Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics

A review with some applications

G. Gentile†, V. Mastropietro††

† IHES, 35 Route de Chartres, 91440 Bures sur Yvette, France.

†† Dipartimento di Matematica, Università di Roma II "Tor Vergata", 00133 Roma, Italia.

Abstract: *This paper consists in a unified exposition of methods and techniques of the renormalization group approach to quantum field theory applied to classical mechanics, and in a review of results: (1) a proof of the KAM theorem, by studing the perturbative expansion (Lindstedt series) for the formal solution of the equations of motion; (2) a proof of a conjecture by Gallavotti about the renormalizability of isochronous hamiltonians, i.e. the possibility to add a term depending only on the actions in a hamiltonian function not verifying the anisochrony condition so that the resulting hamiltonian is integrable. Such results were obtained first by Eliasson; however the difficulties arising in the study of the perturbative series are very similar to the problems which one has to deal with in quantum field theory, so that the use the methods which have been envisaged and developed in the last twenty years exactly in order to solve them allows us to obtain unified proofs, both conceptually and technically. In the final part of the review, the original work of Eliasson is analyzed and exposed in detail; its connection with other proofs of the KAM theorem based on his method is elucidated.*

Keywords: *Classical mechanics, KAM theorem, quantum field theory, renormalization group, multiscale analysis, tree expansion, counterterms, cancellations*

> *Solomon saith:* There is no new thing upon the earth. *So that as Plato had an imagination,* that all knowledge was but remembrance*; so Solomon giveth his sentence,* that all novelty is but oblivion.
>
> Francis Bacon, [Ba]

1. Introduction

Quasiperiodic solutions of the equations of motion obtained under small perturbations of integrable hamiltonian systems can be given formal perturbative expansions, which are known as *Lindstedt series*, [L]. The existence problem for such series was studied by

Poincaré, [P], Vol. II, Ch. XIII, who proved that, *in general*, the series diverge (*Poincaré's triviality theorem*); he was not able to exclude that under suitable hypotheses on the unperturbed solution the convergence would occur, although he found such a possibility very unlikely and advanced the conjecture that the series could never converge, (see in particular Vol. II, Ch. XIII, §149). Doubts on Poincaré's conjecture were raised from Weierstrass, [W].

On the contrary the KAM theorem, stating the conservation of quasiperiodic motions for a large class of hamiltonian systems and initial data, proves the convergence of the Lindstedt series.

The original proof is due to Kolmogorov, [K], and to Arnol′d, [A1], [A2], who extended the theorem to cover cases relevant for the three (or N) body problem. It is based on a convergent iterative technique, which proves the existence of a canonical transformation conjugating, analytically in the perturbative parameter, the motion to a simple rotation on the torus: the tecnique of proof easily implies the convergence of the Lindstedt series.

Moser proved the corresponding theorem in the case of non analytic (small enough) perturbation: an extension requiring deep new ideas and tools (like Nash's implicit function theorem, [N]). In [M1] the theorem was proven for C^p area-preserving mappings in the plane, with $p \geq 333$, and improvements were provided later, [R2], [M5]. The extension of the theorem for hamiltonian systems to the differentiable case ($p > 2 + 2(\tau + 1) > 2(\ell + 1)$, where $\tau > \ell - 1$ is the diophantine constant in (1.3) below and ℓ is the number of degrees of freedom) was discussed in [M2], [M4].

Only in recent times, a new proof of the convergence was provided by Eliasson, [E1], [E3], by studying directly the Lindstedt series and proving the existence of cancellations to all perturbative orders. However Eliasson's work has not enjoyed a wide diffusion and understanding as it would have deserved. This led to a sequence of works presenting technically and conceptually simplified proofs. In [G7], [GM1] and [CF1] simplified models are studied, following the attempt by Thirring, [T], Vol. 1, Ch. 3, §3.6, to find a model which allows us to make easier to explain the KAM theorem.

Our starting point is the analogy of the Lindstedt series with the perturbative series in constructive quantum field theory, pointed out in [FT] and [G9]. It was carried out to a deeper extent in [GGM], where a quantum field model is explicitly exhibited whose Feynman's graphs are exactly the same diagrams which correspond to a natural graphical representation of the Lindstedt series. It is then natural to prove the convergence of the Lindstedt series by exploiting techniques usual in the renormalization group approach to the quantum field theory, like the multiscale decomposition of the propagators, the tree expansion and the introduction of counterterms whose value has to be uniquely fixed in order to make the problem soluble, (see [G5] for a review). A similar analysis is introduced also in [G8], [Ge1] and [Ge2] in order to study the persistence of the stable and unstable manifolds (*whiskers*) of low dimensional tori in a class of almost integrable systems; in particular, in [Ge2], the introduction of suitable counterterms in the hamiltonian is found to be a useful device in order to simplify in a relevant way the proof of existence of low dimensional tori.

In this work we extend the ideas and results contained in [G7], [GG], [GM1] and [GM2], so recovering completely the (analytic) KAM theorem. This means that the only assumptions on the hamiltonian function will be: (1) the anisochrony condition, and (2) the diophantine property.[1]

There are interesting technical advantages in using the methods of quantum field theory, with respect to [E1], [E2], [E3], [CF1] and [CF2]. In fact some problems like the *overlapping divergences* and the approximate cancellations of the resonances are automatically bypassed, so that there is no need to distinguish between contributions which really require a bound improvement and contributions which can only apparently raise problems, (as it has been done first in [E1], where one is led to define several kinds of resonances, such that only the "critical" ones are "dangerous"). This is a standard feature of the techniques used in quantum field theory, and shows that such an approach is "very natural" in order to attack the problem. These and other technical improvements are discussed further and with more details in the next sections.

Besides the mathematical interest and motivation, the ultimate hope is that the newly introduced methods will allow us to solve open problems of the "KAM theory". In [GM2] for instance we perform a resummation in the Lindstedt series, which is used in [GGM] in the heuristic analysis of the universality of the breakdown phenomenon for KAM invariant tori via the study of the singularities of the Lindstedt series.

A further application is provided in the present paper: if the free hamiltonian is isochronous (*i.e.* it describes a system of harmonic oscillators) and its frequencies satisfy the diophantine property, then the Lindstedt series is not convergent. In [G2] Gallavotti advanced the conjecture that in such a case, if a suitable "counterterm", analytic in the perturbative parameter and depending only on the action variables, is added to the hamiltonian, then the modified hamiltonian is integrable. Here we show that the validity of such a conjecture is a byproduct of our method by proving that the Lindstedt series for the modified model is convergent.

In [R1], Rüssmann proved that, if the counterterms make the equations of motion *formally* soluble, then there there exists an analytic solution, (*i.e.* if the series can be formally defined, then it converges), a result reproduced by Gallavotti, in [G2], with a different formalism; however the problem of the formal solubility was left unsolved. A partial result about the conjecture is implied by the papers of Dinaburg and Sinai, [DS], and Rüssmann, [R3], who proved the conjecture to hold for special interactions of the form $\varepsilon\vec{A}\cdot\vec{f}(\vec{\alpha})$ (in the notations of this paper, see §1.1), by a method of variation of constants, inherited from [M3]: in fact they studied the one-dimensional Schrödinger equation with a quasi-periodic potential, but the problem can be shown to be equivalent to a hamiltonian

[1] In [CF2] the general case of the KAM theorem is discussed, by describing explicitly the cancellations, sketching the strategy one has to follow to complete the proof (with the hint that it can be adapted from [CF1]).

problem in classical mechanics of oscillators interacting via a potential linear in the action variables, (see also [G6]).

Gallavotti's conjecture has been first proven in the general case also by Eliasson in a work, [E2], which apparently had the same reception problems of the quoted ones. Another proof with different tools is presented in [EV], where the notions of *mould* and *arborification*, introduced by Ecalle, are used in order to prove the analyticity of the "correction" of any resonant local analytic field, (the hamiltonian case being included), under the so called *Bryuno's diophantine condition*, which is weaker than the property (1.3).

In the remaining part of the introduction, we give a more formal statement of the results which will be proven in next sections, and introduce the basic notations.

1.1. The hamiltonian function is

$$\mathcal{H}(\vec{\alpha},\vec{A}) = \mathcal{H}_0(\vec{A}) + f(\vec{\alpha},\vec{A};\varepsilon)\ , \tag{1.1}$$

where (1) $\mathcal{H}_0(\vec{A})$ and $f(\vec{\alpha},\vec{A};\varepsilon)$ are analytic functions in $\vec{A}$, in a domain $D\subset\mathbb{R}^\ell$, and in $\vec{\alpha}$, for $\mathrm{Re}\,\vec{\alpha}\in\mathbb{T}^\ell$ and $|\mathrm{Im}\,\vec{\alpha}|<\xi$ for some positive constant ξ, and (3) $f(\vec{\alpha},\vec{A};\varepsilon)$ is analytical in ε, for $|\varepsilon|\le\varepsilon_1$, and divisible by ε, (see (A2.1) below).

The *free hamiltonian* $\mathcal{H}_0(\vec{A})$ satisfies the *anisochrony condition* (or *twist condition*)

$$\det\Big(\partial_{A_i}\partial_{A_j}\mathcal{H}_0(\vec{A}_0)\Big)\neq 0\ , \tag{1.2}$$

where $\vec{A}_0\in D$, and the vector $\vec{\omega}_0\equiv\partial_{\vec{A}}\mathcal{H}_0(\vec{A}_0)$ verifies the *diophantine property* with diophantine constants $C_0,\tau>0$; this means that

$$C_0|\vec{\omega}_0\cdot\vec{\nu}|\ge|\vec{\nu}|^{-\tau}\ ,\qquad\qquad \vec{0}\neq\vec{\nu}\in\mathbb{Z}^\ell\ , \tag{1.3}$$

where $\vec{\omega}_0\cdot\vec{\nu}=\sum_{j=1}^\ell\omega_{0j}\nu_j$ is the scalar product in $\mathbb{R}^\ell$. It is easy to see that the *diophantine vectors* have full measure in $\mathbb{R}^\ell$ if τ is fixed $\tau>\ell-1$.

If J_j^{-1}, $j=1,\ldots,\ell$, are the eigenvalues of the matrix $T(\vec{A}_0)\equiv\partial_{\vec{A}}\partial_{\vec{A}}\mathcal{H}_0(\vec{A}_0)$, such that $0<J_j^{-1}<\infty$, $\forall\ j=1,\ldots,\ell$, (because of the twist condition), we define

$$J_m=\min_{j=1,\ldots,\ell}J_j\ ,\qquad J_M=\max_{j=1,\ldots,\ell}J_j\ , \tag{1.4}$$

(note that, if $\mathcal{H}_0(\vec{A})$ is quadratic in the actions, then the J_j's are the principal momenta of inertia), and, because of the analyticity assumption on the interaction potential, we can write

$$f(\vec{\alpha},\vec{A};\varepsilon)=\sum_{\vec{\nu}\in\mathbb{Z}^\ell}e^{i\vec{\nu}\cdot\vec{\alpha}}\,f_{\vec{\nu}}(\vec{A};\varepsilon)\ ,\qquad\qquad |f_{\vec{\nu}}(\vec{A};\varepsilon)|\le F\,e^{-\xi|\vec{\nu}|}\ , \tag{1.5}$$

with ξ defined after (1.1) and

$$F=\sup_{|\varepsilon|\le\varepsilon_1}\ \sup_{\vec{A}\in W(\vec{A}_0,\rho_0)}\ \sup_{\vec{\alpha}\in\mathbb{T}^\ell}\{f(\vec{\alpha},\vec{A};\varepsilon)\}\ . \tag{1.6}$$

where $W(\vec{A}_0,\rho_0)=\{\vec{A}\in\mathbb{R}^\ell\ :\ |\vec{A}-\vec{A}_0|\le\rho_0\}\subset D$. Finally we define

$$E_0=\sup_{\vec{A}\in W(\vec{A}_0,\rho_0)}\{\mathcal{H}_0(\vec{A})\}\,,\qquad E=\max\{F\varepsilon_1^{-1}\,,\,E_0\}\,,\tag{1.7}$$

and introduce the dimensionless frequency $\vec{\omega}=C_0\vec{\omega}_0$, and the dimensionless analyticity radius $\rho=(C_0/J_m)\rho_0$, and set $r=\min\{1,\rho^2\}$.

In this paper we give a proof of the following result, by verifying the convergence of the Lindstedt series.

1.2. Theorem (Kolmogorov-Arnol′d-Moser's theorem). *The hamiltonian model (1.1), with $\mathcal{H}_0(\vec{A})$ verifying the anisochrony condition (1.2) and $\vec{\omega}_0=\partial_{\vec{A}}\mathcal{H}_0(\vec{A}_0)$ satisfying the diophantine property (1.3), admits an ε-analytic family of motions starting at $\vec{\alpha}(0)=\vec{0}$ and having the form*

$$\vec{A}(t)=\vec{A}_0+\vec{H}(\vec{A}_0,\vec{\omega}_0t;\varepsilon)\,,\qquad\qquad \vec{\alpha}(t)=\vec{\omega}_0t+\vec{h}(\vec{A}_0,\vec{\omega}_0t;\varepsilon)\,,\tag{1.8}$$

where:
(1) $\vec{H}(\vec{A},\vec{\psi};\varepsilon)$, $\vec{h}(\vec{A},\vec{\psi};\varepsilon)$ are analytic in $\vec{\psi}$ with $\mathrm{Re}\vec{\psi}\in\mathbb{T}^\ell$, and $|\mathrm{Im}\vec{\psi}|<\xi$, and in $\vec{A}\in W(\vec{A}_0,\rho_0)$;
(2) $\vec{h}(\vec{A},\vec{\psi};\varepsilon)$ has vanishing average in $\mathbb{T}^\ell$;
(3) $\vec{H}(\vec{A}_0,\vec{\psi};\varepsilon)$ and $\vec{h}(\vec{A}_0,\vec{\psi};\varepsilon)$ are analytic for $|\varepsilon|<\varepsilon_0$ with a suitable ε_0 close to 0:

$$\varepsilon_0=\eta_0\,[J_M\,J_m^{-2}\,C_0^2\,E\,r^{-1}]^{-2}\,,\tag{1.9}$$

with η_0 being a dimensionless quantity depending only on ℓ, ξ and τ. This means that the set $(\vec{A}(t),\vec{\alpha}(t))$ described by replacing $\vec{\omega}_0t$ in (1.8) with $\vec{\psi}\in\mathbb{T}^\ell$ is, for ε small enough, an analytic invariant torus for (1.1), which is run quasi periodically with angular velocity vector $\vec{\omega}_0$, and coincides, for $\varepsilon=0$, with the torus $\vec{A}(t)=\vec{A}_0$, $\vec{\alpha}(t)=\vec{\omega}_0t\in\mathbb{T}^\ell$.

The condition in item (2) is imposed in order to simplify the analysis of the equations of motion, and of the recursive relations defining them.

The dependence on ℓ in (1.9) is a factorial to some negative power, (see comments between (5.8) and (5.9) in §5), similar to the estimates derived in the usual KAM proofs, which is of the form $(\ell!)^{-a}$, for some constant $a>0$, (see, *e.g.* , [G3]).

1.3. In a paper by Gallavotti, [G2], the question was raised if, given a perturbation of a free *isochronous* hamiltonian not satisfying the anisochrony condition, it is possible to "renormalize" it by adding to it a function analytic in the perturbative parameter, in such a way that the resulting hamiltonian turns out to be integrable.

More formally let us consider a hamiltonian of the form (1.1) with $\mathcal{H}_0'\neq 0$, (if we denote by $\mathcal{H}_0'$ the derivative of $\mathcal{H}_0$ with respect to its argument) and $\vec{\omega}_0$ being a diophantine vector, (see (1.3)). All the functions are supposed to be analytic in their arguments in suitable domains, as in §1.1.

We know, from trivial examples in [G1], [G4], (see (1.13) below), or from the *genericity theorem*, [S2], [MS], that in such a case the hamiltonian $\mathcal{H}$ is not integrable in general. Nevertheless we can ask ourselves if it is possible to add to $\mathcal{H}$ a function N_f depending analytically on $\vec{A}$ and ε (in a suitable domain to be found), $\vec{\psi}$-independent and such that the hamiltonian $\mathcal{H} - N_f$ is integrable.

If it is so, by analogy with the renomalization problem in quantum field theory and following [G2], we can define "renormalized hamiltonian" the function $\mathcal{H} - N_f$, and we can set

$$: f : = f - N_f \ ,$$

by calling the operator $::$ so introduced the "Wick ordering with respect to $\mathcal{H}_0$".

Let us take the hamiltonian function as in (1.1), and choose the "free hamiltonian" of the form $\mathcal{H}_0(\vec{A}) \equiv \vec{\omega}_0 \cdot \vec{A}$, and $\vec{\omega}_0$ satisfying the diophantine property.

Then we consider the "renormalized" hamiltonian

$$\mathcal{H}(\vec{\alpha}, \vec{A}) = \vec{\omega}_0 \cdot \vec{A} + f(\vec{\alpha}, \vec{A}; \varepsilon) - N_f(\vec{A}; \varepsilon) \ , \tag{1.10}$$

and we look for a function $N_f(\vec{A};\varepsilon)$ such that the system described by the hamiltonian (1.10) is integrable, *i.e.* admits quasiperiodic motions. We will call *counterterm* the function $N_f(\vec{A};\varepsilon)$. This can be regarded as a "control theory theorem" and it might to have some applications in the costruction of stabilizing devices: for instance one is interested to have persistence of tori in stellarators used for plasma confinement via the application of toroidal magnetic fields, (see, *e.g.* , [SHS], [HC]). Of course we expect that there are perturbations in the motion of the particles which should generate chaotic motions, but the following results shows that by tuning in a suitable way proper magnetic fields the motion could remain stable.

The following result, conjectured in [G2] and proven in [E2], [EV], will be shown to fit into our general scheme and proven again.

1.4. THEOREM. *Given the hamiltonian (1.1), where $\mathcal{H}_0(\vec{A}) = \vec{\omega}_0 \cdot \vec{A}$, with the vector $\vec{\omega}_0$ verifying the condition (1.3), and $f(\vec{\alpha}, \vec{A}; \varepsilon)$ is analytic in ε, with $|\varepsilon| \le \varepsilon_1$, for some positive constant ε_1, and*

$$f(\vec{\alpha}, \vec{A}; \varepsilon) = \sum_{\vec{\nu} \in \mathbb{Z}^\ell} e^{i\vec{\nu}\cdot\vec{\alpha}} f_{\vec{\nu}}(\vec{A};\varepsilon) \ , \qquad \sup_{|\varepsilon|\le\varepsilon_1} \sup_{\vec{A}\in D} |f_{\vec{\nu}}(\vec{A};\varepsilon)| < F\, e^{-\xi|\vec{\nu}|} \ ,$$

then it is possible to fix a function $N_f(\vec{A};\varepsilon)$ such that the hamiltonian (1.10) admits a family of motions starting at $\vec{\alpha}(0) = \vec{0}$ and having the form

$$\vec{A}(t) = \vec{A}_0 + \vec{H}(\vec{A}_0, \vec{\omega}_0 t; \varepsilon) \ , \qquad \vec{\alpha}(t) = \vec{\omega}_0 t + \vec{h}(\vec{A}_0, \vec{\omega}_0 t; \varepsilon) \ , \tag{1.11}$$

where:
(1) $\vec{H}(\vec{A}, \vec{\psi}; \varepsilon)$ and $\vec{h}(\vec{A}, \vec{\psi}; \varepsilon)$ are analytic in $\vec{\psi}$ with $\mathrm{Re}\vec{\psi} \in \mathbb{T}^l$, and $|\mathrm{Im}\vec{\psi}| < \xi$, and in

$\vec{A} \in W(\vec{A}_0, \rho_0) \ \equiv \ \{\vec{A} \in D : |\vec{A} - \vec{A}_0| < \rho_0\}$;
(2) $\vec{H}(\vec{A}_0, \vec{\omega}_0 t; \varepsilon)$ and $\vec{h}(\vec{A}_0, \vec{\omega}_0 t; \varepsilon)$ have vanishing average;
(3) $\vec{H}(\vec{A}_0, \vec{\omega}_0 t; \varepsilon)$ and $\vec{h}(\vec{A}_0, \vec{\omega}_0 t; \varepsilon)$ are analytic in ε, for $|\varepsilon| < \varepsilon_0$, with a suitable ε_0 close to zero:

$$\varepsilon_0 = \eta_0 \left[C_0 \, F \, \rho_0^{-1} \, \varepsilon_1^{-1} \right]^{-1} , \tag{1.12}$$

with η_0 a dimensionless constant depending only on ℓ, ξ and τ;
(4) $N_f(\vec{A}; \varepsilon)$ is analytic in $\vec{A} \in W(\vec{A}_0, \rho_0)$ and in ε, for $|\varepsilon| < \varepsilon_0$.

In general the analyticity properties of $N_f(\vec{A}; \varepsilon)$ are related to those of $f(\vec{\alpha}, \vec{A}; \varepsilon)$. For example, let us consider the Birckhoff hamiltonian, for $\ell = 2$,

$$\mathcal{H}(\vec{\alpha}, \vec{A}) = [\omega_0 A_1 + A_2] + \varepsilon \left[A_2 + f_1(\vec{\alpha}_1) \, f_2(\vec{\alpha}_2) \right] \ , \tag{1.13}$$

where the vector $\vec{\omega}_0 = (\omega, 1)$ is a diophantine one. Then the function $N_f(\vec{A}; \varepsilon)$ is trivially defined as $N_f((A_1, A_2); \varepsilon) = \varepsilon A_2$, (and the resulting hamiltonian becomes integrable, see [G7], [GM1], for a proof with the methods used in the present paper), so that it is an entire function of ε. The original conjecture in [G2] was that, under the stronger hypothesis that the perurbation was a polinomial in ε, the function $N_f(\vec{A}; \varepsilon)$ would be *entire* in ε. In the case of (1.13), it is not difficult to verify that our methods, applied to that particular case, give $N_f(\vec{A}; \varepsilon) = \varepsilon A_2$; however, in general, we are not able to prove or refute the conjecture in the stronger form, since we have been successful only in proving an upper bound C^k, for some positive constant C, to the k-th perturbative order. However the validity of this stronger statement seems quite unlikely to us, without extra assumptions.

The paper is self-contained, and, although many ideas and notations are inherited from [G7], [GM1] and [GM2], no one of those works is required to be read in order to understand the present one; on the contrary, with respect to the previous papers, more details are given, as, because of the greater generality of the problem, the notations often become more involved. In §2, recursive formulae defining the coefficients of the formal series expansion in powers of the perturbative parameter of the solution of the equation of motions are given, and in §3 a diagrammatic representation in terms of *tree diagrams* (or simply *trees*) for such coefficients is introduced. In §4 a multiscale decomposition of the propagators is furnished, so that the construction of the labeled trees is completed. In §5 the contributions which are source of convergence troubles are identified, while §6 and §7 are devoted to the problem of dealing with such contributions; in §6 and §7, it is shown that the contributions arising from the resonances can be split into two parts, the first of which vanishes when the sum over all the trees is performed (see §6), while the second one can be easily handled through dimensional arguments (see §7): this completes the proof of Theorem 1.2. In §8 the discussion is adapted in order to prove Theorem 1.4.

In §9 a comparison with the existing literature on the matter is proposed, and in particular we give a detailed translation of Eliasson's original work in our formalism.

Some technical features of the proofs are discussed in Appendices A1 and A2. In Appendix A3 we give a "more natural" multiscale decomposition via a partition of unity with characteristic functions. This is an alternative approach to the one discussed in §2÷8, where the multiscale decomposition is based on a partition of unity built with smooth (C^∞) functions, with respect to which it presents some technical intricacies, but it turns out to be more suitable for the discussion of the breakdown of the invariant tori in [GGM].

2. Recursive formulae

Let us consider first the notationally less involved case in which the perturbation is linear in the perturbative parameter ε:

$$f(\vec{\alpha},\vec{A};\varepsilon)=\varepsilon\, f(\vec{\alpha},\vec{A})\;,\qquad F_0=\sup_{\vec{A}\in W(\vec{A}_0,\rho_0)}\sup_{\vec{\alpha}\in\mathbb{T}^\ell} f(\vec{\alpha},\vec{A})\;,\qquad E=\max\{F_0,E_0\}\;. \tag{2.1}$$

We shall see in Appendix A2 how to extend the discussion to the more general case.

Calling $\vec{H}^{(k)}$ and $\vec{h}^{(k)}$ the k-th order coefficients of the Taylor expansion of $\vec{H}$ and $\vec{h}$ in powers of ε, and writing the equations of motion as

$$\begin{aligned}\frac{d\vec{A}}{dt}&=-\varepsilon\partial_{\vec{\alpha}}f(\vec{\alpha},\vec{A})\;,\\ \frac{d\vec{\alpha}}{dt}&=\partial_{\vec{A}}\mathcal{H}_0+\varepsilon\partial_{\vec{A}}f(\vec{\alpha},\vec{A})\;,\end{aligned} \tag{2.2}$$

we get immediately recursion relations for $\vec{H}^{(k)},\vec{h}^{(k)}$; for $k=1$:

$$\begin{aligned}\vec{\omega}_0\cdot\partial_{\vec{\psi}}\,\vec{H}^{(1)}&=-\partial_{\vec{\alpha}}f\;,\\ \vec{\omega}_0\cdot\partial_{\vec{\psi}}\,\vec{h}^{(1)}&=T(\vec{A}_0)\vec{H}^{(1)}+\partial_{\vec{A}}f\;,\end{aligned} \tag{2.3}$$

(being $T(\vec{A}_0)$ defined before (1.4)), and, for $k>1$:

$$\begin{aligned}\vec{\omega}_0\cdot\partial_{\vec{\psi}}\,\vec{H}^{(k)}=&\sum_{(k-1)}{}^{*}(-\partial_{\vec{\alpha}})\sum_{p\ge0}\sum_{q\ge0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(\vec{h}^{(k_s)}\cdot\partial_{\vec{\alpha}}\right)\prod_{r=1}^{q}\left(\vec{H}^{(k'_r)}\cdot\partial_{\vec{A}}\right)f(\vec{\omega}_0t,\vec{A}_0)\;,\\ \vec{\omega}_0\cdot\partial_{\vec{\psi}}\,\vec{h}^{(k)}=&T(\vec{A}_0)\vec{H}^{(k)}+\sum_{(k)}{}^{*}\partial_{\vec{A}}\sum_{q\ge2}\frac{1}{q!}\prod_{r=1}^{q}\left(\vec{H}^{(k'_r)}\cdot\partial_{\vec{A}}\right)\mathcal{H}_0(\vec{A}_0)\\ &+\sum_{(k-1)}{}^{*}\partial_{\vec{A}}\sum_{p\ge0}\sum_{q\ge0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(\vec{h}^{(k_s)}\cdot\partial_{\vec{\alpha}}\right)\prod_{r=1}^{q}\left(\vec{H}^{(k'_r)}\cdot\partial_{\vec{A}}\right)f(\vec{\omega}_0t,\vec{A}_0)\;,\end{aligned} \tag{2.4}$$

where the $\sum^*_{(K)}$, $K = k, k-1$, denotes summation over the integers $k_s \geq 1$, $k'_s \geq 1$, with: $\sum_{s=1}^p k_s + \sum_{r=1}^q k'_r = K$, and the derivatives are supposed to apply to the functions $f(\vec{\alpha}, \vec{A})$, and then evaluated in $(\vec{\alpha}, \vec{A}) = (\vec{\omega}_0 t, \vec{A}_0)$. If $p = 0$ or $q = 0$, (both cases simultaneously are not possible), the corresponding product is meant as 1.

2.1. *Proof of the formal solubility of the recursive relations.* We proceed inductively through the following steps.
(1) From the equations of motion for the angular momenta, we obtain immediately the first recursive relation in (2.4). Then suppose that $\vec{h}^{(k)}$ has vanishing average in $\vec{\psi}$, for $1 \leq k < k_0$, and that $\vec{h}^{(k)}$ and $\vec{H}^{(k)}$ solve the equations of motion, (and therefore (2.4)), for $1 \leq k < k_0$. Note that for $k = 1$ the statement holds, (as can be trivially verified), provided the $\vec{0}$-th Fourier component of $\vec{H}^{(1)}$ is suitably chosen so that the right hand side of the second equation in (2.3) has vanishing $\vec{0}$-th Fourier component.
(2) Then the first equation in (2.4) can be solved for $k = k_0$, if the right hand side has vanishing average. This can be easily checked, as follows, (see [CG], App. A12, for an analogous discussion; see also [CZ]). Let be $Y(\vec{\psi}) \equiv Y = (\vec{h}(\vec{A}_0, \vec{\psi}\,; \varepsilon), \vec{H}(\vec{A}_0, \vec{\psi}\,; \varepsilon))$, where $\vec{h}$ and $\vec{H}$ are defined by the formal series expansions, respectively, $h = \sum_k \vec{h}^{(k)} \varepsilon^k$ and $H = \sum_k \vec{H}^{(k)} \varepsilon^k$, and let be E the symplectic matrix. Then, by assumption,

$$\frac{dY}{dt} \equiv (\vec{\omega}_0 \cdot \partial_{\vec{\psi}}\,) \, Y = (E \partial \mathcal{H})(Y) \,, \qquad \text{up to order } k_0 - 1 \,, \tag{2.5}$$

where $\partial = (\partial_{\vec{\alpha}}, \partial_{\vec{A}})$, and $\mathcal{H}(\vec{\alpha}, \vec{A})$ is the hamiltonian (1.1). But, for any periodic function $Y(\vec{\psi}\,)$, $\vec{\psi} \in \mathbb{T}^\ell$, (not necessarily the previously considered one), one has

$$\int_{\mathbb{T}^\ell} d\vec{\psi} \; (\partial_{\vec{\psi}} Y)(\vec{\psi}\,) \cdot \Big[E (\vec{\omega}_0 \cdot \partial_{\vec{\psi}}) Y \Big] \, (\vec{\psi}\,) = \vec{0} \,, \qquad \int_{\mathbb{T}^\ell} d\vec{\psi} \; (\partial_{\vec{\psi}} Y)(\vec{\psi}\,) \cdot (\partial \mathcal{H}) \, (\vec{\psi}\,) = \vec{0} \,, \tag{2.6}$$

where the first identity can be easily obtained by integration by parts and depends only on the periodicity of the function $Y(\vec{\psi}\,)$, while the second one is trivial, the integrand being the gradient with respect to $\vec{\psi}$ of the hamiltonian $\mathcal{H}$. Then, if Y is the function $(\vec{h}, \vec{H})$, the fact that the equations of motion are satisfied up to order $k_0 - 1$, (see (2.5)), implies that the sum of the two identities (2.6), *i.e.*

$$\int_{\mathbb{T}^\ell} d\vec{\psi} \; \Big[\partial_{\vec{\psi}} Y(\vec{\psi}\,) \cdot \Big(E(\vec{\omega}_0 \cdot \partial_{\vec{\psi}}\,) \, Y + \partial \mathcal{H}(\vec{\psi}\,) \Big) \Big] = \vec{0} \,, \tag{2.7}$$

to order k_0 gives simply

$$\int_{\mathbb{T}^\ell} d\vec{\psi} \; \Big[\partial_{\vec{\psi}} Y(\vec{\psi}\,) \Big]^{(0)} \cdot \Big[\partial \mathcal{H}(\vec{\psi}\,) \Big]^{(k_0)} = \vec{0} \,, \tag{2.8}$$

but the first term in (2.8) is a constant, since $Y^{(0)}(\vec{\psi}\,) = (\vec{\psi}\,, \vec{A}_0)$, so that we have that the average of the function $\partial_{\vec{\alpha}} \mathcal{H}$ is vanishing to order k_0. Then the assertion follows immediately, by noting that $[-\partial_{\vec{\alpha}}\mathcal{H}]^{(k_0)}$ is exactly the right hand side of the first equation in (2.4).

(3) Then the first equation in (2.4) yields a function $\vec{H}^{(k_0)}(\vec{\psi})$ which is defined up to the constant $\vec{\mu}^{(k_0)} \equiv \vec{H}^{(k_0)}_{\vec{0}}$, (*i.e.* the $\vec{0}$-th Fourier component), which we call "counterterm". Such a constant, however, must be taken so that the equation for the angle variables, *i.e.* the second of (2.4), has zero average, in order to be soluble.

(4) Hence the equation for $\vec{h}^{(k)}$ can be solved and its solution is defined up to an arbitrary constant: such a constant can be chosen to be vanishing, and the procedure can be iterated. This shows that a formal solution of (2.4) can be obtained. ■

If we look at (2.3) and (2.4), we see that $\vec{H}^{(k)}$ is given by a sole contribution which has always (at least) one derivative with respect to $\vec{\alpha}$, whereas $\vec{h}^{(k)}$ is the sum of two or three contributions such that the first one has again a derivative with respect to $\vec{\alpha}$ in front of all, while the other ones have a derivative with respect to $\vec{A}$. Then we can introduce the following notation: $\vec{H}^{(k)}$ is given by a sum of terms which are of the form $H \leftarrow h$, where H denotes that they contribute to $\vec{H}^{(k)}$ and h that the first derivative is with respect to the angle variables. In the same way, $\vec{h}^{(k)}$ is given by three sums (the second one is absent if $k=1$) of terms which are, in the first sum, of the form $h \leftarrow h$, and, in the latter two, of the form $h \leftarrow H$. The terms of the second and the third sums will be distinguished by a label δ, which can be set $\delta = 0$ for the second one, and $\delta = 1$ for the third one; for future convenience we assign a label $\delta = 1$ also to the other terms $H \leftarrow h$ and $h \leftarrow h$.

It can be convenient to write the above recursive formulae in the Fourier space. Then, if we take into account also the compatibility conditions required in order to make the equations (2.3), (2.4) to be soluble (see item (3) in §2.1), we easily find, for $k=1$, from (2.3),

$$
\begin{aligned}
\vec{h}^{(1)}_{\vec{\nu}} &= \frac{-iT(\vec{A}_0)\,\vec{\nu}\, f_{\vec{\nu}}(\vec{A}_0)}{(i\vec{\omega}_0\cdot\vec{\nu})^2} + \frac{\partial_{\vec{A}}\, f_{\vec{\nu}}(\vec{A}_0)}{i\vec{\omega}_0\cdot\vec{\nu}}\,, \qquad \vec{\nu}\neq\vec{0}\,,\\
\vec{H}^{(1)}_{\vec{\nu}} &= \frac{-i\vec{\nu}\, f_{\vec{\nu}}(\vec{A}_0)}{i\vec{\omega}_0\cdot\vec{\nu}}\,, \qquad \vec{\nu}\neq\vec{0}\,,\\
\vec{H}^{(k)}_{\vec{0}} &\equiv \vec{\mu}^{(1)} = -T^{-1}(\vec{A}_0)\,\partial_{\vec{A}}\, f_{\vec{0}}(\vec{A})\,,
\end{aligned}
\tag{2.9}
$$

and, for $k>1$ and $\vec\nu\neq\vec 0$, from (2.4),

$$
\begin{aligned}
(i\vec\omega_0\cdot\vec\nu)\,\vec H^{(k)}_{\vec\nu} =& \sum_{(k-1)}^{\;*}(-i\vec\nu_0)\sum_{p\ge0}\sum_{q\ge0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(i\vec\nu_0\cdot\vec h^{(k_s)}_{\vec\nu_s}\right)\prod_{r=1}^{q}\left(\vec H^{(k'_r)}_{\vec\nu'_r}\cdot\partial_{\vec A}\right)f_{\vec\nu_0}(\vec A_0)\,,\\
(i\vec\omega_0\cdot\vec\nu)\,\vec h^{(k)}_{\vec\nu} =& T(\vec A_0)\vec H^{(k)}_{\vec\nu}+\sum_{(k)}^{\;*}\partial_{\vec A}\sum_{q\ge2}\frac{1}{q!}\prod_{r=1}^{q}\left(\vec H^{(k'_r)}_{\vec\nu'_r}\cdot\partial_{\vec A}\right)\mathcal H_0(\vec A_0)\\
&+\sum_{(k-1)}^{\;*}\partial_{\vec A}\sum_{p\ge0}\sum_{q\ge0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(i\vec\nu_0\cdot\vec h^{(k_s)}_{\vec\nu_s}\right)\prod_{r=1}^{q}\left(\vec H^{(k'_r)}_{\vec\nu'_r}\cdot\partial_{\vec A}\right)f_{\vec\nu_0}(\vec A_0)\,,
\end{aligned}
\tag{2.10}
$$

if, for $k>1$ and $\vec\nu=\vec 0$, we set $\vec H^{(k)}_{\vec 0}=\vec\mu^{(k)}$, with

$$
\vec\mu^{(k)}=\sum_{(k-1)}^{\;*}\left(-T^{-1}(\vec A_0)\partial_{\vec A}\right)\sum_{p\ge0}\sum_{q\ge0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(i\vec\nu_0\cdot\vec h^{(k_s)}_{\vec\nu_s}\right)\prod_{r=1}^{q}\left(\vec H^{(k'_r)}_{\vec\nu'_r}\cdot\partial_{\vec A}\right)f_{\vec\nu_0}(\vec A_0)\,,
\tag{2.11}
$$

where the $\sum^*_{(K)}$, $K=k,k-1$, denotes summation over the integers $k_s\ge1$, $k'_r\ge1$, with: $\sum_{s=1}^p k_s+\sum_{r=1}^q k'_r)=K$, and over the integers $\vec\nu_0$, $\vec\nu_s$, $\vec\nu'_r$, with: $\vec\nu_0+\sum_{s=1}^p\vec\nu_s+\sum_{r=1}^q\vec\nu'_r$ $=\vec\nu$, and in the second contribution of the second equation in (2.10) $\vec\nu_0\equiv\vec 0$ identically. The interpretation of the cases $q=0$ and $p=0$ is as in (2.3).

Again, if we use the terminology introduced after (2.3), $\vec H^{(k)}_{\vec\nu}$ is given by a sum of terms of the form $H\leftarrow h$, and $\vec h^{(k)}_{\vec\nu}$ by three contributions which are sums of terms of the form $h\leftarrow h$ and $h\leftarrow H$, (and $\delta=0,1$ in the second case). The contribution to $\vec\mu^{(k)}$ can be interpreted as a sum of terms of the form $\mu\leftarrow H$, and a label $\delta=1$ can be assigned to them.

3. Diagrammatic expansion

The equations (2.8)÷(2.10) provide an algorithm to evaluate a formal power series solution to our problem: in fact they allow us to carry out a *diagrammatic expansion* of $\vec h^{(k)}_{\vec\nu}$ and $\vec H^{(k)}_{\vec\nu}$: we simply "iterate" it until only $\vec h^{(1)}_{\vec\nu}$, $\vec H^{(1)}_{\vec\nu}$, $\vec\nu\neq\vec 0$, and $\vec\mu^{(1)}$ appear.

If we define the dimensionless quantity $\vec X^{(k)}_{\vec\nu}(\zeta)$ as

$$
\vec X^{(k)}_{\vec\nu}(\zeta)=\begin{cases}\vec h^{(k)}_{\vec\nu}\,, & \text{if } \zeta=h,\ \vec\nu\neq\vec 0\,,\\ (C_0/J_m)\vec H^{(k)}_{\vec\nu}\,, & \text{if } \zeta=H,\ \vec\nu\neq\vec 0\,,\\ (C_0/J_m)\vec\mu^{(k)}\,, & \text{if } \zeta=\mu,\ \vec\nu=\vec 0\,,\end{cases}
$$

then we shall show that it is possible to write $\vec X^{(k)}_{\vec\nu}(\zeta)$ as sum of contributions each of which can be graphically represented as a tree diagram ϑ. In other words, we shall give some rules in order to associate to a suitable diagram ϑ a value $\mathrm{Val}(\vartheta)$, and show that it turns out to be, for each Fourier component,

$$\vec{X}^{(k)}_{\vec{\nu}}(\zeta) = \sum_{\vartheta\in\mathcal{T}_k} \mathrm{Val}(\vartheta) \;, \tag{3.1}$$

where the sum is over all the trees contained in a certain class $\mathcal{T}_k$, (which will be defined later).

3.1. *Topological and semitopological trees.* A tree diagram (or *tree*) ϑ will consist of a family of lines (*branches* or *lines*) arranged to connect a partially ordered set of points (*vertices* or *nodes*), with the higher vertices to the right. The branches are naturally ordered as well; all of them have two vertices at their extremes (possibly one of them is a top vertex), except the lowest or *first branch* which has only one vertex, the *first vertex* v_0 of the tree. The other extreme r of the first branch will be called the *root* of the tree and will not be regarded as a vertex; we shall call the first branch also *root branch.* A possible tree is represented in Fig.3.1.

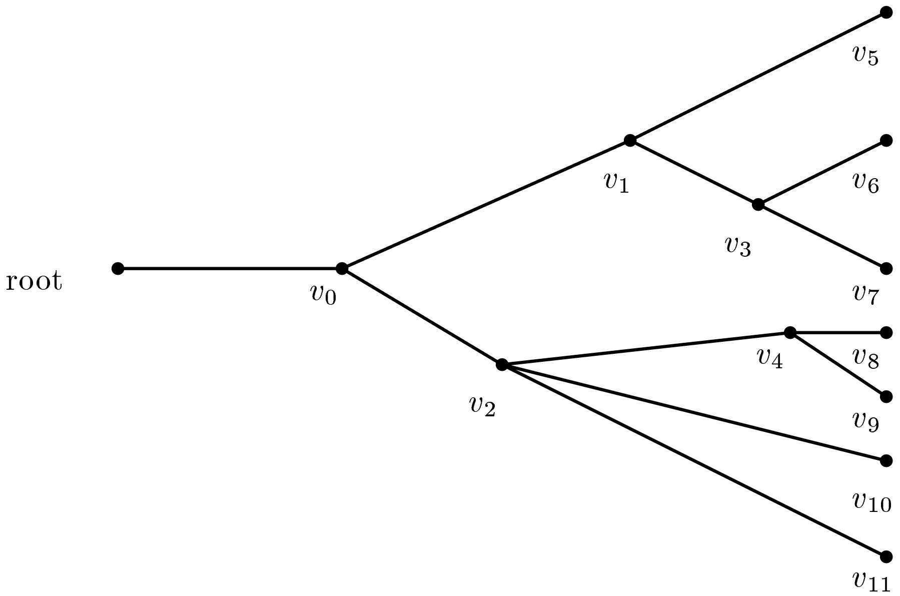

Fig.3.1. A tree ϑ with degree $d = 11$. Each line (branch) is supposed to carry an arrow (which is not explicitly drawn) pointing to the root. If we consider a vertex of the tree, *e.g.* v_3, then we define λ_{v_3}, or equivalently $v_1 \leftarrow v_3$, the line connecting v_3 to v_1, and we write $v_1 = v_3'$. The arrow, if drawn, would point from v_3 to v_1, as one has to cross v_1 in order to reach the root from v_3.

If v_1 and v_2 are two vertices of the tree we say that $v_1 < v_2$ if v_2 follows v_1 in the order established by the tree, *i.e.* if one has to pass v_1 before reaching v_2, while climbing the tree. Since the tree is partially ordered not every pair of vertices will be related by the order relation, (which we are denoting $\leq$): we say that two vertices are comparable if they are related by the order relation.

Given a vertex v we denote by v' the vertex immediately preceding v. We also imagine that each branch of the tree carries an arrow pointing to the root ("gravity direction", opposite to the order): this means that if a line connects two vertices, say v' and v, with $v' < v$, (so that v *follows* v'), then the arrow points from v to v', and we say that the line *emerges from* v and *enters* v'.

Given a tree ϑ with first vertex v_0, each vertex $v > v_0$ can be considered the first vertex of the tree consisting of the vertices following v: such a tree will be called a *subtree* of ϑ.

Let us define the *degree* of a tree as the number of vertices of the tree. Obviously, the degree of a tree ϑ counts also the number of branches of ϑ: in fact there is a correspondence 1-to-1 between vertices and branches, if we associate to each vertex v the branch λ_v emerging from it. We can also represent a branch λ_v as $v' \leftarrow v$, (see Fig.3.1 for definiteness).

A group $\mathcal{G}$ of transformations acts on the trees, generated by the following operations: fix a node $v \in \vartheta$ and permute the subtrees emerging from it. We call *semitopological trees* the trees which are superposable up to a continuous deformation of the branches on the plane, and *topological trees* when the same happens modulo the action of the just defined group of transformations. We shall denote by $\Theta^{(s)}$ the set of semitopological trees, and by $\Theta^{(t)}$ the set of the topological trees. For example, the trees drawn in Fig.3.2 will be regarded as different as semitopological trees and equivalent as topological trees.

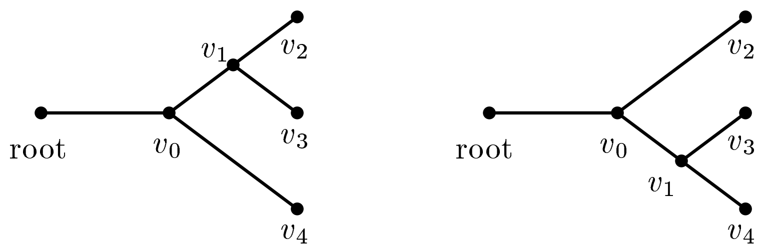

Fig.3.2. Two trees ϑ_1 and ϑ_2 of degree $d = 5$, which are different if regarded as semitopological trees and identical if regarded as topological trees. In fact, if we permute the subtrees emerging from the first vertex, we obtain ϑ_2 from ϑ_1 and *viceversa*.

Note that the number of topological trees and the number of semitopological trees of degree d can be bounded by 2^{2d}, (see, *e.g.* , [HP]).

Given a vertex $v \in \vartheta$ at which the tree ϑ bifurcates into m_v subtrees among which there are only $\mathcal{B}_v$ topologically different subtrees $\vartheta_1, \ldots, \vartheta_{\mathcal{B}_v}$, each of which is repeated $\mathcal{N}_v(\vartheta_i)$ times, and given a function $F(\vartheta)$ whose value depends only on the topological tree, one has

$$\sum_{\vartheta \in \Theta^{(s)}} \prod_{v \in \vartheta} \frac{1}{m_v!} F(\vartheta) = \sum_{\vartheta \in \Theta^{(t)}} \prod_{v \in \vartheta} \prod_{i=1}^{\mathcal{B}_v} \frac{1}{\mathcal{N}_v(\vartheta_i)!} F(\vartheta) \, , \tag{3.2}$$

where we recall that $\Theta^{(s)}$ is the set of semitopological trees and $\Theta^{(t)}$ is the set of topological trees, (see also [G8], §5).

3.2. *Numbered trees.* It can be convenient to introduce also another kind of trees, which we call *numbered trees*, (following [G8]): they are obtained by imagining to have a deposit of d branches numbered from 1 to d and depositing them on the branches of a topological tree with degree d. The numbered trees will be regarded as identical if superposable by the action of a transformation of the group $\mathcal{G}$, in such a way that all the numbers associated to the lines match. One has, for a function $F(\vartheta)$ depending only on the topological trees,

$$\sum_{\vartheta\in\Theta^{(s)}}\prod_{v\in\vartheta}\frac{1}{m_v!}F(\vartheta)=\frac{1}{d!}\sum_{\vartheta\in\Theta^{(n)}}F(\vartheta)\;, \tag{3.3}$$

if $\Theta^{(n)}$ is the set of numbered trees. The number of numbered trees of degree d is bounded by $d!2^{2d}$.

To work with the numbered trees can be very convenient from a combinatorial point of view: this will be particularly evident in §6, where the cancellation mechanisms operating to all perturbative orders will be investigated and will be shown to be easily visualized in terms of numbered trees. Nevertheless it is important to keep in mind that it is absolutely equivalent to study the perturbative expansions in terms of semitopological trees or in terms of numbered trees, and to choose one of the two possibilities is only a matter of convenience. So in §5, it will turn out to be easier to work with semitopological trees.

3.3. *Labeled trees.* To each vertex we associate a finite set of labels, defined as follows.
(1) d_v is the number of vertices w, such that $w\ge v$, *i.e.* the number of vertices of the subtree having v as first vertex, (and it is the degree of such a subtree);
(2) $\delta_v=0,1$;
(3) k_v is defined as $k_v=\sum_{w\ge v}\delta_w$, and it is called the *order* of the subtree having v as first vertex;
(4) $\vec\nu_v\in\mathbb{Z}^\ell$ is called the *mode label*;
(5) ζ_v^1 and ζ_v^2 can assume the symbolic values $\zeta_v^1,\zeta_v^2=h,H,\mu$;
(6) m_v is the number of branches entering the vertex v, (if v is a top vertex, then trivially $m_v=0$); morover, let us define p_v and q_v the number of branches entering v and emerging from vertices w (such that $w'=v$) carrying, respectively, a label $\zeta_w^1=h$ and a label $\zeta_w^1=H,\mu$, so that $m_v=p_v+q_v$.

If a branch λ_v connects to a vertex v' a vertex v, with labels ζ_v^1 and ζ_v^2, we shall classify the branch as a $\zeta_v^1\leftarrow\zeta_v^2$ branch. To each branch λ_v we associate a *momentum* $\vec\nu_{\lambda_v}\equiv\sum_{w\ge v}\vec\nu_w$, and a functional given by the product of an operator $\mathcal{O}_{\lambda_v}$ times a propagator $g(\vec{\omega}\cdot\vec{\nu}_{\lambda_v})$, which are defined as follows:

$$
\begin{array}{ccc}
\text{operator} & \text{propagator} & \text{branch} \\
C_0^2\,[i\vec{\nu}_{v'}\cdot(-iT\,\vec{\nu}_v)] & [i\vec{\omega}\cdot\vec{\nu}_{\lambda_v}]^{-2} & h\leftarrow h \\
C_0\left[i\vec{\nu}_{v'}\cdot(\partial_{\vec{A}_v})\right] & [i\vec{\omega}\cdot\nu_{\lambda_v}]^{-1} & h\leftarrow H \\
C_0\left[\partial_{\vec{A}_{v'}}\cdot(-i\vec{\nu}_v)\right] & [i\vec{\omega}\cdot\vec{\nu}_{\lambda_v}]^{-1} & H\leftarrow h \\
\partial_{\vec{A}_{v'}}\cdot(-T^{-1}\partial_{\vec{A}_v}) & 1 & \mu\leftarrow H
\end{array}
\tag{3.4}
$$

for all the branches distinct from the root branch, and

$$
\begin{array}{ccc}
\text{operator} & \text{propagator} & \text{branch} \\
C_0^2\,[-iT\,\vec{\nu}_v] & [i\vec{\omega}\cdot\vec{\nu}_{\lambda_v}]^{-2} & h\leftarrow h \\
C_0\left[\partial_{\vec{A}_v}\right] & [i\vec{\omega}\cdot\nu_{\lambda_v}]^{-1} & h\leftarrow H \\
C_0^2\,J_m^{-1}\,[-i\vec{\nu}_v] & [i\vec{\omega}\cdot\vec{\nu}_{\lambda_v}]^{-1} & H\leftarrow h \\
C_0\,J_m^{-1}\left[-T^{-1}\partial_{\vec{A}_v}\right] & 1 & \mu\leftarrow H
\end{array}
\tag{3.5}
$$

for the root branch; the first three terms in (3.4) and (3.5) can occur only if $\vec{\nu}_{\lambda_v}\neq\vec{0}$, while the fourth one only if $\vec{\nu}_{\lambda_v}=\vec{0}$. All the other pairs (ζ_v^1,ζ_v^2) are defined to give a vanishing contribution, so that we can get rid of them.

3.4. *Remark.* Note that, in (3.4), the operators corresponding to the lines $h\leftarrow H$ and $H\leftarrow h$ are opposite to each other under the change of the vertices v' and v.

Then we multiply all the above operators $\mathcal{O}_{\lambda_v}$ (we simply regard as multiplication operators the factors in which no derivative appears) to the function

$$
\prod_{\substack{v\in\vartheta\\ \delta_v=1}} f_{\vec{\nu}}(\vec{A}_v)\prod_{\substack{v\in\vartheta\\ \delta_v=0}}\mathcal{H}_0(\vec{A}_v)\;,
\tag{3.6}
$$

and evaluate the result at the points $\vec{A}_v\equiv\vec{A}_0$, $\forall v\in\vartheta$.

We define O_v the (tensor) factor we can associate to each vertex, once the above operators $\mathcal{O}_{\lambda_w}$, $w\in\vartheta$, have been applied to the function (3.6),

$$
\prod_{v\in\vartheta}O_v=\prod_{v\in\vartheta}\mathcal{O}_{\lambda_v}\prod_{\substack{v\in\vartheta\\ \delta_v=1}} f_{\vec{\nu}}(\vec{A}_v)\prod_{\substack{v\in\vartheta\\ \delta_v=0}}\mathcal{H}_0(\vec{A}_v)\Big|_{\vec{A}_v=\vec{A}_0}\;,
\tag{3.7}
$$

where O_v is given by

$$\begin{aligned}O_v =\Big\{ & \Big[C_0^2\Big(-iT\,\vec\nu_v\Big)\,\delta_{\zeta_v^1,h}\,\delta_{\zeta_v^2,h} + C_0\Big(\partial_{\vec A_v}\Big)\,\delta_{\zeta_v^1,h}\,\delta_{\zeta_v^2,H} + C_0^2\,J_m^{-1}\Big(-i\vec\nu_v\Big)\,\delta_{\zeta_v^1,H}\,\delta_{\zeta_v^2,h}\\ & + C_0\,J_m^{-1}\Big(-T^{-1}\,\partial_{\vec A_v}\Big)\,\delta_{\zeta_v^1,\mu}\,\delta_{\zeta_v^2,H}\Big]\cdot\\ & \cdot\prod_{\substack{w\\ w'=v}}\Big[(i\vec\nu_v)\,\delta_{\zeta_w^1,h} + C_0^{-1}J_m\,(\partial_{\vec A_v})\Big(\delta_{\zeta_w^1,H}+\delta_{\zeta_w^1,\mu}\Big)\Big]\Big\}\cdot\\ & \cdot\ \Big[f_{\vec\nu_v}(\vec A_v)\,\delta_{\delta_v,1} + \mathcal{H}_0(\vec A_v)\,\delta_{\delta_v,0}\Big]\Big|_{\vec A_v=\vec A_0}\ ,\end{aligned} \tag{3.8}$$

being the product over all the vertices immediately following v (so that it is missing if v is a top vertex).

3.5. *Remark.* Note that, unlike the operators, the factors O_v do not depend only on the line λ_v, but also on the lines entering v, (which can carry some derivatives with respect to the action variables acting on $f_{\vec\nu_v}(\vec A_v)$ or $\mathcal{H}_0(\vec A_v)$). For this reason we prefer to associate the factors to the vertices rather than to the lines, but, obviously, this is somewhat arbitrary, as there is a correspondence 1-to-1 between vertices and branches, (see §3.1 and Fig.3.1). From (3.8) we see that the derivatives with respect to the action variable $\vec A_v$, $v\in\vartheta$, collected together, give a tensor

$$\Big[(1-\delta_{\zeta_v^2,H})+\delta_{\zeta_v^2,H}\,\partial_{\vec A_v}\Big]\prod_{\substack{w\\ w'=v}}\Big[\partial_{\vec A_v}\Big(\delta_{\zeta_w^1,H}+\delta_{\zeta_w^1,\mu}\Big)\Big]\Big[f_{\vec\nu_v}(\vec A_v)\,\delta_{\delta_v,1}+\mathcal{H}_0(\vec A_v)\,\delta_{\delta_v,0}\Big]\Big|_{\vec A_v=\vec A_0}\ ,$$

each of whose entries can be bounded by

$$(q_v+\delta_{\zeta_v^2,H})!\,E\,\rho_0^{-(q_v+\delta_{\zeta_v^2,H})}\ ,$$

through the Schwarz's lemma, [Ti], §5.2, (which is in turn a trivial application of the Cauchy formula).

3.6. *Remark.* Note that two trees topologically or semitopologically equivalent can become different as labeled trees. They will be considered as identical only if, when superposed (if topologically equivalent) or superposed modulo a transformation of the group $\mathcal{G}$ (if semitopologically equivalent), all their labels match. If the trees are numbered, they will be considered equivalent if also the numbers match.

Now we have all the definitions and notations necessary to represent $\vec X(\zeta)$ as sum of values associated to semitopological trees (in §6, we shall see how to change the definition of tree value in order to express $\vec X(\zeta)$ in terms of numbered trees).

We can define the *value of a labeled tree* ϑ as

$$\mathrm{Val}(\vartheta)=\prod_{v\in\vartheta}\frac{O_v}{m_v!}\,g(\vec\omega\cdot\vec\nu_{\lambda_v})\ , \tag{3.9}$$

and the $\vec{\nu}$-the Fourier component of the function $\vec{X}^{(k)}(\zeta)$ can be expressed as a sum of the form (3.1), where $\mathcal{T}_k = \mathcal{T}_k^{(s)}$, if $\mathcal{T}_k^{(s)}$ is the *collection of all the possible not equivalent labeled semitopological trees* with order $k_{v_0} = k$ and $\vec{\nu}_{v_0} = \vec{\nu}$, if v_0 is the first vertex of the trees, (the notion of equivalence being defined in Remark 3.6).

The proof of such an assumption can be easily obtained, if we recall that any subdiagram emerging from a vertex $v \in \vartheta$ is again a tree having v as first vertex and v' as root, and look at the recursive formulae (2.8), (2.9) and (2.10). Then the interpretation of all the labels listed in 3.3 becomes clear, as they can be related to the formulae (2.8)÷(2.10) and to the notations introduced at the end of §2 (for the labels δ and the symbols $\zeta_v^1 \leftarrow \zeta_v^2$). In particular the following result can be easily proven to hold.

3.7. PROPOSITION. *The labels so defined have to satisfy the following* compatibilty condition*: if $\delta_v = 0$, then it is $\zeta_v^1 = h$, $\zeta_v^2 = H$, $p_v \equiv 0$ and $m_v \equiv q_v \geq 2$. Then it is easy to see that $d_v \leq 2k_v - 1$ for each v. In particular $d \leq 2k - 1$.*

3.8. *Proof of Proposition 3.7.* We note from (2.8) that $\delta_v \equiv 1$ if v is a top vertex. Then Proposition 3.7 follows immediately. ■

Note that each non trivial (*i.e.* not corresponding to the case $\mu \leftarrow H$) propagator can be written as

$$g(\vec{\omega} \cdot \vec{\nu}_{\lambda_v}) = \frac{1}{[\vec{\omega} \cdot \vec{\nu}_{\lambda_v}]^{R_{\lambda_v}}} , \tag{3.10}$$

where $R_{\lambda_v} = 1$ if λ_v is a $H \leftarrow h$ branch, ($\zeta_v^1 = H$ and $\zeta_v^2 = h$), or a $h \leftarrow H$ branch, ($\zeta_v^1 = h$ and $\zeta_v^2 = H$)), $R_{\lambda_v} = 2$ if λ_v is a $h \leftarrow h$ branch, ($\zeta_v^1 = \zeta_v^2 = h$).

4. Multiscale analysis of the tree values

We introduce a multiscale decomposition of the propagator. Let $\chi(x)$ be a C^∞ not increasing function such that $\chi(x) = 0$, if $|x| \geq 2$ and $\chi(x) = 1$ if $|x| \leq 1$, and let $\chi_n(x) = \chi(2^{-n}x) - \chi(2^{-(n-1)}x)$, $n \leq 0$, and $\chi_1(x) = 1 - \chi(x)$: such functions realize a C^∞ partition of unity, for $|x| \in [0, \infty)$, in the following way. Let us write

$$1 = \chi_1(x) + \sum_{n=-\infty}^{0} \chi_n(x) \equiv \sum_{n=-\infty}^{1} \chi_n(x) . \tag{4.1}$$

Then we can decompose the propagator in the following way:

$$g(\vec{\omega} \cdot \vec{\nu}_{\lambda_v}) = \frac{1}{[i\vec{\omega} \cdot \vec{\nu}_{\lambda_v}]^{R_{\lambda_v}}} \equiv \sum_{n=-\infty}^{1} \frac{\chi_n(\vec{\omega} \cdot \vec{\nu}_{\lambda_v})}{[i\vec{\omega} \cdot \vec{\nu}_{\lambda_v}]^{R_{\lambda_v}}} \equiv \sum_{n=-\infty}^{1} g^{(n)}(\vec{\omega} \cdot \vec{\nu}_{\lambda_v}) \tag{4.2}$$

where $g^{(n)}(\vec{\omega} \cdot \vec{\nu}_{\lambda_v})$ is the "propagator at scale n". If $n < 0$, $g^{(n)}(\vec{\omega} \cdot \vec{\nu}_{\lambda_v})$ is a C^∞ compact support function different from 0 for $2^{n-1} < |\vec{\omega} \cdot \vec{\nu}_{\lambda_v}| \leq 2^{n+1}$, while $g^{(1)}(\vec{\omega} \cdot \vec{\nu}_{\lambda_v})$ has support for $1 < |\vec{\omega} \cdot \vec{\nu}_{\lambda_v}|$. In the domain where it is different from zero, the propagator verifies the bound

$$\left| \frac{\partial^p}{\partial x^p} g^{(n)}(x) \right|_{x=\vec{\omega}\cdot\vec{\nu}_{\lambda_v}} \le a_{R_{\lambda_v}}(p)\, 2^{-n(R_{\lambda_v}+p)} \;, \qquad p \in \mathbb{N} \;, \tag{4.3}$$

where $a_{R_{\lambda_v}}(p)$ is a suitable constant, such that $a_{R_{\lambda_v}}(0) = 2^{R_{\lambda_v}}$, which depends on the form of the function $\chi(x)$. The constant $a_{R_{\lambda_v}}(p)$ has a bad dependence on p, (since $g^{(n)}(x)$ is only a C^∞ function), but we shall see that in our bounds p does not increase ever beyond 2.

Proceeding as in quantum field theory, see [G5], given a tree ϑ we can attach a *scale label* n_{λ_v} to each branch λ_v in ϑ, which is equal to the scale of the propagator associated to the branch.

Looking at such labels we identify the connected cluster T of vertices which are linked by a continuous path of branches with the same scale labels n_T or a higher one and which are maximal: we shall say that *the cluster T has scale n_T*. Therefore an inclusion relation is established between the clusters, in such a way that the innermost clusters are the clusters with the highest scale, and so on. Each cluster can have an arbitrary number of branches entering it, (*incoming lines*), but only one branch exiting, (*outgoing line*); we use the fact that the branches carry an arrow pointing to the root: this gives a meaning to the words "incoming" and "outgoing". We call *external lines* the lines which are either outgoing or incoming. A possible situation is described in Fig.4.1.

The multiscale decomposition (4.2) of the propagator allows us to rewrite (3.9)

$$\mathrm{Val}(\vartheta) = \prod_{v\in\vartheta} \frac{O_v}{m_v!}\, g^{(n_{\lambda_v})}(\vec{\omega}\cdot\vec{\nu}_{\lambda_v}) \;, \tag{4.4}$$

and a formula like (3.1) holds still, provided that we count also the scale labels among the tree labels.

Obviously the choise of the partition of unity is not uniquely fixed: a different possibility is envisaged in Appendix A3. Of course the result we are looking for, *i.e.* the proof of Theorem 1.2, is independent on the particular partition, but the technical features can be more or less suitable for the discussion. For instance in [GGM] it turns out to be more convenient to work with the partition which is illustraded in Appendix A3.

5. Resonances and related problems

In this section, we confine ourselves to single out the contributions wich can be source of problems and need a more careful analysis. The discussion of such terms, and the exhibition of the cancellation mechanisms which have to be exploited in order to prove the convergence of the perturbative series are differed to next sections.

5.1. DEFINITION (RESONANCE). *Among the clusters we consider the ones with the property that there is only one incoming line, carrying the same momentum of the outgoing*

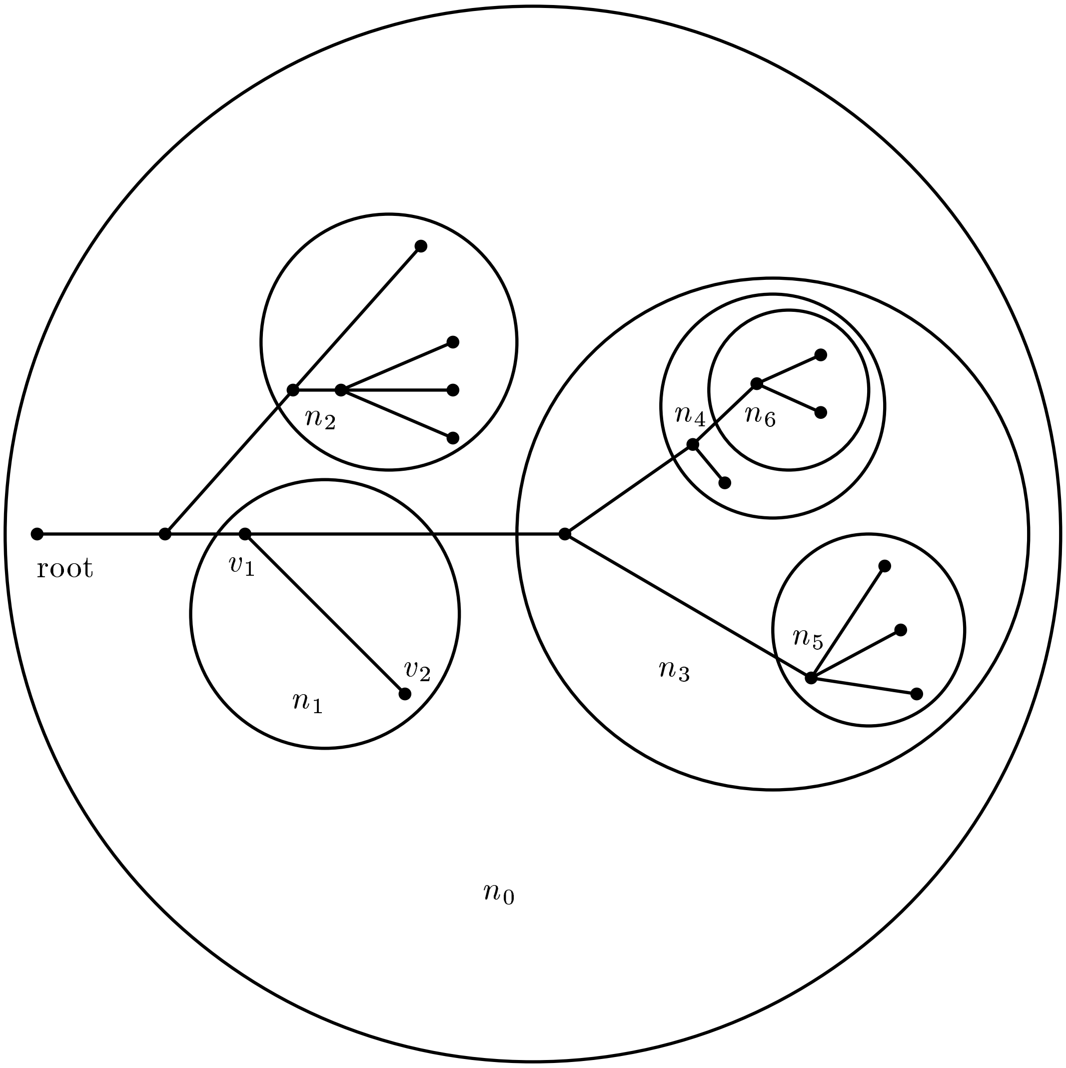

Fig.4.1. A tree with scale labels associated to the lines. There are seven clusters $T_0, \ldots, T_6$ on scale, respectively, $n_0, \ldots, n_6$, which satisfy the ordering relations: $n_0 > n_1$, $n_0 > n_2$, $n_0 > n_3$, $n_3 > n_4$, $n_3 > n_5$, and $n_4 > n_6$. If the external lines of the cluster T_1 carry the same momentum (*i.e.* $\vec{\nu}_{v_1} + \vec{\nu}_{v_2} = \vec{0}$), then T_1 is a resonance (see Definition 5.1), and we write $T_1 = V_1$. Note that there is always a maximal cluster encircling all the tree, and that there is only one outgoing line per cluster.

line, and we define them resonances. *If V is one such cluster we denote by λ_V the incoming line, and by $d(V)$ and $k(V)$, respectively, the number of vertices contained in V (*resonance degree*) and the quantity $\sum_{w\in V}\delta_w$ (*resonance order*). We call n_{λ_V} the* resonance-scale, *and λ_V a* resonant line*: if n_V is the scale of the resonance as a cluster,* i.e. *the lowest scale of the line inside V, one has $n_V \ge n_{\lambda_V} + 1$.*

Given a tree ϑ, let us define $N_n(\vartheta)$ the number of lines with scale $n \le 0$, and $N_n^j(\vartheta)$, $j = 1, 2$, the number of lines λ with scale $n \le 0$ and $R_\lambda = j$. Let us define also $\Delta(\vartheta)$ the collection of vertices v's in ϑ (with $\delta_v = 1$) such that $\vec{\nu}_v \ne \vec{0}$, and $M(\vartheta)$ the quantity $M(\vartheta) = \sum_{v\in\Delta(\vartheta)} |\vec{\nu}_v|$.

Then it is easy to check that the scaling properties of the propagators and the definitions (3.4), (3.5) and (3.7) immediately imply that the contribution to $\vec{X}_{\vec{\nu}}^{(k)}(\zeta)$ arising from a given tree ϑ can be bounded as follows:

$$|\mathrm{Val}(\vartheta)| \le \mathcal{C}^k \, e^{-\xi\, M(\vartheta)} \prod_{n\le 0} 2^{-(2nN_n^2(\vartheta)+nN_n^1(\vartheta))} \prod_{v\in\Delta(\vartheta)} \frac{|\vec{\nu}_v|^{p_v+1}}{p_v!} \,, \tag{5.1}$$

(*dimensional bound*) for a suitable constant $\mathcal{C}$, given by

$$\mathcal{C} = \left[2^3\, J_M\, J_m^{-2}\, C_0^2\, E\, \ell\, \rho^{-2}\right]^2 \,, \tag{5.2}$$

where C_0 is the diophantine constant introduced in (1.3), ρ is introduced after (1.7), the eigenvalues J_m and J_M are defined in (1.4) and E in (1.7).

5.2. *Proof of (5.1) and (5.2).* In (5.2), 2^3 arises from the definition of the compact support of the propagators, (which gives $a_{R_{\lambda_v}}(0) \le a_2(0) = 2^2$), and from the fact that (1) to each vertex v a factorial $(q_v + \delta_{\zeta_v^2,H})!$ is associated when the Cauchy formula is used in order to bound the derivatives, (see Remark 3.5), (2) one has $p_v!\, q_v! \le m_v!$, and (3) one can bound $(q_v+\delta_{\zeta_v^2,H})!/q_v! \le 2^{q_v}$, with $\sum_v q_v \le d_{v_0}$, (which gives another 2).

Moreover each line λ_v, $v_0 < v \in \vartheta$, has associated a scalar product, which, in according to the kind of branch λ_v and if we neglect the factorials, (which have been already taken into account), can be bounded, respectively, by

$$\begin{array}{ll}
C_0^2\, J_m^{-1}\, |\vec{\nu}_{v'}|\, |\vec{\nu}_v|\, E\, e^{-\xi|\vec{\nu}_v|} & \qquad h \leftarrow h \\[6pt]
C_0^2\, J_m^{-1}\, |\vec{\nu}_{v'}|\, \ell\, \rho^{-1}\, E\, e^{-\xi|\vec{\nu}_v|} & \qquad h \leftarrow H \\[6pt]
C_0^2\, J_m^{-1}\, \ell\, \rho^{-1}\, |\vec{\nu}_v|\, E\, e^{-\xi|\vec{\nu}_v|} & \qquad H \leftarrow h \\[6pt]
C_0^2\, J_M\, J_m^{-2}\, \ell\, \rho^{-2}\, E\, e^{-\xi|\vec{\nu}_v|} & \qquad \mu \leftarrow H
\end{array}$$

and then each of the four above quantities can be bounded with

$$J_M\, J_m^{-2}\, C_0^2\, \ell\, r^{-1}\, E\, e^{-\xi|\vec{\nu}_v|}\, \max\{|\vec{\nu}_{v'}|,1\}\, \max\{|\vec{\nu}_v|,1\}\, e^{-\xi|\vec{\nu}_v|} \,,$$

where r is defined after (1.7), $r^{-1} = \max\{1,\rho^{-2}\}$; an analogous bound holds for the root branch too, *i.e.*

$$J_M\, J_m^{-2}\, C_0^2\, r^{-1} \max\{|\vec{\nu}_{v_0}|,1\}\, E\, e^{-\xi|\vec{\nu}_{v_0}|} \,,$$

so that, if we recall Proposition 3.4, (which allows us to bound $d \equiv d_{v_0} \le 2k-1$), then (5.2) immediately follows. ■

The concept of resonance is a very important one: in fact it allows us to identify the terms which need an improvement of the dimensional bound (5.1) in order to prove the convergence of the Lindstedt series. Indeed if there were no resonances, the series would converge, as we can easily show.

In order to obtain a bound on $\vec{X}^{(k)}_{\vec{\nu}}(\zeta)$, in general we have to sum the values of all the labeled semitopological trees. Such a sum can be arranged as

$$\sum_{\vartheta\in\mathcal{T}_k^{(s)}} = \sum_{\vartheta\in\Theta^{(s)}} \sum_{\substack{\{\delta_v\}_{v\in\vartheta} \\ \sum_{v\in\vartheta}\delta_v=k}} \sum_{\{\zeta_v^1,\zeta_v^2\}_{v\in\vartheta}} \sum_{\{\vec{\nu}_v\}_{v\in\vartheta}} \sum_{\{n_v\}_{v\in\vartheta}}$$

where the summations have to be performed by starting from the rightmost one, and run over the sets defined in the following way:
(1) the first one is over all the semitopological trees with no labels;
(2) the second one is over all the possible assignments of the labels δ_v to the vertices $v\in\vartheta$ of a fixed unlabeled semitopological tree, with the constraint $\sum_{v\in\vartheta}\delta_v = k$;
(3) the third one is over all the possible assignments of the labels ζ_v^1, ζ_v^2, to the vertices $v\in\vartheta$;
(4) the fourth one is over the mode labels;
(5) the last one is over the scale labels.

This exhausts all the possibilities, being all the other labels uniquely determined by the just considered ones.

Then we perform the fifth sum, keeping the mode labels (and all the other labels) fixed, so obtaining $2^{2(2k-1)}$ terms.[2] Successively we write, as far the mode labels are concerned,

$$\sum_{\{\vec{\nu}_v\}_{v\in\vartheta}} = \sum_{\Delta(\vartheta)} \sum_{M(\vartheta)=|\Delta(\vartheta)|}^{\infty} \sum_{\substack{\{\vec{\nu}_v\}_{v\in\Delta(\vartheta)} \\ \sum_{v\in\vartheta}|\vec{\nu}_v|=M(\vartheta)}} \tag{5.3}$$

where $|\Delta(\vartheta)|$ denotes the number of vertices in $\Delta(\vartheta)$.

We want to show that, if there are no resonances, fixed $M(\vartheta)$ in the rewriting (5.3) of the sum in item (4), then we obtain a quantity which depends in a summable way on $M(\vartheta)$, and the resulting expression can be bounded by C_2^k for some positive constant C_2. In fact, if we are successful in proving such an estimate, then the remaining sums, *i.e.* the sums in items (1), (2) and (3), are over $< 2^{2(2k-1)}\cdot 2^{2k-1}\cdot 2^{2(2k-1)}$ terms, (as there are four possible pairs (ζ_v^1,ζ_v^2) and two possible values δ_v, for each $v\in\vartheta$, and the number of unlabeled semitopological trees of order k is bounded by $2^{2(2k-1)}$), so that an overall

[2] Because of the compact support of the propagators (see comments between (4.2) and (4.3)), given a value $\vec{\nu}_\lambda$, there are only two consecutive scales $n'=n,n+1$ such that $g^{(n')}(\vec{\omega}\cdot\vec{\nu}_\lambda)$ is different from zero, and there are at most $2k-1$ propagators associated to the lines of the tree. We note also that, given a tree, when all its labels are fixed (in an arbitrary way), most of the contributions we obtain are vanishing: the mode labels fix uniquely the momenta running through the lines of the tree, and, as a consequence of the previous paragraph, only two scale labels associated to that line are possible. This simply means that the sum over the trees in (3.1) is restricted over all the *compatible* trees. This property has been taken into account in the labels counting in the text.

bound G_0^k for some positive constant G_0 follows immediately, *as far as the resonances are neglected*, (see (5.11) below). The claimed estimate of the sum in item (4) is a consequence of the following result.

5.3. LEMMA (SIEGEL-BRYUNO'S LEMMA). *The following bound holds for the number of lines $\lambda \in \vartheta$ with scale $n_\lambda = n \le 0$ and $R_\lambda = j$, which we denote by $N_n^j(\vartheta)$:*

$$N_n^1(\vartheta) + 2N_n^2(\vartheta) \le 8\, M(\vartheta) 2^{(n+2)/\tau} + \sum_{\substack{T \\ n_T = n}} \Big[-2 + \sum_{j=1}^{2} j\, m_T^j(\vartheta) \Big] , \tag{5.4}$$

where $m_T^j(\vartheta)$ is the number of resonances V's of ϑ inside the cluster T, having resonance-scale $n_{\lambda_V} = n_T$ and $R_{\lambda_V} = j$, being R_{λ_V} defined in (3.10).

This is an adaptation of the Bryuno's proof, [B], of the Siegel's lemma, [S1], as it is presented in [Pö] and [G7]: a proof is in Appendix A1.

Therefore, fixed $M(\vartheta)$, if we define $P(\vartheta) = \sum_{v\in\vartheta} p_v$, we can bound

$$\sum_{\substack{\{\vec\nu_v\}_{v\in\Delta(\vartheta)} \\ \sum_{v\in\vartheta} |\vec\nu_v| = M(\vartheta)}} \Big[\prod_{v\in\Delta(\vartheta)} e^{-\xi|\vec\nu_v|} \, \frac{|\vec\nu_v|^{p_v+1}}{p_v!} \Big] \le e^{-\xi M(\vartheta)} \, \frac{[M(\vartheta)]^{P(\vartheta)+(\ell+1)|\Delta(\vartheta)|}}{[\ell\,|\Delta(\vartheta)|]!\, P(\vartheta)!\, |\Delta(\vartheta)|!} , \tag{5.5}$$

where $|\Delta(\vartheta)| \le k$ and $P(\vartheta) \le d < 2k$.

5.4. *Proof of (5.5).* In the left hand side of (5.5), if we set $D = |\Delta(\vartheta)|$ and $P = P(\vartheta)$, we can bound

$$\begin{aligned}
\prod_{v\in\Delta(\vartheta)} \frac{|\vec\nu_v|^{p_v+1}}{p_v!} &\le \prod_{v\in\Delta(\vartheta)} |\vec\nu_v| \sum_{\sum_{v\in\Delta(\vartheta)} p_v = P} \; \prod_{v\in\Delta(\vartheta)} \frac{|\vec\nu_v|^{p_v}}{p_v!} \\
&\le \Bigg[\sum_{\sum_{v\in\Delta(\vartheta)} p_v = D} \; \prod_{v\in\Delta(\vartheta)} \frac{|\vec\nu_v|^{p_v}}{p_v!} \Bigg] \Bigg[\sum_{\sum_{v\in\Delta(\vartheta)} p_v = P} \; \prod_{v\in\Delta(\vartheta)} \frac{|\vec\nu_v|^{p_v}}{p_v!} \Bigg] \\
&\le \frac{M^P}{P!} \, \frac{M^D}{D!} ,
\end{aligned}$$

where $M = M(\vartheta)$ and in the sum in the first brackets the p_v's are only summation labels; moreover we have

$$\sum_{\substack{\{\vec\nu_v\}_{v\in\Delta(\vartheta)} \\ \sum_{v\in\Delta(\vartheta)} |\vec\nu_v| = M}} 1 \le \frac{M^{\ell\, D}}{[\ell\, D]!} ,$$

so that (5.5) follows. ■

Then we proceed in the following way in order to bound the small divisors contribution. Let be n_0 a negative integer value to be fixed later. We can write

$$\prod_{n\le 0}\prod_{j=1}^{2} 2^{-jnN_n^j(\vartheta)} \le 2^{-2n_0 k} \prod_{n\le n_0}\prod_{j=1}^{2} 2^{-jnN_n^j(\vartheta)} \; , \tag{5.6}$$

where the scale label $n=1$ can be forgotten, since the propagators on scale $n=1$ can be bounded by 1. Then the product in the right hand side of (5.6) can be bounded by

$$\begin{aligned} & \prod_{n\le n_0}\prod_{j=1}^{2} 2^{-jnN_n(\vartheta)} \\ & \quad \le \exp\Big[8\,\ln 2\, M(\vartheta) \sum_{n=-\infty}^{n_0} (-n2^{n/\tau})\Big] \prod_{n\le 0} \prod_{\substack{T \\ n_T=n}} 2^{2n} \prod_{j=1}^{2} 2^{-jnm_T^j(\vartheta)} \; , \end{aligned} \tag{5.7}$$

where the product on n in the right hand side is extended also to the scale labels $n>n_0$: obviously this could be avoided, but we keep it so in order to not complicate the analysis of the cancellations, since we are not looking for an optimal bound. The sum in the argument of the exponential, in the right hand side of (5.7), can be bounded by $c_{n_0} M(\vartheta)\, 2^{n_0/\tau}$, for some constant $c_{n_0}>0$ depending on n_0 and explicitly computable:

$$c_{n_0} = 8\,\ln 2 \sum_{p=0}^{\infty} (p-n_0)\, 2^{-p/\tau} \; .$$

Therefore, fixed the value ξ, we can choose $n_0=n_0(\xi)$, such that $c_{n_0(\xi)} 2^{n_0(\xi)/\tau} < \xi/2$, so that we can bound

$$\begin{aligned} & \prod_{v\in\Delta(\vartheta)} e^{-\xi|\vec\nu_v|} \, \frac{|\vec\nu_v|^{p_v+1}}{p_v!} \prod_{n\le 0}\prod_{j=1}^{2} 2^{-jnN_n^j(\vartheta)} \\ & \quad \le e^{-\xi M(\vartheta)/2} \, \frac{[M(\vartheta)]^{P(\vartheta)+(\ell+1)|\Delta(\vartheta)|}}{[\ell\,|\Delta(\vartheta)|]!\,P(\vartheta)!\,|\Delta(\vartheta)|!} \, 2^{-2n_0(\xi)k} \prod_{n\le 0}\prod_{T, n_T=n} 2^{2n} \prod_{j=1}^{2} 2^{-jnm_T^j(\vartheta)} \; , \end{aligned} \tag{5.8}$$

(note that such a $n_0(\xi)$ depends on τ as $a_1\tau\ln\tau$, for some constant $a_1<0$; if $\tau>\ell-1$, then we have $2^{-2n_0(\xi)} \sim (\ell!)^{-a_2}$, for some positive constant a_2). If we take into account that

$$\frac{[M(\vartheta)]^{P(\vartheta)+(\ell+1)|\Delta(\vartheta)|}}{[\ell\,|\Delta(\vartheta)|]!\,P(\vartheta)!\,|\Delta(\vartheta)|!} \le C_4^{-[P(\vartheta)+(\ell+1)|\Delta(\vartheta)|]} \, \exp[C_4\, M(\vartheta)] \le C_4^{-(\ell+3)k} \, \exp[C_4\, M(\vartheta)] \; , \tag{5.9}$$

we can choose C_4 small enough so that $C_4 \le \xi/4$, and write, in (5.3),

$$\sum_{\Delta(\vartheta)} = \sum_{D=0}^{k} \sum_{\substack{\Delta(\vartheta) \\ |\Delta(\vartheta)|=D}} = \sum_{D=0}^{k} \binom{k}{D} = 2^k , \tag{5.10}$$

so obtaining, *if we neglegt the resonances*, a well defined expression which is summable on $M(\vartheta)$ and gives a bound G_0^k for some positive constant G_0, whose (non optimal) value can be deduced from the above discussion:

$$G_0 = 2^{15}\,\mathcal{C}\,(4/\xi)^{\ell+3}\,2^{-2n_0(\xi)} .$$

Then we have obtained a bound on $\vec{X}_{\vec{\nu}}^{(k)}(\zeta)$ of the form

$$|\vec{X}_{\vec{\nu}}^{(k)}(\zeta)| \le e^{-\xi|\vec{\nu}|/4}\,G_1\,G_0^k G_2^k . \tag{5.11}$$

for some constant $G_1 > 0$, (the above discussion gives $G_1 = 2^{-7}[1-\exp(-\xi/4)]^{-1}$), and for $G_2 = 1$ if we neglect the resonances.

However the presence of resonances has the effect that, for each tree ϑ, we have to take into account also the factor

$$\Big[\prod_{n\le 0}\prod_{\substack{T \\ n_T=n}} 2^{2n}\,2^{-(2nm_T^2(\vartheta)+nm_T^1(\vartheta))}\Big] , \tag{5.12}$$

arising from the resonant lines (see Definition 5.1). It is possible to show that there are trees ϑ such that the dimensional bound (5.1) gives a behaviour $C^k(k!)^\alpha$, for some positive constants C and α, (an explicit example can be found in [E1], §II). By taking into account the cancellations occurring between the various tree values contributing to the same perturbative order, it is possible to see that a bound (5.11) is still possible, for some constant $G_2 > 1$. To the proof of such an assertion next two sections are devoted.

6. Approximate cancellations of the resonances

In this section, it will be more convenient to work with the numbered trees, (see 3.2). In fact, if we recall (3.3) and we consider numbered trees, (4.4) has to be replaced with

$$\mathrm{Val}(\vartheta) = \frac{1}{d!}\prod_{v\in\vartheta} O_v\,g^{(n_{\lambda_v})}(\vec{\omega}\cdot\vec{\nu}_{\lambda_v}) , \tag{6.1}$$

and (3.1) becomes

$$\vec{X}_{\vec{\nu}}^{(k)}(\zeta) = \sum_{\vartheta\in\mathcal{T}_k^{(n)}} \mathrm{Val}(\vartheta) , \tag{6.2}$$

where the sum is over all the labeled numbered trees of order k. Note that two trees have to be regarded as identical if they are topologically equivalent, (*i.e.* they are superposable modulo a transformation of the group $\mathcal{G}$ defined in 3.1), and all their labels (included the numbers associated to the branches) match, (see also Remark 3.5). The advantage of dealing with the perturbative expansion in terms of numbered trees is that in such a way each tree is "weighted" with the same combinatorial factor.

Let us introduce some notations to classify the resonances: this will be useful in the following.

6.1. DEFINITION (RESONANCE FACTOR). *Let be given a resonance V; let λ_V be the incoming line, as in Definition 5.1. We denote by w_0 the vertex from which the outgoing line of V comes out, and by w_1 the vertex from which the incoming line of V comes out, (then the vertex w_1 is outside the resonance, but $w_0 \in V$). Let us consider the labels $\zeta^2_{w_0}$ and $\zeta^1_{w_1}$, which can assume only the values H and h, (by construction the value μ is forbidden in such cases). We denote by $\mathcal{P}(w_0, w_1')$ the (unique) path leading from w_0 to w_1', being w_1' the vertex immediately preceding w_1, (i.e. $\lambda_V \equiv \lambda_{w_1}$ is the line $w_1' \leftarrow w_1$). Let us define the* resonance factor $\mathcal{V}^{n_{\lambda_V}}_{\zeta^2_{w_0},\zeta^1_{w_1}}(\vec{\omega}\cdot\vec{\nu}_{\lambda_V})$ *as the quantity*

$$\mathcal{V}^{n_{\lambda_V}}_{\zeta^2_{w_0},\zeta^1_{w_1}}(\vec{\omega}\cdot\vec{\nu}_{\lambda_V}) = \prod_{w\in V} O_w \prod_{\lambda\in V} g^{(n_\lambda)}(\vec{\omega}\cdot\vec{\nu}_\lambda) \ , \tag{6.3}$$

where the first product is over all the $d(V)$ vertices inside the resonance V and the second one is over the $d(V)-1$ lines internal to V. Let us define V_0 as the collection of lines and vertices in V external to the maximal resonances contained inside V, (we say that a line is in V_0 if both its extremes are in V and at least one of them is in V_0).

We modify the rules how to construct the trees by splitting each resonance factor $\mathcal{V}$ as $\mathcal{V} = \mathcal{L}\mathcal{V} + (\mathbb{1} - \mathcal{L})\mathcal{V}$, where

$$\begin{aligned} \mathcal{L}\mathcal{V}^n_{h,h}(\vec{\omega}\cdot\vec{\nu}) &= \mathcal{V}^n_{h,h}(0) + [\vec{\omega}\cdot\vec{\nu}]\,\dot{\mathcal{V}}^n_{h,h}(0) \ , \\ \mathcal{L}\mathcal{V}^n_{H,h}(\vec{\omega}\cdot\vec{\nu}) &= \mathcal{V}^n_{H,h}(0) \ , \\ \mathcal{L}\mathcal{V}^n_{h,H}(\vec{\omega}\cdot\vec{\nu}) &= \mathcal{V}^n_{h,H}(0) \ , \\ \mathcal{L}\mathcal{V}^n_{H,H}(\vec{\omega}\cdot\vec{\nu}) &= 0 \ , \end{aligned} \tag{6.4}$$

where $\dot{\mathcal{V}}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}(0)$ denotes the first derivative of $\mathcal{V}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}$ with respect to $\vec{\omega}\cdot\vec{\nu}$, computed in $\vec{\omega}\cdot\vec{\nu} = 0$. The operator $\mathcal{L}$ will be called *localization operator*.

6.2. *Remark.* Note that the resonance factors depend on $\vec{\omega}\cdot\vec{\nu}$ only through the propagators. Then, for each line λ inside the resonance, the momentum flowing in it is given by $\vec{\nu}_\lambda \equiv \vec{\nu}^0_\lambda + \varepsilon_\lambda\vec{\nu}$, where $\vec{\nu}^0_\lambda$ is the sum of the mode labels corresponding to the vertices following λ but inside the resonance, and $\varepsilon_\lambda = 1$ if $\lambda \in \mathcal{P}(w_0, w_1')$, and $\varepsilon_\lambda = 0$ otherwise. Even if we set

$\vec{\omega}\cdot\vec{\nu}=0$, (*i.e.* $\vec{\omega}\cdot\vec{\nu}_\lambda=\vec{\omega}\cdot\vec{\nu}^0_\lambda$ for each λ inside the resonance), no too small divisor appears because of the presence of the compact support functions $\chi_{n_\lambda}(\vec{\omega}\cdot\vec{\nu}_\lambda)$, $n_\lambda>n$.

Given a tree, on any cluster the $\mathcal{L}$ or $\mathbb{1}-\mathcal{L}\equiv\mathcal{R}$ operators apply. We want to show in this section that the contributions arising from diagrams containing resonances on which the $\mathcal{L}$ operator applies add to zero, so that we can rule out such contributions and consider only trees with resonances on which the operator $\mathcal{R}$ applies. The latter kind of trees will be studied in §7, where the convergence of the Lindstedt series will be eventually proven.

6.3. Definition (Resonance family). *Given a tree ϑ with some resonances, let us consider the family of trees $\mathcal{F}_V(\vartheta)$ obtained from ϑ in the following way. Given a resonance V in ϑ, if $\zeta^1_{w_1}=h$, we add to ϑ the trees we obtain by detaching from the resonance the subtree with root in w_1', then reattaching it to all the remaining vertices $w\in V$ having $\delta_w=1$ and external to the resonances internal to V; if $\zeta^2_{w_0}=h$, to the just considered trees we add all the trees we obtain by detaching the outgoing line of the resonance from the vertex w_0, then reattaching to all the remaining vertices $w\in V$ having $\delta_w=1$ and external to the resonances internal to V. The number of terms so obtained is $\tilde{k}(V)^2$, (if $\zeta^2_{w_0}=\zeta^1_{w_1'}=h$), and $\tilde{k}(V)$, (if only one of the two labels $\zeta^2_{w_0}$, $\zeta^1_{w_1'}$ assumes the value h), where $\tilde{k}(V)\equiv k(V)-\sum_{V'\subset V}k(V')$, being the sum extended over all the maximal resonances internal to V. We call $\mathcal{F}_V(\vartheta)$ a* resonance family *(associated to the resonance V).*

The simplest case of resonance family is drawn in Fig.6.1.

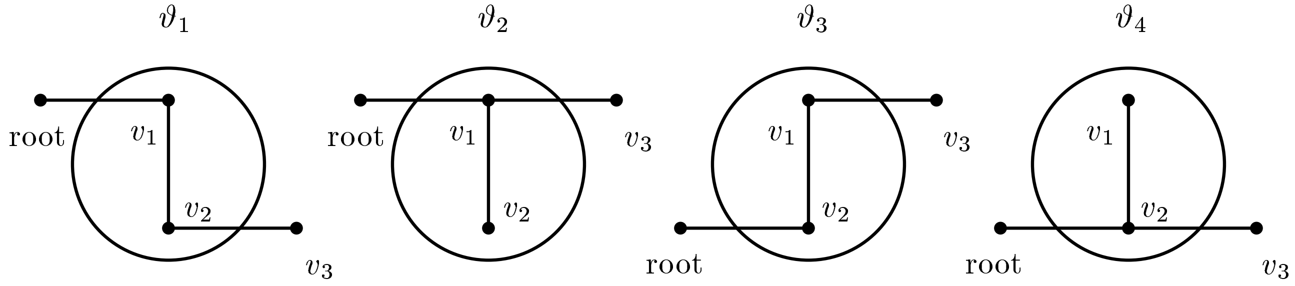

Fig.6.1. The possible resonance families $\mathcal{F}_V(\vartheta)$'s associated to resonance V's with degree $d(V)=2$: (1) if $\zeta^2_{w_0}=\zeta^1_{w_1}=h$, one has only one family with four trees, $\mathcal{F}_V(\vartheta_1)=\{\vartheta_1,\ldots,\vartheta_4\}$, (2) if $\zeta^2_{w_0}=H$ and $\zeta^1_{w_1}=h$, one has two families with two trees, $\mathcal{F}_{V_1}(\vartheta_1)=\{\vartheta_1,\vartheta_2\}$ and $\mathcal{F}_{V_2}(\vartheta_3)=\{\vartheta_3,\vartheta_4\}$, (3) if $\zeta^2_{w_0}=H$ and $\zeta^1_{w_1}=h$, one has two families with two trees, $\mathcal{F}_{V_1}(\vartheta_1)=\{\vartheta_1,\vartheta_4\}$ and $\mathcal{F}_{V_2}(\vartheta_2)=\{\vartheta_2,\vartheta_3\}$, (4) if $\zeta^2_{w_0}=\zeta^1_{w_1}=H$, one has four one-tree families $\mathcal{F}_{V_i}(\vartheta_i)=\{\vartheta_i\}$, $i=1,\ldots,4$. Note that, unlike the labels $\zeta^2_{w_0}$ and $\zeta^1_{w_1}$, the location of the vertices w_0 and w_1' varies inside the family $\mathcal{F}_V(\vartheta)$. For instance, in the case (1), one has $w_0=v_1$ in ϑ_1 and ϑ_2, and $w_0=v_2$ in ϑ_3 and ϑ_4, and, analogously, $w_1'=v_1$ in ϑ_2 and ϑ_3, and $w_1'=v_2$ in ϑ_1 and ϑ_4. Note also that, if $\vartheta'\in\mathcal{F}_V(\vartheta)$, then $\mathcal{F}_V(\vartheta)=\mathcal{F}_V(\vartheta')$: this simply means that a resonance family can be defined with respect to any tree it contains. For instance, in item (1), one can define the resonance family as $\mathcal{F}_V(\vartheta_i)$, $\forall i=1,\ldots,4$. The resonance V, containing the vertices v_1 and v_2, has scale $n_V\ge n_{\lambda_V}+1$, if n_{λ_V} is the scale of the resonant line $\lambda_V\equiv\lambda_{v_3}$ (which is equal to the scale of the line entering the root): the line connecting the vertices v_1 and v_2 is a line $v_1\leftarrow v_2$ in ϑ_1 and ϑ_2, and is a line $v_2\leftarrow v_1$ in ϑ_3 and ϑ_4.

Note that each time we shift the outgoing line of a resonance, we produce an *apparently* deep change on the tree value. In fact all the arrows superposed to the lines point to the tree root, so that all the arrows of the lines inside the resonance have to point to the vertex w_0 to which the outgoing line is reattached, (note that in general, given a tree $\vartheta' \in \mathcal{F}_V(\vartheta)$, we call w_0 the vertex from which the outgoing line of the resonance emerges, and w_1' the vertex which the incoming line $w_1' \leftarrow w_1$ enters: then the location of w_0 and w_1 depends on the particular tree in $\mathcal{F}_V(\vartheta)$, while the values $\zeta_{w_0}^2$ and $\zeta_{w_1'}$ are the same for any tree contained in the resonance family, so that their values do not depend on the vertices w_0 and w_1'; see also Fig.6.1). This means that some arrows superposed to the branches inside the resonance change their direction, and, correspondingly, the lines $h \leftarrow H$ become lines $H \leftarrow h$ and *viceversa.* The reason why we say that the change is only apparently deep will become clear in Lemmata 6.4 and 6.6 below.

In general, fixed the value of the momentum $\vec{\nu}_\lambda$ associated to the line λ, only two scale labels n_λ give a propagator $g^{(n_\lambda)}(\vec{\omega} \cdot \vec{\nu}_\lambda)$ which is not vanishing, (see note 5). It can happen that the shift of the incoming and outgoing lines produces a change of the value of the momenta flowing through the lines inside the resonance. This means that, fixed the scale labels inside the resonance V of a tree ϑ, some of the resonance factors associated to trees in $\mathcal{F}_V(\vartheta)$ (hence obtained by shifting of external lines) can vanish, as containing vanishing propagators.

However we can proceed in a different (and more suitable way): we ignore the fact that, fixed all the labels, there are vanishing contributions, and we consider all the trees inside the family $\mathcal{F}_V(\vartheta)$ as possible (*i.e.* as they were compatible trees). In fact we shall see below that each (vanishing and not vanishing) resonance factor $\mathcal{V}$ can be written as sum of two parts, $\mathcal{V} = \mathcal{L}\mathcal{V} + \mathcal{R}\mathcal{V}$, such that (1) the first one, $\mathcal{L}\mathcal{V}$, is exploited in order to obtain a cancellation, while (2) the latter, $\mathcal{R}\mathcal{V}$, can be easily bounded. The proof of assertion (1) is given in the remaining part of this section and leads to Corollary 6.8, and the to the proof of assertion (2) next section is devoted. Ovbiously, when the resonance factor is vanishing, this simply means that the two parts of the decomposition are equal and opposite,[3] so

[3] The fact that the two parts into which a vanishing resonance factor is decomposed are opposite, but are bounded in a different way is not so surprising. In fact we shall see in §7 that a gain $(2^{n_{\lambda_V}}/2^{n_V})^R$, $R = 1, 2$, is obtained for the part $\mathcal{R}\mathcal{V}$ of the resonance factor associated to a resonance V. If the resonance factor $\mathcal{V}$ is vanishing, then there has to be some vanishing propagator $g^{(n_\lambda)}(\vec{\omega} \cdot \vec{\nu}_\lambda)$ corresponding to a line λ inside V. If $\vec{\omega} \cdot \vec{\nu}_\lambda^0$ is far from the boundary of the support of the χ_{n_λ} function, *i.e.* centered near to 2^{n_λ}, (the momentum $\vec{\nu}_\lambda^0$ is defined in Remark 6.2), then $\vec{\omega} \cdot \vec{\nu}$ has to be large enough so that $\vec{\omega} \cdot \vec{\nu}_\lambda = \vec{\omega} \cdot \vec{\nu}_\lambda^0 + \vec{\omega} \cdot \vec{\nu}$ is outside the interval $[2^{n_\lambda - 1}, 2^{n_\lambda + 1}]$: in other words $\vec{\omega} \cdot \vec{\nu}$ has to be comparable with $\vec{\omega} \cdot \vec{\nu}_\lambda^0$, and so the ratio $2^{n_{\lambda_V}}/2^{n_V}$ is bounded below by a constant and no true gain is obtained with respect to the part $\mathcal{L}\mathcal{V}$ of the decomposition. On the contrary, if $\vec{\omega} \cdot \vec{\nu}_\lambda^0$ is near to the boundary, the considered ratio is small, but in such a case the function χ_{n_λ} appearing in $\mathcal{L}\mathcal{V}$ is much lesser than 1, and the $\mathcal{L}\mathcal{V}$ is much smaller than the dimensional bound can suggest. In other words, in case (1) $\mathcal{R}\mathcal{V}$ is not truely smaller than $\mathcal{L}\mathcal{V}$, in case (2) the $\mathcal{L}\mathcal{V}$ is not truely larger than $\mathcal{R}\mathcal{V}$, and we can conclude that, when the resonance factor is vanishing, the bounds on the two parts into which it is decomposed are only apparently different.

that the first one is not really dangerous, and could be directly bounded without using the cancellations. However one of the advantages of the techniques we are using is that we can avoid to distinguish between dangerous and not dangerous contributions, and we can treat all resonances in the same way, (see also the comments in §1). Note that the number of extra trees we are counting by considering also the vanishing contributions is bounded by C_3^k, for some constant $C_3 > 0$.

Then the following lemmata hold, which permit us to control the differences between the values of the trees contained in the same resonance family.

6.4. LEMMA. *Consider a resonance family $\mathcal{F}_V(\vartheta)$ obtained from a tree ϑ such that $V \in \vartheta$ is a resonance, and set $\vec{\nu} = \vec{0}$, if $\vec{\nu}$ is the momentum flowing through the branch entering the resonance ($\vec{\nu} = \vec{\nu}_{\lambda_V}$). Then the product of the operator associated to any line $\lambda \in V$, when applied to the function (3.6), times the corresponding propagator assumes the same value for each tree $\vartheta' \in \mathcal{F}_V(\vartheta)$.*

6.5. *Proof of Lemma 6.4.* That the product of the operators times the propagators associated to the lines in ϑ do not change by setting $\vec{\nu} = \vec{0}$ and shifting the incoming line $w_1' \leftarrow w_1$, hence the point w_1', is trivial, and is due simply to the fact that (1) if $\zeta_{w_1}^1 = h$, the line $\lambda_{w_1} \equiv \lambda_V$ does not carry any derivative acting on $f_{\vec{\nu}_{w_1'}}(\vec{A}_{w_1'})$, and (2) if $\vec{\nu} = \vec{0}$, the incoming line of the resonance does not contribute anymore to the momenta flowing along the branches $\lambda \in V$, so that no memory is left about the vertex which the line λ_V enters, (see Remark 6.2).

The independence of the product of the operators times the propagators on the location of the vertex w_0 from which the outgoing line of the resonance emerges can be argued in the following way. Note that, *only* if $\zeta_{w_0}^2 = h$, the outgoing line λ_{w_0} is shifted, and it is reattached *only* to vertices with $\delta_w = 1$, (by construction of the family $\mathcal{F}_V(\vartheta)$).

The arrows of the lines inside the resonance point all to the tree vertex w_0'. If we detach the outgoing line from $w_0 \equiv w_2$, and reattach it to another vertex, say w_3, then the direction of the arrows can change or not, in according to the following rule: if the line either is contained simultaneously in both paths $P(w_2, w_1')$ and $P(w_3, w_1')$, or is not contained in none of the two paths, then the arrow direction does not change, while, if the line is in only one of the two paths, its direction changes. Obviously, only the propagators and operators corresponding to lines with reversed arrows can be different.

To treat such lines, we start by noting that each line $v' \leftarrow v$ inside the resonance V realizes a partition of the resonance into two subsets W_v^1 and W_v^2 such that $W_v^1 = \{w \in V : w \geq v\}$ and $W_v^2 = V \setminus W_v^1$. Since $\sum_{w \in V} \vec{\nu}_w = \vec{0}$, we have trivially $\sum_{w \in W_v^1} \vec{\nu}_w = -\sum_{w \in W_v^2} \vec{\nu}_w$. We define $\vec{\nu}(v) = \sum_{w \in W_v^1} \vec{\nu}_w$, and we (arbitrarily) fix the set W_v^1 as the collection of vertices $w \geq v$ in V when the momentum flows from v to v', (the same convention will be followed henceforth).

Then if the arrow direction of the line λ_v changes, this means that the line λ_v should be denoted by $\lambda_{v'}$, and the momentum flowing through it (by setting $\vec{\omega} \cdot \vec{\nu}$ and recalling the definition of $\vec{\nu}^0_\lambda$ given in Remark 6.2) is no more $\vec{\nu}^0_{\lambda_v} = \vec{\nu}(v) = \sum_{w\in W^1_v}\vec{\nu}_w$, but becomes $\vec{\nu}^0_{\lambda_{v'}} = \sum_{w\in W^2_v}\vec{\nu}_w \equiv -\vec{\nu}(v)$, so that it simply changes its direction, *i.e.* its sign. Note that we are taking into account the fact that, when the arrow changes its direction, then the momentum flowing through the line λ_v becomes the sum of the mode labels associated to the vertices in W^2_v, which assumes now the same rôle that W^1_v had before the arrow reversal: then if the arrow points from v to v', we have $\nu^0_{\lambda_v} = \vec{\nu}(v)$, while if it points from v' to v, we have $\vec{\nu}^0_{\lambda_v} = -\vec{\nu}(v)$.

Then, if we look at (3.4), we see that, given a line $\lambda_v \in V$, the case $\zeta^1_v = \zeta^2_v = h$ can be dealt with by considering the product of the operator times the propagator, and by noting that the matrix $T(\vec{A}_0)$ is symmetric, so that the numerator does not change by arrow reversal, and the same the denominator does, depending quadratically on the momentum.

The other two cases $\zeta^1_v = h$, $\zeta^2_v = H$ and $\zeta^2_v = H$, $\zeta^1_v = h$ can be easily dealt with, by noting that if we change the arrow direction, we change the sign of the numerator, (see Remark 3.4), and reverse the direction of the momentum, so changing also the sign of the denominator: then the overall sign remains the same. ■

It can be useful to define the quantity $\mathcal{W}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}(\vec{\omega}\cdot\vec{\nu})$ as

$$\mathcal{V}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}(\vec{\omega}\cdot\vec{\nu}) = (-i\mathcal{M}\,\vec{\nu}_{w_0})\,(i\vec{\nu}_{w'_1})\,\mathcal{W}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}(\vec{\omega}\cdot\vec{\nu})\,, \tag{6.5}$$

where the matrix $\mathcal{M}$ is defined to be $\mathcal{M} = \mathbb{1}$, (if $\zeta^1_{w_0} = H$), or $\mathcal{M} = T(\vec{A}_0)$, (if $\zeta^1_{w_0} = h$).

6.6. Lemma. *Consider a resonance family $\mathcal{F}_V(\vartheta)$, and let $\mathcal{V}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}(\vec{\omega}\cdot\vec{\nu})$ be the resonance factor associated to the resonance V in ϑ. Suppose that $\zeta^2_{w_0} = \zeta^1_{w_1} = h$: this means that, in according to which vertex w_0 the outgoing line of the resonance comes out from and to which vertex w'_1 the incoming line enters, the resonance factor contains the two factors $(-i\mathcal{M}\vec{\nu}_{w_0})$ and $i\vec{\nu}_{w'_1}$. All the others factors appearing in the O_v's, $v\in V$, are the same ones for all trees in $\mathcal{F}_V(\vartheta)$. Then if we sum together all the contributions* to first order in *$\vec{\omega}\cdot\vec{\nu}$ arising from the family $\mathcal{F}_V(\vartheta)$, we see that they differ only as far as the resonance factors are concerned, so that we obtain a factorising term times the following expression*

$$\begin{aligned}\sum_{\vartheta'\in\mathcal{F}_V(\vartheta)} &(\vec{\omega}\cdot\vec{\nu})\dot{\mathcal{V}}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}(0)\\ &= (\vec{\omega}\cdot\vec{\nu})\,\mathcal{W}^n_{h,h}(0)\sum_{v\in V}\Big[-\frac{R_{\lambda_v}}{\vec{\omega}\cdot\vec{\nu}(v)} + \frac{\dot{\chi}_{n_{\lambda_v}}(\vec{\omega}\cdot\vec{\nu}(v))}{\chi_{n_{\lambda_v}}(\vec{\omega}\cdot\vec{\nu}(v))}\Big]\cdot\\ &\cdot\Big[\sum_{\substack{w_0\in W^2_v\cap V_0\\ w'_1\in W^1_v\cap V_0}} - \sum_{\substack{w_0\in W^1_v\cap V_0\\ w'_1\in W^2_v\cap V_0}}\Big](-i\mathcal{M}\,\vec{\nu}_{w_0})(i\vec{\nu}_{w'_1})\,,\end{aligned} \tag{6.6}$$

where:

(1) in the left hand side $\dot{\mathcal{V}}_{\zeta^2_{w_0},\zeta^1_{w_1}}(0)$ depends on the tree ϑ', (though the dependence is not

explicitly shown);
(2) the matrix $\mathcal{M}$ is defined to be $\mathcal{M} = 1\!\!1$ or $\mathcal{M} = T(\vec{A}_0)$;
(3) $\dot{\chi}_{n_{\lambda_v}}(\vec{\omega}\cdot\vec{\nu}(v))$ denotes the first derivative of $\chi_{n_{\lambda_v}}(\vec{\omega}\cdot\vec{\nu}(v))$ with respect to its argument;
(4) the quantity $\mathcal{W}_{h,h}(0)$ is defined as in (6.5), with $\zeta^2_{w_0} = \zeta^1_{w_1'} = h$; and
(5) $\vec{\nu}(v) \equiv \sum_{w\in W_v^1} \vec{\nu}_w$; this means that $\vec{\nu}(v) = \sum_{v\le w\in V_0} \vec{\nu}_w = \vec{\nu}^0_{\lambda_v}$ only if $w_0 \in W_v^2$, if $\vec{\nu}^0_{\lambda_v}$ is defined as in Remark 6.2.

6.7. *Proof of Lemma 6.6.* It is enough to note that only the propagators of the lines inside V can depend on $\vec{\omega}\cdot\vec{\nu}$, so that, when we derive a term of the form $\chi_{n_\lambda}(\vec{\omega}\cdot\vec{\nu}_\lambda)\,[i\vec{\omega}\cdot\vec{\nu}_\lambda]^{-R_\lambda}$ we obtain $-i\,R_\lambda\,\chi_{n_\lambda}(\vec{\omega}\cdot\vec{\nu}_\lambda)\,[i\vec{\omega}\cdot\vec{\nu}_\lambda]^{-R_\lambda-1} + \dot{\chi}_{n_{\lambda_v}}(\vec{\omega}\cdot\vec{\nu}(v))\,[i\vec{\omega}\cdot\vec{\nu}_\lambda]^{-R_\lambda}$, each time the line λ is along the path $\mathcal{P}(w_0, w_1')$ and zero otherwise. Then we sum together all the contributions arising from the family $\mathcal{F}_V(\vartheta)$, taking into account the fact that $\mathcal{V}_{h,h}(0)$ assume the same value for all contribution, (by Lemma 6.4), and rearrange the sum, by (1) fixing the line and considering together all the trees in which the path $\mathcal{P}(w_0, w_1')$ contains that line, and (2) exploiting the fact that the momentum $\vec{\nu}^0_{\lambda_w}$ flowing through the line λ_w is directed from w_1 toward w_0, so that it is $\vec{\nu}(w)$ if $w_0 \in W_v^2$ and $-\vec{\nu}(w)$ otherwise. We are taking into account also the fact that $\chi_{n_\lambda}(\vec{\omega}\cdot\vec{\nu}_\lambda)$ is an even function of its argument (see the definition given in §4), so that the first derivative is odd. The condition $w_0, w_1' \in V_0$ follows from the definition of the resonance family $\mathcal{F}_V(\vartheta)$. ■

From Lemma 6.4 and Lemma 6.6, we deduce immediately the following corollary, which can be considered an extension of the corresponding results of [G7], [CF1] and [GM2].

6.8. COROLLARY. *If we sum together all the values corresponding to the trees ϑ' contained in a resonance family $\mathcal{F}_V(\vartheta)$, then the contributions obtained by applying the $\mathcal{L}$ operator to the resonance factors identically vanish.*

6.9. *Proof of Corollary 6.8.* As we have said, the proof is an easy consequence of Lemma 6.4 and Lemma 6.6. If $\zeta^1_{w_2} = h$, we obtain a vanishing contribution to first order, because we have cancellations between the terms in $\mathcal{F}_V(\vartheta)$ we obtain by shifting the incoming branch. In fact by Lemma 6.4 the various contributions we obtain differ only because of a factor $i\vec{\nu}_{w_1'}$ associated to the vertex w_1': when such an operation is performed we can choose as w_1' vertices only the vertices $w \in V_0$ having $\delta_w = 1$, but

$$\sum_{\substack{w\in V_0\\ \delta_w=1}} \vec{\nu}_w \equiv \sum_{w\in V} \vec{\nu}_w \equiv \vec{0}\,, \tag{6.7}$$

since $\vec{\nu}_w \equiv \vec{0}$ for w with $\delta_w = 0$ and $\sum_{w\in V'} \vec{\nu}_w = \vec{0}$ for any resonance $V' \subset V$, so that $\sum_{w_1'\in V_0} \vec{\nu}_{w_1'} = \vec{0}$.

The same happens if $\zeta^2_{w_0} = h$, and the cancellations are between the terms we obtain by shifting the outgoing branch.

If both $\zeta^2_{w_0} = h$ and $\zeta^1_{w_1} = h$, we have a zero to second order, by Lemmata 6.4 and 6.6. In fact the zero to first order follows from the above discussion, while the zero to second order can be easily deduced from (6.6): if we perform explicitly the sums over the vertices w_0 and w_1, we obtain, *for any fixed* v, $[-i\mathcal{M}\,(-\vec{\nu}(v))]\,[i\vec{\nu}(v)]\;-\;[-i\mathcal{M}\,\vec{\nu}(v)]\,[i(-\vec{\nu}(v))]$, which is trivially zero, (the matrix $\mathcal{M}$ being defined after (6.5)).

Then, if we recall the definition of the localization operator, see (6.4), the statement of Corollary 6.8 follows immediately. ■

7. Convergence of the Lindstedt series for KAM tori

In order to prove that $|\vec{X}^{(k)}_{\vec{\nu}}(\zeta)| \le C^k$, for some constant C, we shall find convenient to modify the definition of the functionals to associate to the lines $h \leftarrow H$.

Obviously the proof can be carried out even if such a change is not introduced: however it will turn out to simplify the analysis in a relevant way. If we consider the graph rules given in §3 (see in particular (3.4)) and the definition of localization operator in (6.4), we can check that the are:
(1) resonances such that $\zeta^2_{w_0} = \zeta^1_{w_1} = H$ (and with external lines λ's having by construction $R_\lambda = 1$), for which no gain is obtained,
(2) and resonances such that $\zeta^2_{w_0} = \zeta^1_{w_1} = h$ and the incoming line is a $h \leftarrow H$ line (*i.e.* $\zeta^2_{w_1} = H$), for which a larger than needed gain is obtained.

The reason why we say that the gain obtained for resonances in item (2) is larger than how it is needed can be understood by considering how the cancellation mechanism described in §6 works.

Each time we have a resonance, there are two lines λ_{w_0} and $\lambda_{w_1} \equiv \lambda_V$ (the external lines of the resonance) such that we can interpret the product of the corresponding propagators as an "effective propagator" $[\vec{\omega}\cdot\vec{\nu}_{\lambda_V}]^{-R_{\lambda_{w_0}}-R_{\lambda_V}}$; if we have a chain of resonances, *i.e.* a sequence of resonances $V_1,\ldots,V_{\mathcal{N}}$, such that the incoming line of V_i, $i=1,\ldots,\mathcal{N}-1$ is the outgoing line of V_{i+1}, we can associate to it an "effective propagator"

$$[\vec{\omega}\cdot\vec{\nu}_{\lambda_{V_1}}]^{-R_{\lambda_{w_0}}}\prod_{i=1}^{\mathcal{N}}[\vec{\omega}\cdot\vec{\nu}_{\lambda_{V_1}}]^{-R_{\lambda_{V_i}}}\ ,$$

where $\vec{\nu}_{\lambda_{V_1}} = \ldots = \vec{\nu}_{\lambda_{V_{\mathcal{N}}}}$ and λ_{w_0} is the outgoing line of the first resonance of the chain, *i.e.* V_1.

Then the cancellation of the localized parts allows us to extend the discussion in §5 in such a way to cover also the resonances, if we can obtain a quantity $O([\vec{\omega}\cdot\vec{\nu}_{\lambda_V}]^{R_{\lambda_V}})$, up to factors which can be controlled, from the bound on the resonance factor associated to any resonance V, once its localized part has been subtracted.

In fact problems arise when a chain of resonances appear in a tree, so that there is a lot of "repeated small divisors": but if a gain $O([\vec{\omega}\cdot\vec{\nu}_{\lambda_V}]^{R_{\lambda_V}})$ is obtained for each resonance V in the chain, then the corresponding effective propagator reduces itself to $O([\vec{\omega}\cdot\vec{\nu}_{\lambda_{V_1}}]^{R_{\lambda_{w_0}}})$.

If we look at (6.4), we see that the required property is not fulfilled by the resonances in item (1) above. On the other hand, the resonances in item (2) have a gain which is $O([\vec{\omega}\cdot\vec{\nu}_{\lambda_V}]^{R_{\lambda_V}+1})$.

However, if we consider a single resonance of the kind in item (1), we can simply say that the effective propagator is quadratic, and no problem arises *unless if a chain of such resonances on the same scale occur*. But by construction this is possible only if resonances as in item (2) are inserted, so that they provide the extra required gain.

In order to simplify the analysis and take into account directly the just described compensations along the chains, we slightly change the definition of the functionals associated to the lines $h \leftarrow H$, as anticipated in the beginning of this section. Then the second row in (3.2) will be replaced with

$$\begin{array}{ccc} \text{operator} & \text{propagator} & \text{branch} \\ C_0 \left[i\vec{\nu}_{v'} \cdot (\partial_{\vec{A}_v})(i\vec{\omega} \cdot \vec{\nu}_{\lambda_v}) \right] & [i\vec{\omega} \cdot \vec{\nu}_{\lambda_v}]^{-2} & h \leftarrow H \end{array} \tag{7.1}$$

(in such a way that the product of the operator times the propagator remains the same, as it has to be). Then the propagators will be always of the form (3.8), but $R_{\lambda_v} = 1$ if $\zeta_v^1 = H$ and $R_{\lambda_v} = 2$ if $\zeta_v^1 = h$. Then (5.12) has to be replaced by

$$\left[\prod_{n \le 0} \prod_{\substack{T \\ n_T = n}} 2^{2n} \, 2^{-(2nm_T^2(\vartheta) + nm_T^1(\vartheta))} \right] \prod_{\substack{v \in \vartheta \\ \zeta_v^1 = h, \zeta_v^2 = H}} 2^{n_{\lambda_v}} \ . \tag{7.2}$$

while a bound G_0^k still holds for the resonanceless trees, with the same constant G_0, because some lines which in §5 were counted among $m_T^1(\vartheta)$ contribute now to $m_T^2(\vartheta)$, (they are exactly the resonant lines $h \leftarrow H$).

The definition (7.1) allows us to change the definition of the $\mathcal{L}$ operator and substitute the equations (6.4) with the following ones:

$$\begin{aligned} \mathcal{L}\mathcal{V}_{h,h}^n(\vec{\omega} \cdot \vec{\nu}) &= \mathcal{V}_{h,h}^n(0) + [\vec{\omega} \cdot \vec{\nu}] \, \dot{\mathcal{V}}_{h,h}^n(0) \ , \\ \mathcal{L}\mathcal{V}_{H,h}^n(\vec{\omega} \cdot \vec{\nu}) &= \mathcal{V}_{H,h}^n(0) + [\vec{\omega} \cdot \vec{\nu}] \, \dot{\mathcal{V}}_{h,h}^n(0) \ , \\ \mathcal{L}\mathcal{V}_{h,H}^n(\vec{\omega} \cdot \vec{\nu}) &= \mathcal{V}_{h,H}^n(0) \ , \\ \mathcal{L}\mathcal{V}_{H,H}^n(\vec{\omega} \cdot \vec{\nu}) &= \mathcal{V}_{H,H}^n(0) \ , \end{aligned} \tag{7.3}$$

such that a result analogous to Corollary 6.8 applies to them. More precisely one can easily prove the following result.

7.1. Corollary. *If the operators and the propagators associated to the lines $h \leftarrow H$ are defined as in (7.1), while the other ones as in (3.4), and if the action of the $\mathcal{L}$ operator on the resonance factors is given by (7.3), then, when we sum together all the values corresponding to trees ϑ' contained in a resonance family $\mathcal{F}_V(\vartheta)$, the contributions we obtain by applying the $\mathcal{L}$ operator to the resonance factors identically vanish.*

7.2. *Proof of Corollary 7.1.* With respect to the previous situation, only the cases with $\zeta_{w_0}^2 = H$ behave in a different way, because of the factor $[i\vec{\omega} \cdot \vec{\nu}_{\lambda_{w_0}}]$ appearing in the operator associated to the line λ_{w_0}, (which has to be a line $h \leftarrow H$): then we have a zero to one order higher for $\mathcal{LV}^n_{H,h}(\vec{\omega}\cdot\vec{\nu})$ and $\mathcal{LV}^n_{H,H}(\vec{\omega}\cdot\vec{\nu})$, $\vec{\nu}=\vec{\nu}_{\lambda_V}$. ■

Therefore we can rule out again all the contributions in which the $\mathcal{L}$ operator applies to any resonance, and we are left with resonances on which only the $\mathcal{R}$ operator can act. Obviously, in (7.2) the last product has to be replaced with

$$\prod_{\substack{v\notin W_0(\vartheta)\\ \zeta^1_v=h,\zeta^2_v=H}} 2^{n_{\lambda_v}} \;, \tag{7.4}$$

if $W_0(\vartheta)$ is defined as the set of the vertices in ϑ from which a resonance outgoing line comes out: in fact in such cases, (*i.e.* $v\in W_0$), the factor $[i\vec{\omega}\cdot\vec{\nu}_{\lambda_v}]$ is used in order to implement the cancellation of the resonance factors.

It is convenient to write the effect of $\mathcal{R}$ on a resonance V as

$$\begin{aligned}
&\mathcal{RV}^n_{\zeta,\zeta'}(\vec{\omega}\cdot\vec{\nu}) = (\vec{\omega}\cdot\vec{\nu})\int_0^1 dt\; \dot{\mathcal{V}}^n_{\zeta,\zeta'}(t\vec{\omega}\cdot\vec{\nu}) \qquad\qquad \text{(first order zero)}\;,\\
&\mathcal{RV}^n_{\zeta,\zeta'}(\vec{\omega}\cdot\vec{\nu}) = (\vec{\omega}\cdot\vec{\nu})^2\int_0^1 dt\; t\; \ddot{\mathcal{V}}^n_{\zeta,\zeta'}(t\vec{\omega}\cdot\vec{\nu}) \qquad \text{(second order zero)}\;,
\end{aligned} \tag{7.5}$$

where $\ddot{\mathcal{V}}^n_{\zeta,\zeta'}$ denotes the second derivative with respect to the variable $\vec{\omega}\cdot\vec{\nu}$.

As there are resonances enclosed in other resonances the above formula can suggest that there are propagators derived up to $\approx k$ times, if k is the order of the graph. This would be of course a source of problems, as $a_{R_{\lambda_V}}(p) > p!$, where $a_{R_{\lambda_V}}(p)$ is defined in (4.3). However it is not so: in fact the propagators are derived at most two times. This can be seen as follows. Let n be the resonance-scale of the maximal resonance V, and recall that V_0 is the collection of lines and vertices in V not contained in any resonance internal to V, (see Definition 6.1). Then we can write $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_V})$, (we do not write explicitly the labels of the resonance factor), as

$$\mathcal{R}\Big(\prod_{\lambda\in V_0} g^{(n_\lambda)}(\vec{\omega}\cdot\vec{\nu}_\lambda)\prod_{V'\subset V}[\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}})]\prod_{v\in V_0} O_v\Big)\;, \tag{7.6}$$

being the second product over the resonances $V'\subset V$ which are maximal; in (7.6), for any resonance $V'\subseteq V$, $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}})$ can be written either as in (7.5) or as a difference $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}}) = \mathcal{V}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}}) - \mathcal{LV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}})$, in according to which expression turns out to be more convenient to deal with.

Then the first step is to write the action of $\mathcal{R}$ on the maximal cluster as in (7.5), leaving the other terms $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}})$ written as differences: so (7.6) can be written by the Leibniz's rule as a sum of terms, and the derivatives of $\mathcal{R}$ apply either on some propagator $g^{(n_\lambda)}$ or on some $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}})$. In the end there can be either no derivative, or one derivative, or two derivatives applied on each $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}})$. If only one derivative acts on $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu})$, $\vec{\nu}=\vec{\nu}_{\lambda_{V'}}$, and the zero is of the second order, then we write, when such a term is not vanishing,

$$\partial\mathcal{RV}(\vec{\omega}\cdot\vec{\nu})=\partial\mathcal{V}(\vec{\omega}\cdot\vec{\nu})-\dot{\mathcal{V}}(0)=(\vec{\omega}\cdot\vec{\nu})\int_0^1 dt\,\ddot{\mathcal{V}}(t\vec{\omega}\cdot\vec{\nu})\;,$$

because the derivative with respect to $\vec{\omega}\cdot\vec{\nu}$ is equal to the derivative with respect to $\vec{\omega}\cdot\vec{\nu}_{\lambda_V}$, while if two derivatives act on $\mathcal{RV}(\vec{\omega}\cdot\vec{\nu}_{\lambda_{V'}})$, then we write

$$\partial^2\mathcal{RV}(\vec{\omega}\cdot\vec{\nu})=\ddot{\mathcal{V}}(\vec{\omega}\cdot\vec{\nu})\;.$$

The case of a first order zero is easier, and can be discussed in the same way. Then no more than two derivatives can act on each resonance V' in any case, and the procedure can be iterated, since the resonances V' can be dealt with as the resonance V.

The effect of the $\mathcal{R}$ operator is to obtain a gain factor either $2^{2+n-n'}$ or $2^{2+n-n'}2^{2+n'-n''}$, where n' and n'' are the scales of two lines λ' and λ'' contained in some clusters T' and T'' inside V, n is the resonance-scale, and the factor 2 is due to the support properties of the propagators; the line λ'' can coincide with λ', or also be absent, if there is a first order zero. So we can rewrite, *e.g.* , the first factor as $2^{2+n-n'}=2^2 2^{n-n_1}\ldots 2^{n_q-n'}$, where n_i is the scale of the cluster $T_i\supset T_{i+1}$, with $T_0=V$ and $T_{q+1}=T'$. Analogous considerations hold for n'', so that we can conclude that:
(1) no more than two derivatives can ever act on any propagators;
(2) a gain $2^{R_{\lambda_{V'}}(2+n_{\lambda_{V'}}-n_{V'})}$ is obtained for any resonance $V'\subseteq V$;
(3) the total number of terms generated by the derivation operations is bounded by $k(V)^2$, if $k(V)$ is the order of the resonance V, (see Definition 5.1), as $\sum_{V'\in\vartheta}\tilde{k}(V')=k(V)$.

Therefore, for the value of the diagram formed by the resonance plus its incoming line, we find the bound

$$\begin{aligned}&2^{-R_{\lambda_V}n_V}\Big[2^{4k}\,\tilde{\mathcal{C}}^k\prod_{\substack{v\notin W_0\\ \zeta_v^1=h,\zeta_v^2=H}}2^{n_{\lambda_v}}\prod_{n\ge n_V}2^{-(2nN_n^2+nN_n^1)}\Big]\cdot\\ &\qquad\cdot\Big[\prod_{n\ge n_V}\prod_{\substack{T\\ n_T=n}}\prod_{j=1}^{2}\prod_{i=1}^{m_T^j(\vartheta)}2^{R_{\lambda_{V_i}}(n-n_{V_i})}\Big]\;,\end{aligned}\tag{7.7}$$

where n_V is the scale of the resonance, $R_{\lambda_V}=1,2$ and the second square bracket is the part coming from the resummations, and follows from the above discussion about the gain factors. The constant $\tilde{\mathcal{C}}$ differs from $\mathcal{C}$ in (5.2) as it takes into account the bound on the derivatives of the propagators: we can set $\tilde{\mathcal{C}}=\mathcal{C}\,e^2\,[a_2(2)2^{-2}]^2$, as the sum over all the outer resonances V's of the factors $[2k(V)]^2$ can be bounded by e^{2k}, and $a_R(p)\le a_2(2)$, for any $R=1,2$, and $0\le p\le 2$, (and the factor 2^{-2} simplifies $a_2(0)$ in $\mathcal{C}$).

Then if we recall (5.11), we see that the $jm_T^j(\vartheta)$ is taken away by the first factor in $2^{R_{\lambda_{V_i}}n}\;2^{-R_{\lambda_{V_i}}n_{V_i}}$, being $n=n_{\lambda_{V_i}}$, while the remaining $2^{-R_{\lambda_{V_i}}n_{V_i}}$ are compensated by factors furnished by the clusters counting. In particular we can get rid of the factor $2^{-R_{\lambda_V} n_V}$ in (7.7). This completes the discussion of the resonances: then Theorem 1.2 is proven, in the case in which the perturbation is of the form (2.1), and the estimate (1.9) follows, with E defined in (2.1). The extension of the proof to the general case is given in Appendix A2, and yields (1.9) with E defined in (1.7).

8. Renormalizability of the isochronous hamiltonians

As done for Theorem 1.2, let us consider first the notationally less involved case (2.1). The general hamiltonian appearing in the statement of Theorem 1.4 will be studied in Appendix A2.

Let us write the solution of the equations of motions, (if there are any), as in (1.11), where $\vec{A}_0$ is a constant vector in D and the set $W(\vec{A}_0, \rho_0)$, defined in item (1) of Theorem 1.4, is contained in D; then let us call $\vec{H}^{(k)}$ and $\vec{h}^{(k)}$ the k-th order coefficients of the (formal) Taylor expansion of $\vec{H}$ and $\vec{h}$ in powers of ε.

We look for a solution of the equation of motion corresponding to the hamiltonian (1.10) of the form (1.11), with $\vec{h}(\vec{\omega}_0 t, \vec{A}_0; \varepsilon)$ and $\vec{H}(\vec{\omega}_0 t, \vec{A}_0; \varepsilon)$ being quasiperiodic functions with vanishing average.

The equations of motion are

$$\begin{aligned} \frac{d\vec{A}}{dt} &= -\varepsilon \partial_{\vec{\alpha}} f(\vec{\alpha}, \vec{A}) \,, \\ \frac{d\vec{\alpha}}{dt} &= \vec{\omega}_0 + \varepsilon \partial_{\vec{A}} f(\vec{\alpha}, \vec{A}) - \partial_{\vec{A}} N_f(\vec{A}; \varepsilon) \,. \end{aligned} \tag{8.1}$$

Then we get immediately recursion relations for $\vec{H}^{(k)}, \vec{h}^{(k)}$:

$$\begin{aligned} \vec{\omega}_0 \cdot \partial_{\vec{\psi}}\, \vec{H}^{(k)} &= \sum_{(k-1)}{}^{*} (-\partial_{\vec{\alpha}}) \Big[\sum_{p\ge 0} \sum_{q\ge 0} \frac{1}{p!\, q!} \prod_{s=1}^{p} \Big(\vec{h}^{(k_s)} \cdot \partial_{\vec{\alpha}} \Big) \prod_{r=1}^{q} \Big(\vec{H}^{(k'_r)} \cdot \partial_{\vec{A}} \Big) \Big] f(\vec{\omega}_0 t, \vec{A}_0) \,, \\ \vec{\omega}_0 \cdot \partial_{\vec{\psi}}\, \vec{h}^{(k)} &= \sum_{(k-1)}{}^{*} \partial_{\vec{A}} \Big[\sum_{p\ge 0} \sum_{q\ge 0} \frac{1}{p!\, q!} \prod_{s=1}^{p} \Big(\vec{h}^{(k_s)} \cdot \partial_{\vec{\alpha}} \Big) \prod_{r=1}^{q} \Big(\vec{H}^{(k'_r)} \cdot \partial_{\vec{A}} \Big) \Big] f(\vec{\omega}_0 t, \vec{A}_0) \\ &\quad - \sum_{(k)}{}^{*} \partial_{\vec{A}} \Big[\sum_{q\ge 0} \frac{1}{q!} \prod_{r=1}^{q} \Big(\vec{H}^{(k'_r)} \cdot \partial_{\vec{A}} \Big) \Big] N_f^{(k_0)}(\vec{A}_0) \,, \end{aligned} \tag{8.2}$$

where the $\sum^*_{(k-1)}$ denotes summation over the integers $k_s \ge 1$, $k'_s \ge 1$, with: $\sum_{s=1}^{p} k_s + \sum_{r=1}^{q} k'_r = k-1$, the $\sum^*_{(k)}$ denotes summation over the integers $k_s \ge 1$, $k'_s \ge 1$, $k_0 \ge 1$, with: $k_0 + \sum_{s=1}^{p} k_s + \sum_{r=1}^{q} k'_r = k$, and the derivatives are supposed to apply to the functions $f(\vec{\alpha}, \vec{A})$, $N_f^{(k_0)}(\vec{A})$ and then evaluated in $(\vec{\alpha}, \vec{A}) = (\vec{\omega}_0 t, \vec{A}_0)$. The terms in brackets corresponding to the values $p, q = 0$ have to be interpreted as 1. In the first two brackets p and q can not be simultaneously vanishing for $k>1$; in the third line of (8.2) $q=0$ yields $k_0=k$.

8.1. *Proof of the formal solubility of the recursive relations.* We proceed inductively as in §2.1, the only difference being that, when we impose that the right hand side of the second equation in (8.2) has vanishing average, then we obtain some constraint on the $N_f^{(k)}(\vec{A}_0)$ coefficients.

In fact, from the equations of motion (8.1) one obtains immediately the equations in (8.2). Then assume that, for a suitable choise of the coefficients $N_f^{(k)}(\vec{A}_0)$, $k<k_0$, the functions $\vec{H}$ and $\vec{h}$, up to order k_0-1, (1) have vanishing average on t (*i.e.* in $\vec{\psi}$), and (2) solve the equations of motion. We show that, under such an asumption, it is possible to prove that it is possible to choose $N_f^{(k_0)}(\vec{A}_0)$ in such a way that $\vec{H}^{(k_0)}$ and $\vec{h}^{(k_0)}$ solve (formally) the equations of motions (8.2) for $k=k_0$, and both functions have vanishing average, *i.e.* their $\vec{0}$-th Fourier components are identically vanishing: $\vec{H}_{\vec{0}}^{(k_0)}=\vec{h}_{\vec{0}}^{(k_0)}=\vec{0}$.

The proof of such an assertion can be carried out as in §2.1. Therefore we do not give explicitly the details: we confine ourselves to note that the right hand side of the first equation in (8.2) is automatically with vanishing average, as a consequence of the assuntions on the functions $\vec{H}^{(k)}$ and $\vec{h}^{(k)}$ for $k<k_0$. The $\vec{0}$-th Fourier component of the right hand side of the second equation can be set equal to zero, by suitably fixing the value of $N_f^{(k_0)}(\vec{A}_0)$ in such a way that it exactly cancels the $\vec{0}$-th Fourier component of the remaining expression. We stress here that, at this level, we obtain simply a formal identity between two expressions: this yields that a formal perturbative expansion is possible, but no further information is provided, and the convergence of the perturbative series has to be proven yet.

Once $N_f^{(k_0)}(\vec{A}_0)$ is fixed, the equations (8.2) can be solved, and $\vec{H}^{(k_0)}$ and $\vec{h}^{(k_0)}$ are obtained up to a constant, which, in both cases, can be chosen to be zero, so that the procedure can be iterated. ■

Then it is easy to check that, in the Fourier space, we obtain, for $\vec{\nu}\neq\vec{0}$,

$$
\begin{aligned}
(i\vec{\omega}_0\cdot\vec{\nu})\,\vec{H}_{\vec{\nu}}^{(k)} &= \sum_{(k-1)}{}^{*}(-i\vec{\nu}_0)\Big[\sum_{p\geq 0}\sum_{q\geq 0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(i\vec{\nu}_0\cdot\vec{h}_{\vec{\nu}_s}^{(k_s)}\right)\prod_{r=1}^{q}\left(\vec{H}_{\vec{\nu}'_r}^{(k'_r)}\cdot\partial_{\vec{A}}\right)\Big]f_{\vec{\nu}_0}(\vec{A}_0)\;,\\
(i\vec{\omega}_0\cdot\vec{\nu})\,\vec{h}_{\vec{\nu}}^{(k)} &= \sum_{(k-1)}{}^{*}\partial_{\vec{A}}\Big[\sum_{p\geq 0}\sum_{q\geq 0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(i\vec{\nu}_0\cdot\vec{h}_{\vec{\nu}_s}^{(k_s)}\right)\prod_{r=1}^{q}\left(\vec{H}_{\vec{\nu}'_r}^{(k'_r)}\cdot\partial_{\vec{A}}\right)\Big]f_{\vec{\nu}_0}(\vec{A}_0)\\
&\quad -\sum_{(k)}{}^{*}\partial_{\vec{A}}\Big[\sum_{q\geq 0}\frac{1}{q!}\prod_{r=1}^{q}\left(\vec{H}^{(k'_r)}\cdot\partial_{\vec{A}}\right)\Big]N_f^{(k_0)}(\vec{A}_0)\;,
\end{aligned}
\tag{8.3}
$$

and, for $\vec{\nu}=\vec{0}$,

$$
\begin{aligned}
N_f^{(k)}(\vec{A}_0) = & \sum_{(k-1)}{}^{*}\Big[\sum_{p\geq 0}\sum_{q\geq 0}\frac{1}{p!\,q!}\prod_{s=1}^{p}\left(i\vec{\nu}_0\cdot\vec{h}^{(k_s)}_{\vec{\nu}_s}\right)\prod_{r=1}^{q}\left(\vec{H}^{(k'_r)}_{\vec{\nu}'_r}\cdot\partial_{\vec{A}}\right)\Big]f_{\vec{\nu}_0}(\vec{A}_0) \\
& -\sum_{(k)}{}^{*}\Big[\sum_{q>0}\frac{1}{q!}\prod_{r=1}^{q}\left(\vec{H}^{(k'_r)}\cdot\partial_{\vec{A}}\right)\Big]N_f^{(k_0)}(\vec{A}_0)\;,
\end{aligned}
\tag{8.4}
$$

where the $\sum^*_{(k-1)}$ denotes summation over the integers $k_s\geq 1$, $k'_r\geq 1$, with: $\sum_{s=1}^p k_s + \sum_{r=1}^q k'_r) = k$, and over the integers $\vec{\nu}_0$, $\vec{\nu}_s$, $\vec{\nu}'_r$, with: $\vec{\nu}_0+\sum_{s=1}^p\vec{\nu}_s+\sum_{r=1}^q\vec{\nu}'_r=\vec{\nu}$, while $\sum^*_{(k)}$ denotes summation over the integers $k_s\geq 1$, $k'_r\geq 1$, $k>k_0\geq 1$ with: $k_0+\sum_{s=1}^p k_s + \sum_{r=1}^q k'_r = k$, and over the integers $\vec{\nu}_0$, $\vec{\nu}_s$, $\vec{\nu}'_r$, with: $\vec{\nu}_0+\sum_{s=1}^p\vec{\nu}_s+\sum_{r=1}^q\vec{\nu}'_r=\vec{\nu}$, and the constraint $q>0$ implies that the case $k_0=k$ has to be excluded. As in (8.2) the case $p=q=0$ in the first two lines of (8.3) and in the first line of (8.4) is possible only if $k=1$.

8.2. *Remark.* Obviously the equation for $N_f^{(k)}$ involves its derivative with respect to the action variable, *i.e.* $\partial_{\vec{A}}N_f^{(k)}(\vec{A}_0)$, but we can trasfer the equation directly on the function $N_f^{(k)}$, because we can define it up to a constant (in $\vec{\alpha}$ and $\vec{A}$), without loss of generality.

For example, if $k=1$, we have from (8.3) and (8.4)

$$
\begin{aligned}
\vec{H}^{(1)}_{\vec{\nu}} &= \frac{-i\vec{\nu}_v\, f_{\vec{\nu}}(\vec{A}_0)}{i\vec{\omega}_0\cdot\vec{\nu}}\;, \qquad \vec{\nu}\neq\vec{0} \\
\vec{h}^{(1)}_{\vec{\nu}} &= \frac{\partial_{\vec{A}}f_{\vec{\nu}}(\vec{A}_0)}{i\vec{\omega}_0\cdot\vec{\nu}}\;, \qquad \vec{\nu}\neq\vec{0} \\
N_f^{(1)} &= f_{\vec{0}}(\vec{A}_0)\;,
\end{aligned}
\tag{8.5}
$$

where the last equation fixes the value of the first coefficient of the series expansion of $N_f(\vec{A}_0;\varepsilon)$.

If we compare the equations (8.4) and (8.5) with the equations (2.9)÷(2.11) in §2, we see that it is still possible to carry out a diagrammatic expansion of the coefficients $\vec{h}^{(k)}_{\vec{\nu}}$, $\vec{H}^{(k)}_{\vec{\nu}}$ and $N_f^{(k)}$.

We give first the rules how to construct the diagrams in general. First of all we introduce the labels to associate to the tree vertices. Obviously we have to take into account that some vertices $v\in\vartheta$ can have associated a factor $N_f^{(k_v)}$, for some $k_v\geq 1$, instead of a factor $f_{\vec{\nu}_v}$: the meaning of the label k_v is now very different from that of the label introduced in item (3) of §3.3 (there it denoted the order of the subtree having v as first vertex, here it is independent on such a subtree). Let us introduce a label δ_v, such that $\delta_v=1$ if a factor $f_{\vec{\nu}_v}$ is associated to v, and $\delta_v=0$ otherwise: in the latter case, also a label $k_v\geq 1$ is assigned to v. The terminology is somewhat reminiscent of that used in §3, where $\delta_v=0$

denoted a factor $\mathcal{H}_0$, which, like $N_f^{(k_v)}$, depends only on the action variables, but note that, whereas in §3.3 $\delta_v = 0$ implied that no ε was associated to that vertex, now $\delta_v = 0$ implies that a factor ε^{k_v} is associated to v.

Then we have the following collection of labels.

(1) $\delta_v = 0, 1$;
(2) $k_v \equiv 1$ if $\delta_v = 1$, and $k_v \geq 1$ if $\delta_v = 0$;
(3) $\vec{\nu}_v \in \mathbb{Z}^\ell$ is the mode label;
(4) ζ_v^1 and ζ_v^2 can assume the symbolic values $\zeta_v^1, \zeta_v^2 = h, H, N$;
(5) $m_v = p_v + q_v$ is defined as in item (6) in §3.3.

With respect to (3.4) and (3.5) of §3, now we have the following functionals associated to the tree branches.

$$\begin{array}{ccc} \text{operator} & \text{propagator} & \text{branch} \\ 0 & 0 & h \leftarrow h \\ C_0 \left[i\vec{\nu}_{v'} \cdot (\partial_{\vec{A}_v}) \right] & [i\vec{\omega} \cdot \nu_{\lambda_v}]^{-1} & h \leftarrow H \\ C_0 \left[\partial_{\vec{A}_{v'}} \cdot (-i\vec{\nu}_v) \right] & [i\vec{\omega} \cdot \vec{\nu}_{\lambda_v}]^{-1} & H \leftarrow h \\ 0 & 0 & N \leftarrow H \end{array} \tag{8.6}$$

for all the branches distinct from the root branch, and

$$\begin{array}{ccc} \text{operator} & \text{propagator} & \text{branch} \\ 0 & 0 & h \leftarrow h \\ C_0 \left[\partial_{\vec{A}_v} \right] & [i\vec{\omega} \cdot \nu_{\lambda_v}]^{-1} & h \leftarrow H \\ C_0^2 \left[-i\vec{\nu}_v \right] & [i\vec{\omega} \cdot \vec{\nu}_{\lambda_v}]^{-1} & H \leftarrow h \\ 1 & 1 & N \leftarrow H \end{array} \tag{8.7}$$

for the root branch.

Then all the above operators are applied to the function

$$\prod_{\substack{v \in \vartheta \\ \delta_v = 1}} f_{\vec{\nu}_v}(\vec{A}_v) \prod_{\substack{v \in \vartheta \\ \delta_v = 0}} [-N_f^{(k_v)}(\vec{A}_v)] \,, \tag{8.8}$$

which replaces (3.6), so obtaining a factor O_v for each vertex v, analogously to what happened in §3. The explicit expression of O_v can be obtained from the above discussion, by reasoning as in §3, (see in particular (3.8)).

The following result (analogous to Proposition 3.7 in §3) holds.

8.3. Proposition. *Let us consider a tree of order k; then*

$$\sum_{\substack{v\in\vartheta\\ \delta_v=1}} 1 + \sum_{\substack{v\in\vartheta\\ \delta_v=0}} k_v = k\ , \tag{8.9}$$

and, if $\delta_v = 0$, $\zeta_v^1 = h$ and $\zeta_v^2 = H$, and $\zeta_w^1 = H$ and $\zeta_w^2 = h$ for any $w \in \vartheta$ such that $v = w'$.

8.4. *Proof of proposition 8.3.* The properties stated in Proposition 8.3 are immediate consequences of the definitions and of (8.3)÷(8.5). ∎

We can not introduce the dimensionless variable $\vec{X}^{(k)}_{\vec{\nu}}$ of §3, because now it would be $J_m = 0$. This does not means that it is impossible to introduce dimensionless variables (of course), but simply the definition given in §3 was suitable only in that case (system verifying the anisochrony condition).

8.5. *Remark.* Note that, if we define $T(\vec{A}_0)$ as the matrix of the second derivatives of the free hamiltonian, then it is trivially $T(\vec{A}_0) \equiv 0$, (see also the previous remark). Therefore, if we compare (8.6) and (8.7) above with the corresponding ones in §3, we see that no matrix $T(\vec{A}_0)$, $T(\vec{A}_0)^{-1}$ appears now, as it has to be if we want that the formulae we write are meaningful.

We note that, as it is possible to check, some contributions corresponding to trees carrying labels $\delta_v = 0$ on some vertices can cancel exactly, when summed together to some other tree values.

To see this, let us consider a tree ϑ_1 of order k, with a vertex v carrying a label $\delta_v = 0$ and having associated a factor $N_f^{(k_v)}$. Now let us consider also the tree ϑ_2 which differ from ϑ_1 because: (1) the vertex v carries a label $\delta_v = 1$, and (2) other subtrees ϑ_1', …, ϑ_s' (which were absent in ϑ_1) emerge from v, such that the sum of their orders add to k_v; all the other labels are the same in ϑ_1 and $\vartheta_2 \setminus \cup_{j=1}^s \vartheta_j'$. If we expand $N_f^{(k_v)}$ in ϑ_1 according to the graph rules, then, among the several possible trees, *we can obtain from ϑ_1 also a tree which has exactly the same shape and the same labels as ϑ_2, hence the same value, but with opposite sign,*[4] so that the two values cancel each other.

In this way one can cancel also the values of some trees with resonances, if (1) the line λ_v carries the same momentum of one of the lines entering v, (2) all the lines of the other subtrees emerging from v have a higher scale label, and (3) the tree representation of $N_f^{(k_v)}$ does not destroy the inclusion relations between the clusters. But it is important the fact that the corresponding resonance factors are of the form $\mathcal{V}^n_{H,H}$, and we recall from §6 (see in particular the fourth term in (6.4)) that *no cancellation was required* in such a case.

[4] Because of the sign minus before $N_f^{(k_v)}$ in (8.8). Note also that the cancellation we are considering occurs only if the combinatorial weights of the two trees are equal: this happens only if no subtree among ϑ_1', …, ϑ_s' is equivalent to some other subtree emerging from v and apearing both in ϑ_1 and ϑ_2.

Therefore the cancellations occurring in according to the just described mechanism are not quite dangerous, and can be easily forgotten: in other words there is no cancellations overlapping and we can reason as if no cancellation involving the factors $N_f^{(k_v)}$ can happen, and only the cancellations described in §6 are taken into account.

We can now proceed as in the proof of Theorem 1.2, with the propagators and operators defined as in (8.6) and (8.7) above.

Then the results of the previous sections apply. Note that it is not really necessary to repeat the discussion (if we are not looking for optimal bounds): it is enough to realize that now the diagrammatic rules are simplified with respect to the case studied before. In fact we can simply (1) write down the perturbative expansion by using the rules of §3 and §4; (2) get rid of the $h \leftarrow h$ lines by imposing that each time such a line appears we obtain a vanishing contribution; (3) analogously get rid of the lines $N \leftarrow H$ which are not root branches; (4) replace the operator in the fourth row in (3.5) with the identity $\mathbb{1}$; (5) the factors $N_f^{(k_v)}$ associated to the vertices $v \in \vartheta$ with $\delta_v = 0$ can be expanded and graphically represented as trees in according to (8.4), so giving a factorizing term which can be treated in the same way. In particular it is easy to check that a bound like (1.12) follows, if we set $F\varepsilon_1^{-1} = F_0$.

This concludes the discussion in the case in which the interaction is as in (2.1). The general case will be discussed in Appendix A2.

9. A comparison with other proofs of the convergence of the Lindstedt series

After Eliasson works, [E1], [E2] and [E3], somewhat hard to read (the fundamental paper [E1] has not been yet published), simplified and clearer proofs of the convergence of the Lindstedt series have been proposed in several papers, all quoted in the introduction. This section is devoted to a comparison between such proofs and the one discussed in the present paper.

9.1. The first proof is the Eliasson's one, [E1]. In [E1], §II, Eliasson writes a tree expansion for the invariant tori, (although the trees are called *simple index sets* which is called *tree structure* in [E2]), and realizes that some terms to order k in this expansion are of order $O(C^k(k!)^\alpha)$, for some positive constants α and C depending on the particular contribution studied: then he concludes that the Lindstedt series would converge to an analytic function only if there are "very sharp compensations of signs between different terms of these series expansion", [E1], page 3. However Eliasson *does not look for an identification of terms between which there are compensations*: on the contrary he replaces the Lindstedt series with a different series whose convergence follows *without the necessity of exploiting any cancellation.* Since the sum of the series does not verify the equations of motions, he introduces in it a free parameter to be fixed so that the series verifies such equations:

if this can be accomplished, then the resulting series has to coincide with the Lindstedt series.

We first propose here a translation of Eliasson's work in our formalism, up to a minor technical modification which will be explained below. We consider, following Eliasson's spirit, the quantities $\vec{h}^*(\vec{\psi}) = \sum_{\vec{\nu}\in\mathbb{Z}^l} e^{i\vec{\nu}\cdot\vec{\psi}}\,\vec{h}^*_{\vec{\nu}}$ and $\vec{H}^*(\vec{\psi}) = \sum_{\vec{\nu}\in\mathbb{Z}^l} e^{i\vec{\nu}\cdot\vec{\psi}}\,\vec{H}^*_{\vec{\nu}}$, defined as

$$\vec{X}^*_{\vec{\nu}}(\zeta) = \sum_k \sum_{\theta\in\mathcal{T}^*_k} \mathrm{Val}\,(\theta)\;, \qquad \zeta\in\{h,H,\mu\}\;, \tag{9.1}$$

(compare with (3.1)), where $\mathrm{Val}(\theta)$ is defined as in (4.4) and $\mathcal{T}^*_k$ is the set of the labeled semitopological trees such that *on the resonances only the $\mathcal{R}$ operators apply*, *i.e.* tree values containing factors $\mathcal{LV}^n_{\zeta^2_{w_0},\zeta^1_{w_1}}(\vec{\omega}\cdot\vec{\nu})$ are discarded.

We know from the analysis in §7 that the convergence of the above series follows simply from the Siegel-Bryuno's bound, while, in order to prove the convergence of the original series one has to exploit the cancellation mechanisms discussed in §6.

The series $\vec{h}^*(\vec{\psi}), \vec{H}^*(\vec{\psi})$ are not the original series for the tori, hence they will not verify the equations of motion (2.2); however they obey to a very simple modification of them. If $\vec{A}^*$ and $\vec{\alpha}^*$ are given by (1.11) with $\vec{h}$ and $\vec{H}$ replaced with $\vec{h}^*$ and $\vec{H}^*$, one has

$$\begin{aligned}\frac{d\vec{H}^*}{dt} =&[-\varepsilon\partial_{\vec{\alpha}^*}f(\vec{\alpha}^*,\vec{A}^*)] - \mathcal{D}_{h,h}\,\vec{h}^* - \mathcal{D}_{h,H}\,\vec{H}^*\;,\\ \frac{d\vec{h}^*}{dt} =&[\partial_{\vec{A}^*}\mathcal{H}_0 + \varepsilon\partial_{\vec{A}^*}f(\vec{\alpha}^*,\vec{A}^*)] - \mathcal{D}_{H,h}\,\vec{h}^*\;,\end{aligned} \tag{9.2}$$

where, using our notations, one can easily check that the $\ell\times\ell$ matrices $\mathcal{D}_{i,j}$ are given by the sums is over all the possible localized resonance factors $\mathcal{LV}_{i,j}(-i\vec{\omega}\cdot\partial_{\vec{\psi}})$, running the sum on the perturbative order and over trees.

The proof of the above equation is quite simple. Let us express $\vec{h}^*$ and $\vec{H}^*$ as sums over trees, and let v_0 be the the first vertex, as in Fig.3.1. A tree in which v_0 does not belong to a resonant cluster contributes to the series expansion of the square brackets in the rigth hand side of (9.2); if v_0 is in a resonant cluster, if we recall that $\mathcal{R} = \mathbb{1} - \mathcal{L}$, we can split the value of the corresponding tree into two parts: the term in which $-\mathcal{L}$ is applied on the resonance factor contribute to the terms outside the square brakets, while the term in which $\mathbb{1}$ is applied on the resonance factor is included in the square brackets contribution.

It is a remarkable fact that $\vec{A}^*, \vec{\alpha}^*$ obey to equations "so similar" to the original equations of motion (2.2). With the notations of §2.1, we can rewrite (9.2) as

$$\frac{dY^*}{dt} = (E\partial\mathcal{H})(Y^*) + ED(Y^* - Y^{(0)})\;, \tag{9.3}$$

where $Y^* = (\vec{h}^*, \vec{H}^*)$, $Y^{(0)} = (\vec{\psi}, \vec{A}_0)$ and D is the $2\ell\times 2\ell$ matrix with entries $D_{ij} = (-1)^{\delta_{i,H}}\,\mathcal{D}_{i,j}$, for $(i,j)\neq(H,H)$, and $D_{H,H}(0) = 0$.

The series $\vec{h}^*, \vec{H}^*$ are not, in principle, the solutions of the equations of motion $dY/dt = (E\partial\mathcal{H})(Y)$, unless one proves that the sum of the localized resonance factors vanishes (*i.e.* $D \equiv 0$ identically). This is exactly what was done in §6; on the other hand we could proceed in a different more abstract way, much closer to Eliasson's approach.

We can suppose that up to order $k_0 - 1$, $D \equiv 0$ and $Y^* = Y$, if Y is the solution of the equation $dY/dt = (E\partial\mathcal{H})(Y)$. Then (9.3) is satisfied also to order k_0, so that $Y^* = Y$ up to order k_0. From the fact that (2.6) holds for any periodic function and from the inductive assumption, (2.8) still holds and, as in §2.1, implies that the average of the function $[\partial_{\vec{\alpha}}\mathcal{H}]^{(k_0+1)}$ vanishes. Then $D_{11}^{(k_0)}\vec{h}_{\vec{0}}^{(1)} + D_{12}^{(k_0)}\vec{H}_{\vec{0}}^{(1)} = \vec{0}$, and the arbitrarity of $\vec{h}_{\vec{0}}^{(1)}$, (see §2.1, item (4)), yields (1) $D_{11}^{(k_0)} = 0$, and (2) $D_{12}^{(k_0)}\vec{H}_{\vec{0}}^{(1)} = \vec{0}$. It turns out that the latter identity yields $D_{12}^{(k_0)} = 0$. Accepting the last statement we have also $D_{21}^{(k_0)} = 0$, as a consequence of Lemma 6.4. Since $D_{22} \equiv 0$, see (6.4), this proves that $D \equiv 0$.

It remains to check that $D_{12}^{(k_0)}\vec{H}_{\vec{0}}^{(1)} = \vec{0}$ implies $D_{12}^{(k_0)} = 0$. To see this, let us write the perturbation $f \equiv f(\vec{\alpha}, \vec{A})$ as $f = f_0 + \tilde{f}$, where

$$f_0 = f_{\vec{0}}(\vec{A}) \ , \quad \tilde{f} = \sum_{\vec{\nu}\neq\vec{0}} e^{i\vec{\nu}\cdot\vec{\alpha}} f_{\vec{\nu}}(\vec{A}) \ ,$$

and let us consider f_0 and $\tilde{f}$ as two independent functions. Then, if we study the dependence on f, we have, essentially by definition, $D_{12}^{(k_0)}(f) \equiv D_{12}^{(k_0)}(f_0, \tilde{f}) = A(f_0, \tilde{f}) + B(\tilde{f})$, where $A(f_0, \tilde{f})$ depends explicitly on f_0 and vanishes for $f_0 = 0$, while $\vec{H}_{\vec{0}}^{(1)}(f_0) \equiv \vec{\mu}^{(1)}(f_0)$ depends only on f_0, (see the third equation in (1.9)), and $\vec{\mu}^{(1)}(f_0) \neq \vec{0}$ for $f_0 \neq 0$, $\vec{\mu}^{(1)}(0) = \vec{0}$. Then $D_{12}^{(k_0)}\vec{H}_{\vec{0}}^{(1)} = \vec{0}$ can be written $\left[A(f_0, \tilde{f}) + B(\tilde{f})\right]\vec{\mu}^{(1)}(f_0) = 0$, which holds for any f_0. If $f_0 \neq 0$, then $A(f_0, \tilde{f}) + B(\tilde{f}) = 0$, and since $B(\tilde{f})$ does not depend on f_0, one has $A(f_0, \tilde{f}) = B(\tilde{f}) = 0$ identically in $\tilde{f}$, *i.e.* $B(\tilde{f}) \equiv 0$. This means that $D_{12}^{(k_0)}(f) = 0$ for any f.

Eliasson's discussion is very similar to the just described one, with the following differences (which seem to us to be unessential). First he performs a change of coordinates (simply a rescaling) on the unperturbed hamiltonian, so that the graph rules are slightly different from the ones described in §3; furthermore he does not introduce any multiscale decompositions, and this makes the discussion of the convergence of the series more involved: he defines "resonance" a generic pair of vertices such that the lines emerging from them carry the same momentum, ([E1], §IV), and then he has to solve the problem of the "overlapping divergences", so distinguishing among several types of resonances, such that only the "critical ones" cannot be controlled through the Siegel-Bryuno's lemma. The series he studies are *not* exactly the above series $\vec{h}^*$ and $\vec{H}^*$, rather they are series for two functions $\vec{h}^{**}$ and $\vec{H}^{**}$, defined as in (9.1), but with the difference that the localization operator is defined by (6.4) as far as the first three lines are concerned, whereas the last line is replaced with $\mathcal{L}\mathcal{V}^n_{H,H}(\vec\omega\cdot\vec\nu)=\mathcal{V}^n_{H,H}(0)$.[5]

Therefore, instead of the series which solve the equations (9.2), Eliasson obtains somewhat different quantities, $\vec h^{**}$ and $\vec H^{**}$, which can be thought as obtained by the above $\vec h^*$ and $\vec H^*$, by replacing $\partial_{\vec A}\partial_{\vec A}f_0$ with $\partial_{\vec A}\partial_{\vec A}f_0+M$, being M a $\ell\times\ell$ matrix whose expression is left free and has to be determined. It is then easy to check that the following equation holds for $\vec A^{**},\vec\alpha^{**}$:

$$\begin{aligned}\frac{d\vec H^{**}}{dt}&=[-\varepsilon\partial_{\vec\alpha^{**}}f(\vec\alpha^{**},\vec A^{**})]-\mathcal{D}_{h,h}\,\vec h^{**}-\mathcal{D}_{h,H}\,\vec H^{**}\;,\\ \frac{d\vec h^{**}}{dt}&=[\partial_{\vec A^{**}}\mathcal{H}_0+\varepsilon\partial_{\vec A^{**}}f(\vec\alpha^{**},\vec A^{**})]-\mathcal{D}_{H,h}\,\vec h^{**}-(\mathcal{D}_{H,H}+M)\,\vec H^{**}\;.\end{aligned}\tag{9.4}$$

Then, by using an argument based on the symplectic structure of the problem, similar to the one used by Poincaré to prove the formal existence of the Lindstedt series, Eliasson proves that it is possible to choose M as an analytic function of ε so that all terms after the first square brakets in (9.4) are vanishing, and, because of the uniqueness of the solution, $\vec h^{**}=\vec h,\vec H^{**}=\vec H$, ([E1], §VI): then the convergence of the Lindstedt series is proven. The choice of M has the effect that $\mathcal{V}_{H,H}$ is not really renormalized, (*i.e.* we could define $\mathcal{L}\mathcal{V}_{H,H}(i\vec\omega\cdot\vec\nu)=0$ as in (6.4)). In our formalism, this corresponds to the fact that, if Eliasson's localization is used, all the sums of localized resonance factors in (9.4) give a vanishing contribution, except the sum corresponding to the resonances with $\zeta^2_{w_0}=\zeta^1_{w_1}=H$: this means that *he subtracts also a contribution, which in fact is not vanishing, but which is known, from the analysis of* § *7, not to give problems.* Then he has to introduce a parameter (the matrix M), in order to recover such a contribution (which, of course, from our point of view, should *not* have been subtracted to begin with): all the other localized resonance factors are automatically vanishing. The symplectic argument is not just the same we reproduced above: Eliasson's original argument is very quickly sketched in [E1] and it is similar the Poincaré's argument for proving the formal existence of the Lindstedt series, [P], Vol. II, §126, while in the above discussion we followed an argument similar to the proof of the formal solubilty in §2.1, based on [CZ], [CG].

In [E2], Eliasson applies the same methods in order to prove the conjecture by Gallavotti stated in the introduction of the present paper, Theorem 1.4.

9.2. We stress that the possibility of fixing the the initial data, in order to obatin a formal power expansion by solving the equations of motion, is standard, and was well

[5] It can also be worth to remark that, in the definition of the series $\vec h^*(\vec\psi)$ and $\vec H^*(\vec\psi)$, (see (9.1)), the quantities which are subtracted from the resonance factors (and which are precisely the localized resonance factors) cannot at all be identified with the resonance factors of some other trees, (as sometimes Eliasson's argument has been erroneously interpreted): simply the contributions which are responsible for the apparent divergence of the perturbative series are deleted in order to obtain a series whose convergence can be proved without exploiting cancellations.

known already to Lindstedt, [L], Newcomb, [Ne], and Poincaré, [P], Vol. II. The first work in which Eliasson ideas have been resumed is a paper by Feldman and Trubowitz, [FT], where the graph representation of the solution of the equations of motion in terms of trees is illustrated (following [E1] and [E3]), and the analogy with quantum field theory is pointed out as to the introduction of the counterterms in order to obtain the formal solubility of the equations of motion, (this corresponds to §2.1 in the present paper). The problem of the convergence of the perturbative series is not touched, (on this point the authors refer to the classical proofs and to Eliasson's work). The possibility of interpreting the Lindstedt series as formal perturbative series of a euclidean field theory on the torus $\mathbb{T}^\ell$ is pointed out in [G9], where an action giving rise to the Lindstedt series is proposed (together with a convergence proof). The idea is further developed in [GGM].

9.3. In an effort to understand Eliasson's work his ideas have been applied to simplified models in which the free hamiltonian is quadratic in the action variables, and the interaction potential is taken to be action-independent and either (1) a trigonometric polynomial in the angles, or (2) or an analytic function in the actions.

The analysis of the model (1) is carried out in [G7], with the further simplification that the potential is an *even* trigonometric polynomial and the rotation vectors satisfy a property stronger than the usual diophantine one. The second hypothesis is completely relaxed in [GG]. A discussion of the model (2) can be found in [CF1].

The cancellation mechanism between the resonances described in the present paper is exactly the same as in [G7] and [CF1], as far as the first order zero is concerned; the second order zero is ensured in [G7] from the parity properties of the interaction potential (a situation very often realized in physics), while in [CF1] a more subtle anlysis of the cancellations is required in order to discuss the not even case. In this regard, we mention, as a relevant fact, that we have taken from [CF1] the idea to shift also the exiting line of the resonances in order to see the cancellations in the not even case. The main difference between [G7] and [CF1] is in the definition of resonance (and therefore in the way the related problems are solved). However both definitions in the end turn out to be somewhat equivalent, because the ultimate aim is to assure that some quantities (essentially the quantities which correspond to our resonance factors), considered as functions of the scalar product between $\vec{\omega}$ and the momentum flowing through the incoming line of the resonance, say $\vec{\omega}\cdot\vec{\nu}_{\lambda_V}$, do not go out from their analyticity domain, when shifting the external lines of the resonances. This is implemented in [CF1] by requiring that the quantity $|\vec{\omega}\cdot\vec{\nu}_{\lambda_V}|$ corresponding to the momentum $\vec{\nu}_{\lambda_V}$ flowing through the resonant line cannot be smaller than a prefixed fraction of the quantity $|\vec{\omega}\cdot\vec{\nu}_\lambda|$ corresponding to the momentum of any line λ inside a resonance, so that the shift of the external lines of the resonances "does not modify too much" the small divisors. In [G7] the same is accomplished by requiring that the number of lines inside the resonance are bounded by a suitable constant. In [G7] use is made also of the *strong diophantine property*, *i.e.* $\vec{\omega}$ is such that $\min_{0\ge p\ge n}|C_0|\vec{\omega}\cdot\vec{\nu}|-2^{-p}|>2^{-(n+1)}$ for $0<|\vec{\nu}|\le(2^{n+3})^{-\tau^{-1}}$: in this way one can forbid that two quantities $|\vec{\omega}\cdot\vec{\nu}_\lambda|$

and $|\vec{\omega} \cdot \vec{\nu}_{\lambda'}|$ with different scales are too much near to each other; the last condition was eliminated in [GG] through a suitable choice of the unity partition used in the multiscale decomposition of the propagator. Both constructions reflect the fact that only resonances V such that the scales n_{λ_V} and n_V are very different can give problems (because it is only when this happens that dangerously small divisors can occur) and cannot be treated by the Siegel-Bryuno's lemma, or any variant of it. We can note that in [G7] and [GG] the analysis is somewhat simpler with respect to [CF1], as far as what concerns the problem of singling out the contributions needing a more careful discussion (in order to show the cancellations). For instance, as far as the non overlapping of resonances is concerned, this is implied by the definition of *critical resonance* in [CF1] (see Proposition 5.2 in [CF1]), while it is automatically satisfied by the construction of clusters in [G7] and [GG], (and, obviously, in the successive related papers, including the present one).

With respect to the just described works, the technique used in the present paper (which is taken from [GM1] and [GM2], where simplified models are studied) is closer to quantum field methods, and in fact reduces the proof of the KAM theorem to the study of a renormalizable field theory, thus allowing us to use all the powerful ideas which have been developed so far in order to treat such kind of problems. We think that the reason why the proof can be given in a form as simple as in [G7], [GM1] (for semplified models) is due precisely to the use of such ideas.

Furthermore some simplifications can be obtained in the proof: first the resonances are defined *tout court* as the clusters such that there is only one entering line, and it carries the same momentum of the exiting one (there is no need to distinguish between resonances, λ-resonances and critical resonances, as in [CF1]: obviously this corresponds to the fact that a gain is truly necessary only for the last ones) and the problem of exiting from certain analyticity domains never arise. This property is easily obtained through the introduction of the compact support functions realizing the partition of the unity (see §4 and App. A3), and it can be understood in the following way.

If the notations in §6 are adopted, for each resonance V, one has that the sum over all the trees contained in the resonance family $\mathcal{F}_V(\vartheta)$ produces a quantity which, considered as a function of $\vec{\omega} \cdot \vec{\nu}_{\lambda_V}$, vanishes to first and second order; then, if such a function is analytic in $\vec{\omega} \cdot \vec{\nu}_{\lambda_V}$ in a ball centered on the origin, the Schwarz's lemma for analytic functions can be used in order to obtain bounds on the small divisors. This is what was done in [G7], [GG] and [CF1]. ¿From a technical point of view, this yields that the quantities $\vec{\omega} \cdot \vec{\nu}_{\lambda_V}$ appearing in the small divisors have to be dealt with as parameters which can assume values larger than their true values, and, if we have several resonances contained in each other, a very careful analysis of the holomorphy domains is needed. For details we refer to the quoted papers and to [GG] in particular (where the problem is discussed to mucxh extent): the main point here is that the technical intricacies are a consequence of the analyticity request. In this paper we do not require analyticity, but we want only a bound on the second derivatives of the small divisors (see §7), once we have proven that no higher order derivative appears (*ibidem*). This means that we do not need a bound on the values of the small divisors in a neighbourhood of the origin, but *only* on the their first and second derivatives appearing in the interpolation formulae (7.5): in other words we only require that the resonance factors are twice differentiable. See also the comments between Definition 6.3 and Lemma 6.4. The deep reason and the drawback of this semplification is that it follows from an overcompensation, *i.e.* we collect together terms producing more cancellations than it would be necessary in order to make the series convergent; the overcompensation is paid by worse final estimates for the convergence radius.

Acknowledgements. This work springs from the purpose of extending to the most general case the results found in simplified models in previous papers, and can be considered as a development of the original ideas contained in the paper [G6] of Giovanni Gallavotti, which we thank for having introduced us to a quantum field theory approach to KAM theorem and for many enlightening discussions. We thank Antonio Giorgilli for a useful discussion about the work [GL], which we do not discuss here, as it is a new proof of KAM theorem in Kolmogorov's spirit, so no directly related to the works we have analyzed in §9. One of us (G.G.) thanks IHES for partial support and hospitality, while part of this work was done. This work is part of the research program of the European Network on "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460.

Appendix A1. Resonant Siegel-Bryuno's bound.

Given a tree ϑ, if we are interested only in the momentum and mode labels, the case in which either there is a line entering a vertex v and carrying a zero momentum or there is no line at all behave exactly in the same way, as far the momenta of the vertices $w < v$ are concerned. This means that (5.6) can be written as a product of $1+\mathcal{N}_{\vec{0}}(\vartheta)$ factorising terms, if $\mathcal{N}_{\vec{0}}(\vartheta)$ is the number of lines λ in ϑ (other than the branch root) such that $\vec{\nu}_\lambda = \vec{0}$.

Therefore we can confine ourselves to the case in which there is no line carrying a vanishing momentum.

We prove by induction on the tree order that, if $N_n^*(\vartheta)$ is defined as the number of non resonant lines in ϑ carring a scale label $\le n$, then $N_n^*(\vartheta) \le 2M(\vartheta)2^{(n+2)/\tau} - 1$, if $N_n(\vartheta) \ne 0$.

Let ϑ be a tree of order k. If ϑ has the root line with scale $> n$ then calling $\vartheta_1, \vartheta_2, \ldots, \vartheta_m$ the subtrees of ϑ emerging from the first vertex of ϑ and with $M(\vartheta_j) > 2^{-(n+2)/\tau}$ lines, it is $N_n^*(\vartheta) = N_n^*(\vartheta_1) + \ldots + N_n^*(\vartheta_m)$ and the statement is inductively implied from its validity for $k' < k$ provided it is true that $N_n^*(\vartheta) = 0$ if $M(\vartheta) < 2^{-(n+2)/\tau}$, which is is certainly the case.

In the other case, (*i.e.* if the root branch has scale label $\le n$), it is $N^*_n(\vartheta) \le 1 + \sum_{i=1}^m N^*_n(\vartheta_i)$, and if $m = 0$ the statement is trivial, or if $m \ge 2$ the statement is again inductively implied by its validity for $k' < k$.

If $m = 1$ we once more have a trivial case unless it is $M(\vartheta_1) > M(\vartheta) - 2^{-1}2^{-(n+2)/\tau}$. But in the latter case, it turns out that the root line of ϑ_1 is either a resonant line or it has scale $> n$.

Accepting the last statement we have: $N^*_n(\vartheta) = 1 + N^*_n(\vartheta_1) = 1 + N^*_n(\vartheta'_1) + \ldots + N^*_n(\vartheta'_{m'})$, with ϑ'_j's being the m' subtrees emerging from the first vertex of ϑ'_1 with $M(\vartheta'_j) > 2^{-(n+2)/\tau}$: this is so because the root line of ϑ_1 will not contribute its unit to $N^*_n(\vartheta_1)$. Going once more through the analysis the only non trivial case is if $m' = 1$ and in that case $N^*_n(\vartheta'_1) = N^*_n(\vartheta''{}_1) + \ldots + N^*_n(\vartheta''{}_{m''})$, *etc.*, until we reach either a trivial case or a tree $\tilde\vartheta$ such that $M(\tilde\vartheta) < M(\vartheta) - 2^{-1}2^{-n/\tau}$.

It remains to check that, if $M(\vartheta_1) > M(\vartheta) - 2^{-1}2^{-(n+2)/\tau}$, then the root line of ϑ_1 has scale $> n$, unless it is entering a resonance.

Suppose that the root line of ϑ_1 has scale $\le n$ and is not entering a resonance. Note that $|\vec\omega \cdot \vec\nu_{\lambda_{v_0}}| \le 2^{n+1}$, $|\vec\omega \cdot \vec\nu_{\lambda_{v_1}}| \le 2^{n+1}$, if v_0 and v_1 are the first vertices of ϑ and ϑ_1 respectively. Hence $\delta \equiv |(\vec\omega \cdot (\vec\nu_{\lambda_{v_0}} - \vec\nu_{\lambda_{v_1}})| \le 2\,2^{n+1}$ and the diophantine assumption implies that $|\vec\nu_{\lambda_{v_0}} - \vec\nu_{\lambda_{v_1}}| > (2^{n+2})^{-1/\tau}$, or $\vec\nu_{\lambda_{v_0}} = \vec\nu_{\lambda_{v_1}}$. The latter case being discarded as we are not considering the resonances,[6] it follows that $M(\vartheta) - M(\vartheta_1) < 2^{-1}\,2^{-(n+2)/\tau}$ is inconsistent: it would in fact imply that $\vec\nu_{\lambda_{v_0}} - \vec\nu_{\lambda_{v_1}}$ is a sum of $k - k_1$ vertex modes such that $|\vec\nu_{\lambda_{v_0}} - \vec\nu_{\lambda_{v_1}}| < 2^{-1}2^{-(n+2)/\tau}$, hence $\delta > 2^{n+3}$ which is contradictory with the above opposite inequality.

Analogously, we can prove that, if $N_n(\vartheta) > 0$, then the number $p_n(\vartheta)$ of clusters of scale n verifies the bound $p_n(\vartheta) \le 2M(\vartheta)2^{(n+2)/\tau}) - 1$. In fact this is true for a tree ϑ such that $M(\vartheta) \le 2^{(n+2)/\tau}$. Otherwise, if the first tree vertex v_0 is not in a cluster of scale n, it is $p_n(\vartheta) = p(\vartheta_1) + \ldots + p_n(\vartheta_m)$, with the above notation, and the statement follows by induction. If v_0 is in a cluster on scale n we call $\tilde\vartheta_1, \ldots, \tilde\vartheta_m$ the subdiagrams emerging from the cluster containing v_0 and such that $M(\vartheta_j) > 2^{-(n+2)/\tau}$, $j = 1, \ldots, m$: it will be $p_n(\vartheta) = 1 + p(\tilde\vartheta_1) + \ldots + p_n(\tilde\vartheta_m)$. Again we can assume $m = 1$, the other cases being trivial. But in such a case there will be only one branch entering the cluster T on scale n containing v_0 and it will have a momentum of scale $n' \le n - 1$. Therefore the cluster T must contain vertices such that at least $\sum_{v\in T} |\vec\nu_v| > 2^{-(n+2)/\tau}$ vertices, (otherwise, if λ is a line on scale n contained in T, and $\vec\nu^0_\lambda$ is the sum of the mode labels corresponding to the vertices following v_0 but inside T, we would have $|\vec\omega \cdot \vec\nu_\lambda| \le 2^{n+1}$ and, simultaneously, $|\vec\omega \cdot \vec\nu_\lambda| \ge 2^{n+3} - 2^{n-1} > 2^{n+2}$, which would lead to a contradiction). This means that $M(\vartheta_1) \le M(\vartheta) - 2^{-(n+2)/\tau}$.

---

[6] Note that $M(\vartheta) - M(\vartheta_1) < 2 - 12^{-(n+2)/\tau}$ implies that $|\vec\omega \cdot \lambda| \ge 2^{1+(n+2)/\tau}$ for all the lines λ preceding the root line and not contained in ϑ_1, so that the set composed by such lines, if $\vec\nu_{\lambda_{v_0}} = \vec\nu_{\lambda_{v_1}}$, is a resonance on scale $n' > n$.

From the above proven results, (5.4) follows, if we note that $\sum_{T,n_T=n} 1 = p_n(\vartheta)$.

Appendix A2. Relaxing of the hypothesis (2.1)

If $f(\vec{\alpha};\vec{A};\varepsilon)$ is analytic in ε, then we can write

$$f(\vec{\alpha},\vec{A};\varepsilon) = \sum_{k=1}^{\infty} f^{(k)}(\vec{\alpha},\vec{A})\,\varepsilon^k \;, \tag{A2.1}$$

where $|f^{(k)}| \le F\,\varepsilon_1^{-k}$ in a domain $|\varepsilon| \le \varepsilon_1$.

Let us consider first the hamiltonian in Theorem 1.2, and trees such that $\delta_v = 1\ \forall v \in \vartheta$. Then a diagrammatic expansion is still possible, and the only difference is that now to order k we have to consider all the possible graphs with p vertices, $p = 1, \ldots, k$, such that (1) to each vertex v a factor $f^{(k_v)}_{\vec{\nu}_v}$ is associated, and (2) the k_v's labels have to satisfy the constraint $\sum_{v\in\vartheta} k_v = k$.

For each tree a bound C_1^p can be obtained: this can be easily argued from the discussion for the interaction (2.1), and the constant C_1 is the same one (up to the factor F, which now is missing as the factors $f_{\vec{\nu}_v}$ are replaced by the new ones $f^{(k_v)}_{\vec{\nu}_v}$), times the product $\prod_{v\in\vartheta} F\,\varepsilon_1^{-k_v}$. Then we have to consider all the possible ways to assign the factors $f^{(k_v)}_{\vec{\nu}_v}$, i.e. the k_v labels, to the vertices of the tree, which gives a sum

$$\sum_{p=1}^{k} \sum_{\substack{k_1+\ldots+k_p=k \\ \{k_i\ge 1\}_{i=1}^p}} C_1^p\,[F\,\varepsilon_1^{-1}]^k \le [F\,\varepsilon_1^{-1}]^k \sum_{p=1}^{k} \frac{C_1^p\,k^p}{p!} \le [e\,C_1\,F\,\varepsilon_1^{-1}]^k \;, \tag{A2.2}$$

so that, like in the case discussed previously, (interaction of the form (2.1)), we find again a bound C_2^k, where now $C_2 = C_0\,e\,\varepsilon_1^{-1}$, if C_0 was the value previously obtained. If there are also vertices $v \in \vartheta$ with $\delta_v = 0$, then the previous discussion has to be restricted to the vertices having $\delta_v = 1$, and the same result holds. This concludes the proof of Theorem 1.2.

In the case of Thorem 1.4, we can repeat the same analysis, by restricting it again to the k vertices v's having $\delta_v = 1$, (as the other ones correspond to factors $\mathcal{H}_0$), and the same result is obtained.

Appendix A3. A partition of unity via characteristic functions

Besides the partition of unity described in §4, there are other possibilities that could be envisaged. A very natural one is discussed in this Appendix.

Let us define

$$\chi_n(x) = \theta(|x| - 2^{n-1}) - \theta(|x| - 2^n) \;, \quad n \le 0 \;, \qquad \chi_1(x) = \theta(|x| - 1) \;, \tag{A3.1}$$

where $\theta(x)$ is the Heaviside function, defined as

$$\theta(x) = \begin{cases} 1\ , & \text{if } x > 0\ , \\ 1/2\ , & \text{if } x = 0\ , \\ 0\ , & \text{if } x < 0\ . \end{cases} \tag{A3.2}$$

Then, for any $x \in \mathbb{R}$, we have

$$\sum_{n=-\infty}^{1} \chi_n(x) = 1\ , \tag{A3.3}$$

so that we can define the "propagator at scale n" as

$$g^{(n)}(\vec{\omega}\cdot\vec{\nu}_\lambda) = \frac{\chi_n(\vec{\omega}\cdot\vec{\nu}_\lambda)}{(\vec{\omega}\cdot\vec{\nu}_\lambda)^{R_\lambda}}\ , \tag{A3.4}$$

which replaces (4.2).

Then the discussion of §5 remains unchanged, the only (irrelevant) difference being that the sum in item (5) before (5.3) is over $\le 2^{2k-1}$ terms. In fact, if the momentum $\vec{\nu}_\lambda$ in (A3.2) is fixed, there is only one scale n such that the propagator is not vanishing, (see note 5 for analogous considerations for the partition (4.1)), except the case in which $\vec{\omega}\cdot\vec{\nu}_\lambda$ is a diadic point 2^n, so that two successive scales are possible, but then (A3.1) and (A3.2) give a factor $1/2$ for each scale.

The discussion in §6 about the approximate cancellations can be easily adapted. Looking at (6.4) and (6.6), one could think that problems arise from derivatives of the functions (A3.1), since delta functions appear:

$$\dot{\chi}_n(x) = \delta(|x| - 2^{n-1}) - \delta(|x| - 2^n)\ . \tag{A3.5}$$

In fact, we can rule out all contributions containing any derivatives $\dot{\chi}_{n_{\lambda_v}}(\vec{\omega}\cdot\vec{\nu}_{\lambda_v})$, because the corresponding localized resonance factor either is vanishing or gives a vanishing contribution when the sum over the scales is performed.

This is a property which follows from the fact that, from definition (A3.5), the derivatives of the characteristic functions in (A3.1) can be different from zero *only* if some $\vec{\omega}\cdot\vec{\nu}_\lambda$ falls on the boundary of some diadic interval $[2^{n_\lambda-1}, 2^{n_\lambda}]$, say or 2^n, if $n = n_\lambda$ or $n = n_\lambda - 1$.

But in such a case, by starting from the outermost (*i.e.* maximal) resonances for which this happens, we consider together the values of the two trees in which the scale label of that line is n and $n+1$. Then, by denoting as usual n_V the scale of the resonance V as a cluster and n_{λ_V} the resonance-scale, if one has not $\vec{\omega}\cdot\vec{\nu}_\lambda = 2^{n_\lambda-1} = 2^{n_V-1} = 2^{n_{\lambda_V}}$, both cases $n_\lambda = n, n+1$ are compatible with the resonance structure, and we see that the two values we obtain by (1) considering the derivative of the χ_{n_λ} and (2) collecting together the $\mathcal{L}$ and $\mathcal{R}$ parts of the resonance factors for all resonances inside V, differ only because the first one contains a delta $-\delta(|\vec{\omega}\cdot\vec{\nu}(v)| - 2^n)$, whereas the latter contains the same delta, *but with opposite sign*, *i.e.* $\delta(|\vec{\omega}\cdot\vec{\nu}(v)|-2^n)$, (see (A3.5)), being all the other factors equal. Therefore the two tree values are opposite, and, when summed together, cancel exactly.

On the contrary, in the case $\vec{\omega}\cdot\vec{\nu}_\lambda = 2^{n_\lambda-1} = 2^{n_V-1} = 2^{n_{\lambda_V}}$, if we assign the scale label $n+1$ to the line λ, then we destroy the cluster structure, and we have no more a resonance. But of course we can again define the localized part of the quantity obtained from the resonance factor by shifting by 1 the scale of λ: then such a part cancels with the localized part of the resonance factor corresponding to V, (exactly as before), while the remaining part can be easily handled as the $\mathcal{R}$ part of the resonance factor.

Once the maximal clusters are treated, we pass to the next-to-maximal resonances, *i.e.* to the maximal resonances contained inside the maximal ones, and we study in the same way the localized parts. And so on until the innermost resonances are dealt with.

A similar analysis is required in order to adapt the discussion in §7, where some $\dot{\chi}_{n_\lambda}(\vec{\omega}\cdot\vec{\nu}_\lambda^0+t\vec{\omega}\cdot\vec{\nu})$ and $\ddot{\chi}_{n_\lambda}(\vec{\omega}\cdot\vec{\nu}_\lambda^0+t\vec{\omega}\cdot\vec{\nu})$ appear ($\vec{\nu}_\lambda^0$ is defined in Remark 6.2, and t is the interpolation parameter introduced in (7.5)).

When t varies in $[0,1]$, $\vec{\omega}\cdot\vec{\nu}_\lambda^0+t\vec{\omega}\cdot\vec{\nu}$ varies in $[\vec{\omega}\cdot\vec{\nu}_\lambda^0,\vec{\omega}\cdot\vec{\nu}_\lambda^0+\vec{\omega}\cdot\vec{\nu}]$. Then we can proceed in the following way, by starting from the maximal resonances as before and by considering all the possible labels assignments inside the resonance. In particular in this way, for each line $\lambda\in V$, we sum all the scales $n_\lambda > n_{\lambda_V}$, so that the functions (A3.1) give a function $\vartheta(|\vec{\omega}\cdot\vec{\nu}_\lambda|-2^{n_{\lambda_V}})$. Then the derivative of such a function can give a delta, only if $\vec{\omega}\cdot\vec{\nu}_\lambda^0+t\vec{\omega}\cdot\vec{\nu}$ falls on $2^{n_{\lambda_V}}$ for some $t\in[0,1]$. But if this occur, then we can perform the integration on t: no gain is obtained in such an operation, but no gain is really needed in such a case (which would correspond to have a scale label $n_{\lambda_V}+1$ on λ). Then we pass to the next-to-maximal resonances, we apply the $\mathcal{L}$ and $\mathcal{R}$ operators to the corresponding resonance factors, and we proceed in the same way, in order to study the $\mathcal{R}$ part of the resonance factor. And so on, until all the resonances are studied.

At this point, we can repeat the discussion in §7, and the same results can be obtained.

References

[A1] V.I. Arnol′d: Small divisor problems in classical and celestial mechanics, *Uspekhi Mat. Nauk* **18**, No. 6, 91-192 (1963); english translation in *Russ. Math. Surv.* **18**, No. 6, 85–191 (1963).

[A2] V.I. Arnol′d: Small divisors II. Proof of a A.N. Kolmogorov's theorem on conservation of conditionally periodic motions under small perturbations of the hamiltonian, *Uspekhi Mat. Nauk* **18**, No. 5, 13–40 (1963); english translation in *Russ. Math. Surv.* **18**, No. 5, 9–36 (1963).

[Ba] F. Bacon: *Essays*, LVIII; see also J.L. Borges: *El inmortal*.

[B] A.D. Bryuno: The analytic form of differential equations I, *Trans. Moscow Math. Soc.* **25**, 131–288 (1971); II, *Trans. Moscow Math. Soc.* **26**, 199–239 (1972).

[CF1] L. Chierchia, C. Falcolini: A direct proof of a theorem by Kolmogorov in hamiltonian systems, *Ann. Scuola Norm. Super. Pisa Cl. Sci. Ser. IV* **21**, No. 4, 541–593 (1995).

[CF2] L. Chierchia, C. Falcolini: Compensations in small divisors problems, preprint, Roma (1994), archived in `mp_arc@math.utexas.edu`, #94-270,

[CG] L. Chierchia, G. Gallavotti: Drift and diffusion in phase space, *Ann. Inst. H. Poincaré Phys. Théor.* **60**, No. 1, 1–144 (1994).

[CZ] L. Chierchia, E. Zehnder: Asymptotic expansions of quasi–periodic motions, *Ann. Scuola Norm. Super. Pisa Cl. Sci. Ser. IV* **16**, – (1989).

[DS] E.I. Dinaburg, Ya.G. Sinai: The one dimensional Schrödinger equation with a quasi periodic potentional, *Func. Anal. and its Appl.* **9**, 279–289 (1975).

[E1] L.H. Eliasson: Absolutely convergent series expansions for quasi-periodic motions, Report 2–88, Department of Mathematics, University of Stockholm (1988).

[E2] L.H. Eliasson: Hamiltonian systems with linear normal form near an invariant torus, in *Nonlinear Dynamics*, Ed. G. Turchetti, Bologna Conference, 30/5 to 3/6 1988, World Scientific, Singapore (1989).

[E3] L.H. Eliasson: Generalization of an estimate of small divisors by Siegel, in *Analysis, et cetera*, Ed. E. Zehnder & P. Rabinowitz, book in honor of J. Moser, Academic press (1990).

[EV] J. Ecalle, B. Vallet: The trimmed and corrected forms of resonant vector fields or diffeomorphisms, preprint, Paris (1994).

[FT] J. Feldman, E. Trubowitz: Renormalization on Classical Mechanics and Many Body Quantum Field Theory, *J. Anal. Math.* **58**, 213–247 (1992).

[G1] G. Gallavotti: *Elements of Mechanics*, Springer, New York (1983), translation of *Meccanica Elementare*, Boringhieri, Torino (1980).

[G2] G. Gallavotti: A criterion of integrability for perturbed harmonic oscillators. "Wick Ordering" of the perturbations in classical mechanics and invariance of the frequancy spectrum, *Comm. Math. Phys.* **87**, 365–383 (1982).

[G3] G. Gallavotti: Perturbation theory for classical hamiltonian systems, in *Scaling and Self-Similarity in Physics*, Ed. J. Frölich, Birkhäuser, Boston (1983).

[G4] G. Gallavotti: Quasi–integrable mechanics sysyems, in *Critical Phenomena, Random Systems, Gauge Theories*, Les Houches, Session XLIII (1984), Vol. II, 539–624, Ed. K. Osterwalder & R. Stora, North Holland, Amsterdam (1986).

[G5] G. Gallavotti: Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods, *Rev. Modern Phys.* **57**, 471–572 (1985).

[G6] G. Gallavotti: Classical Mechanics and Renormalization Group, Lectures notes at the 1983 Erice summer school, Ed. A. Wightman & G. Velo, Plenum, New York (1985).

[G7] G. Gallavotti: Twistless KAM tori, *Comm. Math. Phys.* **164**, 145–156 (1994).

[G8] G. Gallavotti: Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable hamiltonian systems. A review, *Rev. Math. Phys.* **6**, No. 3, 343–411 (1994).

[G9] G. Gallavotti: Perturbation Theory, in *Mathematical Physics towards the XXI Century*, 275–294, Ed. R. Sen & A. Gersten, Ben Gurion University Press, Ber Sheva (1994).

[GG] G. Gallavotti, G. Gentile: Majorant series convergence for twistless KAM tori, *Ergodic Theory Dynam. Systems*, to appear (1995).

[GGM] G. Gallavotti, G. Gentile, V. Mastropietro: KAM theorem and quantum field theory, preprint, Bures sur Yvette, (1995), archived in `mp_arc@math.utexas.edu`, #95-151.

[Ge1] G. Gentile: A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable hamiltonian systems, *Forum Math.*, to appear (1995).

[Ge2] G. Gentile: Whiskered tori with prefixed frequencies and Lyapunov spectrum, *Dynam. Stability of Systems*, to appear (1995).

[GM1] G. Gentile, V. Mastropietro: KAM theorem revisited, preprint, Roma (1994), archived in `mp_arc@ math. utexas.edu`, #94-403.

[GM2] G. Gentile, V. Mastropietro: Tree expansion and multiscale analysis for KAM tori, preprint, Roma (1994), archived in `mp_arc@math. utexas. edu`, #94-405.

[GL] A. Giorgilli, U. Locatelli: Kolmogorov theorem and classical perturbation theory, preprint, Milano (1995).

[HC] J.D. Hanson, J.R. Cary: Elimination of stochasticity in stellarators, *Phys. Fluids* **27**, 767–769 (1984).

[HP] F. Harary, E. Palmer: *Graphical Enumerations*, Academic Press, New York (1973).

[K] A.N. Kolmogorov: On the preservation of conditionally periodic motions, *Dokl. Akad. Nauk* **96**, 527–530 (1954); english translation in G. Casati, J. Ford: *Stochastic behavior in classical and quantum hamiltonian*, Lectures Notes in Physics **93**, Springer, Berlin (1979).

[L] A. Lindstedt: *Mémoires de l'Académique de Saint-Pétersbourg* (1882), quoted in H. Poincaré: *Les méthodes nouvelles de la mécanique céleste*, Vol.II, Gauthier-Villars, Paris (1893), *partim.*

[M1] J. Moser: On invariant curves of an area preserving mapping of the annulus, *Nachr. Akad. Wiss. Göttingen Math.–Phys. Kl. II*, 1–20 (1962).

[M2] J. Moser: A rapidly convergent iteration method and nonlinear differential equations II, *Ann. Scuola Norm. Super. Pisa Cl. Sci. Ser. III*, **20**, 499–535 (1966).

[M3] J. Moser: Convergent series expansions for quasi–periodic motions, *Math. Ann.* **169**, 136–176 (1967).

[M4] J. Moser: On the construction of almost periodic solutions for ordinary differential equations, *Proc. Int. Conf. of functional analysis and related topics*, Tokyo, 60–67 (1969).

[M5] J. Moser: *Stable and Random Motions in Dynamical Systems, with special emphasys on Celestial Mechanics*, Annals of Mathematical Studies 77, Princeton University Press, Princeton (1973).

[MS] J. Moser, C.L. Siegel: *Lectures on Celestial Mechanics*, Springer, Berlin (1971).

[N] J. Nash: The imbedding of Riemannian manifolds, *Ann. of Math.* **63**, 20–63 (1956).

[Ne] S. Newcomb: *Smithsonian contributions to knowledge* (1874), quoted in H. Poincaré: *Les méthodes nouvelles de la mécanique céleste*, Vol.II, Gauthier-Villars, Paris (1893), *partim.*

[P] H. Poincaré: *Les méthodes nouvelles de la mécanique céleste*, Gauthier-Villars, Paris, Vol. I (1892), Vol. II (1893), Vol. III (1899).

[Pö] J. Pöschel: Invariant manifolds of complex analytic mappings, in *Critical Phenomena, Random Systems, Gauge Theories*, Les Houches, Session XLIII (1984), Vol. II, 949–964, Ed. K. Osterwalder & R. Stora, North Holland, Amsterdam (1986).

[R1] H. Rüssmann: Über der Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtlösung, *Math. Ann.* **169**, 55–72 (1967).

[R2] H. Rüssmann: Über invariante Kurven differenzierbaren Abbildungen eines Kreisringes, *Nachr. Akad. Wiss. Göttingen Math.–Phys. Kl. II*, 67–105 (1970).

[R3] H. Rüssmann: One dimensional Schrödinger equation, *Ann. New York Acad. Sci.* **357**, 90–107 (1980).

[S1] C.L. Siegel: Iterations of analytic functions, *Ann. of Math.* **43**, No.4, 607–612 (1943).

[S2] C.L. Siegel: Über die Existenz einer Normalform analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtlösung, *Math. Ann.* **128**, 144–170 (1954).

[SHS] R.M. Sinclair, J.C. Hosea, G.V. Sheffield: A method for mapping a toroidal magnetic field by storage of phase stabilized electrons, *Rev. Sci. Instrum.* **41**, 1552–1559 (1970).

[T] W. Thirring: *Course in Mathematical Physics*, Springer, New York, Vol. 1 (1978), Vol. 2 (1979), Vol. 3 (1981), Vol. 4 (1983), translation of *Lehrbuch der Mathematischer Physik*, Springer, Wien, Vol. 1 (1977), Vol. 2 (1978), Vol. 3 (1979), Vol. 4 (1980).

[Ti] E.C. Titchmarsh: *The theory of functions*, Oxford University Press, London (1932).

[W] K.T.W. Weierstrass: Letters published in *Act. Math.* **35**, 29–65 (1911), quoted in J. Moser: *Stable and Random Motions in Dynamical Systems, with special emphasys on Celestial Mechanics*, Annals of Mathematical studies 77, Princeton University Press, Princeton (1973), Ch. I, §2.