Commun. Math. Phys. 175, 135 – 160 (1996)

Communications in Mathematical Physics

Compensations in Small Divisor Problems

L. Chierchia[1], **C. Falcolini**[2]

[1] Dipartimento di Matematica, Terza Università di Roma, via Corrado Segre 2, I-00146 Roma, Italy. Internet: luigi@matrm3.mat.uniroma3.it
[2] Dipartimento di Matematica, Università di Roma "Tor Vergata", via della Ricerca Scientifica, I-00133 Roma, Italy. Internet: falcolini@vax.mat.utovrm.it*

Received: 11 August 1994/in revised form: 10 March 1995

Abstract: A general direct method, alternative to KAM theory, apt to deal with small divisor problems in the real–analytic category, is presented and tested on several small divisor problems including the construction of maximal quasi-periodic solutions for nearly-integrable non-degenerate Hamiltonian or Lagrangian systems and the construction of lower dimensional resonant tori for nearly-integrable Hamiltonian systems. The method is based on an explicit graph theoretical representation of the formal power series solutions, which allows to prove compensations among the monomials forming such representation.

Contents

1. Introduction

1. Small divisors are ubiquitous in non-linear conservative dynamical systems; they arise, for example, in: conjugacy problems such as linearizations of germs of analytic functions or linearizations of circle maps (see, e.g., [1] and references therein);

* L.C. thanks C. Simó and the *Centre de Recerca Matemàtica* (Bellaterra) for kind hospitality; he also acknowledges partial support by CNR–GNAFA. The authors gratefully acknowledge helpful discussions with C. Liverani.

in constructing invariant tori or curves for Hamiltonian flows or symplectic diffeomorphisms ([2] and references therein); in constructing periodic or quasi-periodic solutions for non-linear PDE's (partial differential equations) carrying a Hamiltonian structure (see, e.g., [25, 32, 12] and references therein) or for systems of elliptic PDE's (see [10]).

Essentially, the only *general technique* used to overcome small divisor problems has been the powerful *KAM theory* started by A.N. Kolmogorov in the early fifties and developed in the early sixties by V.I. Arnold and J.K. Moser (see [2] and references therein).

This paper might be viewed as an attempt to develop *general tools*, which are different from those of KAM theory, *apt to solve small divisor problems in the real analytic category*.

2. The strategy that we shall pursue here goes back to the 1942 work of C.L. Siegel [29], who was the first to overcome a small divisor problem (namely the conjugation to its linear part of an analytic function of the form $e^{i\omega}z + z^2 g(z)$ with ω a real number satisfying suitable generic conditions). *Siegel's method* consisted "simply" in computing the formal power series solution (for the conjugating function) and estimating its k^{th} coefficient by a constant to the k^{th} power, thus showing the convergence of the conjugating function. Generalizations of Siegel's method will henceforth be referred to as *direct methods.*

Recently, such methods have been revived: In 1988 H. Eliasson, [13] (see also [14, 15]), proposed a direct proof of the convergence of maximal quasi-periodic solutions (or maximal invariant tori or Lindstetd series) for nearly-integrable, non-degenerate Hamiltonian systems; in 1993–94 several preprints/papers appeared on the subject: in [9, 17–20] the convergence of maximal quasi-periodic solutions for Hamiltonians of the form $\frac{1}{2}y^2 + \varepsilon f(x)$ (x periodic vector variable, y canonically conjugate vector variable) is proved by direct methods[1]; [21, 22] extend [17] and use direct methods to prove the existence of "whiskered" (see [3]) tori (of dimension $N-1$, if $2N$ is the dimension of the phase space) for certain special Hamiltonians; in [10] direct methods are extended so as to prove convergence of quasi-periodic solutions for elliptic systems of PDE's of the form $\Delta u = \varepsilon f_u(u, y)$ (where u is a vector-valued function of $y \in \mathbb{R}^M$ and f is periodic in all its variables).

3. The obstacle to the extension of Siegel's method to other small divisor problems (motivating KAM theory and explaining the time gap between Siegel's work and the one above mentioned direct proofs) is related to certain *repetitions of small divisors* appearing in the formal series solutions. To describe this phenomenon, let us consider the "mother of all small divisor problems," namely the convergence of the Lindstetd series for nearly-integrable, non-degenerate Hamiltonian systems (see [28]). Let $H = h(y) + \varepsilon f(x, y)$ be a real-analytic Hamiltonian, parametrized by ε, periodic in the "angle variables" $x = (x_1, \ldots, x_N)$ and with the "action variables" y varying in some ball around a point $y_0 \in \mathbb{R}^N$. Assume the following non-degeneracy

[1] The approach in these papers is quite similar and is different from that of Eliasson (for a comparison with Eliasson's method, see [9]). In [17–20] similarities between direct methods in classical mechanics and some aspects of constructive field theory are pointed out. Also, in [17–22], it is assumed that the perturbation f is an even trigonometric polynomial, a fact that makes the analyses in [9, 10] and the present paper, on one side, and in [17–20], on the other, somewhat different. After the completion of this paper, we received a preprint [23] which extend the analysis of [17–20].

conditions: (i) the "frequency" vector $\omega \equiv \partial_y h(y_0)$ is Diophantine, i.e.

$$|\omega \cdot n| \geqq \frac{1}{\gamma |n|^\tau}, \quad \forall\, n \in \mathbb{Z}^N \backslash \{0\}, \tag{1.1}$$

for some positive constants γ, τ (dot denotes the standard inner product in $\mathbb{R}^N$); (ii) the Hessian matrix $\partial_y^2 h(y_0)$ is invertible. Then one can formally compute the Lindstetd series, i.e. the formal series

$$X \sim \sum_{j \geqq 0} X^j \varepsilon^j, \qquad Y \sim \sum_{j \geqq 0} Y^j \varepsilon^j,$$

with X^j, Y^j (vector-valued) real-analytic on the standard N torus $\mathbb{T}^N \equiv \mathbb{R}^N/(2\pi\mathbb{Z}^N)$, $Y^0 \equiv y_0$ and such that $t \to (x(t), y(t)) \equiv (\omega t + X(\omega t), Y(\omega t))$ is a formal solution (in the sense of formal power series in ε) of the (standard) Hamiltonian equations $\dot{x} = \partial_y H$, $\dot{y} = -\partial_x H$ (over dot denoting t derivative). The Fourier coefficients (X_n^k, Y_n^k) may be recursively computed in terms of monomials made up of k Fourier coefficients of the periodic functions $x \to \partial_y^s f(y_0, x)$ (with $s \leqq k$) and in terms of the divisors $\omega \cdot m$, $m \in \mathbb{Z}^N$ (which appear in the denominator of the monomial), and in terms, of course, of $\partial_y^s h(y_0)$. Roughly speaking, what happens is that, in the expression of some of these monomials, there appear products of a given divisor $\omega \cdot m$ *which are not accompanied* by Fourier coefficients of the Hamiltonian with Fourier index "related" to m. Since $|\omega \cdot m|$ may be arbitrarily small (whence the name "small divisors"), it is not clear what can counter-balance such products of small divisors, which may accumulate in such a way to produce single monomials, in the recursive decomposition of (X_n^k, Y_n^k), of size of the factorial $k!$ (compare, e.g., [9], Appendix B). Controlling such formal series is therefore a difficult problem (which lead Poincaré to say that convergence of the Lindstetd series is "fort invraisemblable," [28], vol. II, Sect. XIII).

4. The direct method discussed here rests on two clearly distinct steps: (a) *the algebraic part of the method* (*compensations*), which consists in showing that it is possible to consistently group together the monomials with *extra* repetitions of small divisors[2] into quantities that behave effectively as *if there were no extra repetitions* (on a technical level this will be achieved by showing that certain meromorphic functions have zeroes of high enough degree); (b) *the quantitative part of the method*, which consists in implementing Siegel's estimates so as to prove convergence of the formal series.

Step (b) is practically common to all small divisor problems; a detailed version of the needed estimates can be found, e.g., in [9] (see, in particular, Lemma 5.1 and Lemma 5.2), and one can adapt such estimates (once the algebraic step is settled) to other small divisor problems including the ones discussed in this paper.

The key step, changing from problem to problem, is the algebraic step (a) and the rest of the paper is devoted to develop a formalism general enough to deal, in a unified manner, with different small divisor problems (which, instead, need quite different KAM approaches as explained below). The language we shall use to prove compensations is borrowed from graph theory (see e.g. [4] or Appendix A of [9] for the fundamentals and [9–11, 17–22] for use of graph theoretic language

[2] "Extra" means, as above, that the small divisor $\omega \cdot m$ is not balanced by a proportional number of Fourier coefficients with index m.

in the context of dynamical systems). In particular we will introduce a natural generalization of labeled rooted trees (called below "weighted trees") which will turn out to be quite convenient in order to describe, in a completely explicit way, formal solutions and in order to recognize the compensating families of monomials in the (tree) decomposition of the formal solution.

In this paper we shall analyze the algebraic step (a) in various cases, while we shall not deal with the quantitative part of the method (b) which is a technical issue and, as mentioned above, may be recovered rather straightforwardly from [9].

5. Let us now discuss briefly the models and respective small divisor problems for which we prove compensations.

P1 (*Maximal invariant tori, Hamiltonian case*). The first problem for which we prove compensations is that of convergence of the Lindstedt series for nearly-integrable, non-degenerate Hamiltonian systems (see point 1 above). The compensations (step (a)) plus the estimates of [9] (step (b)) yield a new direct proof of the classical theorem by Kolmogorov (sometimes called "KAM theorem") alternative to that of [13].

P2 (*Maximal invariant tori, Lagrangian case*). We then consider the similar problem of finding maximal quasi-periodic solutions for the Euler–Lagrange equations associated to a real analytic Lagrangian of the form $L(x,y) \equiv L_0(y) + \varepsilon L_1(x,y)$. As above, $x = (x_1, \dots, x_N)$ is a periodic vector-valued variable and y varies in some ball around a point $y_0 \in \mathbb{R}^N$; the analogous non-degeneracy conditions are assumed, namely one is given a vector ω satisfying the Diophantine condition (1.1) and y_0 is such that the Hessian $\partial_y^2 L_0(y_0)$ is invertible. A maximal quasi-periodic solution is, then, a solution $x(t)$ of $\frac{d}{dt} L_y(x(t), \dot{x}(t)) = L_x(x(t), \dot{x}(t))$ of the form $x(t) = \omega t + X(\omega t)$ with $X(\theta)$ periodic in θ. Step (a) is then proved in essentially the same way used in the Hamiltonian case P1. We point out that the KAM theory needed to solve the Lagrangian problem is much more recent (1988) than classical KAM theory and is due to Moser, Salamon and Zehnder (see [30] and references therein). The Lagrangian approach is more convenient than the Hamiltonian one for rigorous stability bounds (see [6, 7]) and also for generalizations to variational PDE cases (see [26, 10]).

P3 (*Lower dimensional resonant tori*). 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 ω rationally independent over $\mathbb{Z}^N$ and $N = \#$ of degrees of freedom. We consider now the problem of analytic continuation (in ε) of "resonant" tori, i.e. invariant unperturbed tori for which there exist $n \in \mathbb{Z}^N \backslash \{0\}$ such that $\omega \cdot n = 0$. We shall argue (see the 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 the following special case of a Hamiltonian of the form

$$H \equiv \frac{1}{2} y^2 + \frac{1}{2} p^2 + \varepsilon f(x,q) \,, \tag{1.2}$$

where (x,y) are symplectic variables as above (i.e. with $x \in \mathbb{T}^N$) and so are (q,p) with $q \in \mathbb{T}^M$; hence the number of degrees of freedom is $N+M$ and we are interested in N-dimensional invariant tori. For $\varepsilon = 0$, N-dimensional invariant tori (up to a trivial linear and symplectic change of coordinates[3]) 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 (1.1)) the evolution equations of H in (1.2) 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 $\mathbb{T}^N$ of f. Motivated by this observation, we consider the Hamiltonian

$$\begin{aligned}\tfrac{1}{2}y^2+\tfrac{1}{2}p^2+\varepsilon f_0(q)+\varepsilon^2\tilde{f}(x,q) &\equiv H_0(q,y,p;\varepsilon)+H_1(x,q;\varepsilon)\\ H_1 &\equiv \varepsilon^2\tilde{f}\ .\end{aligned}\tag{1.3}$$

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 a formal (and eventually convergent with a 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 (1.3)). We also make the following *hyperbolicity assumption*: we assume that q_0 and ε are such that the matrix $\varepsilon\partial_q^2 f_0(q_0)$ is *negative definite*. Under these hypotheses it is easy to check that there exists a formal expansion

$$Z:\mathbb{T}^N\to\mathbb{R}^{2(N+M)}\ ,\qquad Z\sim\sum_{k\geqq 1}Z^k(\theta;\varepsilon)\mu^k\ ,\tag{1.4}$$

such that $t\to\big(x(t),y(t),q(t),p(t)\big)\equiv\big(\omega t,\omega,q_0,0\big)+Z(\omega t)$ is a formal quasi-periodic solution of the Hamiltonian equations governed by $H_0+\mu H_1$. Then, *we can prove compensations for the formal solution* (1.4). 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 $(\omega t,\omega,q_0,0)+\{Z(\theta):\theta\in\mathbb{T}^N,\mu=1\}$ is an invariant N-torus for the Hamiltonian (1.3). We mention that, in fact, these tori are *whiskered* in the sense of [3].

The proof of compensations (together with the estimates of [9]) yields a new (direct) proof of the construction of lower dimensional resonant tori (with particular emphasis on analyticity properties).

The KAM theory concerning "partially hyperbolic tori" (i.e. generalizations of the case $H_0+\mu H_1$ above, with μ as perturbative parameter and ε fixed) goes back to [24]. For a KAM theory for resonant tori see [31]. In [17, 21, 22] a direct proof is given for the existence of resonant tori (and their whiskers) in case (1.3) with p,q scalar ($M=1$), $f_0=\cos q$ and $\tilde{f}$ an even trigonometric polynomial. We finally mention that the resonant tori (and their whiskers) are particularly relevant for the study of "Arnold diffusion" (see [3, 11, 8]).

Different is the story for "partially elliptic tori" corresponding to $\varepsilon\partial_q^2 f_0(q_0)$ being positive definite. Such a case is technically more difficult (due to the presence of "extra small divisors") and it has been overcome, with KAM technique, by Eliasson

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

[16] (see also [27]). We believe that the direct method may be applied also to the "partially elliptic case" and, more in general, to "mixed cases."

P4. Consider again the Hamiltonian (1.2). 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 exist formal quasi-periodic solutions.* However, we shall prove that, in general, these series do not exhibit compensations (at least of the algebraic type considered in this paper). In view of this fact it seems natural to *conjecture that such formal power series are divergent.* As "indirect" (i.e. "without looking at the structure of the formal series") motivation for divergence of these series, we mention that, if $\varepsilon\partial_q^2 f_0(q_0)$ is definite, the phase portrait around the N-dimensional H_0-invariant torus $\mathcal{T}_0$ changes drastically as ε changes sign. It would be nice to give a complete "direct" proof of the above conjecture.

6. Several amplifications of the results obtained here are possible. For example one can treat, with essentially no changes, cases with Hamiltonians (or Lagrangians) explicitly depending on time in a periodic or quasi-periodic way (of course, in such a case, one will have to make the natural assumptions on the frequency vector ω). Also, allowing more general families of weighted trees, one could handle the case in which the perturbation is an analytic function of ε (say $\varepsilon H_1 + \varepsilon^2 H_2 + \cdots$).

As mentioned above, the results concerning lower dimensional resonant tori can certainly be improved and, in fact, it should be possible to deal with "general" situations. Also, in the "partially hyperbolic case," the construction of the associated stable/unstable manifolds (the "whiskers" of the tori) should not present new difficulties.

There is lot of room for improvements in the direction of PDE's. For example one can use the methods presented here in order to extend the results of [10] on the existence of quasi-periodic solutions for elliptic systems (assuming, e.g., the principal part of the operator not in a diagonal form). Also, these methods might be used to construct quasi-periodic solutions for non-linear wave equations, etc.

2. A General Set-up for Small Divisor Problems

Let us describe more formally what we mean by compensations in small divisor problems. Given a "nearly-integrable" real-analytic dynamical system (say, as in P1–P3 of Sect. 1), one is interested in finding invariant surfaces on which the flow is conjugated to a linear flow on a (standard) torus or, equivalently, in finding quasi-periodic solutions. Quasi-periodic solutions are described in terms of functions on the standard N-dimensional torus $\mathbb{T}^N \equiv \mathbb{R}^N/(2\pi\mathbb{Z}^N)$ substituting the phase variable $\theta \in \mathbb{T}^N$ with $\omega t = (\omega_1 t, \dots, \omega_N t)$, where $\omega \in \mathbb{R}^N$ is a given (rationally independent) vector and t denotes time[4]. 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 part of step (a)* (see 4 of Sect. 1). Formal quasi-periodic solutions have already been considered in the last century; for a modern discussion

[4] In the PDE case mentioned in 2 of Sect. 1, $t = y$ is a multidimensional independent variable and ω is a matrix; see [10].

of the formal solutions of the models discussed here we refer the reader to Appendix B of [9][5].

The formal solutions we will be dealing with have the form

$$Z \equiv (Z_1, \ldots, Z_d) \sim \sum_{k \geqq 0} Z^k(\theta)\varepsilon^k \, , \qquad \theta \in \mathbb{T}^N, \tag{2.1}$$

where, for each $k \geqq 1$, the vector-valued function Z^k is a real-analytic function over $\mathbb{T}^N$ having an exponentially fast converging Fourier expansion of the form[6]

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

while for $k = 0$, Z^0 is the solution of the unperturbed problem.

The *second part of step (a)* consists in giving Z a representation in terms of *trees*, i.e. a representation of the form:

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

where, intuitively speaking, $\mathcal{T}_*^k$ is a suitable family of trees (with "set of vertices" V and of "order k") taking care of the combinatorics coming out of Taylor's formula and its repeated applications (see Sect. 3); the sum over the indices α attached to each vertex comes from expanding everything in terms of Fourier series (hence the constraint 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[7]; the summands of these sums are split into two factors: $A(T, \alpha, \beta)$, 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[8], $V \equiv V(T)$ denotes the set of vertices of T and the suffix ("order") k refers to the following estimates on cardinalities

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

(for a suitable constant $c_i > 0$); the second sum in (2.3) runs over all possible functions assigning to each vertex $v \in V$ of a rooted tree T an integer vector $\alpha_v \in \mathbb{Z}^N$ with the constraint $\sum_{v \in V} \alpha_v = n$; the third sum runs over a suitable set of (possibly) vector-valued indices taking a finite number of values ($\#B < \infty$); A is a complex vector depending on T, $\{\alpha_v\}_{v \in V}$, $\{\beta_v\}_{v \in V}$ and on the Hamiltonian (or Lagrangian); finally $\gamma_v \in \mathbb{R}$ are *divisors* that are described as follows. Rooted trees can be naturally equipped with a partial order: we say that $v' \leqq v$ if the path joining the

[5] In [9] the formal expansion for P1 is proved; formal solutions for P2–P4 can be proved in a completely similar way.

[6] 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.

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

[8] See any introductory book on graph theory, such as [4], for the standard terminology or Appendix A in [9].

root r of T with v' contains v; (obviously $v' < v$ means $v' \leqq v$ and $v' \neq v$ and $r \geqq v$ $\forall\, v \in V$, i.e. the root r is the first vertex of the rooted tree T). Given a function α, we define

$$\delta_v \equiv \delta_v(T) \equiv \sum_{v' \in V : v' \leqq v} \alpha_{v'} \ . \tag{2.5}$$

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*

$$|\langle n \rangle|^{-1} \leqq \gamma |n|^{\tau} \ , \quad \forall\, n \in \mathbb{Z}^N \backslash \{0\} \tag{2.6}$$

with suitable positive constants γ, τ.[9] Then

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

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

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

then, for all k and n one has

$$a_n^k \leqq a^k e^{-\xi |n|} \ .$$

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[10]

$$\prod_{v \in T} |\gamma_v| \leqq c_3^k \prod_T (1 + |\alpha_v|^b) \tag{2.9}$$

for some $c_3 > 1$ and $b > 0$, then from (2.4), (2.8) and (2.9) it would follow

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

(for suitable $c_5 > 0$) leading to "absolute convergence" (better: "convergence without compensations") of the formal expansion Z. Indeed Siegel's original proof is based on a similar 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 the Fourier series is replaced by Taylor series and α_v ranges over $\mathbb{Z}_+^N$: in such a case $\delta_v \neq \delta_{v'}$ whenever $v > v'$ and Siegel's method [29] yields the estimates

[9] Typically, in dynamical systems, $\langle n \rangle = \omega \cdot n$ (where the dot denotes the standard inner product in $\mathbb{R}^N$) but in other situations (e.g. [10]) the function $\langle \cdot \rangle$ might be more complicated (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 $\omega \in \mathbb{R}^N$ satisfy 2.6 for some γ.

[10] From now on we will adhere to 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 $vv' \in E(T)$.

(2.9).[11] The problem with $\alpha_v \in \mathbb{Z}^N$ 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. [9], Appendix B) that, in general, one has, for arbitrarily large k, subfamilies $\mathcal{F}_{\text{div}} \subset \mathcal{T}_*^k$ and a choice of $\bar{\alpha}$ and $\bar{\beta}$ (depending only on the subfamily) such that, for suitable $\bar{a}, \bar{b} > 0$,

$$\frac{1}{k!} \sum_{T \in \mathcal{F}_{\text{div}}} \Lambda(T, \bar{\alpha}, \bar{\beta}) \prod_{v \in V} \gamma_v \geqq \bar{a}^k k!^{\bar{b}} . \tag{2.10}$$

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[12] $R \subset T$ 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[13] 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 $R_1, \ldots, R_h$ with R_i adjacent[14] 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

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

(where $v \in C$ means $v \in \bigcup_i V(R_i)$). The examples for which (2.10) holds are based on chains with $\bar{\sigma} \sim k$ and $|\langle n \rangle| \sim k^{-1}$. This phenomenon may be counterbalanced by compensations. To be more precise we introduce the notion of "compensable chain." Consider a chain $C = (R_1, \ldots, R_h)$, $(h \geqq 1)$ 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 $u_1 \geqq w_1 > u_2 \geqq \cdots \geqq w_h$; 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 $t \in \mathbb{C}$,

$$\pi_C(t; T, \alpha, \beta) \equiv \prod_{i=1}^{h} \pi_{R_i}(t; T, \alpha, \beta) , \qquad \pi_{R_i}(t) \equiv \left(\prod_{v \in R_i \backslash P_i} \gamma_v \right) \left(\prod_{\substack{v \in P_i \\ v \neq u_i}} \bar{\gamma}_v(t) \right) , \tag{2.11}$$

---

[11] For a detailed discussion, in the present language, of Siegel's methods see Appendix C of [9]; for a different approach see [5].

[12] 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.

[13] 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).

[14] I.e. connected by one edge.

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

$$\bar{\gamma}_v(t) \equiv \left\langle \sum_{\substack{v' \in R_i \\ v' \leqq v}} \alpha_{v'} + tn \right\rangle^{-\sigma_v} \tag{2.12}$$

if $|R_i| = 1$ (hence $u_i = w_i$ and $\alpha_{u_i} = 0$) we set $\pi_{R_i}(t) \equiv 1$.

If $v \in \bigcup_i P_i$, the function $t \to \pi_C(t; T, \alpha, \beta)$ is a meromorphic function of t and since, by definition of resonance, $\sum_{\substack{v' \in R_i \\ v' \leqq v}} \alpha_{v'} \neq 0$, for any $v \neq u_i$ and $v \in P_i$, $\pi_C(t; T, \alpha, \beta)$ is analytic at $t = 0$. Moreover, $\gamma_v = \bar{\gamma}_v(1)$ and

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

We say that *a chain* $C \subset T \in \mathcal{T}_*^k$ (i.e. $C = (R_1, \ldots, R_h)$ 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 a common chain of resonances, and, for each T', there exists a choice of indices $\beta' = \beta'(T')$, such that the function[15]

$$\bar{\pi}_C(t) \equiv \sum_{T' \in \mathcal{F}_C} A(T', \alpha, \beta') \ \pi_C(t; T', \alpha, \beta') \tag{2.13}$$

has a zero in x of order at least $\bar{\sigma} - 1$.

We can now reformulate analytically the result described in the introduction. Consider first the problem P1 of Sect. 1, and observe that quasi–periodic solutions of the Hamilton equations with frequencies ω, i.e. solutions of the form $(x(t), y(t)) = Z(\omega t)$ with $Z^0(\theta) = (\theta, y_0)$ and $Z^k : \theta \in \mathbb{T}^N \to Z^k(\theta) \in \mathbb{R}^{2N}$, satisfy the equations

$$DZ = J\partial H(Z(\theta)) \ , \tag{2.14}$$

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[16] that there exists a unique formal solution $Z \sim \sum_{k \geqq 0} \varepsilon^k Z^k$ of (2.14) with the normalization condition

$$\int_{\mathbb{T}^N} \pi_1 \circ Z^k d\theta = 0 \quad (k \geqq 1) \ , \tag{2.15}$$

where π_1 is the projection onto the first coordinates: $\pi_1(x, y) = x$. The following result then holds.

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

Consider the Lagrangian problem P2 of Sect. 1. Quasi-periodic solutions $x(t) = Z(\omega t) = \sum_{k \geqq 0} \varepsilon^k Z^k(\omega t)$, where now $Z^0(\theta) = \theta$ and $Z^k : \theta \in \mathbb{T}^N \to Z^k(\theta) \in \mathbb{R}^N$,

---

15 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 $R_2, \ldots, R_h$ with $T \backslash C$, and by (possibly) changing the values of the indices β on C.

16 See [28] or Appendix B of [9].

satisfy the equations

$$D\partial_y L\big(Z(\theta), DZ(\theta)\big) = \partial_x L\big(Z(\theta), DZ(\theta)\big), \quad (D \equiv \omega \cdot \partial_\theta) . \tag{2.16}$$

Mimicking the proof of Appendix B in [9] one can easily show that there exists a unique formal solution $Z \sim \sum_{k \geqq 0} \varepsilon^k Z^k$ of (2.16) with the normalization condition as in (2.15) but without the projection π_1. Then, *Theorem* 2.1 *holds also in this case.*

Consider problem P3 of Sect. 1. It is easy to see that (recall the hyperbolicity assumption) there exists a (unique) formal expansion

$$Z \equiv (Z_1, \ldots, Z_d) \sim \sum_{k \geqq 0} Z^k(\theta; \varepsilon)\mu^k \ , \quad \theta \in \mathbb{T}^N \ , \quad d = 2(N + M)$$

such that $t \to Z(\omega t)$ is a formal quasi-periodic solution of the Hamiltonian equations governed by $H_0 + \mu H_1$. Uniqueness is achieved by requiring (2.15), where π_1 denotes again the projection onto the x variable. Then, *Theorem 2.1 holds also in this case.*

Finally consider problem P4 of Sect. 1. Indeed, *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* (1.1)) *is a formal quasi-periodic solution for* (1.2) *and the set $\{Z^0(\theta) : \theta \in \mathbb{Z}^N\}$ coincides with the torus spanned by $y = \omega$, $p = 0$, $q = q_0$, $x = x_0 + \omega t$.* Uniqueness, again, is enforced by the requirement (2.15). 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 $C = (R_1, \ldots, R_h)$ such that if $\mathcal{F}_0$ denotes the family of *all* trees with chain C then

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

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 Sect. 2, $\mathcal{T}_*^k$ is simply the family of all labeled, rooted trees with k vertices[17], which we will denote by $\mathcal{T}^k$. In this case, as is well known (see e.g. [4]), $\#\mathcal{T}^k = k^{k-1}$ 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

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

[17] The basic terminology can be found in any introductory book on graphs or in Appendix A of [9].

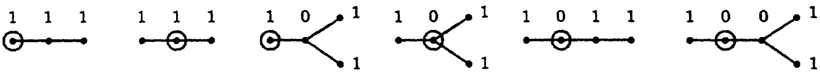

Fig. 1. The elements of $\widetilde{\mathcal{T}}_3$. (The numbers 0, 1 are the values of the index χ for the corresponding vertices and the encircled vertex is the root of the tree).

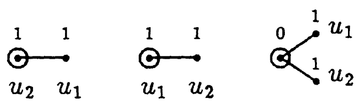

Fig. 2. The elements of $\mathcal{T}_2$ (u_1 and u_2 are the labels of $\mathcal{T}_2$).

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

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

Notice that, in particular, all final vertices (i.e. vertices of degree 1) have weight 1 and that, for any $T \in \widetilde{\mathcal{T}}_k$, $\#V(T) \leqq 2k-1$, as is easy to verify.[18] 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.[19]

The relation between Taylor's formula and trees may be based on the following operation τ (that we now briefly discuss for the case of $\mathcal{T}_k$; for the case $\mathcal{T}^k$ see Appendix B of [9]). 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). Let $\chi \in \{0, 1\}$ and $\ell \geqq 1$, 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

$$T \equiv \tau(T_{h_1}, \ldots, T_{h_\ell}) \equiv \Big(T_{h_1}^0 \cup \cdots \cup T_{h_\ell}^0 \cup \{r\} + \sum_{i=1}^{\ell} r r_i\Big)_r\ ,$$

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

---

[18] 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., [4] 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 $\deg v \geqq 3$, 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 $k_0 \leqq k_1 - 1$ which is the claim. Such an inequality is optimal.

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

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

$$\frac{1}{k!}\sum_{T\in\mathcal{T}_k} F(\widetilde{T}) = \sum_{\chi=0,1}\sum_{\ell=2-\chi}^{k-\chi}\frac{1}{\ell!}\sum_{\substack{h_1+\cdots+h_\ell=k-\chi\\ h_i\geqq 1}}\prod_{i=1}^{\ell}\frac{1}{h_i!}\sum_{T_i\in\mathcal{T}_{h_i}} F\left(\tau(\widetilde{T}_{h_1},\ldots,\widetilde{T}_{h_\ell})\right) .$$

For the proof we refer to [9] (Corollary B.1 of Appendix B).[20]

4. Compensations I (Maximal Hamiltonian Tori)

4.1. Tree expansion. Consider the model introduced in P1 of Sect. 2 and recall that there exists a unique formal solution $Z \sim \sum_{k\geqq 0}\varepsilon^k Z^k$ satisfying (2.14) and (2.15). Denote by $Z^{(1)k}$ the x-component (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^{th} order operator which to a (possibly formal) power series $a \sim \sum a_k\varepsilon^k$ associates its k^{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 (2.14) as

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

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

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

and, for $k = 1$, is $(x,y) = (\theta, y_0)$. We rewrite (4.2) in a more compact way as

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

Notice that while, by (2.15), 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 $\mathbb{T}^N$:

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

Since the average of $[\partial^{(1)}H_1]_{k-1}^{(k-1)}$ vanishes (as it is clear from the second of (4.2)) one can apply to it the operator D^{-1} obtained by inverting the constant coefficient

---

[20] In [9] the case of $\mathcal{T}^k$ is treated (the τ 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.

operator[21] D. Taking the n-Fourier coefficient of (4.4), (4.5) one gets

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

where $D^{(\sigma,\rho)}$ is the vector-valued operator[22]

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

and the $*$ attached to the range of σ means that the following *constraints* have to be satisfied: $\sigma+\rho\in\{2,3\}$ and $\sigma=0 \Longleftrightarrow n=0$ 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 $f: x\in\mathbb{R}^m \to f(x)\in\mathbb{R}$ is a C^∞ function and if $a(\varepsilon)\sim\sum_{s\geqq 1}\varepsilon^s a^{(s)}$ is a $\mathbb{R}^m$-valued (possibly formal) power series, then

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

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 $k\geqq 2$,

$$Z_{nj}^{(\rho)k} = \sum_{\substack{\sigma\in\{0,1,2\}^* \\ \gamma\in\{0,1\}}}\ \sum_{\ell=2-\chi}^{k-\chi}\frac{1}{\ell!}\sum_{\substack{1\leqq h_i\leqq k-1 \\ \Sigma h_i=k-\chi \\ (1\leqq i\leqq\ell)}}\ \sum_{\substack{n_i\in\mathbb{Z}^N \\ \Sigma n_i=n \\ (0\leqq i\leqq\ell)}}\ \sum_{\substack{\rho_1,\dots,\rho_\ell \\ \rho_i\in\{1,2\}}}\ \sum_{\substack{j_1,\dots,j_\ell \\ 1\leqq j_i\leqq N}} \langle n\rangle^{-\sigma} \tag{4.9}$$

$$\left\{\frac{\partial^\ell\, D_j^{(\sigma,\rho)}H_\chi}{\partial z_{j_1}^{(\rho_1)}\cdots\partial z_{j_\ell}^{(\rho_\ell)}}\right\}_{n_0}\prod_{i=1}^{\ell} Z_{n_i j_i}^{(\rho_i)h_i} ,$$

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))

$$Z_{nj_0}^{(\rho_0)k} = \frac{1}{k!}\sum_{T_r\in\mathcal{T}_k}\ \sum_{\substack{\alpha.V\to\alpha_v\in\mathbb{Z}^N \\ \Sigma_v\alpha_v=n}}\ \sum_{\substack{\beta:V\to\beta_v\in B \\ \rho_r=\rho_0}}\ \sum_{\substack{j:V\to j_v\in\{1,\dots,N\} \\ j_r=j_0}}$$

$$\prod_{v\in T_r}\{A_v(T_r,\beta,j)H_{\chi_v}\}_{\alpha_v}\prod_{v\in T_r}\gamma_v , \tag{4.11}$$

---

[21] If f is a (smooth) function on $\mathbb{T}^N$ 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 $g(\theta)=D^{-1}f(\theta) = -i\sum_{n\in\mathbb{Z}^N\setminus\{0\}}\frac{f_n}{\omega\cdot n}\exp(in\cdot\theta)$, where $i=\sqrt{-1}$: the inversion of D introduces the small divisors.

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

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

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

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

$$\Lambda_v(T_r, \beta, j) \equiv D_{j_v}^{(\sigma_v, \rho_v)} \prod_{v' \in \mathcal{N}_v} \partial_{j_{v'}}^{(\rho_{v'})}\ , \quad \mathcal{N}_v \equiv \{v' \in T_r \text{ s.t. } v' < v \text{ and } v' \text{ adjacent to } v\}\ , \tag{4.13}$$

(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

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

where, obviously, V is the set of vertices of T_r.

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 $1, \dots, k-1$. Given $h \geqq 1$, ρ_0, j_0, n, consider the function of (unlabeled) trees $T_r \in \widetilde{\mathcal{T}}_h$, given by

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

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

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

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 Sect. 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 the definition of resonance one has $\delta_u = \delta_{w'} \neq 0$ (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 (2.11))

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

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

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

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

$$\sum_{i=1}^{h} s_{R_i} = \bar\sigma + \rho_{u_1} - \rho_z\ , \tag{4.19}$$

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), (4.19), (2.11) and the definition of Λ (4.14) it follows easily that

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

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 $\hat u > u \geqq w > w'$ (hence $u, w \in R$). Let $\bar R$ be the maximal subtree of R such that $\delta_v(\bar R_u) \neq 0$ $\forall v \neq u$ in $\bar R$. 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 $\bar R$. Hence, $T \in \mathcal{F}'_R$ and if $\rho_{w'} = 1$, $\#\mathcal{F}'_R = \#\bar R$. 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 $\bar R$. As above, $T \in \mathcal{F}''_R$ and if $\sigma_u + \rho_u = 3$, $\#\mathcal{F}'_R = \#\bar R$. 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 a 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 $p \geqq 2$ 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 $v_1v_2,\ldots, v_{p-1}v_p$. We then set

$$\begin{aligned}
\beta'_{v_p} &\equiv \beta_{v_1}\ , \\
\beta'_{v_i} &\equiv \left(\sigma_{v_{i+1}}, 4 - \sigma_{v_{i+1}} - \rho_{v_{i+1}}\right)\ , \quad \text{for } 1 \leqq i \leqq p-1\ , \\
\beta'_v &\equiv \beta_v\ , \qquad\qquad \forall\ v \notin P(u,v)\ .
\end{aligned} \tag{4.20}$$

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

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

Remark 4.1. (i) If $s_R = 0$, one has $\mathcal{F}_R = \{T\}$ 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 (4.20) (the definition which depends on the ordered path $P(u, \bar{u})$) then $\beta'(\beta'(\beta; u, \bar{u}); \bar{u}, u) = \beta$. More in general, it is immediate to check that, for all u, $\bar{u}$, $\bar{v}$ in R, one has

$$\beta'\left(\beta'(\beta; u, \bar{u}), \bar{u}, \bar{v}\right) = \beta'(\beta; u, \bar{v}) , \tag{4.21}$$

which makes transparent the well posedness of the $\beta \to \beta'$ transformation.

(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.

By the above discussion, it is clear that Theorem 2.1 is implied by (4.17). The proof of (4.17) is rather straightforward although a bit technical. We give a detailed proof of (4.17) in Appendix.

5. Compensations II (Maximal Lagrangian Tori)

Consider the model introduced in P2 of Sect. 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 (1.1) and a function $Z(\theta)$ which is the (unique) formal solution $Z \sim \sum_{k \geqq 0} \varepsilon^k Z^k$, with $Z \in \mathbb{R}^N$, of (2.16). 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 $Z^k \in \mathbb{R}^N$ 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 2.16 can then be written as

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

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

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

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

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

and $\{0,1,2\}^*$ is the same set of Eq. (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 Sect. 4.2 and the Appendix for details.

6. Compensations III (Lower Dimensional Tori)

Recall the notations of Sect. 2, P3. Because of the particular form of the Hamiltonian, the formal solution[23] $Z(\theta,\mu)$ is of the form $Z \equiv (X, Q, DX, DQ)$ where, as usual, $D \equiv \omega \cdot \partial_\theta$, $\theta \in \mathbb{T}^N$ and $X \in \mathbb{R}^N$, $Q \in \mathbb{R}^M$. Denote

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

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

$$\Big(-D^2 + (\rho-1)A \Big) Z^{(\rho)k} = \sum_{\chi=0,1} \left[\partial^{(\rho)} H_\chi\right]_{k-\chi}^{(k-1)} , \quad \rho = 1,2 , \qquad (6.1)$$

where the argument of the derivatives of H_χ is $x = \theta$, $q = q_0$. From (6.1) one can see that the average (over $\mathbb{T}^N$) 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 (6.1) we get the analogue of (4.6), namely

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

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

$$\sigma \in \{0,2\}^* \iff \sigma+\rho \in \{2,3\}\ , \qquad n = 0 \implies \sigma = 0\ , \qquad (6.2)$$

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

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

[23] 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.

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

$$N_1 \equiv N \ , \qquad N_2 \equiv M \implies j \in \{1, \dots, N_\rho\} \ . \tag{6.3}$$

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

$$\|(a+A)^{-1}\| \leqq \frac{b}{|\varepsilon|} \tag{6.4}$$

for a suitable constant b depending only on $\partial^2 f_0(q_0)$. We also remark that, by (6.2), $\rho = 2$ 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 case provided we change the following items: N in the fourth sum is replaced (see (6.3)) by N_{ρ_v}; the index set B is defined as

$$B \equiv \Big\{ \beta = (\sigma, \rho) : \ \sigma \in \{0,2\}; \ \rho \in \{1,2\}; \text{ s.t.}$$
$$\sigma + \rho \in \{2,3\} \ , \delta_v = 0 \implies \sigma = 0 \Big\} \ ;$$

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

$$\Lambda_v(T_r, \alpha, \beta, j) \equiv D^{(\rho_v)}_{\delta_v j_v} \prod_{v' \in \mathcal{N}_v} \partial^{(\rho_{v'})}_{j_{v'}} \ . \tag{6.5}$$

Also (4.14) is readily adapted replacing N by N_{ρ_v} and Λ_v with (6.5). 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 Sect. 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 Sect. 4.2)

$$\beta'_{v_p} \equiv \beta_{v_1} \ ,$$
$$\beta'_{v_i} \equiv \beta_{v_{i+1}} \ , \quad \text{for } 1 \leqq i \leqq p-1 \ ,$$
$$\beta'_v \equiv \beta_v \ , \quad \forall \ v \notin P(u,v) \ .$$

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 $\sum_{k \geqq 1} \mu^k Z^k$. It is an easy exercise to adapt the estimates in [9] to the present case and to check how the radius of convergence depends on ε. In fact, observing that, by (6.4), one has

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

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

$$\mu_0 \geqq \text{const} \ \min\{\varepsilon, \gamma^{-2}\} \ .$$

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 (1.2) with

$$f(x,q) = \sum_{s\geqq 1} f_s e^{i(n^{(s)}\cdot x + m^{(s)}\cdot q)}\ , \tag{7.1}$$

where $n^{(s)} \in \mathbb{Z}^N$, $m^{(s)} \in \mathbb{Z}^M$ 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 $f_0 = \sum_{s\geqq 1} f_s e^{im^{(s)}\cdot q}$); then, as for the previous cases, there exists a (unique) formal power series

$$Z \equiv (Z_1,\dots,Z_d) \sim \sum_{k\geqq 0} Z^k(\theta)\varepsilon^k\ , \quad \theta \in \mathbb{T}^N$$

with $d = 2(N+M)$, such that $t \to Z(\omega t)$ is a formal quasi-periodic solution for (1.2) and the set $\{Z^0(\theta) : \theta \in \mathbb{T}^N\}$ coincides with the torus spanned by $y = \omega$, $p = 0$, $q = q_0$, $x = x_0 + \omega t$, where $\omega \in \mathbb{R}^N$ satisfies condition (1.1).

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

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

where $\hat\alpha_v \equiv (\alpha_v, \alpha'_v) \in \mathbb{Z}^{N+M}$; finally $\gamma_v = \langle\delta_v\rangle^{-2} \equiv \left(\omega\cdot \sum_{v'\leqq v}\alpha_{v'}\right)^{-2}$.

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

$$\sum_{s\in I} n^{(s)} = 0 \implies \sum_{s\in I} m^{(s)} = 0 \tag{7.2}$$

$\forall I \subset \mathbb{N}$; in fact in such a case one could repeat word-by-word the arguments in [9]. In general, (7.2) does not hold and 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

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

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 $\omega = (\sqrt{2}, -1)$.

We fix a tree $T \in \widetilde{\mathcal{T}}^k$, with $k = 3h+1$, and a function $\hat\alpha_v : V(T) \to \mathbb{Z}^3$ so that T contains a chain $C = (R_1,\dots,R_h)$ 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 Fig. 3).

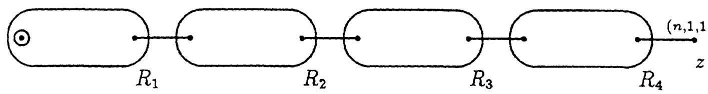

Fig. 3. Divergent contribution: the chain $C = (R_1, R_2, R_3, R_4)$ is made of 4 adjacent resonances with $|R_i| = 3$ (in this symbolic picture only two vertex are drawn).

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 (2.13)) for $h = 1$ is:

$$(\bar{\pi}_C(x))_i \equiv (\bar{\pi}_{R_1}(x))_i \equiv \prod_{v \in V} f_{\hat{\alpha}_v} \sum_{u,w \in R_1} \hat{e}_i \cdot \hat{\alpha}_u \left(\prod_{vv' \in E(R_1)} \hat{\alpha}_v \cdot \hat{\alpha}_{v'} \right) \hat{\alpha}_w \cdot \hat{\alpha}_z \, \pi_{R_1}(x)$$

$$\equiv 2 \sum_{j=1}^{3} (\delta_{i3}\delta_{j3} + x^2 B_{ij}(x)) \hat{\alpha}_{zj} ,$$

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

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

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

$$\bar{\pi}_C(x) = 2^h (A + x^2 B)^h \hat{\alpha}_z , \qquad A_{ij} = \delta_{i3}\delta_{j3} . \tag{7.3}$$

Thus, if $x = \omega \cdot \alpha_z = \sqrt{2}n - 1$, taking the third component of (7.3) (the component for which condition (7.2) is violated) and assuming that $n \geqq h$, one easily checks that

$$\left((A + x^2 B)^h \hat{\alpha}_z \right)_3 \geqq 1 + O\left(\frac{1}{n} \right) ,$$

which implies that C is a non compensable chain.

A. Proof of (4.17)

Using the notations of Sect. 4.2, the possible cases are the following: (i) $s_R = 0$; (ii) $s_R = 1$ and $\sigma_u + \rho_u = 2$, $\rho_{w'} = 1$; (iii) $s_R = 1$ and $\sigma_u + \rho_u = 3$, $\rho_{w'} = 2$; (iv) $s_R = 2$ i.e. $\sigma_u + \rho_u = 3$ and $\rho_{w'} = 1$.

Case (i). Since $t \to \pi_R(t; T, \alpha, \beta)$ is analytic at $t = 0$ [recall Definitions (2.11) and (2.12)], (4.17) follows at once.

Case (ii). Let $T^{(2)} \subset T$ be the unrooted tree with vertices $V(T) \equiv \{v \in T : v \leqq w'\}$, $T^{(0)} \subset T$ be the rooted tree with vertices $V(T^{(0)}) \equiv V(T) \setminus \left(V(R) \cup V(T^{(2)}) \right)$. For any $\bar{u}, \bar{w} \in \overline{R}$, define

$$T(\bar{u}, \bar{w}) \equiv T^{(0)} \cup R \cup T^{(2)} + \hat{u}\bar{u} + \bar{w}w' .$$

Thus, $\mathcal{F}_R \equiv \mathcal{F}'_R = \{T(u,\bar{w}) : \bar{w} \in \overline{R}\}$. When $t = 0$, $\pi_R(t; T(u,\bar{w}),\alpha,\beta)$ does not depend upon $\bar{w}$ and we can write (introducing a new symbol)

$$\pi_R(0; T(u,\bar{w}),\alpha,\beta) \equiv \pi_R(T^{(1)},\alpha,\beta) ,$$

where $T^{(1)} \equiv T^{(0)} \cup R + \hat{u}u$. Furthermore, we can assume that $\chi_{\bar{w}} = 1$: in fact, $\chi_{\bar{w}} = 0$ implies [recall (4.13)] $\Lambda_{\bar{w}}(T(u,\bar{w}),\beta,j)H_0 = 0$ as the operator $\Lambda_{\bar{w}}(T(u,\bar{w}),\beta,j)$ contains $\partial^{(1)}_{j_{w'}}$ (being $\rho_{w'} = 1$) and H_0 is x-independent. Whence

$$\Lambda(T(u,\bar{w}),\alpha,\beta) = \Lambda(T^{(1)},\alpha^{(1)},\beta^{(1)}) \;\; i\alpha_{\bar{w}} \cdot \Lambda(T^{(2)}_{w'},\alpha^{(2)},\beta^{(2)})$$

where $\alpha^{(i)}$ and $\beta^{(i)}$ are the restrictions of α and β to $V(T^{(i)})$. But, since $\sum_{\bar{w}\in\overline{R}} \alpha_{\bar{w}} = 0$, we get:

$$\sum_{T'\in\mathcal{F}_R} \Lambda(T',\alpha,\beta') \; \pi_R(0;T',\alpha,\beta')$$
$$= i\pi_R(T^{(1)},\alpha,\beta)\Lambda(T^{(1)},\alpha^{(1)},\beta^{(1)}) \sum_{\bar{w}\in\overline{R}} \alpha_{\bar{w}} \cdot \Lambda(T^{(2)},\alpha^{(2)},\beta^{(2)}) = 0 ,$$

which proves (4.17) in case (ii).

Case (iii). In this case it is $\mathcal{F}_R \equiv \mathcal{F}''_R = \{T(\bar{u},w) : \bar{u} \in \overline{R}\}$. It will be convenient to keep track of the signs (and the power of $i = \sqrt{-1}$) in the definition of Λ. To do this, let us rewrite the operator $D^{(\sigma,\rho)}$ as follows:

$$D^{(\sigma,\rho)} \equiv \varepsilon_{\sigma,\rho} \; \tilde{D}^{(\sigma,\rho)} ,$$

where

$$\varepsilon_{\sigma,\rho} \equiv (-1)^{1-\sigma\rho} \; i^{-\sigma}, \qquad \tilde{D}^{(\sigma,\rho)} \equiv A^{\sigma-1}\partial^{(4-\sigma-\rho)} .$$

Analogously, we shall denote $\tilde{\Lambda}$ and $\tilde{\Lambda}_v$ the corresponding objects Λ and Λ_v with $D^{(\sigma,\rho)}$ replaced by $\tilde{D}^{(\sigma,\rho)}$. If $T' = T(\bar{u},w)$, $\beta' \equiv (\sigma',\rho') \equiv \beta'(\beta;u,\bar{u})$ (see point (ii) of Remark 4.1), and the value of the index j relative to the root r is kept fixed and equal to j_0, we have

$$\Lambda_{j_0}(T',\alpha,\beta') = i\lambda_0 \cdot \alpha_{\bar{u}} \left(\prod_{v\in R_{\bar{u}}} \varepsilon_{\sigma_v,\rho_v} \right) \lambda(\bar{u}) \cdot \Lambda(T^{(2)}_{w'},\alpha^{(2)},\beta^{(2)}) , \qquad \text{(A.1)}$$

where the vectors λ_0, $\lambda(\bar{u})$ are given by the following formulas[24]:

$$\Lambda'_v(T^{(0)},\beta,j) \equiv \begin{cases} \partial^{(\rho_u)} D^{(\sigma_{\hat{u}},\rho_{\hat{u}})}_{j_{\hat{u}}} \prod_{v'\in\mathcal{N}_{\hat{u}}(T^{(0)})} \partial^{(\rho_{v'})}_{j_{v'}} , & \text{if } v = \hat{u} \\ \Lambda_v(T,\beta,j) , & \text{if } v \in T^{(0)}\backslash\{\hat{u}\} \end{cases} ,$$

$$\lambda_0 \equiv A^{\sigma_u-1} \left(\sum_{\substack{j:V(T^{(0)})\to\{1,\ldots,N\} \\ j_r=j_0}} \prod_{v\in V(T^{(0)})} \{\Lambda'_v(T^{(0)},\beta,j)H_{\chi_v}\}_{\alpha_v} \right) ,$$

[24] Remarks: (a) If T is a rooted tree and $v \in V(T)$, $\mathcal{N}_v(T) = \{v' \in V(T) : v' < v \text{ and } v'v \in E(T)\}$. (b) Note that the operators Λ'_v and Λ''_v are scalar operators if, respectively, $v \neq \hat{u}$ and $v \neq \bar{w}$, while they are vector–valued operators if, respectively, $v = \hat{u}$ and $v = w$ (that is why λ_0 and λ are vectors). (c) $\beta'|_{R_{\bar{u}}}$ and $j|_{R_{\bar{u}}}$ denote the restrictions of, respectively, β' and j to the rooted tree $R_{\bar{u}}$. (d) Note that, since A^{σ_u-1} is a symmetric (real) matrix, if we rewrite λ_0 as $\lambda_0 = A^{\sigma_u-1}\bar{\lambda}_0$ it is $\lambda_0 \cdot \alpha_u = \bar{\lambda}_0 \cdot A^{\sigma_u-1}\alpha_u$.

$$\Lambda''_v(R_{\bar{u}},\beta',j) \equiv \begin{cases} \partial^{(\rho_{w'})}\tilde{D}^{(\sigma_w,\rho_w)}_{j_w} \prod_{v'\in\mathcal{N}_w(R_{\bar{u}})} \partial^{(\rho'_{v'})}_{j_{v'}}, & \text{if } v=w \\ \prod_{v'\in\mathcal{N}_{\bar{u}}(R_{\bar{u}})} \partial^{(\rho'_{v'})}_{j_{v'}}, & \text{if } v=\bar{u} \\ \tilde{\Lambda}_v(R_{\bar{u}},\beta'|_{R_{\bar{u}}},j|_{R_{\bar{u}}}), & \text{if } v\neq\bar{u},w \end{cases},$$

$$\lambda(\bar{u}) \equiv \sum_{j:V(R_{\bar{u}}\setminus\{\bar{u}\})\to\{1,\dots,N\}} \prod_{v\in V(R_{\bar{u}})} \{\Lambda''(R_{\bar{u}},\beta',j)H_{\chi_v}\}_{\alpha_v} .$$

We now claim that

$$\lambda(\bar{u}) = \lambda(\bar{v}) \quad , \quad \forall \bar{u},\bar{v} \in \bar{R} . \tag{A.2}$$

Obviously, it is enough to check (A.2) for $\bar{u},\bar{v}$ adjacent. Let $\bar{\beta}' \equiv \beta'(\bar{\beta};\bar{u},\bar{v})$, with $\bar{\beta} \equiv \beta'(\beta;u,\bar{u})$, and notice that, by (4.21) it is $\bar{\beta}' = \beta'(\beta;u,\bar{v})$. Thus, letting for ease of notation $(\sigma_v,\rho_v) \equiv \bar{\beta}_v$ and $(\sigma'_v,\rho'_v) \equiv \bar{\beta}'_v$, we have (by definition of β')

$$(\sigma'_{\bar{u}},\rho'_{\bar{u}}) = (\sigma_{\bar{v}},4-\sigma_{\bar{v}}-\rho_{\bar{v}}), \quad (\sigma'_{\bar{v}},\rho'_{\bar{v}}) = (\sigma_{\bar{u}},\rho_{\bar{u}}) . \tag{A.3}$$

Fix $\ell \in \{1,\dots,N\}$. Then the ℓ^{th} component of the vector $\lambda(\bar{u})$ has the form:

$$\lambda_\ell(\bar{u}) = \{\partial^{(\rho_{\bar{v}})}g_{\bar{u}}\}_{\alpha_{\bar{u}}} \cdot A^{\sigma_{\bar{v}}-1}\{\partial^{(4-\sigma_{\bar{v}}-\rho_{\bar{v}})}g_{\bar{v}}\}_{\alpha_{\bar{v}}} , \tag{A.4}$$

where, if $z=\bar{u},\bar{v}$, the functions g_z are given by the following formulas. Let $R(\bar{u})$ be the subtree of R rooted at $\bar{u}$ with vertices given by $V(R(\bar{u})) \equiv V(R)\setminus\{v \in R_{\bar{u}} : v \leqq \bar{v}\}$ (the order "$\leqq$" being that of $R_{\bar{u}}$) and let $R(\bar{v})$ be defined analogously (i.e. exchanging the role of $\bar{u}$ and $\bar{v}$); let $G(z)\equiv R(z)\setminus\{z\}$ (in general $G(z)$ is a union of unrooted subtrees), and let

$$d_v \equiv \begin{cases} \partial_\ell^{(\rho_{w'})}, & \text{if } v=w \\ 1, & \text{if } v\neq w ; \end{cases}$$

then (for $z=\bar{u}$ or $\bar{v}$)

$$g_z \equiv \sum_{j:G(z)\to\{1,\dots,N\}} \Big(\prod_{v'\in\mathcal{N}_z(R(z))} d_z\partial^{(\rho_{v'})}_{j_{v'}}H_{\chi_z}\Big) \prod_{v\in G(z)} \Big\{d_v\tilde{\Lambda}_v(R(z),\bar{\beta}|_{R(z)},j|_{R(z)})H_{\chi_v}\Big\}_{\alpha_v} . \tag{A.5}$$

The definition of $\lambda(\bar{v})$ is completely analogous: $\bar{u}$ and $\bar{v}$ are exchanged and $\bar{\beta}$ is replaced by $\bar{\beta}'$; notice, however, that the definition of g_z above depends only on the values of $\bar{\beta}$ on $G(z)$ and $\bar{\beta}_v = \bar{\beta}'_v$ for $v \in G(\bar{u}) \cup G(\bar{v})$, thus in the definitions of $\lambda(\bar{v})$ one may take exactly the definition of g_z given in (A.5). By (A.3) and the symmetry of A one obtains:

$$\begin{aligned} \lambda_\ell(\bar{v}) &= \Big\{\partial^{(\rho'_{\bar{u}})}g_{\bar{v}}\Big\}_{\alpha_{\bar{v}}} \cdot A^{\sigma'_{\bar{u}}-1}\Big\{\partial^{(4-\sigma'_{\bar{u}}-\rho'_{\bar{u}})}g_{\bar{u}}\Big\}_{\alpha_{\bar{u}}} \\ &= \{\partial^{(4-\sigma_{\bar{v}}-\rho_{\bar{v}})}g_{\bar{v}}\}_{\alpha_{\bar{v}}} \cdot A^{\sigma_{\bar{v}}-1}\{\partial^{(\rho_{\bar{v}})}g_{\bar{u}}\}_{\alpha_{\bar{u}}} \\ &= A^{\sigma_{\bar{v}}-1}\{\partial^{(4-\sigma_{\bar{v}}-\rho_{\bar{v}})}g_{\bar{v}}\}_{\alpha_{\bar{v}}} \cdot \{\partial^{(\rho_{\bar{v}})}g_{\bar{u}}\}_{\alpha_{\bar{u}}} \\ &= \lambda_\ell(\bar{u}) ; \end{aligned}$$

and (A.2) is proven. Next we prove that also the function

$$\mu(\bar{u}) \equiv \Big(\prod_{v\in R} \varepsilon_{\sigma_v,\rho_v}\Big)\pi_R(0;T(\bar{u},w),\alpha,\bar{\beta})$$

is independent on $\bar{u} \in \overline{R}$. Again, it is clearly enough to check that $\mu(\bar{u}) = \mu(\bar{v})$ with $\bar{u}$ and $\bar{v}$ adjacent in $\overline{R}$. Since $\sigma_{\bar{u}} + \sigma_{\bar{v}} = \sigma'_{\bar{u}} + \sigma'_{\bar{v}}$, one has

$$\prod_{v\in R} i^{-\sigma_v} = \prod_{v\in R} i^{-\sigma'_v} \; . \tag{A.6}$$

Then, observe that

$$\prod_{v\in P(\bar{v},\bar{u})} (-1)^{(1-\sigma'_v\rho'_v)} = \Big(\prod_{v\in P(\bar{u},\bar{v})} (-1)^{(1-\sigma_v\rho_v)}\Big)(-1)^{\sigma_{\bar{v}}} \tag{A.7}$$

and, since R is a resonance, $\delta_{\bar{u}}(R_{\bar{v}}) = -\delta_{\bar{v}}(R_{\bar{u}})$; thus

$$\langle\delta_{\bar{u}}(R_{\bar{v}})\rangle^{-\sigma'_{\bar{u}}} = \langle -\delta_{\bar{v}}(R_{\bar{u}})\rangle^{-\sigma_{\bar{v}}} = (-1)^{\sigma_{\bar{v}}}\langle\delta_{\bar{v}}(R_{\bar{u}})\rangle^{-\sigma_{\bar{v}}} \; . \tag{A.8}$$

From (A.6), (A.7), (A.8) it follows immediately that μ is constant on R. We are ready to prove (4.17) in case (iii): denoting by λ and μ the common values of, respectively, $\lambda(\bar{u})$ and $\mu(\bar{u})$ for $\bar{u} \in \overline{R}$, we obtain, by (A.1),

$$\begin{aligned}
&\sum_{T'\in\mathcal{F}''_R} A_{j_0}(T',\alpha,\beta')\pi_R(0;T',\alpha,\beta') \\
&\quad \equiv \sum_{\bar{u}\in\overline{R}} A_{j_0}\Big(T(\bar{u},w),\alpha,\beta'(\beta,u,\bar{u})\Big)\pi_R\Big(0;T(\bar{u},w),\alpha,\beta'(\beta,u,\bar{u})\Big) \\
&\quad = i\mu\lambda\cdot A(T^{(2)}_{w'},\alpha^{(2)},\beta^{(2)})\sum_{\bar{u}\in\overline{R}}\lambda_0\cdot\alpha_{\bar{u}} = 0 \; ,
\end{aligned}$$

which, by the analyticity at 0 of $\pi_R(t;T',\alpha,\beta')$ implies (4.17) in case (iii).

Case (iv). In this case $\mathcal{F}_R = \mathcal{F}'_R \cup \mathcal{F}''_R = \{T(\bar{u},\bar{w}) : \bar{u},\bar{w} \in \overline{R}\}$ and $s = 2$. From case (ii) or (iii), it follows immediately that (4.17) holds with $s = 1$ (i.e. the l.h.s. of (4.17) vanishes at $t = 0$). To check (4.17) with $s = 2$ we have to show that the derivative at 0 of the l.h.s. of (4.17) vanishes. By (ii) and (iii) above we see that, if $T(\bar{u},\bar{w}) \in \mathcal{F}_R$, $\bar{\beta} \equiv \beta'(\beta;u,\bar{u})$ and $\bar{\beta}_v \equiv (\sigma_v,\rho_v)$, the j_0 component of $A(T(\bar{u},\bar{w}),\alpha,\bar{\beta})$ may be written as

$$A_{j_0}(T(\bar{u},\bar{w}),\alpha,\bar{\beta}) = -\kappa\;\lambda_0\cdot\alpha_{\bar{u}}\;\alpha_{\bar{w}}\cdot A(T^{(2)}_{w'},\alpha^{(2)},\beta^{(2)})\Big(\prod_{v\in R_{\bar{u}}}\varepsilon_{\sigma_v,\rho_v}\Big) \; , \tag{A.9}$$

where κ is defined by the r.h.s. of (A.4) with g_z defined as in (A.5) but with d_z and d_v replaced by 1. Notice that replacing d by 1 does not affect the independence on $\bar{u}$ of this definition. Notice also that, since $\rho_{w'} = 1$, the d's produce the factor $i\alpha_{\bar{w}}$ in (A.9). Computing the derivative at 0 of $\pi_R(t;T(\bar{u},\bar{w}),\alpha,\bar{\beta})$, one finds:

$$\frac{d}{dt}\bigg|_{t=0}\pi_R(t;T(\bar{u},\bar{w}),\alpha,\bar{\beta}) = \begin{cases} \pi_R(0;T(\bar{u},\bar{w}),\alpha,\bar{\beta})\displaystyle\sum_{\substack{v\in P(\bar{u},\bar{w})\\ v\neq\bar{u}}}(-\sigma_v)\langle\delta_v(R_{\bar{u}})\rangle^{-1}, & \text{if } \bar{u}\neq\bar{w} \\ 0\,, & \text{if } \bar{u}=\bar{w}\,. \end{cases}$$

Thus, using the notations of case (iii) and denoting $\beta'_v(\beta; u, \bar{u}) \equiv (\sigma_v(u, \bar{u}),\ \rho_v(u, \bar{u}))$,

$$\begin{aligned} \frac{d}{dt}\bigg|_{t=0} & \sum_{T' \in \mathcal{F}_R} \Lambda_{j_0}(T', \alpha, \beta') \pi_R(t; T', \alpha, \beta') \\ &\equiv \frac{d}{dt}\bigg|_{t=0} \sum_{\bar{u}, \bar{w} \in \bar{R}} \Lambda_{j_0}(T(\bar{u}, \bar{w}), \alpha, \bar{\beta})\ \pi_R(t; T(\bar{u}, \bar{w}), \alpha, \bar{\beta}) \\ &= \kappa\mu \sum_{\bar{w} \in \bar{R}} \Lambda(T_{w'}^{(2)}, \alpha^{(2)}, \beta^{(2)}) \cdot \alpha_{\bar{w}}\ \ v(\bar{w})\ , \end{aligned} \tag{A.10}$$

where

$$v(\bar{w}) \equiv \sum_{\substack{\bar{u} \in \bar{R} \\ \bar{u} \neq \bar{w}}} \lambda_0 \cdot \alpha_{\bar{u}} \sum_{\substack{v \in P(\bar{u}, \bar{w}) \\ v \neq \bar{u}}} \sigma_v(u, \bar{u}) \langle \delta_v(R_{\bar{u}}) \rangle^{-1}\ .$$

We now claim that $v(\bar{w}) \equiv v$ is independent of $\bar{w}$. To check the claim one may again look at adjacent $\bar{w}$ and $\bar{w}'$. The removal of the edge $\bar{w}\bar{w}'$ disconnects $\bar{R}$ in two subtrees: one, which we denote S, containing $\bar{w}$ and another, S', containing $\bar{w}'$. Observe that by the definition of the map $\beta \to \beta'$ one has:

$$\sigma_{\bar{w}'}(u, \bar{u}) = \sigma_{\bar{w}'}(u, \bar{w})\ ,\ \ \forall \bar{u} \in S\ ; \sigma_{\bar{w}'}(u, \bar{w}) = \sigma_{\bar{w}}(u, \bar{w}')\ ;$$

moreover (by definition of resonance) $\delta_{\bar{w}'}(R_{\bar{w}}) = -\delta_{\bar{w}}(R_{\bar{w}'})$. Thus,

$$\begin{aligned} v(\bar{w}) - v(\bar{w}') &= -\sum_{\substack{\bar{u} \in S \\ \bar{u} \neq \bar{w}}} \lambda_0 \cdot \alpha_{\bar{u}}\ \sigma_{\bar{w}'}(u, \bar{u}) \langle \delta_{\bar{w}'}(R_{\bar{u}}) \rangle^{-1} + \sum_{\substack{\bar{u} \in S' \\ \bar{u} \neq \bar{w}'}} \lambda_0 \cdot \alpha_{\bar{u}}\ \sigma_{\bar{w}}(u, \bar{u}) \langle \delta_{\bar{w}}(R_{\bar{u}}) \rangle^{-1} \\ &\quad + \lambda_0 \cdot \alpha_{\bar{w}'}\ \sigma_{\bar{w}}(u, \bar{w}') \langle \delta_{\bar{w}}(R_{\bar{w}'}) \rangle^{-1} - \lambda_0 \cdot \alpha_{\bar{w}}\ \sigma_{\bar{w}'}(u, \bar{w}) \langle \delta_{\bar{w}'}(R_{\bar{w}}) \rangle^{-1} \\ &= -\sum_{\substack{\bar{u} \in S \\ \bar{u} \neq \bar{w}}} \lambda_0 \cdot \alpha_{\bar{u}}\ \sigma_{\bar{w}'}(u, \bar{w}) \langle \delta_{\bar{w}'}(R_{\bar{w}}) \rangle^{-1} + \sum_{\substack{\bar{u} \in S' \\ \bar{u} \neq \bar{w}'}} \lambda_0 \cdot \alpha_{\bar{u}}\ \sigma_{\bar{w}}(u, \bar{w}') \langle \delta_{\bar{w}}(R_{\bar{w}'}) \rangle^{-1} \\ &\quad + \lambda_0 \cdot \alpha_{\bar{w}'}\ \sigma_{\bar{w}}(u, \bar{w}') \langle \delta_{\bar{w}}(R_{\bar{w}'}) \rangle^{-1} - \lambda_0 \cdot \alpha_{\bar{w}}\ \sigma_{\bar{w}'}(u, \bar{w}) \langle \delta_{\bar{w}'}(R_{\bar{w}}) \rangle^{-1} \\ &= \Big(\sum_{\bar{u} \in \bar{R}} \lambda_0 \cdot \alpha_{\bar{u}} \Big)\ \sigma_{\bar{w}}(u, \bar{w}') \langle \delta_{\bar{w}}(R_{\bar{w}'}) \rangle^{-1} = 0\ . \end{aligned}$$

At this point, the proof of (4.17) in this last case (iv) follows at once from (A.10).

References

1. Arnold, V.I. : Geometrical Methods in the Theory of Ordinary Differential Equations, A series of comprehensive studies in Mathematics. **250**, Berlin, Heidelberg, New York: Springer 1988 (Second Edition)
2. Arnold, V.I. (ed.): Encyclopaedia of Mathematical Sciences, Dynamical Systems. Vol. **3**, Berlin, Heidelberg, New York: Springer 1988
3. Arnold, V.I.: Instability of dynamical systems with several degrees of freedom. Sov. Math. Dokl. **5**, 581–585 (1964)
4. Bollobas, B.: Graph Theory. Graduate text in Math. **63**, Berlin, Heidelberg, New York: Springer, 1979
5. Bruno, 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)

6. Celletti, A., Chierchia, L.: Construction of analytic KAM surfaces and effective stability bounds. Commun. Math. Phys. **118**, 119–161 (1988)
7. Celletti, A., Chierchia, L.: A constructive theory of Lagrangian invariant tori and computer assisted applications. Dynamics Reported **4**, 60–129 (1995)
8. Chierchia, L.: Arnold Instability for Nearly-Integrable Analytic Hamiltonian Systems. To appear in the Proceedings of the Conference "Local and Variational Methods in the Study of Hamiltonian Systems," (October 1994, Trieste), Edt. A. Ambrosetti, G.F. Dell'Antonio
9. Chierchia, L., Falcolini, C.: A direct proof of a theorem by Kolmogorov in Hamiltonian systems. Annali Sc. Norm. Super. Pisa, Cl. Sci. Serie IV, Vol. **XXI** Fasc. 4 (1994)
10. Chierchia, L., Falcolini, C.: A note on quasi-periodic solutions of some elliptic systems. To appear in ZAMP
11. Chierchia, L., Gallavotti, G.: Drift and Diffusion in phase space. Ann. Inst. Henri Poincaré (Physique Théorique) **60**, n^{o} 1, 1–144 (1994)
12. Craig, W, Wayne, C. E.: Periodic solutions of nonlinear waves equations. Commun. Pure and Applied Math., Vol. **XLVI** (1993)
13. Eliasson, L.H.: Absolutely convergent series expansions for quasi periodic motions. Reports Department of Math., Univ. of Stockholm, Sweden, No. 2, 1–31 (1988)
14. Eliasson, L.H.: Hamiltonian systems with linear normal form near an invariant torus. In "Nonlinear Dynamics", G. Turchetti (ed.) Singapore: World Scientific, 1989
15. Eliasson, L.H.: Generalization of an estimate of small divisors by Siegel. In: "Analysis, et cetera", P.H. Rabinowitz, E. Zehnder (eds.), New York: Academic Press, 1990
16. Eliasson, L.H.: Perturbations of Stable Invariant Tori for Hamiltonian Systems. Annali Sc. Norm. Super. Pisa, Cl. Sci., IV Ser. **15**, 115–147 (1988)
17. 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 Math. Phys. **6**, 343–411 (1994)
18. Gallavotti, G.: Twistless KAM tori. Commun. Math. Phys. **164**, 145–156 (1994)
19. Gallavotti, G.: Invariant tori: A field theoretic point of view on Eliasson's work. Preprint (1993)
20. Gallavotti, G., Gentile, G.: Majorant series for the KAM theorem. To appear in Ergodic Theory and Dynamical Systems
21. Gentile, G.: A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable hamiltonian systems. To appear in Forum Mathematicum.
22. Gentile, G.: Whiskered tori with prefixed frequencies and Lyapunov spectrum. Preprint (1994)
23. Gentile, G., Mastropietro, V.: Convergence of the Lindstedt series for KAM tori. Preprint (1995)
24. Graff, S.M.: On the conservation of hyperbolic invariant tori for Hamiltonian systems. J. Diff. Equations **15**, 1–69 (1974)
25. Kuksin, S.: Nearly integrable infinite-dimensional Hamiltonian systems. Lectures Notes in Mathematics **1556**, Berlin, Heidelberg, New York: Springer, 1993
26. Moser, J.: A stability theorem for minimal foliations on a torus. Ergod. Th. Dynam. Sys. **8**, 251–281 (1988)
27. Pöschel, J.: On Elliptic Lower Dimensional Tori in Hamiltonian Systems. Math. Z. **202**, 559–608 (1989)
28. Poincaré, H.: Les méthodes nouvelles de la mécanique céleste. Vols. **1–3**. Paris: Gauthier–Villars. (1892/1893/1899)
29. Siegel, C.L.: Iterations of analytic functions. Ann. of Math. **43**, No. 4, 607–612 (1942)
30. Salamon, D., Zehnder, E.: KAM theory in configuration space. Comm. Math. Helv. **64**, 84 (1988)
31. Treshev, D.V.: The fracture mechanism of resonant tori of Hamiltonian systems. Mat. Sb. **180** n.10, 1325–1346 (1989); english translation in Math. USSR–Sb.
32. Wayne, C.E.: Periodic and quasi–periodic solutions of nonlinear wave equations via KAM theory. Commun. Math. Phys. **127**, 479–528 (1990)

Communicated by J.-P. Eckmann