A Direct Proof of a Theorem by Kolmogorov in Hamiltonian Systems

L. Chierchia
Dipartimento di Matematica
Università di Genova
via L.B.Alberti 4, 16132 Genova (Italy)
(Internet: chierchia@mat.utovrm.it)

C. Falcolini
Dipartimento di Matematica
Università di Roma "Tor Vergata"
via della Ricerca Scientifica, 00133 Roma (Italy)
(Internet: falcolini@mat.utovrm.it)

June 8, 1993

Abstract

We present a direct proof of Kolmogorov's theorem on the persistence of quasi-periodic solutions for nearly integrable, real–analytic Hamiltonian systems with Hamiltonians of the form $\frac{1}{2}y \cdot y - \varepsilon f(x)$ where $(y, x) \in {}^N \times {}^N$ are standard symplectic coordinates. The method of proof consists in constructing, via graph theory, the formal solution as a formal power series in ε and to show that the k^{th} coefficient of such formal series can be bounded by a constant to the k^{th} power. All details are presented in a self contained way (included what is needed from the theory of graphs).

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 [14] 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 ^{N}, $x \in {}^N \equiv {}^N/(2\pi^N)$, $\varepsilon \geq 0$), provided the Hessian matrix $(\frac{\partial h}{\partial y_i \partial y_j})$ is invertible on U and provided $|\varepsilon|$ is sufficiently small. We recall that a solution $(y(t), x(t))$ is called (maximal) quasi–periodic if there exists a rationally independent vector $\omega \in {}^N$ and smooth functions $Y, X : {}^N \to {}^N$ such that $(y(t), x(t)) = (Y(\omega t), \omega t + X(\omega t))$. A detailed proof, different from that outlined by Kolmogorov in [14], was provided, in 1963, by V. I. Arnold in [1] and J. Moser [15] 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 ε–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* [17]. Poincaré, following M. Lindstedt (*Mémoires de l'Académie de Saint–Pétersbourg*, 1882), considered formal quasi–periodic solutions for which ω *depends on* ε and showed that such series are *divergent* ([17], vol. II, §XIII). 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 $n \in {}^N \backslash \{0\}$, and such factors ("small divisors") become arbitrarily small as $|n| \to \infty$. In fact (see Section 4 below), the *repeated occurrence of small divisors* ("resonances") in the k^{th} coefficient of the formal solution leads to contributions of the order of $(k!)^a$ with $a > 0$. However, in 1967 Moser [16] showed *indirectly* (see below) that the formal series *converges* leading to analytic solutions, provided the ε–*independent* vector ω satisfies certain number theoretic assumptions (verified by almost all (with respect to Lebesgue measure) vectors in ^{N}; see the "Diophantine condition" (2.7) below). This means that the k^{th} coefficient of the formal solution contains *many huge contributions which compensate among themselves* producing terms that behaves like a constant to the k^{th} power. Moser's proof, as well as *all* proofs in KAM theory (up to the 1988 Eliasson work [9]), 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 solving recursively 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 has many advantages and can be used very effectively (see [6] for accurate estimates and Appendix 2 in [8] for a five–page proof) but is indirect and hides the mechanism beyond the above mentioned compensations.

In the different context of linearization of germs of analytic diffeomorphisms, C. L. Siegel [20] 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* $n\in{}^N$ 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 [9] (which to the best of our knowledge has not been published), extended Siegel's method so as to cover the Hamiltonian case. We present a different version of Eliasson's proof based on a tree representation of the formal solutions and on the explicit exhibition of compensations of huge contributions. We follow closely [9] in its convenient reformulation of Siegel's method (see Lemma 6.1 below) 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* $H=\frac{1}{2}y\cdot y-\varepsilon f(x)$: The extension of our method to the general situation is rather straightforward but would lead to heavier notations without providing any new insight.

The *starting point* is to express the coefficients of the formal solution in terms of *labeled rooted trees*[1]. This allows to express the $k^{\rm th}$ coefficient in terms of a sum over all possible labeled rooted trees of order k, and over all possible (and "admissible") integers $\alpha_v\in{}^N\backslash\{0\}$ (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 B), f_n denote Fourier coefficients of f and δ_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 $v\in V$, $v\ne r$; 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 $\delta_v\ne0\ \forall\ v$ *i.e.* $\sum_{v'\le v}\alpha_{v'}\ne 0\ \forall\ v\in V$).

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 $k^{\rm th}$ 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

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

for suitable constants c_4 and β_4, determined below, depending on N and on the numerical properties of the vector ω; $\deg_{\mathcal{F}}v$ 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 (1.1) 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[2]. Given Eliasson's

[1] The use of graph theory in connection with formal power series is natural and very old (see *e.g.* [12] and references therein); for connections with KAM theory see [7] and [11].

[2] In [11] the problem analyzed here is also investigated with similar tools; however the families of trees considered there, as well as most of the technical aspects, are different from ours.

version of Siegel's method (which we include for completeness) and the identification of complete families, the convergence of the formal solutions follows very easily (under the Diophantine condition on ω). All the constants are computed explicitly (see Remark 6.1 below).

We close this introduction by mentioning possible directions of future researches.
(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 ε–singularities of the function X (see, *e.g.* , [6], [3], [4], [10] and references therein).
(ii) Let (x^1, x^2) with $x^i \in {}^{N_i}$ and $N_1 + N_2 = N$, let $\omega \in {}^{N_1}$ and let x_0^2 be a nondegenerate critical point of the periodic function $x^2 \in {}^{N_2} \to \int_{{}^{N_1}} f(x^1, x^2) dx^1$. Then, it is not difficult to see that (if ω 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 ([18]) that maximal quasi–periodic solutions are stable under weaker assumptions on the vector ω 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 A). The occurrence of huge terms is explicitly exhibited in Section 4. The construction of complete families is carried out in Section 5 allowing to formulate the results in Section 6 divided in four Lemmas and one Theorem. Detailed proofs are given in Section 7. Graph theory is used here mainly as a (very useful!) language and the basic definitions (enough to read our paper without any knowledge of graph theory) are presented in Appendix B.

2 Formal Quasi–Periodic Solutions

We recall the notions of quasi–periodic and formal quasi–periodic solution for a "spatially periodic" Hamiltonian.

Let $H(y, x)$ be a $C^\infty(U \times {}^N)$ Hamiltonian where U is an open set of ^{N} and ${}^N \equiv {}^N/2\pi^N$ (*i.e.* H can be seen as a function of the $2N$ variables $y_1, ... y_N$,$x_1, ..., x_N$, 2π–periodic in x_i). Consider the standard Hamilton equations:

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

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

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

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

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

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

"Le problème général de la dynamique" according to Henri Poincaré ([17], 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 ε. As we shall see below, if the Hessian matrix $h'' \equiv h_{yy} \equiv (\frac{\partial^2 h}{\partial y_i \partial y_j})$ is invertible, then there are "a lot" of "formal quasi–periodic solutions" of the equations (2.1) with Hamiltonian (2.4).

A *formal ε–power series F, over* ^{N}, is an infinite sequence of $C^\infty({}^N)$ 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^∞ 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

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

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

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

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^∞ in a neighborhood of $\{y_0\} \times {}^N$ 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* [3]

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

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

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

[3] It is well known that for any $\tau > N$, almost all (with respect to Lebesgue measure) $\omega \in {}^N$ satisfy (2.7) with some $\gamma > 0$ (see, *e.g.* , [2], chapter I, §3) while if $\tau = N - 1$ then for any $\omega \in {}^N$ there exist a $\gamma > 0$ and an infinite number of $n \in {}^N$ such that (2.7) is violated (this is a theorem by Dirichlet, see, *e.g.* , [19]).

Proof 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 ^{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 $y_0, Y^{(1)},..., Y^{(k-1)}$, $X^{(1)},...,X^{(k-1)}$, are smooth functions over ^{N} solving (2.6) with $\int_N X^{(h)}=0$. We claim that

$$\int_N \phi^{(k)}(\theta)\equiv\int_N\Big[f_x(\sum_{h=0}^{k-1}\varepsilon^h Y^{(h)},\theta+\sum_{h=0}^{k-1}\varepsilon^h X^{(h)})\Big]_{k-1}=0 \tag{2.9}$$

(note that the claim is obvious for $k=1$ as in such a case (2.9) 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^∞ (vector–valued) function over ^{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

$$D^{-1}\phi^{(k)}\equiv\sum_{\substack{n\in{}^N\\ n\neq 0}}\frac{\phi_n^{(k)}}{i\omega\cdot n}e^{in\cdot\theta} \tag{2.10}$$

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 (2.10), ensuring that $D^{-1}\phi^{(k)}$ is $C^\infty({}^N)$). 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 ^{N}; rewriting such equation we get

$$\begin{aligned} DX^{(k)} &= h_{yy}(y_0)Y^{(k)}+[h_y(\sum_{h=0}^{k-1}\varepsilon^h Y^{(h)})]_k+\Big[f_y(\sum_{h=0}^{k-1}\varepsilon^h Y^{(h)},\theta+\sum_{h=0}^{k-1}\varepsilon^h X^{(h)})\Big]_{k-1}\\ &\equiv h_{yy}(y_0)c_k+\psi^{(k)} \end{aligned} \tag{2.11}$$

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

$$c_k=h_{yy}(y_0)^{-1}\int_N\psi^{(k)} \tag{2.12}$$

in which case (2.11) has a unique solution with $\int_N X^{(k)}=0$.
It remains to prove (2.9). Observe that for any (smooth) functions $\bar X,\bar Y$ over ^{N} one has

$$\int\Big\{(D\bar Y+H_x(\bar Y,\theta+\bar X))\ (I+\partial_\theta\bar X)\ -\ (\omega+D\bar X-H_y(\bar Y,\theta+\bar X))\ \partial_\theta\bar Y\Big\}d\theta=0 \tag{2.13}$$

where $\partial_\theta\bar X$ is the matrix $(\partial_\theta\bar X)_{ij}\equiv\partial_{\theta_j}\bar X_i$ 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 (2.13) follows immediately if one notices that $(H_x)(I+\partial_\theta\bar X)+(H_y)(\partial_\theta Y)$ is just the θ–gradient of $\theta\to H(\bar Y,\theta+\bar X)$ (so that its average vanishes) and that the remaining terms in (2.13) disappear by integration by parts. Now, we use (2.13) with $\bar Y=\sum_{h=0}^{k-1}\varepsilon^h Y^{(h)}$, $\bar X=\sum_{h=1}^{k-1}\varepsilon^h X^{(h)}$ to get an identity in ε; differentiating such identity with respect to ε 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 $h\leq k-1$ we easily obtain (2.9). ∎

3 Tree expansion of formal series

A very explicit representation of the formal solutions described in Proposition 2.1 can be obtained by mean of *trees*, as explained below. For a different representation see [21].

From now on we restrict our attention on Hamiltonian (2.4) *of the simple form*

$$H(y,x;\varepsilon) = \frac{1}{2}\, y\cdot y \;-\; \varepsilon f(x)\ . \tag{3.1}$$

Such a choice (besides the harmless change of sign in front of the perturbation f) has the advantage of simplifying sensibly the notations without introducing any essential modification: in other words, all the arguments and proofs presented below could be extended, at the expense of a heavier formalism, to cover the general case (2.4) (with $\det h'' \neq 0$).

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

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

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

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

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

We recall that a tree T is a connected acyclic graph (see Appendix B for the fundamentals used here or see any introductory book on graph theory such as [13] 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 η 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

$$u \geq v \qquad \text{or} \qquad u \geq_{T_r} v \tag{3.4}$$

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 $\omega \in\ ^N$ satisfying (2.7) and denote the inner product between $n \in\ ^N$ and ω by

$$\langle n \rangle \equiv \omega \cdot n \equiv \sum_{i=1}^{N} \omega_i n_i \tag{3.5}$$

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

$$\delta_v(T_r;\alpha) \equiv \langle \sum_{\substack{v' \in V \\ v' \leq v}} \alpha_{v'} \rangle \tag{3.6}$$

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

$$\alpha_v \neq 0 \; ; \qquad \sum_{v' \leq v} \alpha_{v'} \neq 0 \tag{3.7}$$

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 (see Appendix B for more informations); and, for any integer–valued function α and any subset $S \subset V$ of vertices of a tree T, we denote

$$\alpha(S) \equiv \sum_{v \in S} \alpha_v \tag{3.8}$$

Proposition 3.1 *The j^{th} component of the n–Fourier coefficients of $X^{(k)}$ is given by*

$$\begin{aligned} X^{(k)}_{jn} &= \frac{1}{k!} \sum_{T_r \in \mathcal{T}^k} F_{jn}(T_r) \; , \qquad \text{with :} \\ F_{jn}(T_r) &\equiv (-i) \sum_{\substack{\alpha \in \mathcal{A}(T_r) \\ \alpha(T_r)=n, \alpha_\eta \equiv e_j}} \prod_{v \in V} f_{\alpha_v} \prod_{vv' \in E} \alpha_v \cdot \alpha_{v'} \prod_{v \in V} \delta_v^{-2} \end{aligned} \tag{3.9}$$

where e_j is the unit vector with all zeroes except in the j^{th} entry.

Note that in the above formula $E = E(T_r)$ *includes* the edge ηr and for this reason the function α has been extended on η (which is outside $V = V(T_r)$).

Remark 3.1 The above function $F^{(k)}_{jn}$ 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, (3.9) can be rewritten as follows. Let $\widetilde{T}_r \equiv [T_r]$ denote the rooted tree obtained by removing the labels from T_r (clearly $\widetilde{T}_r$ can be seen as the equivalence class generated by T_r), and let $\ell(\widetilde{T}_r)$ denote the number of ways of putting k labels on $\widetilde{T}_r$ (*i.e.* $\ell(\widetilde{T}_r) = \#\{T'_r \in \mathcal{T}^k : T'_r \in \widetilde{T}_r\}$). For example, if $\widetilde{T}_r$ is the path of order k rooted in one of its endpoints, then $\ell(\widetilde{T}_k) = k!$. The "$\widetilde{T}_r$–admissible" class of ^{N}–valued functions is defined in exactly the same way (see (3.7) replacing T_r by $\widetilde{T}_r$ (as the labeling did not enter in the definition of $\mathcal{A}(T_r)$). Finally denote by $\widetilde{\mathcal{T}}^k$ the class of all rooted trees of order k. With this notations we see that

$$\begin{aligned} X^{(k)}_{jn} &= \sum_{\widetilde{T}_r \in \widetilde{\mathcal{T}}^k} \frac{\ell(\widetilde{T}_r)}{k!} \varphi_{jn}(\widetilde{T}_r) \; , \qquad \text{with :} \\ \varphi_{jn}(\widetilde{T}_r) &\equiv (-i) \sum_{\substack{\alpha \in \mathcal{A}(\widetilde{T}_r) \\ \alpha(\widetilde{T}_r)=n, \alpha_\eta \equiv e_j}} \prod_{v \in V} f_{\alpha_v} \prod_{vv' \in E} \alpha_v \cdot \alpha_{v'} \prod_{v \in V} \delta_v^{-2} \end{aligned} \tag{3.10}$$

Below it will be useful to exchange the sums (over trees and over integers α's) in (3.9), in which case we obtain the following formula:

$$X_{jn}^{(k)} = \frac{1}{k!} \sum_{\substack{n_1,\dots,n_k \\ n_i \in {}^N\setminus\{0\}, \sum n_i = n}} \sum_{T_r \in \mathcal{T}^k} \bar{F}_{j\{n_i\}}(T_r) \tag{3.11}$$

where if $v_1, \dots, v_k$ 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 $\bar{F}$ is defined by

$$\bar{F}_{j\{n_i\}}(T_r) \equiv 0\ , \qquad \text{if} \quad \exists\ v : \sum_{v' \leq v} \alpha_{v'} = 0 \tag{3.12}$$

and otherwise by

$$\bar{F}_{j\{n_i\}}(T_r) \equiv \prod_{v \in V} f_{\alpha_v} \prod_{vv' \in E} \alpha_v \cdot \alpha_{v'} \prod_{v \in V} \delta_v^{-2} \tag{3.13}$$

Expansions like (3.9) are quite general; another tree expansion, which we shall need later, for a series satisfying an equation simpler than (but related to) (3.3) is described in the following

Proposition 3.2 *Let $n \in {}^N \to \phi_n \in$ decay faster than any power of $|n|$ (i.e. $\forall\ s > 0$, $\sup_n |n|^s |\phi_n| < \infty$) and let $g = g(\varepsilon) \sim \sum_{k \geq 1} g_k \varepsilon^k$ be the unique formal solution of the equation:*

$$g \sim \varepsilon \sum_{n \in {}^N} \phi_n |n| \exp(|n| g)\ . \tag{3.14}$$

Then

$$\begin{aligned} g_k &= \frac{1}{k!} \sum_{T_r \in \mathcal{T}^k} \sum_{\substack{\alpha : T_r \to {}^N \\ \alpha_\eta \equiv 1}} \prod_{v \in V} \phi_{\alpha_v} \prod_{vv' \in E} |\alpha_v|\ |\alpha_{v'}| \\ &= \frac{1}{k!} \sum_{T_r \in \mathcal{T}^k} \sum_{\alpha : T_r \to {}^N} \prod_{v \in V} \phi_{\alpha_v} \prod_{v \in V} |\alpha_v|^{\deg v} \end{aligned} \tag{3.15}$$

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

Actually, if the ϕ_n's decay exponentially (*i.e.* $|\phi_n| \leq M \exp(-\xi |n|)$ for some $M, \xi > 0$ and all $n \in {}^N$), it follows immediately from the (analytic) Implicit Function Theorem (applied to the analytic function of two complex variables $(x, \varepsilon) \in {}^2 \to G(x, \varepsilon) \equiv x - \varepsilon \sum \phi_n |n| \exp(|n| x)$) that g converges absolutely near $\varepsilon = 0$ (see also Subsection 7.1 below).

The proofs (by induction on k) of Proposition 3.1 and Proposition 3.2 are of pure combinatorial character. Here we outline the main steps and refer to Appendix A for complete details. The proofs of the two Propositions are very similar but, obviously, the proof of Proposition 3.2 is simpler and we discuss it first.

The starting point is so to get a recursive formula for the coefficients g_k defined implicitly by (3.14): Expanding the right hand side of (3.14) in Taylor series and comparing equal powers of ε one immediately obtain

$$\begin{aligned} g_1 &= \sum_{n\in N} \phi_n |n| \\ g_k &= \sum_{j=2}^{k} \sum_{n\in N} \phi_n \, \frac{|n|^j}{(j-1)!} \sum_{\substack{h_1+\cdots h_{j-1}=k-1 \\ h_i\geq 1}} \prod_{i=1}^{j-1} g_{h_i} \qquad (k\geq 2) \end{aligned} \tag{3.16}$$

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

Definition 3.1 Let $s\geq 1$, let $\widetilde{T}_i \in \widetilde{\mathcal{T}}^{h_i}$ where $i=1,...,s$ and $h_i\geq 1$ with $h\equiv \sum_{i=1}^s h_i$. Let $\widetilde{T}_i^0 \in \widetilde{\mathcal{T}}_0^{h_i}$ be the (unrooted) tree obtained from $\widetilde{T}_i$ by not distinguishing the root. We define the rooted tree $\widetilde{T}_1 * \cdots * \widetilde{T}_s \in \widetilde{\mathcal{T}}^{h+1}$ by setting[4]

$$\widetilde{T}_1 * \cdots * \widetilde{T}_s \equiv \Big(\widetilde{T}_1^0 \cup \cdots \cup \widetilde{T}_s^0 \cup \{r\} + \sum_{i=1}^{s} rr_i\Big)_r \tag{3.17}$$

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

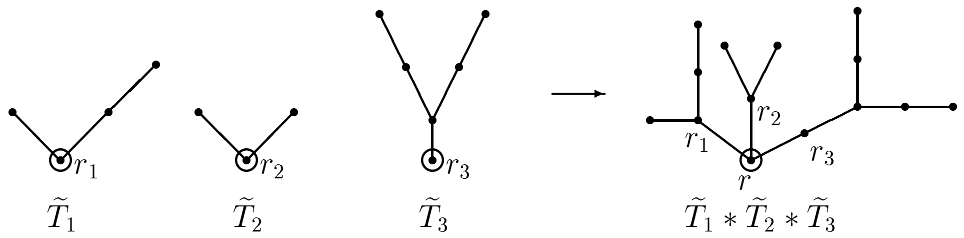

Figure 1: The $*$ operation

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

Lemma 3.1 *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*

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

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

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

[4]For the (standard) notation see Appendix B

The proof is given in Appendix A. 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 3.1 *Let* $G : \widetilde{\mathcal{T}}^k \to$. *Then*

$$\frac{1}{k!}\sum_{T\in\mathcal{T}^k} G(\widetilde{T}) = \sum_{j=2}^{k}\frac{1}{(j-1)!}\sum_{\substack{h_1+\cdots h_{j-1}=k-1\\ h_i\geq 1}}\prod_{i=1}^{j-1}\frac{1}{h_i!}\sum_{T_1\in\mathcal{T}^{h_1},\dots,T_{j-1}\in\mathcal{T}^{h_{j-1}}} G(\widetilde{T}_1 * \cdots * \widetilde{T}_{j-1}) \tag{3.18}$$

¿From (3.18), Proposition 3.2 will follow at once. To get also Proposition 3.1 from Corollary 3.1, one needs only to rewrite (3.9) properly. To do this, let for any $m, n \in {}^N\backslash\{0\}$ and $k \geq 1$

$$\xi^m \equiv i\ m\cdot X\ ,\quad \xi^{(m,k)} \equiv i\ m\cdot X^{(k)}\ ,\quad \xi_n^{(m,k)} \equiv i\ m\cdot X_n^{(k)} \tag{3.19}$$

Then from (3.3) and Taylor expansion (in ε) we obtain easily

$$\begin{aligned}
\xi_n^{(m,1)} &= -\langle n\rangle^{-2}\ m\cdot n\ f_n\\
\xi_n^{(m,k)} &= -\langle n\rangle^{-2}\sum_{j=2}^{k}\sum_{n\in{}^N\backslash\{0\}}\sum_{\substack{n_1+\cdots+n_j=n\\ n_i\in{}^N\backslash\{0\}}} m\cdot n_j\ f_{n_j}\ \times\\
&\qquad\times\frac{1}{(j-1)!}\sum_{\substack{h_1+\cdots h_{j-1}=k-1\\ h_i\geq 1}}\prod_{i=1}^{j-1}\xi_{n_i}^{(n_j,h_i)}\quad (k\geq 2)
\end{aligned} \tag{3.20}$$

The similarity of (3.20) and (3.16) is transparent[5] and it is enough to change the extensions of the α functions on the earth η to generalizes (3.9) to

$$\begin{aligned}
\xi_n^{(m,k)} &= \frac{1}{k!}\sum_{T_r\in\mathcal{T}^k} F_{mn}\ ,\qquad \text{with :}\\
F_{mn}(T_r) &\equiv \sum_{\substack{\alpha\in\mathcal{A}(T_r)\\ \alpha(T_r)=n,\alpha_\eta\equiv m}}\prod_{v\in V} f_{\alpha_v}\prod_{vv'\in E}\alpha_v\cdot\alpha_{v'}\prod_{v\in V}\delta_v^{-2}
\end{aligned} \tag{3.21}$$

For full details see Appendix A.

4 Divergences

From now on f *in* (3.3) *will be assumed real–analytic.* In this section we illustrate the well known mechanism of divergences of series with coefficients of type (3.9), (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 δ_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.

[5] $|n|\phi_n$ corresponds to $\langle n\rangle^{-2} m\cdot n_j f_{n_j}$, while $|n|g_{h_i}$ corresponds to $\xi_{n_i}^{(n_j,h_i)}$.

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

$$A^{(k)}_{\bar{\jmath}n} = \frac{1}{k!} \sum_{T_r \in \mathcal{T}^k} \sum_{\substack{\alpha \in \mathcal{A}(T_r) \\ \alpha(T_r)=n, \alpha_\eta \equiv e_{\bar{\jmath}}}} \mid \prod_{v \in V} f_{\alpha_v} \prod_{vv' \in E} \alpha_v \cdot \alpha_{v'} \mid \prod_{v \in V} \delta_v^{-2} \tag{4.1}$$

we want to show that

$$\sup_{\bar{\jmath},k,n} A^{(k)}_{\bar{\jmath}n} \ e^{-\sigma|n|} \ \varepsilon_0^k = +\infty \ , \qquad \forall \ \varepsilon_0, \ \sigma > 0 \tag{4.2}$$

where in the supremum $1 \leq \bar{\jmath} \leq N$, $k \geq 1$ and $n \in {}^N$. We shall prove (4.2) in the case

$$\begin{aligned} &N = 2 \ , \qquad f(x_1, x_2) \equiv \ \cos x_1 \ + \cos(x_2 - x_1) \ , \\ &\omega_2 > \omega_1 > 0 \ , \qquad \omega_1/\omega_2 \quad \text{irrational} \end{aligned} \tag{4.3}$$

since it will be clear that the same proof can easily be adapted to the general analytic case *as long as f has two Fourier coefficients with linearly independent (in ^{N}) indeces.*

Proof of (4.2) **in case** (4.3) ¿From an elementary number theoretical theorem by Dirichlet (see [19]) it follows that there exist $p_j, q_j \geq 1$, with $q_j \nearrow \infty$ such that

$$\mid \frac{\omega_1}{\omega_2} q_j \ - \ p_j \mid < \frac{1}{q_j} \tag{4.4}$$

and we can assume without loss of generality that for all $j \geq 1$

$$q_j > \frac{\omega_2}{\omega_2 - \omega_1} \tag{4.5}$$

Observing the the left hand side of (4.4) is bigger or equal than $p_j - q_j(\omega_1/\omega_2)$ we see immediately that (4.5) implies that

$$p_j < q_j \tag{4.6}$$

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 α). Let P be the path of order k, $u_k u_{k-1} ... u_1$ rooted at u_k: $u_k = r$ and $u_k > u_{k-1} > ... > u_1$. Let $s_j \equiv q_j - p_j$ and, for any $h \geq 0$ let $k \equiv q_j + 2h$:

$$q_j = p_j + s_j \ , \qquad 1 \leq s_j, p_j < q_j \ , \qquad k = q_j + 2h \tag{4.7}$$

Let $\alpha_{u_i} \equiv n_i \in {}^2$ be defined as follows:

$$n_i \equiv \begin{cases} (1, 0) & \text{if} \ \ 1 \leq i \leq s_j \\ (1, -1) & \text{if} \ \ s_j < i \leq q_j \\ (1, 0) & \text{if} \ \ i = q_j + 1, q_j + 3, ..., q_j + (2h - 1) \\ (-1, 0) & \text{if} \ \ i = q_j + 2, q_j + 4, ..., q_j + 2h \end{cases} \tag{4.8}$$

To this choice of Fourier indeces α, for which

$$n \equiv \sum_{i=1}^{k} n_i = (q_j, -p_j) \tag{4.9}$$

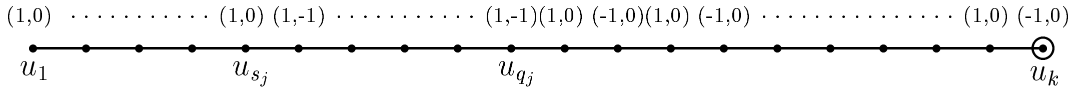

Figure 2: Divergent contributions

correspond the divisors

$$|\delta_{u_i}| = \begin{cases} \omega_1 \ i & \text{if} \ \ 1 \le i \le s_j \\ |\omega_1 i - \omega_2 (i - s_j)| & \text{if} \ \ s_j < i \le q_j \\ |\omega_1 q_j - \omega_2 p_j + \omega_1| & \text{if} \ \ i = q_j + 1, q_j + 3, ..., q_j + (2h-1) \\ |\omega_1 q_j - \omega_2 p_j| & \text{if} \ \ i = q_j + 2, q_j + 4, ..., q_j + 2h \end{cases} \tag{4.10}$$

¿From (4.4), (4.5) and (4.10) it follows that

$$\begin{aligned} & |\omega_1 q_j - \omega_2 p_j + \omega_1| < \omega_2 \\ & |\delta_{u_i}| < 2\omega_2 \ i \ , \qquad \text{for} \quad 1 \le i \le q_j \end{aligned} \tag{4.11}$$

which, together with (4.4) yields

$$\prod_{i=1}^{k} \delta_{u_i}^{-1} > \omega_2^{-k} \ 2^{-q_j} \ \frac{q_j^h}{q_j!} \tag{4.12}$$

Furthermore, since for (4.3) it is

$$|f_{\pm(1,0)}| = |f_{\pm(1,-1)}| = \frac{1}{2} \tag{4.13}$$

we have also

$$\prod_{i=1}^{k} |f_{n_i}| = 2^{-k} \ , \qquad \prod_{E(P)} |\alpha_v \cdot \alpha_{v'}| \ge 1 \tag{4.14}$$

having set $\bar{\jmath} = 1$ *i.e.* $\alpha_\eta \equiv (1,0)$. We now choose

$$h = 3q_j \ , \qquad \text{so} \quad \text{that} \qquad k \equiv k_j = 7q_j \tag{4.15}$$

Then, if in the sum (4.1) (which is finite in the case of (4.3)), 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 $k!$, we obtain (recall that $\bar{\jmath} = 1$ and that $n = (q_j, -p_j)$)

$$A_{1n}^{(k_j)} \ > 2^{-k_j} \ (2\omega_2)^{-2k_j} \ (k_j!)^{4/7} \tag{4.16}$$

from which (4.2) follows at once. Notice that by increasing h, the exponent of $k_j!$ in (4.16) becomes arbitrarily close to 1. ∎

5 Resonances

Repetitions of small divisors correspond to *resonant subtrees i.e.* to subtrees S of a given (labeled rooted) tree T for which $\alpha(S) \equiv \sum_{v\in S} \alpha_v = 0$: In the example of Section 4 resonant subtrees are the subtrees with endpoints u_{q_j+s} and u_{q_j+s+1} for $1 \le s \le k-1$, or the subtrees with endpoints u_{q_j+s} and u_{q_j+s+3} *etc.* 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 $\alpha : V \equiv \{v_1, ..., v_k\} \to {}^N$, $\{\alpha_{v_i}\} \equiv \{n_i\}$, is *admissible* in the usual sense that $n_i \neq 0$ and $\sum_{v' \le v} \alpha_v \neq 0$ for all $v \in V$ (in which case $\delta_v \neq 0$ by the rational independence of ω). In this section we develop the tools needed to *make a partition* of $\mathcal{T}^k$, for a given choice of $\{n_i\}$ (see (3.11)$\div$ (3.13)), 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 (3.11), coming from the sum over trees belonging to the same class of the partition, by a constant to the k^{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 $\alpha_{v_i} \equiv n_i$, $v_1, ..., v_k$ being the labels of $\mathcal{T}^k$ (see Appendix B for the fundamentals from graph theory used here). Let us begin (with a bit of patience).

Definition 5.1 (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 $T \backslash S$ (if T is rooted and $r \in S$ the edge ηr has to be counted).

Definition 5.2 (Resonances) Given a rooted tree T and an admissible function α 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.*

Definition 5.3 (Resonant couples) Given a rooted tree T, $\lambda > 0$ we say that a couple of points $(u, w) \in V \times V$ is λ*–resonant* if: (i) $u > w$; (ii) $\delta_u = \delta_w$; (iii) u and w are not adjacent and $|\delta_w| \le \lambda |\delta_v|$ for all v between u and w.

Such a notion is given also in [9] where the couple (u, w) is called a critical resonance; we reserve such a name for a different object (see Definition 5.5 below).

Definition 5.4 (λ–Resonances) Given a rooted tree T, $\lambda > 0$ and a resonance $R \subset T$ let

$$n \equiv \sum_{v<R} \alpha_v \ , \qquad \Delta(R) \equiv \min_{\substack{R' \subset R \\ \alpha(R') \neq 0}} |\langle \alpha(R') \rangle| \tag{5.1}$$

where $v < R$ means $v < u$ for all $u \in R$. We say that R is a λ*–resonance* (or a λ–resonant subtree) if

$$|\langle n \rangle| \le \lambda \Delta(R) \ . \tag{5.2}$$

Remark 5.1 To two points $u > w$ such that $\delta_u = \delta_w$ but $\delta_v \neq \delta_u$ for any v between u and w (if there are any) it is always possible to associate a resonant subtree $R \equiv R(u,w)$:

$$R(u,w) \equiv \{v \in V: \ v \geq u\} \backslash \{v \in V: \ v \geq w\} \tag{5.3}$$

Notice however that (u,w) may be λ–resonant while the associated resonance $R(u,w)$ is not λ–resonant as shown in the example in Figure 3: 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 $\delta_u = \delta_w$) and $|\langle n\rangle| \leq \lambda |\langle n_2+n_3+n\rangle|$ (implying (u,w) λ–resonant) but $|\langle n\rangle| > \lambda |\langle n_3\rangle|$ (implying $R(u,w)$ not λ–resonant).

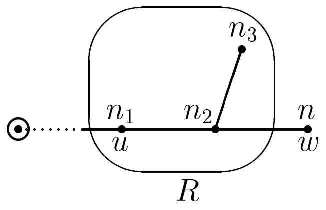

Figure 3: $R = R(u,w)$

Proposition 5.1 (Non overlapping of resonant couples) *Let (u_i, w_i) for $i = 1,2$ be couples of λ_i–resonant points. If $u_1 > u_2 > w_1 > w_2$ then either $\lambda_1 \geq 1$ or $\lambda_2 \geq 1$.*

Proof By contradiction: Assume both λ_1 and λ_2 are less than one.Then:

$$|\delta_{u_2}| = |\delta_{w_2}| \leq \lambda_2 |\delta_{w_1}| < |\delta_{w_1}| \ , \qquad |\delta_{w_1}| \leq \lambda_1 |\delta_{u_2}| < |\delta_{u_2}| \tag{5.4}$$

which is absurd. ∎

Proposition 5.2 (Non overlapping of resonances) *Let R_i for $i = 1,2$ be λ_i–resonances which are not one subtree of the other and with nonempty 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 4.

In drawing pictures we are using the following

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

Since R_i are resonant subtrees $n_2 + n_3 = 0$, $n_3 + n_4 = 0$ in case (a) while $n_1 = -n_2$ in case (b) and (c). We proceed now by contradiction: Assume $\lambda_i < 1/2$.

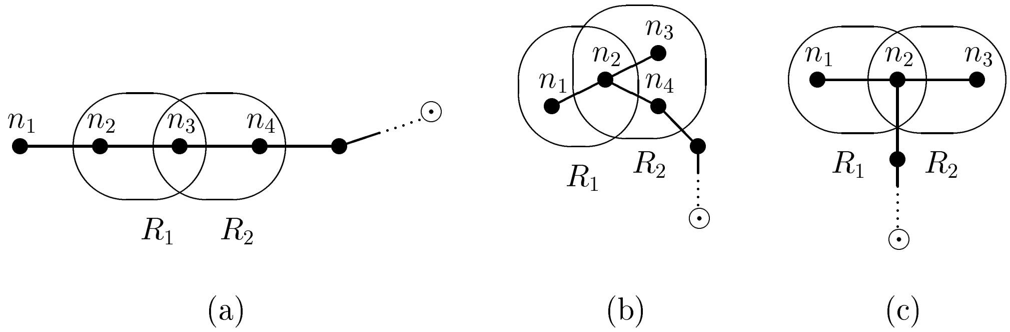

Figure 4: Intersections of resonances

In case (a)

$$|\langle n_1 \rangle| \leq \lambda_1 |\langle n_2 \rangle| \ , \qquad |\langle n_1 + n_2 \rangle| \leq \lambda_2 |\langle n_3 \rangle| = \lambda_2 |\langle n_2 \rangle| \tag{5.5}$$

and

$$|\langle n_1 + n_2 \rangle| \geq (1 - \lambda_1)|\langle n_2 \rangle| \qquad \Longrightarrow \qquad |\langle n_2 \rangle| \leq \frac{\lambda_2}{1 - \lambda_1} \, |\langle n_2 \rangle| < |\langle n_2 \rangle| \tag{5.6}$$

which is absurd.

In case (b) and (c): $|\langle n_1 \rangle| \leq \lambda_1 |\langle n_2 \rangle| = \lambda_1 |\langle n_1 \rangle|$ which is absurd because $\lambda_1 < 1$. ∎

Definition 5.5 (Critical Resonances) A resonance R is called *critical* (or λ–critical) if it is λ–resonant with $\lambda < 1/2$ and if it is maximal *i.e.* it is not properly contained into another λ–resonant subtree.
Obviously: (i) Critical resonances cannot intersect (by definition and by Proposition 5.2). (ii) Critical resonances may contain λ'–resonances with any λ'.

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 5.6 (λ–Subresonances) Let R' be a null subtree of degree two of a λ–critical resonance R and let $u_1 v_1$, $u_2 v_2$, 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

$$m \equiv \sum_{\substack{v \in R_{v_1} \\ v \leq u_1}} \alpha_v = - \sum_{\substack{v \in R_{v_2} \\ v \leq u_2}} \alpha_v \tag{5.7}$$

where R_{v_i} is the rooted tree R with root in v_i and the order $\leq$ in each sum is relative to R_{v_i} (we are using the standard convention that $v \in T$ means $v \in V(T)$). The integer $m \in {}^N \backslash \{0\}$ is defined up to sign (as the roles of the v_i can be exchanged). We then say that R' is a λ–*subresonance* if

$$|\langle m \rangle| \leq \lambda \, \Delta(R') \tag{5.8}$$

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

Proposition 5.3 (Non overlapping of subresonances) *Let R_i' for $i = 1, 2$ be λ–subresonances of a λ–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 5.2 and is left to the reader.

Definition 5.7 (Hierarchies of critical subresonances) Let R be a λ–critical resonance (*i.e.* R is a maximal λ–resonance with $\lambda < 1/2$). Let R_1 be a maximal λ–subresonance (if it exists) of R (maximal means that R_1 is not properly contained in another λ–subresonance of R). Note that, by Proposition 5.3, R_1 cannot overlap with another λ–subresonance of R. We can now *define subresonances of R_1 replacing, in Definition 5.6, R with R_1*. We then let R_2 be a maximal λ–subresonance (if it exists) of R_1. And so on, till R_h, $h \geq 1$, does not contain any λ–subresonance. We consider also the case in which R does not contain any λ–subresonance setting in such a case $h = 0$ and $R_0 \equiv R$. The sequence

$$\mathcal{H} \equiv \mathcal{H}_\lambda(R) \equiv \{R_1, ..., R_h\} \ , \qquad R \supset R_1 \supset \cdots \supset R_h \tag{5.9}$$

is called a *critical hierarchy of λ–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 7 below).

Remark 5.2 In general a critical λ–subresonance need not be a λ–resonance as shown by the following example.

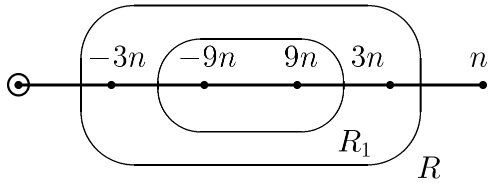

Figure 5: R_1 is not a $\frac{1}{3}$–resonance

Fix $\lambda = 1/3$ and let $n \in {}^N\backslash\{0\}$. Then one checks immediatly that R, in Figure 5, is a λ–critical resonance and that $\mathcal{H}_\lambda(R) = \{R_1\}$. However the critical subresonance R_1 *is not a λ–resonance* since $|\langle 4n\rangle| > \lambda\Delta(R_1) = 3|\langle n\rangle|$.

Remark 5.3 If R' is a λ–critical subresonance then $R' = R_i$ for some $1 \leq i \leq h$ where $\{R_1, ..., R_h\}$ is a λ–hierarchy associated to some λ–critical resonance R. To each R_j we can associated (up to sign) an integer $m_j \in {}^N\backslash\{0\}$ as in (5.7) (see Figure 6).
Then $|\langle m_j\rangle| \leq \lambda\Delta(R_j)$ for any $j = 1, ..., h$ and in particular $|\langle m_j\rangle| \leq \lambda|\langle m_{j+1}\rangle|$ so that

$$|\langle m_i\rangle| \leq \lambda^{j-i}|\langle m_j\rangle| \ , \qquad \forall \ 1 \leq i \leq j \leq h \tag{5.10}$$

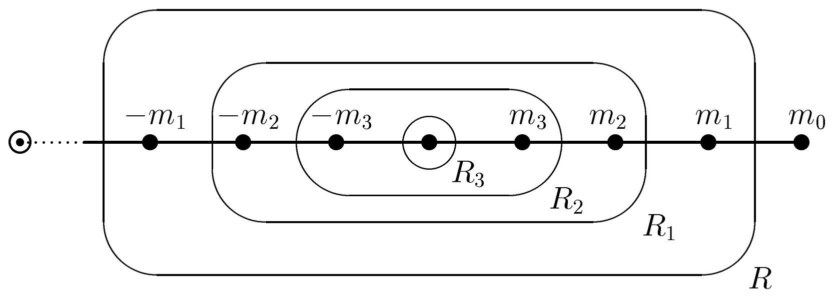

Figure 6: A hierarchy of resonances with $h = 3$

and

$$\sum_{i=1}^{j} |\langle m_i \rangle| < \frac{1}{1-\lambda} \, |\langle m_j \rangle| \leq \frac{\lambda}{1-\lambda} \Delta(R_j) \tag{5.11}$$

Also, if σ_h is plus or minus one and $\sigma_i = 0$ or ± 1 $(i = 1, ...h-1)$ then

$$|\langle \sum_{i=1}^{h} \sigma_i m_i \rangle| \geq \frac{1-2\lambda}{1-\lambda} |\langle m_h \rangle| \tag{5.12}$$

Furthermore, if we let $m_0 \equiv \sum_{v' < R} \alpha_{v'}$ (since $|\langle m_0 \rangle| \leq \lambda |\langle m_1 \rangle|$ we get as above:

$$\sum_{i=0}^{h} |\langle m_i \rangle| < \frac{1}{1-\lambda} |\langle m_h \rangle| \leq \frac{\lambda}{1-\lambda} \Delta(R_h) \tag{5.13}$$

Definition 5.8 (The set of critical (sub)resonance of T) We shall often use the word "(sub)resonance" to indicate either a resonance or a subresonance. Given a rooted tree T, an admissible α and a $\lambda < 1/2$ we let

$$\mathcal{R} \equiv \mathcal{R}(T) \equiv \{ \text{ all } \lambda-\text{critical resonances and } \lambda-\text{critical subresonances of T} \} \tag{5.14}$$

and

$$\mathcal{R}_0 \equiv \mathcal{R}_0(T) \equiv \{ \text{ all } \lambda-\text{critical resonances of T} \} \tag{5.15}$$

If $2 < p \equiv |\mathcal{R}_0| \equiv$ 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 λ–critical subresonances, otherwise

$$h_i \equiv \max_{\mathcal{H}_\lambda \subset R^i} |\mathcal{H}_\lambda| \,, \qquad |\text{ set}| \equiv \text{ cardinality of the set} \tag{5.16}$$

See Figure 7 for an example of critical resonances and subresonances. (In Figure 7 $\mathcal{R}_0 = \{R^1, R^2\}$ and $R_1, ..., R_9$ are critical subresonances; notice that R^1 contains five hierarchies, three of which have $h = 1$).

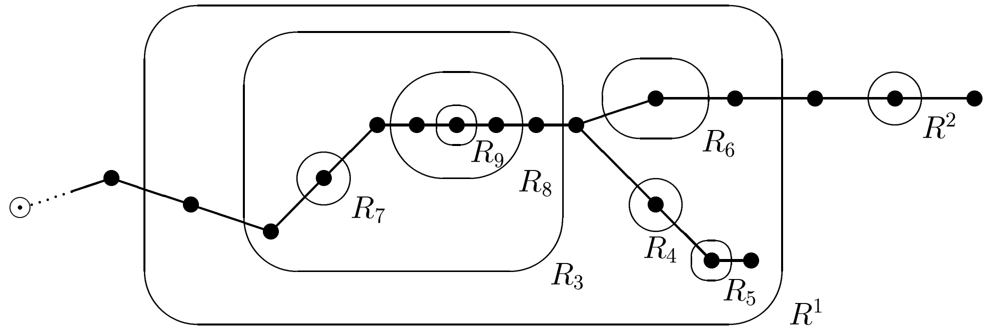

Figure 7: Critical (sub)resonances

Definition 5.9 (Coresonances) To any λ–critical (sub)resonance R in $\mathcal{R}$ we associate *its coresonance*

$$\overline{R} \equiv R \backslash \{ \text{ all } \lambda\text{−critical subresonances contained in R} \} \tag{5.17}$$

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

$$\alpha(\overline{R}) \equiv \sum_{v \in \overline{R}} \alpha_v = 0 \ . \tag{5.18}$$

We denote

$$\overline{\mathcal{R}} \equiv \bigcup_{R \in \mathcal{R}} \overline{R} \tag{5.19}$$

the set of all coresonances. Finally, we let

$$\overline{T} \equiv T \backslash \bigcup_{R \in \mathcal{R}_0} R \tag{5.20}$$

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 η if $r \in R \in \mathcal{R}_0$) and of one or more subtrees.

Remark 5.4 $R \in \mathcal{R}$ is connected to $T \backslash R$ by two edges $u\bar{u}$, $v\bar{v}$ with $u, v \in T \backslash R$ ($u \neq v$ and either u or v may be the earth η) and with $\bar{u}, \bar{v} \in \overline{R}$ ($\bar{u}, \bar{v}$ may coincide).

Definition 5.10 (Subtree contractions) Let T ba 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 $u\bar{u}$, $v\bar{v}$, ($\bar{u}, \bar{v} \in R$), be the two edges connecting R with $T \backslash R$; we then set

$$T/R \equiv T \backslash R + uv \tag{5.21}$$

See Figure 8 for examples.

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

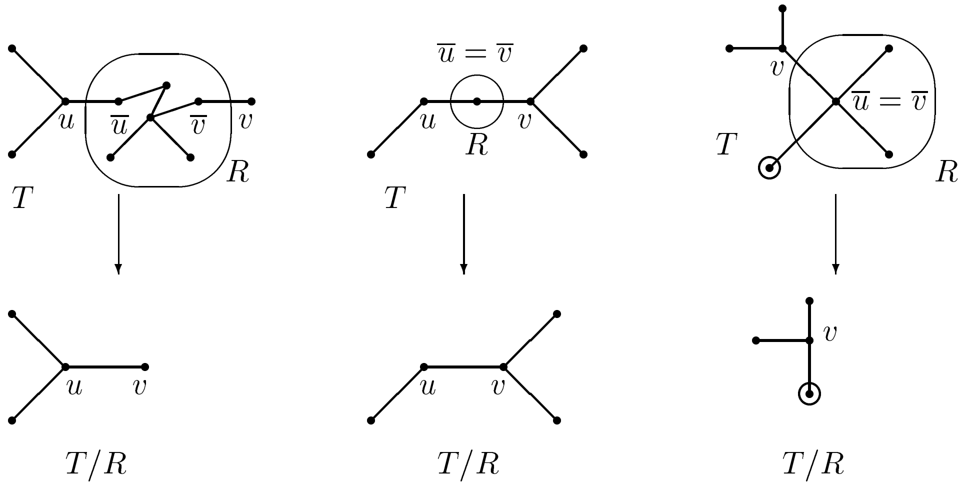

Figure 8: Tree contractions

Definition 5.11 (Critical chains and connecting points) *A (λ)–critical chain C is an ordered s–ple of coresonances,* $(\overline{R}_1, ..., \overline{R}_s)$, such that $\overline{R}_i$ is connected with $\overline{R}_{i+1}$ for all $1 \le i \le s-1$ and such that $\overline{R}_1$ and $\overline{R}_s$ are also connected with either $\overline{T}$ or with a coresonance $\overline{R} \in \overline{\mathcal{R}}$ which contains the whole chain C (*i.e.* $R_1 \cup ... \cup R_s \subset R$). 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 $C \equiv (\overline{R}_1, ..., \overline{R}_s)$ 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

$$\mathcal{C} \equiv \{ \text{ the set of all chains } \} = \overline{\mathcal{R}} \tag{5.22}$$

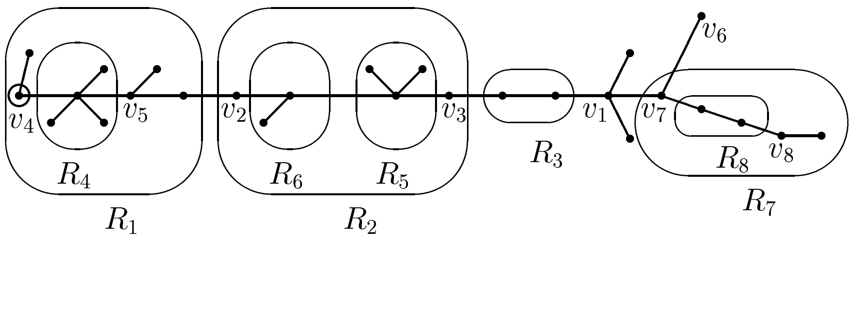

Figure 9: Critical chains $\mathcal{C} \equiv \{(\overline{R}_1, \overline{R}_2, R_3), (R_5, R_6),\ R_4,\ \overline{R}_7,\ R_8\ \}$

Given a chain $C \equiv (\overline{R}_1, ..., \overline{R}_s)$ 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 $u\bar{u}$ and $w\bar{w}$ are the edges connecting, respectively,

$\overline{R}_1$ and $\overline{R}_s$ with $T\backslash C$: thus $\bar{u} \in \overline{R}_1$, $\bar{w} \in \overline{R}_s$, u and w are distinct (and one of them may be the earth η). The points u, w will be called *the base points of the chain* C.
In the example shown in Figure 9, 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 $u(R_4) = v_5$, $w(R_4) = v_6$, $u(\overline{R}_7) = v_6$, $w(\overline{R}_7) = v_1$, $u(R_8) = v_7$, $w(R_8) = v_8$.
We denote by

$$U \equiv U(T) \equiv \{u(C), w(C) : \; C \in \mathcal{C}\} \tag{5.23}$$

the *set of base points of* T.

Remark 5.5 The order of a chain $(\overline{R}_1, ..., \overline{R}_s)$ 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 α.*

Definition 5.12 (Complete families) Given $T \in \mathcal{T}^k$, α T–admissible and $\lambda < 1/2$ we define

$$\mathcal{F}(T) \equiv \{T\} \tag{5.24}$$

if $\mathcal{R}(T) = \emptyset$ (*i.e.* if there are no λ–critical resonances in T), otherwise, letting $\mathcal{R} \equiv \mathcal{R}(T)$, $\mathcal{C} \equiv \mathcal{C}(T)$, we define

$$\begin{aligned}\mathcal{F}(T) \quad &\equiv \{T' \in \mathcal{T}^k : \quad T' = \overline{T} \cup \bigcup_{\overline{R} \in \overline{\mathcal{R}}} \overline{R} + \\ &\qquad \sum_{C = (\overline{R}_1, ..., \overline{R}_s) \in \mathcal{C}} u(C)\bar{u}_1 + w(C)\bar{w}_s + \sum_{i=1}^{s-1} \bar{w}_i \bar{u}_{i+1} \;, \\ &\qquad \text{where } \bar{u}_i, \bar{w}_i \text{ vary in } \overline{R}_i\}\end{aligned} \tag{5.25}$$

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

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 5.4 (Properties of $\mathcal{F}$) *Let $T \in \mathcal{T}^k$, α a T–admissible function and $\lambda < 1/2$. Then:*

$$\begin{aligned}
&\text{(i)} && \mathcal{F}(T) = \{T\} \iff \mathcal{R}(T) = \emptyset \;; \\
&\text{(ii)} && \mathcal{F}(T') = \mathcal{F}(T) \qquad \forall\, T' \in \mathcal{F}(T) \;; \\
&\text{(iiii)} && \mathcal{F}(T) \cap \mathcal{F}(T') \neq \emptyset \implies T' \in \mathcal{F}(T) \;; \\
&\text{(iv)} && \overline{T}, \mathcal{R}(T), U(T) \quad \text{is a complete and minimal} \\
& && \text{set of invariants for } \mathcal{F}(\mathrm{T}) \;; \\
&\text{(v)} && |\mathcal{F}| = \prod_{\overline{R} \in \overline{\mathcal{R}}} |V(\overline{R})|^2 \;; \\
&\text{(vi)} && \deg_{T'} v \leq \deg_{T''} v + 2 \qquad \forall\, T', T'' \in \mathcal{F}(T)
\end{aligned}$$

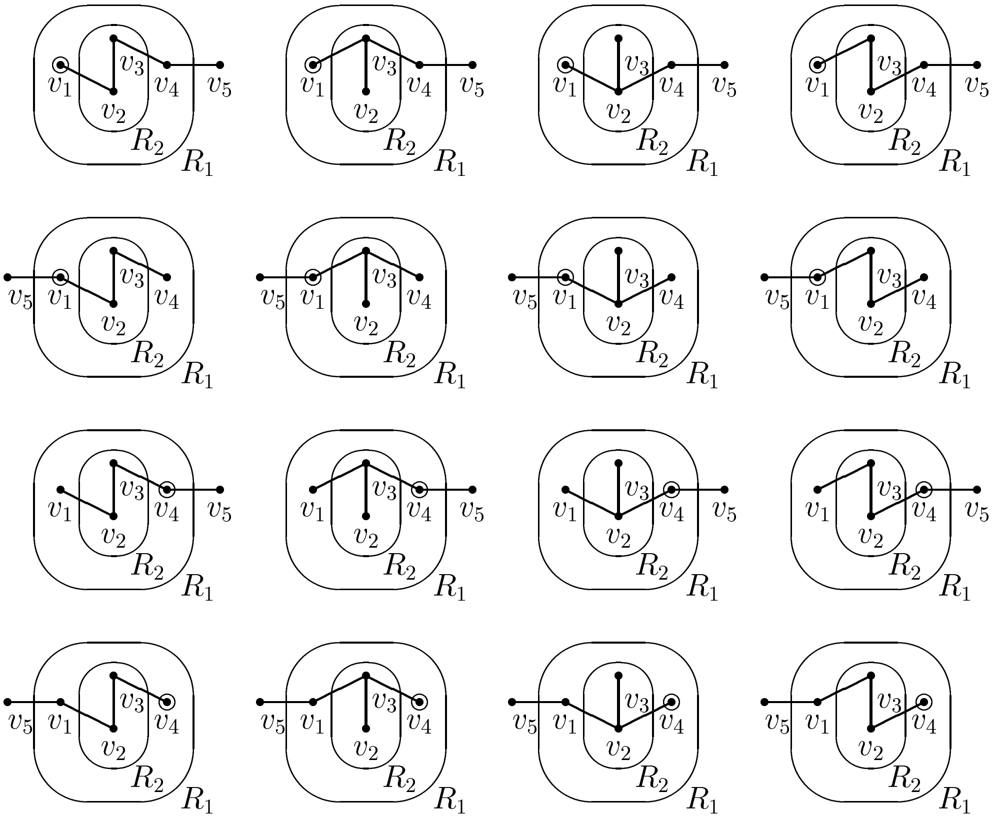

Figure 10: A complete family $\mathcal{F}$ (Example 1)

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

Remark 5.6 ¿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 α is not admissible for some T we are omitting the (meaningless) contribution coming from that tree.

A final comment: the set of labels $\{v_1, ..., v_k\}$ 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

$$\deg_{\mathcal{F}} v = \max_{T' \in \mathcal{F}} \deg_{T'} v \tag{5.26}$$

and from (vi) of Proposition 5.4 it follows immediately that:

$$\deg_{\mathcal{F}} v - 2 \leq \deg_{T'} v \leq \deg_{\mathcal{F}} v \ , \qquad \forall\, T' \in \mathcal{F} \tag{5.27}$$

In Figure 10 and Figure 11 we show the complete families associated to order 5 trees with two resonances.
In Figure 10 we have $\mathcal{R}_0 = R_1$, $\mathcal{R} = \{R_1, R_2\}$, $C_1 \equiv \overline{R}_1$, $C_2 \equiv R_2$, $\overline{T} = \{\eta\} \cup \{v_5\}$, $\overline{R}_1 \equiv \{v_1\} \cup \{v_4\}$, $u(C_1) = \eta$, $w(C_1) = v_5$, $u(C_2) = v_1$, $w(C_2) = v_4$, $|\mathcal{F}| = |V(\overline{R}_1)|^2 |V(R_2)|^2 = 16$.

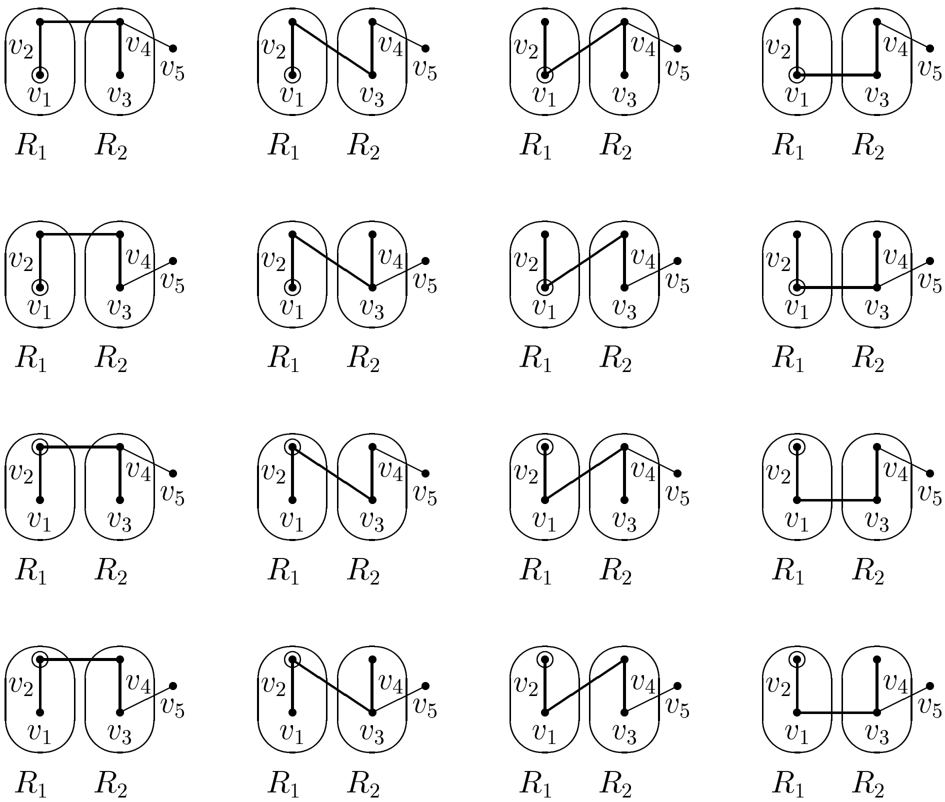

Figure 11: A complete family $\mathcal{F}$ (Example 2)

In Figure 11 we have $\mathcal{R}_0 = \mathcal{R} = \{R_1, R_2\}$, $C_1 = (R_1, R_2)$, $u(C_1) = u(R_1) = \eta$, $w(C_1) = w(R_2) = v_5$, $|\mathcal{F}| = |V(R_1)|^2|V(R_2)|^2 = 16$.

6 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 [9]. 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 occurrence has been discussed in Section 4 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 6.1 below. The proofs are presented in the next section.

Recall (2.7) and the definition of admissible functions (3.7).

Lemma 6.1 (Siegel, Eliasson) *Let $T \in \mathcal{T}^k$ be a rooted tree of order k (i.e. $k \equiv$ cardinality of $V(T)$), α a T–admissible function and $0 < \lambda < 1$. Assume that*

$$\text{if } \delta_u = \delta_w \text{ for some } u > w \;\text{ then } \exists\, v \text{ s.t. } u > v > w \text{ and } |\delta_v| > \lambda |\delta_u| \tag{6.1}$$

Then there exist constants $c_1 > \gamma$, $\beta_1 > \tau$ such that

$$\prod_{v\in T} |\delta_v|^{-1} \le c_1^k \prod_{v\in T} |\alpha_v|^{\beta_1} \tag{6.2}$$

The constant c_1, β_1 can be taken to be

$$c_1 = \gamma\; 2^{6\tau}\, \frac{1+\lambda}{\lambda}\,, \qquad \beta_1 = 3\tau \tag{6.3}$$

Next Lemma extend the above type of estimate to products of small divisors arising in trees with (possible) resonant subtrees which are non λ–resonant (for some prefixed $\lambda < 1/2$); recall that such trees may contain couples of points which violate (6.1) (see Remark 5.1) so that the result described in the next Lemma is a genuine generalization of the estimates described in Lemma 6.1.

Lemma 6.2 *Let T and α be as in Lemma 6.2 and let $0 < \lambda < 1/2$. Assume that there are no λ–resonance (Definition 5.3). Then there exist constants $c_2 > \gamma$ and $\beta_2 > \tau$ such that*

$$\prod_{v\in T} |\delta_v|^{-1} \le c_2^k \prod_{v\in T} |\alpha_v|^{\beta_2} \tag{6.4}$$

The constant c_2, β_2 can be taken to be

$$c_2 = c_1\; 2^{\tau}\, \frac{1-\lambda}{1-2\lambda}\,, \qquad \beta_2 = \beta_1 + \tau \tag{6.5}$$

Lemma 6.3 (Small divisor compensations)
Let T, α and λ be as in Lemma 6.2 and let $0 < \bar\lambda < 1$. Let R be a minimal λ–(sub)resonance of T (i.e. R does not contain any λ–subresonance, see Definition 5.6)

and let $n \in {}^N\backslash\{0\}$ be such that $|\langle n\rangle| \le \bar\lambda\Delta(R)$. 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[6] *$R \cup \{z\} + wz$ rooted at u. Then there exist constants $c_3 > \gamma^2$, $\beta_3 > 2\tau$ such that for all $1 \le i, j \le N$*

$$|\sum_{u,w\in R} \alpha_{ui}\alpha_{wj} \prod_{v\in R} \delta_v(R_u^w)^{-2}| \le c_3^{|V(R)|} \prod_{v\in R} |\alpha_v|^{\beta_3} \tag{6.6}$$

where α_{ui} (respectively α_{uj}) denote the i^{th} (respectively the j^{th}) component of the integer vector α_u. The constant c_3, β_3 can be taken to be

$$c_3 = (c_2\ 2^{\tau+1}(1-\bar\lambda)^{-1})^2\ , \qquad \beta_3 = 2(\beta_2+\tau+1) \tag{6.7}$$

The estimates (6.6) lead easily to control the contribution to (3.9) coming from complete families.

Lemma 6.4 *Let T, α and $\lambda < 1/2$ be as above. Let $\mathcal{F} \equiv \mathcal{F}(T)$ be the complete family defined in Definition 5.12, $V \equiv V(T) = V(\mathcal{F})$ and recall the definition of $\deg_{\mathcal{F}}$ given in (5.26). Then there exist constants $c_4 > \gamma^2$, $\beta_4 > 2\tau$ such that*

$$|\sum_{T'\in\mathcal{F}} \prod_{vv'\in E(T')} \alpha_v\cdot\alpha_{v'} \prod_V \delta_v^{-2}| \le \prod_{v\in V} |\alpha_v|^{\deg_{\mathcal{F}} v}\ c_4^{|V|} \prod_{v\in V} |\alpha_v|^{\beta_4} \tag{6.8}$$

The constants c_4, β_4 can be taken to be

$$c_4 \equiv (c_2\ 2^{\tau+1}\ \frac{1-\lambda}{1-2\lambda})^2\ , \qquad \beta_4 = \beta_3 \tag{6.9}$$

Kolmogorov's theorem follows now easily. We introduce the following norms on analytic functions $g : {}^N \to$:

$$\|g\|_\sigma \equiv \sum_{n\in{}^N} |g_n| e^{|n|\sigma} \tag{6.10}$$

Clearly, for any analytic function on ^{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 $g = (g_1, ..., g_M) : {}^N \to {}^M$ we set

$$\|g\|_\sigma \equiv \sup_{1\le j\le M} \|g_j\|_\sigma \tag{6.11}$$

Theorem 6.1 *Let f in (3.2) be real–analytic with $\|f\|_{\bar\sigma} < \infty$ for some $\bar\sigma > 0$. Then the formal solution $X(\theta, \varepsilon)$ described in Proposition 2.1 is real–analytic. Furthermore, fix $0 < \sigma < \bar\sigma$, $0 < \lambda < 1/2$ and set*

$$\varepsilon_0 \equiv \frac{(\bar\sigma-\sigma)^{\beta_5}}{c_5\|f\|_{\bar\sigma}} \quad \text{with} \quad \begin{cases} \beta_5 \equiv \beta_4 + 4 \\ c_5 \equiv c_4 2^{\beta_5+1}\ \beta_5! \end{cases} \tag{6.12}$$

where the constants β_4, c_4 are defined above. Then $X(\theta, \varepsilon)$ is jointly analytic in the domain $\{|\mathrm{Im}\theta_i| \le \sigma\} \times \{|\varepsilon| \le \varepsilon_0\}$ and for any (complex) ε with $|\varepsilon| < \varepsilon_0$ the following bound holds

$$\|X(\cdot, \varepsilon)\|_\sigma \le \frac{\bar\sigma-\sigma}{2}\ \frac{|\varepsilon|}{\varepsilon_0 - |\varepsilon|} \tag{6.13}$$

Remark 6.1 Choosing $\lambda = 1/5$ (so as to minimize the constant c_5) one obtains

$$\beta_5 = 10\tau + 12\ , \qquad c_5 < \gamma^2\ 2^{26\tau+23}\ (10\tau+12)! \tag{6.14}$$

Such constants are certainly not optimal.

[6]If R is rooted, $R = R_r$, first remove the edge ηr so as to obtain an unrooted subtree.

7 Proofs

In the following five Subsections we give the proofs of the four Lemmas and of the Theorem of Section 6. In the first Subsection we show how Kolmogorov's Theorem can be easily derived from Lemma 6.4 and in the remaining Subsections we give the proofs of the four Lemmas.

7.1 Proof of Theorem 6.1

Recall the form of the coefficients $X_{jn}^{(k)}$ given in (3.11) $\div$ (3.13). By Remark 5.6 we know that, for a given choice of $\{n_1, ..., n_k\}$ with $n_i \in {}^N\backslash\{0\}$,the labeled rooted trees in $\mathcal{T}^k$ can be partitioned into complete families having as common admissible α function $\alpha_{v_i} = n_i$. Denote by $\mathcal{N} \equiv \mathcal{N}(n_1, ..., n_k)$ the set of all such complete families. Then, by (3.11) $\div$ (3.13), by Lemma 6.4 and by (5.26), (5.27) we see that, for any $0 < \sigma < \bar{\sigma}$,

$$\begin{aligned}
&\sum_{n \in {}^N\backslash\{0\}} e^{\sigma|n|}|X_{jn}^{(k)}| \\
&= \sum_n e^{\sigma|n|} \frac{1}{k!} \Big| \sum_{\substack{n_1,...,n_k \\ n_i\neq 0, \sum n_i = n}} \sum_{\mathcal{F}\in\mathcal{N}} \prod_{i=1}^k f_{n_i} \sum_{T\in\mathcal{F}} \prod_{vv'\in E(T)} \alpha_v \cdot \alpha_{v'} \prod_V \delta_v^{-2} \Big| \\
&\leq \sum_n \frac{1}{k!} \sum_{\substack{n_1,...,n_k \\ n_i\neq 0, \sum n_i = n}} \sum_{\mathcal{F}\in\mathcal{N}} \prod_{i=1}^k (f_{n_i} e^{\sigma|n_i|}) \prod_{v\in V} |\alpha_v|^{\deg_{\mathcal{F}} v} c_4^k \prod_V |\alpha_v^{\beta_4}| \\
&\leq \frac{1}{k!} \sum_{T\in\mathcal{T}^k} \sum_{\alpha:T\to{}^N} \prod_V (c_4 |f_{\alpha_v}| |\alpha_v|^{\beta_4+2} e^{\sigma|\alpha_v|}) \prod_V |\alpha|^{\deg v} \qquad (7.1)
\end{aligned}$$

In view of Proposition 3.2 we see that, if we define

$$\phi_n \equiv c_4 |f_n| e^{\sigma|n|} |n|^{\beta_4+2} \qquad (7.2)$$

then, last line of (7.1) is exactly the formal solution g of the equation (3.14) with coefficient g_k given in (3.15). But since (by the analyticity of f) the nonnegative numbers ϕ_n decay exponentially fast as $|n| \to \infty$, the function of two complex variables

$$G(x, \varepsilon) \equiv x - \varepsilon \sum_{n\in{}^N} \phi_n |n| e^{|n|x} \qquad (7.3)$$

is holomorphic and bounded in the complex 2–disk

$$D \equiv \{x \in \;:\; |x| \leq s\} \times \{\varepsilon \in \;:\; |\varepsilon| \leq \varepsilon_0\} \qquad (7.4)$$

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

$$\begin{aligned}
\sup_D |G| < \sup_D |G_x| \;&=\; \varepsilon_0 \sum_{n\in{}^N} \phi_n |n|^2 e^{\sigma|n|} \\
&\leq\; \varepsilon_0 \, c_4 \, (\bar{\sigma} - \sigma - s)^{-\beta_5} \beta_5! \, \|f\|_{\bar{\sigma}} \qquad (7.5)
\end{aligned}$$

Taking $s \equiv (\bar\sigma - \sigma)/2$ and ε_0 as in (6.12) 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

$$G(g(\varepsilon), \varepsilon) = 0 \ , \qquad \sup_{\substack{\varepsilon\in \\ |\varepsilon|\le\varepsilon_0}} |g(\varepsilon)| \le s = \frac{\bar\sigma - \sigma}{2} \tag{7.6}$$

Thus $g_k \le \frac{1}{2}(\bar\sigma - \sigma)\varepsilon_0^{-k}$ and therefore

$$\|X^{(k)}\|_\sigma \le \frac{1}{2}(\bar\sigma - \sigma)\varepsilon_0^{-k} \tag{7.7}$$

and, for any complex ε with $|\varepsilon| < \varepsilon_0$

$$\|X(\cdot, \varepsilon)\|_\sigma \le \frac{\bar\sigma - \sigma}{2} \frac{\varepsilon/|\varepsilon_0|}{1 - \varepsilon/|\varepsilon_0|} \tag{7.8}$$

completing the proof of Theorem 6.1. ∎

7.2 Proof of Lemma 6.1

We follow [9]. The first two steps are stated in the next two Sublemmas.

Sublemma 7.1 *Let $P \equiv v_1...v_k$ be the path with k vertices rooted at v_1. Let $\alpha : P \to {}^N\backslash\{0\}$ be a P–admissible function and fix $0 < \lambda < 1$. Assume that*

$$\text{if } \delta_u = \delta_w \ \text{ for some } u > w \ \text{ then } \exists\ v \text{ between } u \text{ and } w \text{ s.t. } |\delta_w| > \lambda|\delta_v| \tag{7.9}$$

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

$$\prod_P |\delta_v|^{-1} \le B_0 D_0^{k-1} \prod_V |\alpha_v|^\tau \tag{7.10}$$

The constants B_0, D_0 can be taken to be

$$B_0 = \gamma \ , \qquad D_0 = \gamma\ 2^\tau\ \frac{1+\lambda}{\lambda} \tag{7.11}$$

Proof By induction on k. If $k = 1$, (7.10) follows at once from the Diophantine inequality (2.7). Let $k \ge 2$ and assume the statement true for all rooted paths with up to $k-1$ vertices satisfying (7.9). Let (P, α) as above. Define s to be the greatest integer between 1 and k such that

$$|\delta_{v_s}| \ge \max_{1\le i\le k} |\delta_{v_i}| \tag{7.12}$$

There are four cases:

$$\begin{array}{ll} \text{(i) } s = 1 \ , & \text{(ii) } \alpha_{v_s} + \alpha_{v_{s-1}} \ne 0 \ , \\ \text{(iii) } \alpha_{v_s} + \alpha_{v_{s-1}} = 0 \ , s+1 < k \ , & \text{(iv) } \alpha_{v_s} + \alpha_{v_{s-1}} = 0 \ , \ s+1 = k \end{array} \tag{7.13}$$

Note that in case (iii) and (iv) we have $s < k$ because α 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

$$|\langle \alpha_{v_i} \rangle| = |\delta_{v_i} - \delta_{v_{i+1}}| \le |\delta_{v_i}| + |\delta_{v_{i+1}}| \le 2|\delta_{v_s}| \tag{7.14}$$

hence, in case (i), by (2.7),

$$|\delta_{v_1}|^{-1} \le 2|\langle \alpha_{v_1} \rangle|^{-1} \le 2\gamma |\alpha_{v_1}|^{\tau} \tag{7.15}$$

while in case (ii)

$$|\delta_{v_s}| \ge \frac{1}{2} \max\{|\langle \alpha_{v_s} \rangle|, |\langle \alpha_{v_{s-1}} \rangle|\} \tag{7.16}$$

In case (iii) and (iv) we have that $\delta_{v_{s+1}} = \delta_{v_{s-1}}$, and by (7.9)

$$|\delta_{v_{s+1}}| > \lambda |\delta_{v_s}| > \lambda |\delta_{v_{s+2}}| \tag{7.17}$$

where last inequality holds only in case (iii). In all cases mimicking (7.14) replacing 2 by $(1+\lambda)^{-1}$ we obtain

$$|\delta_{v_{s+1}}| \ge \frac{\lambda}{1+\lambda} \max\{|\langle \alpha_{v_{s+1}} \rangle|, |\langle \alpha_{v_s} \rangle|\} \ . \tag{7.18}$$

Let now $\bar v \equiv v_s$ in case (i), (ii) and $\bar v \equiv v_{s+1}$ in case (iii), (iv) and let, in the cases (ii), (iii), (iv), $v' > \bar v$ be adjacent to $\bar v$. Then in case (ii), (iii), (iv) we obtain from (7.16) and (7.18), using the Diophantine inequality (2.7) (and the fact that $\lambda < 1$), that

$$|\delta_{\bar v}|^{-1} \ \frac{|\alpha_{\bar v} + \alpha_{v'}|^{\tau}}{|\alpha_{\bar v}|^{\tau} |\alpha_{v'}|^{\tau}} \le \frac{1+\lambda}{\lambda} 2^{\tau} \gamma \tag{7.19}$$

Now, if we let $\bar P \equiv P/\{\bar v\}$ (recall Definition 5.10) and $\bar\alpha : \bar P \to {}^N \backslash \{0\}$ defined by

$$\bar a_v \equiv \alpha_v \text{ in case (i) ; and in case (ii), (iii), (iv) :} \quad \bar\alpha_v \equiv \begin{cases} \alpha_v & \text{if } v \ne v' \\ \alpha_v + \alpha_{\bar v} & \text{if } v = v' \end{cases} \tag{7.20}$$

Then

$$\prod_P |\delta_v|^{-1} = |\delta_{\bar v}|^{-1} \prod_P |\delta_v(\bar P; \bar\alpha)|^{-1} \tag{7.21}$$

If $(\bar P, \bar\alpha)$ verifies (7.9) then we can apply the inductive hypothesis and, using (7.16) or (7.17), the Sublemma would follow immediately. It remains to check that $(\bar P, \bar\alpha)$ satisfy (7.9).

Case (i) is obvious as $\bar\alpha$ coincide with α on $\bar P = v_2 ... v_k$. Also case (iv) is easily checked as any couple of resonant points (*i.e.* points where the divisors δ coincide) in $\bar P$ would also be resonant for P. Consider case (ii) and assume that $u > w$ form a couple of resonant points for $\bar P$. If either $w > v_{s-1}$ or $v_{s-1} > u$, then (7.9) follows immediately from the validity of (7.9) 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 $v_i \ne v_s$, (7.9) holds for $\bar P$ as $v_i \in \bar P$; on the other hand if $v_i = v_s$ then

$$|\delta_{v_s}| \ge |\delta_{v_{s-1}}| \quad \Longrightarrow \quad |\delta_w| > \lambda |v_s| \ge \lambda |\delta_{v_{s-1}}| \tag{7.22}$$

showing that (7.9) holds for $\bar P$ 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

$$\sum_{\substack{v\in P\\ v_s>v>w}} \alpha_v = \sum_{\substack{v\in P\\ v_{s+1}>v>w}} \alpha_v = -\alpha_{v_{s+1}} \tag{7.23}$$

Hence $v_{s+1} > w$ for a resonant couple in P and there exists by (7.9) a point v_j between v_{s+1} and w ($\implies v_j \in \bar{P}$) such that $|\delta_w| > \lambda|\delta_{v_j}|$ and we see that (7.9) holds for $(\bar{P}, \bar{\alpha})$ also in this final case. ∎

Sublemma 7.2 *Let $r \geq 0$ and $s \geq 2$ be integers and let $x_1, ..., x_r$, $y_1, ..., y_s$ be real numbers greater or equal than 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*

$$\sum_{j=1}^{s} y_j \geq \frac{z}{2} \qquad \text{and} \qquad y_j \leq \frac{z}{2} \quad \forall\, j \tag{7.24}$$

then

$$\prod_{i=0}^{r} x_i^{-1} \prod_{j=1}^{s} y_j^{-2} \leq 2^{r+2s} z^{-3} \tag{7.25}$$

Proof The proof is taken from [9]: 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

$$(y_1 - t)(y_2 - t) \leq y_1 y_2\ , \qquad \forall\ \ 0 \leq t \leq y_1 + y_2 \tag{7.26}$$

and taking $t = y_1 - 1$ we see that we can assume that $y_1 = 1$. 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 (7.26) we can again assume that either $y_1 = 1$ or $y_2 = z/2$. The remaining cases are checked directly. ∎

We are now now ready for the

Proof of Lemma 6.1 We shall prove, by induction on the number $|V|$ of vertices of T, the following estimate

$$\prod_V |\delta_v|^{-1} \leq BD^{|V|-1} \prod_V |\alpha_v|^{3\tau} (\sum_V |\alpha_v|)^{-2\tau} \tag{7.27}$$

for suitable constants $D \geq B \geq \gamma$ determined below. Lemma 6.1 follows immediately from (7.27) by taking $\beta_1 = 3\tau$ and $c_1 \geq D$.

If $|V| = 1$, (7.27) follows from (2.7) (as $B \geq \gamma$). Now let $k \geq 2$, assume that (7.27) holds for any (T', α') satisfying the hypotheses of Lemma 6.1 and with $|V(T')| \leq k - 1$ and let (T, α) be as in Lemma 6.1 with $|V(T)| = k$. We need some notation. Let

$$S_v \equiv \text{ the subtree } \ \{v' \in T:\ v' \leq v\} \ \text{ rooted at } v \tag{7.28}$$

and for any subset of vertices $W \subset V$ let

$$|\alpha|(W) \equiv \sum_{v \in W} |\alpha_v| \tag{7.29}$$

Define

$$P \equiv \{v \in T : \ |\alpha|(S_v) > \frac{1}{2}|\alpha|(T)\} \tag{7.30}$$

It is easy to check that P *is a path*: $P = u_0 \ldots u_p$ 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

$$\begin{aligned}
&y_{ij} \equiv |\alpha|(S_{ij}) \ , \quad y_i \equiv \sum_{j=1}^{m_i} y_{ij} \ , \quad x_i \equiv |\alpha_{u_i}| \\
&I \equiv [0, \ldots, p] \ , \quad I_0 \equiv \{i \in I : \ m_i = 0\} \ , \quad I_1 \equiv \{i \in I : \ m_i > 0\} \\
&m \equiv \sum_I m_i \ , \quad k_{ij} \equiv |S_{ij}| \ , \quad k_i \equiv \sum_{j=1}^{m_i} k_{ij} \ , \quad p_0 \equiv |I_0| \ , \ p_1 \equiv |I_1|
\end{aligned} \tag{7.31}$$

It then follows

$$|\alpha|(T) = \sum_I x_i + \sum_{I_1} y_i \ , \quad p_0 + p_1 = p + 1 \ , \quad k = p + 1 + \sum_{i \in I_1} k_i \ , \quad m \geq p_1 \tag{7.32}$$

and, by the definition of P

$$\begin{aligned}
&\sum_{j \geq i} x_j + \sum_{\substack{j \geq i \\ j \in I_1}} > \frac{1}{2}|\alpha|(T) \ , \qquad \forall \ i \in I \\
&x_i \leq \frac{1}{2}|\alpha|(T) \ \ \forall \ i < p \ , \quad y_{ij} \leq \frac{1}{2}|\alpha|(T) \ \ \forall i \in I_1, \ 1 \leq j \leq m_i
\end{aligned} \tag{7.33}$$

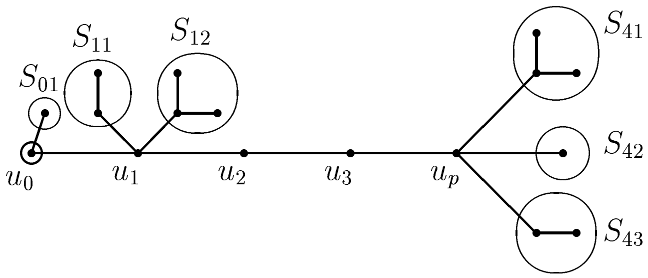

Figure 12: $m_0 = 1$, $m_1 = 2$, $p = 4$ and $m_p = 3$

If $I_1 = \emptyset$ then $P \equiv T$ and we can apply directly Sublemma 7.1 assuming that

$$B \geq B_0 \ , \qquad D \geq 2^{2\tau} \ D_0 \tag{7.34}$$

¿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

$$\bar{\alpha}_{u_i} \equiv \alpha_{u_i} + \alpha(S_{ij}) \qquad \Longrightarrow \qquad |\bar{\alpha}_{u_i}| \leq x_i + y_i \tag{7.35}$$

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

$$\prod_T |\delta_v|^{-1} = \prod_P |\delta_{u_i}(P;\bar{\alpha})|^{-1} \prod_{i\in I_1} \prod_{j=1}^{m_i} \prod_{v\in S_{ij}} |\delta_v|^{-1} \tag{7.36}$$

and we can use Sublemma 7.1 to bound the product over P (as $(P,\bar{\alpha})$ obviously satisfy (7.9)) and the inductive hypotheses to bound the product over S_{ij} (note that $\delta_v(T) = \delta_v(S_{ij})$). Recalling the definitions in (7.31), we obtain from (7.36) and (7.35)

$$\prod_{v\in T} |\delta_v|^{-1} \leq MBD^{k-1}(\prod_T |\alpha_v|^{3\tau})|\alpha|(T)^{-2\tau} \tag{7.37}$$

with

$$M \equiv \frac{B_0}{B}(\frac{B}{D})^m(\frac{D_0}{D})^p[\prod_{I_0} x_i^{-2} \cdot \prod_{I_1}(\frac{x_i+y_i}{x_i^3}\prod_j y_{ij}^{-1}) \cdot |\alpha|(T)^2]^\tau \tag{7.38}$$

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 (7.33) with $i = p$ we have

$$x_p > \frac{1}{2}|\alpha|(T) \tag{7.39}$$

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 (7.38) by $4x_p^2$ we get

$$M \leq \frac{B_0}{B}(\frac{D_0}{D})^p[\prod_{\substack{i\in I\\ i\neq p}} x_i^{-2} 2^{p_1+2}\, 2^{m-p_1} \prod_{i\in I_1,j} y_{ij}^{-1}]^\tau \leq 1 \tag{7.40}$$

provided

$$D \geq 2^\tau \max\{B, 2^\tau D_0\}\ , \qquad B \geq B_0 2^{2\tau} \tag{7.41}$$

(which is stronger than (7.34)).

Suppose now that $p \in I_1$. We can assume that $x_p \leq |\alpha|(T)/2$ (otherwise (7.39) holds also in this case and we can proceed as above). Then

$$M \leq \frac{B_0}{B}(\frac{B}{D})^m(\frac{D_0}{D})^p[2^{2m} \prod_{\substack{i\in I\\ i\neq p}} x_i^{-1} \cdot \prod_{\substack{i\in I_1\\ i=neqp}} y_i^{-1} \cdot x_p^{-2} \prod_{j=1}^{m_p} y_{pj}^{-2} \cdot y_p|\alpha|(T)^2]^\tau \tag{7.42}$$

and we see that we are in position to apply Sublemma 7.2 (with $r = p+p_1-1$, $s = m_p+1$, $z = |\alpha|(T) = \sum_I x_i + \sum_{I_1} y_i$) obtaining

$$M \leq \frac{B_0}{B}(\frac{B}{D})^p(\frac{D_0}{D})^p[2^{2m}\, 2^{p+p_1-1}\, 2^{2(m_p+1)} \frac{y_p}{|\alpha|(T)}]^\tau \leq 1 \tag{7.43}$$

provided we take

$$B \equiv B_0\, 4^\tau \;, \qquad D \equiv 4^\tau \max\{D_0, B_0 2^{4\tau}\} \tag{7.44}$$

(which satisfy also (7.41)). Thus, by (7.11), we can take

$$\beta_1 \equiv 3\tau \;, \qquad c_1 = \gamma\; 2^{6\tau}\, \frac{1+\lambda}{\lambda} \tag{7.45}$$

This conclude the proof of Lemma 6.1. ∎

7.3 Proof of Lemma 6.2

We prove the Lemma by induction on $k \equiv |V|$. For $k = 1$, (6.4) follows from (2.7) (as $c_2 > \gamma$ and $\beta_2 > \tau$). Let $k \geq 2$, assume that the Lemma holds for all trees with up to $k-1$ vertices and let $T \in \mathcal{T}^k$ have no λ–resonances. Recall that even though T does not contain λ–resonances, it may contain either λ–resonant couples of points (compare Remark 5.1), or *adjacent points* u, w such that $\delta_u = \delta_w$ (recall that the main hypothesis of Lemma 6.1, (6.1), excludes the presence of adjacent points of constancy for δ). For any couple (u, w) of λ–resonant points let $R(u, w)$ be the associated resonant subtree (see (5.3)). Let R be a minimal resonant tree associated to a couple of λ–resonant points (minimal meaning that R does not contain λ–resonant points); if there are no λ–resonant points we set $R = \emptyset$. There are two cases:
(i) either $R = \emptyset$ or R contains two adjacent points $u > w$ such that $\delta_u = \delta_w$, or
(ii) $R \neq \emptyset$ and R does not contain any couple of adjacent points on which δ is constant.
Consider case (i): if there are no couples of adjacent points u, w such that $\delta_u = \delta_v$, then Lemma 6.2 follows from Lemma 6.1, otherwise let $S = S(v_1, v_2)$ be a minimal resonant subtree such that v_1 and v_2 are adjacent and of constancy for δ (minimal means that S does not contain any couple of adjacent points on which δ is constant). Since $\alpha_v \neq 0$ for all vertices v, S contains at least two points and by hypothesis S is not λ–resonant. Thus observing that

$$\prod_T |\delta_v|^{-1} = \prod_{T/S} |\delta_v|^{-1} \prod_S |\delta_v|^{-1} \equiv \prod_{T/S} |\delta_v|^{-1} \prod_S |\delta_v(S_{v_1})|^{-1} \tag{7.46}$$

we see that we can apply Lemma 6.1 to the tree S_{v_1} while to the tree T/S we can use the inductive hypothesis[7] so that (6.4) holds in case (i) provided

$$c_2 \geq c_1 \;, \qquad \beta_2 \geq \beta_1 \tag{7.47}$$

Consider case (ii). It is easy to see that the tree T/R does not contain any λ–resonance and since

$$\prod_T |\delta_v|^{-1} = \prod_{T/R} |\delta_v|^{-1} \prod_R |\delta_v|^{-1} \tag{7.48}$$

we see that we can estimate, by the inductive hypothesis, the product over T/R. To bound the product over $R = R(u, w)$, let $n \equiv \sum_{v' \leq w} \alpha_{v'}$. Then, since (u, w) form a couple of λ–resonant points, for any $u < v < w$ we have

$$\begin{aligned} |\delta_w| \quad \equiv |\langle n \rangle| \;\; &\leq \lambda |\delta_v| = \lambda |\delta_v(R_u) + \langle n \rangle| \\ &\leq \lambda (|\delta_v(R_u)| + \langle n \rangle|) \end{aligned} \tag{7.49}$$

[7] Note that contracting resonant trees associated to adjacent points where δ is constant cannot introduce new resonances, thus all resonances in T/S are not λ–resonant.

(where R_u is the tree R rooted at u) and therefore

$$|\langle n\rangle| \le \frac{\lambda}{1-\lambda}|\delta_v(R_u)| \ , \quad \forall \ u < v < w \tag{7.50}$$

Let $\bar{u}$ be the point in R adjacent to w (which is outside R) and let P the path (in R) joining u and $\bar{u}$. Then

$$\begin{aligned}\prod_R |\delta_v|^{-1} &= |\langle n\rangle|^{-1} \prod_{\substack{v\in R\\ v\notin P}} |\delta_v(R_u)|^{-1} \prod_{\substack{v\in P\\ v\neq u}} |\delta_v(R_u)+\langle n\rangle|^{-1} \\ &\le |\langle n\rangle|^{-1} \prod_{\substack{v\in R\\ v\notin P}} |\delta_v(R_u)|^{-1} \prod_{\substack{v\in P\\ v\neq u}} \frac{1-\lambda}{1-2\lambda}|\delta_v(R_u)|^{-1} \\ &\le (\frac{1-\lambda}{1-2\lambda})^k \ |\langle n\rangle|^{-1} \prod_{v\in R, v\neq u} |\delta_v(R_u)|^{-1} \end{aligned} \tag{7.51}$$

Now we can apply Lemma 6.1 to the rooted tree(s) $R\backslash\{u\}$ and, since R is not a λ–resonance

$$\begin{aligned}|\langle n\rangle|^{-1} &< \lambda^{-1}\Delta(R)^{-1} \le \lambda^{-1}\gamma(\sum_{v\in R}|\alpha_v|)^\tau \\ &\le \lambda^{-1}2^{(|V(R)|-1)\tau}\gamma\prod_R|\alpha_v|^\tau \end{aligned} \tag{7.52}$$

Thus

$$\prod_R |\delta_v|^{-1} \le (2^\tau c_1)^{|V(R)|-1} \ \lambda^{-1}\gamma(\frac{1-\lambda}{1-2\lambda})^{|V(R)|}\prod_R|\alpha_v|^{\beta_1+\tau} \tag{7.53}$$

and we see that Lemma 6.2 holds also in this case if we take

$$\beta_2 = \beta_1 + \tau \ , \qquad c_2 = c_1 \ 2^\tau \ \frac{1-\lambda}{1-2\lambda} \tag{7.54}$$

(as $2^\tau c_1 > \lambda^{-1}\gamma$). ∎

7.4 Proof of Lemma 6.3

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

Proposition 7.1 (Compensations) *Let T, R, n be as in Lemma 6.3 and set (for prefixed $1 \le i,j \le N$)*

$$\sigma(x) \equiv \sum_{u,v\in R}\alpha_{ui}\alpha_{uj}\prod_{\substack{v\in P(u,w)\\ v\neq u}}(\delta_v+x)^{-2}\prod_{v\in R\backslash P(u,w)}\delta_v^{-2} \tag{7.55}$$

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

$$\sigma(0) = \sigma'(0) = 0 \qquad (' \ \equiv \frac{d}{dx}) \ . \tag{7.56}$$

Note that

$$\sigma(\langle n\rangle) = \sum_{u,w\in R} \alpha_{ui}\alpha_{uj} \prod_{R} \delta_v(R_u^w)^{-2} \tag{7.57}$$

(the definition of R_u^w is given in Lemma 6.3).

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

$$\begin{aligned}
\mu(u) &\equiv \prod_{\substack{v\in R\\ v\neq u}} |\delta_v(R_u)| \\
\nu(w) &\equiv \sum_{u\in S, u\neq w} \alpha_{wi} \sum_{\substack{v\in P(u,w)\\ v\neq u}} |\delta_v(R_u)|^{-1}
\end{aligned} \tag{7.58}$$

We claim that

$$\begin{aligned}
&\text{the functions } \mu \text{ and } \nu \textit{ are independent of } u \textit{ and } w : \\
&\mu(u)\equiv\mu\equiv\mu_R\ , \qquad \nu(u)\equiv\nu\equiv\nu_{R,S}
\end{aligned} \tag{7.59}$$

The independence of $\mu(u)$ from u comes immediately from the identities

$$\delta_v(R_u)=\delta_v(R_w)\ , \quad \forall\ v\notin\ P(u,w) \tag{7.60}$$

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

$$\prod_{\substack{v\in P(u,w)\\ v\neq u}} \delta_v(R_u) = \prod_{\substack{v\in P(u,w)\\ v\neq w}} (-\delta_v(T_w))\ , \quad \forall\ u\neq w \tag{7.61}$$

Note that (7.60) holds for any rooted tree (*i.e.* not necessarily null) while in (7.61) it is important that R is null. Now, for any u and w with $u\neq w$ (7.60) and (7.61) imply

$$\begin{aligned}
\mu(u) &= \prod_{\substack{v\in P(u,w)\\ v\neq u}} |\delta_v(R_u)| \prod_{v\notin P(u,w)} |\delta_v(R_u)| \\
&= \prod_{\substack{v\in P(u,w)\\ v\neq w}} |\delta_v(R_w)| \prod_{v\notin P(u,w)} |\delta_v(R_w)| = \mu(w)
\end{aligned} \tag{7.62}$$

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

$$\begin{aligned}
\nu(w)-\nu(w') &= \sum_{\substack{u\in S\\ u\neq w,w'}} \alpha_{ui}\Big[\sum_{\substack{v\in P(u,w)\\ v\neq u}} \delta_v(R_u)^{-1} - \sum_{\substack{v\in P(u,w')\\ v\neq u}} \delta_v(R_u)^{-1}\Big] + \\
&\qquad +\chi_S(w')\alpha_{w'i}\delta_w(R_{w'})^{-1} - \chi_S(w)\alpha_{wi}\delta_{w'}(R_w)^{-1} \\
&= \sum_{\substack{u\in S\\ u\neq w,w'}} \alpha_{ui}\delta_w(R_{w'})^{-1} + \chi_S(w')\alpha_{w'i}\delta_w(R_{w'})^{-1} + \\
&\qquad +\chi_S(w)\alpha_{wi}\delta_w(R_{w'})^{-1} \\
&= \Big(\sum_{u\in S}\alpha_{ui}\Big)\ \delta_w(R_{w'})^{-1} = 0
\end{aligned} \tag{7.63}$$

where in the second equality we used $\delta_{w'}(R_w) = -\delta_w(R_{w'})$ which is a particular case of (7.61) and χ_S is the characteristic function of the set S. This finishes the proof of the claim (7.59).
The check of (7.56) is now trivial:

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

and

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

finishing the proof of Proposition 7.1. ∎

Remark 7.1 In fact, from the above proof it follows that (7.56) 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 6.3.

Since $\alpha(R) = 0$ we can rewrite the left hand side of (6.6) as follows

$$
\begin{aligned}
\sum_{u,w\in R} \alpha_{ui}\alpha_{wj} \prod_{R} \delta_v(R_u^w)^{-2} &= \langle n\rangle^{-2} \sum_{u,w\in R} \alpha_{ui}\alpha_{wj} \prod_{\substack{v\in R\\ v\neq u}} \delta_v(R_u^w)^{-2} \\
&\equiv \langle n\rangle^{-2}\sigma(\langle n\rangle) \qquad (7.64)
\end{aligned}
$$

see (7.57). By Proposition 7.1 we see that

$$
\begin{aligned}
\sigma(x) &= x^2 \int_0^1 \sigma''(tx)(1-t)\ dt \\
&\equiv x^2\bar{\sigma}(x) \qquad (7.65)
\end{aligned}
$$

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

$$
|\langle n\rangle| \leq \bar{\lambda}\ \Delta \equiv \bar{\lambda}\Delta(R) \qquad (7.66)
$$

and we see that σ *is holomorphic and bounded* in the complex disk $\{x \in\ : |x| \leq \frac{1+\bar{\lambda}}{2}\ \Delta\}$.
Recall the *"Cauchy estimate"*:

$$
\sup_{D'} |g^{(k)}| \leq\ k!\ r^{-k} \sup_{D} |g|\ , \qquad r \equiv \operatorname{dist}\{D', \partial D\} \qquad (7.67)
$$

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 ∂ denotes "boundary of")[8]. Then, letting

$$D \equiv \{x \in \ : \ |x| \leq \frac{1+\bar{\lambda}}{2}\,\Delta\}\ , \quad D' \equiv \{x \in \ : \ |x| \leq \bar{\lambda}\,\Delta\} \tag{7.68}$$

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

$$\sup_{D'} |\sigma''| \leq 8\,(1-\bar{\lambda})^{-2}\,\Delta^{-2}\,\sup_{D} |\sigma| \tag{7.69}$$

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 $\Delta \leq \delta_v(R_u)$).

$$\begin{aligned} \Big(\delta_v(R_u) + x\Big)^{-2} &\leq \Big(\delta_v(R_u) - \frac{1+\bar{\lambda}}{2}\,\Delta\Big)^{-2} \\ &\leq 4(1-\bar{\lambda})^{-2}\,\delta_v(R_u)^{-2} \end{aligned} \tag{7.70}$$

Thus, by (7.65), (7.66), (7.69), (7.70), (7.55) and Lemma 6.2

$$\begin{aligned} \langle n\rangle^{-2}|\sigma(\langle n\rangle)| &= |\bar{\sigma}(\langle n\rangle)| \leq \frac{1}{2}\sup_{D'}|\sigma''| \\ &\leq 4(1-\bar{\lambda})^{-2}\,\Delta^{-2}\,\sup_{D}|\sigma| \\ &\leq \Big(4(1-\bar{\lambda})^{-2}\Big)^{|V(R)|}\Delta^{-2}\sum_{u,w\in R}|\alpha_u|\,|\alpha_w|\prod_{v\neq u}|\delta_v(R_u)|^{-2} \\ &\leq \Big(4(1-\bar{\lambda})^{-2}\Big)^{|V(R)|}\Delta^{-2}\,\Big(|V(R)|\prod_{v\in R}|\alpha_v|\Big)^2\,c_2^{2(|V(R)|-1)}\prod_{\substack{v\in R\\ v\neq u}}|\alpha_v|^{2\beta_2} \\ &\leq \Big((1-\bar{\lambda})^{-2}\,2^{\tau+1}\,c_2\Big)^{2|V(R)|}\prod_{v\in R}|\alpha_v|^{2(\beta_2+\tau+1)} \end{aligned} \tag{7.71}$$

where in the last estimate we have applied Lemma 6.2 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 the estimate

$$\Delta^{-1} \leq \gamma\,(\sum_R |\alpha_v|)^\tau \leq \gamma\,2^{(|V(R)|-1)\tau}\prod_R |\alpha_v|^\tau \tag{7.72}$$

The proof of Lemma 6.3 is complete. ∎

7.5 Proof of Lemma 6.4

If $T' \in \mathcal{F}(T)$, denote by

$$\Gamma(T') \equiv \{vv' \in E(T') : \ \exists\ R \in \mathcal{R}(T) \ \text{ with } \ v' \in \overline{R} \text{ and } v \notin R\} \tag{7.73}$$

[8]The proof of (7.67) is standard: it follows immediately by expressing the k^{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).

and by $E(\overline{\mathcal{R}}) \equiv \cup_{\overline{R}\in\overline{\mathcal{R}}} E(\overline{R})$. Then

$$\prod_{vv'\in E(T')} \alpha_v \cdot \alpha_{v'} = \prod_{vv'\in\Gamma(T')} \alpha_v \cdot \alpha_{v'} \prod_{vv'\in E(\overline{\mathcal{R}})\cup\overline{T}} \alpha_v \cdot \alpha_{v'} \tag{7.74}$$

and the second product does not depend on T' (*i.e.* it is common to all elements of $\mathcal{F}$). Furthermore, recalling (5.25), we see that

$$\prod_{vv'\in\Gamma(T')} \alpha_v \cdot \alpha_{v'} = \prod_{C=(\overline{R}_1,...,\overline{R}_s)\in\mathcal{C}} \alpha_{u(C)} \cdot \alpha_{\bar{u}_1}\ \alpha_{w(C)} \cdot \alpha_{\bar{w}_s} \prod_{i=1}^{s-1} \alpha_{\bar{w}_i} \cdot \alpha_{\bar{u}_{i+1}} \tag{7.75}$$

last product being absent if $s = 1$. Fixing a set of indices (depending upon C and taking the values $1, ..., N$), $j_i = j_i(C)$, for $i = 0, ..., s$, we can make more explicit (7.74) by writing out the scalar products:

$$\prod_{vv'\in\Gamma(T')} \alpha_v \cdot \alpha_{v'} = \sum_{\substack{\{j_0,...,j_s\}_{C\in\mathcal{C}}\\ 1\le j_i\le N}} \prod_{C\in\mathcal{C}} \alpha_{u(C)j_0}\alpha_{w(C)j_s} \prod_{C\in\mathcal{C}} \prod_{i=1}^{s} \alpha_{\bar{u}_i j_{i-1}}\alpha_{\bar{w}_i j_i} \tag{7.76}$$

Now observe that $|\overline{\mathcal{R}}| \le (|V|-1)/2$, which implies

$$\sum_{\{j_o,...,j_s\}_{C\in\mathcal{C}}} 1 \le N^{|V|}\ , \tag{7.77}$$

and notice that, for all $T' \in \mathcal{F}$ (recall (5.26))

$$\Big| \prod_{vv'\in E(T')} \alpha_v \cdot \alpha_{v'} \Big| \le \prod_{v\in V} |\alpha_v|^{\deg_{T'} v} \le \prod_{v\in V} |\alpha_v|^{\deg_{\mathcal{F}} v} \tag{7.78}$$

Then

$$\begin{aligned} &\Big| \sum_{T'\in\mathcal{F}} \prod_{vv'\in E(T')} \alpha_v \cdot \alpha_{v'} \prod_V \delta_v^{-2} \Big| \\ &\quad \le \prod_{vv'\in E(\overline{\mathcal{R}})\cup E(\overline{T})} |\alpha_v|\ |\alpha'_v| \sum_{\{j_0,...,j_s\}} \prod_{C\in\mathcal{C}} |\alpha_{u(C)j_0}|\ |\alpha_{w(C)j_s}|\ \times \\ &\qquad \times \prod_{\overline{R}\in\overline{\mathcal{R}}} \Big| \sum_{\bar{u},\bar{w}\in\overline{R}} \alpha_{\bar{u}i_0}\alpha_{\bar{w}i'_0} \prod_{v\in T'} \delta_v(T')^{-2} \Big| \\ &\quad \le \prod_V |\alpha_v|^{\deg_{\mathcal{F}} v}\ N^{|V|} \sup_{\substack{T'\in\mathcal{F}\\ 1\le i(\overline{R}),i'(\overline{R})\le N}} \Big| \sum_{\bar{u},\bar{w}\in\overline{R}} \alpha_{\bar{u}i}\alpha_{\bar{w}i'} \prod_V \delta_v(T')^{-2} \Big| \end{aligned} \tag{7.79}$$

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 $R_0 \in \mathcal{R}_0$: $\mathcal{H}_\lambda(R_0) = \{R_1, ..., R_h\}$ and $R_* = R_h$. Now, observing that

$$\overline{T} = \overline{T'/R_*}\ , \qquad \overline{\mathcal{R}}(T'/R_*) \equiv \overline{\mathcal{R}}\backslash\{R_*\} \tag{7.80}$$

we have

$$\prod_{R\in\overline{\mathcal{R}}(T')}\sum_{\bar u,\bar w\in\overline{\mathcal{R}}}\alpha_{\bar u i}\alpha_{\bar w i'}\prod_{v\in T'}\delta_v(T')^{-2}$$
$$=\Big(\prod_{R\in\overline{\mathcal{R}}(T'/R_*)}\sum_{\bar u,\bar w\in\overline{\mathcal{R}}}\alpha_{\bar u i}\alpha_{\bar w i'}\prod_{v\in T'}\delta_v(T'/R_*)^{-2}\Big)\Big(\sum_{\bar u,\bar w\in R_*}\alpha_{\bar u i}\alpha_{\bar w i'}\prod_{v\in R_*}\delta_v(T')^{-2}\Big) \tag{7.81}$$

If $h>0$ (*i.e.* R_* is a critical subresonance) and if m_i, for $i=0,1,...,h$, are the integer vectors associated to the hierarchy $\{R_1,...,R_h\}$ (see Remark 5.3 of Section 5) we have for some $\sigma_h=\pm 1$ and some $\sigma_i=0,\pm 1$, (for $i=0,...,h-1$)

$$m_0\equiv\sum_{v<R_0}\alpha_v=\sum_{i=0}^{h}\sigma_i m_i\ ,\qquad (|\sigma_h|=1) \tag{7.82}$$

so that, by (5.13), we have

$$|\langle m_0\rangle|<\frac{\lambda}{1-\lambda}\ \Delta(R_*) \tag{7.83}$$

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 6.3 to R_* with $\bar\lambda\equiv\lambda/(1-\lambda)<1$, obtaining

$$\Big|\sum_{\bar u,\bar w\in R_*}\alpha_{\bar u i}\alpha_{\bar w i'}\prod_{V}\delta_v(T')^{-2}\Big|\le c_3^{|R_*|}\prod_{v\in R_*}|\alpha_v|^{\beta_3} \tag{7.84}$$

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 6.2 can be applied. Thus Lemma 6.4 follows with

$$c_4\equiv N\ \max\{\ c_3\ ,\ c_2^2\}\ ,\qquad \beta_4\equiv\ \max\{\ \beta_3\ ,\ 2\beta_2+2\} \tag{7.85}$$

which by (6.7) yield (6.9). ∎

A Combinatorics

Here we give the proofs of Lemma 3.1, Corollary 3.1, Proposition 3.2 and Proposition 3.1.

Proof of Lemma 3.1

Proof of (i). Denote by $\widetilde{\mathcal{T}}_0^k$ the set of labeled (unrooted) trees of order k. Then

$$t_{kj}=k\ \#\{T\in\widetilde{\mathcal{T}}_0^k:\ \deg v_k=j-1\} \tag{A.1}$$

Now, recalling (Appendix B) that given integers $d_i\ge 1$ such that $\sum_{i=1}^k d_i=2k-2$, one has

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

we obtain

$$\begin{aligned} t_{kj} &= k \sum_{T\in\widetilde{\mathcal{T}}_0^k:\ \deg v_k=j-1} 1 = k \sum_{\substack{d_k=j-1,d_i\geq 1\\ \sum d_i=2k-2}} \frac{(k-2)!}{\prod_{i=1}^k (d_i-1)!} \\ &= k\binom{k-2}{j-2} \sum_{\substack{s_i\geq 0\\ s_1+\cdots+s_{k-1}=k-j}} \frac{(k-j)!}{\prod_{i=1}^{k-1} s_i!} = k\binom{k-2}{j-2}(k-1)^{k-j} \end{aligned}$$

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 $(\sum h_i)!/(h_1!\cdots h_{j-1}!) = (k-1)!/(h_1!\cdots h_{j-1}!)$ ways of doing such a choice. Now, we attach the $j-1$ trees T_i according to the rule described in (3.17) (with $s=j-1$) so as to form a generic element of $\mathcal{T}^k$: obviously, since changing the order in (3.17) 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.* $j=2,...,k$. ∎

Proof of Corollary 3.1 is an immediate consequence of the argument used to prove (ii) above. ∎

Proof of Proposition 3.2 We prove (3.15) by induction on $k\geq 1$. For $k=1$ the claim is trivially true. Let $k\geq 2$ and assume that (3.15) holds for $1,...,k-1$. Denoting, for $T\in\widetilde{\mathcal{T}}^h$

$$F(T) \equiv \sum_{\alpha:T\to N}\prod_{v\in V}\phi_{\alpha_v}|\alpha_v|^{\deg v} \tag{A.3}$$

in view of (3.18) we find

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

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

$$\sum_{n\in N}\phi_n|n|^j F(\widetilde{T}_1)\cdots F(\widetilde{T}_{j-1}) = F(\widetilde{T}_1 * \cdots * \widetilde{T}_{j-1}) \tag{A.4}$$

This finishes the proof of Proposition 3.2. ∎

Remark A.1 In fact, it is easy to see from (3.15) that

$$g_k = \frac{1}{k!} \sum_{n_1,\dots,n_k} \Big(\prod_{i=1}^{k} \phi_{n_i} |n_i| \Big) (|n_1| + \cdots |n_k|)^{k-1} \tag{A.5}$$

However, a direct (arithmetic) proof of (3.15), starting from the explicit expression (A.5), might not be simpler than the proof presented above.

Proof of Proposition 3.1 Notice that F_{mn} in (3.21) satisfies, for any $\widetilde{T}_i \in \widetilde{\mathcal{T}}^{h_i}$, $(i = 1, \dots, j-1)$,

$$\begin{aligned} \langle n \rangle^{-2} \sum_{\substack{n_1,\dots,n_{j-1} \\ n_1+\cdots n_{j-1}=n}} & m \cdot n_j \; f_{n_j} \; F_{n_j n_1}(\widetilde{T}_1) \cdots F_{n_j n_{j-1}}(\widetilde{T}_{j-1}) \\ &= F_{mn}(\widetilde{T}_1 * \cdots * \widetilde{T}_{j-1}) \end{aligned} \tag{A.6}$$

Now, mimic the above proof using (A.6) in place of (A.4). ∎

B 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.* [13] 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 $0 \leq |E(G)| \leq \binom{k}{2}$ ($|E| = 0 \iff E = \emptyset$). 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 correspondently 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 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 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*.

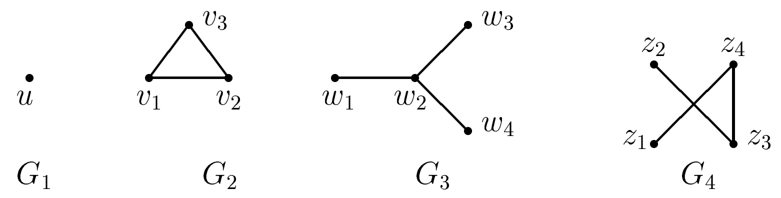

Figure 13: $G = G_1 \cup G_2 \cup G_3 \cup G_4$

The set of vertices of the graph G in Figure 13 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 ($\iff$ 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 is 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 13 and the graphs depicted in Figure 14 all represent the same graph.

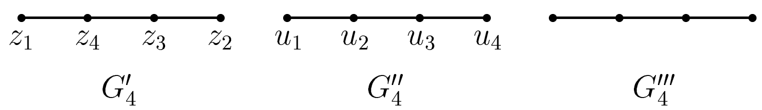

Figure 14: Graphs isomorphic to G_4

An isomorphism 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 thorough paths: A *path* P is a graph whose vertices can be given an order such that if $V(P) = \{v_1, ..., v_p\}$ then $E(P)$ is given by $\{v_1v_2, ..., v_{p-1}v_p\}$ (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 *closed path* or *cycle* is defined as a path but with $v_1 = v_p$. A graph G is *connected* if any two (different) points can be joined by a path in G. G_4 in Figure 13 (and in Figure 14) is a path.

A *tree* is a connected acyclic (*i.e.* with no cycles) graph. In Figure 13, 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.* [13]): (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.

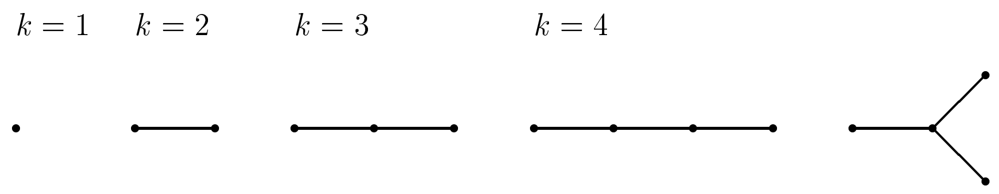

Figure 15: Trees up to order $k = 4$

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*, ηr, in $E(T)$: Thus for rooted trees it is $|E(T)| = |V(T)|$ and the unique edge of the form ηv identifies the root v (usually denoted r). Consistently, we let the degree of the root r be the number of edges incident with r, ηr *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 ηv): $\deg_{T_v} v = \deg_T v + 1$. 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

$$\begin{aligned} \mathcal{T}^k &\equiv \{ \text{ labeled rooted trees of order } k\} \ , \qquad & \mathcal{T}_0^k &\equiv \{ \text{ labeled trees of order } k\} \\ \widetilde{\mathcal{T}}^k &\equiv \{ \text{ rooted trees of order } k\} \ , \qquad & \widetilde{\mathcal{T}}_0^k &\equiv \{ \text{ trees of order } k\} \end{aligned} \tag{B.1}$$

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 $t_k = k t_k^{(0)}$. Less obvious are the following statements

$$\text{given } d_i \geq 1 \text{ such that } \textstyle\sum d_i = 2k-2, \quad \#\{T \in \widetilde{\mathcal{T}}^k : \ \deg_T v_i = d_i\} = \frac{(k-2)!}{\prod_{i=1}^k (d_i - 1)!}$$

$$\text{given } d_i \geq 1 \text{ such that } \textstyle\sum d_i = 2k-1, \quad \#\{T \in \mathcal{T}^k : \ \deg_T v_i = d_i\} = \frac{(k-1)!}{\prod_{i=1}^k (d_i - 1)!}$$

$$t_k \equiv \#\mathcal{T}^k = k^{k-1} \ , \qquad t_k^{(0)} \equiv \#\mathcal{T}_0^k = k^{k-2} \tag{B.2}$$

Proofs of these relations can be found in [5] (chapter VII, §3).

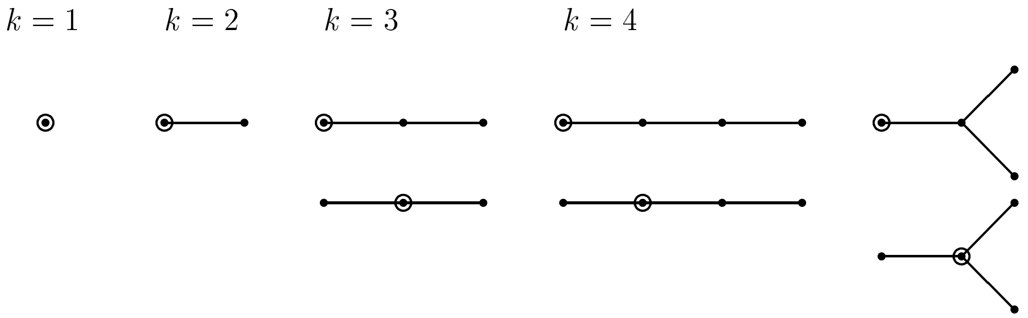

Figure 16: Rooted trees up to order $k = 4$

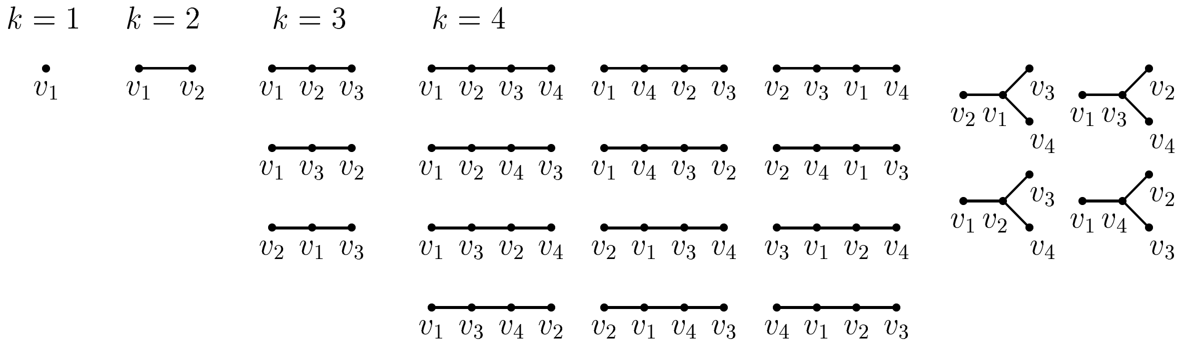

Figure 17: Labeled trees on $\{v_1, ..., v_k\}$, for $k = 1, ..., 4$

References

[1] Arnold, V.I.: *Proof of A.N. Kolmogorov's theorem on the preservation of quasi–periodic motions under small perturbation of the Hamiltonian.* Usp. Mat. Nauk **18**, No. 5, 13–40 (1963) (Russian); English trans.: Russ. Math. Surv. **18**, No. 5, 9–36 (1963).

[2] Arnold, V.I.: *Small denominators and problems of stability of motions in classical and celestial mechanics.* Usp. Mat. Nauk **18**, No. 6, 91–192 (1963) (Russian); English trans.: Russ. Math. Surv. **18**, No. 6, 85–192 (1963).

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

[4] Berretti, A.; Celletti, A.; Chierchia, L.; Falcolini, C.: *Natural boundaries for area–preserving twist maps*, Journal of Statistical Physics, **66** 1613–1630 (1992).

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

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

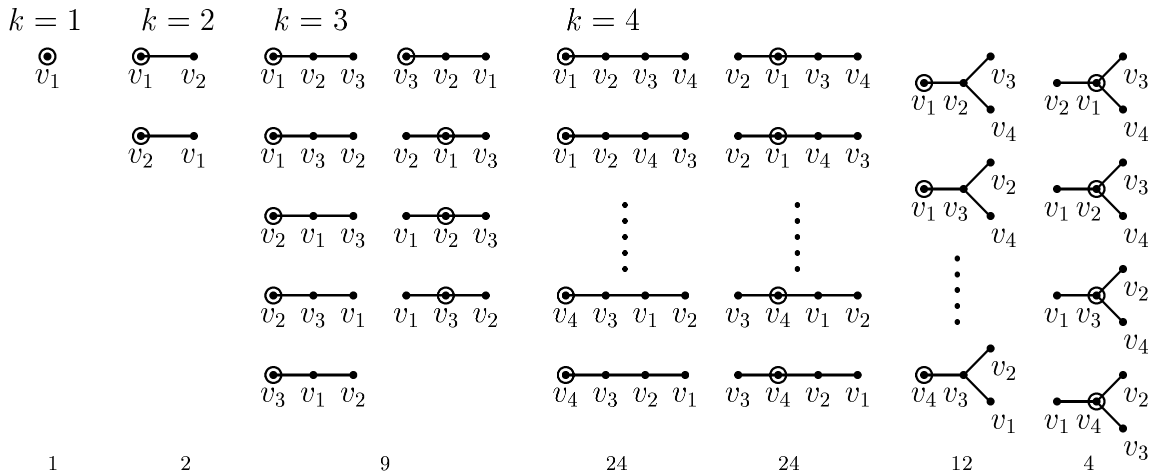

Figure 18: Labeled rooted trees on $\{v_1, ..., v_k\}$, for $k = 1, ..., 4$

[7] Chierchia, L.; Gallavotti, G.: *Drift and Diffusion in phase space*, Preprint, 1–135 (1992), to appear in Ann. Inst. Henri Poincaré (Physique Théorique).

[8] Chierchia, L.; Perfetti, P.: *Second order Hamiltonian equations on $^\infty$ and almost–periodic solutions*, Preprint, 1–30 (1992), to appear in Journal of Differential Equations.

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

[10] Falcolini, C.; de la Llave, R.: *Numerical calculation of domains of analyticity for perturbation theories in the presence of small divisors*, Journal of Statistical Physics, **67** No. 3/4, (1992).

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

[12] Goulden, I. P.; Jackson, D. M.: *Combinatorial Enumeration*, Wyley Interscience Series in Discrete Math., 1983.

[13] Harary, F.: *Graph Theory*, Addison–Wesley, 1969.

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

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

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

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

[18] Rüssmann, H.: *Kleine Nenner. I: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes.* Nachr. Akad. Wiss. Gött., II. Math.–Phys. Kl. 1970, 67–105 (1970).

[19] Schmidt, W. M.: *Diophantine Approximation*, Lecture Notes in Math. **785**, 1980.

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

[21] Vittot, M.: *Lindstedt perturbation series in Hamiltonian mechanics: explicit formulation via a multidimensional Burmann–Lagrange formula*, Preprint 1–20 (1991)