ANNALI DELLA
SCUOLA NORMALE SUPERIORE DI PISA
*Classe di Scienze*

L. CHIERCHIA
C. FALCOLINI

**A direct proof of a theorem by Kolmogorov in hamiltonian systems**

*Annali della Scuola Normale Superiore di Pisa, Classe di Scienze 4e série*, tome 21, nº 4 (1994), p. 541-593.

<http://www.numdam.org/item?id=ASNSP_1994_4_21_4_541_0>

# A Direct Proof of a Theorem by Kolmogorov in Hamiltonian Systems

L. CHIERCHIA - C. FALCOLINI (*)

**Contents**



## 1. - Introduction

Each Section of this paper begins with a short summary of what is done in that Section: Here we present a bit of history of Kolmogorov's Theorem on the stability of quasi-periodic solutions in Hamiltonian systems followed by a rough outline of a novel proof, based on a tree representation of formal series.

A 1954 theorem by Kolmogorov [19] guarantees the existence of infinitely many quasi-periodic solutions for the (standard) Hamilton equations associated to real-analytic, "spatially periodic" Hamiltonians of the form $H(y, x; \varepsilon) = h(y) + \varepsilon f(y, x)$, ($y \in U$ open subset of $\mathbb{R}^N$, $x \in \mathbb{T}^N \equiv \mathbb{R}^N/(2\pi\mathbb{Z}^N)$, $\varepsilon \geq 0$), provided the Hessian matrix $\left(\dfrac{\partial h}{\partial y_i \partial y_j}\right)$ is invertible on $U$ and provided $|\varepsilon|$ is sufficiently

(*) The authors are indebted with one of the referees for having pointed out a mistake in the proof of Lemma (5.2) of a previous version of this paper. They are also grateful to C. Liverani for helpful discussions.

Pervenuto alla Redazione il 14 Giugno 1993 e in forma definitiva il 7 Febbraio 1994.

small. We recall that a solution $(y(t), x(t))$ is called (maximal) quasi-periodic if there exists a rationally independent vector $\omega \in \mathbb{R}^N$ and smooth functions $Y, X : \mathbb{T}^N \to \mathbb{R}^N$ such that $(y(t), x(t)) = (Y(\omega t), \omega t + X(\omega t))$. A detailed proof, different from that outlined by Kolmogorov in [19], was provided, in 1963, by V.I. Arnold in [1] and J. Moser [20] proved, in the same period, an analogous theorem for symplectic diffeomorphisms removing the hypothesis of analyticity of the perturbation. The *corpus* of results and methods stemmed out from Kolmogorov's original ideas is now known as "KAM (Kolmogorov-Arnold-Moser) theory".

It is not difficult to write down *formal $\varepsilon$-expansion* of quasi-periodic solutions; the problem is then turned into whether such series are convergent or not. This question was extensively and thoroughly investigated by H. Poincaré in his *méthodes nouvelle de la mécanique céleste* [23]. Poincaré, following Lindstedt (*Mémoires de l'Académie de Saint-Pétersbourg*, 1882), considered formal quasi-periodic solutions concluding that such series are likely to be divergent[1]. The main problem in this context is that the formal solution has Fourier coefficients which are divided by terms of the type $\omega \cdot n \equiv \sum_{i=1}^{N} \omega_i n_i$ with $n \in \mathbb{Z}^N \backslash \{0\}$, and such factors ("small divisors") become arbitrarily small as $|n| \to \infty$. In fact (see Appendix B below), the *repeated occurrence of small divisors* ("resonances") in the $k^{\text{th}}$ coefficient of the formal solution leads to contributions of the order of $(k!)^a$ with $a > 0$. However, in 1967 Moser [21] showed *indirectly* (see below) that the formal serie *converges* leading to analytic solutions, provided the *$\varepsilon$-independent* vector $\omega$ satisfies certain number theoretic assumptions verified by almost all (with respect to Lebesgue measure) vectors in $\mathbb{R}^N$; see the "Diophantine condition" (2.7) below. This means that the $k^{\text{th}}$ coefficient of the formal solution contains *many huge contributions which compensate among themselves* producing terms that behave like a constant to the $k^{\text{th}}$ power. Moser's proof, as well as *all* proofs in KAM theory (up to the 1988 Eliasson work [11]), are based on a "rapidly convergent" iteration technique. The strategy, similarly to Newton's method of tangents, consists in finding solutions of a nonlinear differential equation $\mathcal{N}(u) = 0$ by recursively solving a sequence of approximate equations of the form $\mathcal{N}(u_j) = e_j$ where the size of the "error function" $e_j$ becomes *quadratically smaller* at each step of the procedure. Such an approach is very powerful and can be used very effectively (see [22] for applications to partial differential equations, [7] for accurate estimates and Appendix 2 in [9] for a five-page proof) but is indirect and hides the mechanism beyond the above mentioned compensations.

[1] "M. Lindstedt ne démontrait pas la convergence des développements qu'il avait ainsi formés, et, en effet, ils sont divergents" ([23], vol. II, §IX, n° 123); and later: "Il semble donc permis de conclure que le séries (2) ne convegent pas. Toutfois le raisonnement qui précède ne suffit pas pour établir ce point avec une rigueur complète [...] Tout ce qu'il m'est permis de dire, c'est qu'il est fort invraisemblable." ([23], vol. II, §XIII entitled "Divergence des séries de M. Lindstedt", n° 149).

In the different context of linearization of germs of analytic diffeomorphisms, C.L. Siegel [26] succeeded in 1942 in proving directly the convergence of formal power series involving small divisors: the crucial difference being that the small divisors are of the form $\omega \cdot n$ *with* $n \in \mathbb{N}^N$ so that a given divisor *occurs at most once* in each coefficient of the formal solution (*i.e.* "there are no resonances").

In 1988 Eliasson, in a Report of the University of Stockholm [11] (which to the best of our knowledge has not been published elsewhere[2]), extended Siegel's method so as to cover the Hamiltonian case. We present a different version of Eliasson's proof[3] based on a tree representation of the formal solutions and on the explicit exhibition of compensations of huge contributions. We follow closely [11] in its convenient reformulation of Siegel's method (Appendix C) which allows to bound product of (possibly) small divisors whenever (certain dangerous) repetition do not occur; however for the crucial part (grouping together the huge contributions) we adopt a different approach which now we outline.

In order to simplify the presentation *we consider the particular model with Hamiltonian* $H = \frac{1}{2} y \cdot y - \varepsilon f(x)$: Such a choice has the advantage of making (we hope) more transparent the presentation without introducing (we believe) any essential modification.

The *starting point* is to express the coefficients of the formal solution in terms of *labeled rooted trees*[4]. This allows to express the $k^{\text{th}}$ coefficient in terms of a sum over all possible labeled rooted trees of order $k$, and over all possible (and "admissible") integers $\alpha_v \in \mathbb{Z}^N \backslash \{0\}$ ($v$ denoting a vertex of a tree) of

$$\prod_V f_{\alpha_v} \prod_E \alpha_v \cdot \alpha_{v'} \prod_V \delta_v^{-2} \tag{1.1}$$

where $V$ and $E$ denote, as customary, the set of vertices $v$ and edges $vv'$ of a given tree (see Appendix A), $f_n$ denote Fourier coefficients of $f$ and $\delta_v$ represents the divisor associated to the vertex $v$ *i.e.* $\delta_v \equiv \omega \cdot \sum_{v' \le v} \alpha_{v'}$ (rooted trees have a natural order according to which the root $r$ is $> v$ for all vertices

[2] See, however, the excerpts from the proceedings [12] and [13].

[3] For a short comparison between our approach and that of Eliasson, see the end of this introduction; more technical comments, at this regard, are given throughout the rest of the paper (see, especially, Remark 4.1).

[4] The use of graph theory in connection with formal power series is natural and very old (see *e.g.* [17] and references therein); for connections with KAM theory see [8] and [15]. In [15] the problem analyzed here is also investigated with similar tools, emphasizing, in particular, the similarities with renormalization group approaches to quantum field theory. However, the families of trees considered there, as well as most of the technical aspects are different from ours; in particular stronger assumptions on $f$ (assumed to be an even trigonometric polynomial) and $\omega$ (assumed to be a "strong Diophantine" vector) are made; in [16] the "strong Diophantine" hypothesis has been removed.

$v \in V$, $v \neq r$; the square in (1.1) comes from the fact that the Hamilton equations for our model can be immediately written as the second order system $\ddot{x} = f_x$; "admissible" means that $\delta_v \neq 0$ $\forall v$ *i.e.* $\sum_{v' \leq v} \alpha_{v'} \neq 0$ $\forall v \in V$).

Now, the *main point* is to make a *partition* of the trees of order $k$ into families $\mathcal{F}$ (called below "complete families") so as to be able to bound sums over such families by a constant to the $k^{\text{th}}$ power (recall that even single contributions given by (1.1) may have size of order $(k!)^a$ as already mentioned). The main technical estimate is

$$\left| \sum_{T \in \mathcal{F}} \prod_{vv' \in E(T)} \alpha_v \cdot \alpha_{v'} \prod_V \delta_v^{-2} \right| \leq \prod_{v \in V} |\alpha_v|^{\deg_{\mathcal{F}} v} c_4^k \prod_{v \in V} |\alpha_v|^{\beta_4} \tag{1.2}$$

for suitable constants $c_4$ and $\beta_4$, determined below, depending on $N$ and on the numerical properties of the vector $\omega$; $\deg_{\mathcal{F}} v$ is defined as the maximum degree of $v$ when $T$ varies in the family $\mathcal{F}$. Obviously, it is crucial that complete families either do not intersect or coincide (in fact it is much easier to find families of trees for which (1.2) holds but which do not form a partition: clearly such families are of no use for our purposes). The construction of the partition of complete families is the delicate part of our paper. Given Siegel's method (which we include for completeness in Appendix C) and the identification of complete families, the convegence of formal solutions follows very easily (under the Diophantine condition on $\omega$). All the constants are computed explicitly (see Remark 5.1 below).

Let us now point out possible directions of future investigations.

(i) The method of proof presented here, being based on a very direct approach, seems particularly suitable for computer-aided implementations and might shed some new light on the difficult problem of the break-down of stability of quasi-periodic solutions in connection with the $\varepsilon$-singularities of the function $X$ (see, *e.g.*, [7], [3], [4], [14] and references therein).

(ii) Let $(x^1, x^2)$ with $x^i \in \mathbb{T}^{N_i}$ and $N_1 + N_2 = N$, let $\omega \in \mathbb{R}^{N_1}$ and let $x_0^2$ be a nondegenerate critical point of the periodic function $x^2 \in \mathbb{T}^{N_2} \to \int_{\mathbb{T}^{N_1}} f(x^1, x^2) dx^1$. Then, it is not difficult to see that (if $\omega$ satisfies a Diophantine condition) there are formal (non maximal) quasi-periodic solutions whose first term ($\varepsilon = 0$) is given by $(\omega t, x_0^2)$. It would be nice to extend the proof in this paper so as to establish convergence for such formal series.

(iii) It is well known ([24]) that maximal quasi-periodic solutions are stable under weaker assumptions on the vector $\omega$ than the classical one made here. It might be interesting to see how far, in this direction, can lead the technique worked out in this paper.

The paper is organized as follows. In Section 2 existence and uniqueness of formal quasi-periodic solutions are established. In Section 3 the tree expansion of formal quasi-periodic solutions is given (the purely combinatorics aspects are proven in Appendix B where the occurrence of huge terms is also explicitly exhibited). The construction of complete families is carried out in Section 4 allowing to formulate the results in Section 5 divided in four Lemmas and one Theorem. Detailed proofs are given in Section 6 and in the appendices B and C. Graph theory is used here mainly as a (very useful!) language and the standard definitions (enough to read our paper without any knowledge of graph theory) are presented in Appendix A.

We close this introduction with a few words of comment about the relation of our approach with the work in [11]. The reason why Siegel strategy does not work in the Hamiltonian case is due to repetitions of the value of some divisor $\delta_w$ in the product $\prod_V \delta_v$ in (1.1); to such repetitions correspond subtrees $R \subset T$, called below resonances, such that $\sum_{v \in R} \alpha_v = 0$. It is actually easy to extend Siegel's original strategy so as to admit resonances $R$ such that the absolute value of the corresponding repeated divisor $|\delta_w|$ is "not too small" compared to some divisor $\delta_v$ with $v \in R$ (Appendix C). When this is not the case ("critical resonances"), single contributions (1.1) can, as already mentioned, have a size of $\sim k!$. Eliasson's approach is then the following. The formal solution is in general not unique since, for each $k$, one can choose an arbitrary constant ("phase shift"). In [11] use of such freedom is made by adding, at each order $k$, suitable "counter-terms" which are chosen to balance the divergences due to critical resonances. Such counter-terms are not identified *a priori* with suitable terms of the formal expansion, instead it is *a posteriori* shown that the convergent series thus obtained is still solution of the original problem. The strategy followed in this paper, as already mentioned above, is different: We fix from the beginning a formal solution choosing all the phase shifts equal to 0 (in fact this *identifies uniquely* the formal solution). We then group together suitable single contributions, coming from different trees, and, keeping track of the signs, we then prove that the sum of such groups of contributions can be bounded by a constant to the $k^{\text{th}}$ power showing the convergence of the series. In our opinion such strategy (which is also at the basis of [15], [16]) might be a relevant simplification, at least from a conceptual – if not from a technical – point of view, with respect to [11].

## 2. - Formal Quasi-Periodic Solutions

We review some known facts recalling, in particular, the notions of quasi-periodic and formal quasi-periodic solution for a "spatially periodic" Hamiltonian.

Let $H(y, x)$ be a $C^\infty(U \times \mathbb{T}^N)$ Hamiltonian where $U$ is an open set of $\mathbb{R}^N$ and

$\mathbb{T}^N \equiv \mathbb{R}^N/2\pi\mathbb{Z}^N$ (*i.e.* $H$ can be seen as a function of the $2N$ variables $y_1, \dots, y_N$, $x_1, \dots, x_N$, $2\pi$-periodic in $x_i$). Consider the standard Hamilton equations:

$$\dot{y} = -H_x, \quad \dot{x} = H_y \tag{2.1}$$

where, as usual, $H_x \equiv \partial_x H \equiv \left(\dfrac{\partial H}{\partial x_1}, \dots, \dfrac{\partial H}{\partial x_N}\right)$ and $H_y \equiv \partial_y H \equiv \left(\dfrac{\partial H}{\partial y_1}, \dots, \dfrac{\partial H}{\partial y_N}\right)$. A solution $(y(t), x(t))$ of (2.1) is called (*maximal*) *quasi-periodic* with frequency $\omega \in \mathbb{R}^N$ if $\omega$ is rationally independent (*i.e.* $\omega \cdot n \equiv \sum_{i=1}^N \omega_i n_i = 0$ for some $n \in \mathbb{Z}^N$ implies $n = 0$) and if there exist smooth functions $Y$, $X : \mathbb{T}^N \to \mathbb{R}^N$ such that:

$$y(t) = Y(\omega t), \quad x(t) = \omega t + X(\omega t). \tag{2.2}$$

The rational independence of $\omega$ easily implies that the functions $\theta \in \mathbb{T}^N \to Y(\theta)$, $X(\theta)$ satisfy the system of equations:

$$\begin{aligned} &DY = -H_x(Y, \theta + X), \qquad \left(D \equiv \sum_{i=1}^N \omega_i \partial_{\theta_i}\right) \\ &\omega + DX = H_y(Y, \theta + X) \end{aligned} \tag{2.3}$$

Viceversa, given a solution of (2.3), (2.2) (or more generally (2.2) with $\omega t$ replaced by $\theta + \omega t$ for any $\theta \in \mathbb{T}^N$) is a quasi-periodic solution of (2.1).

"Le problème général de la dynamique" according to Henri Poincaré ([23], vol. I, chapter I, §13) is the study of the equations governed by the "nearly-integrable" Hamiltonian

$$H(y, x; \varepsilon) \equiv h(y) + \varepsilon f(y, x) \tag{2.4}$$

for small values of the parameter $\varepsilon$. As we shall see below, if the Hessian matrix $h'' \equiv h_{yy} \equiv \left(\dfrac{\partial^2 h}{\partial y_i \partial y_j}\right)$ is invertible, then there are "a lot" of "formal quasi-periodic solutions" of the equations (2.1) with Hamiltonian (2.4).

A *formal $\varepsilon$-power series $F$, over* $\mathbb{T}^N$, is an infinite sequence of $C^\infty(\mathbb{T}^N)$ functions $\{F_k\}_{k \geq 0}$, $F_k = F_k(\theta)$, and it is customary to write $F \sim \sum_{k \geq 0} F_k \varepsilon^k$. If $g$ is a $C^\infty$ function and $F \sim \sum_{k \geq 0} F_k \varepsilon^k$ a formal series, one naturally defines the formal series $g \circ F \sim \sum_{k \geq 0} G_k \varepsilon^k$ by setting

$$G_k = \left[g\left(\sum_{h=0}^k F_h \varepsilon^h\right)\right]_k \equiv \frac{1}{k!}\frac{d^k}{d\varepsilon^k} g\left(\sum_{h=0}^k F_h \varepsilon^h\right)\Bigg|_{\varepsilon=0} \tag{2.5}$$

A formal solution of (2.3) with $H$ as in (2.4) is a couple of (vector-valued) formal $\varepsilon$-power series over $\mathbb{T}^N$, $Y \sim \sum_{k\geq 0} Y^{(k)}\varepsilon^k$, $X \sim \sum_{k\geq 0} X^{(k)}\varepsilon^k$, $(Y^{(k)}, X^{(k)} : \mathbb{T}^N \to \mathbb{R}^N)$, verifying (2.3) in the sense of formal series (*i.e.* "=" should be replaced by "$\sim$") or equivalently:

$$
\begin{aligned}
&DY^{(0)} = 0, \quad DY^{(k)} = \left[-f_x\left(\sum_{h=0}^{k-1} Y^{(h)}\varepsilon^h, \theta + \sum_{h=0}^{k-1} X^{(h)}\varepsilon^h\right)\right]_{k-1} \\
&\omega + DX^{(0)} = h_y(Y^{(0)}), \quad DX^{(k)} = \left[H_y\left(\sum_{h=0}^{k} Y^{(h)}\varepsilon^h, \theta + \sum_{h=0}^{k-1} X^{(h)}\varepsilon^h\right)\right]_{k}
\end{aligned}
\tag{2.6}
$$

where $k \geq 1$; notice the different ranges of indices in the equation for $DX^{(k)}$ due to the particular form of (2.4).

PROPOSITION 2.1. *Let $H$ as in* (2.4) *be $C^\infty$ in a neighborhood of $\{y_0\}\times\mathbb{T}^N$ with $y_0$ such that $h''(y_0)$ is invertible and $\omega \equiv h'(y_0) \equiv h_y(y_0)$ is "Diophantine" i.e. such that*[5]

$$
|\omega \cdot n| \geq \frac{1}{\gamma|n|^\tau}, \quad \forall n \in \mathbb{Z}^N\backslash\{0\} \textit{ for some } \gamma,\ \tau > 0. \tag{2.7}
$$

*Then there exists a unique formal quasi-periodic solution $Y$, $X$ of* (2.6) *with*

$$
\int_{\mathbb{T}^N} X^{(k)} d\theta = 0, \quad \forall k. \tag{2.8}
$$

A proof is given in Appendix B.

## 3. - Tree Expansion of Formal Series

A very explicit representation of the formal solutions described in Proposition 2.1 can be obtained by means of *trees*, as explained below. For a different representation see [27].

*From now on we restrict our attention to Hamiltonians* (2.4) *of the form*

$$
H(y, x; \varepsilon) = \frac{1}{2} y \cdot y - \varepsilon f(x), \quad \int_{\mathbb{T}^N} f(x)dx = 0. \tag{3.1}
$$

[5] It is well known that for any $\tau>N$, almost all (with respect to Lebesgue measure) $\omega\in\mathbb{R}^N$ satisfy (2.7) with some $\gamma>0$ (see, *e.g.*, [2], chapter I, §3) while if $\tau=N-1$ then for any $\omega\in\mathbb{R}^N$ there exist a $\gamma>0$ and an infinite number of $n\in\mathbb{Z}^N$ such that (2.7) is violated (this is a theorem by Dirichlet, see, *e.g.*, [25]).

Notice that, obviously, the average of $f$, as well as its sign, plays no role in Hamilton equations.

Thus, in the present case (3.1), Hamilton equations and the equations characterizing quasi-periodic solutions with frequencies $\omega$ are given respectively (see (2.1), (2.3)) by

$$\ddot{x} = \varepsilon f_x(x), \quad D^2 X = \varepsilon f_x(\theta + X) \tag{3.2}$$

The formal power series $X \sim \sum_{k\geq 1} X^{(k)}\varepsilon^k$, with $\int_{\mathbb{T}^N} X^{(k)} = 0$, whose existence and uniqueness (for $f \in C^\infty(\mathbb{T}^N)$) is guaranteed by Proposition 2.1, satisfies

$$D^2 X \sim \varepsilon f_x(\theta + X) \text{ or } D^2 X^{(k)} = \left[ f_x\left(\theta + \sum_{h=1}^{k-1} X^{(h)}\varepsilon^h\right)\right]_{k-1} \tag{3.3}$$

The rest of this section is devoted to a *tree representation* of the formal solution $X$ of (3.3) (with $\int_{\mathbb{T}^N} X^{(k)} = 0$).

We recall that a tree $T$ is a connected acyclic graph (see Appendix A for the fundamentals used here or see any introductory book on graph theory such as [18] or [5]). We denote respectively by $V = V(T)$ and $E = E(T)$ the set of *vertices* and *edges* (or *points* and *lines*) of the tree $T$. A *rooted tree* is a tree with one distinguished vertex called *root*; we shall usually denote $r$ such a point and $T_r$ the rooted tree obtained by selecting, as root, the vertex $r$ of the tree $T$. It will also be useful to regard a rooted tree $T_r$ as a tree with one extra point $\eta \notin V(T_r)$, called the *earth*, and one more edge $\eta r \in E(T_r)$ connecting the earth $\eta$ to the root $r$. It is natural to define on rooted trees a *partial ordering*: Given $T_r$ and $u, v \in V$ we say that $u \geq v$ or $u \geq_{T_r} v$ if the path with endpoints $r$ and $v$ passes through $u$; $u > v$ means obviously $u \geq v$ and $u \neq v$. In particular $r \geq v$ for any $v \in V$.

*We fix once and for all a Diophantine vector* $\omega \in \mathbb{R}^N$ *satisfying* (2.7) and denote the inner product between $n \in \mathbb{Z}^N$ and $\omega$ by

$$\langle n \rangle \equiv \omega \cdot n \equiv \sum_{i=1}^N \omega_i n_i. \tag{3.4}$$

Given a rooted tree $T_r$ and a function $\alpha : v \in V \to \alpha_v \in \mathbb{Z}^N$ we denote by $\delta_v(T_r; \alpha)$ (or $\delta_v(T_r)$ or $\delta_v$ when there is no ambiguity) the quantity

$$\delta_v(T_r; \alpha) \equiv \left\langle \sum_{\substack{v' \in V(T_r) \\ v' \leq v}} \alpha_{v'} \right\rangle . \tag{3.5}$$

Given a rooted tree $T_r$ and an integer-valued function $\alpha : V \to \mathbb{Z}^N$ we say that $\alpha$ is *$T_r$-admissible* if, for all vertices of $T_r$, one has

$$\alpha_v \neq 0; \quad \sum_{v' \leq v} \alpha_{v'} \neq 0. \tag{3.6}$$

We shall denote by $\mathcal{A}(T_r)$ the set of all $T_r$-admissible functions over $T_r$.

Finally, we let $\mathcal{T}^k$ denote the set of *rooted, labeled trees* with $k$ vertices and, for any integer-valued function $\alpha$ and any subset $S \subset V$ of vertices of a tree $T$, we denote $\alpha(S) \equiv \sum_{v \in S} \alpha_v$.

PROPOSITION 3.1. *The $j^{\text{th}}$ component of the n-Fourier coefficients of $X^{(k)}$ is given by*

$$\begin{aligned} X^{(k)}_{jn} &= \frac{1}{k!} \sum_{T_r \in \mathcal{T}^k} F_{jn}(T_r), \quad \text{with:} \\ F_{jn}(T_r) &\equiv (-i) \sum_{\substack{\alpha \in \mathcal{A}(T_r) \\ \alpha(T_r)=n, \alpha_\eta \equiv e_j}} \prod_{v \in V} f_{\alpha_v} \prod_{vv' \in E} \alpha_v \cdot \alpha_{v'} \prod_{v \in V} \delta_v^{-2} \end{aligned} \tag{3.7}$$

*where $e_j$ is the unit vector with all zeros except in the $j^{\text{th}}$ entry.*

The proof is given in Appendix B. Note that in the above formula $E = E(T_r)$ *includes* the edge $\eta r$ and for this reason the function $\alpha$ has been extended on $\eta$ (which is outside $V = V(T_r)$).

REMARK 3.1. The above function $X^{(k)}_{jn}$ is obviously a function of rooted trees rather than labeled rooted trees and the introduction of the labels has the only task of simplifying the combinatorial factors entering in the formula. In terms of rooted trees, (3.7) can be rewritten as follows. Let $\tilde{T}_r \equiv [T_r]$ denote the rooted tree obtained by removing the labels from $T_r$ (clearly $\tilde{T}_r$ can be seen as the equivalence class generated by $T_r$), and let $\ell(\tilde{T}_r)$ denote the number of ways of putting $k$ labels on $\tilde{T}_r$ (*i.e.* $\ell(\tilde{T}_r) = \#\{T'_r \in \mathcal{T}^k : T'_r \in \tilde{T}_r\}$). For example, if $\tilde{T}_r$ is the path of order $k$ rooted in one of its endpoints, then $\ell(\tilde{T}_k) = k!$. The "$\tilde{T}_r$-admissible" class of $\mathbb{Z}^N$-valued functions is defined in exactly the same way (see (3.6) replacing $T_r$ by $\tilde{T}_r$ (as the labeling did not enter in the definition of $\mathcal{A}(T_r)$). Finally denote by $\tilde{\mathcal{T}}^k$ the class of all rooted trees of order $k$. With this notations we see that

$$\begin{aligned} X^{(k)}_{jn} &= \sum_{\tilde{T}_r \in \tilde{\mathcal{T}}^k} \frac{\ell(\tilde{T}_r)}{k!} \varphi_{jn}(\tilde{T}_r), \quad \text{with:} \\ \varphi_{jn}(\tilde{T}_r) &\equiv (-i) \sum_{\substack{\alpha \in \mathcal{A}(\tilde{T}_r) \\ \alpha(\tilde{T}_r)=n, \alpha_\eta \equiv e_j}} \prod_{v \in V} f_{\alpha_v} \prod_{vv' \in E} \alpha_v \cdot \alpha_{v'} \prod_{v \in V} \delta_v^{-2}. \end{aligned} \tag{3.8}$$

Below it will be useful to exchange the sums (over trees and over integers $\alpha$'s) in (3.7), in which case we obtain the following formula:

$$X_{jn}^{(k)} = \frac{1}{k!} \sum_{\substack{n_1,\ldots,n_k \\ n_i \in \mathbb{Z}^N \setminus \{0\}, \Sigma n_i = n}} \sum_{T_r \in \mathcal{T}^k} \overline{F}_{j\{n_i\}}(T_r) \tag{3.9}$$

where if $v_1, \ldots, v_k$ are the labels of $\mathcal{T}^k$ and if we set $\alpha_{v_i} \equiv n_i$ for any choice of $\{n_i\}$, the function $\overline{F}$ is defined by $\overline{F}_{j\{n_i\}}(T_r) \equiv 0$, if $\exists v : \sum_{v' \leq v} \alpha_{v'} = 0$ and otherwise by

$$\overline{F}_{j\{n_i\}}(T_r) \equiv \prod_{v \in V} f_{\alpha_v} \prod_{vv' \in E} \alpha_v \cdot \alpha_{v'} \prod_{v \in V} \delta_v^{-2}. \tag{3.10}$$

## 4. - Resonances

Repetitions of small divisors correspond to *resonant subtrees i.e.* to subtrees $S$ of a given (labeled rooted) tree $T$ for which $\alpha(S) \equiv \sum_{v \in S} \alpha_v = 0$. Throughout the rest of the paper (unless otherwise stated), given $T \in \mathcal{T}^k$ (we shall often omit the explicit indication of the root $r$ appearing elsewhere as index of $T$ when this does not lead to confusion), the function $\alpha : V \equiv \{v_1, \ldots, v_k\} \to \mathbb{Z}^N$, $\{\alpha_{v_i}\} \equiv \{n_i\}$, is *admissible* in the usual sense that $n_i \neq 0$ and $\sum_{v' \leq v} \alpha_v \neq 0$ for all $v \in V$ (in which case $\delta_v \neq 0$ by the rational independence of $\omega$). In this section we develop the tools needed to *make a partition* of $\mathcal{T}^k$, for a given choice of $\{n_i\}$ (see (3.9) $\div$ (3.10)), into "*complete families*". In the next section we shall see that the main property of complete families is that (possibly) huge terms (coming from repetition of small divisors) *compensate among themselves* within the same class of the partition. This fact will allow us to bound the contribution to (3.9), coming from the sum over trees belonging to the same class of the partition, by a constant to the $k^{\text{th}}$ power.

The rest of this section consists basically in a sequel of definitions and checks of elementary properties of trees with a given admissible $\alpha_{v_i} \equiv n_i$, $v_1, \ldots, v_k$ being the labels of $\mathcal{T}^k$. Let us begin (with a bit of patience).

DEFINITION 4.1 (Degree of a subtree). Given a (possibly rooted) tree $T$ and $S \subset T$ (*i.e.* $S$ is a subtree of $T$ *i.e.* $S$ is a connected subgraph of $T$) we call *degree of* $S$, $\deg S$, the number of edges connecting $S$ with $T \setminus S$ (if $T$ is rooted and $r \in S$ the edge $\eta r$ has to be counted; see Figure 1 below where three subtrees of degree two are encircled).

DEFINITION 4.2 (Resonances). Given a rooted tree $T$ and an admissible function $\alpha$ on it, we say that the subtree $R \subset T$ is *resonant* if: (i) $\deg R = 2$;

(ii) $R$ is null *i.e.* $\alpha(R) = 0$; (iii) $R$ cannot be disconnected by removal of one edge into two null subtrees. A resonant subtree will also be called a *resonance*. It follows that a resonance $R$ is connected to $T\backslash R$ by two edges $u\overline{u}$ and $w\overline{w}$ with $\overline{u}$, $\overline{w}$ (possibly coincident) belonging to $R$ and $u > w$ outside $R$ ($u$ may coincide with the earth $\eta$). If $\overline{u} = \overline{w}$ we shall call $R$. following [11], a *short resonance* (the second resonance in Figure 1 is a short one).

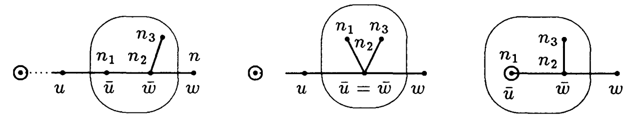



Fig. 1: $n_1 + n_2 + n_3 = 0$; (in the third example $u = \eta$)

DEFINITION 4.3 (Resonant couples). Given a rooted tree $T$ and a real number $\lambda > 0$, we say that a couple of points $(v, w) \in V \times V$ is $\lambda$-*resonant* if: (i) $v > w$; (ii) $\delta_v = \delta_w$; (iii) $v$ and $w$ are not adjacent and $|\delta_w| \leq \lambda|\delta_{v'}|$ for all $v'$ between $v$ and $w$.

Such a notion[6] is given also in [11] where the couple $(v, w)$ is called a critical resonance; we reserve such a name for a different object (Definition 4.5) closely related to:

DEFINITION 4.4 ($\lambda$-Resonances). Given a rooted tree $T$, $\lambda > 0$ and a resonance $R \subset T$, let $u\overline{u}$, $w\overline{w}$ be the edges connecting $R$ with $T\backslash R$, with[7] $\overline{u}$, $\overline{w} \in R$ and $u > \overline{u} \geq \overline{w} > w$; let also

$$n \equiv \sum_{v \leq w} \alpha_v, \quad \Delta(R) \equiv \min_{\substack{v_1, v \in R \\ v_1 \neq v}} |\delta_v(R_{v_1})|. \tag{4.1}$$

We say that $R$ is a $\lambda$-*resonance* (or a $\lambda$-resonant subtree) if

$$|\delta_w| = |\langle n \rangle| \leq \lambda \Delta(R). \tag{4.2}$$

It is easy to check that, since $\alpha(R) = 0$, $\Delta(R) = \min_{\substack{v \in R \\ v \neq v_1}} |\delta_v(R_{v_1})|$, for *any fixed* $v_1 \in R$.

REMARK 4.1. (i) To two points $v > w$ such that $\delta_v = \delta_w$ but $\delta_{v'} \neq \delta_w$ for any $v'$ between $v$ and $w$ (if there are any) it is always possible to as-

[6] Which is *not* fundamental in our approach but may help the reader in the comparison with [11].

[7] Here (and below) we make the common abuse of notation using "$v \in R$" in place of the more correct notation "$v \in V(R)$"; to read (4.1) recall that $\delta_v(R_u) = \sum_{\substack{v' \leq v \\ v' \in R}} \langle \alpha_{v'} \rangle$ where the order $\leq$ is that of the tree $R$ rooted at $u$.

sociate a resonant subtree $R \equiv R(v,w)$ of $T$ whose vertices are given by $V(R(v,w)) \equiv \{v' \in V : v' \leq v\} \backslash \{v' \in V : v' \leq w\}$. Notice however that $(v,w)$ may be $\lambda$-resonant while the associated resonance $R(v,w)$ is not $\lambda$-resonant. Consider, in fact, the first example in Figure 1 and set $v \equiv \overline{u}$: it is easy to see that one can choose $n_1$, $n_2$, $n_3$ so that $n_1 + n_2 + n_3 = 0$ (implying $\delta_v = \delta_w = \langle n \rangle$) and $|\langle n \rangle| \leq \lambda |\langle n_2 + n_3 + n \rangle|$ (implying $(v,w)$ $\lambda$-resonant) but $|\langle n \rangle| > \lambda |\langle n_3 \rangle|$ (implying $R(v,w)$ not $\lambda$-resonant).

To compare with the approach in [11] we note the following. Roughly speaking, *complete families* will be constructed by considering certain "critical" resonances $R$ and by grouping together all trees obtained by replacing the connecting edges $u\overline{u}$, $w\overline{w}$ with edges $uu'$, $ww'$ with $u'$, $w'$ *arbitrary points in* $R$ (e.g. the first two trees in Figure 1 might belong to the same complete family). It is therefore clear that the notion of "criticality" must be invariant under such operation and also that, in view of the above example, the notion of criticality introduced in [11] *is not adequate for our purposes.*

(ii) If $R$ is a non short $\lambda$-resonance and $u\overline{u}$, $w\overline{w}$ are the edges connecting $R$ with $T \backslash R$ (with $u > \overline{u} > \overline{w} > w$), for any $v'$ such that $\overline{u} > v' > w$, one has

$$\delta_{v'} \equiv \delta_{v'}(T) = \delta_{v'}(R_{\overline{u}}) + \delta_w$$

therefore, by (4.2),

$$|\delta_{v'}| \geq (1-\lambda)|\delta_{v'}(R_{\overline{u}})| \Longrightarrow |\delta_w| \leq \frac{\lambda}{1-\lambda}|\delta_{v'}|. \tag{4.3}$$

Hence, if $\lambda < 1$, the points $\overline{u}$ and $w$ *form a couple of* $\lambda'$*-resonant points with* $\lambda' \equiv \dfrac{\lambda}{1-\lambda}$.

(iii) We may classify resonances as follows. Let $R$ be a resonance and $u\overline{u}$, $w\overline{w}$ be the edges connecting $R$ with $T \backslash R$ (with $u > \overline{u} \geq \overline{w} > w$). We shall call $R$ a

(a) *Siegel resonance* (with parameter $\lambda$) if there exists a vertex $v'$ such that $\overline{u} > v' > w$ and $|\delta_w| > \lambda |\delta_{v'}|$;

(b) *Eliasson resonance* (with parameter $\lambda$) if $(\overline{u}, w)$ form a couple of $\lambda$-resonant points;

(c) $\lambda$*-resonance*, as above, if $R$ satisfies Definition 4.4[8].

[8] The reason for the names is the following: Siegel resonances can be treated basically by Siegel's original technique (a fact easy to justify intuitively since the repetition of the divisor $\delta_w = \delta_v$ due to a Siegel resonance $R(v,w)$ can be controlled in terms of some divisor $\delta_{v'}$ with $v' \in R$; see Appendix C). Eliasson resonances are those which are called critical by Eliasson and controlling the repetition coming from such resonances constitutes the main difficulty in [11]; with our approach (Lemma 5.2) we shall be able to extend Siegel's method so as to control those Eliasson resonances which are not $\lambda$-resonant (however in order to do this a rather careful "topological" analysis is needed: see Subsection 6.2). Finally $\lambda$-resonances (and $\lambda$-subresonaces: see below) are responsible for repetitions which *cannot be bounded effectively by looking at a fixed tree* (as done

Observe that a short resonance ($\overline{u} = \overline{w}$) may be a $\lambda$-resonance or not but is never Siegel or Eliasson. Other relations have already being pointed out above: e.g. point (i) of this Remark can be rephrased by saying that an Eliasson resonance with parameter $\lambda$ need not be a $\lambda$-resonance (and, in fact, need not be a $\lambda'$-resonance for any prefixed $\lambda'$), while point (ii) says that a non-short $\lambda$-resonance is an Eliasson resonance with parameter $\dfrac{\lambda}{1-\lambda}$.

PROPOSITION 4.1 (Non overlapping of resonant couples). *Let $(v_i, w_i)$ for $i = 1, 2$ be couples of $\lambda_i$-resonant points. If $v_1 > v_2 > w_1 > w_2$ then either $\lambda_1 \geq 1$ or $\lambda_2 \geq 1$.*

PROOF. By contradiction: Assume both $\lambda_1$ and $\lambda_2$ are less than one. Then: $|\delta_{v_2}| = |\delta_{w_2}| \leq \lambda_2 |\delta_{w_1}| < |\delta_{w_1}|$, $|\delta_{w_1}| \leq \lambda_1 |\delta_{v_2}| < |\delta_{v_2}|$ which is absurd. □

PROPOSITION 4.2 (Non overlapping of resonances). *Let $R_i$ for $i = 1, 2$ be $\lambda_i$-resonances which are not one subtree of the other and with non empty intersection (i.e. with at least one common vertex). Then either $\lambda_1 \geq 1/2$ or $\lambda_2 \geq 1/2$.*

PROOF. Two subtrees of degree 2 which are not one a subtree of the other can intersect in three ways, see Figure 2.

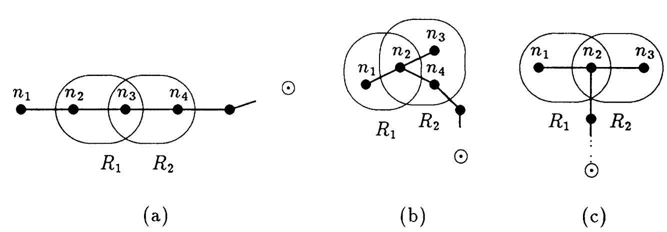



Fig. 2: Intersections of resonances

In the pictures we are using the following:

NOTATION REMARK. *Thick points correspond in general to subtrees (a subtree of degree $d$ collapses to a thick point of the same degree) and the $n$ written above a thick point correspond to the sum of $\alpha_v$ when $v$ varies in the corresponding (collapsed) subtree. In the figures, null subtrees (indicated usually by $R$) are encircled (while the root, as customary, is distinguished by a small circle around it.)*

with Siegel's method) and, instead, their contributions can be controlled only grouping together similar (in size) contributions coming from other trees *keeping track of signs and compensations*: see Lemma 5.3 and Proposition 6.1).

Since $R_i$ are resonant subtrees[9] $n_2 + n_3 = 0$, $n_3 + n_4 = 0$ in case (a) while $n_1 = -n_2$ in case (b) and (c). We proceed now by contradiction: Assume $\lambda_i < 1/2$. In case (a) $|\langle n_1\rangle| \le \lambda_1|\langle n_2\rangle|$, $|\langle n_1+n_2\rangle| \le \lambda_2|\langle n_3\rangle| = \lambda_2|\langle n_2\rangle|$ and $|\langle n_1 + n_2\rangle| \ge (1-\lambda_1)|\langle n_2\rangle| \Longrightarrow |\langle n_2\rangle| \le \dfrac{\lambda_2}{1-\lambda_1}|\langle n_2\rangle| < |\langle n_2\rangle|$ which is absurd. In case (b) and (c): $|\langle n_1\rangle| \le \lambda_1|\langle n_2\rangle| = \lambda_1|\langle n_1\rangle|$ which is absurd because $\lambda_1 < 1$. $\square$

DEFINITION 4.5 (Critical Resonance). A resonance $R$ is called *critical* (or $\lambda$-critical) if it is $\lambda$-resonant with $\lambda < 1/2$ and if it is maximal *i.e.* it is not properly contained into another $\lambda$-resonant subtree.

Obviously: (i) Critical resonances cannot intersect (by definition and by Proposition 4.2). (ii) Critical resonances may contain $\lambda'$-resonances with any $\lambda'$.

Critical resonances may appear in sequels where each resonance is a subtree of another resonance (creating "hierarchies of subresonances"). In order to classify them, we introduce the following concept.

DEFINITION 4.6 ($\lambda$-Subresonances). Let $R'$ be a null subtree of degree two of a $\lambda$-critical resonance $R$ and let $u_1v_1$, $u_2v_2$, with $v_i \in R'$, be the two edges connecting $R'$ with $R\backslash R'$ (obviously $u_i \in R$ and $u_1 \neq u_2$). Define

$$m \equiv \sum_{\substack{v\in R_{v_1}\\ v\le u_1}} \alpha_v = -\sum_{\substack{v\in R_{v_2}\\ v\le u_2}} \alpha_v \tag{4.4}$$

where $R_{v_i}$ is the rooted tree $R$ with root in $v_i$ and the order $\le$ in each sum is relative to $R_{v_i}$. The integer $m \in \mathbb{Z}^N\backslash\{0\}$ is defined up to sign (as the roles of the $u_iv_i$ can be exchanged); see Figure 3 for an example. We then say that $R'$ is a *$\lambda$-subresonance* if

$$|\langle m\rangle| \le \lambda\Delta(R'). \tag{4.5}$$

As in case of resonances we have a simple non overlapping criterion for subresonances:

PROPOSITION 4.3 (Non overlapping of subresonances). *Let $R_i'$ for $i = 1,2$ be $\lambda$-subresonances of a $\lambda$-critical resonance $R$. Assume that $R_i'$ are not one subtree of the other. Then $V(R_1') \cap V(R_2') = \emptyset$.*

The proof is very similar to the proof of Proposition 4.2 and is left to the reader.

[9] Observe that in case (a) of Figure 2, $\langle n_2\rangle = \delta_v(R_{1,u})$ for some $u \neq v$ in $R_1$; $\langle n_1\rangle = \delta_w$ with $w\bar{w}$ connecting $R_1$ with $T\backslash R_1$; etc.

DEFINITION 4.7 (Hierarchies of critical subresonances). Let $R$ be a $\lambda$-critical resonance (*i.e.* $R$ is a maximal $\lambda$-resonance with $\lambda < 1/2$). Let $R_1$ be a maximal $\lambda$-subresonance (if it exists) of $R$ (maximal means that $R_1$ is not properly contained in another $\lambda$-subresonance of $R$). Note that, by Proposition 4.3, $R_1$ cannot overlap with another $\lambda$-subresonance of $R$. We can now *define subresonances of $R_1$ replacing, in Definition 4.6, $R$ with $R_1$*. We then let $R_2$ be a maximal $\lambda$-subresonance (if it exists) of $R_1$. And so on, till $R_h$, $h \geq 1$, does not contain any $\lambda$- subresonance. We consider also the case in which $R$ does not contain any $\lambda$-subresonance setting in such a case $h = 0$ and $R_0 \equiv R$. The sequence $\mathcal{H} \equiv \mathcal{H}_\lambda(R) \equiv \{R_1, \dots, R_h\}$, $R \supset R_1 \supset \cdots \supset R_h$ is called a *critical hierarchy of $\lambda$-subresonances* (or simply a hierarchy of subresonances) and the elements of the hierarchy, $R_i$ with $1 \leq i \leq h$, are called *critical subresonances*. Thus by definition a critical subresonance of the rooted tree $T$ is an element of a hierarchy associated to some critical resonance. Obviously a critical resonance may contain more than one hierarchy of subresonances (see Figure 5 below).

REMARK 4.2. In general a critical $\lambda$-subresonance need not be a $\lambda$-resonance (as shown by the following example) even though *it is always a $\lambda'$-resonance with* $\lambda' \equiv \dfrac{\lambda}{1-\lambda}$: see (4.6).

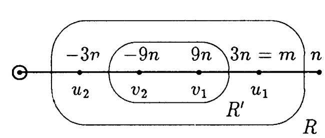



Fig. 3: A subresonance $R'$

Fix $\lambda = 1/3$ and let $n \in \mathbb{Z}^N \backslash \{0\}$. Then one checks immediately that $R$, in Figure 3, is a $\lambda$-critical resonance and that $\mathcal{H}_\lambda(R) = \{R'\}$. However the critical subresonance $R'$ *is not a $\lambda$-resonance* since $|\langle 4n \rangle| > \lambda\Delta(R') = 3|\langle n \rangle|$ (but it is a $\dfrac{1}{2}$-resonance).

REMARK 4.3. If $R'$ is a $\lambda$-critical subresonance then $R' = R_i$ for some $1 \leq i \leq h$ where $\{R_1, \dots, R_h\}$ is a $\lambda$-hierarchy associated to some $\lambda$-critical resonance $R$. To each $R_j$ we can associated (up to sign) an integer $m_j \in \mathbb{Z}^N \backslash \{0\}$ as in (4.4) (see Figure 4).

Then $|\langle m_j \rangle| \leq \lambda\Delta(R_j)$ for any $j = 1, \dots, h$ and in particular $|\langle m_j \rangle| \leq \lambda |\langle m_{j+1} \rangle|$ so that $|\langle m_i \rangle| \leq \lambda^{j-i} |\langle m_j \rangle|$, $\forall 1 \leq i \leq j \leq h$ and

$$\sum_{i=1}^{j} |\langle m_i \rangle| < \frac{1}{1-\lambda} |\langle m_j \rangle| \leq \frac{\lambda}{1-\lambda} \Delta(R_j).$$

Also, if $\sigma_h$ is plus or minus one and $\sigma_i = 0$ or $\pm 1$ $(i = 1, \dots, h-1)$ then

$$\left| \left\langle \sum_{i=1}^{h} \sigma_i m_i \right\rangle \right| \geq \frac{1-2\lambda}{1-\lambda} \, |\langle m_h \rangle|$$

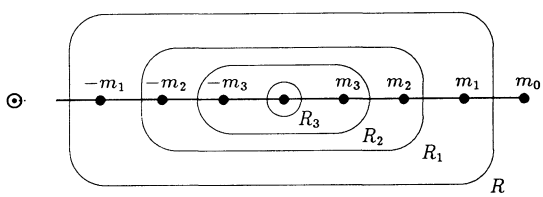



Fig. 4: A hierarchy of resonances with $h = 3$

Furthermore, let[10] $m_0 \equiv \sum_{v'<R} \alpha_{v'}$. If we call $w$ the first vertex following $R_h$, then one has $\delta_w = \left\langle m_0 + \sum_{i=1}^{h} \sigma_i m_i \right\rangle$ and (since $|\langle m_0 \rangle| \leq \lambda |\langle m_1 \rangle|$):

$$(4.6) \qquad |\delta_w| \leq \sum_{i=0}^{h} |\langle m_i \rangle| < \frac{1}{1-\lambda} |\langle m_h \rangle| \leq \frac{\lambda}{1-\lambda} \Delta(R_h)$$

The same bound holds if $h$ is replaced by any $j$; thus *each subresonance $R_j$ is a $\frac{\lambda}{1-\lambda}$-resonance.*

We shall often use the word "(sub)resonance" to indicate either a resonance or a subresonance.

DEFINITION 4.8 (The set of critical (sub)resonance of $T$). Given a rooted tree $T$, an admissible $\alpha$ and a $\lambda < 1/2$ we let $\mathcal{R} \equiv \mathcal{R}(T)$ be the set of all $\lambda$-critical resonances and $\lambda$-critical subresonances of $T$ and $\mathcal{R}_0 \equiv \mathcal{R}_0(T)$ be the set of all $\lambda$-critical resonances of $T$. If $2 \leq p \equiv |\mathcal{R}_0| \equiv$ cardinality of $\mathcal{R}_0$ then $R^i \cap R^j = \emptyset$ for all $R^i$, $R^j \in \mathcal{R}_0$ with $i \neq j$. To each element $R^i$ of $\mathcal{R}_0$ we can associate an integer $h_i$ which is zero if $R^i$ does not contain $\lambda$-critical subresonances, otherwise $h_i \equiv \max_{\mathcal{H}_\lambda \subset R^i} |\mathcal{H}_\lambda|$, where $|A|$ denotes the cardinality of the set $A$. See Figure 5 for an example of critical resonances and subresonances[11].

[10] If $u > \bar{u} \geq \bar{w} > w$ and $u\bar{u}$, $w\bar{w}$ connect $R$ with $T \backslash R$, $v < R$ means $v < \bar{w}$.

[11] In Figure 5 $\mathcal{R}_0 = \{R^1, R^2\}$ and $R_3, \dots, R_9$ are critical subresonances; notice that $R^1$ contains five hierarchies, three of which have $h=1$.

DEFINITION 4.9 (Coresonances). To any $\lambda$-critical (sub)resonance $R$ in $\mathcal{R}$ we associate *its coresonance* $\overline{R} \equiv R\backslash\{$all $\lambda$-critical subresonances contained in $R\}$. If $R$ is minimal (*i.e.* does not contain any element of $\mathcal{R}$) then $\overline{R} = R$ otherwise $\overline{R}$ is the disjoint union of two or more subtrees of $T$. In any case, for any coresonance, $\overline{R}$, we always have $\alpha(\overline{R}) \equiv \sum_{v\in\overline{R}} \alpha_v = 0$. We denote $\overline{\mathcal{R}} \equiv \{\overline{R} : R \in \mathcal{R}\}$ the set of all coresonances. Finally, we let $\overline{T} \equiv T\backslash \left( \bigcup_{R\in\mathcal{R}_0} R \right)$ so that $\overline{T} = T$ if $\mathcal{R}_0 = \emptyset = \mathcal{R}$ and otherwise $\overline{T}$ is the disjoint union of one rooted tree (or only of the earth $\eta$ if $r \in R \in \mathcal{R}_0$) and of one or more subtrees.

REMARK 4.4. $R \in \mathcal{R}$ is connected to $T\backslash R$ by two edges $u\overline{u}$, $v\overline{v}$ with $u$, $v \in T\backslash R$ ($u \neq v$ and either $u$ or $v$ may be the earth $\eta$) and with $\overline{u}$, $\overline{v} \in \overline{R}$ ($\overline{u}$, $\overline{v}$ may coincide).

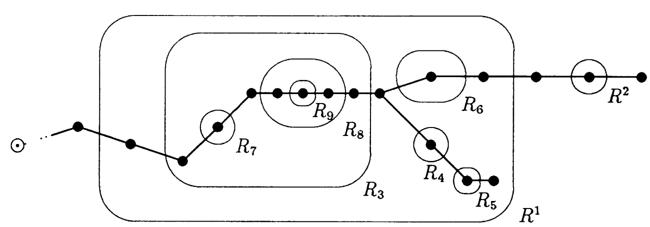



Fig. 5: Critical (sub)resonances

DEFINITION 4.10 (Subtree contractions). Let $T$ be a (possibly rooted) tree and $R \subset T$ a subtree of degree two. We define *the contraction of $T$ over $R$* denoted $T/R$ as follows. Let $u\overline{u}$, $v\overline{v}$, ($\overline{u}, \overline{v} \in R$), be the two edges connecting $R$ with $T\backslash R$; we then set

$$T/R \equiv T\backslash R + uv$$

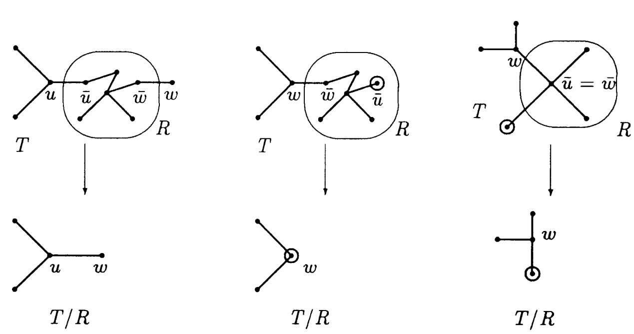



Fig. 6: Tree contractions; (in the second example $u = \eta$)

Coresonances $\overline{R}$ are crucial for constructing "complete families": they will be the invariants of all the trees (with a common $\alpha$ function) belonging to the same complete family. We describe with more detail the set $\overline{\mathcal{R}}$.

DEFINITION 4.11 (Critical chains and connecting points). *A $(\lambda)$-critical chain $C$ is an ordered s-tuple of coresonances*, $(\overline{R}_1, \ldots, \overline{R}_s)$, such that $\overline{R}_i$ is connected by one edge with $\overline{R}_{i+1}$ for all $1 \leq i \leq s-1$ and such that $\overline{R}_1$ and $\overline{R}_s$ are also connected by one edge with either $\overline{T}$ or with a coresonance $\overline{R} \in \overline{\mathcal{R}}$ which contains the whole chain $C$ (*i.e.* $R_1 \cup \ldots \cup R_s \subset R$). The case $s = 1$, *trivial chain*, is admitted: in such a case $C = \overline{R}_1$ is connected with either $\overline{T}$ or with a coresonance $\overline{R}$ such that $R_1 \subset R$. Given a chain $C$, $\overline{R} \in C$ means that $C \equiv (\overline{R}_1, \ldots, \overline{R}_s)$ and $\overline{R} = \overline{R}_i$ for some $i$; analogously $v \in C$ means that $v$ is a vertex $\in \overline{R}$ for some $\overline{R} \in C$. With this convention we see that $\mathcal{C} \equiv \{$the set of all chains$\} = \overline{\mathcal{R}}$.

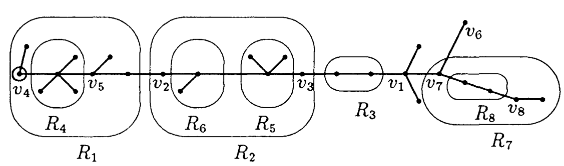



Fig. 7: Critical chains $\mathcal{C} \equiv \{(\overline{R}_1, \overline{R}_2, R_3), (R_5, R_6), R_4, \overline{R}_7, R_8\}$

Given a chain $C \equiv (\overline{R}_1, \ldots, \overline{R}_s)$ we let $u \equiv u(\overline{R}_1) \equiv u(C)$, $w \equiv w(\overline{R}_s) \equiv w(C)$ be the unique points outside $C$ such that $u\overline{u}$ and $w\overline{w}$ are the edges connecting, respectively, $\overline{R}_1$ and $\overline{R}_s$ with $T \backslash C$: thus $\overline{u} \in \overline{R}_1$, $\overline{w} \in \overline{R}_s$, $u$ and $w$ are distinct (and one of them may be the earth $\eta$). The points $u$, $w$ will be called *the base points of the chain $C$*.

In the example shown in Figure 7, if $C_1 \equiv (\overline{R}_1, \overline{R}_2, R_3)$ and $C_2 \equiv (R_5, R_6)$ we have the connecting points: $u(C_1) = \eta$, $w(C_1) = v_1$, $u(C_2) = v_3$, $w(C_2) = v_2$ and $u(R_4) = v_5$, $w(R_4) = v_4$, $u(\overline{R}_7) = v_6$, $w(\overline{R}_7) = v_1$, $u(R_8) = v_7$, $w(R_8) = v_8$.

We denote by $U \equiv U(T) \equiv \{u(C), w(C) : C \in \mathcal{C}\}$ the *set of base points of $T$*.

REMARK 4.5. The order of a chain $(\overline{R}_1, \ldots, \overline{R}_s)$ is related to the order in $T$ as follows: *either* $\overline{R}_1 > \cdots > \overline{R}_s$ (as in the chain $C_1$ of the above example) *or* $\overline{R}_1 < \cdots < \overline{R}_s$ (as in $C_2$).

Finally, we are ready to define *complete families of trees generated by a tree $T$ and a given $T$-admissible function $\alpha$.*

DEFINITION 4.12 (Complete families). Given $T \in \mathcal{T}^k$, $\alpha$ $T$-admissible and $\lambda < 1/2$ we define $\mathcal{F}(T) \equiv \{T\}$, if $\mathcal{R}(T) = \emptyset$ (*i.e.* if there are no $\lambda$-critical resonances in $T$), otherwise, letting $\mathcal{R} \equiv \mathcal{R}(T)$, $\mathcal{C} \equiv \mathcal{C}(T)$, we define

$$\mathcal{F}(T) \equiv \Big\{ T' \in \mathcal{T}^k : T' = \overline{T} \cup \bigcup_{\overline{R} \in \overline{\mathcal{R}}} \overline{R} + \sum_{C=(\overline{R}_1,\ldots,\overline{R}_s)\in\mathcal{C}} u(C)\overline{u}_1 + w(C)\overline{w}_s + \sum_{i=1}^{s-1} \overline{w}_i \overline{u}_{i+1}, \ \forall\, \overline{u}_i, \overline{w}_i \in \overline{R}_i \Big\} \tag{4.7}$$

if $s = 1$ for some $C$ the sum over $i$ has to be omitted. *$\mathcal{F}$ is called the complete family associated to $(T, \alpha)$.*

Obviously, the integers $n_i \equiv \alpha_{v_i}$, which together with the "topological" structure of the tree $T$ generates $\mathcal{F}$, may be thought of as fixed "attributes" of the labels $v_i$ and, *by construction, $\{\alpha_{v_i}\}$ is admissible for any tree $T' \in \mathcal{F}(T)$.* The following Proposition collects a few elementary properties of complete families.

PROPOSITION 4.4 (Properties of $\mathcal{F}$). *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function and $\lambda < 1/2$. Then:* (i) $\mathcal{F}(T) = \{T\} \iff \mathcal{R}(T) = \emptyset$; (ii) $\mathcal{F}(T') = \mathcal{F}(T)$, $\forall T' \in \mathcal{F}(T)$; (iii) $\mathcal{F}(T) \cap \mathcal{F}(T') \neq \emptyset \implies T' \in \mathcal{F}(T)$; (iv) $\overline{T}$, $\mathcal{R}(T)$, $U(T)$ *is a complete and minimal set of invariants for* $\mathcal{F}(T)$; (v) $|\mathcal{F}| = \prod_{\overline{R}\in\overline{\mathcal{R}}} |V(\overline{R})|^2$; (vi) $\deg_{T'} v \le \deg_{T''} v + 2$, $\forall T', T'' \in \mathcal{F}(T)$.

The proof of the Proposition is a simple consequence of the various definitions.

REMARK 4.6. From (i), (ii), (iii) above it follows immediately that the property $T' \in \mathcal{F}(T)$ is an *equivalence relation*. Thus the set of all labeled rooted trees, given $\{n_i\} \equiv \{\alpha_{v_i}\}$, can be *partitioned into disjoint complete families*: here we have preassigned the function $\alpha_{v_i} \equiv n_i$ and we use the convention that if $\alpha$ is not admissible for some $T$ we are omitting the (meaningless) contribution coming from that tree.

A final comment: the set of labels $\{v_1, \ldots, v_k\}$ of $\mathcal{T}^k$ can be thought of as the set of vertices $V$ of $\mathcal{F}$ and for any $v \in V$ we set

$$\deg_{\mathcal{F}} v = \max_{T' \in \mathcal{F}} \deg_{T'} v \tag{4.8}$$

and from (vi) of Proposition 4.4 it follows immediately that:

$$\deg_{\mathcal{F}} v - 2 \le \deg_{T'} v \le \deg_{\mathcal{F}} v, \quad \forall T' \in \mathcal{F}. \tag{4.9}$$

In Appendix D we show two examples of complete families associated to order 5 trees with two resonances.

## 5. - Estimates

In this section we formulate the estimates on (sums) of products of small divisors needed to prove Kolmogorov's (KAM) theorem following Siegel's strategy. The first two lemmas are a simple generalization of Siegel's argument and deal with situations without resonances (or better with non critical resonances). The first lemma is taken from [11] and we include for completeness its proof in Appendix C; estimates of products of "small divisors" for trees with no resonances have been also considered in [6]. The second lemma extends Siegel's argument so as to cover the case of trees with no $\lambda$-resonances (but which may contain Eliasson resonances *i.e.* resonances which are called critical in [11][12]). The third and fourth lemmas are the crucial point of our paper: it is shown that sums over complete families obey the same bounds that hold for products of small divisors without resonances (*i.e.* without repetitions). In other words, the "divergent terms" (whose actual occurrence is discussed in Appendix B) compensate (and in fact *they do not cancel exactly*) within complete families. The four lemmas will easily yield Kolmogorov's theorem for the case under consideration (3.2); see Theorem 5.1 below. The proofs are presented in the next section (except for the proof of Lemma 5.1 which is given in Appendix C). In formulating the results we shall make certain requirements on the parameter $\lambda$ which cannot be guessed by simply looking at the explicit form of the various constants appearing in the estimates: such requirements are needed in order to perform inductive proofs (often based on contractions of resonant subtrees: see, in particular, Subsection 6.2).

Recall (2.7) and the definition of admissible functions (3.6).

LEMMA 5.1 (Siegel, Eliasson). *Let $T \in \mathcal{T}^k$ be a rooted tree of order*[13] *$k$, $\alpha$ a $T$-admissible function and $0 < \lambda < 1$. Assume that all resonances (if any) are Siegel, (recall* (iii) *of Remark* 4.1*). Then there exist constants $c_1 > \gamma$, $\beta_1 > \tau$ such that*

$$\prod_{v \in T} |\delta_v|^{-1} \leq c_1^k \prod_{v \in T} |\alpha_v|^{\beta_1}. \tag{5.1}$$

*The constants $c_1$, $\beta_1$ can be taken to be*

$$c_1 = \gamma 2^{6\tau} \frac{1+\lambda}{\lambda}, \quad \beta_1 = 3\tau \tag{5.2}$$

The next Lemma extends the above type of estimate to products of small divisors arising in trees with (possible) resonant subtrees which are non $\lambda$-resonant (for some prefixed $\lambda$); recall that such trees may contain couples of points which violate the main hypothesis in Lemma 5.1 *i.e.* they may contain either short resonances or Eliasson resonances (which are not $\lambda$-resonant, see Remark 4.1)

[12] Recall the definitions given in (iii) of Remark 4.1.

[13] *i.e.* $k$= cardinality of $V(T)$.

so that the result described in the next Lemma is a genuine generalization of the estimates described in Lemma 5.1.

LEMMA 5.2. *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function and $0 < \lambda \leq 2 - \sqrt{3}$. Assume that there are no $\lambda$-resonances (Definition 4.4). Then there exist constants $c_2 > \gamma$ and $\beta_2 > \tau$ such that*

$$\prod_{v \in T} |\delta_v|^{-1} \leq c_2^k \prod_{v \in T} |\alpha_v|^{\beta_2}. \tag{5.3}$$

*The constants $c_2$, $\beta_2$ can be taken to be*

$$c_2 = c_1 2^{2\tau} \frac{1-\lambda}{1-2\lambda}, \quad \beta_2 = \beta_1 + 2\tau. \tag{5.4}$$

LEMMA 5.3 (Small divisor compensations). *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function, $0 < \lambda \leq 1/2$ and let $0 < \overline{\lambda} < 1$. Let $R$ be a minimal $\lambda$-(sub)resonance of $T$ (i.e. $R$ does not contain any $\lambda$-subresonance, see Definition 4.6) and let $n \in \mathbb{Z}^N \backslash \{0\}$ be such that $|\langle n \rangle| \leq \overline{\lambda}\Delta(R)$. Let $z$ be an extra vertex (not in $T$) and set $\alpha_z = n$. Finally, for any $u$, $w \in R$, denote by $R_u^w$ the tree[14] $R \cup \{z\} + wz$ rooted at $u$. Then there exist constants $c_3 > \gamma^2$, $\beta_3 > 2\tau$ such that for all $1 \leq i$, $j \leq N$*

$$\left| \sum_{u,w \in R} \alpha_{ui} \alpha_{wj} \prod_{v \in R} \delta_v(R_u^w)^{-2} \right| \leq c_3^{|V(R)|} \prod_{v \in R} |\alpha_v|^{\beta_3} \tag{5.5}$$

*where $\alpha_{ui}$ (respectively $\alpha_{uj}$) denote the $i^{\text{th}}$ (respectively the $j^{\text{th}}$) component of the integer vector $\alpha_u$. The constants $c_3$, $\beta_3$ can be taken to be*

$$c_3 = (c_2 2^{\tau+1}(1-\overline{\lambda})^{-1})^2, \quad \beta_3 = 2(\beta_2 + \tau + 1). \tag{5.6}$$

Estimate (5.5) easily leads to control the contributions to (3.7) coming from complete families.

LEMMA 5.4. *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function, $0 < \lambda \leq 1/4$. Let $\mathcal{F} \equiv \mathcal{F}(T)$ be the complete family defined in Definition 4.12, let $V \equiv V(T) = V(\mathcal{F})$ and recall the definition of $\deg_{\mathcal{F}}$ given in (4.8). Then there exist constants $c_4 > \gamma^2$, $\beta_4 > 2\tau$ such that*

$$\left| \sum_{T' \in \mathcal{F}} \prod_{vv' \in E(T')} \alpha_v \cdot \alpha_{v'} \prod_V \delta_v^{-2} \right| \leq \prod_{v \in V} |\alpha_v|^{\deg_{\mathcal{F}} v} c_4^{|V|} \prod_{v \in V} |\alpha_v|^{\beta_4}. \tag{5.7}$$

*The constants $c_4$, $\beta_4$ can be taken to be*

$$c_4 \equiv \left( c_2 2^{\tau+1} \frac{1-\lambda}{1-2\lambda} \right)^2, \quad \beta_4 = \beta_3. \tag{5.8}$$

[14] If $R$ is rooted, $R = R_r$, first remove the edge $\eta r$ so as to obtain an unrooted subtree.

Kolmogorov's theorem now easily follows. We introduce the following norms on analytic functions $g : \mathbb{T}^N \to \mathbb{C} : \|g\|_\sigma \equiv \sum_{n\in\mathbb{Z}^N} |g_n| e^{|n|\sigma}$. Clearly, for any analytic function on $\mathbb{T}^N$ there exist a $\sigma > 0$ such that the above norm is finite; viceversa if $g$ is such that $\|g\|_\sigma < \infty$ for some $\sigma > 0$ then $|g_n| \le \|g\|_\sigma \exp(-|n|\sigma)$ so that $g$ is analytic. If $g = (g_1, \dots, g_M) : \mathbb{T}^N \to \mathbb{C}^M$ we set $\|g\|_\sigma \equiv \sup_{1\le j\le M} \|g_j\|_\sigma$.

THEOREM 5.1. *Let $f$ in* (3.2) *be real-analytic with $\|f\|_{\overline{\sigma}} < \infty$ for some $\overline{\sigma} > 0$. Then the formal solution $X(\theta, \varepsilon)$ described in Proposition* 2.1 *is real-analytic. Furthermore, fix $0 < \sigma < \overline{\sigma}$, $0 < \lambda < 1/2$ and set*

$$\varepsilon_0 \equiv \frac{(\overline{\sigma} - \sigma)^{\beta_5}}{c_5 \|f\|_{\overline{\sigma}}} \quad \text{with} \quad \begin{cases} \beta_5 \equiv \beta_4 + 4 \\ c_5 \equiv c_4 2^{\beta_5+1} \beta_5! \end{cases} \tag{5.9}$$

*where the constants $\beta_4$, $c_4$ are defined above. Then $X(\theta, \varepsilon)$ is jointly analytic in the domain $\{|\Im(\theta_i)| \le \sigma\} \times \{|\varepsilon| \le \varepsilon_0\}$ and for any (complex) $\varepsilon$ with $|\varepsilon| < \varepsilon_0$ the following bound holds*

$$\|X(\,\cdot\,, \varepsilon)\|_\sigma \le \frac{\overline{\sigma} - \sigma}{2} \frac{|\varepsilon|}{\varepsilon_0 - |\varepsilon|} \tag{5.10}$$

REMARK 5.1. Choosing $\lambda = 1/5$ (so as to minimize the constant $c_5$) one obtains

$$\beta_5 = 12\tau + 6, \quad c_5 < \gamma^2 2^{30\tau+16} (12\tau + 6)!$$

Such constants are certainly not optimal.

## 6. - Proofs

In the following five Subsections we give the proofs of Lemmas 5.2, 5.3, 5.4 and of the Theorem of Section 5. In the first subsection we show how Kolmogorov's Theorem can be easily derived from Lemma 5.4. In the second subsection we discuss the hypotheses under which certain resonant subtrees can be contracted without generating, in the contracted tree, new $\lambda$-resonances: this analysis allow to perform inductive proofs of Lemma 5.2 and 5.4. A proof of Lemma 5.1 is given, for sake of completeness, in Appendix C.

### 6.1 Proof of Theorem 5.1

Recall the form of the coefficients $X_{jn}^{(k)}$ given in (3.9) $\div$ (3.10). By Remark 4.6 we know that, for a given choice of $\{n_1, \dots, n_k\}$ with $n_i \in \mathbb{Z}^N \backslash \{0\}$, the labeled rooted trees in $\mathcal{T}^k$ can be partitioned into complete families having as common admissible $\alpha$ function $\alpha_{v_i} = n_i$. Denote by $\mathcal{N} \equiv \mathcal{N}(n_1, \dots, n_k)$ the set of all such complete families. Then, by (3.9) ÷ (3.10), by Lemma 5.4 and by (4.8), (4.9) we see that, for any $0 < \sigma < \overline{\sigma}$,

$$
\begin{aligned}
&\sum_{n\in\mathbb{Z}^N\setminus\{0\}} e^{\sigma|n|}|X_{jn}^{(k)}| \\
&= \sum_n e^{\sigma|n|}\frac{1}{k!}\left|\sum_{\substack{n_1,\dots,n_k\\ n_i\neq 0,\Sigma n_i=n}} \sum_{\mathcal{F}\in\mathcal{N}}\prod_{i=1}^k f_{n_i}\sum_{T\in\mathcal{F}}\prod_{vv'\in E(T)}\alpha_v\cdot\alpha_{v'}\prod_V \delta_v^{-2}\right| \\
&\le \sum_n \frac{1}{k!}\sum_{\substack{n_1,\dots,n_k\\ n_i\neq 0,\Sigma n_i=n}}\sum_{\mathcal{F}\in\mathcal{N}}\prod_{i=1}^k (f_{n_i}e^{\sigma|n_i|})\prod_{v\in V}|\alpha_v|^{\deg_{\mathcal{F}} v}c_4^k\prod_V|\alpha_v|^{\beta_4} \\
&\le \frac{1}{k!}\sum_{T\in\mathcal{T}^k}\sum_{\alpha:T\to\mathbb{Z}^N}\prod_V (c_4|f_{\alpha_v}|\,|\alpha_v|^{\beta_4+2}e^{\sigma|\alpha_v|})\prod_V|\alpha_v|^{\deg v}.
\end{aligned}
\tag{6.1}
$$

It is easy to check (see Proposition B.1) that, if we let $\phi_n \equiv c_4|f_n|e^{\sigma|n|}|n|^{\beta_4+2}$ then, last line of (6.1) *is exactly the* $k^{\text{th}}$ *coefficient,* $g_k$*, of the formal solution* $g$ *of*

$$g \sim \varepsilon\sum_{n\in\mathbb{Z}^N}\phi_n|n|\exp(|n|g). \tag{6.2}$$

But since (by the analyticity of $f$) the nonnegative numbers $\phi_n$ decay exponentially fast as $|n|\to\infty$, the function of two complex variables

$$G(x,\varepsilon) \equiv x - \varepsilon\sum_{n\in\mathbb{Z}^N}\phi_n|n|e^{|n|x} \tag{6.3}$$

is holomorphic and bounded in the complex 2-disk

$$D \equiv \{x\in\mathbb{C} : |x|\le s\}\times\{\varepsilon\in\mathbb{C} : |\varepsilon|\le\varepsilon_0\} \tag{6.4}$$

for any $\varepsilon_0 > 0$ and any $0 < s < \overline{\sigma}-\sigma$. In fact, recalling that $\beta_5\equiv\beta_4+4$ we find

$$
\begin{aligned}
\sup_D |G| < \sup_D |G_x| &= \varepsilon_0\sum_{n\in\mathbb{Z}^N}\phi_n|n|^2e^{\sigma|n|} \\
&\le \varepsilon_0 c_4(\overline{\sigma}-\sigma-s)^{-\beta_5}\beta_5!\|f\|_{\overline{\sigma}}.
\end{aligned}
\tag{6.5}
$$

Taking $s\equiv(\overline{\sigma}-\sigma)/2$ and $\varepsilon_0$ as in (5.9) we see that $\sup_D |G_x|\le 1/2$, and, since $G(0,0)=0$, by the complex Implicit Function Theorem, we conclude that there exists a unique analytic function $g(\varepsilon)$ such that

$$G(g(\varepsilon),\varepsilon)=0,\qquad \sup_{\substack{\varepsilon\in\mathbb{C}\\ |\varepsilon|\le\varepsilon_0}}|g(\varepsilon)|\le s=\frac{\overline{\sigma}-\sigma}{2}. \tag{6.6}$$

Thus $g_k \leq \frac{1}{2}(\overline{\sigma} - \sigma)\varepsilon_0^{-k}$ and therefore $\|X^{(k)}\|_\sigma \leq \frac{1}{2}(\overline{\sigma}\sigma)\varepsilon_0^{-k}$ and, for any complex $\varepsilon$ with $|\varepsilon| < \varepsilon_0$

$$\|X(\cdot\,,\varepsilon)\|_\sigma \leq \frac{\overline{\sigma} - \sigma}{2} \frac{|\varepsilon|/|\varepsilon_0|}{1 - |\varepsilon|/|\varepsilon_0|} \tag{6.7}$$

completing the proof of Theorem 5.1. □

## 6.2 Contractions

Here we discuss the circumstances under which it is possible to contract certain resonant subtrees without generating new $\lambda$-(sub)resonances. The reason why one wants to contract resonances is simply that, if $R$ is a resonance of $T$,

$$\prod_T |\delta_v|^{-1} = \prod_{T/R} |\delta_v(T/R)|^{-1} \prod_R |\delta_v|^{-1} \tag{6.8}$$

so that one can reduce the order of the tree (passing from $T$ to $T/R$).

We start with a "negative" example *i.e.* an example of "forbidden" contraction. Consider Figure 8: it easy to see that, given any $\lambda' \ll \lambda \leq 1/2$, if $m$, $n$ and $p$ are suitably chosen, then: $R$ is an Eliasson resonance[15] (with parameter $\lambda$) but not $\lambda$-resonant; $R'$ is a short resonance not $\lambda$-resonant; $R/R'$ is $\lambda'$-resonant (hence, in particular, is a $\lambda$-resonance)[16].

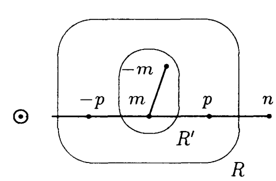



Fig. 8: A forbidden contraction if $|\langle m\rangle| \ll |\langle n\rangle| \ll |\langle p\rangle|$

The first simple result says that Siegel resonances cannot become $\lambda$-resonant by contraction of Eliasson or short resonances:

LEMMA 6.1. *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function, $0 < \lambda \leq 1/2$. If $R' \subset R$ are resonances and $R'$ is either Eliasson or short and $R$ is Siegel, then $R/R'$ is not $\lambda$-resonant.*

[15] Recall (iii) of Remark 4.1.

[16] The parameters can be chosen as follows: fix $0<\lambda'\leq\frac{\lambda}{1+\lambda}\leq\lambda\leq 1/2$; fix $p\in\mathbb{Z}^N\setminus\{0\}$ and choose $m=n$ such that $0<|\langle n\rangle|\leq\lambda'|\langle p\rangle|$. Then one checks immediately that $|\langle n\rangle|\leq\lambda|\langle n+p\rangle|$ (*i.e.* $R$ is Eliasson); $|\langle n\rangle|>\lambda|\langle m\rangle|$ (*i.e.* $R$ is not $\lambda$-resonant); $|\langle n+p\rangle|>\lambda|\langle m\rangle|$ (*i.e.* $R'$ is not $\lambda$-resonant); but $|\langle n\rangle|\leq\lambda'|\langle p\rangle|$ (*i.e.* $R/R'$ is $\lambda'$-resonant).

PROOF. Let $R' = R(\overline{u}', w')$ and $R = R(\overline{u}, w)$. The statement is obvious if $R'$ is short or if $R'$ does not intersect the path $P(\overline{u}, w)$. Assume therefore that $R'$ is Eliasson (with parameter $\lambda$) with $\overline{u} > \overline{u}' > w' > w$ and assume also (by contradiction) that $R/R'$ is $\lambda$-resonant. $R$ Siegel implies that there exists $\overline{u} > \overline{v} > w$ such that $|\delta_w| > \lambda|\delta_{\overline{v}}|$; in fact $\overline{u}' > \overline{v} > w'$ (otherwise $R/R'$ would clearly be Siegel). Since $R'$ is Eliasson, $|\delta_{w'}| \leq \lambda|\delta_{\overline{v}}|$, so that $|\delta_{w'}| < |\delta_w|$. But $\delta_{w'} = \delta_{w'}(R_{\overline{u}}) + \delta_w$ and, since $|\delta_w| \leq \lambda|\delta_{w'}(R_{\overline{u}})|$ (having assumed that $R/R'$ is $\lambda$-resonant), it is $|\delta_{w'}| \geq (1-\lambda)|\delta_{w'}(R_{\overline{u}'})|$. Putting these bounds together one obtains $\lambda|\delta_{w'}(R_{\overline{u}})| \geq |\delta_w| > |\delta_{w'}| \geq (1-\lambda)|\delta_{w'}(R_{\overline{u}})|$ which is a contradiction since $\lambda \leq 1/2$. □

The next lemma shows that (sub)resonances can always be contracted without introducing new critical (sub)resonances, provided $\lambda \leq 1/4$.

LEMMA 6.2. *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function, $0 < \lambda \leq 1/4$. Let $R_*$ be a minimal critical (sub)resonance then $\overline{\mathcal{R}}(T/R_*) = \overline{\mathcal{R}}(T)\backslash\{R_*\}$.*

Such a result is an immediate corollary of the following

SUBLEMMA 6.1. *Let $\lambda \leq 1/3$, let $R'$ be a $\lambda$-resonance and let $R \supset R'$ be a resonance which is non $\lambda$-resonant. Then $R/R'$ is not $\lambda$-resonant.*

PROOF (of Lemma 6.2 given Sublemma 6.1). If $R_*$ is a critical resonance, the claim follows directly from Sublemma 6.1 (as critical resonances are maximal $\lambda$-resonances). If $R_*$ is a subresonance then it is a $\lambda'$-resonance with $\lambda' = \lambda/(1-\lambda)$ (recall Remark 4.1). Let $R \supset R_*$ be a resonance with $R \notin \mathcal{R}(T)$. Then by the lemma (since $\lambda \leq 1/4$ is equivalent to $\lambda' \leq 1/3$), $R/R_*$ cannot be a $\lambda'$-resonance and hence cannot be a critical (sub)resonance of $T$. □

PROOF (of Sublemma 6.1). Let $R = R(u, z)$ and $\overline{u}u'$, $w'\overline{w}$ the edges connecting $R'$ with $\overline{R} \equiv R/R'$ and $u', w' \in R'$, $u' \geq w'$, $\overline{u} > \overline{w}$. $R$ not $\lambda$-resonant means that $\exists v_1 \in R$, $v_1 \neq u$ such that

$$|\delta_z| > \lambda|\delta_{v_1}(R_u)|. \tag{6.9}$$

If $v_1 \notin R'$ (or $v_1 = u'$ in which case $\delta_{v_1}(R_u) = \delta_{\overline{w}}(R_u)$ and $\overline{w} \notin R'$) then the claim is obvious. Assume that $u' \neq v_1 \in R'$ and let us proceed by contradiction, assuming that $\overline{R}$ is $\lambda$-resonant. Because $\overline{R}$ and $R'$ are $\lambda$-resonant, one has

$$|\delta_z| \leq \lambda|\delta_{\overline{w}}(\overline{R}_u)|, \quad |\delta_{\overline{w}}| \leq \lambda|\delta_{v_1}(R'_{u'})|. \tag{6.10}$$

Observe also that $\exists\sigma, \sigma' \in \{0, 1\}$ such that

$$\delta_{v_1}(R_u) = \delta_{v_1}(R'_{u'}) + \sigma\,\delta_{\overline{w}}(\overline{R}_u), \quad \delta_{\overline{w}} = \delta_{\overline{w}}(\overline{R}_u) + \sigma'\delta_z \tag{6.11}$$

($\sigma = 1$ if $v_1$ is on the path $P(u', w')$ and $\sigma = 0$ otherwise; $\sigma' = 1$ if $\overline{w} \geq z$ and $\sigma' = 0$ otherwise). Now, using (6.10) and (6.11) we get

$$\lambda|\delta_{v_1}(R'_{u'})| \geq |\delta_{\overline{w}}| \geq (1-\lambda)|\delta_{\overline{w}}(\overline{R}_u)| \tag{6.12}$$

whence

$$|\delta_{v_1}(R_u)| \geq |\delta_{v_1}(R'_{u'})| - |\delta_{\overline{w}}(\overline{R}_{v_2})| \geq \frac{1-2\lambda}{1-\lambda}\,|\delta_{v_1}(R'_{u'})|. \tag{6.13}$$

Putting (6.12), (6.10), (6.13) and (6.9) together we get

$$\frac{\lambda^2}{1-\lambda}\,|\delta_{v_1}(R'_{u'})| \geq |\delta_z| > \lambda\,\frac{1-2\lambda}{1-\lambda}\,|\delta_{v_1}(R'_{u'})| \tag{6.14}$$

which implies $\lambda > 1/3$ which is a contradiction. □

Let now $R'$ be Eliasson (with parameter $\lambda$) and $R' \subset R$ with $R$ not $\lambda$-resonant. With the same notations used in the above proof, we see that, if $u' > v_1 \geq w'$, in place of the second of (6.10), it holds $|\delta_{\overline{w}}| \leq \frac{\lambda}{1-\lambda}\,|\delta_{v_1}(R'_{u'})|$. Mimicking the rest of the estimates leading to (6.14), we obtain

$$\frac{\lambda^2}{(1-\lambda)^2}\,|\delta_{v_1}(R'_{u'})| \geq |\delta_z| > \lambda\,\frac{1+\lambda^2-3\lambda}{(1-\lambda)^2}\,|\delta_{v_1}(R'_{u'})| \tag{6.15}$$

which implies $\lambda > 2-\sqrt{3} = 0.26\ldots$. We have thus proved the following statement:

SUBLEMMA 6.2. *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function, $0 < \lambda \leq 2-\sqrt{3}$. Let $R'$ be an Eliasson resonance (with parameter $\lambda$) and let $R \equiv R(u,z) \supset R'$ be such that $|\delta_z| > \lambda|\delta_{v_1}(R_u)|$ for some $u > v_1 > z$. Then $R/R'$ is not $\lambda$-resonant.*

We are now ready to give a criterion which allows to contract resonances which are not critical (sub)resonances.

LEMMA 6.3. *Let $T \in \mathcal{T}^k$, $\alpha$ a $T$-admissible function, $0 < \lambda \leq 2-\sqrt{3}$, $h \geq 2$. Let $R_1, \ldots, R_h$, $R'_h$ be resonances such that $R_h \cap R'_h = \emptyset$; $R_h \cup R'_h \subset R_{h-1} \subset \cdots \subset R_1$; $R_h$ and $R'_h$ are either Eliasson or short; $R_i$, for $1 \leq i \leq h-1$, are not $\lambda$-resonant. Then one can choose $R_*$ equal either to $R_h$ or to $R'_h$ so that, for $1 \leq i \leq h-1$, $R_i/R_*$ is not $\lambda$-resonant.*

PROOF. Let, for $1 \leq i \leq h$, $R_i = R(\overline{u}_i, w_{i-1})$ and let $\overline{w}_i > w_{i-1}$ be adjacent to $w_i$ (so that $\overline{w}_i \in R_i$ and $\overline{w}_i w_{i-1}$ is one of the edges connecting $R_i$ with $T\backslash R_i$); set similar definitions for $R'_h = R(\overline{u}'_h, w'_{h-1})$. If $R_i/R_h$ is not $\lambda$-resonant for all $i$, we put $R_* \equiv R_h$ and we are done. Assume now that for some $1 \leq j \leq h-1$, $R_j/R_h$ is $\lambda$-resonant. Since $R_j$ is not $\lambda$-resonant, $\exists v_h \in R_j$ such that $|\delta_{w_{j-1}}| > \lambda|\delta_{v_h}(R_j)|$ (where $R_j$ is rooted at $\overline{u}_j$). The vertex $v_h$ cannot be outside $R_h$ (otherwise $\delta_{v_h}(R_j) = \delta_{v_h}(R_j/R_h)$ and $R_j/R_h$ would not be $\lambda$-resonant); thus $v_h \in R_h$ (and $v_h \neq \overline{u}_h$). In fact, $v_h \in R_h \backslash P(\overline{u}_h, \overline{w}_h)$ (otherwise $R_h$, which is assumed to be either Eliasson or short, would be Eliasson and, by Sublemma 6.2, $R_j/R_h$ would not be $\lambda$-resonant). Hence $\delta_{v_h}(R_j) = \delta_{v_h}$ and

$$|\delta_{w_{j-1}}| > \lambda|\delta_{v_h}|. \tag{6.16}$$

Since $R_j/R_h$ is $\lambda$-resonant, it is $|\delta_{w_{j-1}}| \le \lambda|\delta_v(R_j)|$ for all $v \in R_j/R_h$; in particular

$$|\delta_{w_{j-1}}| \le \lambda|\delta_v(R_j/R_h)| = \lambda|\delta_v|, \quad \forall v \in R'_h \backslash P(\overline{u}'_h, \overline{w}'_h). \tag{6.17}$$

We claim that the Lemma holds also in the present case if we put $R_* \equiv R'_h$. In fact, if $R'_h = P(\overline{u}'_h, \overline{w}'_h)$ then the claim is true (arguing as above and using Sublemma 6.2). Let $R'_h \backslash P(\overline{u}'_h, \overline{w}'_h) \neq \emptyset$ and assume (by contradiction) that $R_{j'}/R'_h$ is $\lambda$-resonant for some $j' < h$. Then (repeating the arguments given above) there exists $v'_h \in R'_h \backslash P(\overline{u}'_h, \overline{w}'_h)$ such that $|\delta_{w_{j'-1}}| > \lambda|\delta_{v'_h}|$ and by (6.17) (used with $v = v'_h$) and (6.16), $|\delta_{v'_h}| > |\delta_{v_h}|$. Thus $|\delta_{w_{j'-1}}| > \lambda|\delta_{v_h}|$ so that $R_{j'}/R'_h$ is not $\lambda$-resonant which is a contradiction. □

### 6.3 Proof od Lemma 5.2

If all resonances are Siegel we are in the hypothesis of Lemma 5.1, and Lemma 5.2 follows since $\beta_2 > \beta_1$ and $c_2 > c_1$. Assume now that $T$ contains short and/or Eliasson resonances. We shall proceed by induction on $k \equiv |V|$ and to reduce the order $k$ we shall *contract suitable resonances* as advertised above.

The main technical estimate is the following.

SUBLEMMA 6.3. *Let $h \ge 1$ and assume that $R_h \subset \cdots \subset R_1 \equiv R$ are resonances of $T$ such that:* (i) *$R$ is not $\lambda$-resonant;* (ii) *each $R_i$ is either Eliasson or short;* (iii) *if we define $\overline{R}_h \equiv R_h$ and, for $i < h$, $\overline{R}_i \equiv R_i/R_{i+1}$, then $\overline{R}_i$ contains at most Siegel resonances. Then*

$$\prod_R |\delta_v|^{-1} \le c_2^{|R|} \prod_R |\alpha_v|^{\beta_2} \tag{6.18}$$

*where $c_2$, $\beta_2$ are as in* (5.4).

PROOF. Let (for $i = 1, \dots, h$) $u_{i-1}\overline{u}_i$, $\overline{w}_i w_{i-1}$ be the edges connecting $R_i$ with $T\backslash R_i$, with $\overline{u}_i$, $\overline{w}_i \in R_i$ and $u_{i-1} > w_{i-1}$. Clearly: $\overline{u}_i$, $u_i$, $\overline{w}_i$, $w_i \in R_i$ for all $1 \le i < h$; $\overline{u}_h$, $\overline{w}_h \in R_h$; $u_0$ is outside $R$ and might be the earth $\eta$; $w_0$ is the first vertex following $R$; $u_{i-1} > \overline{u}_i \ge \overline{w}_i > w_{i-1}$; $\overline{u}_i \ge u_i$ (for $i < h$). If $v \neq \overline{u}_i$ and $v \in R_i$ is not on the path $P(\overline{u}_i, \overline{w}_i)$ joining $\overline{u}_i$ with $\overline{w}_i$ (which is always the case if $R_i$ is short) then $\delta_v = \delta_v(R_{i,\overline{u}_i})$. If $R_i$ is Eliasson and $\overline{u}_i < v \le \overline{w}_i$, then $\delta_v = \delta_v(R_i) + \delta_{w_{i-1}}$ (where $R_i$ is rooted at $\overline{u}_i$) and, from the definition of Eliasson resonance, it follows that $|\delta_{w_{i-1}}| \le \dfrac{\lambda}{1-\lambda}|\delta_v(R_i)|$ so that, in all cases, if $v \in R_i$ and $v \neq \overline{u}_i$ it holds

$$|\delta_v|^{-1} \le c|\delta_v(R_i)|^{-1}, \quad c \equiv \frac{1-\lambda}{1-2\lambda}. \tag{6.19}$$

Thus, since $\delta_{\overline{u}_i} = \delta_{w_{i-1}}$, by (6.19) we get

$$\begin{aligned}\prod_R |\delta_v|^{-1} &= \prod_{i=1}^h \prod_{\overline{R}_i} |\delta_v|^{-1} = \prod_{i=1}^h |\delta_{\overline{u}_i}|^{-1} \prod_{\substack{\overline{R}_i \\ v \neq \overline{u}_i}} |\delta_v|^{-1} \\ &\leq \prod_{i=1}^h |\delta_{w_{i-1}}|^{-1} c^{|\overline{R}_i|-1} \prod_{\substack{\overline{R}_i \\ v \neq \overline{u}_i}} |\delta_v(\overline{R}_i)|^{-1}\end{aligned} \tag{6.20}$$

($\overline{R}_i$ rooted at $\overline{u}_i$). Now, using (6.19) with $v = w_i$ and the Diophantine inequality (2.7), we get, for $1 \leq i < h$,

$$|\delta_{w_i}|^{-1} \leq c|\delta_{w_i}(R_i)|^{-1} = c|\delta_{w_i}(\overline{R}_i)|^{-1} \leq c\lambda^{-1}\gamma 2^{|\overline{R}_i|-1} \prod_{\overline{R}_i} |\alpha_v|^\tau. \tag{6.21}$$

Since $R$ is not $\lambda$-resonant, $|\delta_{w_0}| > \lambda\Delta(R)$, thus (6.21) *holds also for $i = 0$ if we define $\overline{R}_0 \equiv R$*. Since by (iii) we can use Lemma 5.1 to estimate the products over $\overline{R}_i$ in (6.20), we get easily (6.18) (to get the constants straight, observe that $2^\tau c_1 > \lambda^{-1}\gamma$). □

PROOF (of Lemma 5.2). Let us call, here, "ES" a resonance which is either Eliasson or short. Let us also call (for the purpose of the present proof) a "resonant hierarchy" a family of ES resonances $R_h \subset \cdots \subset R_1$ $(h \geq 1)$ satisfying (i), (ii) and (iii) of Sublemma 6.3 and such that $R_1$ is either maximal (*i.e.* is not contained in another ES resonance) or is contained into some ES resonance $S$ which contains another ES resonance disjoint from $R_1$. Pick a maximal ES resonance, $R \subset T$. Then either $R = R_1$ for some resonant hierarchy or $R$ contains more than one resonant hierarchies. In the first case, by Lemma 6.1, $T/R$ will not contain $\lambda$-resonances and we can use the inductive hypothesis and the estimate in Sublemma 6.3, to obtain Lemma 5.2. In the second case, call $S_i$ (for $i = 1, \dots, p$ for some $p > 1$) all the ES resonances contained in $R$ which do not belong to any resonant hierarchy (in particular $R$ is one of such resonances and we shall put $S_1 = R$). Pick a minimal $S_*$ among the $\{S_i\}$ (*i.e.* $S_*$ does not contain any $S_j$). By construction $S_*$ contains at least two resonant hierarchies, say, $R_1 \supset \cdots \supset R_h$ and $R'_1 \supset \cdots \supset R'_{h'}$ (with $R_1 \cap R'_1 = \emptyset$). By Lemma 6.3 (and Lemma 6.1) we can contract either $R_1$ or $R'_1$ so that the remaining tree does not contain any $\lambda$-resonance and use the estimate in Sublemma 6.3 to estimate the product over the contracted hierarchy. □

### 6.4 Proof od Lemma 5.3

The heart of the matter is the following remarkable *algebraic fact*.

PROPOSITION 6.1 (Compensations). *Let $R$ be a tree and $\alpha$ be a $R_u$-admissible function, $\forall u \in R$. Assume that $\alpha(R) = 0$ and set (for prefixed $1 \leq i$, $j \leq N$)*

$$\sigma(x) \equiv \sum_{u,w \in R} \alpha_{ui}\alpha_{wj} \prod_{\substack{v \in P(u,w) \\ v \neq u}} (\delta_v(R_u) + x)^{-2} \prod_{v \in R \setminus P(u,w)} \delta_v(R_u)^{-2} \tag{6.22}$$

*where $P(u,w)$ is the path joining $u$ and $w$ (if $u = w$ the first product is missing, while if $R$ is a path the second product is missing). Then $\sigma(0) = \sigma'(0) = 0$ $\left(' \equiv \dfrac{d}{dx}\right)$.*

PROOF. Recall that $\alpha(R) = 0$ and define, for any $u$, $w$ in $R$ and any $S \subset V(R)$ such that $\alpha(S) = 0$

$$\mu(u) \equiv \prod_{\substack{v \in R \\ v \neq u}} |\delta_v(R_u)|, \quad \nu(w) \equiv \sum_{u \in S, u \neq w} \alpha_{ui} \sum_{\substack{v \in P(u,w) \\ v \neq u}} \delta_v(R_u)^{-1}. \tag{6.23}$$

We claim that *the functions $\mu$ and $\nu$ are independent of $u$ and $w$*: $\mu(u) \equiv \mu \equiv \mu_R$, $\nu(u) \equiv \nu \equiv \nu_{R,S}$. The independence of $\mu(u)$ from $u$ comes immediately from the identities

$$\delta_v(R_u) = \delta_v(R_w), \quad \forall v \notin P(u,w) \tag{6.24}$$

(where as usual $P(u,w)$ denotes the path joining $u$ and $w$) and

$$\prod_{\substack{v \in P(u,w) \\ v \neq u}} \delta_v(R_u) = \prod_{\substack{v \in P(u,w) \\ v \neq w}} (-\delta_v(R_w)), \quad \forall u \neq w. \tag{6.25}$$

Note that (6.24) holds for any rooted tree (*i.e.* not necessarily null) while in (6.25) it is important that $R$ is null. Now, for any $u$ and $w$ with $u \neq w$ (6.24) and (6.25) imply

$$\begin{aligned} \mu(u) &= \prod_{\substack{v \in P(u,w) \\ v \neq u}} |\delta_v(R_u)| \prod_{v \notin P(u,w)} |\delta_v(R_u)| \\ &= \prod_{\substack{v \in P(u,w) \\ v \neq w}} |\delta_v(R_w)| \prod_{v \notin P(u,w)} |\delta_v(R_w)| = \mu(w) \end{aligned} \tag{6.26}$$

which proves the independency of $\mu$ from points in $R$. Next, observe that to prove the independence of $\nu$ from points in $R$ it is enough to check that $\nu(w) = v(w')$ for *adjacent points* $w$ and $w'$. Thus, let $w$, $w'$ be adjacent points in $R$. Then

$$\nu(w)-\nu(w')=\sum_{\substack{u\in S\\ u\neq w,w'}}\alpha_{ui}\left[\sum_{\substack{v\in P(u,w)\\ v\neq u}}\delta_v(R_u)^{-1}-\sum_{\substack{v\in P(u,w')\\ v\neq u}}\delta_v(R_u)^{-1}\right]+$$

$$+\chi_S(w')\alpha_{w'i}\delta_w(R_{w'})^{-1}-\chi_S(w)\alpha_{wi}\delta_{w'}(R_w)^{-1}$$

$$(6.27)\qquad =\sum_{\substack{u\in S\\ u\neq w,w'}}\alpha_{ui}\delta_w(R_{w'})^{-1}+\chi_S(w')\alpha_{w'i}\delta_w(R_{w'})^{-1}+$$

$$+\chi_S(w)\alpha_{wi}\delta_w(R_{w'})^{-1}$$

$$=\left(\sum_{u\in S}\alpha_{ui}\right)\delta_w(R_{w'})^{-1}=0$$

where in the second equality we used $\delta_{w'}(R_w)=-\delta_w(R_{w'})$ which is a particular case of (6.25) and $\chi_S$ is the characteristic function of the set $S$. This finishes the proof of the above claim.

The check of the Proposition is now trivial:

$$\sigma(0)=\sum_{u,w\in R}\alpha_{ui}\alpha_{wj}\prod_{\substack{v\in R\\ v\neq u}}\delta_v(R_u)^{-2}=\mu_R^{-2}\sum_{u,w\in R}\alpha_{ui}\alpha_{wj}=0$$

and

$$\frac{d\sigma}{dx}(0)=-2\sum_{\substack{u,w\in R\\ u\neq v}}\alpha_{ui}\alpha_{wj}\prod_{\substack{v\in R\\ v\neq u}}\delta_v(R_u)^{-2}\sum_{\substack{v\in P(u,w)\\ v\neq u}}\delta_v(R_u)^{-1}$$

$$=-2\mu_R^{-2}\sum_{w\in R}\alpha_{wj}\left(\sum_{u\neq w}\alpha_{ui}\sum_{\substack{v\in P(u,w)\\ v\neq u}}\delta_v(R_u)^{-1}\right)$$

$$=-2\mu_R^{-2}\nu_{R,R}\sum_{w\in R}\alpha_{wj}=0$$

finishing the proof of Proposition 6.1. □

REMARK 6.1. In fact, from the above proof it follows that Proposition 6.1 holds also if the sum over $R$ is replaced by the sum over any null subset of $R$.

We can now proceed with the Proof of Lemma 5.3. First observe that (letting in Proposition 6.1 $R$ be as in Lemma 5.3)

$$(6.28)\qquad \langle n\rangle^{-2}\sigma(\langle n\rangle)=\sum_{u,w\in R}\alpha_{ui}\alpha_{wj}\prod_R\delta_v(R_u^w)^{-2}$$

so that

$$\sigma(x) = x^2 \int_0^1 \sigma''(tx)(1-t)dt \equiv x^2\overline{\sigma}(x) \tag{6.29}$$

where $\overline{\sigma}$ is defined here. By assumption we have

$$|\langle n\rangle| \leq \overline{\lambda}\Delta \equiv \overline{\lambda}\Delta(R) \tag{6.30}$$

and we see that $\sigma$ *is holomorphic and bounded* in the complex disk $\left\{x \in \mathbb{C} : |x| \leq \frac{1+\overline{\lambda}}{2}\Delta\right\}$. Recall the "*Cauchy estimate*":

$$\sup_{D'} |g^{(k)}| \leq k!r^{-k} \sup_{D} |g|, \quad r \equiv \operatorname{dist}\{D', \partial D\} \tag{6.31}$$

valid for any holomorphic function $g$ bounded in a complex domain $D$, for any $k \geq 0$ and for any subdomain $D'$ whose closure is contained in $D$ (above $\partial$ denotes "boundary of")[17]. Then, letting

$$D \equiv \left\{x \in \mathbb{C} : |x| \leq \frac{1+\overline{\lambda}}{2}\Delta\right\}, \quad D' \equiv \{x \in \mathbb{C} : |x| \leq \overline{\lambda}\Delta\} \tag{6.32}$$

(recall that $\overline{\lambda} < 1$ so that $D'$ is smaller that $D$) we see that

$$\sup_{D'} |\sigma''| \leq 8(1-\overline{\lambda})^{-2}\Delta^{-2} \sup_{D} |\sigma|. \tag{6.33}$$

Furthermore, for any $x \in D$ and any $u,\ v \in R$ we have (recall that from the definition of $\Delta(R)$ it follows that $\Delta \leq \delta_v(R_u)$):

$$(\delta_v(R_u)+x)^{-2} \leq \left(\delta_v(R_u) - \frac{1+\overline{\lambda}}{2}\Delta\right)^{-2} \leq 4(1-\overline{\lambda})^{-2}\delta_v(R_u)^{-2}. \tag{6.34}$$

[17] The proof of (6.31) is standard: it follows immediately by expressing the $k^{\text{th}}$ derivative of $g$ at a point $x \in D'$ in terms of Cauchy's integral formula (taking as path of integration a disk centered at $x$ and of radius $r$).

Thus, by (6.28), (6.29), (6.30), (6.32), (6.33), (6.34), and Lemma 5.2

$$
\begin{aligned}
&\langle n\rangle^{-2}|\sigma(\langle n\rangle)| = |\overline{\sigma}(\langle n\rangle)| \le \frac{1}{2}\sup_{D'}|\sigma''| \\
&\le 4(1-\overline{\lambda})^{-2}\Delta^{-2}\sup_{D}|\sigma| \\
(6.35)\qquad &\le (4(1-\overline{\lambda})^{-2})^{|V(R)|}\Delta^{-2}\sum_{u,w\in R}|\alpha_u|\,|\alpha_w|\prod_{v\neq u}|\delta_v(R_u)|^{-2} \\
&\le (4(1-\overline{\lambda})^{-2})^{|V(R)|}\Delta^{-2}\left(|V(R)|\prod_{v\in R}|\alpha_v|\right)^2 c_2^{2(|V(R)|-1)}\prod_{\substack{v\in R\\ v\neq u}}|\alpha_v|^{2\beta_2} \\
&\le ((1-\overline{\lambda})^{-2}2^{\tau+1}c_2)^{2|V(R)|}\prod_{v\in R}|\alpha_v|^{2(\beta_2+\tau+1)}
\end{aligned}
$$

where in the estimate before last we have applied Lemma 5.2 to the tree(s) $R\backslash\{u\}$ (if $\deg_{R_u} u > 2$ then $R\backslash\{u\}$ is the disjoint union of $\deg_{R_u} - 1$ trees) and in the last estimate

$$
(6.36)\qquad \Delta^{-1} \le \gamma\left(\sum_R |\alpha_v|\right)^\tau \le \gamma 2^{(|V(R)|-1)\tau}\prod_R |\alpha_v|^\tau .
$$

The proof of Lemma 5.3 is complete. □

### 6.5 Proof of Lemma 5.4

If $T' \in \mathcal{F}(T)$, denote by $\Gamma(T') \equiv \{vv' \in E(T') : \exists R \in \mathcal{R}(T)$ with $v' \in \overline{R}$ and $v \notin R\}$ and by $E(\overline{\mathcal{R}}) \equiv \bigcup_{\overline{R}\in\overline{\mathcal{R}}} E(\overline{R})$. Then

$$
(6.37)\qquad \prod_{vv'\in E(T')}\alpha_v\cdot\alpha_{v'} = \prod_{vv'\in\Gamma(T')}\alpha_v\cdot\alpha_{v'}\prod_{vv'\in E(\overline{\mathcal{R}})\cup\overline{T}}\alpha_v\cdot\alpha_{v'}
$$

and the second product does not depend on $T'$ (*i.e.* it is common to all elements of $\mathcal{F}$). Furthermore, recalling (4.7), we see that

$$
\prod_{vv'\in\Gamma(T')}\alpha_v\cdot\alpha_{v'} = \prod_{C=(\overline{R}_1,\dots,\overline{R}_s)\in\mathcal{C}}\alpha_{u(C)}\cdot\alpha_{\overline{u}_1}\alpha_{w(C)}\cdot\alpha_{\overline{w}_s}\prod_{i=1}^{s-1}\alpha_{\overline{w}_i}\cdot\alpha_{\overline{u}_{i+1}}
$$

the last product being absent if $s = 1$. Fixing a set of indices (depending upon $C$ and taking the values $1,\dots,N$), $j_i = j_i(C)$, for $i = 0,\dots,s$, we can make more explicit (6.37) by writing out the scalar products:

$$
\prod_{vv'\in\Gamma(T')}\alpha_v\cdot\alpha_{v'} = \sum_{\substack{\{j_0,\dots,j_s\}_{C\in\mathcal{C}}\\ 1\le j_i\le N}}\prod_{C\in\mathcal{C}}\alpha_{u(C)j_0}\alpha_{w(C)j_s}\prod_{C\in\mathcal{C}}\prod_{i=1}^{s}\alpha_{\overline{u}_i j_{i-1}}\alpha_{\overline{w}_i j_i}.
$$

Now observe that $|\overline{\mathcal{R}}| \le (|V|-1)/2$, which implies $\sum_{\{j_0,\dots,j_s\}} 1 \le N^{|V|}$, and notice that, for all $T' \in \mathcal{F}$ (recall (4.8))

$$\left|\prod_{vv'\in E(T')} \alpha_v \cdot \alpha_{v'}\right| \le \prod_{v\in V} |\alpha_v|^{\deg_{T'} v} \le \prod_{v\in V} |\alpha|^{\deg_{\mathcal{F}} v}.$$

Then

$$\begin{aligned}
&\left|\sum_{T'\in\mathcal{F}} \prod_{vv'\in E(T')} \alpha_v \cdot \alpha_{v'} \prod_V \delta_v^{-2}\right| \\
&\le \prod_{vv'\in E(\overline{\mathcal{R}})\cup E(\overline{T})} |\alpha_v|\,|\alpha_v'| \sum_{\{j_0,\dots,j_s\}} \prod_{C\in\mathcal{C}} |\alpha_{u(C)j_0}|\,|\alpha_{w(C)j_s}| \\
&\qquad \times \prod_{\overline{R}\in\overline{\mathcal{R}}} \left|\sum_{\overline{u},\overline{w}\in\overline{R}} \alpha_{\overline{u}i_0}\alpha_{\overline{w}i_0'} \prod_{v\in T'} \delta_v(T')^{-2}\right| \\
&\le \left(\prod_V |\alpha_v|^{\deg_{\mathcal{F}} v}\right) N^{|V|} \sup_{\substack{T'\in\mathcal{F}\\ 1\le i(\overline{R}),i'(\overline{R})\le N}} \left|\prod_{\overline{R}\in\overline{\mathcal{R}}} \sum_{\overline{u},\overline{w}\in\overline{R}} \alpha_{\overline{u}i}\alpha_{\overline{w}i'} \prod_V \delta_v(T')^{-2}\right|
\end{aligned}$$

where the supremum is taken over all choices of the indices $\{i(\overline{R}), i'(\overline{R})\}_{\overline{R}\in\overline{\mathcal{R}}}$ and $i_0$, $i_0'$ coincide with suitable indices $j$'s. Fix $T' \in \mathcal{F}$, fix indices $i \equiv i(\overline{R})$ and $i' \equiv i'(\overline{R})$ $(1 \le i, j \le N)$ and let $R_*$ be a *minimal critical (sub)resonance: i.e.* either $R_* \equiv R_0 \in \mathcal{R}_0$ is a critical resonance which does not contain any critical subresonance or $R_*$ is the smallest element of a hierarchy $\mathcal{H}_\lambda(R_0)$ for some $R_0 \in \mathcal{R}_0 : \mathcal{H}_\lambda(R_0) = \{R_1, \dots, R_h\}$ and $R_* = R_h$. Now, by Lemma 6.2, $\overline{T} = \overline{T'/R_*}$ and $\overline{\mathcal{R}}(T'/R_*) \equiv \overline{\mathcal{R}}\backslash\{R_*\}$. Thus

$$\begin{aligned}
&\prod_{\overline{R}\in\overline{\mathcal{R}}(T')} \sum_{\overline{u},\overline{w}\in\overline{R}} \alpha_{\overline{u}i}\alpha_{\overline{w}i'} \prod_{v\in R} \delta_v(T')^{-2} \\
&= \left(\prod_{R\in\overline{\mathcal{R}}(T'/R_*)} \sum_{\overline{u},\overline{w}\in\overline{R}} \alpha_{\overline{u}i}\alpha_{\overline{w}i'} \prod_{v\in R} \delta_v(T'/R_*)^{-2}\right)\left(\sum_{\overline{u},\overline{w}\in R_*} \alpha_{\overline{u}i}\alpha_{\overline{w}i'} \prod_{v\in R_*} \delta_v(T')^{-2}\right).
\end{aligned}$$

If $h > 0$ (*i.e.* $R_*$ is a critical subresonance) and if $m_i$, for $i = 0, 1, \dots, h$, are the integer vectors associated to the hierarchy $\{R_1, \dots, R_h\}$ (see Remark 4.3 of Section 4) we have for some $\sigma_h = \pm 1$ and some $\sigma_i = 0, \pm 1$, (for $i = 0, \dots, h-1$), $m_0 \equiv \sum_{v<R_0} \alpha_v = \sum_{i=0}^{h} \sigma_i m_i$, $(|\sigma_h| = 1)$, so that, by (4.6), we have $|\langle m_0\rangle| < \frac{\lambda}{1-\lambda}\Delta(R_*)$. Notice that such a bound holds also in the case

$R_* = R_0 \in \overline{\mathcal{R}}_0$ (and in fact holds without the factor $(1-\lambda)^{-1}$). We can apply Lemma 5.3 to $R_*$ with $\overline{\lambda} \equiv \lambda/(1-\lambda) < 1$, obtaining

$$\left| \sum_{\overline{u}, \overline{w} \in R_*} \alpha_{\overline{u}_*} \alpha_{\overline{w}i'} \prod_{V} \delta_v(T')^{-2} \right| \le c_3^{|R_*|} \prod_{v \in R_*} |\alpha_v|^{\beta_3}.$$

Repeating the above procedure to the tree $T'/R_*$ it is clear that *we can inductively contract all critical resonances* ending up with the tree $T/\overline{\mathcal{R}}_0$ to which Lemma 5.2 can be applied. Thus Lemma 5.4 follows with $c_4 \equiv N \max\{c_3, c_2^2\}$, $\beta_4 \equiv \max\{\beta_3, 2\beta_2 + 2\}$ which by (5.6) yield (5.8). □

## A. - Trees

In this appendix we collect the basic facts from graph theory that are used in the main text. The material is standard and can be found in the introductory sections of most elementary books on graph theory (see *e.g.* [18] or [5]).

Given a finite set $V$, a graph $G$ on $V$ is a couple $(V, E)$ where $E$ is a subset of *unordered* couples of different elements of $V$. If needed, we specify $V \equiv V(G)$ and $E \equiv E(G)$. One can already start to count: if $k \equiv |V|$ is the cardinality of $V$, also called the *order* of $G$, then $0 \le |E(G)| \le \binom{k}{2}$ ($|E| = 0 \Longleftrightarrow E = \emptyset$). Elements $v$ of $V$ are called equivalently either *vertices* or *points* or *nodes*, while elements of $E$, denoted by $vv' \equiv v'v$ (as the couple are taken disregarding the order), are respectively called *edges* or *lines* or *branches*. It is customary to write $v \in G$ in place of the proper notation $v \in V(G)$ and, analogously, $vv' \in G$ in place of $vv' \in E(G)$. The edge $vv'$ is said to be *incident* with the vertices $v$ and $v'$. Two vertices $v, v'$ are called *adjacent* if $vv' \in E$. Two edges are *adjacent* if they are incident with the same vertex. The number of edges incident with a vertex $v$ is called the *degree of* $v$ and denoted $\deg v$ or $\deg_G v$ when needed. A point with degree 1 is called an *endpoint* (or *endvertex*). Clearly, $\sum_{v \in G} \deg v = 2|E(G)|$. Standard set theoretic notations are used in graph theory with the obvious meaning clear from context. For example: a *subgraph* $G'$ of $G$, denoted $G' \subset G$, is a graph such that $V(G') \subset V(G)$ and such that $E(G')$ is a subset of edges $vv'$ of $E(G)$ with $v$, $v' \in V(G')$; if $G_1$ and $G_2$ are two graphs, their union $G_1 \cup G_2$ is always meant as *disjoint union* $(V(G_1 \cup G_2) \equiv V(G_1) \dot{\cup} V(G_2)$, $E(G_1 \cup G_2) \equiv E(G_1) \dot{\cup} E(G_2))$; if $G'$ is a subgraph of $G$, $G \backslash G'$ is the graph whose vertices are given by $V(G) \backslash V(G')$ and whose edges are given by $E(G) \backslash \{vv' : v' \in V(G')\}$ (notice that the set of edges subtracted is, in general, strictly larger than $E(G')$); if $v$, $v' \in G$ (but $vv' \notin G$), $G + vv'$ denotes the graph obtained by including the edge $vv'$.

All the above definitions, and in general everything concerning graphs, are better understood by looking at *diagrams*.

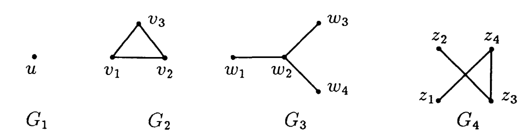



Fig. 9: $G = G_1 \cup G_2 \cup G_3 \cup G_4$

The set of vertices of the graph $G$ in Figure 9 is given by $\{u, v_1, v_2, v_3, w_1, w_2, w_3, w_4, z_1, z_2, z_3, z_4\}$, and the set of edges is $\{v_1v_3, v_1v_2, v_3v_2, w_1w_2, w_2w_3, w_2w_4, z_1z_4, z_4z_3, z_3z_2\}$. The graph $G$ is *disconnected* (see below) and the connected components are given by the subgraphs $G_1$, $G_2$, $G_3$, $G_4$; the degree of $u$ is 0 ($\Longleftrightarrow$ $u$ is an *isolated* point) the degree of $w_2$ is 3; some endpoints are $w_1$ and $z_2$; the lines incident with $z_4$ are $z_4z_1$ and $z_3z_4$; $z_1$ and $z_4$ are adjacent; the graph $G'$ given by $V(G') \equiv \{v_3, v_2\}$ and $E(G') \equiv \{v_3v_2\}$ is a subgraph of $G$; the order of $G$ is 12 and the number of edges in 9. Two graphs $G$, $G'$ (and their corresponding diagrams) are *identified* if there exists a bijection $j$ from $V(G)$ onto $V(G')$ such that $E(G') = \{j(v)j(v') : vv' \in E(G)\}$; $j$ is called a *graph isomorphism.* Thus $G_4$ in Figure 9 and the graphs depicted in Figure 10 all represent the same graph.

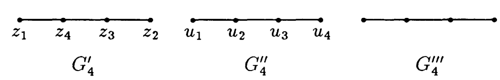



Fig. 10: Graphs isomorphic to $G_4$

An ismorphism between $G_4'$ and $G_4''$ is $j(z_1) = u_1$, $j(z_4) = u_2$, $j(z_3) = u_3$, $j(z_2) = u_4$ (another isomorphism is $j(z_1) = u_4$, $j(z_4) = u_3$, $j(z_3) = u_2$, $j(z_2) = u_1$); notice that graphically $G_4'''$ (*no labels*) is a complete description of the graph $G_4$. Here the "labels" $z_i$ or $u_i$ are used only for notational convenience as the only relevant aspect is the "topology" of lines and points; below we shall introduce instead the concept of *labeled trees* where points are distinguished by labels attached to them.

*Connectedness* is defined through paths: A *path* $P$ is a graph whose vertices can be given an order such that if $V(P) = \{v_1, \ldots, v_p\}$ then $E(P)$ is given by $\{v_1v_2, \ldots, v_{p-1}v_p\}$ (if $p = 1$, $E(P) = \emptyset$ and the path is *trivial*). The *length* of a path is given by its order minus 1 ($\equiv$ number of lines in $P$); a path $P$ is denoted by $v_1v_2 \cdots v_p$. A graph $G$ is *connected* if any two (different) points can be joined by a path in $G$. $G_4$ in Figure 9 (and in Figure 10) is a path. A *closed path* or *cycle* is a connected graph whose vertices have degree 2.

A *tree* is a connected acyclic (*i.e* with no cycles) graph. In Figure 9, $G_1$, $G_3$ and $G_4$ are trees while $G_2$ is not a tree (being itself a cycle). Equivalent characterizations of trees are the following (see *e.g.* [18]): (i) *$T$ is a tree*; (ii) *$T$ is connected and* $|E(T)| = |V(T)| - 1$ (as above we set $|\emptyset| = 0$); (iii) *Every two vertex of $T$ are joined by a unique path.* A *subtree* $T'$ of $T$ is a connected subgraph of $T$ (hence $T'$ is also a tree); unless otherwise specified, if $T$ is a tree, $T' \subset T$ means that $T'$ is a subtree of $T$.

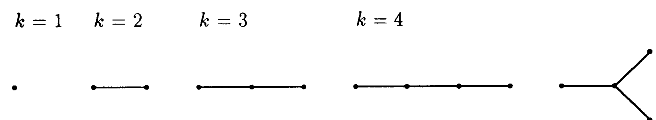



Fig. 11: Trees up to order $k = 4$

A *rooted tree* is a tree with a distinguished vertex called the *root.*

A graph (tree) on a set of *distinct* points (or *labeled points* or simply *labels*) is called a *labeled graph* (*tree*).

*Labeled rooted graphs* (*trees*) are labeled graphs (trees) with one distinguished (labeled) vertex.

A useful way of identifying the root $r$ of a rooted tree $T$ is to introduce an extra point $\eta \notin V(T)$, called the "earth", and to *add a new edge*, $\eta r$, in $E(T)$: Thus for rooted trees it is $|E(T)| = |V(T)|$ and the unique edge of the form $\eta v$ identifies the root $v$ (usually denoted $r$). Consistently, we let the degree of the root $r$ be the number of edges incident with $r$, *$\eta r$ included.* If $T$ is an unrooted tree and $v \in V(T)$, we denote by $T_v$ the rooted tree obtained from $T$ putting the root in $v$ (equivalently, adding the edge $\eta v$): $\deg_{T_v} v = \deg_T v + 1$. Notice that for rooted trees of order $k$ it is $\sum \deg v = 2k - 1$.

We conclude by counting some classes of trees. Denote by $\mathcal{T}^k \equiv \{$labeled rooted trees of order $k\}$, $\mathcal{T}_0^k \equiv \{$labeled trees of order $k\}$, $\tilde{\mathcal{T}}^k \equiv \{$rooted trees of order $k\}$, $\tilde{\mathcal{T}}_0^k \equiv \{$trees of order $k\}$ and by $t_k$, $t_k^{(0)}$, $\tilde{t}_k$. $\tilde{t}_k^{(0)}$ the respective cardinalities. The only trivial relation among the $t$'s is that $t_k = kt_k^{(0)}$. Less obvious are the following statements: (a) given $d_i \geq 1$ such that $\sum d_i = 2k - 2$, $\#\{T \in \tilde{\mathcal{T}}^k : \deg_T v_i = d_i\} = (k-2)! \left(\prod_{i=1}^{k} (d_i - 1)!\right)^{-1}$; (b) given $d_i \geq 1$ such that $\sum d_i = 2k - 1$, $\#\{T \in \mathcal{T}^k : \deg_T v_i = d_i\} = (k-1)! \left(\prod_{i=1}^{k} (d_i - 1)!\right)^{-1}$; (c) $t_k \equiv \#\mathcal{T}^k = k^{k-1}$, $t_k^{(0)} \equiv \#\mathcal{T}_0^k = k^{k-2}$. Proofs of (a) $\div$ (d) can be found in [5] (chapter VII, §3).

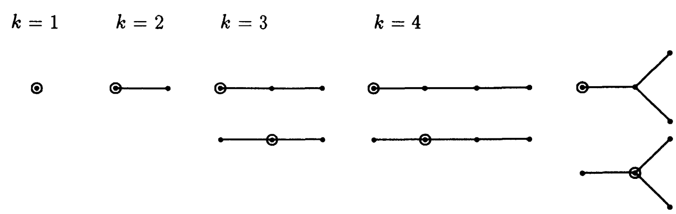



Fig. 12: Rooted trees up to order $k = 4$

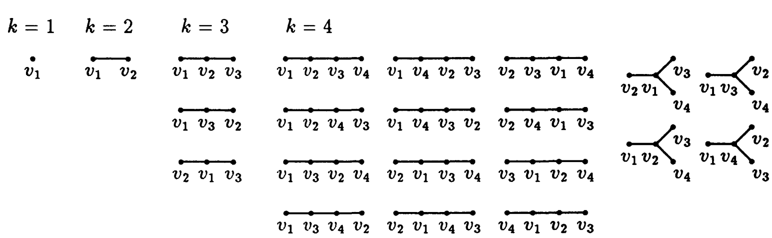



Fig. 13: Labeled trees in $\{v_1, \ldots, v_k\}$, for $k = 1, \ldots, 4$

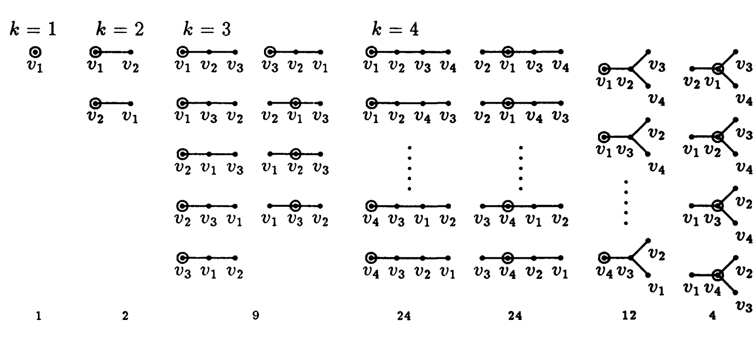



Fig. 14: Labeled rooted trees on $\{v_1, \ldots, v_k\}$, for $k = 1, \ldots, 4$

## B. - Formal Solutions, Combinatorics, Divergences

In this appendix we provide details on some known facts, mentioned in the main text, whose proofs would not be trivially traced back in the literature. We will start by giving the proofs of Proposition 2.1 on the existence and uniqueness of formal solutions and of Proposition 3.1 on the tree representation of the formal solution. Finally we shall illustrate with a concrete example the presence of single terms of size $\sim k!$ in the tree expansion of the $k^{\text{th}}$ coefficient of the formal solution.

PROOF OF PROPOSITION 2.1. We construct the formal solution by induction over $k$. For $k = 0$ we get immediately from (2.6) that $Y^{(0)}$ and $X^{(0)}$ are constant vectors and that $h'(Y^{(0)}) = \omega$ (as if $g$ is a smooth function over $\mathbb{T}^N$ and $Dg = a$ with some constant $a$, then $g \equiv 0 = a$) so that, from our hypotheses it follows that $Y^{(0)} = y_0$; requirement (2.8) fixes the constant vector $X^{(0)}$ to be zero. Let, now, $k \geq 1$ and assume that $y_0$, $Y^{(1)}, \dots, Y^{(k-1)}$, $X^{(1)}, \dots, X^{(k-1)}$, are smooth functions over $\mathbb{T}^N$ solving (2.6) with $\int_{\mathbb{T}^N} X^{(h)} = 0$. We claim that

$$\int_{\mathbb{T}^N} \phi^{(k)}(\theta) \equiv \int_{\mathbb{T}^N} \left[ f_x \left( \sum_{h=0}^{k-1} \varepsilon^h Y^{(h)}, \theta + \sum_{h=0}^{k-1} \varepsilon^h X^{(h)} \right) \right]_{k-1} = 0 \tag{B.1}$$

(note that the claim is obvious for $k = 1$ as in such a case (B.1) is just the average of the gradient of a periodic function). If the claim is true, then the proposition follows easily: $\phi^{(k)}(\theta)$ would be a $C^\infty$ (vector-valued) function over $\mathbb{T}^N$ with zero average and therefore the solutions of the equation $DY^{(k)} = \phi^{(k)}$ are given by $Y^{(k)} = D^{-1}\phi^{(k)} + c_k$ with $c_k$ constant and $D^{-1}\phi^{(k)}$ the smooth function with zero average given in Fourier expansion by

$$D^{-1}\phi^{(k)} \equiv \sum_{\substack{n \in \mathbb{Z}^N \\ n \neq 0}} \frac{\phi_n^{(k)}}{i\omega \cdot n} e^{in\cdot\theta} \tag{B.2}$$

where $\phi_n^{(k)}$ are the Fourier coefficients of $\phi^{(k)}$ (notice that the fast decay of the Fourier coefficients of the smooth function $\phi^{(k)}$ yields the fast decay, in view of (2.7), of the coefficients in (B.2), ensuring that $D^{-1}\phi^{(k)}$ is $C^\infty(\mathbb{T}^N)$). The constant $c_k$ is not arbitrary, as the right hand side of the second of (2.6), in order to make sense, must have vanishing mean value over $\mathbb{T}^N$; rewriting such equation we get

$$
\text{(B.3)}\qquad
\begin{aligned}
DX^{(k)} &= h_{yy}(y_0)Y^{(k)} + \left[h_y\left(\sum_{h=0}^{k-1}\varepsilon^h Y^{(h)}\right)\right]_k \\
&\quad + \left[f_y\left(\sum_{h=0}^{k-1}\varepsilon^h Y^{(h)}, \theta + \sum_{h=0}^{k-1}\varepsilon^h X^{(h)}\right)\right]_{k-1} \equiv h_{yy}(y_0)c_k + \psi^{(k)}
\end{aligned}
$$

where $\psi^{(k)}$ is a smooth function (depending on $H$ and $X^{(h)}$, $Y^{(h)}$ for $h \le k-1$) so that

$$
\text{(B.4)}\qquad c_k = h_{yy}(y_0)^{-1}\int_{\mathbb{T}^N}\psi^{(k)}
$$

in which case (B.3) has a unique solution with $\int_{\mathbb{T}^N} X^{(k)} = 0$.

It remains to prove (B.1). Observe that for any (smooth) functions $\overline{X}$, $\overline{Y}$ over $\mathbb{T}^N$ one has

$$
\text{(B.5)}\qquad \int \{(D\overline{Y} + H_x(\overline{Y},\theta+\overline{X}))(I+\partial_\theta\overline{X}) - (\omega + D\overline{X} - H_y(\overline{Y},\theta+\overline{X}))\partial_\theta\overline{Y}\}d\theta = 0
$$

where $\partial_\theta\overline{X}$ is the matrix $(\partial_\theta\overline{X})_{ij} \equiv \partial_{\theta_j}\overline{X}_i$ and we adopted the standard convention about row-by-column multiplication of matrices (interpreting vectors as (respectively) $1\times N$ ($N\times 1$) matrices if they are to the left (right) of an $N\times N$-matrix); identity (B.5) follows immediately if one notices that $(H_x)(I+\partial_\theta\overline{X}) + (H_y)(\partial_\theta Y)$ is just the $\theta$-gradient of $\theta \to H(\overline{Y},\theta+\overline{X})$ (so that its average vanishes) and that the remaining terms in (B.5) disappear by integration by parts. Now, we use (B.5) with $\overline{Y} = \sum_{h=0}^{k-1}\varepsilon^h Y^{(h)}$, $\overline{X} = \sum_{h=1}^{k-1}\varepsilon^h X^{(h)}$ to get an identity in $\varepsilon$; differentiating such identity with respect to $\varepsilon$ $k$ times and setting $\varepsilon = 0$ (*i.e.* evaluating $[\cdot]_k$ of the identity) and using the fact that $Y^{(h)}$, $X^{(h)}$ solve (2.6), for $h \le k-1$ we easily obtain (B.1). For a similar proof see [10]. □

PROOF OF PROPOSITION 3.1. In order to prove Proposition 3.1 we discuss first another tree expansion, used in the proof of Theorem 5.1, for a series satisfying an equation simpler than (but related to) (3.3).

PROPOSITION B.1. *Let* $n \in \mathbb{Z}^N \to \phi_n \in \mathbb{C}$ *decay faster than any power of* $|n|$ *(i.e.* $\forall s > 0$, $\sup_n |n|^s|\phi_n| < \infty$*) and let* $g = g(\varepsilon) \sim \sum_{k\ge 1} g_k\varepsilon^k$ *be the unique* *formal solution of the equation* (6.2) $g \sim \varepsilon \sum_{n\in\mathbb{Z}^N} \phi_n |n| \exp(|n| g)$. *Then*

$$\text{(B.6)} \qquad \begin{aligned} g_k &= \frac{1}{k!} \sum_{T_r \in \mathcal{T}^k} \sum_{\substack{\alpha: T_r \to \mathbb{Z}^N \\ \alpha_\eta \equiv 1}} \prod_{v\in V} \phi_{\alpha_v} \prod_{vv'\in E} |\alpha_v| \, |\alpha_{v'}| \\ &= \frac{1}{k!} \sum_{T_r \in \mathcal{T}^k} \sum_{\alpha: T_r \to \mathbb{Z}^N} \prod_{v\in V} \phi_{\alpha_v} \prod_{v\in V} |\alpha_v|^{\deg v} \end{aligned}$$

*where* $\deg v$ *denotes the degree of* $v$ *(i.e. the number of edges incident with* $v$*).*

(Actually, if the $\phi_n$'s decay exponentially (*i.e.* $|\phi_n| \le M \exp(-\xi|n|)$ for some $M$, $\xi > 0$ and all $n \in \mathbb{Z}^N$), it follows immediately from the (analytic) Implicit Function Theorem (applied to the analytic function of two complex variables $(x, \varepsilon) \in \mathbb{C}^2 \to G(x,\varepsilon) \equiv x - \varepsilon \sum \phi_n |n| \exp(|n| x)$) that $g$ converges absolutely near $\varepsilon = 0$).

PROOF. The starting point is so to get a recursive formula for the coefficients $g_k$ defined implicitly by (6.2): Expanding the right hand side of (6.2) in Taylor series and comparing equal powers of $\varepsilon$ one immediately obtain

$$\text{(B.7)} \qquad \begin{aligned} g_1 &= \sum_{n\in\mathbb{Z}^N} \phi_n |n|, \\ g_k &= \sum_{j=2}^{k} \sum_{n\in\mathbb{Z}^N} \phi_n \frac{|n|^j}{(j-1)!} \sum_{\substack{h_1+\cdots+h_{j-1}=k-1 \\ h_i \ge 1}} \prod_{i=1}^{j-1} g_{h_i}, \quad (k \ge 2). \end{aligned}$$

Now, there is a natural way of linking (labeled rooted) trees of order $k$ with all possible trees of order $h_1, \ldots, h_{j-1}$ with $h_1 + \cdots + h_{j-1} = k-1$: This link is based on the following construction. Let $\tilde{\mathcal{T}}^k$ denote the rooted trees of order $k$ (non labeled) and let $\tilde{\mathcal{T}}_0^k$ denote the trees of order $k$ (no labels, no root).

DEFINITION B.1. Let $s \ge 1$, let $\tilde{T}_i \in \tilde{\mathcal{T}}^{h_i}$ where $i = 1, \ldots, s$ and $h_i \ge 1$ with $h \equiv \sum_{i=1}^{s} h_i$. Let $\tilde{T}_i^0 \in \tilde{\mathcal{T}}_0^{h_i}$ be the (unrooted) tree obtained from $\tilde{T}_i$ by not distinguishing the root. We define the rooted tree $\tau(\tilde{T}_1, \ldots, \tilde{T}_s) \in \tilde{\mathcal{T}}^{h+1}$ by setting[18]

$$\text{(B.8)} \qquad \tau(\tilde{T}_1, \ldots, \tilde{T}_s) \equiv \left( \tilde{T}_1^0 \cup \cdots \cup \tilde{T}_s^0 \cup \{r\} + \sum_{i=1}^{s} r r_i \right)_r$$

where $r$ is the root of $\tau(\tilde{T}_1, \ldots, \tilde{T}_s)$ ($r$ is an extra vertex *i.e.* $r \notin \tilde{T}_i$) and the $r_i$'s are the roots of $\tilde{T}_i$ *i.e.* $(\tilde{T}_i^0)_{r_i} = \tilde{T}_i$.

[18] For the (standard) notation see Appendix A.

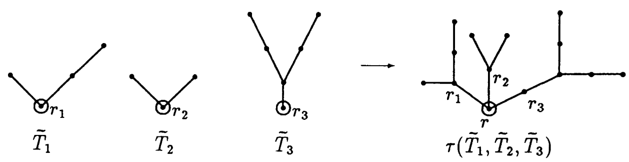



Fig. 15: The $\tau$-operation

And here it is the combinatorics (recall that the number of labeled rooted trees of order $h$ is $h^{h-1}$ (see Appendix A)):

LEMMA B.1. *For $k \geq j \geq 2$, denote by $t_{kj}$ the number of labeled rooted trees of order $k$ with root of degree $j$. Then*

(i)
$$t_{kj} = k \binom{k-2}{j-2} (k-1)^{k-j}$$

(ii)
$$t_{kj} = \frac{k!}{(j-1)!} \sum_{\substack{h_1+\cdots+h_{j-1}=k-1 \\ h_i \geq 1}} \prod_{i=1}^{j-1} \frac{h_i^{h_i-1}}{h_i!}$$

(iii)
$$\frac{k^{k-1}}{k!} = \sum_{j=2}^{k} \frac{1}{(j-1)!} \sum_{\substack{h_1+\cdots+h_{j-1}=k-1 \\ h_i \geq 1}} \prod_{i=1}^{j-1} \frac{h_i^{h_i-1}}{h_i!}$$

PROOF. Proof of (i). Denote by $\tilde{\mathcal{T}}_0^k$ the set of labeled (unrooted) trees of order $k$. Then $t_{kj} = k\#\{T \in \tilde{\mathcal{T}}_0^k : \deg v_k = j-1\}$. Now, recalling (Appendix A) that given integers $d_i \geq 1$ such that $\sum_{i=1}^{k} d_i = 2k-2$, one has

(B.9)
$$\#\{T \in \tilde{\mathcal{T}}_0^k : \deg v_i = d_i\} = \frac{(k-2)!}{\prod_{i=1}^{k} (d_i - 1)!}$$

we obtain

$$t_{kj} = k \sum_{T\in\tilde{\mathcal{T}}_0^k:\deg v_k=j-1} 1 = k \sum_{\substack{d_k=j-1,d_i\geq 1\\ \sum d_i=2k-2}} \frac{(k-2)!}{\prod_{i=1}^{k}(d_i-1)!}$$

$$= k\binom{k-2}{j-2} \sum_{\substack{s_i\geq 0\\ s_1+\cdots+s_{k-1}=k-j}} \frac{(k-j)!}{\prod_{i=1}^{k-1} s_i!} = k\binom{k-2}{j-2}(k-1)^{k-j}$$

proving (i).

Proof of (ii). It is just a matter of counting: For any choice of $j-1$ trees $T_i \in \mathcal{T}^{h_i}$ with $\sum_{i=1}^{j-1} h_i = k-1$, pick one among $k$ labels (it will be the label of the root of an element of $\mathcal{T}^k$) and choose $h_i$ labels among the $k-1$ labels at your disposal. There are $\left(\sum h_i\right)!/(h_1!\cdots h_{j-1}!) = (k-1)!/(h_1!\cdots h_{j-1}!)$ ways of doing such a choice. Now, we attach the $j-1$ trees $T_i$ according to the rule described in (B.8) (with $s=j-1$) so as to form a generic element of $\mathcal{T}^k$ with root of degree $j$: obviously, since changing the order in (B.8) give rise to the same tree, we have to divide by $(j-1)!$. Summing over all possible choices of $h_i$ and over all possible trees in $\mathcal{T}^{h_i}$ we obtain (ii).

(iii) is immediately obtained by summing (ii) over all possible degrees of the root *i.e.* $j=2,\ldots,k$. □

The main point of the above Lemma is (ii): (i) is just a curiosity and will not be used and (iii) is simply obtained by summing (ii) over $j$. An immediate corollary of this Lemma (better: "of the proof of this Lemma") is the following Corollary.

COROLLARY B.1. *Let* $G:\tilde{\mathcal{T}}^k\to\mathbb{C}$. *Then*

$$\text{(B.10)}\qquad \frac{1}{k!}\sum_{T\in\mathcal{T}^k} G(\tilde{\mathcal{T}}) = \sum_{j=2}^{k}\frac{1}{(j-1)!}\sum_{\substack{h_1+\cdots+h_{j-1}=k-1\\ h_i\geq 1}}\prod_{i=1}^{j-1}\frac{1}{h_i!}\sum_{T_1\in\mathcal{T}^{h_1},\ldots,T_{j-1}\in\mathcal{T}^{h_{j-1}}} G(\tau(\tilde{T}_1,\ldots,\tilde{T}_{j-1}))$$

PROOF. It is an immediate consequence of the argument used to prove (ii) above. □

We are now ready to prove (B.6) by induction on $k\geq 1$. For $k=1$ the claim is trivially true. Let $k\geq 2$ and assume that (B.6) holds for $1,\ldots,k-1$.

Denoting, for $T \in \tilde{\mathcal{T}}^h$

$$F(T) \equiv \sum_{\alpha:T\to\mathbb{Z}^N} \prod_{v\in V} \phi_{\alpha_v} |\alpha_v|^{\deg v} \tag{B.11}$$

in view of (B.10) we find

$$\begin{aligned}
g_k &= \sum_{j=2}^{k} \frac{1}{(j-1)!} \sum_{\substack{h_i\geq 1 \\ h_1+\cdots+h_{j-1}=k-1}} \prod_{i=1}^{j-1} \frac{1}{h_i!} \\
&\qquad \times \sum_{T_1\in\mathcal{T}^{h_1},\ldots,T_{j-1}\in\mathcal{T}^{h_{j-1}}} \sum_{n\in\mathbb{Z}^N} \phi_n |n|^j F(\tilde{T}_1)\cdots F(\tilde{T}_{j-1}) \\
&= \sum_{j=2}^{k} \frac{1}{(j-1)!} \sum_{\substack{h_i\geq 1 \\ h_1+\cdots+h_{j-1}=k-1}} \prod_{i=1}^{j-1} \frac{1}{h_i!} \\
&\qquad \times \sum_{T_1\in\mathcal{T}^{h_1},\ldots,T_{j-1}\in\mathcal{T}^{h_{j-1}}} F(\tau(\tilde{T}_1,\ldots,\tilde{T}_{j-1})) \\
&= \frac{1}{k!} \sum_{T\in\mathcal{T}^k} F(\tilde{T}).
\end{aligned}$$

Notice that in the second equality we have used the crucial property of $F$:

$$\sum_{n\in\mathbb{Z}^N} \phi_n |n|^j F(\tilde{T}_1)\cdots F(\tilde{T}_{j-1}) = F(\tau(\tilde{T}_1,\ldots,\tilde{T}_{j-1})). \tag{B.12}$$

This finishes the proof of Proposition B.1. □

REMARK B.1. In fact, it is easy to see from (B.6) that

$$g_k = \frac{1}{k!} \sum_{n_1,\ldots,n_k} \left( \prod_{i=1}^{k} \phi_{n_i} |n_i| \right) (|n_1| + \cdots + |n_k|)^{k-1}. \tag{B.13}$$

However, a direct (arithmetic) proof of (B.6), starting from the explicit expression (A.5), might not be simpler than the proof presented above.

To get also the **proof of Proposition 3.1** from Corollary B.1, one needs only to rewrite (3.7) properly. To do this, let for any $m$, $n \in \mathbb{Z}^N \backslash \{0\}$ and $k \geq 1$

$$\xi^m \equiv im \cdot X, \quad \xi^{(m,k)} \equiv im \cdot X^{(k)}, \quad \xi_n^{(m,k)} \equiv im \cdot X_n^{(k)}.$$

Then from (3.3) and Taylor expansion (in $\varepsilon$) we obtain easily

$$\begin{aligned}\xi_n^{(m,1)} &= \langle n\rangle^{-2} m\cdot n f_n\\ \xi_n^{(m,k)} &= \langle n\rangle^{-2}\sum_{j=2}^{k}\sum_{n\in\mathbb{Z}^N\setminus\{0\}}\sum_{\substack{n_1+\cdots+n_j=n\\ n_i\in\mathbb{Z}^N\setminus\{0\}}} m\cdot n_j f_{n_j}\\ &\qquad\times\frac{1}{(j-1)!}\sum_{\substack{h_1+\cdots+h_{j-1}=k-1\\ h_i\geq 1}}\prod_{i=1}^{j-1}\xi_{n_i}^{(n_j,h_i)}\quad(k\geq 2).\end{aligned}\tag{B.14}$$

The similarity of (B.14) and (B.7) is transparent ($|n|\phi_n$ corresponds to $\langle n\rangle^{-2}m\cdot n_j f_{n_j}$, while $|n|g_{h_i}$ corresponds to $\xi_{n_i}^{(n_j,h_i)}$). It is enough to change the extensions of the $\alpha$ functions on the earth $\eta$ to generalizes (3.7) to

$$\begin{aligned}\xi_n^{(m,k)} &= \frac{1}{k!}\sum_{T_r\in\mathcal{T}^k} F_{mn},\quad\text{with:}\\ F_{mn}(T_r) &\equiv \sum_{\substack{\alpha\in\mathcal{A}(T_r)\\ \alpha(T_r)=n,\alpha_\eta\equiv m}}\prod_{v\in V} f_{\alpha_v}\prod_{vv'\in E}\alpha_v\cdot\alpha_{v'}\prod_{v\in V}\delta_v^{-2}.\end{aligned}$$

Notice that $F_{mn}$ satisfies, for any $\tilde{T}_i\in\tilde{\mathcal{T}}^{h_i}$, $(i=1,\dots,j-1)$,

$$\begin{aligned}&\langle n\rangle^{-2}\sum_{\substack{n_1,\dots,n_{j-1}\\ n_1+\cdots+n_{j-1}=n}} m\cdot n_j f_{n_j} F_{n_j n_1}(\tilde{T}_1)\cdots F_{n_j n_{j-1}}(\tilde{T}_{j-1})\\ &= F_{mn}\tau(\tilde{T}_1,\dots,\tilde{T}_{j-1}).\end{aligned}\tag{B.15}$$

Now, mimic the proof of Proposition B.1 using (B.15) in place of (B.12). $\square$

We proceed now to illustrate the well-known **mechanism of divergences** of series with coefficients of type (3.7), (for an $f$ in (3.3) real analytic), *if signs are disregarded i.e.* when absolute values are introduced within the sums. This fact, based on the repetition of particularly small divisors $\delta_v$ (resonances), shows the need for detecting the compensations among all the terms whose size in absolute value grows faster than exponentially in $k$ and whose actual occurrence will readily be seen.

More explicitly, let, for $1\leq\bar{\jmath}\leq N$:

$$A_{\bar{\jmath}n}^{(k)} = \frac{1}{k!}\sum_{T_r\in\mathcal{T}^k}\sum_{\substack{\alpha\in\mathcal{A}(T_r)\\ \alpha(T_r)=n,\alpha_\eta\equiv e_{\bar{\jmath}}}}\left|\prod_{v\in V} f_{\alpha_v}\prod_{vv'\in E}\alpha_v\cdot\alpha_{v'}\right|\prod_{v\in V}\delta_v^{-2}\tag{B.16}$$

we want to show that

$$\sup_{\jmath,k,n} A^{(k)}_{\jmath n} e^{-\sigma|n|} \varepsilon_0^k = +\infty, \quad \forall \varepsilon_0,\ \sigma > 0 \tag{B.17}$$

where in the supremum $1 \le \jmath \le N$, $k \ge 1$ and $n \in \mathbb{Z}^N$. In fact, it is enough to check (B.17) in the case $N = 2$, $f(x_1, x_2) \equiv \cos x_1 + \cos(x_2 - x_1)$, $\omega_2 > \omega_1 > 0$, $\omega_1/\omega_2$ irrational, since it will be clear that the same argument can easily be adapted to the general analytic case *as long as $f$ has two Fourier coefficients with linearly independent (in $\mathbb{Z}^N$) indices*. From an elementary number theoretical theorem by Dirichlet (see [25]) it follows that there exist $p_j$, $q_j \ge 1$, with $q_j \nearrow \infty$ such that $\left|\frac{\omega_1}{\omega_2} q_j - p_j\right| < \frac{1}{q_j}$ and we can assume without loss of generality that, for all $j \ge 1$, $q_j > \frac{\omega_2}{\omega_2 - \omega_1}$. We shall now select a particular (unlabeled) rooted tree and associate to its vertices a particular choice of Fourier indices (*i.e.* a particular choice of $\alpha$). Let $P$ be the path of order $k$, $u_k u_{k-1} \dots u_1$ rooted at $u_k$: $u_k = r$ and $u_k > u_{k-1} > \dots > u_1$. Let $s_j \equiv q_j - p_j$ and, for any $h \ge 0$ let $k \equiv q_j + 2h : q_j = p_j + s_j,\ 1 \le s_j,\ p_j < q_j,\ k = q_j + 2h$. Let now $\alpha_{u_i} \equiv n_i \in \mathbb{Z}^2$ be defined as follows:

$$n_i \equiv \begin{cases} (1,0) & \text{if } 1 \le i \le s_j \\ (1,-1) & \text{if } s_j < i \le q_j \\ (1,0) & \text{if } i = q_j + 1,\ q_j + 3, \dots, q_j + (2h-1) \\ (-1,0) & \text{if } i = q_j + 2,\ q_j + 4, \dots, q_j + 2h \end{cases}$$

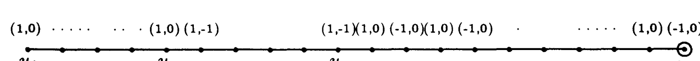



Fig. 16: Divergent contributions

and choose $h = 3q_j$ (so that $k \equiv k_j = 7q_j$).

Then, one can easily check that if in the sum (B.16) (with $\jmath = 1$, $\alpha_\eta \equiv (1,0)$, $n = (q_j, -p_j)$) we keep only the terms corresponding to the unlabeled rooted tree $P$ with the above choice of Fourier indices, recalling from Remark 3.1 that the number of labeled rooted trees corresponding to $P$ are exactly $k!$, it is

$$A^{(k_j)}_{1n} > 2^{-k_j} (2\omega_2)^{-2k_j} (k_j!)^{4/7} \tag{B.18}$$

from which (B.17) follows at once. Notice that by increasing $h$, the exponent of $k_j!$ in (B.18) becomes arbitrarily close to 1. $\square$

## C. - Siegel's Lemma

We follow [11]. The first two steps are stated in the next two Sublemmas.

SUBLEMMA C.1. *Let* $P \equiv v_1 \dots v_k$ *be the path with* $k$ *vertices rooted at* $v_1$. *Let* $\alpha : P \to \mathbb{Z}^N \backslash \{0\}$ *be a* $P$*-admissible function and fix* $0 < \lambda < 1$. *Assume that*

$$\text{if } \delta_u = \delta_w \text{ for some } u > w \text{ then there exists } v \text{ between } u \text{ and } w \text{ such that } |\delta_w| > \lambda |\delta_v|. \tag{C.1}$$

*Then there exist positive constants* $D_0 \geq B_0 \geq \gamma$ *such that*

$$\prod_P |\delta_v|^{-1} \leq B_0 D_0^{k-1} \prod_V |\alpha_v|^\tau. \tag{C.2}$$

*The constants* $B_0$, $D_0$ *can be taken to be*

$$B_0 = \gamma, \quad D_0 = \gamma 2^\tau \frac{1+\lambda}{\lambda}. \tag{C.3}$$

PROOF. By induction on $k$. If $k = 1$, (C.2) follows at once from the Diophantine inequality (2.7). Let $k \geq 2$ and assume the statement true for all rooted paths with up to $k - 1$ vertices satisfying (C.1). Let $(P, \alpha)$ as above. Define $s$ to be the greatest integer between 1 and $k$ such that $|\delta_{v_s}| \geq \max_{1 \leq i \leq k} |\delta_{v_i}|$. There are four cases:

$$\text{(i) } s = 1, \qquad \text{(ii) } \alpha_{v_s} + \alpha_{v_{s-1}} \neq 0, \\ \text{(iii) } \alpha_{v_s} + \alpha_{v_{s-1}} = 0, \ s+1 < k, \quad \text{(iv) } \alpha_{v_s} + \alpha_{v_{s-1}} = 0, \ s+1 = k \tag{C.4}$$

Note that if $\alpha_{v_s} + \alpha_{v_{s-1}} = 0$ we have $s < k$ because $\alpha$ is $P$-admissible. If $i = s$ when $s < k$ or $i = s - 1$ when $s > 1$, by the maximality of $|\delta_{v_s}|$ we have

$$|\langle \alpha_{v_i} \rangle| = |\delta_{v_i} - \delta_{v_{i+1}}| \leq |\delta_{v_i}| + |\delta_{v_{i+1}}| \leq 2 |\delta_{v_s}| \tag{C.5}$$

hence, in case (i), by (2.7), $|\delta_{v_1}|^{-1} \leq 2 |\langle \alpha_{v_1} \rangle|^{-1} \leq 2\gamma |\alpha_{v_1}|^\tau$ while in case (ii)

$$|\delta_{v_s}| \geq \frac{1}{2} \max\{|\langle \alpha_{v_s} \rangle|, |\langle \alpha_{v_{s-1}} \rangle|\}. \tag{C.6}$$

In case (iii) and (iv) we have that $\delta_{v_{s+1}} = \delta_{v_{s-1}}$, and by (C.1)

$$|\delta_{v_{s+1}}| > \lambda |\delta_{v_s}| > \lambda |\delta_{v_{s+2}}| \tag{C.7}$$

where last inequality holds only in case (iii); hence mimicking (C.5) replacing 2 by $(1+\lambda)^{-1}$ we obtain

$$|\delta_{v_{s+1}}| \geq \frac{\lambda}{1+\lambda} \max\{|\langle \alpha_{v_{s+1}} \rangle|, |\langle \alpha_{v_s} \rangle|\}. \tag{C.8}$$

Let now $\overline{v} \equiv v_s$ in case (i), (ii) and $\overline{v} \equiv v_{s+1}$ in case (iii), (iv) and let, in the cases (ii), (iii), (iv), $v' > \overline{v}$ be adjacent to $\overline{v}$. Then in case (ii), (iii), (iv) we obtain from (C.6) or (C.8), using the Diophantine inequality (2.7) (and the fact that $\lambda < 1$), that

$$|\delta_{\overline{v}}|^{-1} \frac{|\alpha_{\overline{v}} + \alpha_{v'}|^\tau}{|\alpha_{\overline{v}}|^\tau |\alpha_{v'}|^\tau} \leq \frac{1+\lambda}{\lambda} 2^\tau \gamma. \tag{C.9}$$

Now, if we let $\overline{P} \equiv P/\{\overline{v}\}$ (recall Definition 4.10) and $\overline{\alpha} : \overline{P} \to \mathbb{Z}^N \backslash \{0\}$ defined by

$$\overline{\alpha}_v \equiv \alpha_v \text{ in case (i); and in case (ii), (iii), (iv):}$$

$$\overline{\alpha}_v \equiv \begin{cases} \alpha_v & \text{if } v \neq v' \\ \alpha_v + \alpha_{\overline{v}} & \text{if } v = v' \end{cases} \tag{C.10}$$

Then

$$\prod_P |\delta_v|^{-1} = |\delta_{\overline{v}}|^{-1} \prod_{\overline{P}} |\delta_v(\overline{P}; \overline{\alpha})|^{-1}. \tag{C.11}$$

If $(\overline{P}, \overline{\alpha})$ verifies (C.1) then we can apply the inductive hypothesis and, using (C.9), the Sublemma would follow immediately. It remains to check that $(\overline{P}, \overline{\alpha})$ satisfy (C.1).

Case (i) is obvious as $\overline{\alpha}$ coincide with $\alpha$ on $\overline{P} = v_2 \ldots v_k$. Also case (iv) is easily checked as any couple of resonant points (*i.e.* points where the divisors $\delta$ coincide) in $\overline{P}$ would also be resonant for $P$. Consider case (ii) and assume that $u > w$ form a couple of resonant points for $\overline{P}$. If either $w > v_{s-1}$ or $v_{s-1} > u$, then (C.1) follows immediately from the validity of (C.1) for $P$. If $u > v_{s-1} > w$, let $v_i \in P$ be such that $u > v_i > w$ and such that $|\delta_w| > \lambda |\delta_{v_i}|$. Then, if $v_i \neq v_s$, (C.1) holds for $\overline{P}$ as $v_i \in \overline{P}$; on the other hand if $v_i = v_s$ then $|\delta_{v_s}| \geq |\delta_{v_{s-1}}| \Longrightarrow |\delta_w| > \lambda |v_s| \geq \lambda |\delta_{v_{s-1}}|$ showing that (C.1) holds for $\overline{P}$ also in this case. Consider case (iii). Arguing as above we see that we only have to check the case when, in $P$, we have $\delta_u = \delta_w$ with $u > v_s > w$. Let as above $v_i \in P$ be such that $|\delta_w| > \lambda |\delta_{v_i}|$. If $v_i = v_{s+1}$ then either $u > v_{s-1}$ or $u = v_{s-1}$. If $u > v_{s-1}$ then $|\delta_w| > \lambda |\delta_{v_{s+1}}| = \lambda |\delta_{v_{s-1}}|$. If $u = v_{s-1}$, then

$$\sum_{\substack{v \in P \\ v_s > v > w}} \alpha_v = \sum_{\substack{v \in P \\ v_{s+1} > v > w}} \alpha_v = -\alpha_{v_{s+1}}. \tag{C.12}$$

Hence $v_{s+1} > w$ for a resonant couple in $P$ and there exists by (C.1) a point $v_j$ between $v_{s+1}$ and $w$ ($\Longrightarrow v_j \in \overline{P}$) such that $|\delta_w| > \lambda |\delta_{v_j}|$ and we see that (C.1) holds for $(\overline{P}, \overline{\alpha})$ also in this final case. □

SUBLEMMA C.2. *Let $r \geq 0$ and $s \geq 2$ be integers and let $x_1, \ldots, x_r$, $y_1, \ldots, y_s$ be real numbers greater than or equal to 1 and let $z$ be their sum*

$z \equiv \sum x_i + \sum y_j$ *($r = 0$ means that there are no x's). If* $\sum_{j=1}^{s} y_j \geq \frac{z}{2}$ *and* $y_j \leq \frac{z}{2}$ $\forall j$ *then*

$$\prod_{i=1}^{r} x_i^{-1} \prod_{j=1}^{s} y_j^{-2} \leq 2^{r+2s} z^{-3}.$$

PROOF. The proof is taken from [11]: we include it for completeness. We first observe that it is enough to prove the statement for $r \leq 1$ (as $\sum x_i \leq r \prod x_i \leq 2^{r-1} \prod x_i$). We can also assume that $s \leq 3$ (in fact, since $(y_i + y_j)^2 \leq 4(y_i y_j)^2$, we can replace $y_i$ and $y_j$ by their sum if such sum is $\leq z/2$; hence we can assume $y_i + y_j > z/2$ for all $i \neq j$ and repeating the argument we end up with at most three $y_j$). Now, let $s = 3$. For $z = 3$ the result holds therefore we assume that $z > 3$ and $y_1 \leq y_2 \leq y_3$. Clearly $y_2 < z/2$ (otherwise $y_3$ would also be $\geq z/2$ contradicting the positivity of $y_1$). Furthermore

$$(y_1 - t)(y_2 + t) \leq y_1 y_2 \quad \forall t \geq 0 \tag{C.13}$$

and taking $t = \min\left\{y_1 - 1, \frac{z}{2} - y_2\right\}$ we see that we can assume either $y_1 = 1$ or $y_2 = \frac{z}{2}$. Repeating the argument for $y_2$ and $y_3$ we can assume that either $y_2$ or $y_3$ is equal to $z/2$. But then $y_1 + y_2 \leq z/2$ and by the above argument we can reduce to the case $s = 2$. We are thus lead to consider only the case $s = 2$. By (C.13) we can again assume that either $y_1 = 1$ or $y_2 = z/2$. The remaining cases are checked directly. □

We are now ready for the:

PROOF OF LEMMA 5.1. We shall prove, by induction on the number $|V|$ of vertices of $T$, the following estimate

$$\prod_V |\delta_v|^{-1} \leq B D^{|V|-1} \prod_V |\alpha_v|^{3\tau} \left(\sum_V |\alpha_v|\right)^{-2\tau} \tag{C.14}$$

for suitable constants $D \geq B \geq \gamma$ determined below. Lemma 5.1 follows immediately from (C.14) by taking $\beta_1 = 3\tau$ and $c_1 \geq D$.

If $|V| = 1$, (C.14) follows from (2.7) (as $B \geq \gamma$). Now let $k \geq 2$, assume that (C.14) holds for any $(T', \alpha')$ satisfying the hypotheses of Lemma 5.1 and with $|V(T')| \leq k - 1$ and let $(T, \alpha)$ be as in Lemma 5.1 with $|V(T)| = k$. We need some notation. Let $S_v$ be the subtree $\{v' \in T : v' \leq v\}$ rooted at $v$ and for any subset of vertices $W \subset V$ let $|\alpha|(W) \equiv \sum_{v \in W} |\alpha_v|$. Define $P \equiv \left\{v \in T : |\alpha|(S_v) > \frac{1}{2}|\alpha|(T)\right\}$. It is easy to check that $P$ *is a path*: $P = u_0 \dots u_p$ with $p \geq 0$, $u_0$ being the root $r$ of the tree $T$. To each $u_i$ we associate $m_i \geq 0$ subtrees $S_{ij}$ as follows. Let $m_i \equiv \deg u_i - 2$ when $i \leq p - 1$,

$m_p \equiv \deg u_p - 1$ and let, for $m_i > 0$, $v_{ij}$ be the $m_i$ vertices adjacent to $u_i$ with $v_{ij} \notin P$. We then set $S_{ij} \equiv S_{v_{ij}}$ when $m_i \neq 0$ and $S_{ij} = \emptyset$ otherwise. We also define:

$$y_{ij} \equiv |\alpha|(S_{ij}), \; y_i \equiv \sum_{j=1}^{m_i} y_{ij}, \; x_i \equiv |\alpha_{u_i}| \tag{C.15}$$

$$\begin{aligned} &I \equiv [0, \dots, p], \; I_0 \equiv \{i \in I : m_i = 0\}, \; I_1 \equiv \{i \in I : m_i > 0\} \\ &m \equiv \sum_I m_i, \; k_{ij} \equiv |S_{ij}|, \; k_i \equiv \sum_{j=1}^{m_i} k_{ij}, \; p_0 \equiv |I_0|, \; p_1 \equiv |I_1|. \end{aligned} \tag{C.16}$$

It then follows

$$|\alpha|(T) = \sum_I x_i + \sum_{I_1} y_i, \; p_0 + p_1 = p + 1, \; k = p + 1 + \sum_{i \in I_1} k_i, \; m \geq p_1$$

and, by the definition of $P$

$$\begin{aligned} &\sum_{j \geq i} x_j + \sum_{\substack{j \geq i \\ j \in I_1}} y_j > \frac{1}{2} |\alpha|(T), \quad \forall i \in I \\ &x_i \leq \frac{1}{2} |\alpha|(T) \; \forall i < p, \; y_{ij} \leq \frac{1}{2} |\alpha|(T) \; \forall i \in I_1, \; 1 \leq j \leq m_i. \end{aligned} \tag{C.17}$$

If $I_1 = \emptyset$ then $P \equiv T$ and we can apply directly Sublemma C.1 assuming that

$$B \geq B_0, \quad D \geq 2^{2\tau} D_0. \tag{C.18}$$

From now on we assume that $I_1 \neq \emptyset$ *i.e.* $m \geq p_1 > 0$. We shall adopt the convention that *sums or products over empty set of indices have to be disregarded* (*i.e.* replaced, respectively, by 0 or 1). Next, let

$$\overline{\alpha}_{u_i} \equiv \alpha_{u_i} + \sum_{j=1}^{m_i} \alpha(S_{ij}) \Longrightarrow |\overline{\alpha}_{u_i}| \leq x_i + y_i \tag{C.19}$$

then $\delta_{u_i} \equiv \delta_{u_i}(T; \alpha) = \delta_{u_i}(P; \overline{\alpha})$, whence

$$\prod_T |\delta_v|^{-1} = \prod_P |\delta_{u_i}(P; \overline{\alpha})|^{-1} \prod_{i \in I_1} \prod_{j=1}^{m_i} \prod_{v \in S_{ij}} |\delta_v|^{-1} \tag{C.20}$$

and we can use Sublemma C.1 to bound the product over $P$ (as $(P, \overline{\alpha})$ obviously satisfy (C.1)) and the inductive hypotheses to bound the product over $S_{ij}$ (note that $\delta_v(T) = \delta_v(S_{ij})$). Recalling the definitions in (C.16), we obtain from (C.20) and (C.19)

$$\prod_{v \in T} |\delta_v|^{-1} \leq MBD^{k-1} \left( \prod_T |\alpha_v|^{3\tau} \right) |\alpha|(T)^{-2\tau}$$

with

$$M \equiv \frac{B_0}{B} \left(\frac{B}{D}\right)^m \left(\frac{D_0}{D}\right)^p \left[ \prod_{I_0} x_i^{-2} \cdot \prod_{I_1} \left( \frac{x_i + y_i}{x_i^3} \prod_j y_{ij}^{-2} \right) \cdot |\alpha|(T)^2 \right]^\tau \tag{C.21}$$

and we have to show that $M \leq 1$ if we choose suitably $D$, $B$. Suppose first that $p \in I_0$. Then by the first inequality in (C.17) with $i = p$ we have

$$x_p > \frac{1}{2} |\alpha|(T). \tag{C.22}$$

Then using repeatedly the bound $\sum_{i=1}^m b_i \leq m \prod_{i=1}^m b_i \leq 2^{m-1} \prod_{i=1}^m b_i$ (valid for real numbers $b_i \geq 1$) and estimating $|\alpha|(T)^2$ in (C.21) by $4x_p^2$ we get

$$M \leq \frac{B_0}{B} \left(\frac{B}{D}\right)^m \left(\frac{D_0}{D}\right)^p \left[ \prod_{\substack{i \in I \\ i \neq p}} x_i^{-2} 2^{p_1+2} 2^{m-p_1} \prod_{i \in I_1, j} y_{ij}^{-1} \right]^\tau \leq 1$$

provided

$$D \geq 2^\tau \max\{B, 2^\tau D_0\}, \quad B \geq B_0 2^{2\tau} \tag{C.23}$$

(which is stronger than (C.18)).

Suppose now that $p \in I_1$. We can assume that $x_p \leq |\alpha|(T)/2$ (otherwise (C.22) holds also in this case and we can proceed as above). Then

$$M \leq \frac{B_0}{B} \left(\frac{B}{D}\right)^m \left(\frac{D_0}{D}\right)^p \left[ 2^{2m} \prod_{\substack{i \in I \\ i \neq p}} x_i^{-1} \cdot \prod_{\substack{i \in I_1 \\ i \neq p}} y_i^{-1} \cdot x_p^{-2} \prod_{j=1}^{m_p} y_{pj}^{-2} \cdot y_p |\alpha|(T)^2 \right]^\tau$$

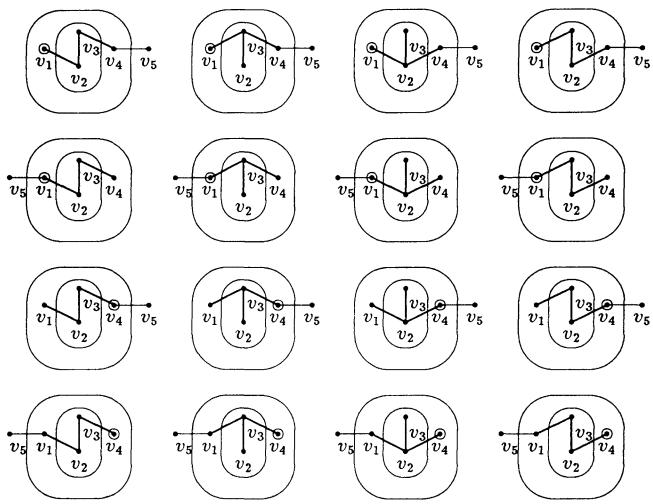
and we see that we are in position to apply Sublemma C.2 (with $r = p + p_1 - 1$, $s = m_p + 1$, $z = |\alpha|(T) = \sum_I x_i + \sum_{I_1} y_i$) obtaining

$$M \leq \frac{B_0}{B} \left(\frac{B}{D}\right)^m \left(\frac{D_0}{D}\right)^p \left[ 2^{2m} 2^{p+p_1-1} 2^{2(m_p+1)} \frac{y_p}{|\alpha|(T)} \right]^\tau \leq 1$$

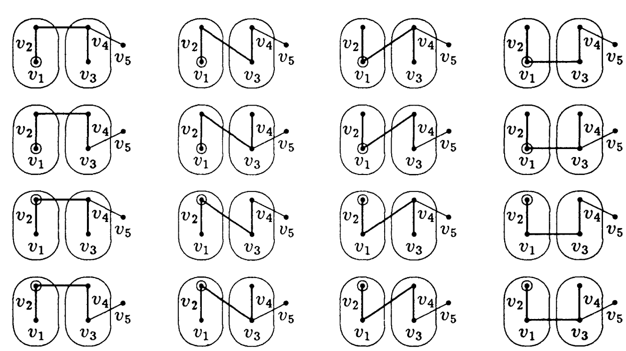
provided we take $B \equiv B_0 4^\tau$, $D \equiv 4^\tau \max\{D_0, B_0 2^{4\tau}\}$ (which satisfy also (C.23)). Thus, by (C.3), we can take $\beta_1 \equiv 3\tau$, $c_1 = \gamma 2^{6\tau} \dfrac{1+\lambda}{\lambda}$. $\square$

## D. - Two Examples of Complete Families ($k = 5$)

1)



2)



## REFERENCES


[1] V.I. ARNOLD, *Proof of A.N. Kolmogorov's theorem on the preservation of quasi-periodic motions under small perturbation of the Hamiltonian.* Uspekhi Mat. Nauk **18**, No. 5 (1963), 13-40 (Russian); English translation: Russian Math. Surveys **18**, No. 5 (1963), 9-36.

[2] V.I. ARNOLD, *Small denominators and problems of stability of motions in classical and celestial mechanics.* Uspekhi Mat. Nauk **18**, No. 6 (1963), 91-192 (Russian); English translation: Russian Math. Surveys **18**, No. 6 (1963), 85-192.

[3] A. BERRETTI - L. CHIERCHIA, *On the complex analytic structure of the golden-mean invariant curve for the standard map*. Nonlinearity **3** (1990), 39-44.

[4] A. BERRETTI - A. CELLETTI - L. CHIERCHIA - C. FALCOLINI, *Natural boundaries for area-preserving twist maps*. J. Statist. Phys., **66** (1992), 1613-1630.

[5] B. BOLLOBAS, Graph Theory. Springer-Verlag (Graduate text in mathematics, 63), Berlin-Heidelberg-New York, 1979.

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

[7] A. CELETTI - L. CHIERCHIA, *Construction of analytic KAM surfaces and effective stability bounds*, Comm. Math. Phys. **118** (1988), 119-161.

[8] L. CHIERCHIA - G. GALLAVOTTI, *Drift and Diffusion in phase space*, Preprint (1992). To appear in Ann. Inst. H. Poincaré Phys. Théor.

[9] L. CHIERCHIA - P. PERFETTI, *Second order Hamiltonian equations on* $\mathbb{T}^\infty$ *and almost-periodic solutions*. Preprint (1992). To appear in J. Differential Equations.

[10] L. CHIERCHIA - E. ZEHNDER, *Asymptotic expansions of quasiperiodic solutions*. Ann. Scuola Norm. Sup. Pisa, Cl. Sci. (4), **16** (1989), 245-258.

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

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

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

[14] C. FALCOLINI - R. DE LA LLAVE, *Numerical calculation of domains of analyticity for perturbation theories in the presence of small divisors*, J. Statist. Phys. **67** (1992).

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

[16] G. GALLAVOTTI - G. GENTILE, *Non recursive proof of the KAM theorem*. Preprint (1993).

[17] I.P. GOULDEN - D.M. JACKSON, Combinatorial Enumeration. Wyley Interscience Series in Discrete Math., 1983.

[18] F. HARARY, Graph Theory. Addison-Wesley, 1969.

[19] A.N. KOLMOGOROV, *On conservation of conditionally periodic motions under small perturbations of the Hamiltonian*. Dokl. Akad. Nauk SSSR **98**, No. 4 (1954), 527-530 (Russian).

[20] J. MOSER, *On invariant curves of area-preserving mappings of an annulus*. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II, **1962** (1962), 1-20.

[21] J. MOSER, *Convergent series expansions for quasi-periodic motions*. Math. Ann. **169** (1967), 136-176.

[22] J. MOSER, *A rapidly convergent iteration method and nonlinear partial differential equations*, I-II. Ann. Scuola Norm. Super. Pisa Cl. Sci. (3) **20** (1966), 265-315 and 499-535.

[23] H. POINCARÉ, *Les méthodes nouvelles de la mécanique céleste*. Vols. 1-3. Gauthier-Villars, Paris 1892/1893/1899.

[24] H. RÜSSMANN, *Kleine Nenner. I: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes*. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II, **1970** (1970), 67-105.

[25] W.M. SCHMIDT, *Diophantine Approximation*, Springer-Verlag (Lecture Notes in Math. 785), Berlin-Heidelberg-New York, 1980.

[26] C.L. SIEGEL, *Iterations of analytic functions*, Ann. of Math. **43** (1942), 607-612.

[27] M. VITTOT, *Lindstedt perturbation series in Hamiltonian mechanics: explicit formulation via a multidimensional Burmann-Lagrange formula*, Preprint (1991).

Dipartimento di Matematica
Università di Roma "Tor Vergata"
Via della Ricerca Scientifica
00133 Roma (Italy)