 f) + \partial_{00} f\,\mathcal{O}(\partial_{220} f)\,\mathcal{O}(\partial_0 f) +\\
&\quad + \partial_{00} f\,\mathcal{O}(\partial_{20} f)\,\mathcal{O}(\partial_{20} f) + \frac{1}{2}\partial_{2200} f\,\mathcal{O}(\partial_0 f)^2 + \partial_{200} f\mathcal{O}(\partial_{20} f)\,\mathcal{O}(\partial_0 f)\Big]
\end{aligned}\tag{5.11}$$

where the derivatives of the f's are evaluated at $\varphi(t)$, $\underline{\alpha} + \underline{\omega}\,t$, with $\varphi(t) \equiv \varphi^0(t)$ the free separatrix motion, see (4.10). We set $\partial_0 \stackrel{def}{=} \partial_\varphi$ and $\partial_i \stackrel{def}{=} \partial_{\alpha_i}$; the $\underline{\alpha}$'s have to be set equal to 0 after evaluating derivatives. The expressions contain improperly convergent integrals (in general) and must be understood by thinking $\int_{-\infty}^{+\infty}$ as $\int_{-\infty}^{0}+\int_{0}^{+\infty}$ and by using the definition (5.1), see [G3].

It is convenient to split the operation $\mathcal{O}$, see (6.5) in [G3], as:

5.12

$$\begin{aligned}\mathcal{O}(F) =&\mathcal{O}_0(F)+|w_{03}(t)|\,G^{(0)}(F)+\,w_{00}(t)\,G^{(1)}(F)\\ \mathcal{O}_0(F) =&\frac{1}{2}\Big(\int_{-\infty}^{t}d\tau(w_{03}(t)\,w_{00}(\tau)-w_{00}(t)\,w_{03}(\tau))\,F(\tau)+\int_{+\infty}^{t}d\tau\,(same)\Big)\\ G^{(0)}(F) =&\frac{1}{2}\int_{+\infty}^{-\infty}d\tau\,w_{00}(\tau)\,F(\tau),\qquad G^{(1)}(F)=\frac{1}{2}\int_{+\infty}^{-\infty}d\tau\,|w_{03}(\tau)|\,F(\tau)\end{aligned}\tag{5.12}$$

The identity, see (6.12) and Appendix A2 of [G3]:

5.13

$$\int_{-\infty}^{+\infty}dt\,F(t)\mathcal{O}(H)(t)=\int_{-\infty}^{+\infty}dt\,H(t)\mathcal{O}(F)(t)\tag{5.13}$$

implies symmetry of the above matrices D_{ij}, at least for the first three orders (see (5.11)): symmetry follows to all orders as said after (4.2), or as it will be seen in §6.

In the anisochronous case, *i.e.* in §7, we shall also use the splitting:

5.14

$$\begin{aligned}&\overline{\mathcal{I}}^2(F)=\mathcal{I}_0^2(F)+|t|G^{(2)}(F)+G^{(3)}(F)\\ &\mathcal{I}_0^2(F)=\frac{1}{2}\Big(\int_{-\infty}^{t}d\tau\,(t-\tau)\,F(\tau)+\int_{+\infty}^{t}d\tau\,(same)\Big)\\ &G^{(2)}(F)=\frac{1}{2}\int_{+\infty}^{-\infty}d\tau\,F(\tau),\qquad G^{(3)}(F)=\frac{1}{2}\int_{+\infty}^{-\infty}d\tau\,|\tau|\,F(\tau)\end{aligned}\tag{5.14}$$

see eq. (6.3),(6.6) in [G3].

The key remark, to understand the asymptotic behaviour of (5.11) as $\eta\to 0$, is that whenever the integrand is analytic it becomes possible to shift the integrations, over t and the τ's, to an axis close to $\mathrm{Im}\,t$, $\mathrm{Im}\,\tau=\pm(\frac{\pi}{2}-\eta^{1/2})$: see §8 in [G3] (choosing the free parameter δ appearing in [G3] as $\eta^{1/2}$).

Using that $G^{(1)}(F)=G^{(0)}(F)=0$ if F is odd (as the odd derivatives of f are, when evaluated at $\underline{\alpha}=\underline{0}$) and using also that $\mathcal{O}_0$ leaves parity unchanged, in general, we shall find that the *non analytic* terms (*i.e.* those containing integrals of a non analytic function, *e.g.* $|w_{03}(\tau)|$) *cancel each other* in their contribution to the determinant of D_{ij} *to all orders* $k\ge 1$. The result will be the proof of the following theorem:

Theorem 2 *("Large angles theorem"): Consider a system described by the hamiltonian (2.1) or (7.1) with f an even trigonometric polynomial. Let μ be small ($|\mu|<\mu_0$) and c large, enough. Consider an invariant torus with diophantine rotation $\underline{\omega}$ among those described in theorem 1 above. At the homoclinic point with $\varphi=\pi$, $\underline{\alpha}=\underline{0}$ the intersection matrix determinant is exponentially small as $\eta\to 0$ and it is generically asymptotic to its first order value ("Melnikov' formula"). The choice given by (5.10) is a concrete example of this result, generically holding for (5.9).*

Remark: The result in the case f is given by (5.10) is $\det D=24\pi\eta^{-1/2}\varepsilon^2 e^{-\frac{\pi}{2}\eta^{-1/2}}$ in both cases (isochronous, (2.1), and anisochronous, (7.1)) with $a>0$. This theorem has an analogue with and $a=-\frac{1}{2}$, *i.e.* only two time scales: an early review is in [G3]. In the latter case it has been extended to analytic perturbations, *i.e.* beyond the trigonometric case, and to cover the exact asymptotic value of the splitting, *i.e.* far beyond [G3], see [DGJS], [RW]. The above three scales case, (2.1), is quite different from the two scales case (discussed in [G3]): but arises naturally in celestial mechanics problems near a double resonance, as in the case of the precession problem in [CG] to which we hope to apply, eventually, the results of this paper.

The case in (2.1) is that of a pair of "clocks" and a pendulum. The case of a "clock", a "rotator" and a pendulum is exemplified by the hamiltonian obtained by adding $\frac{1}{2}\eta A^2$ to (2.1), see (7.1) below. Both cases are treated in full detail in §6,§7.

§6. Nonperturbative splitting analysis in presence of fast and slow rotations.

It would be easy to show that the determinant of the intersection matrix is exponentially small to order ε^4: this requires evaluating the intersection matrix only to third order. But it cannot be done without due care, as the error in [CG] was precisely due to the belief that it was not necessary to evaluate the matrix element D_{22} because it was exponentially small. In fact it is not exponentially small and it has the right value to make, instead, the whole determinant exponentially small.

The real problem is to compute the determinant to all orders and to show that *to all orders* it is exponentially small: *i.e.* to all orders the determinant is a sum of "large" terms (not exponentially small) which "cancel each other" with a result that is exponentially small. So small that the first order calculation dominates in the limit as $\eta \to 0$.

It is remarkable that in fact one can give an exact expression for the leading corrections to the first order of perturbation expansion. See (6.12),(7.19).

This section relies, and in fact it follows almost immediately, from the general theory of the intersection matrix in [G3]. We cannot repeat here the general theory and therefore refer the reader to [G3] for details on the main definitions: we try nevertheless to make what follows readable at least from a formal viewpoint and as a guide to [G3]. The point of [G3] is that the intersection matrix can be quite explicitly calculated to all orders by using a graphical formalism very similar to that used in quantum field theory when the Schwinger functions are expressed via Feynman's graphs.

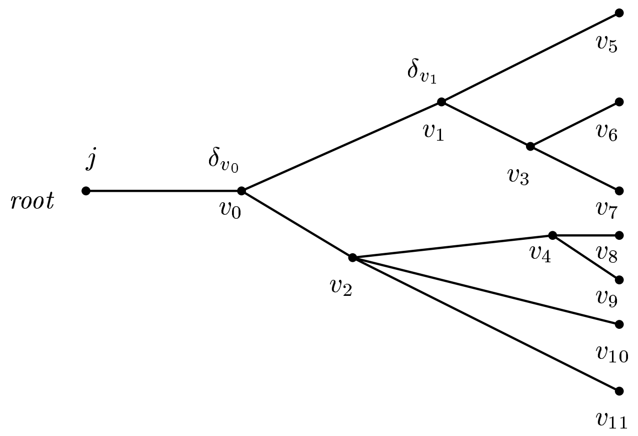
In the present case the graphs will be, topologically, trees: very unusual graphs from the viewpoint of field theory, where loops are often the main source of interest and non triviality. On the other hand the graphs have nodes with arbitrarily large coordination number: also unusual in quantum field theories (with polynomial interactions).

Let ϑ be a tree built with oriented lines all pointing towards a "highest" node r that we call the *root* and that we suppose to have only one "incoming" line, see figure below.

The graph will bear a label $\delta_v = 0, 1$ on each node: if $\delta_v = 1$ it represents $f \stackrel{def}{=} f_1$ while if $\delta_v = 0$ it represents $f_0(\varphi) = g^2 \cos\varphi$. And the labels can be given arbitrarily with the restrictions that all endnodes bear a label $\delta = 1$ and that all nodes bearing a label $\delta = 0$ have at least two incoming lines. Each node will also bear a "time" label τ_v.

We define the *value* of a graph by building the following symbol. We first lay down a set of parentheses () ordered hierarchically and reproducing the tree structure: in fact any tree partially ordered towards the root can be represented as a set of matching parentheses corresponding to the tree nodes. Matching parentheses corresponding to a node v will be made easy to see by appending to them a label v. The root will not be associated with a parenthesis.

Inside the parenthesis $(_v$ and next to it we write $-\partial_0^{p_v+1} f_{\delta_v}$ for all nodes v lower than the node v_0 preceding the root ("first node"). For $v = v_0$ we write $-\partial_j \partial_0^{p_v} f_{\delta_{v_0}}$ where $\partial_0 \stackrel{def}{=} \partial_\varphi$ and $\partial_j \stackrel{def}{=} \partial_{\alpha_j}$, $j = 1, 2$ (this implies that $\delta_{v_0} = 1$): the functions have to be evaluated at $(\varphi^0(t), \underline{\alpha} + \underline{\omega}\, t)$.

6.1 

(6.1)

A graph ϑ with $p_{v_0} = 2, p_{v_1} = 2, p_{v_2} = 3, p_{v_3} = 2, p_{v_4} = 2$ and $k = 12$, $\prod p_v! = 2^4 \cdot 6$, and some decorations. The line numbers, distinguishing the lines, and their orientation pointing at the root, are not shown. The lines length should be the same but it is drawn of arbitrary size. The nodes labels δ_v are indicated only for two nodes.

Outside the parenthesis $(_v$ we write $\mathcal{O}$ for all the $v < v_0$ and we add to the right of the matching parenthesis the symbol (τ_v); for the first node we simply integrate over τ_{v_0} from $+\infty$ to $-\infty$.

The symbol thus defined has the meaning of a linear combination of products of multiple integrals if one uses the definitions of the symbols $\mathcal{O}$, see (5.12). We multiply it by $n!^{-1}$ if n is the number of lines in the graph and we shall regard all the lines different (*i.e.* labeled); *however* two graphs that can be superposed, *labels included*, by successively rotating rigidly around the nodes subtrees that are attached to them have to regarded as identical. This defines the *value* of a graph (it is a function of $\underline{\alpha}$). The reader can see that the above is a rather natural construction by working out patiently the definition in the case of trees with one, two or three nodes.

The sum over all graphs of "order" $k = \sum_v \delta_v$ of the graph values gives the coefficient Q_j^k of order k of the splitting vector $\underline{Q}(\underline{\alpha})$: see [G3] where the above construction is performed in Fourier transform to obtain directly an expression for $\underline{Q}\,^k_{\underline{\nu}}$.

It is convenient to make this more explicit by using the decomposition of $\mathcal{O}$ in the first line of (5.12). This can be easily done by simply attaching to each node lower than the first a label $\beta = O, D, R$ signifying that we select the first of the three terms in the decomposition of $\mathcal{O}$ (see first line in (5.12)) or the second or the third. We can alternatively imagine drawing a circle around each node v enclosing only the subtree with that node v as first node and then to attach the label to the circle.

By (5.12) we see that a circle with a D or R label encircling a node w linked to the higher node v (external to the circle) represents just a function $|w_{03}(\tau_v)|$ or $w_{00}(\tau_v)$ multiplied by a number that in order to be evaluated requires essentially the same operations required to evaluate the value of a graph ϑ.

This allows us to give a *nonperturbative* expression for the splitting vector $\underline{Q}(\underline{\alpha})$. We simply consider the sum of all the values of the graphs bearing a label O on all the nodes except perhaps the endnodes that can bear also D, R labels. We evaluate the graphs values and in the end we replace the number associated with the R, D–labeled endnodes by the *full perturbation series* that is obtained by imagining that inside the circles with R and D labels there is the most general graph with O, R, D labels in all possible ways.

The sum of such perturbation series will be denoted $G^{(0)}(\underline{\alpha})$ for D–labeled circles, and $G^{(1)}(\underline{\alpha})$ for R–labeled circles. The new representation of $\underline{Q}(\underline{\alpha})$ is therefore a representation in terms of trees with a few "*fruits*" around some of the endnodes (possibly none or all) that can be D–labeled or R–labeled (*dry* or *ripe*, to follow the names of [G3]).

Furthermore $G^{(1)}$ and $G^{(0)}$ verify a simple recursion relation that can be found by a more explicit representation of the quantities defined in the same way as $G^{(0)}, G^{(1)}$, $\underline{Q}$ *but evaluated by considering only trees deprived of fruits*, see (6.4) below.

Let ϑ be a graph whose nodes v carry indices $\delta_v = 0, 1$ and τ_v; let v_0 be the first node of ϑ and $v' > v$ be (if $v < v_0$) the node following v.

Fixed ϑ and setting $w(\tau', \tau) = w_{03}(\tau')w_{00}(\tau) - w_{03}(\tau)w_{00}(\tau')$ define the function $A(\{\tau_v\})$ of the time labels τ_v associated with the nodes by:

6.2
$$A = \frac{(-1)^n}{2^n n!} \Big(\prod_{v<v_0} w(\tau_{v'}, \tau_v)\, \varepsilon^{\delta_v}\, \partial_0^{p_v+1} f_{\delta_v}(\tau_v) \Big)\, \partial_0^{p_{v_0}} f_{\delta_{v_0}}(\tau_{v_0}) \tag{6.2}$$

where p_v is the number of lines entering the node v; $f_\delta(\tau) \equiv f_\delta(\varphi(t), \underline{\alpha} + \underline{\omega}\,\tau)$ and n is the number of nodes of the graph ϑ; $\partial_0^q f_\delta(\tau)$ means $(\partial_\varphi^q f_\delta)(\varphi^0(t), \underline{\alpha} + \underline{\omega}\,\tau)$.

We define also $A^{i,j,m,\ldots}_{x,y,w,\ldots}$ by the same expression as A with extra derivatives $\partial_i, \partial_j, \partial_m, \ldots$ acting on the function $f_{\delta_x}, f_{\delta_y}, f_{\delta_m}, \ldots$ in (6.2). We shall say that $i, j, m, \ldots$ are *marks* on the nodes $x, y, w, \ldots$. In terms of graphs we simply represent $A^{i,j,m,\ldots}_{x,y,w,\ldots}$ by affixing *also* "marks" $i, j, m, \ldots$ to the nodes $x, y, w, \ldots$. Then let, for $j = 1, 2$:

6.3
$$\begin{aligned} &\Gamma^{(1)}(\underline{\alpha}) = \textstyle\sum_\vartheta \int |w_{03}(\tau_{v_0})| A^0_{v_0}(\{\tau_v\}), \\ &\Gamma^{(0)}(\underline{\alpha}) = \textstyle\sum_\vartheta \int w_{00}(\tau_{v_0}) A^0_{v_0}(\{\tau_v\}), \qquad \overline{\Gamma}_j(\underline{\alpha}) = 2\textstyle\sum_\vartheta \int A^j_{v_0}(\{\tau_v\}) \end{aligned} \tag{6.3}$$

the integrals ranging over $\tau_{v_0} \in (+\infty, -\infty)$ and over $\tau_v \in (-\infty, \tau_{v'}) \cup (\infty, \tau_{v'})$ for $v < v_0$.

Given the above definitions, $\underline{G} = (G^{(0)}, G^{(1)})$ and $\underline{\Gamma} = (\Gamma^{(0)}, \Gamma^{(1)})$ verify the following (analogues of the well known field theoretic) "Dyson equations" (see also [GGM]):

6.4
$$\begin{aligned} G^{(0)}(\underline{\alpha}) &= \Gamma^{(0)}(\underline{\alpha}) + \sum_{\vartheta; y\in\vartheta} \int_\infty^{-\infty} A^{0,0}_{v_0,y} w_{00}(\tau_{v_0}) \Big(|w_{03}(\tau_y)| G^{(0)}(\underline{\alpha}) + w_{00}(\tau_y) G^{(1)}(\underline{\alpha}) \Big) + .. \\ G^{(1)}(\underline{\alpha}) &= \Gamma^{(1)}(\underline{\alpha}) + \sum_{\vartheta; y\in\vartheta} \int_\infty^{-\infty} A^{0,0}_{v_0,y} |w_{03}|(\tau_{v_0}) \Big(|w_{03}(\tau_y)| G^{(0)}(\underline{\alpha}) + w_{00}(\tau_y) G^{(1)}(\underline{\alpha}) \Big) + .. \end{aligned} \tag{6.4}$$

where the r.h.s. represents the contributions of the trees with no fruit or with just one fruit and the .. denote contributions from the trees with more than just one fruit. Equation (6.4) is obvious if one writes the r.h.s. as a sum of graphs and sees its graphical interpretation: note that (6.4) is a "non perturbative" identity since $\underline{\Gamma}, \underline{G}$ are infinite series in ε (*i.e.* they contain all the contributions from graphs of any order). Then:

6.5
$$Q_j = \overline{\Gamma}_j + 2 \sum_{\vartheta; w\in\vartheta} \int_\infty^{-\infty} A^{j,0}_{v_0,w} \Big(|w_{03}(\tau_w)| G^{(0)}(\underline{\alpha}) + w_{00}(\tau_w) G^{(1)}(\underline{\alpha}) \Big) + \ldots \tag{6.5}$$

with the above meaning of the $\ldots$ (*i.e.* up to sums involving at least two fruits).

The convergence of all the above series and the estimates about their remainders are nontrivial and an *open problem*: they are, however, finite order by order as it follows from the analysis of Appendix A1 of [G3]. Therefore what follows can be interpreted as a collection of identities valid order by order in the expansion and as a way of deriveing and expressing compactly infinitely many identities. This will be sufficient for all our conclusions; nevertheless one would get better bounds if one could prove convergence.

Consider $G^{(0)}_i, \Gamma^{(0)}_i, G^{(1)}_i, \Gamma^{(1)}_i, \overline{\Gamma}_{ij}$ defined as the ∂_i derivatives of $G^{(0)}, \Gamma^{(0)}, G^{(1)}, \Gamma^{(1)}, \overline{\Gamma}_j$, see (6.3), (6.4), *evaluated at* $\underline{\alpha} = \underline{0}$. The terms hidden in the $\ldots$ contribute 0 since *at the homoclinic point by parity* $G^{(0)} \equiv G^{(0)}(\underline{0}) = 0$ *and, as well,* $G^{(1)} \equiv G^{(1)}(\underline{0}) = 0$. Differentiating (6.4) and evaluating at $\underline{\alpha} = \underline{0}$ we get the remarkable *exact* relations:

6.6
$$G^{(0)}_j = \Gamma^{(0)}_j + M_{01}\, G^{(0)}_j + M_{11}\, G^{(1)}_j, \quad G^{(1)}_j = \Gamma^{(1)}_j + M_{00}\, G^{(0)}_j + M_{10}\, G^{(1)}_j \tag{6.6}$$

with the matrix $M=\begin{pmatrix} M_{00} & M_{01}\\ M_{10} & M_{11}\end{pmatrix}$ defined by:

6.7
$$\begin{pmatrix} \sum_{\vartheta;\, y\in\vartheta}\int_{\infty}^{-\infty} A^{0,0}_{v_0,y}|w_{03}(\tau_{v_0})|\,|w_{03}(\tau_y)| & \sum_{\vartheta;\, y\in\vartheta}\int_{\infty}^{-\infty} A^{0,0}_{v_0,y}w_{00}(\tau_{v_0})\,|w_{03}(\tau_y)| \\ \sum_{\vartheta;\, y\in\vartheta}\int_{\infty}^{-\infty} A^{0,0}_{v_0,y}|w_{03}(\tau_{v_0})|\,w_{00}(\tau_y) & \sum_{\vartheta;\, y\in\vartheta}\int_{\infty}^{-\infty} A^{0,0}_{v_0,y}w_{00}(\tau_{v_0})\,w_{00}(\tau_y) \end{pmatrix} \tag{6.7}$$

M is symmetric: this is implied by the symmetry of the operator $\mathcal{O}_0$, which follows immediately from that of $\mathcal{O}$ (see (5.13), and Appendix A2 of [G3]: or proceed inductively). More generally the symmetry of $\mathcal{O}_0$, repeatedly applied (as done with $\mathcal{O}$ in [G3]), implies that $\sum_{\vartheta,y}\int A^{0,0}_{v_0,y}F(\tau_{v_0})H(\tau_y)=\sum_{\vartheta,y}\int A^{0,0}_{v_0,y}H(\tau_{v_0})F(\tau_y)$ for all F,H: *used several times below.*

Going back to (6.5), in order to evaluate the derivatives $\partial_i Q_j(\underline{\alpha})$, $i,j=1,2$ at $\underline{\alpha}=\underline{0}$ we need only the linear terms and the $\ldots$ contribute exactly 0. Therefore, if we set:

6.8
$$\overline{\Gamma}_{ij}=2\sum_{\vartheta;\, y\in\vartheta}\int A^{j,i}_{v_0,y} \tag{6.8}$$

and note that $\underline{G}=(1-\sigma M)^{-1}\,\underline{\Gamma}$ with $\sigma=\begin{pmatrix}0&1\\1&0\end{pmatrix}$, then the intersection matrix D is:

6.9
$$D_{ij}=\overline{\Gamma}_{ij}+2(\Gamma^{(1)}_j G^{(0)}_i+\Gamma^{(0)}_j G^{(1)}_i) \tag{6.9}$$

which, setting $C=\sigma(1-\sigma M)^{-1}$, becomes:

6.10
$$D=\begin{pmatrix}\overline{\Gamma}_{11}+2(\underline{\Gamma}_1,C\,\underline{\Gamma}_1) & \overline{\Gamma}_{12}+2(\underline{\Gamma}_2,C\,\underline{\Gamma}_1)\\ \overline{\Gamma}_{21}+2(\underline{\Gamma}_1,C\,\underline{\Gamma}_2) & \overline{\Gamma}_{22}+2(\underline{\Gamma}_2,C\,\underline{\Gamma}_2)\end{pmatrix} \tag{6.10}$$

The matrix C is symmetric, because M is symmetric; hence D is a symmetric matrix (to all orders as (6.9) is *exact*): this agrees with the mentioned result (see comment following (4.2)) of Lochak and Sauzin on the symmetry of D. Noting that $\Gamma^{(0)}_2=\sum_1^\infty\varepsilon^n\Gamma^{(0)(n)}_2$, $\overline{\Gamma}_{22}=\sum_1^\infty\varepsilon^n\overline{\Gamma}^{(n)}_{22}$, with an explicit calculation yielding in the case (5.10):

6.11
$$\Gamma^{(0)(1)}_2=\frac14\eta^{-1/2}\overline{\Gamma}^{(1)}_{22}=2\pi\eta^{-1}e^{-\frac{\pi}{2}\eta^{-1/2}}(1+O(e^{-2\pi\eta^{-1/2}})) \tag{6.11}$$

An elementary (and *somewhat surprising*) calculation of the determinant yields a remarkably simple expression in which the first non trivial order of $\det D$ (*i.e.* the second) and the possibly non exponentially small contributions are evaluated *without any approximation*:

6.12
$$\det D=\overline{\Gamma}^{(1)}_{11}\overline{\Gamma}^{(1)}_{22}\;-2\,(\overline{\Gamma}_{11}M_{11}-2{\Gamma^{(0)}_1}^2){\Gamma^{(1)}_2}^2\,\det C\,+\,\ldots \tag{6.12}$$

the last $\ldots$ collects contributions to $\det D$ from the products in $\det D$ involving $\overline{\Gamma}_{12},\overline{\Gamma}_{22}$ (for instance $-4\overline{\Gamma}_{12}(\underline{\Gamma}_1,C\,\underline{\Gamma}_2)$ or $2\overline{\Gamma}_{22}(\underline{\Gamma}_1,C\,\underline{\Gamma}_1)$, $-\overline{\Gamma}^2_{12}$ and a few others) or the higher orders of $\overline{\Gamma}_{11}\overline{\Gamma}_{22}$.

The terms not written in (6.12) contain as factors (see Appendix A2 below) derivatives ∂_2 ("fast angle derivatives") of integrals of an analytic function of the τ_v's: if we imagine developing the kernels A as Fourier series in $\underline{\alpha}$ we see that only Fourier components $\underline{\nu}=(\nu_1,\nu_2)$ with $\nu_2\ne0$ can contribute. This means that the dependence on the τ_v will be via the matrix elements $w(\tau_{v'},\tau_v)$ of the pendulum wronskian matrix and via $e^{i\sum_v\tau_v\underline{\omega}\cdot\underline{\nu}_v}$ with $\sum_v\underline{\nu}_{v2}\ne0$ (we say that "the dependence on time is via a total fast rotation").

So that by shifting the integration over τ_v to $\sim\pm i(\pi/2-\eta^{1/2})$, *i.e.* almost as far out in the complex as the singularities of the w–functions allow, one sees that such terms contribute a quantity bounded to all orders proportionally to $e^{-\frac{\pi}{2}\eta^{-1/2}}$ because

$|\underline{\omega}\cdot\underline{\nu}|>\eta^{-1/2}-Nh\eta^{1/2}$ if $\nu_2\neq 0$: this gives a good bound for h not too large, *e.g.* $h<\eta^{-12}$ (actually even for $h\simeq\eta^{-1}$), see Appendix A2 for details.

For $h>\eta^{-1/2}$ one can invoke the convergence of the series for $\det D$ and obtain a bound $(\eta^{-4d}\varepsilon)^{\eta^{-1/2}}$ (much smaller than $e^{-\frac{\pi}{2}\eta^{-1/2}}$ because $\varepsilon=\mu\eta^c$, provided c is large enough). The conclusion is that the terms not written in (6.12) can be bounded by $\varepsilon^3 e^{-\frac{\pi}{2}\eta^{-1/2}}O(\eta^{-3\beta})$, provided c is large enough so that the sum of the bounds of the orders from 3 to $\eta^{-1/2}$ is dominated by the third order bound: see Appendix A2 for details.

The value of β can be taken $2(N+1)+4d$ in terms of the order of the degree N of f and of the diophantine constant d in (2.2), it is explained by the singularities of the elementarily computable Fourier transforms of $\cos N\varphi(t)$ and $\sin N\varphi(t)$, see [GR]. A similar argument is in §8 of [G3].

The above argument is good but it gives a rather poor bound: see Appendix A2 for some improvements.

Hence the leading value of $\det D$ *is given, as* $\eta\to 0$, *by its first order expression* $\overline{\Gamma}^{(1)}_{11}\overline{\Gamma}^{(1)}_{22}$ plus the *apparently much larger* $-2(\overline{\Gamma}_{11}M_{11}-2\Gamma^{(0)2}_1)\,\Gamma^{(1)2}_2\det C$. But we shall show that:

6.13
$$(-\overline{\Gamma}_{11}M_{11}+2{\Gamma^{(0)}_1}^2)\,{\Gamma^{(1)}_2}^2\det C=O(\eta^{-4\beta})\varepsilon^4 e^{-\frac{\pi}{2}\eta^{-1/2}} \tag{6.13}$$

because of the first factor being of $O(\varepsilon^2\eta^{-2\beta}e^{-\frac{\pi}{2}\eta^{-1/2}})$, *i.e.* essentially 0, again assuming (temporarily) convergence of the series for $\underline{G}\,,\underline{\Gamma}\,,\overline{\Gamma},M$.

The discussion of the convergence for the series for $\overline{\Gamma}$, $\underline{\Gamma}\,,M$ is *very non trivial*, while one could show the convergence of the series for $\underline{G}\,$, following [Ge1]. However the series for $\det D$ converges and its convergence, *which is absolutely essential*, follows immediately from Theorem 1 above.

It is remarkable that we can *avoid* proving the convergence of the power series for $\overline{\Gamma}$, $\underline{\Gamma}\,,M,\underline{G}$ and get away with only the easily established convergence of $\det D$: in fact we can just use the above series as formal power series and that is all we really need (*together* with the convergence of $\det D$).

After all the identity (6.12) as an identity between formal power series and the formal bound (6.13) show that *to all orders the* $\det D$ *is exponentially bounded* and this is almost enough. In fact up to order $\eta^{-1/2}$ one needs only bounds on finitely many terms ($O(\eta^{-1/2})$) in the series for $\overline{\Gamma}$, $\underline{\Gamma}\,,M,\underline{G}\,$, as explained above. For the remainder we use that it is bounded by the much smaller quantity $(\varepsilon\eta^{-\overline{c}})^{\eta^{-1/2}}$ derived from the convergence of the series for $\det D$ (if μ in $\varepsilon=\mu\eta^c$ is small enough and c is large enough). The details are given in Appendix A2 below.

We suggest to proceed however by, *at first*, assuming convergence of the series for $\overline{\Gamma}$, $\underline{\Gamma}\,,M,\underline{G}$ for $\varepsilon\eta^{-c}$ small enough, and only on a second reading check that formal power series considerations (plus analyticity of $\det D$) suffice: a technique used in [G3], §8. Here *the difficulty is not the bounds but the cancellations* and assuming convergence removes unessential worries and clarifies the algebra.

We also take $a=\frac{1}{2}$ *in the following calculations* to simplify notations: the general case is obtained by replacing the coefficients *explicitly* appearing in the following formulae ω by η^a and ω^2 by $\eta^{\frac{1}{2}+a}$ respectively. We set $\varphi(t)\equiv\varphi^0(t)$.

For instance in the case (5.10) (with $a=\frac{1}{2}$) $\overline{\Gamma}^{(1)}_{11}=4+O(\eta^{1/2})$ by direct computation and, by (6.11), $\overline{\Gamma}^{(1)}_{22}=8\varepsilon\pi\eta^{-1/2}e^{-\frac{1}{2}\pi\eta^{-1/2}}$ the leading term in (6.12) is:

6.14
$$\det D=24\,\pi\eta^{-1/2}\varepsilon^2 e^{-\frac{\pi}{2}\eta^{-1/2}} \tag{6.14}$$

The reason why (6.13) holds is that if $\omega\stackrel{def}{=}\eta^{1/2}$:

6.15
$$\overline{\Gamma}_{11}=\frac{4}{\omega}\Gamma^{(0)}_1-\frac{1}{\omega^2}\overline{\Gamma}_{12},\qquad M_{11}=\frac{\omega}{2}(\Gamma^{(0)}_1+\frac{1}{\omega^2}\Gamma^{(0)}_2) \tag{6.15}$$

To prove (6.15) note that $\sum_{\vartheta,y}\int A^{10}_{v_0 y} w_{00}(\tau_y) = \sum_{\vartheta,y}\int A^{01}_{v_0 y} w_{00}(\tau_{v_0})$ by the symmetry of $\mathcal{O}_0$; hence:

6.16
$$\overline{\Gamma}_{11} = 2\sum_{\vartheta,y}\int A^{11}_{v_0 y}, \quad \Gamma^{(0)}_1 = \sum_{\vartheta,y}\int A^{10}_{v_0 y} w_{00}(\tau_y) \tag{6.16}$$

Thinking (6.16) as sums of graphs we see that to each graph with n nodes contributing to $\overline{\Gamma}_{11}$ with the node y marked 1 (see definition preceding (6.3)) as in (6.16) there correspond *two* graphs contributing, if $y < v_0$, to $\Gamma^{(0)}_1$. Namely the one with the node y marked 0 and the one obtained by deleting the mark 0 on y, adding a new node y' marked 0 on the line coming out of y and with index $\delta_{y'} = 0$. If $y = v_0$ we associate with it only the first of the above graphs. Note that the second graph has $n+1$ nodes.

Suppose that y is an endnode in a graph ϑ associated with $\overline{\Gamma}_{11}$: by A–kernels definition we must evaluate, when computing the contribution to $\overline{\Gamma}_{11}$ from such graph, the quantity:

6.17
$$\mathcal{O}_0(\partial_{10} f) = \frac{1}{\omega}\mathcal{O}_0(\partial_\tau \partial_0 f) + \frac{1}{\omega}\mathcal{O}_0(-\dot\varphi\,\partial_{00} f) - \frac{1}{\omega^2}\mathcal{O}_0(\partial_{20} f) \tag{6.17}$$

where $\omega \stackrel{def}{=} \eta^{1/2}$(recall also that we take $a = \frac{1}{2}$) and having used $\partial_1 \equiv \frac{1}{\omega}(\partial_\tau - \dot\varphi\,\partial_0) - \frac{1}{\omega^2}\partial_2$ (because the derivatives act on functions of spacial form $F(\varphi(t), \underline{\omega}\, t)$ with $F(\varphi, \underline{\alpha})$ suitable).

Since $\dot\varphi = -2w_{00}$ we see that the second term is $\frac{2}{\omega}\mathcal{O}_0(w_{00}(\tau_v)\partial_{00} f)$: hence it appears in the evaluation of the first among the two corresponding graph contributions to $\frac{4}{\omega}\Gamma^{(0)}_1$. The second corresponding graph contribution to $\frac{4}{\omega}\Gamma^{(0)}_1$ will require computing:

6.18
$$-\frac{4}{\omega}\mathcal{O}_0(w_{00}(\tau_{y'})\partial^3_0 f_0 \mathcal{O}_0(\partial_0 f)) = \frac{2}{\omega}\mathcal{O}_0(\dot\varphi\,\partial^3_0 f_0 \mathcal{O}_0(\partial_0 f)) \tag{6.18}$$

The combinatorial coefficient associated with this graph is $(n+1)!$ rather than $n!$: but one checks that this is what is necessary to verify that the difference between the contributions to $\overline{\Gamma}_{11}$ and $\frac{4}{\omega}\Gamma^{(0)}_1$ due to the three graphs considered requires computing:

6.19
$$\frac{1}{\omega}\mathcal{O}_0\big(\partial_{\tau_y}\partial_0 f - \dot\varphi\,\partial^3_0 f_0 \mathcal{O}_0(\partial_0 f)\big) - \frac{1}{\omega^2}\mathcal{O}_0\big(\partial_{02} f\big) \tag{6.19}$$

Below we prove for all odd $F(t)$ the following commutation relation:

6.20
$$\partial_{\tau_{y'}}\mathcal{O}_0(F) = \mathcal{O}_0\big(\partial_{\tau_y} F - \dot\varphi\,\partial^3_0 f_0 \mathcal{O}_0(F)\big) \tag{6.20}$$

which, noting that below we only apply (6.20) for odd F's, allows us to rewrite (6.19) as:

6.21
$$\frac{1}{\omega}\partial_{\tau_{y'}}\mathcal{O}_0(\partial_0 f) - \frac{1}{\omega^2}\mathcal{O}_0(\partial_{02} f) \tag{6.21}$$

We proceed by picking a new endnode (if any) and by repeating the construction until all endnodes are exhausted. The difference between the contributions to $\overline{\Gamma}_{11}$ and $\frac{4}{\omega}\Gamma^{(0)}_1$ considered so far will therefore be given by a collection of graph values contributing to $\overline{\Gamma}_{11}$ with *no marked node* and with one of the endnodes requiring the evaluation of:

6.22
$$\frac{1}{\omega}\partial_{\tau_{y'}}\mathcal{O}_0(\partial_0 f) - \frac{1}{\omega^2}\mathcal{O}_0(\partial_{20} f) \tag{6.22}$$

Consider now a 1–marked node y which is next to an endnode; and suppose that there are p lines linking it to the endnodes. The differences between the contribution of the considered graphs to $\overline{\Gamma}_{11}$ and of the corresponding ones to $\frac{4}{\omega}\Gamma^{(0)}_1$ *plus* the already evaluated differences due to the previously considered graphs will be, proceeding as before:

6.23
$$\begin{aligned}&\frac{1}{\omega}\mathcal{O}_0\Big((\partial_\tau\partial_0^{p+1}f)\mathcal{O}_0(\partial_0 f)^p\Big)+\frac{1}{\omega}\mathcal{O}_0\Big(\partial_0^{p+1}f\sum_{j=1}^{p}\mathcal{O}_0(\partial_0 f)^{p-1}\partial_{\tau_y}\mathcal{O}_0(\partial_0 f)\Big)+\\&+\frac{1}{\omega}\mathcal{O}_0\Big(-\dot\varphi\,\partial_0^3 f_0\mathcal{O}_0\big(\partial_0^{p+1}f\mathcal{O}_0(\partial_0 f)^p\big)\Big)-\frac{1}{\omega^2}\mathcal{O}_0\Big(\partial_2\partial_0^{p+1}f\mathcal{O}_0(\partial_0 f)^p\Big)-\\&-\frac{1}{\omega^2}\mathcal{O}_0\Big(\partial_0^{p+1}f\sum_{j=1}^{p}\mathcal{O}_0(\partial_0 f)^{p-1}\mathcal{O}_0(\partial_{20}f)\Big)\end{aligned}\tag{6.23}$$

which is, by (6.20):

6.24
$$\frac{1}{\omega}\partial_{\tau_{y'}}\mathcal{O}_0(\partial_0^{p+1}f\mathcal{O}_0(\partial_0 f)^p)-\frac{1}{\omega^2}\partial_2\mathcal{O}_0(\partial_0^{p+1}f\mathcal{O}_0(\partial_0 f)^p)\tag{6.24}$$

We proceed in the same way until we reach the first node and, in fact, we can treat in the same way *also* the first node except that in this case there is only one corresponding graph in $\Gamma_1^{(0)}$. No need, this time, of using the commutation relation (6.20) because the τ_{v_0}–integral is a simple integral not involving $\mathcal{O}_0$ operations any more. Hence eventually we find that the differences between the contributions to $\overline{\Gamma}_{11}$ and $\frac{4}{\omega}\Gamma_1^{(0)}$ add up to:

6.25
$$\Big(\frac{2}{\omega}\int\partial_t A^1_{v_0}\Big)-\Big(\frac{2}{\omega^2}\partial_2\int A^1_{v_0}\Big)=-\frac{2}{\omega^2}\partial_2\int A^1_{v_0}=-\frac{1}{\omega^2}\overline{\Gamma}_{12}\tag{6.25}$$

where one should note that the first term in the l.h.s. vanishes: this does not immediately follow from it being an integral of a derivative because the integrals are improper: one checks this by (5.1) and that $A^1_{v_0}$ is even (at $\underline{\alpha}=\underline{0}$).

In a similar way one shows that $M_{11}=\frac{\omega}{2}(\Gamma_1^{(0)}+\frac{1}{\omega^2}\Gamma_2^{(0)})$. One proceeds by writing $\Gamma_1^{(0)}$ as $\Gamma_1^{(0)}=\sum_{\vartheta,y}\int A^{01}_{v_0y}w_{00}(\tau_{v_0})$, and marking 0 the first node for both $\Gamma_1^{(0)}$ and M_{11}. Then note that to each graph contributing to $\Gamma_1^{(0)}$ with the node y marked 1 there correspond *two* graphs contributing to M_{11} also for $y=v_0$, when $\delta_{v_0}=1$, and they are constructed by following the same prescription given after (6.16); moreover to each graph contributing to $\Gamma_1^{(0)}$ with $\delta_{v_0}=0$ there correspond two graphs contributing to M_{11}, namely the one with the node v_0 marked 0 twice and the one obtained by adding a new node v_0' (which becomes the new first node), marked 0 twice and with $\delta_{v_0'}=0$, on the line coming out of y. Then proceed as before, with the only difference that, at the last step, when dealing with the contributions arising from the graphs with $\delta_{v_0}=0$, one has to use the identity $\mathcal{O}_0(w_{00}^2\sin\varphi)=-\frac{1}{2}\frac{\sinh t}{\cosh^2 t}=\frac{1}{2}\frac{d}{dt}\frac{1}{\cosh t}=-\frac{1}{4}\ddot\varphi$. This can be proved in the same way as the commutation relation (6.20).

Let $\underline{F}=\begin{pmatrix}0\\F\end{pmatrix}$ and $\underline{T}=\begin{pmatrix}\mathcal{O}_0(F)\\\partial_t\mathcal{O}_0(F)\end{pmatrix}$: the vector $\underline{T}$ verifies $\dot{\underline{T}}=L(t)\underline{T}+\underline{F}$ where $L(t)=\begin{pmatrix}0&1\\\cos\varphi&0\end{pmatrix}$. By differentiating with respect to t both sides we get $\ddot{\underline{T}}=L(t)\dot{\underline{T}}+\dot{\underline{F}}+\begin{pmatrix}0&0\\-\partial_0^3 f_0&0\end{pmatrix}\dot\varphi\,\underline{T}$ which means (6.20), up to a homogeneous solution $t\to W(t)X$, $X\in R^2$ of the latter linear equation, when written for the first component; the oddness of F is necessary to check that $X=0$. Choosing $\underline{T}=\begin{pmatrix}w_{00}\\\dot w_{00}\end{pmatrix}$ and $F=0$ one gets in the same way the above mentioned identity (recall $\dot\varphi=-2w_{00}$). A similar identity is obtained by choosing $\underline{T}=\begin{pmatrix}w_{03}\\\dot w_{03}\end{pmatrix}$, useful in §7.

This completes the proof that the cancellation discussed in §5 is in fact taking place to all orders of perturbation theory.

§7. The splitting in anisochronous cases.

We now proceed to the analysis of an anisochronous case:

7.1 $$H = \eta^a A + \frac{1}{\eta^{1/2}} B + \eta^{2a} \frac{A^2}{2} + \frac{I^2}{2} + g^2 (\cos\varphi - 1) + \varepsilon\, f(\varphi, \alpha, \lambda) \tag{7.1}$$

with $\varepsilon = \mu\,\eta^c$ and c will be chosen large enough. This model belongs, see [T],[G2], to a well studied class of models ("Thirring models") and it is a simplified version of the hamiltonian considered in [CG] in §12. If we added (7.1) a further "*monochromatic*" term $f_0(\varphi, \alpha, \lambda)$ which has Fourier harmonics $\underline{\nu}$, for (α, λ), multiples of a given *fast* harmonic $\underline{\nu}^0$ (*i.e.* a harmonic with the component ν_2^0 different from 0) it would offer all the difficulties of that case in spite of being much simpler analytically. We shall not attempt the analysis of such more general hamiltonian: the error in [CG] eliminated almost completely this problem which, hence, has to be studied again if one wants to recover the results of that paper.

Consider, *as an example*, the case (5.10). This time the first problem is to find for how many values of A one can have, for ε small, an invariant hyperbolic torus with rotation number:

7.2 $$\underline{\omega} = (\eta^a + \eta^{2a} A, \eta^{-1/2}) \tag{7.2}$$

and ε–close to an unperturbed one.

For hamiltonians like (7.1) the average position in the A variables of the torus with rotation (7.2) will be exactly A (this is a general property of "twistless tori", as discussed in [G3],[G4]). The average position in the B variables will be chosen so that the energy of the torus is a fixed value, *e.g.* 0: this can be done because the fast variable λ conjugate to B is still isochronous (see §5 in [CG] for the more general case in which also λ is anisochronous when the analysis is somewhat more involved).

Of course $\underline{\omega}$ must verify a diophantine condition: but in view of using the result to show that the tori resisting the perturbation, *i.e.* the ones described by remark (7) to theorem 1 above, are dense enough to build long chains of tori joined by heteroclinic trajectories, "*heteroclinic chains*", we must consider tori that verify a *very generous* diophantine property, compared to (2.2).

Since we shall not discuss the ambitious application attempted in [CG] we just allow rotations vectors verifying (2.2) with a fixed, possibly very large, d.

¿From lemmata 1,1' in §5 of [CG], all tori with rotation $\underline{\omega} \stackrel{def}{=} \underline{\omega}\,(A) \stackrel{def}{=} (\eta^a + \eta^{2a} A, \eta^{-1/2})$ verifying:

7.3 $$|\,\underline{\omega} \cdot \underline{\nu}\,| > C^{-1} \eta^d |\,\underline{\nu}\,|^{-3} \stackrel{def}{=} C(\eta) |\,\underline{\nu}\,|^{-3} \tag{7.3}$$

will survive the perturbation if the parameter c in the definition $\varepsilon = \mu\eta^c$ of the coupling constant is large enough: so that $\varepsilon C(\eta)^{-q} < \varepsilon_0$ for some ε_0 and some $q > 0$

The splitting theory is "insensitive", at given $\underline{\omega}$, to the presence or absence of the isochrony breaking term $\frac{\eta^{2a}}{2} A^2$ in (7.1) provided $\lambda = \eta^{1-a} \xrightarrow[\eta\to 0]{} 0$. We discuss this delicate point below, *for general perturbation f*, see (5.9). The homoclinic splitting is given by (6.14) *with no extra leading terms*. The only effect of the anisochrony is to introduce a few gaps in the foliations of phase space into stable and unstable manifolds: but it has also the advantage that we no longer must be careful about the values of η. it Anisochrony guarantees that the diophantine conditions holds for "many" values of A.

Turning to the *main point* of this section (and of the whole paper) *we prove that usually the first order ("Melnikov's integral") dominates the splitting.* Again the technique will be based on the general theory of [G3].

In the anisochronous case the graph labels have to be extended, see (5.7), (5.14) and [G3]. On each node v one adds a further node label $j_v = 0, 1$ (which in the isochronous case would be $j_v \equiv 0$) and this has the effect that in the definition of A one replaces:

7.4

$$\begin{aligned} &\partial_0^{p_v+1} f_{\delta_v} \ \to\ \partial_{j_v}\partial_{j_{v_1}}\ldots\partial_{j_{v_{p_v}}} f_{\delta_v} \qquad \text{if } v < v_0 \\ &\partial_0^{p_{v_0}} f_{\delta_{v_0}} \ \to\ \partial_{j_{v_1}}\ldots\partial_{j_{v_{p_{v_0}}}} f_{\delta_{v_0}} \qquad \text{if } v = v_0 \end{aligned} \tag{7.4}$$

where p_v is the number of nodes $v_1, \ldots, v_{p_v}$ preceding v. Furthermore the kernels $w(\tau'_v, \tau_v)$ become node dependent $w_v(\tau'_v, \tau_v)$ and equal to $w(\tau'_v, \tau_v)$ if $j_v = 0$ and $(\tau_{v'} - \tau_v)$ if $j_v = 1$.

The first component $X_1(\underline{\omega}\,t, t, \underline{\alpha})$ of $\underline{X}_\downarrow$ will not vanish. Let us define $\alpha(t)$ as $\alpha + \omega t + \sum_{k=0}^\infty \varepsilon^k X_1^k(\underline{\omega}\,t, t, \underline{\alpha})$ where X_1 is the first (and only non vanishing) component of $\underline{X}_\downarrow$.

The contributions to the splitting Q_j due to fruitless trees will be $2\sum_\vartheta A^j_{v_0}$, with the same notations of §6. The full splitting will be:

7.5

$$Q_j = 2\sum_\vartheta \int A^j_{v_0} + \sum_{\vartheta, y, r} \int A^{j,[r]}_{v_0,y} w_r(\tau_y) G^{(r)}(\underline{\alpha}) + .. \tag{7.5}$$

here $[r] = 0$ if $r = 0, 1$ and $[r] = 1$ if $r = 2, 3$, and:

7.6

$$w_0(\tau) = w_{00}(\tau), \quad w_1(\tau) = |w_{03}(\tau)|, \quad w_2(\tau) = \eta, \quad w_3(\tau) = \eta|\tau| \tag{7.6}$$

with $\underline{G} = (G^{(0)}, G^{(1)}, G^{(2)}, G^{(3)})$ representing the fruit values, defined as in §6, for fruits which now can carry also a label $2, 3$ on the first node: the latter values correspond to the fruits carrying label $j_v = 1$ (2 corresponds to a dry fruit and 3 to a ripe fruit): the choices of the w_2, w_3 arise from the form of the operator corresponding to $\mathcal{O}$ for the nodes with the new labels, called $\overline{\mathcal{I}}^2$ in [G3].

In complete analogy with §6 the $\underline{G}$ verify Dyson's equations. If we set:

7.7

$$\Gamma^{(r)}(\underline{\alpha}) = \sum_\vartheta \int A^{[r]}_{v_0} w_r(\tau_{v_0}), \qquad M_{rs} = \sum_{\vartheta, y} \int A^{[r],[s]}_{v_0,y} (\sigma\,\underline{w}\,)_r(\tau_{v_0})(\sigma\,\underline{w}\,)_s(\tau_y) \tag{7.7}$$

where the matrix σ is defined to be $\sigma_{rs} = 0$ except for the matrix elements $\sigma_{01} = \sigma_{10} = \sigma_{23} = \sigma_{32} = 1$:

7.8

$$G^{(r)}(\underline{\alpha}) = \Gamma^{(r)}(\underline{\alpha}) + \sum_{\vartheta, y, s} \int (\,\underline{w}\,)_r(\tau_{v_0}) A^{[r],[s]}_{v_0,y} (\sigma\,\underline{w}\,)_s(\tau_y) G^{(s)}(\underline{\alpha}) + \ldots \tag{7.8}$$

where the dots represent contributions from the graphs with more than one fruit, while the terms explicitly written represent the contributions from the graphs with no fruits or with just one fruit.

At the homoclinic point the derivatives $\underline{G}_j = \partial_j$, $\underline{\Gamma}_j = \partial_j \underline{\Gamma}$ verify exactly:

7.9

$$\underline{G}_j = \underline{\Gamma}_j + \sigma M\,\underline{G}_j, \qquad \underline{G}_j = (1 - \sigma M)^{-1}\,\underline{\Gamma}_j \tag{7.9}$$

(compare with (6.6)).

The intersection matrix will be, setting as in §6, $\overline{\Gamma}_{ij} = 2\sum_{\vartheta,y} \int A^{j,i}_{v_0,y}$, $i, j = 1, 2$ and $C = (\sigma - \sigma M \sigma)^{-1}$:

7.10

$$D_{ij} = \overline{\Gamma}_{ij} + 2(\,\underline{\Gamma}_j, \sigma\,\underline{G}_i) = \overline{\Gamma}_{ij} + 2(\,\underline{\Gamma}_j, C\,\underline{\Gamma}_i) \tag{7.10}$$

The convergence of the above series , (7.5)÷(7.8) and the estimate of their remainders is discussed as in §6: see Appendix A2.

The above equation is not sufficient this time: there are in fact too many variables. There are however several relation between the matrix elements of M and $\underline{\Gamma}$, $\underline{G}$. In fact M is symmetric for the same reasons as the corresponding matrix in §6: *i.e.* by using the symmetry of $\mathcal{O}_0$ and of $\overline{\mathcal{I}}_0^2$ operators defined in (5.14) (the symmetry of $\overline{\mathcal{I}}_0^2$ is proved in the same way as that of $\mathcal{O}_0$). If $\lambda \stackrel{def}{=} \frac{2\eta^{2a}}{\omega}, \omega = \eta^a$ the relations are, *up to terms of order* η^∞:

7.11

$$\begin{aligned} M_{11} = \frac{\omega}{2}\Gamma_1^{(0)}, \quad \Gamma_1^{(2)} = \lambda\Gamma_1^{(0)}, \quad G_1^{(2)} = Z\,\lambda\, G_1^{(0)} \\ M_{03} = \lambda M_{01} - \lambda\Lambda_0, \quad M_{23} = \lambda M_{12} - \lambda\Lambda_1 \\ M_{33} = \lambda^2 M_{11}, \quad M_{31} = \lambda M_{11} \end{aligned} \tag{7.11}$$

where $\Lambda_0 = (2\omega)^{-2}\partial_2 \sum \int |w_{03}(\tau_{v_0})| A^0_{v_0}$ and $\Lambda_1 = (2\omega)^{-2}\partial_2 \sum \int |\tau_{v_0}| A^0_{v_0}$ and $Z = \frac{1-\Lambda_0}{1+\lambda\Lambda_1}$. The relations among the M elements are proved by the same argument discussed in §6 for the first of (7.11). One should use the relation $\mathcal{O}_0(w_{00}|w_{03}|\sin\varphi) = -2|\dot w_{03}|$, see the final remark in §6 and note that $\mathcal{O}_0(\operatorname{sign}\tau\, F)(t) = \operatorname{sign} t\, \mathcal{O}_0(F)(t)$. The constant Z arises solving by iteration (7.9): the structure of the matrix σM and of its powers is, given the relations between the M_{ij} in (7.11), such that the first and third components of $\underline{G}$ are proportional via the constant λZ.

Equation (7.11) allows us to reduce the size of the vectors $\underline{G}$, $\underline{\Gamma}$ and of the matrix M. We shall denote with a tilde the new vectors and matrices. Introduce $\underline{\tilde G} = (G^{(0)}, G^{(1)}, G^{(3)})$, $\underline{\tilde\Gamma} = (\Gamma^{(0)}, \Gamma^{(1)}, \Gamma^{(3)})$, and $\tilde M, N$ as:

7.12

$$\tilde M = \begin{pmatrix} M_{10} + Z\lambda M_{12} & M_{11} & \lambda M_{11} \\ M_{00} + Z\lambda M_{02} & M_{01} & \lambda M_{01} - \lambda\Lambda_0 \\ M_{20} + Z\lambda M_{22} & M_{21} & \lambda M_{21} - \lambda\Lambda_1 \end{pmatrix}, \quad N = \begin{pmatrix} 0 & 1 & \lambda \\ 1 & 0 & 0 \\ Z\lambda & 0 & 0 \end{pmatrix} \tag{7.12}$$

The equations (7.9), (7.10) become respectively:

7.13

$$\underline{\tilde G} = (1-\tilde M)^{-1}\,\underline{\tilde\Gamma}, \qquad D_{ij} = \overline{\Gamma}_{ij} + 2(\,\underline{\tilde\Gamma}_i, N(1-\tilde M)^{-1}\,\underline{\tilde\Gamma}_j\,) \tag{7.13}$$

noting that the matrix $\tilde C = N(1-\tilde M)^{-1}$ is symmetric (because C in (7.10) is symmetric) and that it has the second and third row proportional one deduces that:

7.14

$$\begin{aligned} \det D = \overline{\Gamma}_{11}\overline{\Gamma}_{22} + 2(\Gamma_2^{(1)})^2\big(\overline{\Gamma}_{11}\tilde C_{11} + 2(\Gamma_1^{(0)})^2\Delta_{00,11}\big) + \\ + 2(\Gamma_2^{(3)})^2\big(\overline{\Gamma}_{11}\tilde C_{33} + 2(\Gamma_1^{(0)})^2\Delta_{00,33}\big) + 4\Gamma_2^{(1)}\Gamma_2^{(3)}\big(\overline{\Gamma}_{11}\tilde C_{13} + 2(\Gamma_1^{(0)})^2\Delta_{00,13}\big) \end{aligned} \tag{7.14}$$

where $\Delta_{00,11}, \Delta_{00,33}, \Delta_{00,13}$ denote the determinants of the matrices:

7.15

$$\begin{pmatrix} \tilde C_{00} & \tilde C_{01} \\ \tilde C_{10} & \tilde C_{11} \end{pmatrix}, \quad \begin{pmatrix} \tilde C_{00} & \tilde C_{03} \\ \tilde C_{30} & \tilde C_{33} \end{pmatrix}, \quad \begin{pmatrix} \tilde C_{00} & \tilde C_{03} \\ \tilde C_{10} & \tilde C_{13} \end{pmatrix} \tag{7.15}$$

To compute all the above quantities we note that if we set (see (7.12)) $a = M_{11}, b = M_{01}, c = M_{12}$ and $x = M_{10} + Z\lambda M_{12} = b + Z\lambda c$, $y = M_{00} + Z\lambda M_{02}$, $z = M_{20} + Z\lambda M_{22}$, $\Lambda_0 = -b', \Lambda_1 = -c'$, $Z = (1+b')(1-\lambda c')^{-1}$:

7.16

$$1 - \tilde M = \begin{pmatrix} 1-x & -a & -\lambda a \\ -y & 1-b & -\lambda b - \lambda b' \\ -z & -c & 1 - \lambda c - \lambda c' \end{pmatrix} \tag{7.16}$$

hence $\det(1-\tilde M) = -\Big((y + \lambda z Z)a - (1-x)^2\Big)(1-\lambda c') \stackrel{def}{=} \Delta$, and $(1-\tilde M)^{-1}$ is:

7.17

$$\frac{1}{\Delta}\begin{pmatrix} (1-b)(1-\lambda\tilde c) - \lambda\tilde b c & a(1-\lambda c') & \lambda a(1+b') \\ y(1-\lambda\tilde c) + \lambda z\tilde b & (1-x)(1-\lambda\tilde c) - \lambda a z & (1-x)\lambda\tilde b + \lambda a y \\ yc + (1-b)z & (1-x)c + az & (1-x)(1-b) - ay \end{pmatrix} \tag{7.17}$$

where $\tilde b \stackrel{def}{=} b+b'$, $\tilde c = c+c'$. Thus the matrix $\tilde C = N(1-\tilde M)^{-1}$ is:

7.18
$$\tilde C = \frac{1}{\Delta}\begin{pmatrix} y(1-\lambda c')+\lambda z(1+b') & (1-x)(1-\lambda c') & \lambda(1-x)(1+b')\\ (1-x)(1-\lambda c') & a(1-\lambda c') & \lambda a(1+b')\\ \lambda(1-x)(1+b') & \lambda a(1+b') & \lambda^2 a(1+b')Z \end{pmatrix} \tag{7.18}$$

Noting that $\Delta_{00,11} = -\Delta^{-1}(1-\lambda c')$, $\Delta_{00,33} = Z^2\lambda^2\Delta_{00,11}$, $\Delta_{00,13} = Z\lambda\Delta_{00,11}$ and $\tilde C_{11} = \frac{a(1-\lambda c')}{\Delta}$, $\tilde C_{33} = (Z\lambda)^2\tilde C_{11}, \tilde C_{13} = Z\lambda\tilde C_{11}$ with $a = M_{11}$ (not to be confused with a in (7.1)) we get, for a suitable $\beta > 0$, for $\det D$:

7.19
$$\begin{aligned}&\overline\Gamma^{(1)}_{11}\overline\Gamma^{(2)}_{22} + 2\frac{(\Gamma^{(1)}_2 + Z\lambda\Gamma^{(3)}_2)^2}{(1-\lambda c')^{-1}\Delta}\Big(\overline\Gamma_{11}M_{11} - 2(\Gamma^{(0)}_1)^2\Big) + O(\varepsilon^3\eta^{-\beta}e^{-\frac{\pi}{2}\eta^{-1/2}}) =\\ &= \overline\Gamma^{(1)}_{11}\overline\Gamma^{(2)}_{22} + O(\varepsilon^3\eta^{-\beta}e^{-\frac{\pi}{2}\eta^{-1/2}}) = \det D\end{aligned} \tag{7.19}$$

by the argument leading to (6.12), (6.14) and having taken the parameter a in (7.1) equal to $\frac12$ (not to be confused with the $a = M_{11}$ in (7.18)): this completes the analysis of the remarkable cancellations for separatrices splitting in the anisochronous case. The leading order remains *exactly the same* as in the isochronous case: anisochrony only alters the final result by a factor of order $(1+O(\eta^a))$, as it should have been expected *a priori* once understood the isochronous case.

The proof of the domination of the first order now follows the same path as the corresponding in §8 og [G3]: one uses the above results to treat the first $\eta^{-1/2}$ orders of perturbation theory and for the remainder one just use that the series for the splitting is convergent.

§8. Heteroclinic chains.

For completeness we give the argument for the existence of heteroclinic chains, following [CG], *in the easy case* of isochronous systems. Below we imagine to have fixed μ and to take $\eta\to 0$ (so that $\varepsilon = O(\eta^c)$).

It is worth noting that *no gaps* (*i.e. all actions* $\underline A$ are the average position of an invariant torus) are present in the isochronous cases (2.1), which is, therefore, very similar to the original example proposed by Arnold (also gapless).

Let $A_0 = 0 < A_1 < \ldots < A_{\mathcal N} = \overline A_0$ and choose correspondingly $B_0,\ldots,B_{\mathcal N}$ so that the sequence of action variables (A_j,B_j) describes the time averaged location of invariant tori for (2.1) with energy 0 (say).

We consider a perturbation like (5.9) for which the splitting is given by *Melnikov's formula* $\sigma = O(\eta^b e^{-\frac{\pi}{2}\eta^{-1/2}})$ for some $b>0$. Since there are no gaps we can choose the sequence A_i can be chosen so that $A_{i+1}-A_i < e^{-\delta\frac{\pi}{2}\eta^{-1/2}}$ for a prefixed $\delta > 1$ and for all i's. Hence the number $\mathcal N$ has size $O(\overline A_0 e^{+\delta\frac{\pi}{2}\eta^{-1/2}})$.

We want to show that there are heteroclinic intersections H_i between the unstable manifold of the torus $\underline A_i$ and the stable manifold of $\underline A_{i+1}$. Since by construction the tori have the same energy this simply means finding a solution for the equations: $X^-_\uparrow(\pi,\underline\alpha;\underline A_i) - X^+_\uparrow(\pi,\underline\alpha;\underline A_{i+1}) = \underline 0$ (the energy being equal, this equality then implies $X^-_+(\pi,\underline\alpha;\underline A_i) - X^+_+(\pi,\underline\alpha;\underline A_{i+1}) = 0$, *i.e.* also the pendulum momenta match).

The tori equations depend linearly on their average positions, *i.e.* $\underline X^\pm_\uparrow(\pi,\underline\alpha,\underline A) \stackrel{def}{=} \underline A + \underline Y^\pm(\underline\alpha)$ (see theorem 1) where $Y^\pm_\uparrow$ is defined here. We can regard the equation for the heteroclinc intersection $\underline X^-_\uparrow(\pi,\underline\alpha;\underline A_i) - \underline X^+_\uparrow(\pi,\underline\alpha;\underline A_{i+1}) = \underline 0$ as an implicit function problem which for $\underline A_{i+1} = \underline A_i$ has $\underline\alpha_i = \underline 0$ as a solution.

The linearization of the equation at $\underline A_i$ involves the intersection matrix D at $\underline A_i$, $\underline\alpha = \underline 0$ (which in the isochronous case is $\underline A$–independent):

8.1 $$D\,\underline{\alpha}_i=\underline{A}_i-\underline{A}_{i+1} \tag{8.1}$$

showing that the implicit functions problem of determining the heteroclinic point $\underline{\alpha}_i$ can be solved for η small enough because $\det D=O(\eta^b e^{-\frac{\pi}{2}\eta^{-1}})$ and $|\underline{A}_{i+1}-\underline{A}_i|=O(e^{-\delta\frac{\pi}{2}\eta^{-1}})$, with $\delta>1$.

It might be surprising, at first, that the equation for $\underline{\alpha}$ can be solved *without an explicit estimate of the* $\underline{\alpha}$*–derivatives* of the $\underline{Y}_\uparrow(\underline{\alpha})$ at points $\underline{\alpha}$ near $\underline{0}$. Such estimates can be made directly from the existence theorem: however they give bounds on derivatives values that are *much larger* than σ, *i.e.* they have size $O(\eta^{-\beta})$ for some $\beta>0$. This may seem to undermine the foundations of the implicit functions methods, that rely on the solubility of the linear equation.

However the corrections to (8.1) are bounded by $O(\eta^{-\beta}\,\underline{\alpha}^2)$; and $|\underline{A}_{i+1}-\underline{A}_i|\le e^{-\delta\frac{\pi}{2}\eta^{-1/2}}$. The solution of the linear equation (8.1) has size $O(\eta^{-b}e^{-\delta\frac{\pi}{2}\eta^{-1/2}}\cdot e^{+\frac{\pi}{2}\eta^{-1/2}})$. Hence near such $\underline{\alpha}$ the higher order corrections have roughly still size $O(e^{-\delta\pi\eta^{-1/2}})$: much smaller than the linear contribution. This shows that our knowledge of the smoothness of X *suffices*, together with the basic estimate on the homoclinic angles, to deduce that the linear approximation dominates and to claim that the solutions for the heteroclinic point do exist and are very close to those of (8.1).

Therefore there is a chain of heteroclinic points $H_0,H_1,\ldots,H_{\mathcal{N}-1}$ "connecting" a neighborhood of $\underline{A}_0$ to one of $\underline{A}_{\mathcal{N}}$. The "length" of the chain is $\mathcal{N}=O(e^{+\delta\frac{\pi}{2}\eta^{-1/2}})$, *i.e.* in some sense it is the inverse of the splitting.

We now consider the case in which the system in (7.1) is further perturbed by a *monochromatic* perturbation $\beta f_0(\lambda,\varphi)$. The radius of convergence of the whiskers series in β,μ (recall that $\varepsilon=\mu\eta^c$) can be shown to have size of order $|\beta|<O(\eta^{-\frac{1}{2}})$, $|\mu|<O(1)$, see Appendix A10 of [CG].

The splitting $\det D$ is analytic in β,μ for $\beta\ll\eta^{-1/2}$, see Appendix A10 in [CG], and it is different from 0 for $|\beta|\ll O(\eta^c),\mu\ne0$, provided the generically non zero splitting at $\beta=0$ is not zero. In the latter case the splitting can only vanish finitely many times, at μ fixed, in the domain of convergence of the series if $\mu\ne0$.

Hence for all values of β close to 1 (including $\beta=1$ generically in f_0), there exist heteroclinic chains as long as we wish. However not having an estimate for the size of the splitting *we cannot infer from the above argument how long the chains are.*

The remark is interesting in view of the general fact, [A2], that heteroclinic chains imply diffusion in phase space, *i.e.* existence of motions starting near the torus at one end of the chain and reaching the one at the other hand in due (finite) time, see [G5] for the discussion of the method of proof in [CG].

The extension of the above simple analysis is harder in the case of anisochronous systems: it was attempted in [CG] but it is invalid because it relies on the error mentioned above. For a review on diffusion see [Lo1]. It is likely that the study of this problem will require using the techniques of Nekhorossev's theory, see [Lo2] for the best results available, and [CG], Appendix A10.

§9. Other results. Comments.

(A) General theory of splitting.

We call *diophantine* with diophantine constants $C_0,\tau>0$ a vector $\underline{\omega}\in R^\ell$ such that:

9.1 $$|\underline{\omega}\cdot\underline{\nu}|>C_0^{-1}|\underline{\nu}|^{-\tau},\qquad\forall\underline{\nu}\in Z,\ \underline{\nu}\ne\underline{0} \tag{9.1}$$

compare with (2.2) (here there is no extra parameter η). We use the notations of §2 for the other symbols that are not redefined.

Theorem 3: *Suppose that* $\underline{\omega}\in R^{\ell-1}$ *is diophantine and consider the hamiltonian:*

9.2
$$H=\frac{\underline{A}^{\,2}}{2J_1}+\underline{\omega}\cdot\underline{A}+\frac{I^2}{2J_0}+g^2J_0(\cos\varphi-1)+\mu f(\underline{\alpha},\varphi)\tag{9.2}$$

with $(\underline{A},\underline{\alpha})$, (I,φ) *being* ℓ *pairs of canonically conjugated action–angle variables. Let* f *be a even trigonometric polynomial of degree* N *and, for simplicity,* $J_0\le J_1$*. Then:*
(1) the separatrix splitting, for the torus with rotation vector $\underline{\omega}$ *into which the unperturbed torus* $\underline{A}=\underline{0}$ *evolves with* μ*, is analytic in* μ *near* $\mu=0$.
(2) the power series expansion of the splitting $\underline{Q}(\underline{\alpha})$ *in powers of* μ *has coefficients with* 0*–average; their Fourier components* $\underline{Q}^{\,k}_{\underline{\nu}}$ *are bounded, for any* $\delta<1$*, by:*

9.3
$$|\underline{Q}^{\,k}_{\underline{\nu}}|\le\begin{cases}J_0gD\delta^{-\beta}(B\delta^{-\beta})^{k-1}k!^p\varepsilon(k,\underline{\nu}) & \underline{\nu}\ne\underline{0}\\ J_0gD\eta^{-\beta}(B\eta^{-\beta})^{k-1}\end{cases}\tag{9.3}$$

where D,B *are suitable constants and* β,p *can be taken* $\beta=4(N+1)$, $p=4\tau+4$ *if* τ *is the diophantine constant of* $\underline{\omega}$*, and:*

9.4
$$\varepsilon(k,\underline{\nu})=\max_{\substack{h\le k;\,\{\underline{\nu}_j\}_{j=1,\ldots,h}\\ \underline{0}\ne\underline{\nu}'=\sum_{j=1}^h\underline{\nu}_j}}\Big([\prod_{j=1}^h|f_{\underline{\nu}_j}|]e^{-g^{-1}|\underline{\nu}'\cdot\underline{\omega}|(\frac{\pi}{2}-\delta)}\Big)\tag{9.4}$$

Remarks: (1) this theorem is the main result of [G3]; note that (9.4) is stronger than the form in which it is quoted in eq. (6) of [RW] which refers to the theorem stated in [G3] but not to its proof (which gives in fact (9.4)). The method of proof in [G3] could yield in fact the result for f analytic by using the ideas in [GM] but extra work in necessary: see [DGJS],[RW] for alternative proofs; see also [BCG], where the stronger form (with f analytic) was derived, in a similar problem.
(2) There are many instances in which the first order expression (called *Melnikov's integral*) of the splitting vector $\underline{Q}^{\,1}(\underline{\alpha})$, gives in fact the *leading behavior* (as $\mu\to0$) in the calculation of the splitting. In the case of fixed $\underline{\omega}$, *i.e.* for the *one time scale problem*, this follows from the classical results of Melnikov, [Me].
(3) Another interesting question arises when $\underline{\omega}=\gamma\underline{\omega}_0$ with $\underline{\omega}_0$ diophantine and γ a parameter that we let to ∞: this is a *two time scales* problem. In the case $\ell=2$ (hence $\underline{\omega}_0$ is a constant ω_0) and with f a trigonometric polynomial the above theorem proves that the splitting is (generically) $O(e^{-\frac{\pi}{2}g^{-1}\omega_0\gamma})$: in fact this result was the main purpose of the theory in [G3] (see §8 in [G3] and, in particular, (8.6) and the related discussion). It should be stressed that the latter reference simply provides a new proof of a result already obtained, in a slightly different case, by [HMS] or, in the same case, by [Gl], [GLT], LST]. The interest of [G3] lies in the technique.
The Melnikov's "approximation", *i.e.* the dominance of the first order value of the splitting, is more delicate if $\ell\ge3$. The techniques of [G3] are inadequate to deal with this case and they only show that the splitting is smaller than any power of γ^{-1} while the first order value is $O(e^{c\gamma})$ for some (computable) $c>0$: in the case $\ell=3$ this has been studied in [DGJS], [RW], where the (9.2) is improved by replacing $\varepsilon(k,\underline{\nu})$ by the much better $e^{-\frac{\pi}{2}g^{-1}\gamma|\underline{\omega}_0\cdot\underline{\nu}|}$. Several examples of first order dominance are provided in the latter references. However all examples are constructed with f analytic: it would be nice to find a model with a trigonometrical polynomial f for which the first order theory gives the asymptotic result.

(B) Three time scales. Anisochrony strength. Homoclinic scattering.

The three time scales condition for the first order dominance ("Melnikov's formulae") includes the case $a=0$ which is in fact a 2 time scales problem: denoting always $\eta^{-1/2}$

the fast velocity scale, from the analysis of §7 we see that the slow scale could be η^a with $a \ge 0$.

This means that the above theory provides a class of models in which the Melnikov's formula gives the exact asymptotics as $\eta \to 0$ *and the perturbation is a trigonometric polynomial*, of which (5.10) is a concrete example. This does not seem to contradict the results of [RW] who show that the Melnikov's formula does not necessarily give the leading asymptotics as $\eta \to 0$ in cases corresponding to their $n = 3, s = 2$.

In the only almost overlapping case $a = 0$, however, the above question is not treated in [RW] (they consider the *very* different case $n = 3, s = 2$ in which $a = -\frac{1}{2}$, *i.e.* two fast rotators and a pendulum). This illustrates also that there are several "2 time scales problems", depending on which pair among the three time scales is identified.

The value $a = 0$ is a case considered with other techniques in the paper [RW] (it corresponds to their $n = 3, s = 1$): there the attention is dedicated to a wider question (namely the leading order of the splitting *everywhere* on the section $\varphi = \pi$ rather than just at the homoclinic point). Our asymptotic result is consistent with their theorem 2.1. We also get the complete asymptotics in the case of trigonometric polynomial perturbations: but they do not seem interested in this point and deal only with other cases ($n = 3, s = 2$ and non trigonometric peturbations); their technique seems to apply to our (special) case $a = 0$ as well (in fact a simpler case).

The case $a > 0$ is not considered in [RW] except, perhaps, for a remark at the end of the abstract and following eq.(15): we do not know whether this case, that is explicitly excluded in the paper, can be treated with their techniques. In the end the main difference between our work and that of [RW] might just lie in the technique: we have shown that the work in [G3] provides all the necessary technical tools for the analysis of the splitting and even leads to an "exact" expression for it. It is however limited to the splitting at the homoclinic point $\varphi = \pi$, $\underline{\alpha} = \underline{0}$ (unlike [G3], and [RW] where the splitting is measured at $\underline{\alpha}$ arbitrary, on the section $\varphi = \pi$).

The work [RW] is the last in a series of papers (like [DGJS]) which are inextricably linked with each other (and with [G3], [BCG]). The above comments therefore are easily presented in connection with [RW]: but we are fully aware of the role of the other papers quoted in [RW].

Fixing $a = \frac{1}{2}$ the anisochrony coefficient (of A^2) in (7.1) is η^β with $\beta = 2a$. The value $\beta > a$ is necessary if one wants that the anisochronous and the isochronous splittings coincide to leading order as $\eta \to 0$ (at given rotators velocities): however the analysis above does not seem yet sharp enough for such an improvement (*i.e.* taking $\beta < 1$). Finally the physical interpretation of the precession problem (*i.e.* diffusion in presence of a double resonance for a *a priori* stable system) requires $\beta = 1, a = \frac{1}{2}$.

Extensions of the cancellations theory of §7 to $\ell > 3$ seem only a matter of patience. And they would be interesting as they can be conceivably used to treat a variety of systems and one should expect that the results will be quite different when the number of fast scales exceeds 1, as shown in the "maximal case" in which it is $\ell - 1$ ([DGJS], [RW]). However a general theory of *a priori* stable systems, with a free hamiltonian without free parameters and a perturbation consisting of terms of equal order of magnitude *seems to require substantial new ideas.*

In §10 of [CG] there is also a statement about the *homoclinic scattering*: the techniques of this paper apply to its theory as well. We have not worked out, however, the corresponding details (*the statement was not used anywhere in [CG]*) and at the moment it is still an open question for us whether the homoclinic phase shifts are exponentially small or not at the homoclinic point (as claimed in [CG] on the basis of the computational error mentioned above).

Appendix A1: Computation of the pendulum wronskian.

The pendulum hamiltonian: $H = I^2/2J_0 + g_0^2 J_0(\cos\varphi - 1)$ generates a separatrix motion

$t \to \varphi^0(t)$ which is exactly computable. One finds, starting at $\varphi = \pi$ at $t = 0$:

$$\begin{aligned} \sin \varphi^0(t)/2 &= 1/\cosh g_0 t, & \sin \varphi^0(t) &= 2 \sinh g_0 t \, (\cosh g_0 t)^{-2} \\ \cos \varphi^0(t)/2 &= \tanh g_0 t, & \cos \varphi^0(t) &= 1 - 2\, (\cosh g_0 t)^{-2} \end{aligned}$$

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:

A1.1
$$\begin{aligned} & I^0 = \frac{-2g_0 J_0}{\cosh g_0 t} = -2g_0 J_0 \sin \frac{\varphi^0}{2}, \quad \partial_E I^0 = J_0 (I^0)^{-1} \Big(1 + J_0 g_0^2 (\partial_E \varphi^0) \sin \varphi^0 \Big) \\ & \varphi^0 = 4 \,\mathrm{arctg}\, e^{-g_0 t}, \qquad \partial_E \varphi^0 = \frac{-1}{8 g_0^2 J_0} \big(2 g_0 t + \sinh 2 g_0 t\big) \sin \frac{\varphi^0}{2} \end{aligned} \tag{A1.1}$$

exhibiting the analyticity properties in the complex t plane that are useful in discussing the size of the homoclinic angles. The (A1.1) allows us to compute the wronskian matrix of the above separatrices, *i.e.* the solution of the pendulum equation, namely $\dot\varphi = \frac{I}{J_0}$, $\dot I = J_0 g^2 \sin \varphi$, linearized on the separatices:

A1.2
$$\dot W = L(t) W, \quad W(0) = 1, \quad L(t) = \begin{pmatrix} 0 & J_0^{-1} \\ J_0 g_0^2 \cos \varphi^0(t) & 0 \end{pmatrix} \tag{A1.2}$$

and we get:

A1.3
$$W(t) = \begin{pmatrix} \dot\varphi^0/c_2 & \partial_E \varphi^0/c_1 \\ \dot I^0/c_2 & \partial_E I^0/c_1 \end{pmatrix}, \qquad \begin{aligned} c_1 &= \partial_E I^0(0) \\ c_2 &= \dot\varphi^0(0) \end{aligned} \tag{A1.3}$$

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:

A1.4
$$W(t) = \begin{pmatrix} \frac{1}{\cosh g_0 t} & \frac{\overline{w}}{4 J_0 g_0} \\ -J_0 g_0 \frac{\sinh g_0 t}{\cosh^2 g_0 t} & (1 - \frac{\overline{w}}{4} \frac{\sinh g_0 t}{\cosh^2 g_0 t}) \cosh g_0 t \end{pmatrix}, \qquad \overline{w} \equiv \frac{2 g_0 t + \sinh 2 g_0 t}{\cosh g_0 t} \tag{A1.4}$$

The theory of the jacobian elliptic functions shows how to perform a complete calculation of the functions R_0, S_0 in (3.2): see [CG], Appendix A9.

Appendix A2: Convergence of the "form factors" $\underline{G}, \underline{\Gamma}, \overline{\Gamma}$ and remainders bounds.

Integrals for $\underline{\Gamma}, \overline{\Gamma}, M$, see (6.3),(6.7),(6.8) and the analogous ones in §7, are precisely the object of the analysis of Appendix A1 of [G3]. Hence we adhere closely to it.

Consider any of the form factors, *i.e.* any of the series in §6 or §7. Following [G3], word by word, we obtain a bound on the sum of the contributions of the values of all trees ϑ with m nodes and order h, $m \le 2h$, as $D_0 B_0^{h-1} \varepsilon^h m!^2 \max_{0<|\underline{\nu}| \le N h} |\underline{\omega} \cdot \underline{\nu}|^{-4h}$ with D_0, B_0 constants (see (A1.13) and (A1.4) of [G3]: noting that in (A1.4) $m!^2$ is missing due to a typo). This factorial becomes $h!^4$ in the bound (8.2) of [G3] (and consequently on the form factors and $\det D$ at order h, which interest us here and which are expressed essentially by the same integrals) because $m \le 2h$ and it is responsible of the first value 4 in the variable called p, in [G3], $p = 4 + 4\tau$.

We know from theorem 1 above, that $\det D$ is given by a convergent series but we do not know whether the series for $\underline{\Gamma}, \overline{\Gamma}$ and for the matrix M converge. For the "dressed form factors" $\underline{G}$ one could in fact show that the series in ε converge for small ε by the method developed in Appendix A1 of [G3] (consisting in reexpressing the $\underline{G}$ in terms of the unsplit operators $\mathcal{O}$): but a lot of work can be saved because as we show now we do not need to know such convergence properties and the analyticity of $\det D$ suffices.

One can find some improvements over the bounds (8.2) of [G3] by using the intermediate bound (A1.13) in [G3] on the integrals expressing the h-th order coefficient of $\det D$: $\tilde D_1 \tilde B_1^{h-1} \varepsilon^h h!^4 \max_{0<|\underline{\nu}|\le Nh} |\underline{\omega}\cdot\underline{\nu}|^{-4h}$ with $\tilde D_1, \tilde B_1$ constants.

One can then bound small divisors by $(\eta^{-d}(hN)^\tau)^{4h}$, or $\eta^{-4dh}h!^{4\tau}$ via (2.2): which accounts for part of the constant β and for the 4τ in the value of the constant $p = 4\tau+4$ in the second bound in (8.2) of [G3]. It also accounts for Q in the first bound.

In fact one has a better bound as long as $Nh < \eta^{-1}$ because it is only for $Nh \ge 1$ that one can find really small divisors: for $Nh < \gamma\eta^{-1}$, with any prefixed $\gamma < 1$, we can bound $|\underline{\omega}\cdot\underline{\nu}|$ below by $\eta^{1/2}$, if η is small enough. Hence for such values of h we have the bound: $\overline{D}_1 \overline{B}_1^{h-1} (\varepsilon\eta^{-2})^h$ without any factorial, with $\overline{D}_1, \overline{B}_1$ constants and for $h < N^{-1}\eta^{-1}\gamma$. This bound holds for all form factors and also for $\det D$ (note that it does not imply any convergence).

On the other hand we know that $\det D$ is analytic in ε for $\varepsilon < \eta^c$ and c large enough. Hence, to order h, $\det D$ (but *a priori not* individual form factors) can be bounded by:

A2.1
$$D_2 B_2^{h-1} (\varepsilon\eta^{-\overline{c}})^h, \qquad \text{for all } h \ge 1 \tag{A2.1}$$

for some $\overline{c} > 0$. It was conjectured in [G3], see Appendix A1, that this bound could be obtained directly from the graphical expansions. This has been proved in [Ge1] (getting $\overline{c} = 4d$ if d is the diophantine constant in (2.2)); the same proof shows convergence of the series in §6,§7 for the form factors and for M by bounding them, order by order, by (A2.1) (with different constants D_2, B_2): but we are showing that such stronger result is not needed here.

Finally *in the case of* $\Gamma_2^{(0)}, \Gamma_2^{(2)}, \overline{\Gamma}_{i2}$, *i.e.* in the case of the *"bare" or "analytic form factors"* , which are expressed as integrals of analytic functions, one can further improve the bound by the usual τ_v–variables integrations shift to $\operatorname{Im}\tau_v = \pm i(\frac{\pi}{2} - \eta^{1/2})$, choosing the quantity called d in [G3] as $\eta^{1/2}$, a natural but quite arbitrary choice.

One checks directly (as explained in [G3], Appendix A1) that this simply introduces a factor $\eta^{-\beta'}$ wth $\beta' = 2(N+1)$ due to closeness of the singularities of the wronskian functions or of the $f(\varphi(\tau_v), \underline{\omega}\,\tau_v)$ (located at the same places because f is assumed to be a trigonometric polynomial); *it introduces also an exponentially small factor*: $\varepsilon_h = \min_{0<|\underline{\nu}|\le Nh,\, 0<|\nu_2|} e^{-\frac{\pi}{2}|\underline{\omega}\cdot\underline{\nu}|}$ so that the bound of the h–th order, for such form factors (and hence for $\det D$ by §7) becomes :

A2.2
$$D_3\, h!^4\, B_3^{h-1} (\varepsilon\eta^{-\beta})^h \max_{0<|\underline{\nu}|\le Nh, 0<|\nu_2|} e^{-\frac{\pi}{2}|\underline{\omega}\cdot\underline{\nu}|}, \qquad \text{for all } h < \gamma(N\eta)^{-1} \tag{A2.2}$$

with $\beta = 2(N+1)+2$ and B_3, D_3 constants. Note that in order to deduce this all we really need is that the *exact* expressions (6.12) and (7.10), regarded as formal power series, be true order by order in the expansion. We do not really need the (yet unknown) convergence of the series for the Γ's and for the matrix M, which we think should be nevertheless true.

Since $(\varepsilon\eta^{-\overline{c}})^h = (2\varepsilon\eta^{-\overline{c}})^h\, 2^{-h}$ and for η small enough and $h > (N\eta)^{-1}\gamma$ the quantity 2^{-h} is $< e^{-\frac{\pi}{2}\eta^{-1/2}}$ the bounds (A2.2) and (A2.3) can be combined into the bound for the h–th order value of $\det D$:

A2.3
$$D_4 (B_4 \varepsilon\eta^{-c})^h e^{-\frac{\pi}{2}\eta^{-1/2}} \qquad \text{for all } h \ge 1 \tag{A2.3}$$

with $c = \max(2(N+1)+6, \overline{c})$ and B_4, D_4 suitable constants, immediate consequence of (A2.2) (and $|\underline{\omega}\cdot\underline{\nu}| \ge \eta^{-1/2} - Nh\eta^{1/2}$ for $\nu_2 \ne 0$).

Acknowledgements: We are indebted to P. Lochak for many discussions and for encouraging one of us to revise the previous work [CG] in order to present a simplified version. One of us (GiG) is deeply indebted to V. Gelfreich for pointing out, in a meeting organized and led by P. Lochak, the error in [CG] that is corrected in the present paper. This work is part of the research program of the European Network on: "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460.

References

[A1] Arnold, V.: *Proof of a A.N. Kolmogorov theorem on conservation of conditionally periodic motions under small perturbations of the hamiltonian function*, Uspeki Matematicheskii Nauk, 18, 13– 40, 1963.

[A2] 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.

[BCG] Benettin, G., Carati, A.: *A rigorous implementation of the Jeans–Landau–Teller approximation for adiabatic invariants*, Nonlinearity, **10**, 479–507, 1997.

[BGGM] F. Bonetto, G. Gentile, Gallavotti, G., V. Mastropietro: *Lindstedt series, ultraviolet divergences and Moser's theorem*, mp_arc@ math. utexas. edu, #95-506, 1995. And *Quasi linear flows on tori: regularity of their linearization*, in mp_arc@ math. utexas. edu, #96-251,

[CG] Chierchia, L., Gallavotti, G.: *Drift and diffusion in phase space*, Annales de l' Institut Poincarè, B, **60**, 1–144, 1994.

[DGJS] S. Delshams, V. Gelfreich, A. Jorba, T.M. Seara: *Exponentially small splitting of separatrices under fast quasiperiodic forcing*, Preprint, 1996.

[E] Eliasson, L.H.: *Absolutely convergent series expansions for quasi-periodic motions*, Mathematical Physics Elctronic Journal, MPJE, **2**, 1996.

[G1] Gallavotti, G.: *The elements of Mechanics*, Springer, 1983. See also Gallavotti, G.: *Quasi integrable mechanical systems*, Les Houches, XLIII (1984), vol. II, p. 539– 624, Ed. K. Osterwalder, R. Stora, North Holland, 1986.

[G2] Gallavotti, G.: *Twistless KAM tori*, Communications in Mathematical Physics **164**, 145–156, (1994).

[G3] Gallavotti, G.: *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable hamiltonian systems. A review*, Reviews on Mathematical Physics, **6**, 343– 411, 1994.

[G4] Gallavotti, G.: *Methods in the theory of quasi periodic motions*, expanded version of a talk at the Conference in honor of Lax and Nirenberg, Venezia, june 1996, in stampa, mp_arc@math. utexas.edu #96-498.

[G5] Gallavotti, G.: *Arnold's diffusion in isochronous systems*, preprint 1997; and Gallavotti, G., Gentile, G., Mastropietro, V.: *Arnold's diffusion in anisochronous systems*, in preparation.

[Ge1] Gentile, G.: *A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable systems*, Forum Mathematicum, **7**, 709–753, 1995.

[Ge2] Gentile, G.: *Whiskered tori with prefixed frequencies and Lyapunov spectrum*, Dynamics and Stability of Systems, **10**, 269–308, (1995).

[Gl] Gelfreich, V. G.: *Melnikov method and exponentially small splitting of separatrices*, Physica D, **101**, 227–248, 1996.

[Gr] Graff, S. M.: *On the conservation for hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations 15, 1-69, 1974.

[GG] G. Gallavotti, G. Gentile: *Majorant series convergence for twistless KAM tori*, Ergodic theory and dynamical systems, **15**, 857–869, (1995).

[GM] G. Gentile, V. Mastropietro: KAM theorem revisited, *Physica D* **90**, 225–234 (1996). And *Tree expansion and multiscale analysis for KAM tori*, Nonlinearity, **8**, 1159–1178 (1995). And G. Gentile, V. Mastropietro: Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications, *Reviews in Mathematical Physics*, **8**, 393–444 (1996).

[GR] Gradshteyn, I.S., Ryzhik, I.M.: *Table of integrals series and products*, Academic Press, 1965.

[GGM] G. Gallavotti, G. Gentile, V. Matropietro: *Field theory and KAM tori*, Mathematical Physics Electronic Journal, **1**, (1995).

[GLT] Gelfreich, V. G., Lazutkin, V.F., Tabanov, M.B.: *Exponentially small splitting in Hamiltonian systems*, Chaos, 1 (2), 1991.

[HMS] Holmes, P., Marsden, J., Scheurle,J: *Exponentially small splittings of separatrices with applications to KAM theory and degenerate bifurcations*, Contemporary Mathematics, **81**, 213–244, 1989.

LST] Lazutkin, V.F., Schachmannski, I.G., Tabanov, M.B.: *Splitting of separatrices for standard and semistandard mappings*, Physica D, **40**, 235–248, 1989.

[Lo1] Lochak, P.: *Arnold's diffusion: a compendium of remarks and questions*, Proceedings of 3DHAM, s'Agaro, 1995.

[Lo2] Lochak, P.: *Stability of hamiltonian systems over exponentially long times: the near linear case*, IMA vol. 63, "Hamiltonian dynamical systems, History, Theory and Applications, 221–229, 1995. And *Canonical perturbation theory via simoultaneous approximation*, Russian Mathematical Surveys, **47**, 57–133, 1992.

[Me] Melnikov, V.K.: *On the stability of the center for time periodic perturbations*, Trans. Moscow Math Math. Soc., 12, 1-57, 1963.

[Pö] Pöschel, J.: *Integrability of Hamiltonian systems on Cantor sets*, Communications Pure Appl. Math, 35, 653-696, 1982.

[RW] Rudnev, M., Wiggins, S.: *Existence of exponentially small sepratrix splittings and homoclinic connections between whiskered tori in weakly hyperbolic near integrable hamiltonian systems*, Preprint, january 1997.

[T] Thirring, W.: *Course in Mathematical Physics*, vol. 1, p. 133, Springer, Wien, 1983.

Internet: Authors' preprints downloadable (latest version) at:
`http://chimera.roma1.infn.it`
`http://www.math.rutgers.edu/∼giovanni`
Mathematical Physics Preprints (mirror) pages.
e-mail: giovanni@ipparco.roma1.infn.it, gentileg@ipparco.roma1.infn.it, vieri@@ipparco.roma1.infn.it

Addresses:
G.Ga.: Dipartimento di Fisica, Università di Roma 1, P.le Moro 2, 00185, Italy
G.Ge.: Dipartimento di Matematica, Università di Roma 3, Largo S. Leonardo Murialdo 1, 00146, Roma, Italy
V.Ma.: Dipartimento di Matematica, Università di Roma 2, Viale Ricerca Scientifica, 00133, Roma, Italy
Paper archived also in chao-dyn@xyz.lanl.gov 9709004