Compensations in small divisor problems

L. Chierchia and C. Falcolini

Dipartimento di Matematica

Università di Roma "Tor Vergata"

via della Ricerca Scientifica, 00133 Roma (Italy)

(Internet: chierchia@mat.utovrm.it and falcolini@mat.utovrm.it) *

August 1994

Abstract

Several small divisor problems arising in the perturbative theory of Hamiltonian and Lagrangian systems are considered. A general method that allows to prove compensations among the elementary contributions of the formal power series expansions associated to invariant surfaces is presented.

Contents

*The authors gratefully acknowledge helpful discussions with C. Liverani.

1 Introduction

C.L. Siegel [20] was the first to solve a small divisor problem (linearization of germs of analytic functions). His solution was based on a "direct method" consisting in computing the k^{th} coefficient of the formal power series solution and estimating such coefficient by a constant to the k^{th} power. The solution of small divisor problems arising in perturbative series in the theory of Hamiltonian systems has been one of the major contributions to dynamical systems of this century: the techniques employed for such a task, started by A.N. Kolmogorov in the early fifties and developed in the early sixties by V.I. Arnold and J.K. Moser, are known as "Kolmogorov–Arnold–Moser (KAM) theory" (see *e.g.* [1] and references therein). Such methods show *indirectly* the convergence of the formal expansions of the perturbative series (longly before considered by astronomers and especially by H. Poincarè [19]).

In 1988 H. Eliasson proposed a direct proof of the convergence of formal expansions in the Hamiltonian context ([9], [10], [11]). More recently direct proofs (in the Hamiltonian context) have been reconsidered in several papers: [6], [12], [13], [14] [15], [16], [17]; a small divisor problem arising in elliptic systems of PDE's has been solved by similar methods in [7].

In this paper we consider several small divisor problems and discuss a general method that allows to prove compensations among "elementary" contributions to the k^{th} order of the associated formal expansion. Such compensations, modulo technical estimates (which have been thoroughly discussed in [6]), yield new proofs of the *convergence* of the formal power series.

The starting point in this subject is a formal power series whose coefficients are *bona fide* functions determined by linear recursive equations. "Recursive" means that the k^{th} coefficient can be expressed (typically by Taylor's expansions and by inversion of suitable differential operator D) in terms of the derivatives of a given function ("the Hamiltonian" or "the Lagrangian") and in terms of the h^{th} coefficients with $h < k$. If one keeps on expressing each coefficient in terms of the preceding ones, one will end up with a finite (but huge) decomposition of the k^{th} coefficient involving only the derivatives of the given function or, more precisely, involving simple operations, coming from inverting D, on such derivatives; the inversion of D brings in the small divisors. We shall call the terms obtained through this procedure "elementary contributions". In writing out the elementary contributions in the most explicit way, it appears readily clear that a graphical representation is of big help. On a more formal level this is done using graph theory (more specifically the theory of labeled, rooted trees). Estimating the k^{th} coefficient by inserting absolute values within the sum of the elementary contributions leads, in general (and inevitably), to divergences (*i.e.* to estimates of the k^{th} coefficients by a power of the factorial $k!$). But if the formal series were convergent (and in many examples one *knows* by KAM that this is the case) it would mean that it is possible to group together the elementary contributions in such a way that huge (size $k!$) terms *compensate* among themselves leaving out well behaving (exponential in k) quantities. Graph theory provides a very useful tool in order to recognizes the families of elementary contributions that compensate. Most of the known small divisor problems and some new ones[1] can be solved by this approach.

The small divisor problems we shall consider in this paper (always in the real–analytic category) are: P1) maximal quasi–periodic solutions (or "invariant tori") for nearly–integrable non–degenerate Hamiltonian systems; P2) maximal quasi–periodic solutions for nearly–integrable non–degenerate Lagrangians; P3) lower dimensional hyperbolic (whiskered) tori for "second order" nearly–integrable Hamiltonian systems. We shall also consider a small divisor problem, P4), (lower dimensional "indefinite" tori for second order Hamiltonian systems) which is "non compensable" (in the precise sense described below), strongly supporting that the associated formal power series are divergent; in fact, it would be nice to give a direct proof of the divergence of such series.

Our approach to P1) (integrated by the estimates of [6]) yields a direct proof of the classical "KAM theorem" alternative to that of Eliasson [9]; the direct proofs in the context of P2) and P3) are new[2].

In § 2 we give a precise description of the models P1) ÷ P4) and of the results obtained. In § 3 a new family of trees (simply related to the standard labeled, rooted trees) is introduced. In § 4, § 5, § 6 the compensation mechanism for the models P1), P2), P3) is discussed. Finally, in § 7 we exhibit an example where no compensations (of the type described in this paper) take place.

2 Notations and Results

Let us describe more formally what we mean by compensations in small divisor problems. Given a "nearly–integrable" dynamical system (say, as in P1) ÷ P3)), one is interested in finding invariant surfaces run by a linear flow or, equivalently, in finding quasi–periodic solutions. Quasi–periodic solutions are described in terms of functions on the standard N–dimensional torus $^N \equiv {}^N/(2\pi^N)$ substituting the phase variable $\theta \in {}^N$ with $\omega t = (\omega_1 t, ..., \omega_N t)$ where $\omega \in {}^N$ is a given (rationally independent) vector and t denotes time[3]. The system being "nearly–integrable" means that there is a perturbative parameter, say ε, such that for $\varepsilon = 0$ the system is completely "solvable". It is therefore natural to try to establish the existence of formal solutions: this constitutes the *first step*. Formal quasi–periodic solutions have been considered already in the last century; for a modern discussion of the formal solutions of the models discussed here we refer the reader to Appendix B of [6][4].

[1] e.g. the elliptic systems considered in [7].

[2] The KAM theory in the context of P2) is relatively recent and is sometimes referred to as "KAM in configuration space" (see [21], [5]). For the KAM theory in the context of P3) see [18]. The small divisor problem P3) is particularly relevant in connection with the so–called Arnold diffusion (see [2], [8], [22]).

[3] In the PDE case t is a multidimensional independent variable and ω is a matrix; see [7].

[4] In [6] the formal expansion for P1) is proved; formal solutions for P2) ÷ P4) can be proved in a completely similar way.

The formal solutions we will be dealing with have the form

$$Z \equiv (Z_1, ..., Z_d) \sim \sum_{k \geq 0} Z^k(\theta)\varepsilon^k \ , \qquad \theta \in {}^N \tag{2.1}$$

where, for each $k \geq 1$, the vector–valued function Z^k is a real–analytic function over ^{N} having an exponentially fast converging Fourier expansion of the form[5]

$$Z^k = \sum_{n \in {}^N} Z_n^k e^{in\cdot\theta} \ . \tag{2.2}$$

while for $k = 0$ Z^0 is the solution of the unperturbed problem.
The *second step* is to give Z a representation in terms of *trees i.e.* a representation of the form:

$$Z_n^k = \frac{1}{k!} \sum_{T \in \mathcal{T}_*^k} \sum_{\substack{\alpha: V \to {}^N \\ \Sigma\alpha_v = n}} \sum_{\beta: V \to B} \Lambda(T, \alpha, \beta) \prod_{v \in V} \gamma_v(T, \alpha, \beta) \ , \tag{2.3}$$

where, intuitively speaking, $\mathcal{T}_*^k$ is a suitable family of trees of "order k" taking care of the combinatorics coming out of Taylor's formula and its repeated applications (see § 3); the sum over the indices α attached to each vertex comes from expanding everything in terms of Fourier series (hence the constrain of the total sum of such indices to be n since on the left hand side we have the n–Fourier coefficient of the solution); the final sum is finite and encodes all possible (case–by–case depending) indices which may help in writing out "in the most explicit way" the solution[6]; the addenda of these sums are split into two terms: $\Lambda(T, \alpha, \beta)$, which are complex vectors related to the derivatives of the given function (the Hamiltonian or the Lagrangian) ruling the evolution of the dynamical system and the products of γ_v's which are the small divisors. More formally: $\mathcal{T}_*^k$ is a suitable family of labeled rooted trees[7], $V \equiv V(T)$ denotes the set of vertices of T and the suffix ("order") k refers to the following estimates on cardinalities

$$\# \ \mathcal{T}_*^k \leq k! \ c_1^k \ , \qquad \# \ V(T) \leq c_2 \ k \ , \qquad (\forall \ T \in \mathcal{T}_*^k) \ , \tag{2.4}$$

(for a suitable constant $c_i > 0$); the second sum runs over all possible functions assigning to each vertex $v \in V$ of a rooted tree T an integer vector $\alpha_v \in {}^N$ with the constrain $\sum_{v \in V} \alpha_v = n$; the third sum runs over a suitable set of (possibly) vector–valued indices taking a finite number of values ($\#B < \infty$); Λ is a complex vector depending on T, $\{\alpha_v\}_{v \in V}$, $\{\beta_v\}_{v \in V}$ and on the Hamiltonian (or Lagrangian); finally $\gamma_v \in$ are *divisors* that are described as follows. Rooted trees can be naturally equipped with a partial order: we say that $v' \leq v$ if the path joining the root r of T with v' contains v; (obviously $v' < v$

[5] Note that we are using the suffix k with different meaning as ε^k denotes the k^{th} power of the number ε while Z^k is a vector valued function and k is used as an index. The h^{th} power of the j^{th} component of Z^k will be denoted $(Z_j^k)^h$; the j^{th} component of the Fourier coefficients of Z^k will be denoted Z_{nj}^k.

[6] *e.g.* β_v will typically be related to the degree of the small divisors, see (2.7) below.

[7] See any introductory book on graph theory, such as [3], for the standard terminology or Appendix A in [6].

means $v' \leq v$ and $v' \neq v$ and $r \geq v \ \forall \ v \in V$ *i.e.* the root r is the first vertex of the rooted tree T). Given a function α, we define

$$\delta_v \equiv \delta_v(T,\alpha) \equiv \sum_{v' \in V:\ v' \leq v} \alpha_v \ . \tag{2.5}$$

The divisors γ_v, which may assume arbitrarily small values, are defined in terms of a real–valued function $\langle \cdot \rangle$ which satisfies the *Diophantine condition*

$$|\langle n \rangle|^{-1} \leq \gamma |n|^{\tau} \ , \qquad \forall \ n \in {}^N \backslash \{0\} \tag{2.6}$$

with suitable positive constants γ, τ.[8] Then

$$\gamma_v \equiv \gamma_v(T,\alpha,\beta) \equiv \begin{cases} \langle \delta_v \rangle^{-\sigma_v} & \text{if } \delta_v \neq 0 \\ 1 & \text{if } \delta_v = 0. \end{cases} \tag{2.7}$$

where σ_v is, say, the first component of the index β_v and takes value 0, 1 or 2. In (2.3) only the second sum runs over an infinite set of indices: therefore we assume that there exist positive numbers $\xi, \xi', a > 0$ such that, if we set

$$a_n^k \equiv \frac{1}{k!} \sum_{T \in \mathcal{T}_*^k} \sum_{\substack{\alpha : V \to {}^N \\ \Sigma \alpha_v = n}} \sum_{\beta : V \to B} |\Lambda(T,\alpha,\beta)| \prod_{v \in V} e^{\xi' |\alpha_v|} \ , \tag{2.8}$$

then, for all k and n one has

$$a_n^k \leq a^k e^{-\xi |n|} \ . \tag{2.9}$$

In the models considered here, such an assumption is an immediate consequence of the well–posedness of the (formal) problem and of the analyticity assumptions on the Hamiltonian (Lagrangian). ¿From (2.8) it follows at once that if one could bound the product of the divisors as[9]

$$\prod_{v \in T} |\gamma_v| \leq c_3^k \ \prod_T (1 + |\alpha_v|^b) \tag{2.10}$$

for some $c_3 > 1$ and $b > 0$, then from (2.4), (2.8) and (2.10) it would follow

$$|Z_n^k| \leq \frac{1}{k!} \sum_{T \in \mathcal{T}_*^k} \sum_{\substack{\alpha : V \to {}^N \\ \Sigma \alpha_v = n}} \sum_{\beta : V \to B} |\Lambda(T,\alpha,\beta)| \ c_4^k \ \prod_{v \in V} (1 + |\alpha_v|^b) \leq c_5^k e^{-\xi |n|}, \tag{2.11}$$

(for suitable $c_1 > 0$) leading to "absolute" convergence (*i.e.* convergence without compensations) of the formal expansion Z. Indeed Siegel's original proof is based on a similar

[8] Typically, in dynamical systems, $\langle n \rangle = \omega \cdot n$ (where the dot denotes the standard inner product in ^{N}) but in other situations (*e.g.* [7]) the function $\langle \cdot \rangle$ might be more complicate (*e.g.* non linear); in the case $\langle n \rangle = \omega \cdot n$ it is well known that, if $\tau > N - 1$, up to a set of Lebesgue measure zero, all $\omega \in {}^N$ satisfy (2.6) for some γ.

[9] From now on we will adhere to the the common abuse of notation $v \in T$ in place of the more proper $v \in V(T)$ and also if $vv' = v'v$ denotes an edge of T, we shall denote $vv' \in T$ rather than the more proper $vv' \in E(T)$.

argument, even though the set up is slightly different (and simpler). Technically Siegel's problem corresponds to θ varying in a small (complex) ball so that Fourier series is replaced by Taylor series and α_v ranges over ${}^N_+$: in such a case $\delta_v \neq \delta_{v'}$ whenever $v > v'$ and Siegel's method [20] yields the estimates (2.10)[10]. The problem with $\alpha_v \in {}^N$ is that one can have "resonances" *i.e.* $\delta_v = \delta_{v'}$ for $v > v'$ which may lead to obstinate repetitions of particularly small divisors. It is well known (see *e.g.* [6], Appendix B) that, in general, one has, for arbitrarily large k, subfamilies $\mathcal{F}_{\rm div} \subset \mathcal{T}^k_*$ and a choice of $\bar\alpha$ and $\bar\beta$ (depending only on the subfamily) such that, for suitable $\bar a, \bar b > 0$,

$$\frac{1}{k!}\sum_{T\in\mathcal{F}_{\rm div}} \Lambda(T,\bar\alpha,\bar\beta) \quad \prod_{v\in V}\gamma_v \geq \bar a^k k!^{\bar b} \ . \tag{2.12}$$

Such families are obtained by taking *chains of resonances* which are defined as follows. Given $T \in \mathcal{T}^k_*$ and α (*i.e.* $\{\alpha_v\}_{v\in T}$), a *resonance* is a subtree[11] $R \subset T$ such that: i) R is of degree two (*i.e.* R is connected to $T\backslash R$ by two edges); ii) if u is the first vertex[12] in R and z is the first vertex following R, then $\delta_u = \delta_z \neq 0$; iii) R cannot be disconnected, by removal of one edge, into two subtrees of degree two satisfying i) and ii). It will be important to consider different choices of the index β (*i.e.* $\{\beta_v\}_{v\in T}$): in particular given a resonance R and given β we call *order of the resonance* the number (see (2.7)) $\sigma \equiv \sigma_u$ (u being the first vertex of R). A *chain of resonances* is a maximal series of resonances $R_1, ..., R_h$ with R_i adjacent[13] to R_{i+1}; given a choice of β, the *order of the chain* is defined to be $\bar\sigma \equiv \sigma_1 + \cdots + \sigma_h$ where $\sigma_i \equiv \sigma_{u_i}$, u_i being the first vertex of R_i. ¿From these positions it follows that if C is a chain of order $\bar\sigma$, if $n = \delta_z$ where z is the first vertex following the chain (*i.e.* following the last resonance, which by convention will be R_h), then

$$\prod_{v\in C}\gamma_v = \langle n\rangle^{-\bar\sigma} \prod_{\substack{v\in C\\ v\neq u_i}}\gamma_v \tag{2.13}$$

(where $v \in C$ means $v \in \bigcup_i V(R_i)$). The examples for which (2.12) holds are based on chains with $\bar\sigma \sim k$ and $|\langle n\rangle| \sim k^{-1}$. This phenomenon may be cured through compensations. To be more precise we introduce the notion of "compensable chain". Consider a chain $C = (R_1, ..., R_h)$, $(h \geq 1)$ and fix the indices β_v. Let, as above, u_i be the first vertex in R_i, let R_i be connected to R_{i+1} by the edge $w_i u_{i+1}$ with $w_i \in R_i$ and

[10] For a detailed discussion, in the present language, of Siegel's methods see Appendix C of [6]; for a different approach see [4].

[11] When referred to trees the notation $T' \subset T$ will always mean "T' (unrooted) subtree of T". Other special conventions we are adopting are the following: a) for rooted trees the root (usually denoted r) may be identified by adding an extra edge ηr where η is a symbol (not a vertex of the tree) sometimes called the "earth"; with these positions one has, for trees, $\#E = \#V - 1$ and, for rooted trees, $\#E = \#V$; b) the degree of a vertex v is the number of edges vv' incident with v and if $v = r$ is the root, the edge ηr is included in the count; c) the degree of a subtree $T' \subset T$ is the number of edges connecting T' with $T\backslash T'$; if T is rooted and the root r belongs to T' the edge ηr must be included in the count.

[12] Recall that T is a rooted tree and hence partially ordered (the order being such that the first vertex of T is always its root).

[13] *i.e.* connected by one edge.

$u_1 \geq w_1 > u_2 \geq ... \geq w_h$; let z be the first vertex following the chain (*i.e.* following w_h) and $n \equiv \delta_z$; and, finally, let P_i be the path joining u_i with w_i. Consider the following function of $x \in$

$$\pi_C(x; T, \alpha, \beta) \equiv \prod_{i=1}^{h} \pi_{R_i}(x; T, \alpha, \beta) \ , \qquad \pi_{R_i}(x) \equiv \prod_{v \in (R_i \backslash P_i)} \gamma_v \prod_{\substack{v \in P_i \\ v \neq u_i}} \bar{\gamma}_v(x) \ , \tag{2.14}$$

where, if $u_i = w_i$ (*i.e.* $P_i = \{u_i\}$), the product over P_i is absent, while if $v \in P_i \neq \{u_i\}$ we set

$$\bar{\gamma}_v(x) \equiv \langle \sum_{\substack{v' \in R_i \\ v' \leq v}} \alpha_{v'} + xn \rangle^{-\sigma_v} \ . \tag{2.15}$$

Thus, $\gamma_v = \bar{\gamma}_v(1)$ and

$$\prod_{v \in C} \gamma_v = \langle n \rangle^{-\bar{\sigma}} \pi_C(1) \ . \tag{2.16}$$

We say that *a chain* $C \subset T \in \mathcal{T}_*^k$ (*i.e.* $C = (R_1, ..., R_h)$ with $R_i \subset T$) *is compensable* if there exists a family of trees $\mathcal{F}_C \subset \mathcal{T}_*^k$ whose elements T' have C as common chain of resonance, and, for each T', there exists a choice of indices $\beta' \equiv \beta'(T')$, such that the function[14]

$$\bar{\pi}_C(x) \equiv \sum_{T' \in \mathcal{F}_C} \Lambda(T', \alpha, \beta') \ \pi_C(x; T', \alpha, \beta') \tag{2.17}$$

has a zero in x of order at least $\bar{\sigma} - 1$. ¿From the detailed estimates in [6], it follows easily that *if in a small divisor problem all chains of resonances are compensable then the associated formal series* (2.1) *is in fact convergent*. In the rest of the paper we shall prove the following statements.

P1) Let $H = H_0(y) + \varepsilon H_1(x, y)$ be a real–analytic Hamiltonian with $y \in B(y_0) \subset {}^N$, $x \in {}^N$ and (x, y) standard symplectic coordinates[15]; here $B(y_0)$ denotes some N–sphere centered at y_0. Assume the following standard "non degeneracy conditions" on the "integrable part" H_0: i) the Hessian matrix $\partial_y^2 H_0(y_0)$ is invertible; ii) $\omega \equiv \partial_y H_0 \in {}^N$ is a Diophantine vector *i.e.* there exists $\gamma, \tau > 0$ such that[16]

$$|\omega \cdot n|^{-1} \leq \gamma |n|^\tau \ , \qquad \forall \ n \in {}^N \backslash \{0\} \ . \tag{2.18}$$

Then, quasi–periodic solutions with frequencies ω, *i.e.* solutions of the form $(x(t), y(t)) = Z(\omega t)$ with $Z^0(\theta) = (\theta, y_0)$ and $Z^k : \theta \in {}^N \to Z^k(\theta) \in {}^{2N}$, satisfy the equations

$$DZ = J \partial H\Big(Z(\theta)\Big) \ , \tag{2.19}$$

[14]In other words, the elements T' of $\mathcal{F}_C$ are obtained from T and C by (possibly) changing the edges connecting $T \backslash C$ with R_1, R_1 with R_2,...,R_h with $T \backslash C$, and by (possibly) changing the values of the indices β on C.

[15]This simply means that the Hamiltonian (evolution) equations are given by $\dot{x} = \partial_y H$, $\dot{y} = -\partial_x H$.

[16]Compare with (2.6).

where $D \equiv \omega \cdot \partial_\theta$, J is the standard symplectic matrix $\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ and ∂ is the gradient $\partial_{(x,y)}$ with respect to the variables (x, y). Then it is well known[17] that there exists a unique formal solutions $Z \sim \sum_{k\geq 0} \varepsilon^k Z^k$ of (2.19) with the normalization condition

$$\int_{{}^N} \pi_1 \circ Z^k d\theta = 0 \quad (k \geq 1), \tag{2.20}$$

where π_1 is the projection onto the first coordinates: $\pi_1(x, y) = x$. We can then prove

Theorem 2.1 *There exists a tree expansion* (2.3) *for* Z *such that all chains of resonances are compensable.*

P2) The formulation for the Lagrangian case is very similar. We let $L \equiv L_0(y) + \varepsilon L_1(x, y)$ be a real–analytic Lagrangian[18] with L_i satisfying the same hypotheses assumed above for H_i and ω, except that now $\omega \equiv y_0$. Quasi–periodic solutions $x(t) = Z(\omega t) = \sum_{k\geq 0} \varepsilon^k Z^k(\omega t)$, where now $Z^0(\theta) = \theta$ and $Z^k : \theta \in {}^N \to Z^k(\theta) \in {}^N$, satisfy the equations

$$D\, \partial_y L(Z(\theta), DZ(\theta)) = \partial_x L(Z(\theta), DZ(\theta)) \ , \qquad (D \equiv \omega \cdot \partial_\theta) \ . \tag{2.21}$$

Mimicking the proof of Appendix B in [6] one can show that there exists a unique formal solution $Z \sim \sum_{k\geq 0} \varepsilon^k Z^k$ of (2.21) with the normalization condition as in (2.20) but without the projection π_1. Then, *Theorem 2.1 holds also in this case.*

P3) Maximal invariant tori correspond to analytic continuation (in ε) of unperturbed tori having "all frequencies excited" *i.e.* tori run by linear flow $t \to \omega t$ with $\omega \equiv \partial_y H_0(y_0)$ rationally independent over ^{N} and $N = \#$ of degrees of freedom. More delicate is what happens to unperturbed "resonant" tori[19] *i.e.* invariant unperturbed tori for which there exist $n \in {}^N \backslash \{0\}$ such that $\omega \cdot n = 0$. We shall argue (see next item P4)) that *in general, such tori are not analytically continuable for* $\varepsilon \neq 0$. Nevertheless we will show, under suitable conditions, how to construct, with the (intrinsically analytic) methods outlined above, lower dimensional invariant (unstable) tori when $\varepsilon \neq 0$. For simplicity, we shall discuss only a special case of "second order" nearly integrable Hamiltonian systems with Hamiltonian function

$$H \equiv \frac{1}{2} y^2 + \frac{1}{2} p^2 + \varepsilon f(x, q) \tag{2.22}$$

where (x, y) are symplectic variables as above (*i.e.* with $x \in {}^N$) and so are (q, p) with $q \in {}^M$; hence the number of degrees of freedom is $N + M$ and we shall study N–dimensional invariant tori. The specification "second order" is due to the particular form in which can be cast the Hamilton equations for H:

$$\ddot{x} = -\varepsilon \partial_x f \ , \qquad \ddot{q} = -\varepsilon \partial_q f \ . \tag{2.23}$$

[17] See [19] or Appendix B of [6].

[18] The Lagrangian (evolution equations) are $\frac{d}{dt} \partial_y L(x, \dot{x}) = \partial_x L(x, \dot{x})$.

[19] Here, the word "resonant" which is related to the "resonances" of celestial mechanics, is used in a technically different meaning from the one used throughout this paper.

For $\varepsilon = 0$, N–dimensional invariant tori (up to a trivial linear and symplectic change of coordinates[20]) are spanned by solutions of the form $(y, p) = (\omega, 0)$, $(x, q) = (x_0 + \omega t, q_0)$. ¿From classical transformation theory it follows that (if ω satisfies (2.18)) the evolution equations of H in (2.22) are equivalent to the evolution equations for a Hamiltonian of the form $\frac{y^2}{2} + \frac{p^2}{2} + \varepsilon f_0(q) + o(\varepsilon)$ where f_0 is the average (w.r.t. x) over ^{N} of f. Motivated by this observation, we consider the Hamiltonian

$$\frac{1}{2}y^2 + \frac{1}{2}p^2 + \varepsilon f_0(q) + \varepsilon^2 \tilde{f}(x, q) \equiv H_0(q, y, p; \varepsilon) + H_1(x, q; \varepsilon) \ , \qquad H_1 \equiv \varepsilon^2 \tilde{f} \ . \quad (2.24)$$

Even though H_0 is not, in general, integrable, *if q_0 is a critical point for f_0*, H_0 still admits the invariant N–torus $\mathcal{T}_0$ spanned by $(y, p) = (\omega, 0)$, $q = q_0$ and $x = x_0 + \omega t$ and we want to study the persistence of such torus for the full Hamiltonian. To attack the problem perturbatively, we introduce a new analyticity parameter μ with respect to which we shall make formal (and eventually convergent with radius of convergence greater than one) power series expansion and consider the Hamiltonian $H_0 + \mu H_1$ (so that for $\mu = 1$ we recover the Hamiltonian (2.24)). We also make the following *hyperbolicity assumption*: we assume that q_0 is such that the matrix $\varepsilon \partial_q^2 f_0(q_0)$ is *positive definite*. Under these hypotheses it is easy to prove that there exists a (unique) formal expansion

$$Z \equiv (Z_1, ..., Z_d) \sim \sum_{k \geq 0} Z^k(\theta; \varepsilon)\mu^k \ , \qquad \theta \in {}^N \ , \qquad d = 2(N + M) \quad (2.25)$$

such that $t \to Z(\omega t)$ is a formal quasi–periodic solution of the Hamiltonian equations governed by $H_0 + \mu H_1$ and such that the set $\{Z^0(\theta) : \theta \in {}^N\}$ coincides with the unperturbed torus $\mathcal{T}_0$ introduced above. Uniqueness is achieved by requiring (2.20) where π_1 denotes the projection onto the x variable. Then, *Theorem 2.1 holds also in this case*. ¿From this result, as already remarked, it follows that the formal power series is actually convergent but, what is more interesting in this case, one can show that for $\varepsilon \neq 0$ small enough, *the radius of convergence (in μ) is greater than one* so that the set $\{Z(\theta) : \theta \in {}^N, \mu = 1\}$ is an invariant N–torus for the Hamiltonian (2.24). We finally mention that from [18] it follows easily that these tori are *whiskered* in the sense of [2].

P4) Consider again the Hamiltonian (2.22). Indeed, one can consider formal power series in ε and one can show that *if q_0 is a non degenerate critical point of the x–average of f then there exists a (unique) formal power series* (2.1) *(with $d = 2(N + M)$) such that $t \to Z(\omega t)$ (ω satisfying* (2.18)*) is a formal quasi–periodic solution for* (2.22) *and the set $\{Z^0(\theta) : \theta \in {}^N\}$ coincide with the torus spanned by $y = \omega$, $p = 0$, $q = q_0$, $x = x_0 + \omega t$.* Uniqueness, again, is enforced by the requirement (2.20). Also for such a formal series one can write down a tree expansion completely analogous to those referred to in P1)÷P3) above. However, we shall prove that there exist chains $C = (R_1, ..., R_h)$ such that if $\mathcal{F}_0$ denotes the family of *all* trees with chain C then

$$\sum_{T' \in \mathcal{F}_0} \Lambda(T') \ \pi_C(0) \neq 0 \ . \quad (2.26)$$

[20]A transformation of (standard) symplectic coordinates (q, p) (of a $2d$–dimensional phase space) is called *symplectic* if it preserves the (standard) two form $\sum_{i=1}^d dq_i \wedge dp_i$.

In view of this fact it seems natural to *conjecture that in the present case the formal power series Z is divergent.*

3 Weighted trees

Here we describe the tree family $\mathcal{T}_*^k$, which appears in the basic formula (2.3).

For the models P3) and P4) introduced in § 2, $\mathcal{T}_*^k$ is simply the family of all labeled, rooted trees with k vertices[21], which we will denote by $\mathcal{T}^k$. In this case, as is well known (see *e.g.* [3]), $\#\mathcal{T}^k = k^{k-1}$ and (2.4) is clearly satisfied.

To treat the cases P1) and P2) one has to distinguish, in the Taylor's expansion, the contributions coming from H_0 and L_0 from those coming from H_1 and L_1. To do this we introduce the following family of "weighted trees". Given a rooted (unlabeled) tree T we call a function of the vertices of T

$$\chi : v \in V(T) \to \chi_v \in \{0, 1\} \tag{3.1}$$

a *weight function*. A *weighted rooted tree* is a couple (T, χ) with T a rooted tree and χ a weight function. We now denote $\widetilde{\mathcal{T}}_k$ the set of weighted rooted trees satisfying

$$(i) \quad \deg v \leq 2 \implies \chi_v = 1 \ , \qquad (ii) \quad \sum_{v \in T} \chi_v = k \ . \tag{3.2}$$

Notice that, in particular, all final vertices (*i.e.* vertices of degree 1) have weight 1 and that, for any $T \in \widetilde{\mathcal{T}}_k$, $\#V(T) \leq 2k - 1$ as is easy to verify[22]. We now define the class $\mathcal{T}_k$ of *labeled, weighted rooted trees* obtained from $\widetilde{\mathcal{T}}_k$ by labelling with k different labels the k vertices with weight 1.

In cases P1) and P2) we let $\mathcal{T}_*^k = \mathcal{T}_k$; it is easy to check that (2.4) holds also in this case[23].

The relation between Taylor's formula and trees may be based on the following operation $*$ (that we now briefly discuss for the case of $\mathcal{T}_k$; for the case $\mathcal{T}^k$ see Appendix B of [6]). If $T \in \mathcal{T}_k$ we denote by $\widetilde{T}$ the tree in $\widetilde{\mathcal{T}}_k$ obtained by removing the labels from T; we also denote T^0 (or $\widetilde{T}^0$) the unrooted tree obtained from T (or $\widetilde{T}$) by removing the edge ηr *i.e.* by not distinguishing any more the root from the other vertices; finally if T (or $\widetilde{T}$) is an unrooted labeled (or unlabeled) tree and r is one of its vertices, we denote by T_r (or $\widetilde{T}_r$) the rooted tree obtained by adding the edge ηr (*i.e.* by decreeing that r is the root).

[21] The basic terminology can be found in any introductory book on graphs or in Appendix A of [6].

[22] Let $V_i = \{v \in V : \chi_v = i\}$, and let $k_i = \#V_i$, so that $k = k_1$. It is well known (see, *e.g.* , [3] and recall our convention on degree of the root) that $\sum_{V_0} \deg v + \sum_{V_1} \deg v = 2(k_0 + k_1) - 1$. If $v \in V_0$ then $\deg v \geq 3$, thus the sum over V_0 can be bounded from below by $3k_0$ while the sum over V_1 can be bounded from below by k_1. This leads to $k_0 \leq k_1 - 1$ which is the claim. Such inequality is optimal.

[23] Since ([3]) $\#\widetilde{\mathcal{T}}^h \leq 4^h$ and $\#V(T) \leq 2k-1$ for any $T \in \widetilde{\mathcal{T}}_k$ one sees that $\#\widetilde{\mathcal{T}}_k \leq 4^{2k}$. Since the labels are attached to k vertices, it is $\#\mathcal{T}_k \leq k!4^{2k}$.

Let $\chi \in \{0,1\}$ and $\ell \geq 1$, let h_i be ℓ positive integers such that $h_1 + \cdots + h_\ell = k - \chi$ and pick trees $T_i \in \widetilde{\mathcal{T}}_{h_i}$. We can form a tree $T \in \widetilde{\mathcal{T}}_k$ with root r (r being a vertex different from the vertices of T_i, $\forall$ i) of weight $\chi_r \equiv \chi$ by setting

$$T \equiv T_{h_1} * \cdots * T_{h_\ell} \equiv \Big(T^0_{h_1} \cup \cdots \cup T^0_{h_\ell} \cup \{r\} + \sum_{i=1}^{\ell} r r_i \Big)_r \tag{3.3}$$

where r_i is the root of T_i (and summing an edge e to a tree S means, obviously, to add e to $E(S)$). Then, one has the following

Proposition 3.1 *Let F be a complex valued function defined on trees in $\widetilde{\mathcal{T}}_k$ (for any k). Then*

$$\frac{1}{k!} \sum_{T \in \mathcal{T}_k} F(\widetilde{T}) = \sum_{\chi=0,1} \sum_{\ell=2-\chi}^{k-\chi} \frac{1}{\ell!} \sum_{h_1+\cdots+h_\ell=k-\chi} \prod_{i=1}^{\ell} \frac{1}{h_i!} \sum_{T_i \in \mathcal{T}_{h_i}} F(\widetilde{T}_{h_1} * \cdots * \widetilde{T}_{h_\ell}) \ . \tag{3.4}$$

For the proof we refer to [6] (Corollary B.1 of Appendix B)[24] .

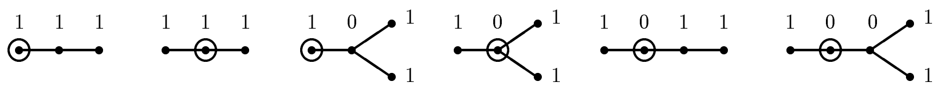

Figure 1: The elements of $\widetilde{\mathcal{T}}_3$. (The numbers 0, 1 are the values of the index χ for the corresponding vertices and the encircled vertex is the root of the tree).

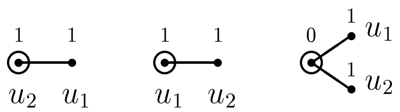

Figure 2: The elements of $\mathcal{T}_2$ (u_1 and u_2 are the labels of $\mathcal{T}_2$).

4 Compensations I (Maximal Hamiltonian tori)

4.1 Tree expansion

Consider the model introduced in P1) of § 2 and recall that there exists a unique formal solutions $Z \sim \sum_{k \geq 0} \varepsilon^k Z^k$ satisfying (2.19) and (2.20). Denote by $Z^{(1)k}$ the x–component

[24]In [6] it is treated the case of $\mathcal{T}^k$ (the $*$ operation in $\widetilde{\mathcal{T}}^k$ is defined as above, replacing χ systematically by 1); adapting the proof to $\mathcal{T}_k$ is a trivial exercise.

($i.e.$ the first N components) of the vector Z^k and by $Z^{(2)k}$ the y–component; consistently, let $\partial^{(1)} \equiv \partial_x$ and $\partial^{(2)} \equiv \partial_y$. We also let

$$[\,\cdot\,]_k \equiv \frac{1}{k!}\frac{d^k}{d\varepsilon^k}\,(\cdot)|_{\varepsilon=0}$$

denote the $k^{\rm th}$ order operator which to a (possibly formal) power series $a \sim \sum a_k \varepsilon^k$ associates its $k^{\rm th}$ order coefficient: $[a]_k \equiv a_k$. Finally let

$$A \equiv \partial_y^2 H_0(y_0)\ . \tag{4.1}$$

With these definitions, we can rewrite (2.19) as

$$DZ^{(1)k} = AZ^{(2)k} + [\partial^{(2)}H_0]_k^{(k-1)} + [\partial^{(2)}H_1]_{k-1}^{(k-1)}\ , \qquad DZ^{(2)k} = -[\partial^{(1)}H_1]_{k-1}^{(k-1)}\ , \tag{4.2}$$

where the suffix ${}^{(k-1)}$ means that the argument of the function within square brackets is, for $k \geq 2$, the polynomial in ε of degree $(k-1)$ given by

$$x = \theta + \sum_{h=1}^{k-1} \varepsilon^h Z^{(1)k}\ , \qquad y = y_0 + \sum_{h=1}^{k-1} \varepsilon^h Z^{(2)k}\ , \tag{4.3}$$

and, for $k=1$, is $(x,y) = (\theta, y_0)$. Since the average of $[\partial^{(1)}H_1]_{k-1}^{(k-1)}$ vanishes (as is clear from the second of (4.2)) we can apply to it the operator D^{-1} obtained by inverting the constant coefficient operator[25] D and rewrite (4.2) in a more compact way as

$$DZ^{(\rho)k} = (2-\rho)AZ^{(2)k} + \sum_{\chi=0,1} (-1)^{3-\rho}[\partial^{(3-\rho)}H_\chi]_{k-\chi}^{(k-1)}\ , \qquad (\rho = 1,2)\ . \tag{4.4}$$

Notice that while, by (2.20), the average of $Z^{(1)k}$ vanishes, the average of $Z^{(2)k}$ can be read (for $\rho = 1$) from (4.4) by integrating over ^{N}:

$$\int_{{}^N} Z^{(2)k} = -A^{-1} \sum_{\chi=0,1} \int_{{}^N} [\partial^{(2)}H_\chi]_{k-\chi}^{(k-1)}\ . \tag{4.5}$$

Taking the n–Fourier coefficient of (4.4), (4.5) one gets

$$Z_n^{(\rho)k} = \sum_{\substack{\sigma\in\{0,1,2\}^* \\ \chi\in\{0,1\}}} \langle n\rangle^{-\sigma} \left\{[D^{(\sigma,\rho)}H_\chi]_{k-\chi}^{(k-1)}\right\}_n\ , \qquad (\langle n\rangle \equiv \omega\cdot n)\ , \tag{4.6}$$

where $D^{(\sigma,\rho)}$ is the vector–valued operator[26]

$$D^{(\sigma,\rho)} \equiv (-1)^{1-\sigma\rho}\ i^{-\sigma}\ A^{\sigma-1}\ \partial^{(4-\sigma-\rho)}\ , \tag{4.7}$$

[25]If f is a (smooth) function on ^{N} with vanishing mean value we denote by $D^{-1}f$ the unique solution with vanishing mean value of the equation for g: $Dg = f$. Expanding in Fourier series one has $g(\theta) = D^{-1}f(\theta) = -i \sum_{n\in{}^N\setminus\{0\}} \frac{f_n}{\omega\cdot n}\exp(in\cdot\theta)$, where $i = \sqrt{-1}$: the inversion of D introduces the small divisors.

[26]For example the $j^{\rm th}$ component of $D^{(2,1)}$ is given by $D_j^{(2,1)} = \sum_{s=1}^N \frac{\partial^2 H_0}{\partial y_j \partial y_s}(y_0)\frac{\partial}{\partial x_s}$.

and the $*$ attached to the range of σ means that the following *constraints* have to be satisfied: $\sigma+\rho\in\{2,3\}$ and $\sigma=0 \iff n=0$ in which case we adopt the convention that $\langle n\rangle^\sigma = 0^0 \equiv 1$. Notice that $D^{(\sigma,\rho)}H_\chi = 0$ if $\chi=0$ and $\sigma+\rho=3$ (as in such a case $\partial^{(4-\sigma-\rho)} = \partial_x$ and H_0 is independent of x).

We shall now use Taylor's formula in the following form. If $f : x \in {}^m \to f(x) \in$ is a C^∞ function and if $a(\varepsilon) \sim \sum_{s\geq 1}\varepsilon^s a^{(s)}$ is a ^{m}–valued (possibly formal) power series, then

$$f(a(\varepsilon)) \sim f(0) + \sum_{h\geq 1}\varepsilon^h \sum_{\ell=1}^{h}\frac{1}{\ell!}\sum_{\substack{h_1+\cdots+h_\ell=h\\ 1\leq h_i\leq \ell}}\ \sum_{\substack{j_1,\ldots,j_\ell\\ 1\leq j_i\leq m}} \frac{\partial^\ell f}{\partial x_{j_1}\cdots\partial x_{j_\ell}}(0)\ a_{j_1}^{(h_1)}\cdots a_{j_\ell}^{(h_\ell)}\ . \tag{4.8}$$

Thus, if $z\equiv(x,y)$ and $z^{(1)}\equiv x$, $z^{(2)}\equiv y$, by (4.6) and (4.8) we get for the j^{th} component of the vector $Z_n^{(\rho)k}$, and for $k\geq 2$,

$$Z_{nj}^{(\rho)k} = \tag{4.9}$$
$$\sum_{\substack{\sigma\in\{0,1,2\}^*\\ \chi\in\{0,1\}}}\ \sum_{\ell=2-\chi}^{k-\chi}\frac{1}{\ell!}\sum_{\substack{1\leq h_i\leq k-1\\ \Sigma h_i=k-\chi\\ (1\leq i\leq \ell)}}\ \sum_{\substack{n_i\in{}^N\\ \Sigma n_i=n\\ (0\leq i\leq\ell)}}\ \sum_{\substack{\rho_1,\ldots,\rho_\ell\\ \rho_i\in\{1,2\}}}\ \sum_{\substack{j_1,\ldots,j_\ell\\ 1\leq j_i\leq N}} \langle n\rangle^{-\sigma}\Big\{\frac{\partial^\ell\ D_j^{(\sigma,\rho)}H_\chi}{\partial z_{j_1}^{(\rho_1)}\cdots\partial z_{j_\ell}^{(\rho_\ell)}}\Big\}_{n_0}\prod_{i=1}^{\ell} Z_{n_i j_i}^{(\rho_i)h_i}\ ,$$

where the derivatives of H_χ are evaluated at (θ,y_0) and then one takes the n_0–Fourier coefficient (with respect to θ); for $k=1$, since $[D^{(\sigma,\rho)}H_0]_1^{(0)}=0$, one has the simple formula

$$Z_{nj}^{(\rho)1} = \sum_{\sigma\in\{0,1,2\}^*}\langle n\rangle^{-\sigma}\{D_j^{(\sigma,\rho)}H_1\}_n\ . \tag{4.10}$$

We are ready to prove the following *tree expansion formula* (recall (2.7), (2.5))

$$Z_{nj_0}^{(\rho_0)k} = \frac{1}{k!}\sum_{T_r\in\mathcal{T}_k}\ \sum_{\substack{\alpha:V\to\alpha_v\in{}^N\\ \Sigma_v\alpha_v=n}}\ \sum_{\substack{\beta:V\to\beta_v\in B\\ \rho_r=\rho_0}}\ \sum_{\substack{j:V\to j_v\in\{1,\ldots,N\}\\ j_r=j_0}}\ \prod_{v\in T_r}\{\Lambda_v(T_r,\beta)H_{\chi_v}\}_{\alpha_v}\prod_{v\in T_r}\gamma_v\ , \tag{4.11}$$

where: the index set B, which depends on the function δ_v, is defined as

$$\begin{aligned} B\ &\equiv \Big\{\beta=(\sigma,\rho):\ \sigma\in\{0,1,2\};\ \rho\in\{1,2\};\ \text{s.t.}\\ &\qquad \sigma+\rho\in\{2,3\}\ ,\sigma=0\iff\ \delta_v=0\Big\}\ ; \end{aligned} \tag{4.12}$$

the scalar operator $\Lambda_v(T_r,\beta)$ is given by

$$\Lambda_v(T_r,\beta)\equiv D_{j_v}^{(\sigma_v,\rho_v)}\prod_{v'\in\mathcal{N}_v}\partial_{j_{v'}}^{(\rho_{v'})}\ ,\qquad \mathcal{N}_v\equiv\{v'<v; v' \text{ adjacent to } v\}\ , \tag{4.13}$$

(if $\mathcal{N}_v=\emptyset$, *i.e.* $\deg v=1$, the product is omitted). To recover (2.3) from (4.11), one simply defines the component j_0 of $\Lambda(T_r,\alpha,\beta)$ as

$$\Lambda(T_r,\alpha,\beta)\equiv\sum_{\substack{j:V\to\{1,\ldots,N\}\\ j_r=j_0}}\ \prod_{v\in T_r}\{\Lambda_v(\beta)H_{\chi_v}\}_{\alpha_v}\ . \tag{4.14}$$

The proof of (4.11) is by induction. For $k = 1$ (4.11) is an immediate consequence of (4.10). Assume that (4.11) holds with k replaced by $1, ..., k-1$. Given $h \geq 1$, ρ_0, j_0, n, consider the function of (unlabeled) trees $T_r \in \widetilde{\mathcal{T}}_h$, given by

$$F_{nj_0}^{(\rho_0)h}(T_r) \equiv \sum_{\substack{\alpha:V\to^N:\Sigma\alpha_v=n\\ \beta\in B\\ j:V\to\{1,...,N\}:j_r=j_0}} \prod_{v\in T_r} \gamma_v \prod_{v\in T_r} \{\Lambda_v(\beta) H_{\chi_v}\}_{\alpha_v} \tag{4.15}$$

and observe that if $T_r = T_1 * \cdots * T_\ell$ with $T_i \in \widetilde{\mathcal{T}}_{h_i}$ and $h_1 + \cdots + h_\ell = h$, then (fixing χ, n, ρ_0, j_0)

$$\begin{aligned} &F_{nj_0}^{(\rho_0)h}(T_1 * \cdots * T_\ell) \\ &= \sum_{\sigma\in\{0,1,2\}^*} \sum_{n_0+\cdots+n_\ell=n} \sum_{\substack{\rho_1,...,\rho_\ell\\ j_1,...,j_\ell}} \langle n_0\rangle^{-\sigma} \Big\{ \frac{\partial^\ell\ D_j^{(\sigma,\rho)} H_\chi}{\partial z_{j_1}^{(\rho_1)} \cdots \partial z_{j_\ell}^{(\rho_\ell)}}(\cdot, y_0) \Big\}_{n_0} \prod_{i=1}^{\ell} F_{n_i j_i}^{(\rho_i)h_i}(T_i) \end{aligned} \tag{4.16}$$

Equation (4.11) follows now from (4.9), (4.16) and Proposition 3.1.

4.2 Compensations

Here, we show how to choose families of trees $\mathcal{F}$ and corresponding indices so that all chains of resonances are compensable (recall the definitions given in § 2).

Given $T = T_r$, α and β, consider a resonance R and let u denote its first vertex and $w' < u$ the first vertex following R. By definition of resonance it is $\delta_u = \delta_{w'} \neq 0$ (*i.e.* $\sum_{v\in R} \alpha_v = 0$) and $\delta_v \neq \delta_{w'}$ if $u > v > w'$. We shall classify resonances by assigning to them an integer $s \equiv s_R \in \{0, 1, 2\}$, which we shall call *index of the resonance* R. Then, to each resonance $R \subset T$ we shall associate a family $\mathcal{F}_R$ of trees T' obtained by (possibly) changing the edges connecting R with $T\backslash R$ and choosing a suitable set of indices . The family $\mathcal{F}_R$ and the index $\beta' \equiv \beta'(T')$ will be chosen so that (recall (2.14))

$$\sum_{T'\in\mathcal{F}_R} \Lambda(T', \alpha, \beta')\ \pi_R(x; T', \alpha, \beta') = O(x^s) \tag{4.17}$$

and the family $\mathcal{F}_C$ will simply be given by

$$\mathcal{F}_C = \bigcup_{i=1}^{h} \mathcal{F}_{R_i}\ . \tag{4.18}$$

Let R be a resonance and let u be its first vertex and w' the first vertex following R. We define the *index of* R as

$$s_R \equiv \sigma_u + \rho_u - \rho_{w'}\ . \tag{4.19}$$

Thus if $C = (R_1, ..., R_h)$ is a chain of resonances and if $\bar{\sigma}$ denotes its order (see § 2), then

$$\sum_{i=1}^{h} s_{R_i} = \bar{\sigma} + \rho_{u_1} - \rho_z \tag{4.20}$$

where z is the first vertex following R_h (which, by convention, is the last resonance in the chain C). Hence, if (4.17) holds, from (4.18), (4.20), (2.14) and the definition of Λ (4.14) it follows easily that

$$\sum_{T'\in\mathcal{F}_C} \Lambda(T',\alpha,\beta')\ \pi_C(x;T',\alpha,\beta') = O(x^{\bar\sigma+\rho_{u_1}-\rho_z}) \tag{4.21}$$

which implies that the chain C is compensable.

Let $R \subset T$ be a resonance and let $\hat{u}u$ and ww' be the edges connecting R with $T\backslash R$, with $\hat{u} > u \geq w > w'$ (hence $u, w \in R$). We proceed by constructing the family $\mathcal{F}_R$. If $\rho_{w'} \neq 1$ we set $\mathcal{F}'_R = \{T\}$; if $\rho_{w'} = 1$ we let $\mathcal{F}'_R$ be the family of all trees T' obtained from T by replacing the edge ww' with the edge $\bar{w}w'$, as $\bar{w}$ varies in R. Hence, $T \in \mathcal{F}'_R$ and if $\rho_{w'} = 1$, $\#\mathcal{F}'_R = \#R$. If $\sigma_u + \rho_u \neq 3$ we set $\mathcal{F}''_R = \{T\}$; if $\sigma_u + \rho_u = 3$ (*i.e.* $(\sigma_u,\rho_u) = (1,2)$ or $(\sigma_u,\rho_u) = (2,1)$) we let $\mathcal{F}''_R$ be the family of all trees T' obtained from T by replacing the edge $\hat{u}u$ with the edge $\hat{u}\bar{u}$, as $\bar{u}$ varies in R. As above, $T \in \mathcal{F}''_R$ and if $\sigma_u + \rho_u = 3$, $\#\mathcal{F}'_R = \#R$. In the first case (*i.e.* for $T' \in \mathcal{F}'_R$) we do not modify the values β_v (*i.e.* $\beta'(T') \equiv \beta$). In the second case (*i.e.* for $T' \in \mathcal{F}''_R$) we define $\beta' \equiv \beta'(T')$ as follows. If $v \notin R$ then $\beta'_v = \beta_v$. Recall that, by definition, the first vertex of R considered as subtree of T' is $\bar{u}$ while the first vertex of R considered as subtree of T is u. Let $\bar{u} \neq u$ (otherwise, obviously, we set $\beta' \equiv \beta$) and consider the path $P(u,\bar{u})$ connecting u with $\bar{u}$. The path P will be formed by $p \geq 2$ ordered vertices that we denote v_i: the order is such that $v_1 = u$, $v_p = \bar{u}$ and the edges of the path are v_1v_2,...,$v_{p-1}v_p$. We then set

$$\begin{aligned} \beta'_{v_p} &\equiv \beta_{v_1}\ , \\ \beta'_{v_i} &\equiv (\sigma_{v_{i+1}}, 4-\sigma_{v_{i+1}}-\rho_{v_{i+1}})\ , \quad \text{for } 1 \leq i \leq p-1 \\ \beta'_v &\equiv \beta_v\ , \qquad\qquad\qquad\qquad \forall\ v \notin P(u,v)\ . \end{aligned} \tag{4.22}$$

It is easy to see that this definition is well posed (see also (ii) of the following Remark). Finally, we define

$$\mathcal{F}_R \equiv \mathcal{F}'_R \cup \mathcal{F}''_R\ . \tag{4.23}$$

Remark 4.1 (i) If $s_R = 0$, it is $\mathcal{F}_R = \{T\}$ and $\beta' \equiv \beta$.
(ii) The map $\beta \to \beta'$ is involutive: more precisely, if we denote by $\beta'(\beta;u,\bar{u})$ the map defined in (4.22) (definition which depends on the ordered path $P(u,\bar{u})$) then $\beta'(\beta'(\beta;u,\bar{u});\bar{u},u) = \beta$.
(iii) One might say that the families $\mathcal{F}'_R$ and $\mathcal{F}''_R$ are constructed "going around" the resonance R with a "discrete curve" obtained by moving, respectively, the edge connecting R with w' and the edge connecting $\hat{u}$ with R. This interpretation might explain the name "index" given to the quantity s_R: $\mathcal{F}_R$ is obtained by "going around" R exactly s_R times.
(iv) (On the definition of $\mathcal{F}'_R$) In practice, moving around the edge $\bar{w}z$ produces a factor proportional to $\alpha_{\bar{w}}$ in (4.14) coming from the $\partial^{(\rho_z)} = \partial^{(1)}$ appearing in (4.13) (as $z \in \mathcal{N}_{\bar{w}}$). It is easy to see that, for $x = 0$, summing $\Lambda\ \pi_R(x)$ over the family $\mathcal{F}'_R$ produces a *common* factor $\sum_{\bar{w}\in R} \alpha_{\bar{w}}$ which vanishes by definition of resonance.

(v) (On the definition of $\mathcal{F}''_R$) The idea is similar: one wants to produce a factor proportional to $\alpha_{\bar{u}}$ when moving around the "first" edge $\hat{u}\bar{u}$ connecting R with $\hat{u}$. But now the situation is more delicate as changing the connection $\hat{u}\bar{u}$ *changes the order* in R and, consequently, change the small divisors and also the structure of the derivatives (since both γ_v and Λ_v depend on the order). Since one wants eventually to set $x = 0$ and collect the factor $\sum_{\bar{u}\in R} \alpha_{\bar{u}}$ one sees the necessity of changing the values of the indices β. In fact β' is defined in such a way that, when $x = 0$, one can factor out the *product* of divisors and the *products* of the operators (derivatives) Λ_v.

With the above remarks in mind, it is not difficult to verify that (4.17) holds, proving Theorem 2.1 in case P1).

5 Compensations II (Maximal Lagrangian tori)

Consider the model introduced in P2) of § 2 for a real–analytic Lagrangian $L \equiv L_0(y) + \varepsilon L_1(x, y)$. Formal quasi–periodic solutions of the Lagrangian equations have the form $x(t) = Z(\omega t)$ for an ω satisfying (2.18) and a function $Z(\theta)$ which is the (unique) formal solution $Z \sim \sum_{k\geq 0} \varepsilon^k Z^k$, with $Z \in {}^N$, of (2.21). The vector valued function Z^k can be put in a form which is amazingly similar to the Hamiltonian case P1).
Let us denote by $Z^{(1)k}$ the full vector $Z^k \in {}^N$ and by $Z^{(2)k}$ the vector $DZ^k = DZ^{(1)k}$ (where $D \equiv \omega \cdot \partial_\theta$) and, as before, let $\partial^{(1)} \equiv \partial_x$, $\partial^{(2)} \equiv \partial_y$ and $A_L \equiv \partial_y^2 L_0(y_0)$. Equation (2.21) can then be written as

$$A_L D Z^{(2)k} = -D[\partial^{(2)} L_0]_k^{(k-1)} - D[\partial^{(2)} L_1]_{k-1}^{(k-1)} + [\partial^{(1)} L_1]_{k-1}^{(k-1)} \; . \qquad (5.1)$$

adopting the same convention used in (4.2): in particular the arguments of the functions within square brackets are as in (4.3). As for the previous case (see (4.6)), the n–Fourier coefficient of (5.1) is

$$Z_n^{(\rho)k} = \sum_{\substack{\sigma\in\{0,1,2\}^* \\ \chi\in\{0,1\}}} \langle n\rangle^{-\sigma} \left\{ [D_L^{(\sigma,\rho)} L_\chi]_{k-\chi}^{(k-1)} \right\}_n \qquad (5.2)$$

where $D_L^{(\sigma,\rho)}$ is now the vector–valued operator

$$D_L^{(\sigma,\rho)} \equiv (-1)^{1-(\sigma+\rho)} \; i^{-\sigma} \; A_L^{-1} \; \partial^{(4-\sigma-\rho)} \; , \qquad (5.3)$$

and $\{0, 1, 2\}^*$ is the same set of equation (4.6).
The Lagrangian problem is now in a form which is identical to the Hamiltonian case, except for the definition of $D_L^{(\sigma,\rho)}$. Therefore the tree expansion formula for the component j_0 of the n–Fourier coefficient of Z^k, with fixed values of ρ_0 and j_0 at the root r, is given again by (4.11) with the only proviso of replacing $D_{j_v}^{(\sigma_v,\rho_v)}$ in (4.13) with $D_{L,j_v}^{(\sigma_v,\rho_v)}$.
Also, the families $\mathcal{F}$ of trees for which there are compensations are found exactly as in the Hamiltonian case and we refer to § 4.2 for details.

6 Compensations III (Lower dimensional tori)

Recall the notations of § 2, P3). Because of the particular form of the Hamiltonian, the formal solution[27] $Z(\theta,\mu)$ is of the form $Z\equiv(X,Q,DX,DQ)$ where, as usual, $D\equiv\omega\cdot\partial_\theta$, $\theta\in{}^N$ and $X\in{}^N$, $Q\in{}^M$. Denote

$$Z^{(1)}\equiv X\ ,\quad Z^{(2)}\equiv Q\ ,\quad \partial^{(1)}\equiv\partial_x\ ,\quad \partial^{(2)}\equiv\partial_q\ ,\quad A\equiv\varepsilon\partial_q^2 f(q_0)\ (>0)\ . \tag{6.1}$$

One checks immediately that the recursive equations for $Z^{(\rho)k}$ ($\rho=1,2$) are

$$\Big(-D^2+(\rho-1)A\Big)Z^{(\rho)k}=\sum_{\chi=0,1}[\partial^{(\rho)}H_\chi]_{k-\chi}^{(k-1)}\ ,\qquad \rho=1,2 \tag{6.2}$$

where the argument of the derivatives of H_χ is $x=\theta$, $q=q_0$. From (6.2) one can see that the average (over ^{N}) of $Z^{(1)k}$ vanishes, while the average of $Z^{(2)k}$ is as in (4.5) but with the minus sign replaced by a plus. Taking Fourier coefficients of (6.2) we get the analogous of (4.6), namely

$$Z_n^{(\rho)k}=\sum_{\substack{\sigma\in\{0,2\}^*\\ \chi\in\{0,1\}}}\langle n\rangle^{-\sigma}\left\{[D_n^{(\rho)}H_\chi]_{k-\chi}^{(k-1)}\right\}_n\ ,\qquad (\langle n\rangle\equiv\omega\cdot n)\ , \tag{6.3}$$

which differs from (4.6) for the range of σ and relative constraints:

$$\sigma\in\{0,2\}^*\iff\sigma+\rho\in\{2,3\}\ ,\qquad n=0\implies\sigma=0\ , \tag{6.4}$$

and for the definition of the vector valued operator $D_n^{(\rho)}$:

$$D_n^{(\rho)}\equiv\Big(\langle n\rangle^2+A\Big)^{1-\rho}\partial^{(\rho)}\ . \tag{6.5}$$

Notice that the components $D_{nj}^{(\rho)}$ are N if $\rho=1$ and M if $\rho=2$; we therefore let

$$N_1\equiv N\ ,\qquad N_2\equiv M\implies j\in\{1,...,N_\rho\}\ . \tag{6.6}$$

Recall that $A=\varepsilon\partial^2 f_0(q_0)$ is positive definite, and so is $a+A$ for any $a\geq 0$; thus

$$\|(a+A)^{-1}\|\leq\frac{b}{|\varepsilon|} \tag{6.7}$$

for a suitable constant b depending only on $\partial^2 f_0(q_0)$. We also remark that, by (6.4), $\rho=2$ implies $\sigma=0$ (hence no divisors) while if $\rho=1$ then $\sigma=2$ and $n\neq 0$, in which case the divisors are $\langle n\rangle^2$. In view of these remarks, we see that (4.11) holds also in the present

[27] Recall that here ε is a fixed real number different from zero, while the (complex) perturbation parameter appearing in the formal power series is μ.

case provided we change the following items: N in the fourth sum is replaced (see (6.6)) by N_{ρ_v}; the index set B is defined as

$$
\begin{aligned}
B \quad & \equiv \Big\{\beta = (\sigma,\rho) : \ \sigma \in \{0,2\}; \ \rho \in \{1,2\}; \ \text{s.t.} \\
& \qquad \sigma+\rho \in \{2,3\} \ , \delta_v = 0 \quad \Longrightarrow \quad \sigma = 0\Big\} \ ;
\end{aligned} \tag{6.8}
$$

finally the operator Λ_v now depends also on α: $\Lambda_v(T_r,\beta)$ in (4.11) is now replaced by

$$
\Lambda_v(T_r,\alpha,\beta) \equiv D^{(\rho_v)}_{\delta_v j_v} \prod_{v' \in \mathcal{N}_v} \partial^{(\rho_{v'})}_{j_{v'}} \ . \tag{6.9}
$$

Also (4.14) is readily adapted replacing N by N_{ρ_v} and Λ_v with (6.9). We can proceed to define the families of trees $\mathcal{F}$ and relative indices β' which exhibit compensations. Given a resonance R (and a choice of α and β), we define the index of R and the family $\mathcal{F}''_R$ and relative indices β' exactly in the same way we did in § 4.2. Also the family $\mathcal{F}'_R$ is defined in the same way but the relative indices β' are now defined in a slightly different way (due to the different definition of $D^{(\rho)}_n$): we let (same notations as in § 4.2)

$$
\begin{aligned}
\beta'_{v_p} &\equiv \beta_{v_1} \ , \\
\beta'_{v_i} &\equiv \beta_{v_{i+1}} \ , \quad \text{for } 1 \le i \le p-1 \\
\beta'_{v} &\equiv \beta_{v} \ , \qquad \forall \ v \notin P(u,v) \ .
\end{aligned} \tag{6.10}
$$

With these definitions it is easy to check that (4.17) holds and hence that Theorem 2.1 is valid in case P3) too.

We close by a remark on the μ–radius of convergence of $\sum_{k\ge1}\mu^k Z^k$. It is an easy exercise to adapt the estimates in [6] to the present case and to check how the radius of convergence depends on ε. In fact, observing that, by (6.7), one has

$$
|\langle n\rangle^{-\sigma}| \ |(\langle n\rangle^2 + A)^{1-\rho}| \le \max\{\gamma^2, \frac{b}{|\varepsilon|}\} \quad \max\{|n|^{2\tau} \ , \ 1\} \ , \tag{6.11}
$$

leading to an estimate on the radius of convergence μ_0 of the type

$$
\mu_0 \ge \text{const} \ \min\{\varepsilon \ , \ \gamma^{-2}\} \ . \tag{6.12}
$$

Thus $\mu = \varepsilon^2$ (or $\mu = \varepsilon^c$ with any $c > 1$) is within the domain of analyticity provided ε is small enough.

7 An example with no compensations

Consider the Hamiltonian (2.22) with

$$
f(x,q) = \sum_{s\ge1} f_s e^{i(n^{(s)}\cdot x + m^{(s)}\cdot q)} \tag{7.1}
$$

where $n^{(s)} \in {}^N$, $m^{(s)} \in {}^M$ are given integer vectors (with $|n^{(s)}|+|m^{(s)}| > 0$) and the Fourier coefficients f_s decay exponentially fast with $|n^{(s)}| + |m^{(s)}|$. Let q_0 be a non degenerate critical point of the x–average of f (i.e. of $f_0 = \sum_{s\geq 1} f_s e^{im^{(s)}\cdot q}$); then, as for the previous cases, there exists a (unique) formal power series

$$Z \equiv (Z_1, \dots, Z_d) \;\sim\; \sum_{k\geq 0} Z^k(\theta)\varepsilon^k \;,\;\; \theta \in {}^N \tag{7.2}$$

with $d = 2(N + M)$, such that $t \to Z(\omega t)$ is a formal quasi–periodic solution for (2.22) and the set $\{Z^0(\theta) : \theta \in {}^N\}$ coincide with the torus spanned by $y = \omega$, $p = 0$, $q = q_0$, $x = x_0 + \omega t$, where $\omega \in {}^N$ satisfies condition (2.18).
Expanding Z^k in Fourier series also in the variable q, besides the variable x (see (2.2)), it is still possible to write Z^k_n as in (2.3) provided one makes the following changes. $\mathcal{T}^k_* = \mathcal{T}^k$; B is the trivial set $B = \{\beta = \sigma = 2\}$; $\Lambda(T, \alpha, \beta)$ is replaced by

$$\Lambda(T, \hat\alpha, \beta) = \sum_{\alpha' : V \to \alpha'_v \in {}^M} \prod_{v\in V} f_{\hat\alpha_v} \prod_{vv'\in E} \hat\alpha_v \cdot \hat\alpha_{v'} \tag{7.3}$$

where $\hat\alpha_v \equiv (\alpha_v, \alpha'_v) \in {}^{N+M}$; finally $\gamma_v = \langle\delta_v\rangle^{-2} \equiv \left(\omega \cdot \sum_{v'\leq v} \alpha_{v'}\right)^{-2}$.
It is then clear that in order to have compensations it would be sufficient to have resonances in the variable $\hat\alpha$ whenever there are resonances in the variable α. In other words compensations take place if the Fourier modes in (7.1) are such that

$$\sum_{s\in I} n^{(s)} = 0 \qquad \Longrightarrow \sum_{s\in I} m^{(s)} = 0 \tag{7.4}$$

$\forall I \subset$; in fact in such a case one could repeat word–by–word the arguments in [6].
In general, if (7.4) does not hold, compensations (of the type described in this paper) do not occur as it is shown by the following example.

Take $N = 2$, $M = 1$, fix a scalar integer $n \neq 0$ and let

$$f(x_1, x_2, q) = 2\{\cos x_1 + \cos x_2 + \cos(x_1 + x_2 - q) + \cos(nx_1 + x_2 + q)\} \;. \tag{7.5}$$

Hence, the range of $\hat\alpha$ is the set $\{\pm(1, 0, 0),\ \pm(0, 1, 0),\ \pm(1, 1, -1),\ \pm(n, 1, 1)\}$; for definiteness we also fix $\omega = (\sqrt{2}, -1)$.

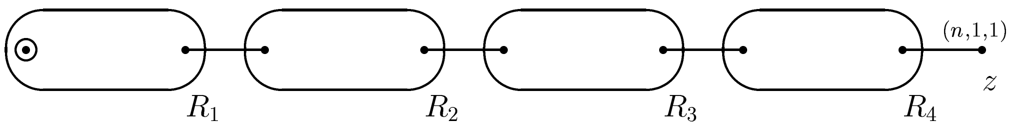

Figure 3: Divergent contribution: the chain $C = (R_1, R_2, R_3, R_4)$ is made of 4 adjacent resonances with $|R_i| = 3$ (in this symbolic picture only two vertex are drawn).

We fix a tree $T \in \widetilde{\mathcal{T}}^k$, with $k = 3h+1$, and a function $\hat{\alpha}_v : V(T) \to {}^3$ so that T contains a chain $C = (R_1, ..., R_h)$ made of h identical resonances $R_i \equiv \{v_1, v_2, v_3\}$, with $\hat{\alpha}_{v_1} = (-1,-1,1)$, $\hat{\alpha}_{v_2} = (0,1,0)$, $\hat{\alpha}_{v_3} = (1,0,0)$, and the last vertex z, following the chain, with $\hat{\alpha}_z = (n,1,1)$ (see Figure 3).
Let $\mathcal{F}_C$ be the family of all trees $T \in \mathcal{T}^k$ which contains the chain C. A lengthy but straightforward computation shows that the i^{th} component of the vector valued function $\bar{\pi}_C(x)$ (see (2.17)) for $h = 1$ is:

$$\begin{aligned}(\bar{\pi}_C(x))_i \equiv (\bar{\pi}_{R_1}(x))_i \quad &\equiv \prod_{v\in V} f_{\hat{\alpha}_v} \sum_{u,w\in R_1} \hat{e}_i \cdot \hat{\alpha}_u \Big(\prod_{vv'\in E(R_1)} \hat{\alpha}_v \cdot \hat{\alpha}_{v'} \Big)\, \hat{\alpha}_w \cdot \hat{\alpha}_z \; \pi_{R_1}(x) \\ &= 2\sum_{j=1}^{3} (\delta_{i3}\delta_{j3} + x^2 B_{ij}(x))\hat{\alpha}_{zj} \qquad (7.6)\end{aligned}$$

where δ_{ij} is the Kronecker's symbol, $\hat{e}_i$ is the vector of j^{th} component δ_{ij} and

$$B(x) = \begin{pmatrix} \frac{-2(6-x^2)}{(2-x^2)^2} & \frac{8\sqrt{2}+18x^2-9x^4+x^6}{(2-x^2)^2(1-x^2)^2} & \frac{6-x^2}{(2-x^2)^2} \\ \frac{8\sqrt{2}+18x^2-9x^4+x^6}{(2-x^2)^2(1-x^2)^2} & \frac{-2(3-x^2)}{(1-x^2)^2} & \frac{3-x^2}{(1-x^2)^2} \\ \frac{6-x^2}{(2-x^2)^2} & \frac{3-x^2}{(1-x^2)^2} & 0 \end{pmatrix}. \qquad (7.7)$$

Hence $\bar{\pi}_{R_1}(0) \neq 0$. For $h > 1$, one simply has:

$$\bar{\pi}_C(x) = 2^h (A + x^2 B)^h \hat{\alpha}_z \quad , \quad A_{ij} = \delta_{i3}\delta_{j3} \; . \qquad (7.8)$$

Thus, if $x = \omega \cdot \alpha_z = \sqrt{2}n - 1$, taking the third component of (7.8) (the component for which condition (7.4) is violated) and assuming that $n \geq h$, one easily checks that

$$((A + x^2 B)^h \hat{\alpha}_z)_3 \geq 1 + O(\frac{1}{n}) \qquad (7.9)$$

which implies that C is a non compensable chain.

References

[1] Arnold, V.I. (Ed.): *Encyclopaedia of Mathematical Sciences, Dynamical Systems*, Vol. 3, Springer–Verlag, 1988

[2] Arnold, V.I.: *Instability of dynamical systems with several degrees of freedom* Sov. Math. Dokl. **5** 581–585 (1964)

[3] Bollobas, B.: *Graph Theory*, Springer (Graduate text in mathematics: 63), 1979

[4] Brjuno, A. D.: *Convergence of transformations of differential equations to normal form*, Dokl. Akad. Nauk SSSR **165**, 987–989 (1965); *Analytic form of differential equations*, Trans. Moscow Math. Soc. **25**, 131–288 (1971) and **26**, 199–239 (1972)

[5] Celletti, A.; Chierchia, L.: *Construction of analytic KAM surfaces and effective stability bounds*, Commun. Math. Phys. **118**, 119–161 (1988)

[6] Chierchia, L.; Falcolini, C.: *A direct proof of a theorem of Kolmogorov in Hamiltonian systems*, to appear in Annali Sc. Norm. Super. Pisa, Cl. Sci.

[7] Chierchia, L.; Falcolini, C.: *Quasi–periodic solutions of some elliptic systems*, Preprint (1993)

[8] Chierchia, L.; Gallavotti, G.: *Drift and Diffusion in phase space*, Ann. Inst. Henri Poincaré (Physique Théorique) **60**, n° 1, 1–144 (1994)

[9] Eliasson, L. H.: *Absolutely convergent series expansions for quasi periodic motions*, Reports Department of Math., Univ. of Stockholm, Sweden, No. 2, 1–31 (1988)

[10] Eliasson, L. H.: *Hamiltonian systems with linear normal form near an invariant torus*, in "Nonlinear Dynamics", G. Turchetti (Ed.) World Scientific, Singapore (1989)

[11] Eliasson, L. H.: *Generalization of an estimate of small divisors by Siegel*, in "Analysis, et cetera", P.H. Rabinowitz and E. Zehnder (Eds.), Academic Press (1990)

[12] Gallavotti, G.: *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable hamiltonian systems. A review.* Preprint (1993)

[13] Gallavotti, G.: *Twistless KAM tori* to appear in Commun. Math. Physics

[14] Gallavotti, G.: *Invariant tori: a field theoretic point of view on Eliasson's work* Preprint (1993)

[15] Gallavotti, G.; Gentile, G.: *Non recursive proof of the KAM theorem*, to appear in Ergodic Theory and Dynamical Systems

[16] Gentile, G.: *A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable hamiltonian systems* Preprint (1994)

[17] Gentile, G.: *Whiskered tori with prefixed frequencies and Lyapunov spectrum* Preprint (1994)

[18] Graff, S.M.: *On the conservation of hyperbolic invariant tori for Hamiltonian systems*, J. Diff. Equations **15**, 1–69 (1974)

[19] Poincaré, H.: *Les méthodes nouvelles de la mécanique céleste.* Vols. 1–3. Paris: Gauthier–Villars. (1892/1893/1899)

[20] Siegel, C. L.: *Iterations of analytic functions*, Annals of Math. **43**, No. 4, 607–612 (1942)

[21] Salamon, D.; Zehnder, E.: *KAM theory in configuration space*, Comm. Math. Helv., **64**, 84 (1988)

[22] Xia, Z.: *Arnold diffusion and oscillatory solutions in the three–body problem*, Preprint (1993)