Whiskered tori with prefixed frequencies and Lyapunov spectrum *

Guido Gentile[1]

Abstract: *A classical mechanics problem, as the existence of whiskered tori for an almost integrable hamiltonian system, is analyzed with techniques reminiscent of the quantum field theory, following the strategy developed in recent works about the matter. The system consists in a collection of rotators interacting with a pendulum via a small potential depending only on the angle variables. The proof of the existence of the stable and unstable manifolds ("whiskers") of the rotators invariant tori corresponding to diophantine rotation numbers is simplified by setting the Lyapunov spectrum to prefixed values via the introduction, in the hamiltonian function, of "counterterms" depending on the strength of the interaction; this is a feature usual in quantum field theory, and emphasizes the analogy between the the field theory and the KAM framework pointed out already in the mentioned works.*

Key words: *KAM, perturbation theory, classical mechanics, quantum field theory, renormalization group*

1. Introduction

Until the Eliasson's work, [E], proofs of the KAM theorem, ([K], [A1], [Mo1]), assuring the existence of infinetely many invariant tori for perturbed systems, were all based on a rapidly convergent iteration technique leading to solve recursively a sequence of approximate equations. Recently, direct (*i.e.* non recursive) proofs of the KAM theorem have been proposed, (see [E], [G1], [G2], [CF]), by exhibiting explicitly cancellations occuring between contributions to the formal expansion series of the quasiperiodic solutions. By extending the important ideas of Melnikov, [Me], the same techniques have been applied to the study of low dimensional tori and their stable and unstable manifolds, [G1]. In particular in [G1], [G2], [GGe], [G5], the analogy between the methods implemented in the direct proofs and those used in quantum field theory, expecially in the renormalization group approach, have been pointed out.[2] In this work such ideas will be developed further, by giving a proof of the existence of the whiskered tori in a class of almost integrable hamiltonian systems alternative to that proposed in [Ge], and based on the introduction of "counterterms" (a technique usual in quantum field theory) in the hamiltonian function; the advantage of such a procedure is that some symmetry properties of the solutions, discussed in [CG] within the KAM theory framework, can be directly visualized for the modified hamiltonian, so simplifying in a remarkable way the topological structure of the diagrams (trees, see below) in terms of which the solutions can be graphically represented. Once the problem is solved, the solutions of the original hamiltonian equations can be recovered in a trivial way, and, by construction, they exhibit explicitly the symmetry properties required by the results found in [CG]. Actually, the original feature with respect to [Ge] resides on the simplification in the proof and on the connection with the quantum field theory methods; therefore the present paper can be considered an introductory work in view of an extension to other similar dynamical systems problems, whose treatment requires such methods.

We consider a model consisting of a family of rotators, say $l-1$ in number, interacting with a pendulum via a conservative force, (see [CG], [G1], [GGe]). The inertia moments J_j, $j=1,\ldots,l-1$, of the rotators form a matrix J which is diagonal, and are supposed to be $J_j \ge J_0 > 0$, if J_0 is the inertia of the pendulum, so setting a scale for the size of the inertia moments; for simplicity's sake we assume $J_j = J$, for all $j=1,\ldots,l-1$.

---

* This paper is deposited in the archive `mp_arc@math.utexas.edu`, #94-??.

[1] E-mail: `gentileg%39943.hepnet@lbl.gov`: Dipartimento di Fisica, Università di Roma, "La Sapienza", P. Moro 5, 00185 Roma, Italia.

[2] In [G5] it has been shown that the function describing the invariant tori of a system of rotators interacting via a potential which is a trigonometrical polynomial in the angle variables, (*i.e.* the hamiltonian system (1.1) without the pendulum), can be represented as the one point Schwinger function of a quantum field model.

The model can be described by the l degrees of freedom hamiltonian $H_\mu \equiv H_0 + \mu f$ given by

$$\vec{\omega}\cdot\vec{A} + \frac{1}{2}J^{-1}\vec{A}\cdot\vec{A} + \frac{I^2}{2J_0} + g^2 J_0(\cos\varphi - 1) + \mu\sum_{|\nu|\le N} f_\nu \cos(\vec{\alpha}\cdot\vec{\nu} + n\varphi) \tag{1.1}$$

where $(I,\varphi)\in\mathbf{R}^2, (\vec{A},\vec{\alpha})\in\mathbf{R}^{2(l-1)}$ are canonically conjugated variables, $\vec{\omega}\in\mathbf{R}^{l-1}$, $\nu\equiv(n,\vec{\nu})\in\mathbf{Z}^l$, $|\nu| = |n| + |\vec{\nu}| = |n| + \sum_{i=1}^{l-1}|\nu_i|$, $g>0$ (g^2 is the "gravity"), $\vec{\omega},\mu$ are parameters, ($\vec{\omega}$ is the rotation vector, and μ is the perturbation strenght), and f_ν are fixed constants. A natural *energy scale* for the model will be $E = g^2 J_0$. We suppose *a priori* that:

Hypothesis H: *The rotation vector* $\vec{\omega}$ *is a diophantine vector, i.e.* $C_0|\vec{\omega}\cdot\vec{\nu}| \ge |\vec{\nu}|^{-\tau}$, *for all* $\vec{0}\ne\vec{\nu}\in\mathbf{Z}^{l-1}$, *for some* diophantine constant C_0 *and some* diophantine exponent $\tau>0$.

For $\mu=0$, the hamiltonian equations generated by (1.1), (*i.e.* $\dot{I} = -\partial_\varphi H_\mu$, $\dot{\varphi} = \partial_I H_\mu$, $\dot{\vec{A}} = -\partial_{\vec{\alpha}} H_\mu$, $\dot{\vec{\alpha}} = \partial_{\vec{A}} H_\mu$), admit $(l-1)$–dimensional invariant tori:

$$\mathcal{T}_0 \equiv \{I = 0 = \varphi\}\times\{\vec{A}\equiv\vec{A}^0\ ,\vec{\alpha}\in\mathbf{T}^{l-1}\} \tag{1.2}$$

possessing homoclinic stable and unstable manifolds, called *whiskers*. The manifolds equations are:

$$W_0^\pm \equiv W_0 \equiv \{\tfrac{I^2}{2J_0} + g^2 J_0(\cos\varphi - 1) = 0\}\times\{\vec{A}\equiv\vec{A}^0\ ,\vec{\alpha}\in\mathbf{T}^{l-1}\} \tag{1.3}$$

so that the non trivial Lyapunov exponents of $\mathcal{T}_0$ are $\pm g$.

Under perturbation we expect the tori to be deformed and to remain unstable with only two non trivial Lyapunov exponents $\pm g'(\mu)$, where $g'(\mu)$ is a suitable analytic function of μ; this is indeed proved in [CG] and [Ge]. The change of the Lyapunov exponents has the consequence that the solutions of the perturbed motion equations depend naturally on t, (besides via $\exp[\pm i\vec{\omega}\cdot\vec{\nu}t]$), via $e^{\pm g'(\mu)t} = e^{\pm gt}[1 \pm \mu g_1 t + \mu^2(\pm g_2^2 t + g_1^2 t^2/2) + O(\mu^3)]$, where g_i, $i>0$, are the coefficients of the series expansion of the function $g'(\mu)$. Therefore, in a perturbation theory, this will generate contributions to the solutions of the equations of motion which depend on time via $e^{\pm gt}\,t^k$ for various values of $k\in\mathbf{N}$. This indeed does happen, [Ge], and it is the source of rather deep technical intricacies: then one has to exhibit some cancellations which ultimately are probably related to the fact that the terms $e^{\pm gt}\,t^k$ appearing in the solutions of the equations of motion can be resummed into "pure exponentials" like $e^{\pm g'(\mu)t}$.

It is therefore natural to ask if one can fix the Lyapunov exponents *a priori*, as one does with the frequencies. This should allow us to get rid of the terms $e^{\pm gt}\,t^k$ and leave us with "pure exponentials", eliminating at the same time also the necessity of the analysis of cancellations between terms with dependence on t through powers of t.

Then, instead of studying the model (1.1), we consider the modified model:

$$\vec{\omega}\cdot\vec{A} + \frac{1}{2}J^{-1}\vec{A}\cdot\vec{A} + \frac{I^2}{2J_0} + g^2(\mu) J_0(\cos\varphi - 1) + \mu\sum_{|\nu|\le N} f_\nu \cos(\vec{\alpha}\cdot\vec{\nu} + n\varphi) \tag{1.4}$$

where $g^2(\mu)$ is a function of μ admitting the series expansion:

$$g^2(\mu) = g^2\sum_{h=0}^{\infty}\gamma_h\mu^h \tag{1.5}$$

with $\gamma_0 = 1$ and the coefficients γ_h, $h\ge 1$, to be determined, as it will be explained later, (see §5), in such a way to guarantee that the frequencies and the non trivial Lyapunov exponents corresponding to the perturbed system whiskered tori are, respectively, $\vec{\omega}$ and $\pm g$, (and we shall see that, if such a request is fulfilled, then the series convergence automatically follows).

We shall prove that the perturbed system still admits $(l-1)$-dimensional invariant tori $\mathcal{T}_\mu$, obtained by analytic continuation from the unpertured ones, with their stable and unstable whiskers

$$W_\mu^\pm = \{(I,\vec{A},\varphi,\vec{\alpha}) = (I_\mu^\pm(\vec{\alpha},\varphi), \vec{A}_\mu^\pm(\vec{\alpha},\varphi),\varphi,\vec{\alpha}) : \vec{\alpha}\in\mathbf{T}^{l-1}, |\varphi|<2\pi\} \tag{1.6}$$

characterized by the sets of initial data $X^\pm_\mu(0)$ such that the distance$(S^t_\mu X^\pm_\mu(0),\mathcal{T}_\mu)\to 0$ fast as $e^{-g'(\mu)|t|}$ $t\to\pm\infty$; here S^t_μ is the hamiltonian flow generated by (1.4).

The paper is selfcontained: §2 has a definitory nature and partially overlaps with §2, 3, 4 of [Ge], (which in turn were often literally taken from the review article [G1], with some abstraction effort), while §3 introduces the diagrammatic formalism, which is similar but not quite identical to the corresponding one of [Ge], and in fact is less involved, (less definitions have to be given). The original work is in §4÷§5 (and in the appendices). The theory of the homoclinic splitting, discussed in [G1] and briefly mentioned in [Ge], will be not even barely touched: here we confine ourselves to the problem of existence of the whiskered tori.

Acknowledgements : The author would thank G. Gallavotti for having proposed the argument and indicated to him the way how to approach it.

2. Analytic formalism

In this section we review the formalism which will be used in the proof of existence of the whiskered tori. With respect to [G1] and [Ge], there are a few differences here and there, but only about the notations.

2.1 Recursive formulae

We derive simple recursive formulae for the functions $I^\pm_\mu$, $\vec A^\pm_\mu$ in (1.6) and their time evolution (see also [G1], §2, and [CG], Appendix A10).

Let us consider the unperturbed motion:

$$X^0(t)\equiv(\varphi^0(t),\vec\alpha+\vec\omega t,I^0(t),\vec 0) \tag{2.1}$$

where $(\varphi^0(t),I^0(t))$ is the separatrix motion, generated by the pendulum in (1.1) starting, say, at $t=0$ in $\varphi=\pi$, and $\varphi^0(t)=4\arctan e^{-gt}$. Let $X^\sigma_\mu(t;\alpha)$, $\sigma=\operatorname{sign}t=\pm$, be the evolution, under the flow generated by (1.1), of the point on W^σ_μ given by $(\pi,\vec\alpha,I^\sigma_\mu(\vec\alpha,\pi),\vec A^\sigma_\mu(\vec\alpha,\pi))$, see (1.6); let:

$$X^\sigma_\mu(t)\equiv X^\sigma_\mu(t;\vec\alpha)\equiv\sum_{h\ge 0}X^{h\sigma}(t;\vec\alpha)\mu^h=\sum_{h\ge0}X^{h\sigma}(t)\mu^h,\qquad \sigma=\pm \tag{2.2}$$

be the power series in μ of X^σ_μ, (which we will show to be convergent for μ small); note that $X^{0\sigma}\equiv X^0$ is the unperturbed whisker. We shall often not write explicitly the $\vec\alpha$ variable among the arguments of various $\vec\alpha$ dependent functions, to simplify the notations, and we shall regard the two functions $X^{h\sigma}(t)$, as forming a single function $X^h(t)$, which is $X^{h+}(t)$ if $\sigma=+$, and $X^{h-}(t)$ if $\sigma=-$.

Inserting (2.2) into the Hamilton equation associated with (1.4), we see that the coefficients $X^{h\sigma}(t)$ satisfy the hierarchy of equations:

$$\frac{d}{dt}X^{h\sigma}\equiv\dot X^{h\sigma}=LX^{h\sigma}+F^{h\sigma} \tag{2.3}$$

where:

$$L(t)=\begin{pmatrix}0&\vec0&J_0^{-1}&\vec0\\ \vec0&0&\vec0&J^{-1}\\ g^2J_0\cos\varphi^0(t)&\vec0&0&\vec0\\ \vec0&0&\vec0&0\end{pmatrix},\qquad F^1(t)=\begin{pmatrix}0\\0\\-\partial_\varphi f(\varphi^0(t),\vec\alpha+\vec\omega t)+g^2\gamma_1J_0\sin(\varphi^0(t))\\-\partial_{\vec\alpha}f(\varphi^0(t),\vec\alpha+\vec\omega t)\end{pmatrix} \tag{2.4}$$

and where $F^{h\sigma}$ depends upon $X^0,...,X^{h-1\sigma}$ but not on $X^{h\sigma}$; here (as everywhere else) the arrows denote $(l-1)$–vectors. The entries of the $(2l\times 2l)$ matrix L have different meaning according to their position: the $\vec0$'s in the first and third row are $(l-1)$ (row) vectors, the $\vec0$'s in the first and third column are $(l-1)$ (column) vectors, and the 0's and J^{-1} in the second and fourth column are $(l-1)\times(l-1)$ matrices, while the 0's in the first and third columns are scalars.

Then, if we number the components of X with a label j, $j=0,\dots,2l-1$, with the convention that:

$$X_0=X_-,\quad (X_j)_{j=1,\dots,l-1}=\vec X_\downarrow,\quad X_l=X_+,\quad (X_j)_{j=l+1,\dots,2l-1}=\vec X_\uparrow \tag{2.5}$$

($i.e.$ we write first the angle and then the action components; first the pendulum and then the rotators), we see that (2.3) takes the form:

$$\begin{aligned}
&\frac{d}{dt}X_+^{h\sigma} = (g^2 J_0 \cos\varphi^0) X_-^{h\sigma} + F_+^{h\sigma}\ , \qquad && \frac{d}{dt}\vec X_\uparrow^{h\sigma} = \vec F_\uparrow^{h\sigma}\\
&\frac{d}{dt}X_-^{h\sigma} = J_0^{-1} X_+^{h\sigma}\ , && \frac{d}{dt}\vec X_\downarrow^{h\sigma} = J^{-1}\vec X_\uparrow^{h\sigma}
\end{aligned} \tag{2.6}$$

as $F_-^{h\sigma}$, $\vec F_\downarrow^{h\sigma}$ vanish identically, for $h \ge 1$. And, for all $h \ge 1$, we can write the following formula for $F^{h\sigma}$ in terms of the coefficients $X^0, \ldots, X^{h-1\sigma}$ and of the derivatives of H_0 and f:

$$\begin{aligned}
&F_-^{h\sigma} \equiv 0\ , \qquad \vec F_\downarrow^{h\sigma} \equiv 0\ ,\\
&\vec F_\uparrow^{h\sigma} = -\sum_{|\vec m|\ge 0} (\partial_{\vec\alpha} f)_{\vec m}(\varphi^0, \vec\alpha + \vec\omega t) \sum_{(h_j^i)_{\vec m, h-1}} \prod_{i=0}^{l-1}\prod_{j=1}^{m_i} X_i^{h_j^i \sigma}\ ,\\
&F_+^{h\sigma} \equiv \sum_{|\vec m|\ge 2} (g^2 J_0 \sin\varphi)_{\vec m}(\varphi^0) \sum_{(h_j^0)_{\vec m, h}} \prod_{j=1}^{m} X_-^{h_j^0\sigma} + \sum_{|\vec m|\ge 0}\sum_{p=1}^{h} (g^2\gamma_p J_0 \sin\varphi)_{\vec m}(\varphi^0) \sum_{(h_j^0)_{\vec m, h-p}} \prod_{j=1}^{m} X_-^{h_j^0\sigma}\\
&\qquad - \sum_{|\vec m|\ge 0} (\partial_\varphi f)_{\vec m}(\varphi^0, \vec\alpha + \vec\omega t) \sum_{(h_j^i)_{\vec m, h-1}} \prod_{i=0}^{l-1}\prod_{j=1}^{m_i} X_i^{h_j^i\sigma}\ ,
\end{aligned} \tag{2.7}$$

where $(G)_{\vec m}(\cdot)$, with $G = \partial_{\vec\alpha} f, \partial_\varphi f, g^2 J_0 \sin\varphi$, and $(h_j^i)_{\vec m, q}$, with $h_j^i \ge 1$, $m_i \ge 0$, $\vec m = (m_0, \ldots, m_{2l-1})$, are defined as:

$$\begin{aligned}
&(G)_{\vec m}(\cdot) \equiv \Big(\frac{\partial_\varphi^{m_0}\partial_{\alpha_1}^{m_1}\ldots\partial_{\alpha_{l-1}}^{m_{l-1}}\partial_I^{m_l}\partial_{A_1}^{m_{l+1}}\ldots\partial_{A_{l-1}}^{m_{2l-1}}\, G}{m_0!\, m_1!\,\ldots\, m_{l-1}!\, m_l!\, m_{l+1}!\,\ldots\, m_{2l-1}!}\Big)(\cdot)\\
&(h_j^i)_{\vec m, q} \equiv (h_1^0, \ldots, h_{m_0}^0, h_1^1, \ldots, h_{m_1}^1, \ldots, h_1^{2l-1}, \ldots, h_{m_{2l-1}}^{2l-1}) \qquad \text{s.t.} \sum h_j^i = q
\end{aligned} \tag{2.8}$$

Note that the first two sums in the expression for $F_+^{h\sigma}$ can only involve vectors $\vec m$ with $m_j = 0$ if $j \ge 1$, because the function $J_0 g^2 \sin\varphi$ depends only on φ and not on $\vec\alpha$). The evolution of X^h is determined by integrating (2.6), if the initial data are known. The $h = 1$ case requires a suitable interpretation of the symbols, in according to equation (2.4).

We introduce the dimensionless quantities related to the perturbed motions by:

$$X_j^{h\sigma} = \Xi_j^{h\sigma}, \quad (0 \le j \le l-1)\ , \qquad X_l^{h\sigma} = g J_0\, \Xi_l^{h\sigma}, \qquad X_j^{h\sigma} = g J\, \Xi_j^{h\sigma}, \quad (l+1 \le j \le 2l-1),$$

and to the functions $F^{h\sigma}$ through the transformations $\vec F_\uparrow^{h\sigma} = (g^2 J)^{-1}\vec\Phi_\uparrow^{h\sigma}$, $F_+^{h\sigma} = (g^2 J_0)^{-1}\Phi_+^{h\sigma}$, (obviously: $\Phi_j^{h\sigma} = F_j^{h\sigma} \equiv 0$, for $j = 0, \ldots, l-1$).

We recall that the *wronskian matrix* $W(t)$ of a solution $t \to x(t)$ of a differential equation $\dot x = f(x)$ in $\mathbf{R}^n$ is a $n \times n$ matrix whose columns are formed by n linearly independent solutions of the linear differential equation obtained by linearizing f around the solution x and assuming $W(0) =$ identity.

The solubility by elementary quadrature of the free pendulum equations on the separatrix leads after a well known classical calculation to the following expression for the wronskian $W(t)$ of the separatrix motion of the pendulum appearing in (1.1), with initial data at $t = 0$ given by $\varphi = \pi, I = 2gJ_0$, ($i.e.$ $\Xi_+^0 = 2$):

$$W(t) = \begin{pmatrix} \frac{1}{\cosh gt} & \frac{\bar w(t)}{4} \\ -\frac{\sinh gt}{\cosh^2 gt} & (1 - \frac{\bar w(t)}{4}\frac{\sinh gt}{\cosh^2 gt})\cosh gt \end{pmatrix}, \qquad \bar w(t) \equiv \frac{2gt + \sinh 2gt}{\cosh gt} \tag{2.9}$$

And the evolution of the $\pm$ ($i.e.$ I, φ) components can be determined by using the above wronskian:

$$\begin{pmatrix} \Xi_-^{h\sigma} \\ \Xi_+^{h\sigma} \end{pmatrix} = W(t)\begin{pmatrix} 0 \\ \Xi_+^{h\sigma}(0) \end{pmatrix} + W(t)\int_0^{gt} W^{-1}(\tau)\begin{pmatrix} 0 \\ \Phi_+^{h\sigma}(\tau) \end{pmatrix}\, d\,g\tau \tag{2.10}$$

Thus, denoting by w_{ij} $(i,j=0,l)$ the entries of W we see immediately that:

$$\begin{aligned}
\Xi_+^{h\sigma}(t) &= w_{ll}(t)\Xi_+^{h\sigma}(0) + w_{ll}(t)\int_0^{gt} w_{00}(\tau)\Phi_+^{h\sigma}(\tau)\,d\,g\tau - w_{l0}(t)\int_0^{gt} w_{0l}(\tau)\Phi_+^{h\sigma}(\tau)\,d\,g\tau \\
\Xi_-^{h\sigma}(t) &= w_{0l}(t)\Xi_+^{h\sigma}(0) + w_{0l}(t)\int_0^{gt} w_{00}(\tau)\Phi_+^{h\sigma}(\tau)\,d\,g\tau - w_{00}(t)\int_0^{gt} w_{0l}(\tau)\Phi_+^{h\sigma}(\tau)\,d\,g\tau
\end{aligned} \tag{2.11}$$

The integration of the equations (2.6) for the $\uparrow,\downarrow$ components yields:

$$\begin{aligned}
\vec\Xi_\uparrow^{h\sigma}(t) =& \vec\Xi_\uparrow^{h\sigma}(0) + \int_0^{gt} \vec\Phi_\uparrow^{h\sigma}(\tau)\,d\,g\tau \\
\vec\Xi_\downarrow^{h\sigma}(t) =& \Big(gt\vec\Xi_\uparrow^{h\sigma}(0) + \int_0^{gt} g(t-\tau)\,\vec\Phi_\uparrow^{h\sigma}(\tau)\,d\,g\tau\Big)
\end{aligned} \tag{2.12}$$

having used that the $\vec\Xi_\downarrow^{h\sigma}(0)\equiv\vec 0$ because the initial datum is fixed and μ independent; and (2.11), (2.12) can be used to find a reasonably simple algorithm to represent the whiskers equations to all orders $h\ge1$ of the perturbation expansion.

2.2 The improper integration $\mathcal{I}$.

We introduce some integrations operations which can be performed on the functions introduced in §2.1. The operation is simply the integration over t from $\sigma\infty$ to t, $\sigma=\operatorname{sign} t$. In general such an operation cannot be defined as an ordinary integral of a summable function, because the functions on which it has to operate (typically the integrands in (2.11) and (2.12)) do not, in general, tend to 0 as $t\to\infty$. But the simplicity of the initial hamiltonian has the consequence that the functions $\Xi^h(t)$, and the matrix elements w_{ij} in (2.9), belong to a very special class of analytic functions on which the integration operations that we need can be given a meaning.

To describe such class we introduce various spaces of functions; all of them are subspaces of the space $\hat{\mathcal{M}}$ of the functions of t defined as follows.

Definition 2.1: *Let $\hat{\mathcal{M}}$ be the space of the functions of t which can be represented, for some $h\ge0$, as:*

$$M(t)=\sum_{j=0}^{h}\frac{(\sigma tg)^j}{j!}M_j^\sigma(x,\vec\omega t)\ ,\quad x\equiv e^{-\sigma gt}\ ,\quad \sigma=\operatorname{sign} t \tag{2.13}$$

with $M_j^\sigma(x,\vec\psi)$ a trigonometric polynomial in $\vec\psi$ with coefficients holomorphic in the x-plane in the annulus $0<|x|<1$, with: 1) possible singularities, outside the open unit disk, in a closed cone centered at the origin, with axis of symmetry on the imaginary axis and half opening $d<\frac{\pi}{2}$; 2) possible polar singularities at $x=0$; 3) $M_h^\sigma\ne0$. The number h will be called the t–degree of M. The smallest cone containing the singularities will be called the singularity cone *of M.*

Definition 2.2: *Let $\hat{\mathcal{M}}_0$ be the subspace of the functions $M\in\hat{\mathcal{M}}$ such that the residuum at $x=0$ of $x^{-1}\langle M_j^\sigma(x,\cdot)\rangle$ is zero (here the average is over $\vec\psi$, i.e. it is an "angle average").*

Definition 2.3: *Let $\mathcal{M}$ and $\mathcal{M}_0$ be the subspaces of the functions $M\in\hat{\mathcal{M}}$ and, respectively, $M\in\hat{\mathcal{M}}_0$ bounded near $x=0$.*

Definition 2.4: *Let $\hat{\mathcal{M}}^h,\hat{\mathcal{M}}_0^h,\mathcal{M}^h,\mathcal{M}_0^h$ denote the subspaces of $\hat{\mathcal{M}},\hat{\mathcal{M}}_0,\mathcal{M},\mathcal{M}_0$, respectively, containing the functions of t–degree $\le h$.*

In the following part of this section we describe briefly the properties of the functions contained in the above defined spaces, referring to [G1] for details:

1) If a function admits a representation like (2.13), with the above properties, then such a representation is unique (see [CG], §10).

2) If $M\in\mathcal{M}$, or $M\in\mathcal{M}_0$, then M_j^σ have no pole at $x=0$ and, furthermore, $M_j^\sigma(0,\vec\psi)=0$ if $j>0$.

3) $M\in\hat{\mathcal{M}}$ can be written as $M=P+M'$ with P being a polynomial in σt (with σ dependent coefficients)

and with $M' \in \hat{\mathcal{M}}_0$: this can be done in only one way and we call P the "polynomial component" of M, and M' the "non singular" component of M.
4) $M \in \mathcal{M}$ can be written as $M = p + M'$ with p being a constant function (with constant value depending on σ) and $M' \in \mathcal{M}_0$: p will be called the "constant component" of M, and M' will be the "non singular" component of M.
5) The functions in $\hat{\mathcal{M}}$ can be expanded as sums of the following monomials:

$$\sigma^\chi \, \frac{(\sigma t g)^j}{j!} x^h e^{i\vec{\omega}\cdot\vec{\nu}\, t} \tag{2.14}$$

where $\chi = 0, 1$ (*i.e.* the (2.14) span the space $\hat{\mathcal{M}}$).
6) The coefficients of the above mentioned expansions and polynomials depend on $\sigma = \pm$, *i.e.* each $M \in \hat{\mathcal{M}}$ is, in general, a pair of functions M^σ defined and holomorphic for $t > 0$ and $t < 0$, respectively (and, more specifically, in a domain $\{\sigma\,\mathrm{Re}\, t > 0, \, |\,\mathrm{Im}\, gt| < \pi/2 - d \equiv \xi\}$). The functions $M^\sigma(t)$ might sometimes (as in our cases below) be continued analytically in t but in general $M^+(-t) \neq M^-(-t)$ even when it makes sense (by analytic continuation) to ask whether equality holds.
7) If $M \in \mathcal{M}$ the points with $\mathrm{Re}\, t = 0$ and $|\,\mathrm{Im}\, gt| < \xi$ ($gt = \pm i\pi/2$ corresponds to $x = \mp i$) are, (by our hypothesis on the location of the singularities of the M_j functions), regularity points so that the values at $t^\pm$, "to the right" and "to the left" of t, will be regarded as well defined and given by $M(t^\pm) \equiv \lim_{t' \to t,\, \mathrm{Re}\, t' \to \mathrm{Re}\, t^\pm} M(t')$; in particular $M^\pm(0^\pm) \equiv M_0^\pm(1^-, \vec{0})$.
8) Since f in (1.1), or (1.4), is a trigonometric polynomial, the function F^1, see (2.4), belongs to $\mathcal{M}$ and, in fact, the component $\vec{F}^1_\uparrow$ belongs to $\mathcal{M}_0$ (as accidentally does F^1_+ as well).

On the class $\hat{\mathcal{M}}$ we can define the following operation.

Definition 2.5: *If $M \in \hat{\mathcal{M}}$, and $t = \tau + i\theta$, with τ, θ real, and $\tau = \mathrm{Re}\, t \neq 0$, $\sigma = \mathrm{sign}\ \mathrm{Re}\, t$, the function:*

$$\mathcal{I}_R M(t) \equiv \int_{\sigma\infty + i\theta}^{gt} e^{-Rg\sigma z} M(z)\, d\, gz \tag{2.15}$$

is defined for $\mathrm{Re}\, R > 0$ and large enough, the integral being on an axis parallel to the real axis. If $M \in \hat{\mathcal{M}}$ then the function of R in (2.15) admits an analytic continuation to $\mathrm{Re}\, R < 0$ with possible poles at the integer values of R and at the values $ig^{-1}\vec{\omega}\cdot\vec{\nu}$ with $|\vec{\nu}| <$ (trigonometric degree of M in the angles $\vec{\psi}$); and we can then set:

$$\mathcal{I} M(t) \equiv \oint \frac{dR}{2\pi i R} \mathcal{I}_R M(t) \tag{2.16}$$

where the integral is over a small circle of radius $r < 1$ and $r < \min |g^{-1}\vec{\omega}\cdot\vec{\nu}|$, the minimum being taken over the $\vec{\nu} \neq \vec{0}$ which appear in the Fourier expansion of M.

¿From the above definition one can immediately derive an expression for the action of $\mathcal{I}$ on the monomial (2.14) and check, in particular, that the radius of convergence in x of $\mathcal{I}M$, for a general M, is the same of that of M (but in general the singularities at $\pm i$ will no longer be polar, even if those of the M_j's were such). In general, $\mathcal{I} : \hat{\mathcal{M}}^h \to \hat{\mathcal{M}}^{h+1}$; but we note that the $\mathcal{I}$ operation does not increase the degree in t when $|h| + |\vec{\nu}| > 0$, (see [G1]).

One readily checks that $\mathcal{I}M$ is a primitive of M (*i.e.* the increment of $\mathcal{I}M$ between t_0 and t is the integral of M between the same extremes). The similarities of the $\mathcal{I}$ operation with a definite integral justify the use of the notation:

$$\fint_{(\sigma)}^{gt} M(\tau)\, d\, g\tau \equiv \mathcal{I}M(t) \ , \qquad M \in \hat{\mathcal{M}}, \ \sigma = \mathrm{sign}\ \mathrm{Re}\, t \tag{2.17}$$

In fact many standard properties of integration are, in such a way, extended to the space $\hat{\mathcal{M}}$, see [G1]. In particular we can define:

$$\fint_{\sigma\infty}^{gt} M(\tau)\, d\, g\tau \equiv \mathcal{I}M(0^\sigma) + \int_0^{gt} M(\tau)\, d\, g\tau \quad . \tag{2.18}$$

2.3 Analytic expressions of the expansion coefficients for the whiskers

We will show that the Ξ^h's defined through (2.2) admit rather simple expressions in terms of the operation $\mathcal{I}$ (and other related operations introduced below). Recall that in §2.1 we have fixed $\vec{\alpha} \in \mathbf{T}^{l-1}$ and $\varphi = \pi$, and we are looking for the motions, on the stable ($\sigma = +$) or unstable ($\sigma = -$) whisker, which start with the given $\vec{\alpha}$ and $\varphi = \pi$ at $t = 0$; in the following $\vec{\alpha}$ is kept constant and usually notationally omitted.

We suppose inductively that $\Xi^{h'} \in \mathcal{M}^{2h'-1}$, $h' < h$, and $\Phi^{h'} \in \mathcal{M}^{2(h'-1)}$, $\vec{\Phi}^{h'}_{\uparrow} \in \mathcal{M}_0^{2(h'-1)}$, $h' \le h$, and, furthermore, that the singularity cone consists of just the imaginary axis, *i.e.* the singularities of the functions defining Ξ^h, Φ^h are on the segments on the imaginary axis $(-i\infty, -i]$ and $[+i, +i\infty)$.

This means, in particular, that $\Phi^{h'}, \Xi^{h'}$ can be represented as:

$$\begin{aligned} \Phi^{h'}(x, \vec{\psi}, t) &= \sum_{j=0}^{2(h'-1)} \frac{(\sigma t g)^j}{j!} \Phi_j^{h'\sigma}(x, \vec{\psi}), \qquad h' = 1, \ldots, h \\ \Xi^{h'}(x, \vec{\psi}, t) &= \sum_{j=0}^{2h'-1} \frac{(\sigma t g)^j}{j!} \Xi_j^{h'\sigma}(x, \vec{\psi}), \qquad h' = 1, \ldots, h-1 \end{aligned} \tag{2.19}$$

by setting $\vec{\psi} = \vec{\omega} t$, $\sigma = \operatorname{sign} t$, $x = e^{-g\sigma t}$, with $\Phi_j^{h'\sigma}, \Xi_j^{h'\sigma}$ holomorphic at $x = 0$ and vanishing at $x = 0$ if $j > 0$. Hence if $x = e^{-g\sigma t}$ and $\vec{\psi}$ is kept fixed, the $\Phi_j^{h'}$'s, $\Xi_j^{h'}$'s tend exponentially to zero as $t \to \sigma\infty$, if $j > 0$; while if $j = 0$ they tend exponentially fast to a limit as $t \to \sigma\infty$ (*i.e.* as $x \to 0$), which we denote $\Phi^{h'}(\vec{\psi}, \sigma\infty)$ dropping the subscript 0 as there is no ambiguity.

Furthermore the inductive hypothesis is enriched by:

$$\vec{\Phi}^{h'\sigma}_{\uparrow\vec{0}}(\sigma\infty) = \vec{0}, \qquad \text{for all } h' \le h \tag{2.20}$$

recalling that, in general, a subscript $\vec{\nu}$ affixed to a function denotes the Fourier component of order $\vec{\nu} \in \mathbf{Z}^{l-1}$ of the considered function: $\Xi^{h'\sigma}_{j\vec{\nu}}(t)$ and $\Phi^{h'\sigma}_{j\vec{\nu}}(t)$ are the Fourier transforms in $\vec{\psi}$ of $\Xi_j^{h'\sigma}(t, \vec{\psi})$ and $\Phi_j^{h'\sigma}(t, \vec{\psi})$, respectively.

Let us suppose, just as an assumption for the time being, that $\Xi^{h'\sigma}(t)$ and, from (2.7), hence also $\Phi^{h'\sigma}(t)$ are bounded as $t \to \sigma\infty$ for all h', so that $\Xi_j^{h'\sigma}(0, \vec{\psi}) = 0$ if $j \ge 1$: we show then that the latter information is very strong and permits us to determine Ξ^h. This does not imply the convergence of the series: however in §5.2 such a result is proven, so justifying the boundedness hypothesis and completing the research of bounded motions.

We note that, since $\Phi^{h\sigma} \in \mathcal{M}^{2(h-1)}$ and $\vec{\Phi}^{h\sigma}_{\uparrow\vec{0}}(\sigma\infty) = \vec{0}$ hold, the function $\vec{\Xi}^{h\sigma}_{\uparrow}(t)$, given by the first of (2.12), is in fact in $\mathcal{M}^{2(h-1)}$ (by integration). But of course we do not know (yet) the initial data $\Xi^{h\sigma}(0)$.

To find expressions for $\Xi^h_{\uparrow}$ we start from the equations (2.6) with initial time at some instant T. And we use that $\mathcal{I}\Phi(t)$ is a primitive of the function $\Phi(t)$, see comment preceding (2.17), so that:

$$\vec{\Xi}^{h\sigma}_{\uparrow}(t) = \vec{\Xi}^{h\sigma}_{\uparrow}(T) + \mathcal{I}\vec{\Phi}^{h\sigma}_{\uparrow}(t) - \mathcal{I}\vec{\Phi}^{h\sigma}_{\uparrow}(T) \tag{2.21}$$

where $\sigma = \operatorname{sign} t$, and T has the same sign of t.

The function $\vec{\Xi}^{h\sigma}_{\uparrow}(T)$ tends to become quasi periodic with exponential speed as $T \to \sigma\infty$: in fact it becomes asymptotic to the $j = 0$ component, see (2.19), at $x = 0$: $\vec{\Xi}^{h\sigma}_{0\uparrow}(0, \vec{\omega} T)$, (in the sense that the difference tends to 0, bounded proportionally to $(g|T|)^{2h-1} e^{-g|T|}$). The function $\mathcal{I}\vec{\Phi}^{h\sigma}_{\uparrow}(T)$ also becomes asymptotically quasi periodic with exponential speed *and* $\vec{0}$ *average*, because $\vec{\Phi}^{h\sigma}_{\uparrow} \in \mathcal{M}_0^{2(h-1)}$ and by the definition of $\mathcal{I}$: therefore the two quasi periodic functions of T must cancel modulo a constant equal to $\langle \vec{\Xi}^{h\sigma}_{0\uparrow}(0, \cdot) \rangle \equiv \vec{\Xi}^{h\sigma}_{\uparrow\vec{0}}(\sigma\infty)$.

Hence it follows that:

$$\vec{\Xi}^{h\sigma}_{\uparrow}(t) = \vec{\Xi}^{h\sigma}_{\uparrow\vec{0}}(\sigma\infty) + \mathcal{I}\vec{\Phi}^{h\sigma}_{\uparrow}(t) \tag{2.22}$$

and, by inserting (2.22) into the second of (2.12), (considering also that $\int_0^{gt} g\tau \vec{\Phi}^{h\sigma}_{\uparrow}(\tau)\, d\, g\tau = gt\mathcal{I}\vec{\Phi}^{h\sigma}_{\uparrow}(t) +$ a t-bounded function), we see that the $\vec{\Xi}^{h\sigma}_{\downarrow}(t)$ can be bounded only if:

$$\vec{\Xi}^{h\sigma}_{\uparrow\vec{0}}(\sigma\infty) = \vec{0}, \qquad \text{hence:} \qquad \vec{\Xi}^{h\sigma}_{\uparrow}(t) = \mathcal{I}\vec{\Phi}^{h\sigma}_{\uparrow}(t) \tag{2.23}$$

yielding, setting $t = 0^\sigma$, the initial values of $\Xi^h_{\uparrow}$ *and* a simple form for its time evolution. Analogously, recalling that $\vec{\Xi}^{h\sigma}_{\downarrow}(0) = \vec{0}$, essentially by definition, one finds:

$$\vec{\Xi}^{h\sigma}_{\downarrow}(t) = \left(\mathcal{I}^2 \vec{\Phi}^{h\sigma}_{\uparrow}(t) - \mathcal{I}^2 \vec{\Phi}^{h\sigma}_{\uparrow}(0^\sigma)\right) \equiv \bar{\mathcal{I}}^2 \vec{\Phi}^{h\sigma}_{\uparrow}(t) \tag{2.24}$$

which gives a simple form to the time evolution of the $\vec\alpha$ (*i.e.* $\downarrow$) component of Ξ^h in terms of the operator $\overline{\mathcal{I}}^2$ defined by the r.h.s. of (2.24).

Likewise considering the (2.11) and the behaviour at $\sigma\infty$ of W in (2.9), if $\Xi^{h\sigma}(t)$ has to be bounded at $\sigma\infty$, we see from the second of (2.11) that:

$$\Xi^{h\sigma}_+(0) = -\mathop{\int\!\!\!\!\!\!-}\nolimits_0^{\sigma\infty} w_{00}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g\tau \tag{2.25}$$

Thus we get (defining at the same time also $\mathcal{O}$ and $\mathcal{O}_+$):

$$\begin{aligned}\Xi^{h\sigma}_+(t) &= w_{ll}(t)\mathop{\int\!\!\!\!\!\!-}\nolimits_{(\sigma)}^{gt} w_{00}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g\tau - w_{l0}(t)\int_0^{gt} w_{0l}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g\tau \equiv \mathcal{O}_+(\Phi^{h\sigma}_+)(t)\\ \Xi^{h\sigma}_-(t) &= w_{0l}(t)\mathop{\int\!\!\!\!\!\!-}\nolimits_{(\sigma)}^{gt} w_{00}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g\tau - w_{00}(t)\int_0^{gt} w_{0l}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g\tau \equiv \mathcal{O}(\Phi^{h\sigma}_+)(t)\end{aligned} \tag{2.26}$$

The (2.23), (2.24) and (2.26), and the boundedness request imply (2.19) for $h' = h+1$, as we can show by reasoning as in [G1]. As already remarked before (2.21) we note again that, since $\Phi^{h'\sigma}_{\uparrow\vec 0}(\sigma\infty) = \vec 0$ for $h' \le h$, the $\vec\Xi^h_\uparrow, \vec\Xi^h_\downarrow$ functions are in fact in $\mathcal{M}^{2(h-1)}$, as the $\mathcal{I}$ operation, on such $\vec\Phi^h_\uparrow$ functions, does not increase the degree. Also, if one looks carefully at the $\Xi^{h'\sigma}_\pm$–evaluation in terms of $\Phi^{h'\sigma}_+$, one realizes that the $\mathcal{O},\mathcal{O}_+$ operations may increase the degree but by at most 1. Thus the inductive hypothesis made in connection with (2.19) is proved for Ξ^h, and it remains to check it for Φ^{h+1}. This follows from the expression of Φ^{h+1}, see (2.7), in terms of the $\Xi^{h'}$ with $h' \le h$: see (2.7). One treats separately the sums in (2.7) with $|\vec m| \ge 2$ and $|\vec m| \ge 0$: one just has to consider that in the first case, which might look dangerous for the inductive hypothesis, the products of Ξ's contains at least two factors (which therefore have order labels smaller than h and verify the inductive hypothesis); and, furthermore, the coefficients $(g^2J_0)^{-1}(\partial_\varphi f)_{\vec m}(\varphi_0,\vec\omega t)$ or $(g^2J)^{-1}(\partial_{\vec\alpha} f)_{\vec m}(\varphi_0,\vec\omega t)$ or $\sin\varphi_0$ or $\cos\varphi_0$ do not contain any terms that can possibly increase the degree. Hence $\Phi^{h+1} \in \mathcal{M}^{2h}$. To see that $\vec\Phi^{(h+1)\sigma}_\uparrow \in \mathcal{M}^{2h}_0$, *i.e.* $\vec\Phi^{(k+1)\sigma}_{\uparrow\vec 0}(\sigma\infty) = \vec 0$, we simply remark that otherwise the second of (2.12) could not be bounded in t as $t\to\infty$.

We can summarize the above considerations as:

$$\vec\Phi^{h\sigma}_{\uparrow\vec 0}(\sigma\infty) \equiv \int_{\mathbf{T}^{l-1}} \vec\Phi^{h\sigma}_\uparrow(\vec\psi,\sigma\infty)\frac{d\vec\psi}{(2\pi)^{l-1}} \equiv \langle\vec\Phi^{h\sigma}_\uparrow(\cdot,\sigma\infty)\rangle = \vec 0 \tag{2.27}$$

for all $h \ge 1$, and, still for all $h \ge 1$, as:

$$\begin{aligned}\Xi^h_-(t) =& w_{0l}(t)\mathcal{I}(w_{00}\Phi^h_+)(t) - w_{00}(t)\big(\mathcal{I}(w_{0l}\Phi^h_+)(t) - \mathcal{I}(w_{0l}\Phi^h_+)(0^\sigma)\big) \equiv \mathcal{O}(\Phi^h_+)(t)\\ \vec\Xi^h_\downarrow(t) =& \Big(\mathcal{I}^2(\vec\Phi^h_\uparrow)(t) - \mathcal{I}^2(\vec\Phi^h_\uparrow)(0^\sigma)\Big) \equiv \overline{\mathcal{I}}^2(\vec\Phi^h_\uparrow(t))\\ \Xi^h_+(t) =& w_{ll}(t)\mathcal{I}(w_{00}\Phi^h_+)(t) - w_{l0}(t)\Big(\mathcal{I}(w_{0l}\Phi^h_+)(t) - \mathcal{I}(w_{0l}\Phi^h_+)(0^\sigma)\Big) \equiv \mathcal{O}_+(\Phi^h_+)(t)\\ \vec\Xi^h_\uparrow(t) =& \mathcal{I}(\vec\Phi^h_\uparrow)(t)\end{aligned} \tag{2.28}$$

where $\mathcal{O},\mathcal{O}_+,\overline{\mathcal{I}}^2,\mathcal{I}$ are defined here and in §2.2; and $\Xi^h \equiv (\Xi_-,\vec\Xi_\downarrow,\Xi_+,\vec\Xi_\uparrow) = (\Xi^h_j)$, $j = 0,\ldots 2l-1$, $\Phi^h = (0,\vec 0,\Phi^h_+,\vec\Phi^h_\uparrow)$. Note that, while Ξ^h has non zero components over both the *angle* ($j = 0,\ldots,l-1$) and over the *action* ($j = l,\ldots,2l-1$) componenents, the Φ^h has only the action components non zero.

We can give the above discussion a more formal statement through the following propositions:

Proposition 2.1: *The series defining the functions $\vec\psi \to X^\sigma(x,\vec\psi,t) = \sum_{h=0}^\infty \mu^h X^{h\sigma}(x,\vec\psi,t)$ are convergent for μ small enough and $|x| \le 1, \sigma t \ge 0$. And if $x = e^{-g\sigma t}$ the surfaces $(\vec\psi,t) \to X^\sigma(x,\vec\psi,t)$ are stable and unstable whiskers $W^\pm_\mu$, (respectively, if $\sigma = \pm$). The functions $\vec\psi \to X^\sigma(0,\vec\psi,\sigma\infty)$ describe invariant tori $\mathcal{T}$, on which the motion is $\vec\psi \to \vec\psi + \vec\omega t$. The two tori coincide as sets, although they may be parameterized differently (i.e. points with the same $\vec\psi$ may be different in the two parametrizations).*

Remark : The map on such torus defined by the correspondence established by having the same $\vec{\psi}$ leads to the notion of homoclinic splitting, homoclinic scattering and homoclinic phase shifts, see [CG], [G1].

Proposition 2.2: *If $(I,\vec{A},\varphi,\vec{\alpha})\in W^{\pm}_{\mu}$, i.e. if $(I,\vec{A},\varphi,\vec{\alpha})=X^{\sigma}_{\mu}$, then the evolution $S_t(I,\vec{A},\varphi,\vec{\alpha})$ converges to a quasiperiodic motion on the torus $\mathcal{T}$ of Proposition 2.1. And in fact the convergence is exponential in the sense that, for $\sigma t\ge 0$, $\big|X^{\sigma}(x,\vec{\psi}+\vec{\omega}t,t)-X^{\sigma}(0,\vec{\psi},\sigma\infty)\big|\le Ce^{-\frac{1}{2}g\sigma t}$, for some constant $C>0$, and for μ small enough.*

The above propositions are immediate consequences of the previous discussion: the only result we have not yet is the convergence of the series (2.2), but this will be obtained in §5.2.

Remark : The reason for the above bound of the exponential damping constant by $\frac{1}{2}g$ is that the true decay is $g(\mu)=g+O(\mu)$, see [CG], §5, Lemma 1. In fact, in next sections, we exploit the μ-dependence of the function $g^2(\mu)$, whose coefficients γ_p, $p\ge 1$, will be set up precisely to make the powers of the time do not arise: in other words the terms to be integrated will have the form (2.14), and the coefficients with $|h|+|\vec{\nu}|=0$ will be automatically vanishing, so that the degree of the functions will never increase, and, since, it was originally zero, it will remain that. This means that a suitable choise of the "counterterms" γ_p, $p\ge 1$, yields that the functions $\Xi^{h\sigma}$, $\Phi^{h\sigma}$ are in fact in $\mathcal{M}^0$, for any $h\ge 1$.

3. Tree formalism

In this section we introduce the graphical formalism, partially developed in [G1], §5, [Ge], §5, §7, in order to represent, via equations (2.28) and (2.7), the generic h-th order contribution to the solutions of the perturbed motion equations. *For the time being we ignore the presence of the second sum in (2.7), (i.e. we reason as it was $\gamma_p\equiv 0$, $\forall\, p\ge 1$); we shall see in §5.1 how the discussion has to be modified when also such terms are taken into account.*

We introduce a label ν to split the functions appearing in (2.7) as sums of their Fourier components; let:

$$\begin{aligned}
&f^{\delta}(\varphi,\vec{\alpha})\equiv\sum_{\nu=(n,\vec{\nu})}\frac{f^{\delta}_{\nu}}{2}\,e^{i(n\varphi+\vec{\nu}\cdot\vec{\alpha})},\qquad \delta=0,1\\
&f^{0}(\varphi,\vec{\alpha})\equiv J_0g^2\cos\varphi=\sum_{\substack{\nu,\,\vec{\nu}=\vec{0}\\ n=\pm1}}\frac{f^{0}_{\nu}}{2}\,e^{in\varphi},\qquad f^{1}(\varphi,\vec{\alpha})\equiv f(\varphi,\vec{\alpha})=\sum_{\nu}\frac{f^{1}_{\nu}}{2}\,e^{i(n\varphi+\vec{\nu}\cdot\vec{\alpha})}
\end{aligned}\tag{3.1}$$

(the above Fourier representation is convenient as it eliminates the derivatives with respect to $\varphi,\vec{\alpha}$ in the coefficients of (2.7)).

A *tree diagram* (or simply *tree*) θ will consist of a family of lines (*branches*) arranged to connect a partially ordered set of points (*nodes*), with the higher nodes to the right. The branches are naturally ordered as well; all of them have two nodes at their extremes (possibly one of them is a top node) except the lowest or *first* branch which has only one node, the first node v_0 of the tree. The other extreme r of the first branch will be called the *root* of the tree and it will not be regarded as a node; moreover we shall call *root branch* the branch connecting r to v_0. If v_1 and v_2 are two nodes we say that $v_1<v_2$, if v_2 follows v_1 in the order established by the tree: *i.e.* if on has to pass v_1 before reaching v_2, while climbing the tree. If v is a node, we denote by v' the (uniquely determined) node immediately preceding v; we set $v_0'\equiv r$. Given a tree θ with first node v_0, each node $v>v_0$ can be considered the first node of the tree constisting of the nodes following v: such a tree will be called a subtree of θ.

To each node v we attach a finite set of labels τ_v, $\nu_v\equiv(n_v,\vec{\nu}_v)$, δ_v and j_v, that we call, respectively, the *time label*, the *mode label*, the *order label* and the *action label*, and to each branch λ_v leading to v we attach a *branch label* j_{λ_v}. The labels are so defined that $\nu_v\in\mathbf{Z}^l$, $|\nu_v|\le N$, $j_v=l,\ldots,2l-1$, $\delta_v=0,1$. Each branch different from the root branch, and leading to v, carries an *angle label*, $j_{\lambda_v}\equiv j_v-l=0,\ldots,l-1$; the root branch label can be either an angle label, or else an *action label* $j_{\lambda_v}\ge l$, and in this case $j_{\lambda_v}=j_v$.

The *order* $h\equiv h_{v_0}$ of the tree θ with first node v_0 is $h=\sum_{v\ge v_0}\delta_v$, *i.e.* the sum of the order labels of the nodes. The number of branches emerging from the node v is $1+m_v$, if m_v is the number nodes immediately following the considered node v (we have to count also the branch leading to v): then $m=1+\sum_{v\ge v_0}m_v$, if m is the number of nodes in θ. Of course, as the order label $\delta_v=0,1$ and as each node v with $\delta_v=0$ *must* have $m_v\ge 2$, it is $h\le m<2h$. In order to dispose of a label counting the number of nodes of a subtree, we

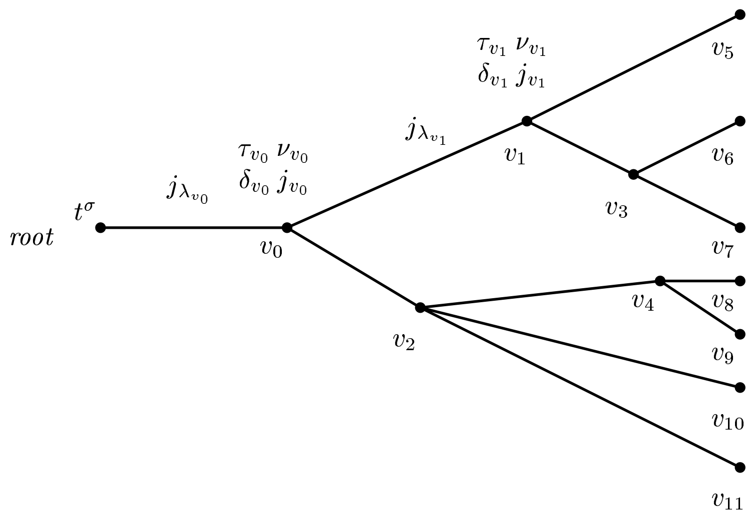

Fig.3.1: A tree θ with $m_{v_0} = 2, m_{v_1} = 2, m_{v_2} = 3, m_{v_3} = 2, m_{v_4} = 2$ and $m = 12$; the root branch label is defined to be $j_{\lambda_{v_0}} = j$.

introduce an extra label, (uniquely determined by the above ones), by defining the *degree* of a node v, d_v, as the number of nodes of the subtree having v as first node: then $d_v = 1 + \sum_{\bar v \geq v} m_{\bar v}$, $d_{v_0} = m$.

We imagine that all the branches have the same lenght (even though they are drawn with arbitrary lenght). A group acts on the sets of trees, generated by the permutations of the subtrees having the same root. Two trees that can be superposed by the action of a transformation of the group will be regarded as identical (recall however that the branches are numbered, *i.e.* are regarded as distinct, and the superposition has to be such that all the decorations of the tree match.[3] We shall imagine that each branch carries also an arrow pointing to the root ("gravity direction", opposite to the ordering).

We define the *momentum* of a node v or of the branch λ_v leading to v as $\vec\nu(v) = \sum_{w\geq v} \vec\nu_w$, if $\nu_v = (n_v, \vec\nu_v)$ is the *mode label* of v. The *total momentum* is $\vec\nu(v_0) = \sum_{v\geq v_0} \vec\nu_v$; we say also that $\vec\nu_v$ is the momentum "emitted" by the node v.

Then to each node v there corresponds a factor:

$$\frac{1}{2}(-i\nu_v)_{j_v-l}\, c_{\nu_v}\ e^{i(n_v\varphi^0(\tau_v)+(\vec\alpha+\vec\omega\tau_v)\cdot\vec\nu_v)} \prod_{s=0}^{l-1} (i\nu_{vs})^{m_v^s} \tag{3.2}$$

(the integers m_v^s decompose m_v and count the number of branches emerging from v and carrying the labels $s = 0, \ldots, l-1$, and the last product in (3.2) is missing if no nodes follow v), where:

$$c_{\nu_v} \equiv [(J_0 g^2)^{-1}\delta_{j_v,l} + (Jg^2)^{-1}\big(1-\delta_{j_v,l}\big)\delta_v] f_{\nu_v}^{\delta_v}\ ,$$

(where $\delta_{j_v,l}$ is 1 if $j_v = l$, and 0 otherwise, *i.e.* $j_v = l+1, \ldots, 2l-1$), and to each branch λ we associate an improper integration operation with upper limit t, denoted $\mathcal{O}$, $\overline{\mathcal{I}}^2$, $\mathcal{O}_+$, $\mathcal{I}$, like in (2.28), and the branch label will be $j_\lambda = 0$ when representing $\mathcal{O}$, $j_\lambda = 1, \ldots, l-1$ for $\overline{\mathcal{I}}^2$, $j_\lambda = l$ for $\mathcal{O}_+$, and $j_\lambda = l+1, \ldots, 2l-1$ for $\mathcal{I}$.

Given all the above decorations on a labeled tree θ we define its value $\tilde V_j(t;\theta)$ via the following operations:
(1) We first lay down a set of parentheses () ordered hierarchically and reproducing the tree structure (in fact any ordered (topological) tree can be represented as a set of matching parentheses representing the tree nodes). Matching parentheses corresponding to a node v will be made easy to see by appending to them a label v. The root will not be represented by a (unnecessary) parenthesis.
(2) Inside the parenthesis $(_v$ and next to it we write the factor (3.2).

[3] If we use the terminology of [G1], we can say that we are considering only *labeled trees*, (and not *topological* or *semitopological trees*).

(3) Furthermore out of $(_v$ and next to it we write a symbol $\mathcal{E}_v^T$ which we interpret differently, depending on the label j_{λ_v} on λ_v:

$$\mathcal{E}_v^T\Big(_v \cdot \Big)_v \equiv \begin{cases} \mathcal{O}\Big(_v \cdot \Big)_v(\tau_{v'}) \ , & \text{if } v \geq v_0 \ , \quad j_{\lambda_v} = 0 \ , \\ \overline{\mathcal{I}}^2\Big(_v \cdot \Big)_v(\tau_{v'}) & \text{if } v \geq v_0 \ , \quad 1 \leq j_{\lambda_v} \leq l-1 \ , \\ \mathcal{O}_+\Big(_v \cdot \Big)_v(\tau_{v'}) \ , & \text{if } v = v_0 \ , \quad j_{\lambda_v} = l; \\ \mathcal{I}\Big(_v \cdot \Big)_v(\tau_{v'}) \ , & \text{if } v = v_0 \ , \quad l+1 \leq j_{\lambda_v} \leq 2l-1 \end{cases} \tag{3.3}$$

being $\tau_{v_0'}$ the root time label t^σ of the tree and the superscript σ attached to t is important only if $t = 0$: in such case (3.3), if $v = v_0$, has to be interpreted as the limit as $t \to 0^\sigma$.

Let us denote $\Xi_{j\vec{\nu}}^{h\sigma}(t)$ the coefficient of order h in the Taylor expansion in powers of μ and of order $\vec{\nu}$ in the Fourier expansion in $\vec{\alpha}$ of the function $(\mu, \vec{\alpha}) \to \Xi_\mu^\sigma(t; \vec{\alpha})$. Then it follows that $\Xi_{j\vec{\nu}}^{h\sigma}(t)$ can be written as:[4]

$$\Xi_{j\vec{\nu}}^{h\sigma}(t) = \sum_{\theta \in trees} \frac{1}{m(\theta)!} \sum_{labels: \sum_v \delta_v = h} \tilde{V}_{j\vec{\nu}}(t;\theta) \tag{3.4}$$

where $m(\theta) =$ number of branches of $\theta =$ number of nodes of θ, and, if $j \geq l$:

$$\begin{aligned} \tilde{V}_{j\vec{\nu}}(t;\theta) = & \prod_{v_0 \leq v \in \theta} \oint \frac{dR_v}{2\pi i R_v} \sum_{\rho_v = 0,1} \int_{\sigma_{v'}\infty}^{\rho_v g \tau_{v'}} d\, g\tau_v \, e^{-\sigma_v g R_v \tau_v} \, w_{j_v}^{\rho_v}(\tau_{v'}, \tau_v) \\ & \cdot \Big[\frac{(-i\nu_v)_{j_v - l}}{2} \, c_{\nu_v} \, e^{i(n_v \varphi^0(\tau_v) + \vec{\nu}_v \cdot \vec{\omega} \tau_v)} \prod_{s=0}^{l-1} (i\nu_{vs})^{m_v^s} \Big] \end{aligned} \tag{3.5}$$

where $\tau_{v_0'} = t$, ($t = 0$ is meant as $t \to 0^\sigma$, $\sigma = \pm$, so that $\sigma = \sigma_{v_0'}$ is well defined also for $t = 0$, see (3.3)), $j_{v_0} = j$, and we have used (2.28), by setting:

$$\begin{aligned} w_{j_v}^0(\tau_{v'}, \tau_v) &= \begin{cases} w_{00}(\tau_{v'}) w_{0l}(\tau_v), & v > v_0 \ , j_v = l \\ g\tau_v, & v > v_0 \ , j_v > l \end{cases} \\ w_{j_{v_0}}^0(t, \tau_{v_0}) &= \begin{cases} w_{l0}(t) w_{0l}(\tau_{v_0}), & j_v = l \\ 0, & j_v > l \end{cases} \\ w_{j_v}^1(\tau_{v'}, \tau_v) &= \begin{cases} w_{0l}(\tau_{v'}) w_{00}(\tau_v) - w_{00}(\tau_{v'}) w_{0l}(\tau_v), & v > v_0 \ , j_v = l \\ g(\tau_{v'} - \tau_v), & v > v_0 \ , j_v > l \end{cases} \\ w_{j_{v_0}}^1(t, \tau_{v_0}) &= \begin{cases} w_{ll}(t) w_{00}(\tau_{v_0}) - w_{l0}(t) w_{0l}(\tau_{v_0}), & j_{v_0} = l \\ 1, & j_{v_0} > l \end{cases} \end{aligned} \tag{3.6}$$

If $j < l$, then (3.5) still holds, but $w_{j_{v_0}}^{\rho_{v_0}}$, in (3.6), is defined as $w_{j_v}^{\rho_v}$, $v > v_0$.

Remark 1 : If we do not perform the operation $\mathcal{E}_{v_0}^T$ relative to the time τ_{v_0} of the first node v_0 and set it to be equal to t, setting also $j \equiv j_{v_0}$, we see that the result is a representation of $\Phi_j^h(t)$.

Remark 2: Note that the tree value $\tilde{V}_j(t;\theta)$, defined through (3.5), is the same one introduced in [G1]. The analysis of [G1] applies, and allows us, fixed the perturbative order h, to give a bound on $\Xi_{j\vec{\nu}}^{h\sigma}(t)$, via (3.4). Nevertheless, in order to obtain our stronger bound C^h, for some constant $C > 0$, some improvement is needed. The special form of the kernels in (3.6) has to be exploited, and the terms from which convergence problems arise have to be singled out: then we perform essentially an exact calculation on such terms, so that the involved cancellation mechanisms can be implemented, while the other ones, harmless with respect to the estimates, can be easily bounded. In order to achieve such a goal and distinguish between "dangerous" and "harmless" contributions, some new labels will be introduced: only when such labels will assume some particular values, a more careful analysis is required.

[4] The only non trivial property of the representation (3.4) is the combinatorics; however it is not difficult to check it, say inductively.

We can split $w^{\rho_v}_{j_v}(\tau_{v'},\tau_v)$, $v > v_0$, as follows: if $j_v > l$ we do nothing, otherwise we decompose it as sum of two (if $\rho_v = 0$) or three (if $\rho_v = 1$) terms:

$$
\begin{aligned}
w^0_{j_v}(\tau_{v'},\tau_v) =& \frac{1}{2}\left\{\frac{g\tau_v}{\cosh g\tau_{v'}\ \cosh g\tau_v} + \frac{\sinh g\tau_v}{\cosh g\tau_{v'}}\right\} \\
w^1_{j_v}(\tau_{v'},\tau_v) =& \frac{1}{2}\left\{\frac{g(\tau_{v'}-\tau_v)}{\cosh g\tau_{v'}\ \cosh g\tau_v} + \frac{\sinh g\tau_{v'}}{\cosh g\tau_v} - \frac{\sinh g\tau_v}{\cosh g\tau_{v'}}\right\}
\end{aligned}
\tag{3.7}
$$

Then we can write:

$$
\begin{aligned}
w^0_{j_v}(\tau_{v'},\tau_v)\, e^{in_v\varphi^0(\tau_v)} &= \begin{cases} g\tau_v\, y^{(0)}_v(\tau_{v'},\tau_v) + y^{(-1)}_v(\tau_{v'},\tau_v)\,, & \text{if } j_v = l \\ g\tau_v\, y^{(2)}_v(\tau_v)\,, & \text{if } j_v > l \end{cases} \\
w^1_{j_v}(\tau_{v'},\tau_v)\, e^{in_v\varphi^0(\tau_v)} &= \begin{cases} g(\tau_{v'}-\tau_v)\, y^{(0)}_v(\tau_{v'},\tau_v) + y^{(1)}_v(\tau_{v'},\tau_v) - y^{(-1)}_v(\tau_{v'},\tau_v)\,, & \text{if } j_v = l \\ g(\tau_{v'}-\tau_v)\, y^{(2)}_v(\tau_v)\,, & \text{if } j_v > l \end{cases}
\end{aligned}
\tag{3.8}
$$

where the functions $y^{(\alpha)}_v$, $\alpha = -1,0,1,2$, are elements of a finite set of functions:

$$
\begin{aligned}
y^{(-1)}_v(\tau_{v'},\tau_v) &= \frac{1}{2}\frac{\sinh g\tau_v}{\cosh g\tau_{v'}}\ e^{in_v\varphi^0(\tau_v)} & y^{(1)}_v(\tau_{v'},\tau_v) &= \frac{1}{2}\frac{\sinh g\tau_{v'}}{\cosh g\tau_v}\ e^{in_v\varphi^0(\tau_v)} \\
y^{(0)}_v(\tau_{v'},\tau_v) &= \frac{1}{2}\frac{1}{\cosh g\tau_v \cosh g\tau_{v'}}\ e^{in_v\varphi^0(\tau_v)} & y^{(2)}_v(\tau_{v'},\tau_v) &= e^{in_v\varphi^0(\tau_v)}
\end{aligned}
\tag{3.9}
$$

and admit the following Laurent series expansion:

$$
\begin{aligned}
y^{(-1)}_v(\tau_{v'},\tau_v) &= \sum_{k'_v=1}^{\infty}\sum_{k_v=-1}^{\infty} y^{(-1)}_v(k'_v,k_v) x_{v'}^{k'_v} x_v^{k_v} & y^{(1)}_v(\tau_{v'},\tau_v) &= \sum_{k'_v=-1}^{\infty}\sum_{k_v=1}^{\infty} y^{(1)}_v(k'_v,k_v) x_{v'}^{k'_v} x_v^{k_v} \\
y^{(0)}_v(\tau_{v'},\tau_v) &= \sum_{k'_v=1}^{\infty}\sum_{k_v=1}^{\infty} y^{(0)}_v(k'_v,k_v) x_{v'}^{k'_v} x_v^{k_v} & y^{(2)}_v(\tau_v) &= \sum_{k_v=0}^{\infty} y^{(2)}_v(0,k_v) x_v^{k_v}
\end{aligned}
\tag{3.10}
$$

with $x_v = \exp[-\sigma_v g\tau_v]$, $\sigma_v = \operatorname{sign}\tau_v$, and $x_{v'} = \exp[-\sigma_{v'} g\tau_{v'}]$, $\sigma_{v'} = \operatorname{sign}\tau_{v'}$. We use the fact that $[\cosh g\tau]^{-1} = 2x/(1+x^2)$, $\sinh g\tau = \sigma(1-x^2)/(2x)$, $\cos\varphi^0(\tau) = 1 - 8x^2/(1+x^2)^2$, and $\sin\varphi^0(\tau) = 4\sigma x(1-x^2)/(1+x^2)^2$, if $x = \exp[-\sigma g\tau]$. We can compute some coefficients of the above expansions, which will turn out to be useful in the following: $y^{(-1)}_v(1,-1) = \sigma_v/2$, $y^{(-1)}_v(1,0) = 2in_v$, $y^{(-1)}_v(1,1) = -\sigma_v/2$, $y^{(0)}_v(1,1) = 2$, $y^{(0)}_v(1,2) = 8in_v\sigma_v$, $y^{(1)}_v(-1,1) = \sigma_{v'}/2$, $y^{(1)}_v(0,1) = 0$, $y^{(1)}_v(1,1) = -\sigma_{v'}/2$, $y^{(2)}_v(0,0) = 1$, $y^{(2)}_v(0,1) = 4in_v\sigma_v$. We define the sets Λ_α, $\alpha = -1,0,1,2$, as: $\Lambda_\alpha = \{v \in \theta : \alpha_v = \alpha\}$.

If $j < l$, a decomposition analogous to (3.7) is still possible also for the term corresponding to the first node v_0:

$$
\begin{aligned}
w^0_{j_{v_0}}(t,\tau_{v_0}) =& -\frac{1}{2}\left\{\frac{g\tau_{v_0}\sinh g\tau_{v_0}}{\cosh^2 gt\ \cosh g\tau_{v_0}} + \frac{\sinh gt \sinh g\tau_{v_0}}{\cosh^2 gt}\right\} \\
w^1_{j_{v_0}}(t,\tau_{v_0}) =& -\frac{1}{2}\left\{\frac{g(t-\tau_{v_0})\sinh gt}{\cosh^2 gt\ \cosh g\tau_{v_0}} + \frac{1+\cosh^2 gt}{\cosh gt \cosh g\tau_{v_0}} - \frac{\sinh gt \sinh g\tau_{v_0}}{\cosh^2 gt}\right\}
\end{aligned}
\tag{3.11}
$$

and expressions similar to (3.8) can be easily found, being the functions $y^{(\alpha_v)}_v$ replaced with new functions, say $\tilde{y}^{(\alpha_v)}_v$, admitting a series expansion differing from (3.10) only in the values of the coefficients $y^{(\alpha_v)}_v(k'_{v_0},k_{v_0})$; we do not give the details, as the analysis proceeds along the same lines, and no relevant difference is introduced.[5]

Then, for each tree node, we have four more labels, $k_v, k'_v, \rho_v, \alpha_v$, to add to the previous ones $\tau_v, \nu_v, \delta_v, j_v$, and, in the end, we have to sum over all the possible consistent collections of such labels, (note that the just introduced labels are not quite independent on each other: *e.g.* $\alpha_v = 1$ is possible only if $\rho_v = 1$, and if an action label is $j_v > l$, then necessarily it is $\alpha_v = 2$). Therefore the tree value $\tilde{V}_{j\vec{\nu}}(t;\theta)$ introduced in (3.3) can be replaced with a new tree value, $V_{j\vec{\nu}}(t;\theta)$, taking into account also the new labels, and (3.4) holds still

[5] We only stress, (as it will turn out to be useful in the discussion of the renormalization procedure of §5), that, as far the functions $y^{(-1)}_v$ and $\tilde{y}^{(-1)}_v$ are concerned, the dependence on the τ_{v_0} variable is through the same factor $\sinh g\tau_{v_0}\exp[in_v\varphi^0(\tau_{v_0})]$.

provided $\tilde{V}_{j\vec{\nu}}(t;\theta)$ is replaced with $V_{j\vec{\nu}}(t;\theta)$. The generic contribution $(1/m!)\ V_{j\vec{\nu}}(t;\theta)$ to $\Xi^{h\sigma}_{j\vec{\nu}}$, corresponding to a given tree θ, with $m(\theta)=m$, is:

$$\frac{1}{m!}\ V_{j\vec{\nu}}(t;\theta)=\frac{1}{m!}\prod_{v_0\le v\in\theta}\oint\frac{dR_v}{2\pi iR_v}\int_{\sigma_{v'}\infty}^{\rho_v g\tau_{v'}}d\,g\tau_v\ \mathcal{V}_v(\theta) \tag{3.12}$$

where we have defined the *node function* $\mathcal{V}_v(\theta)$, (depending on the tree to which the node v belongs), as:

$$\mathcal{V}_v(\theta)\equiv F_{\nu_v}\ T_v(g\tau_{v'},g\tau_v)\ e^{-\sigma_v R_v g\tau_v}\ e^{i\omega_v\tau_v}\ x_v^{k_v}\Big[\prod_{j=1}^{m_v}x_v^{k'_{v_j}}\Big]\,, \tag{3.13}$$

where:
1) $\omega_v=\vec{\omega}\cdot\vec{\nu}_v$;
2) m_v is the number of branches emerging from v, and $v_1,\ldots,v_{m_v}$ are the nodes immediately following v moving along the tree (so that the product in square brackets in (3.13) is missing if v is a top node);
3) $T_v(g\tau_{v'},g\tau_v)$ is defined as:

$$T_v(g\tau_{v'},g\tau_v)=(\delta_{\alpha_v,2}+\delta_{\alpha_v,0})\,[(1-\rho_v)g\tau_v+\rho_v g(\tau_{v'}-\tau_v)]+(\delta_{\alpha_v,-1}+\delta_{\alpha_v,1}) \tag{3.14}$$

(note that $T_v(g\tau_{v'},g\tau_v)\equiv T_v(g\tau_v)$, if $\rho_v=0$, and $T_v(g\tau_{v'},g\tau_v)\equiv T_v(g\tau_{v'}-g\tau_v)$, if $\rho_v=1$);
4) F_{ν_v} is given by:

$$F_{\nu_v}=\frac{(-i\nu_v)_{j_v-l}}{2}\ c_{\nu_v}\ \Big[\prod_{s=0}^{l-1}(i\nu_{vs})^{m_v^s}\Big]\,(-1)^{\delta_{\alpha_v,-1}\delta_{\rho_v,1}}\ y_v^{(\alpha_v)}(k'_v,k_v)\equiv\bar{F}_{\nu_v}\ (-1)^{\delta_{\alpha_v,-1}\delta_{\rho_v,1}}\ y_v^{(\alpha_v)}(k'_v,k_v) \tag{3.15}$$

where the coefficients $\bar{F}_{\nu_v}$ satisfy the following bound:

$$\Big|\prod_{v\ge v_0}\bar{F}_{\nu_v}\Big|\le\Big(\frac{N}{2}F_0N\Big)^m\equiv\mathcal{C}^m \tag{3.16}$$

with $F_0=(J_0g^2)^{-1}\max_\nu\{f_\nu\}$, and the coefficients $y_v^{(\alpha_v)}(k'_v,k_v)$ satisfy the bound:

$$\left|\prod_{v\ge v_0}y_v^{(\alpha_v)}(k'_v,k_v)\right|\le M^{2m}\prod_{v\ge v_0}\lambda^{k_v+k'_v} \tag{3.17}$$

if the arguments of the $y_v^{(a)}$'s are all inside an annulus $0<|x|\le\lambda<1$, so that the Laurent series defining the functions appearing in the $y_v^{(v)}$'s converge: therefore, to order $k\ge 0$, the coefficients can be bounded by a common value M_1 on the maxima of such functions (there are a finite number of them) in a disk of radius $\lambda<1$ times λ^{-k}, and, for $k=-1$, their absolute values are known to be 1, so that we can set $M=\max\{M_1,\lambda\}$.[6]

For each v, once we have integrated over the τ_v variable, we have still to evaluate the residue of the resulting expression at $R_v=0$, so that, if we consider together the two operations of integration over the time and of evaluation of the residue, we can imagine to handle a sequence of hierarchically ordered integrals. This means that we first integrate with respect either to the $(\tau_v-\tau_{v'})$'s, (if $\rho_v=1$), or to the τ_v's, (if $\rho_v=0$), the v's being the top nodes, in an arbitrary order, then we evaluate the corresponding residues, an so on until we reach the tree root.

Definition 3.1 : *Given a tree θ, let us define the* reduced tree $\bar{\theta}$ *in the following way. Let us draw a bubble B_v encircling each node $v>v_0$ with $\rho_v=0$ and the entire subtree emerging from it, and let us delete all the so obtained bubbles, but the outer ones; each remaining bubble encloses a subtree with first node v and ρ_v label fixed to be zero. Then, inside each bubble B_v, we consider all the possible trees with the same labels*

[6] The request that *all* the x satisfy the property $|x|<\lambda$ will turn out to be not very strong: in the cases in which it will be used, the time variables will be ordered so that, if $|x_{v_0}|\le\lambda$, then $|x_v|\le\lambda$ for all $v>v_0$ (see §5.2 below), and this will suffice.

attached to the node v, (in particular with the same h_v and $\rho_v = 0$), and we sum their values (3.12):[7] *the so obtained quantity $L^{h_v \sigma_{v'}}_{j_v \vec{\nu}(v)}(\tau_{v'})$ will be associated with a fat point, replacing the original bubble, which will be called a* leaf *(of the reduced tree). We call* free nodes *the reduced tree nodes different from the leaves; the leaves will be considered a particular type of top nodes, but they will be distinguished from the free nodes. We define the* reduced degree *and the* reduced order *of a reduced tree, respectively, as the number of free nodes and as the sum of their order labels, and the* order of a leaf *as the label h_v associated with the fat point representing it. We can associate to a reduced tree $\bar\theta$ a value $V_{j\vec{\nu}}(t;\bar\theta)$, where, corresponding to each free node v, there is a factor $\mathcal{V}_v(\bar\theta) \equiv \mathcal{V}_v(\theta)$ as in (3.12), and, corresponding to each leaf v, there is factor $L^{h_v \sigma_{v'}}_{j_v \vec{\nu}(v)}(\tau_{v'})$.*

Remark 1 : The reduced degree is so defined that the degree of a reduced tree turns out to be equal to the reduced degree increased by the sum of the degrees of its leaves, as it is natural to set. The analogous property holds for the reduced order.
Remark 2 : Note that an integration time variable is associated only to the free nodes. This could be found a little misleading with respect to the notion of node in the usual terminology, (see [G1], [G2], [GGe]); nevertheless we use the name node also for the leaves for convenience, since we want to affix to the leaves too the labels $k_v = 0$ and k'_v, (see, in particular, the first paragraph in §4 below).
Remark 3 : With respect to [Ge], we do not introduce the notion of generalized reduced tree; this allows us to lighten the notations, (and to avoid using trees which are not easy to visualize), but requires a refinement of the proof. However we shall see in the following that not too much work must be done in order to achieve such a goal.

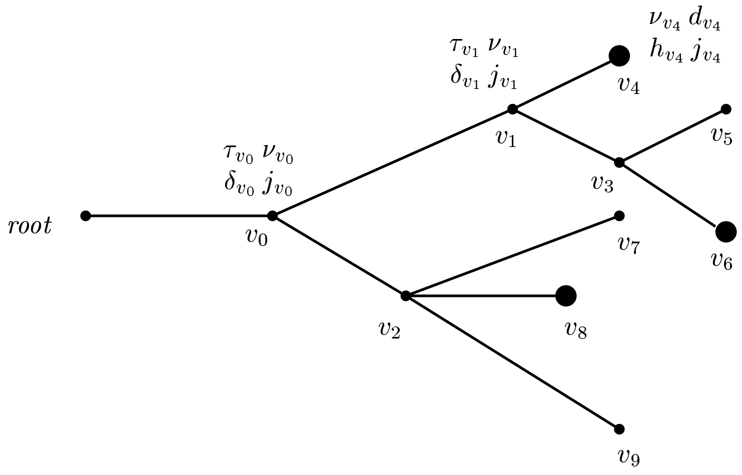

Fig.3.2: A reduced tree $\bar\theta$ with $\mathcal{N}_L = 3$ leaves, $m_{v_0} = 2, m_{v_1} = 2, m_{v_2} = 3, m_{v_3} = 2$, and reduced degree $d_{v_0} = 7$; the branch label is defined to be $j_\lambda = j$. Each fat point represents a leaf.

By construction all the free nodes have $\rho_v = 1$, except the first node v_0 which can have $\rho_{v_0} = 0, 1$, while the leaves have, by definition, $\rho_v = 0$. Given a reduced tree $\bar\theta$, we define $\bar\theta_f \equiv \{v \in \bar\theta : v \text{ is a free node }\}$ and $\bar\theta_L \equiv \{v \in \bar\theta : v \text{ is a leaf}\}$; then $\bar\theta = \bar\theta_f \cup \bar\theta_L$ and $\bar\theta_f \cap \bar\theta_L = \emptyset$. Note that, since $\rho_v = 1$, $\forall$ free node $v > v_0$, the time variables of a reduced tree are ordered: if $\sigma_{v_0} = \sigma$, then $\sigma_v = \sigma$, $\forall\ v > v_0$, $v \in \bar\theta_f$, and $\sigma_v \tau_v > \sigma_{v'} \tau_{v'}$ for any pair of nodes v, v', with v' immediately preceding v.

A leaf v represents a factor contributing to $\Xi^{h_v \sigma_{v'}}_{j_{\lambda_v} \vec{\nu}(v)}(\tau_{v'})$, $j_{\lambda_v} = j_v - l$, (see note 7), whose dependence on $\tau_{v'}$ reveals itself only through the factor, (see the third line in (3.7)):

$$\xi_v(\tau_{v'}) = [w_{00}(\tau_{v'})\delta_{j_v,l} + (1 - \delta_{j_v,l})] \qquad (3.18)\ ,$$

so that we can write $L^{h_v \sigma_{v'}}_{j_v \vec{\nu}(v)}(\tau_{v'}) = \xi_v(\tau_{v'})\ L^{h_v \sigma_{v'}}_{j_v \vec{\nu}(v)}(0)$, being $L^{h_v \sigma_{v'}}_{j_v \vec{\nu}(v)}(0)$ interpretated as the limit as $\tau_{v'} \to 0^{\sigma_{v'}}$. We define $L^{h_v \sigma_{v'}}_{j_v \vec{\nu}(v)}(0)$ as the *value of the leaf* v of the reduced tree. Also the factor (3.18) admits a series

[7] A leaf value $L^{h_v \sigma_{v'}}_{j_v \vec{\nu}(v)}(\tau_{v'})$ contributes to $\Xi^{h_v \sigma_{v'}}_{j_{\lambda_v} \vec{\nu}(v)}(\tau_{v'})$, $j_{\lambda_v} = j_v - l$, where $\vec{\nu}(v)$ the momentum of the node v, and $\sigma_{v'}$ is the sign of the time variable corresponding to the node v'.

expansion like the functions $y_v^{(\alpha_v)}$'s in (3.10):

$$\xi_v(\tau_{v'}) = \sum_{k'_v=1}^{\infty} \xi_v(k'_v, 0) x_{v'}^{k'_v} \tag{3.19}$$

We can use explicitly the order of the integration variables, and define:

$$\omega(v) = \sum_{\bar\theta_f \ni w \geq v} \omega_w \ , \qquad k(v) = \sum_{\bar\theta_f \ni w \geq v} k_w \ , \qquad k'(v) = \sum_{\bar\theta \ni w > v} k'_w \ , \qquad p(v) = k(v) + k'(v)$$

so that we can write:

$$\begin{aligned}
\prod_{\bar\theta_f \ni v \geq v_0} e^{-k_v g \sigma \tau_v} &= e^{-k(v_0) g \sigma \tau_{v_0}} \cdot \prod_{\bar\theta_f \ni v > v_0} e^{-k(v) g \sigma (\tau_v - \tau_{v'})} \\
\prod_{\bar\theta \ni v \geq v_0} e^{-k'_v g \sigma \tau_v} &= e^{-[k'(v_0) + k'_{v_0}] g \sigma \tau_{v_0}} \cdot \prod_{\bar\theta \ni v > v_0} e^{-k'(v) g \sigma (\tau_v - \tau_{v'})} \\
\prod_{\bar\theta_f \ni v \geq v_0} e^{-R_v g \sigma \tau_v} &= e^{-\sum_{w \geq v_0} R_w g \sigma \tau_{v_0}} \cdot \prod_{\bar\theta_f \ni v > v_0} e^{-\sum_{w \geq v} R_w g \sigma (\tau_v - \tau_{v'})} \\
\prod_{\bar\theta_f \ni v \geq v_0} e^{i\omega_v \tau_v} &= e^{i\omega(v_0) \tau_{v_0}} \cdot \prod_{\bar\theta_f \ni v > v_0} e^{i\omega(v)(\tau_v - \tau_{v'})}
\end{aligned} \tag{3.20}$$

since $\sigma_v = \sigma_{v'_0} \equiv \sigma$, $\forall\ v \geq v_0$, $v \in \bar\theta_f$. We have used the fact that each leaf v contributes to the reduced tree a value $L^{h_v \sigma_{v'}}_{j_v \vec\nu(v)}(0)$, which is independent on $\tau_{v'}$, times a factor (3.18), which one has to take into account in the computation of $p(\tilde v)$, for each $\tilde v < v$. The leaves with $j_v = l$ are such that, in (3.20), $k'_v \geq 1$, see (3.19), (3.10), while, if $j_v > l$, it is $k'_v \equiv 0$; in both cases we can define k_v to be identically vanishing, so attaching such a label, for convenience, also to the leaves. Note that only the free nodes contribute to $k(v)$ and $\omega(v)$; if we define the *free momentum* of the reduced tree with first node v_0 as $\vec\nu(v_0) = \sum_{\bar\theta_f \ni w \geq v_0} \vec\nu_w$, then we can write $\omega(v) = \vec\omega \cdot \vec\nu_0(v)$. Note also that, if $(1/m!) V_j(t; \bar\theta)$ is a contribution to $\Xi^{h\sigma_{v'_0}}_{j\vec\nu}(t)$, $\vec\nu \equiv \vec\nu(v_0)$, then in general the free momentum $\vec\nu_0(v_0)$ is different from the "total momentum" $\vec\nu$, since $\vec\nu_0(v_0)$ takes into account only the free nodes of $\bar\theta$, while $\vec\nu$ depends also on the momentum labels affixed to the leaves.

Definition 3.2 : *Given a reduced tree $\bar\theta$, we define the* stripped value *of the reduced tree $V^S_{j\vec\nu}(t; \bar\theta)$ as the value we obtain by associating to each free node a factor $\mathcal{V}_v(\bar\theta) \equiv \mathcal{V}_v(\theta)$ as in (3.12), but retaining for each leaf only the factor $\xi_v(\tau_{v'})$ in (3.18). Note that the discarded contribution of the leaf v is nothing else but its value, $L^{h_v \sigma_{v'}}_{j_v \vec\nu(v)}(0)$, as it is defined after (3.18).*

Remark : Note that the contribution of a leaf $v \in \bar\theta$ to a stripped value $V^S_{j\vec\nu}(t; \bar\theta)$ does not depend on its order h_v, but only on the label $j_{\lambda_v} = j_v - l$ of the branch leading to it, (see (3.18)).

4. Some topological properties of the reduced trees

As we shall see in §5, if it was $p(v) \neq 0\ \forall\ v \geq v_0$, no convergence problem would arise. However, obviously, the case $p(v) = 0$ is possible and cannot be ruled out: to deal with it, we need a very accurate analysis of the integrals appearing in (3.12).

Basically the reason why the bounds can be improved is the following. It is true that the case $p(v) = 0$ is critical, but, when such value of $p(v)$ occurs for some $v \in \theta$, then the values of the k_w and k'_w labels corresponding to the nodes $w \geq v$ cannot be arbitrary: on the contrary they have to be arranged in a very special way. And the fact that the cancellation mechanisms described in Appendix A1 work is strictly connected to the possible special configurations: thereby in this section we study the arrangement of the k_w and k'_w labels, $w \geq v$, when $p(v) = 0$.

Let us consider a reduced tree $\bar\theta$, with first node v_0 and $j \equiv j_{\lambda_{v_0}} < l$. For each $v \geq v_0$, $v \in \bar\theta$, it is $p(v) = k_v + \sum_{\bar\theta \ni w > v} (k_w + k'_w)$, see (3.20), where $k_w + k'_w \geq 0$, for each w, see (3.10), and $k_w \equiv 0$ if w is a leaf, see (3.19). Therefore $p(v_0)$ can vanish only if either $k_{v_0}=0$ and $k_w=-k'_w$ for each $w>v_0$, or $k_{v_0}=-1$ and $k_w=-k'_w$ for each $w>v'$, except one single node $\tilde{w}$ such that $k_{\tilde{w}}+k'_{\tilde{w}}=1$.

Let us discuss first the case $k_{v_0}=0$. Then, if $w\in\Lambda_{-1}$, the above property requires $k'_w=-k_w=1$, because $k_w\geq -1$ and $k'_w\geq 1$; if $w\in\Lambda_1$, then $k'_w=-k_w=-1$, because $k_w\geq 1$ and $k'_w\geq -1$; otherwise, if $w\in\Lambda_2$, it must be $k_w=k'_w=0$; the possibility $w\in\Lambda_0$ has to be excluded as it would imply $k_w+k'_w>0$, and, for the same reason, if w is a leaf, it must be $j_w>l$, so that $k'_w=0$. Then we can write:

$$\sum_{\theta} V_{j\vec{\nu}}(t;\theta)=\sum_{\bar{\theta}} V^S_{j\vec{\nu}}(t;\bar{\theta})\prod_{i=1}^{\mathcal{N}_L(\bar{\theta})} L^{h_{v_i}\sigma_{v'_i}}_{j_i\vec{\nu}(v_i)}(0) \tag{4.1}$$

where $\mathcal{N}_L(\bar{\theta})$ is the number of leaves of the reduced tree $\bar{\theta}$, and $j_i\equiv j_{\lambda_{v_i}}$, where v_i is the i-th leaf. Note that (4.1) is the product of factorizing terms, which can be treated separately, being independent on each other; each $L^{h_{v_i}\sigma_{v'_i}}_{j_i\vec{\nu}(v_i)}(0)$, $i=1,\ldots,\mathcal{N}_L(\bar{\theta})$, corresponds to a leaf and has as first node a node v_i with $\rho_{v_i}=0$, while $V^S_{j\vec{\nu}}(t;\bar{\theta})$ can have either ρ_{v_0} or $\rho_{v_0}=1$. Moreover each $L^{h_{v_i}\sigma_{v'_i}}_{j_i\vec{\nu}(v_i)}(0)$ can have $p(v_i)=0$ only if $k_{v_i}=0$ too; otherwise it is $k_{v_i}=\pm 1$, and, correspondingly, $p(v_i)=\pm 1$. Then we confine ourselves to the study of $V^S_{j\vec{\nu}}(t;\bar{\theta})$, being the other terms either of the same form, (and so admitting the same bound), or of a different type, since $p(v_i)\neq 0$, (and so requiring a different discussion, which we delay). If $j_w=l$, $w\geq v_0$, we consider together the cases $w\in\Lambda_{-1}$ and $w\in\Lambda_1$: they give a contribution to (3.12), containing, as far as the w node is concerned, a factor $\bar{F}_{\nu_w}\exp[i\omega(w)(\tau_w-\tau'_w)]$ times $e^{-g\sigma(\tau_w-\tau_{w'})}y^{(1)}_w(-1,1)-e^{g\sigma(\tau_w-\tau_{w'})}y^{(-1)}_w(1,-1)=(\sigma/2)[e^{-g\sigma(\tau_w-\tau_{w'})}-e^{g\sigma(\tau_w-\tau_{w'})}]$.

Let us consider now the case $p(v_0)=0$, $k_{v_0}=-1$, (we note that such a case arise only if $j_{v_0}=l$). Let us call $\tilde{w}$ the node such that $k_{\tilde{w}}+k'_{\tilde{w}}=1$, (it is $k_w=-k'_w$ for each $w>v_0$, $w\neq\tilde{w}$), and let us denote $\mathcal{P}$ the path leading from v_0 to $\tilde{w}$, and $z_i, i=1,\ldots,m_{\mathcal{P}}$ (with $z_1=v_0$, and $z_{m_{\mathcal{P}}}=\tilde{w}$) the nodes crossed by $\mathcal{P}$.

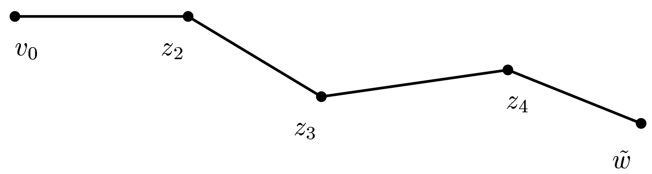

Fig.4.1: A path $\mathcal{P}$ connecting the first node v_0 of the reduced tree $\bar{\theta}$, with the node $\tilde{w}$, (defined as the node verifying the condition $k_{\tilde{w}}+k'_{\tilde{w}}=1$), with $m_{\mathcal{P}}=5$, $z_1=v_0$ and $z_5=\tilde{w}$.

For each $z_i\in\mathcal{P}$, it is $p(z_i)=k_{z_i}+1$, so that the possible values are $p(z_i)=0,1,2$, corresponding, respectively, to the cases: $z_i\in\Lambda_{-1}$, $z_i\in\Lambda_2$, $z_i\in\Lambda_1$.

Note that nodes w with $p(w)=0$ and $\alpha_w=-1$ can occur only along the path $\mathcal{P}$, as can be seen by *reductio ad absurdum*: in fact, if such a w is not in $\mathcal{P}$, it contributes $k'_w\geq 1$ to each $p(\tilde{v})$, $\tilde{v}<w$, so that, in particular, it produces a value $p(v_0)\geq 1$, which is not possible under our assumption.

As a consequence of what has been said, we see that, in order to obtain the contribution to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$, with $p(v_0)=0$, $k_{v_0}=-1$, we have to consider the sum of products of several factorizing terms, as in (4.1), which are of the same type of before, up to the first factor, which is given by the stripped value of a reduced tree with a fixed shape, and labels $p(v_0)=0$, $k_{v_0}=-1$. Therefore, with respect to the previous situation, only this term is new.

Let us consider a reduced tree, with first node v_0 having $p(v_0)\neq 0$, with given shape and collection of indices, and let us consider the $p(v)$ labels, $v>v_0$. Let us single out the nodes v's, with $p(v)=0$: then each such node will be enclosed, together with all the reduced subtree emerging from it, inside a bubble β_v. Each branch leading to a so characterized node v will be called the *stem* of the corresponding bubble. Let us delete all the bubbles, but the outer ones, after summing the values of all the possible reduced subtrees of fixed order h_v and fixed $p(v),k_v$ labels attached to the first node v represented by the end point of the bubble stem.

We can call *flowers* the bubbles; unlike the leaves, the flowers will not be considered nodes. A reduced tree with first node v_0 having $p(v_0)\neq 0$ is decorated with flowers and leaves, and, by construction, all its free nodes, (*i.e.* the nodes which are not leaves), have $p(v)\neq 0$. Each flower β_v will be characterized by the labels j_v, h_v, (h_v will be the *order of the flower*), and by a *flower function*, which is given by (4.1), where the sum is over the reduced trees having the first node v with $p(v) = 0$, and $k_v = 0, -1$.

The degree of a reduced tree is given by the number of its free nodes plus the sum of the degrees of its flowers, and of its leaves; analogously, the order of a reduced tree is given by the sum of the order labels of its nodes, (*i.e.* free nodes and leaves), plus the sum of the orders of its flowers.

Let us consider now the case $j \geq l$. The only change we have with respect to the previous situation is that the function $w_{j_{v_0}}^{\rho_{v_0}}(t, \tau_{v_0})$ which has to be associated with the first node is not equal to those of the other nodes. Nevertheless, as we have said after (3.11), a decomposition like the one in (3.8) can be obtained, with the only difference that the functions $y_v^{(\alpha_v)}$ are repalced with some new functions, *i.e.* $\tilde{y}_v^{(\alpha_v)}$, which admit Laurent series expansions as (3.10) and, therefore, can be treated in the same way. In other words, no further difficulty is introduced.

5. Analyticity of the whiskers

If all the nodes v had $p(v) \neq 0$, then all the integrals would trivially factorize, and give an explicitly computable result bounded by C^m, for some constant C. Yet it can happen that $p(v) = 0$, for some v: then, if $\omega(v) = 0$, the integration would increase by one the power of the time variable, and, moving backwards until the first node is reached, in the end we could meet dangerously high powers of the time, say $\tau_{v_0}^p$, $p \leq 2m$, so that the last integration would give a $p!$-contribution. Also the case $\omega(v) \neq 0$ would give problems, since the result of the integration on the corresponding time variable would be of the form $1/[i\omega(v)]^{-n_v}$, for some integer $n_v \geq 1$, if n_v is the power of τ_v arising as a consequence of the mechanism previously described. In fact both cases can be handled: the first one by setting the value of the "counterterms" γ_p, $p \geq 1$, so that the case $p(v) = \omega(v) = 0$ can never occur, (and, therefore, $n_v \leq 2$ because of the form of the function (3.14)), the latter by exploiting some cancellation mechanisms related to the particular structure of the kernels (3.7), which are partially taken from [G1] and [Ge], and partially introduced in this work, (always by following the same strategy of the quoted references).

The idea is the following. Let us consider only trees without leaves, for the time being. We have seen that the only terms we have to handle carefully are those with label $p(v) = 0$; because of the structure of the kernels (3.7), $p(v)$ can never be "too negative", and, in fact, it is always $p(v) \geq -1$; moreover $p(v)$ can be vanishing only if all the $p(w)$ labels of the following w nodes are equal either to 0 or to ± 1, (according to the rules stated in §4). If $p(v) = 0$, as we have seen, k_v can only assume the values either $k_v = 0$ or $k_v = -1$: then the integrals over the τ_w's, $w \geq v$, can be bounded by using the theory of the twistless KAM tori and the Eliasson's cancellations, once the values of the "counterterms" γ_p, $p \geq 1$ have been suitably fixed. It remains to study the cases $p(v) \neq 0$, but they are quite easily dealt with, by explicit calculations, if we use the results for $p(v) = 0$. As far as the leaf values are concerned, it is enough to note that a leaf v can be viewed as a contribution to $\Xi_{j_{\lambda_v} \vec{\nu}(v)}^{h_v \sigma_{v'}}(0)$, so that it can be studied in the same way as the other terms, and, therefore, admits the same bound.

5.1 Renormalization

In this section we confine ourselves to the first problem, *i.e.* the "elimination" of the powers of the time variables. This will lead to a slightly modified definition of the trees, (and therefore of their values): to be more precise, a restriction on the the compatibility of the node labels will be introduced: no node v with label $j_v = l$, $\alpha_v = -1$, $\rho_v = 1$, $p(v) = 0$, $\vec{\nu}_0(v) = \vec{0}$ will be allowed.[8] In §5.2, we shall show that the values of the so modified trees, to degree m, admit the bound C^m, so that the convergence of the formal series (2.2) follows.

Note also that, as it appears clear from the discussion in §4, if $j_v = l$, then the value $p(v) = 0$ is possible only if $\alpha_v = -1$, *i.e.* if the contribution arising from $y_v^{(-1)}(\tau_{v'}, \tau_v)$, (or $\tilde{y}_v^{(-1)}(\tau_{v'}, \tau_v)$, see the end comments in §4), is studied. The dependence of both such functions on the time variable τ_v is through the factor $\sinh g\tau_v \exp[in_v\varphi^0(\tau_v)]$ (see note 5 after (3.11)); this will imply that both cases can be handled at once. Therefore it will be not restrictive to impose the renormalization condition only to the first function: the

---

[8] Note that when $\rho_v = 0$, even if $p(v) = 0$ and $\vec{\nu}(v) = \vec{0}$, no problem arises, as, in such a case, the integral is automatically vanishing.

latter will turn out to satisfy it automatically.

Recall Remark 1 after (3.6): the functions $\Phi^{h\sigma}_{j\vec{\nu}}(t)$ admit a tree representation like the functions $\Xi^{h\sigma}_{j\vec{\nu}}(t)$, the only difference being that the operation $\mathcal{E}^T_{v_0}$ relative to the time τ_{v_0} of the first node v_0 is not performed: on the contrary τ_{v_0} is set equal to t, and j_{v_0} equal to j. Consider the l-th component of $\Phi^{h\sigma}_{\vec{\nu}}(t)$, *i.e.* $\Phi^{h\sigma}_{l\vec{\nu}}(t) \equiv \Phi^{h\sigma}_{+\vec{\nu}}(t)$, and let us define $\Phi^{h\sigma}_{+\vec{\nu}}(t;\gamma_1,\ldots,\gamma_h)$ the contribution to $\Phi^{h\sigma}_{+\vec{\nu}}(t)$ corresponding to fixed values of the parameters $\gamma_1,\ldots,\gamma_h$; from the first equation in (3.9), we deduce that $y^{(-1)}_v(\tau_{v'},\tau_v)$ can be written as $y^{(-1)}_v(\tau_{v'},\tau_v) = [\cosh g\tau_{v'}]^{-1}\, Y^{(-1)}_v(\tau_v)$, with $Y^{(-1)}_v(\tau_v) = 2^{-1}\sinh g\tau_v\ \exp[in_v\varphi^0(\tau_v)]$ admitting a Laurent series in powers of $x=\exp[-\sigma_v g\tau_v]$: then the following result holds.

Proposition 5.1 : *By (2.17), (3.8) and (3.10), we can write* $Y^{(-1)}_v(\tau_v)\,\Phi^{h\sigma}_{+\vec{\nu}}(\tau_v;\gamma_1,\ldots,\gamma_h)$ *as a function of* $x=\exp[-g\tau_v]$, $\vec{\psi}=\vec{\omega}\tau_v$, *and consider its expansion in* x *and* $\vec{\psi}$. *In general, given a function* $f(t)$, *with* $f(t)=\sum_{p,\vec{q}}\hat{f}(p,\vec{q})\,x^p\,e^{i\vec{\psi}\cdot\vec{q}}$, *let us denote by* $f(t)\Big|_{p=\vec{\omega}\cdot\vec{q}=0}$ *the coefficient* $\hat{f}(0,\vec{0})$ *corresponding to the contribution* $(p,\vec{q})=(0,\vec{0})$ *in the above powers expansion, (recall the hypothesis* H *of* §*1 on* $\vec{\omega}$*); note that* $f(t)\Big|_{p=\vec{\omega}\cdot\vec{q}=0}$ *is a constant in* t. *Then the set of equations:*

$$\left\{ Y^{(-1)}_v(\tau_v)\,\Phi^{h\sigma}_{+\vec{\nu}}(\tau_v;\gamma_1,\ldots,\gamma_h)\Big|_{\substack{p(v_0)=0\\ \omega(v_0)=0}} \right\} = 0 \tag{5.1}$$

with $h\ge 1$, *defines recursively the "counterterms"* γ_p, $p\ge 1$, *satisfying the inequalities* $\gamma_p \le C^p$, *for some constant* $C>0$.

Remark 1 : The expression in left hand side of (5.1) means that the function:

$$f(\tau_v;\gamma_1,\ldots,\gamma_h)\ \equiv\ Y^{(-1)}_v(\tau_v)\,\Phi^{h\sigma}_{+\vec{\nu}}(\tau_v;\gamma_1,\ldots,\gamma_h)\ ,$$

considered as a function of $x_v=\exp[-\sigma_v g\tau_v]$ and $\psi_v=\vec{\omega}\tau_v$, has to be developed into a power series:

$$f(\tau_v;\gamma_1,\ldots,\gamma_h) = \sum_{p(v_0)=-1}^{\infty}\ \sum_{\substack{\vec{\nu}_0(v_0)\\ |\vec{\nu}(v_0)|\ge 0}} \hat{f}(p(v_0),\vec{\nu}(v_0);\gamma_1,\ldots,\gamma_h)$$

and the contribution $p(v_0)=\vec{\omega}\cdot\vec{\nu}(v_0)=0$, (which is τ_v–independent), has to be taken.

Remark 2 : Note that (5.1), with $h=1$, fixes γ_1 to a well defined value, then (5.1), with $h=2$, can be resolved giving γ_2 as a function of γ_1 which is now known, and so on; therefore we dispose of a costructive algorithm to compute the "counterterms". Note that, although (5.1) has the appearance of an implicit equation for the γ_h, this is not really the case. In fact, if we recall (2.7) and explicitly write down the terms in which γ_h appears, (5.1) takes the form:

$$\sum_{n=\pm1} n^2 f^0_n\gamma_h + \tilde{R}^{h\sigma}_{\vec{\nu}}(\gamma_1,\ldots,\gamma_{h-1}) = 0 \tag{5.2}$$

where $\sum_{n=\pm1} n^2 f^0_n \equiv 1$, (see (3.1)), for a suitable function constant $\tilde{R}^{h\sigma}_{\vec{\nu}}(\gamma_1,\ldots,\gamma_{h-1})$, taking into account all the other terms. In fact the only contribution to (5.1) depending on γ_h arises from the term of the second sum in the r.h.s. of (2.7) corresponding to $p=h$, and, if we take into account also the explicit expression of the coefficients $Y^{(-1)}_v(-1)$ and $Y^{(-1)}_v(0)$, (which can be deduced, respectively, from $y^{(-1)}_v(k_{v'},-1)$ and $y^{(-1)}_v(k_{v'},0)$), see (3.9) and the list of coefficients below (3.10)), necessary to produce an exponent $p(v_0)=0$, one immediately sees that such a contribution is given by the first term in (5.2).

Proof of Proposition 5.1 : The statement of the proposition is a rewriting in terms of the tree expansion of the property:

$$\left(Y^{(-1)}_v\ F^{h\sigma}_{+,\vec{0}}\right)(\sigma\infty) \equiv 0 \tag{5.3}$$

analogous to (2.27). The only difference is that (2.27) is imposed by the motions boundedness request, while (5.2) is imposed *a priori* in order to cancel all the terms which could generate some time powers. In fact, when $\Xi^{h_v \sigma_{v'}}_{j\vec{\nu}(v)}(t)$, $j = 0, l$, is evaluated from (2.28), according to the equations (3.12) and (3.20), then the condition that the case $p(v) = \omega(v) = 0$ can never occur implies that the degree of the time power is never increased by one, and, since $F^{1\sigma}_{+}(t) \in \mathcal{M}^0$, then $F^{h\sigma}_{+}(t), X^{h\sigma}_{\pm}(t) \in \mathcal{M}^0$ for all $h > 1$. It is easy to check that (5.1) is of the form (5.2), say inductively, (it is an obvious consequence of (2.7); see also the above remark), so that the only property that remains to be checked is that γ_p, $p \geq 1$, admits the bound C^p. But this will follow from the discussion of §5.2, and from Remark 1 after (3.6), (or from the explicit formula (2.7) expressing the functions $F^{h\sigma}$'s in terms of the functions $X^{h\sigma}$'s). So that Proposition 5.1 is proven. ∎

The formal series expansion of $g^2(\mu)$ will turn out to be convergent in μ; then the only dependence on t of the functions (2.19), taking $\vec{\psi}$ to be a parameter, is through the factor $\exp[-g\sigma t]$. If we recall that the original (true) model had g^2 instead of $g^2(\mu)$, *i.e.* the parameter appearing in the hamiltonian is not $g(\mu) = g(1 + \Gamma(\mu))$, with $\Gamma(\mu) = \sum_{p=1}^{\infty} \Gamma_p \mu^p$, but $g \equiv g(\mu) - g\Gamma(\mu)$, we can conclude that the t-dependence of the corresponding functions of the model (1.1) reveals itself through the factor $\exp[-g(1 - \Gamma(\mu))\sigma t]$.[9] Obviously we can expand such a factor in powers of μ, so reobtaining the powers of t, like in [Ge].

5.2 Bound on the tree values

In this section we prove the fundamental result of this paper, which assures the convergence of the series defining the whiskered tori, and so completes the proof of Proposition 2.1 and Proposition 2.2:

Theorem 5.1: *Let us denote by $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ the dimensionless perturbed motion, $0 \leq j < 2l$, $\sigma = \text{sign}t$. We can always write it in the form $\Xi^{h\sigma}_{j\vec{\nu}}(t) = \tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(x, \vec{\omega}t)$, where $|\vec{\nu}| \leq (2h-1)N$, $x = e^{-\sigma g t}$, and $\tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(x, \vec{\omega}t)$ is an analytic function in x, $\tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(x, \vec{\omega}t) = \sum_{p=0}^{\infty} \tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(p, \vec{\omega}t)x^p$, satisfying the bound $|\tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(x, \vec{\omega}t)| \leq \bar{D}\bar{C}^{2h-1}$, for some constants $\bar{C}, \bar{D} > 0$, and for any $\sigma t \geq 0$.*

Proof of Theorem 5.1 : Let us start by studying the term $V^S_{j\vec{\nu}}(t; \bar{\theta})$ in (5.1), where the first node v_0 of θ has $p(v_0) = 0$, and $k_{v_0} = 0, -1$; we can associate to such a tree a path $\mathcal{P}$, with the convention that $\mathcal{P} \equiv \emptyset$ if $k_{v_0} = 0$, $j_{v_0} > l$, and $\mathcal{P} \equiv v_0$ if $k_{v_0} = 0$, $j_{v_0} > l$. ¿From (3.12) and (3.20) we can obtain a sequence of factorizing integrals; then, for the top nodes $v \notin \mathcal{P}$ different from the leaves (top free nodes), we have

$$\oint \frac{dR_v}{2\pi i R_v} \int_{\sigma\infty}^{0} d\, g\tau_v \;\; T_v(-g\tau_v)\, e^{-gR_v \sum_{w \leq v} \sigma\tau_w}\; e^{i\tau_v \omega(v)}\; e^{-gp(v)\sigma\tau_v} \tag{5.4}$$

where $p(v) = k(v) = k_v$ and $\omega(v) = \omega_v$, and $T_v(-g\tau_v) = (-g\tau_v)^{1-\delta_{j_v,l}}$, see (3.14). The time integration is trivial and yields:

$$(-\sigma)^{\delta_{j_v,l}} \oint \frac{dR_v}{2\pi i R_v} \frac{e^{-gR_v \sum_{w<v} \sigma\tau_w}}{\left(R_v + p(v) - i\sigma g^{-1}\omega(v)\right)^{2-\delta_{j_v,l}}}$$

The case $\omega(v) = p(v) = 0$ can be excluded, since if $j_v = l$ then $p(v) = \pm 1$, and if $j_v > l$ then $p(v) = 0$, but the property remarked in connection with (2.27) requires in such a case $\omega(v) \neq 0$. If $j_v = l$, as we have said before, we sum together the two contributions $k_v = \pm 1$; if $j_v > l$, we have a factor $y^{(2)}_v(0,0) = 1$. Therefore the residue at $R_v = 0$ is

$$\begin{cases} \left[ig^{-1}\omega(v)\right]^{-2} & \text{if } j_v > l \\ \left[1 + g^{-2}\omega^2(v)\right]^{-1} & \text{if } j_v = l \end{cases} \tag{5.5}$$

(a factor 1/2 could be introduced in the second expression, in order to remind us not to overcount the labels $p(v) = \pm 1$, when the sum over the trees is performed).

Next we pass to the nodes immediately preceding the top ones, which can be seen as top ends of a new reduced tree obtained from $\bar{\theta}$ by deleting the original top free nodes. For each $v \in \bar{\theta}/\mathcal{P}$, (*i.e.* $v \in \bar{\theta}$, $v \notin \mathcal{P}$), we have again to consider an expression like (5.4), so that all the integrations can be performed in the same way, if only we bear in mind that the cases $p(v) = 0$, $\omega(v) = 0$ can be excluded, for the same reasons as before: this simply means that the residues are always of the form (5.5). If $v \in \mathcal{P}$, (*i.e.* $v = z_i$, $i = 1, \ldots, m_{\mathcal{P}}$), then the integration (5.4) yields:

$$(-\sigma)^{\delta_{j_v,l}} [p(z_i) - i\sigma g^{-1}\omega(z_i)]^{-(2-\delta_{j_v,l})} \tag{5.6}$$

[9] Note that $g(1 - \Gamma(\mu)) = g'(\mu)$, where $g'(\mu)$ is the function introduced in §1.

where $p(z_i)=0$ is possible only if $z_i\in\Lambda_{-1}$; in such a case it must be $\omega(z_i)\neq 0$, because of the renormalization procedure introduced in §5.1, so that (5.6) corresponds to a factor which is bounded by 1, except for the case $j_{z_i}=l$, $\alpha_{z_i}=-1$, (*i.e.* $p(z_i)=0$), which gives:

$$[ig^{-1}\omega(z_i)]^{-1} \tag{5.7}$$

with $\omega(z_i)\neq 0$.

In the end, only the node v_0 is left. If $k_{v_0}=0$, $j_{v_0}>l$, we have a coefficient $y^{(2)}_{v_0}(0,0)=1$, so we have to integrate the function $g(t-\tau_{v_0})$, if $\rho_{v_0}=1$, or $g\tau_{v_0}$, if $\rho_{v_0}=0$, times $\exp[i\omega(v_0)\tau_{v_0}]$; if $j_{v_0}=l$, then $k_{v_0}=0$ requires $v_0\in\Lambda_{-1}$, and (5.1) imposes $\omega(v_0)\neq 0$. If $k_{v_0}=-1$, again (5.1) requires $\omega(v_0)\neq 0$.[10]

We can summarize the results found so far and state the fundamental convergence bound in the following lemma.

Lemma 5.1 : *Let us consider the contribution to $\Xi^{h\sigma}_{j\vec\nu}(t)$, $\vec\nu\in\mathbf{Z}^{l-1}$, $\sigma=\pm$, $j<l$, corresponding to the stripped value $V^S_j(t;\bar\theta)$ of a reduced tree $\bar\theta$ with label $p(v_0)=0$; then, if we sum together the contributions with $\alpha_v=-1$ and $\alpha_v=1$, for each $v>v_0$, the following results hold for the sum.*
1) Such a sum can be written as:

$$A_{k_{v_0}}(t)\,e^{i\vec\omega\cdot\vec\nu_0(v_0)\rho_{v_0}t}\prod_{\bar\theta_f\ni v\geq v_0}\bar F_{\nu_v}G_v[\omega(v)] \tag{5.8}$$

where $\bar F_{\nu_v}$ is defined in (3.15), $0<|\vec\nu_0(v_0)|\leq m_0N$, m_0 being the number of free nodes in $\bar\theta$, $A_{k_{v_0}}(t)$ is the function:

$$A_{k_{v_0}}(t)=(-1)^{\rho_{v_0}}\,in_{v_0}\,[\cosh gt]^{-1}\,\delta_{k_{v_0},-1}+\delta_{k_{v_0},0}\ ,$$

and $G_v[\omega(v)]$ is defined to be:

$$G_v[\omega(v)]=\begin{cases}[ig^{-1}\omega(v)]^{-2} & \text{if } v\notin\mathcal{P}\text{ and } j_v>l\\ [1+g^{-2}\omega^2(v)]^{-1} & \text{if } v\notin\mathcal{P}\text{ and } j_v=l\\ (\sigma)^{\delta_{j_v,l}}[p(v)-i\sigma g^{-1}\omega(v)]^{-(2-\delta_{j_v,l})} & \text{if } v\in\mathcal{P}\text{ and } \alpha_v\neq -1\\ [ig^{-1}\omega(v)]^{-1} & \text{if } v\in\mathcal{P}\text{ and } \alpha_v=-1\end{cases} \tag{5.9}$$

with the third term always bounded by 1, since $|p(v)|\geq 1$ in such a case.
2) The sum over all the reduced trees with label $p(v_0)$ fixed to be zero, of the expression (5.6), admits the bound $D_0C_0^{m_0-1}$ for some constants $C_0,D_0>0$, if m_0 is the number of free nodes, $m_0<2h_0$, with $h_0\leq h$ being the reduced order of $\bar\theta$.

Proof of Lemma 5.1 : Note that the first statement is an easy consequence of the definitions, as it has been shown, while the second one is rather deep, being essentially equal to the KAM theorem, as it appears from the proof, (see also [G1], [G2], [GGe], [Ge]). So Lemma 5.1 is proven if we show that the bound $D_0C_0^{m_0-1}$, in the statement 2), holds. The sums of the stripped values of all the reduced tree can be easily controlled. If m_0 is the reduced degree of the reduced tree, the number of addends is bounded by the number of tree shapes, $(\leq 2^{2m_0}m_0!)$, see [HP], times the number of ways of attaching the ν_v, ρ_v, α_v and $p(v)$ labels, $(\leq (3N)^{lm_0}\cdot 2^{m_0}\cdot 3^{m_0}\cdot 3^{m_0})$. It remains to check that the "small divisors" in (5.7) give no problems. This is the more subtle, and will be done in Appendices A1, A2 and A3. ■

Now we pass to the reduced trees whose first node has $p(v_0)\neq 0$. For the time being, let us neglect the leaf values. If $p(v_0)=-1$, then it is $k_{v_0}=-1$, and $k_w+k'_w=0$, $\forall\ w>v_0$, so that the case can be treated as the case $p(v_0)=k_{v_0}=0$ of Lemma 5.1, with respect to which only the first node v_0 behaves in a different way; the analysis can be carried out quite unchanged, and so we do not repeat it here. Therefore in the following we can suppose $p(v_0)\neq -1$.

¿From each flower a contribution (5.8) arises, and we can explicitly perform the integrations over the time variables of the free nodes: each integration is a proper one, and gives a factor bounded by 1, so that no new "small divisor" can arise.

[10] In fact the term with $k_{v_0}=0$, and $p(v_0)=\omega(v_0)=0$ vanishes when summed to the term having $k_{v_0}=-1$, and $p(v_0)=\omega(v_0)=0$, for a suitable choise of the "counterterms" γ_p, $p\geq 1$, as it has been shown in §4.

Nevertheless we must be careful, because we still have to sum over the labels $p(v)$, $v \geq v_0$, (the sum over the other labels can be treated as in the previous case). We can resolve this (apparent) problem as follows. If $\rho_{v_0} = 1$, $\sigma t \leq g^{-1}$, we split the integral over τ_{v_0}:

$$\int_{\sigma\infty}^{gt} d\,g\tau_{v_0}\,(\ldots) = \int_{\sigma\infty}^{\sigma} d\,g\tau_{v_0}\,(\ldots) + \int_{\sigma}^{gt} d\,g\tau_{v_0}\,(\ldots) \equiv I_m + \int_{\sigma}^{gt} d\,g\tau_{v_0}\,(\ldots) \tag{5.10}$$

and we consider the first term. Once all the integrations are performed, we are left with a contribution which is the product of a factor admitting a "good m-bound" times a factor of the form $\exp[-p(v_0)]$. Then we can choose $\lambda = 1/2$ in (3.17) in order to get a convergent bound: at worst for every node v we have a factor $2^{k_v+k_{v'}}$ and a factor $e^{-k_v-k_{v'}}$ so that we can perform the summation over the indices k_v, $k_{v'} \geq -1$, (see (3.10)), and the convergence follows. We have left the term in (5.10) in which the first time variable τ_{v_0} has to be integrated between σg^{-1} and t, but one finds that, in the more general case, the integrals can be written as:

$$I_{m_1}\ldots I_{m_p} \prod_{v\in\tilde{\theta}_f} \fint_{\sigma}^{g\tau_{v'}} dg\tau_v(\ldots)$$

(all the free nodes v's have $p(v) \neq 0$, so that $\rho_v = 1$) where $\tilde{\theta}$ is a subtree of $\bar{\theta}$ with first node v_0 and $\tilde{m}$ nodes, with $\tilde{m} + m_1 + \ldots + m_p = m$, and the last integral is manifestly bounded (see also [G1]), so that we see that the only very problem is to show that $I_m \leq C^m$, for some constant C. If $\sigma t > g^{-1}$, we obtain from the last integration, (the one corresponding top the first node v_0), the factor $\exp[-p(v_0)g\sigma t]$, so that, since $\exp[-p(v_0)g\sigma t] \leq \exp[-p(v_0)]$ we can repeat the above argument to deduce the convergence. Eventually, if $\rho_{v_0} = 0$, the same discussion applies, and, in particular, only the first case has to be treated.

Obviously we have to take into account also the values of the leaves. However, if we are interested, say, in the contribution to order h, the reduced order h_0 of the reduced tree and the orders h_i, $i = 1, \ldots, \mathcal{N}_L$, of the $\mathcal{N}_L$ leaves have to be such that $h = h_0 + \sum_{i=1}^{\mathcal{N}_L} h_i$. So we can arrange the sums as follows: fixed h, we sum over $h_0 = 1, \ldots, h$, and, fixed h_0, we sum over the orders of the leaves with the constraint $\sum_{i=1}^{\mathcal{N}_L} h_i = h - h_0$; then we sum over all the reduced trees of fixed order h_0 with $\mathcal{N}_L$ leaves of fixed orders, respectively, h_i, $i = 1, \ldots, \mathcal{N}_L$. Since the value of a leaf of order h_v represents a contribution to $\Xi^{h_v\sigma_v}_{j_{\lambda_v}\vec{\nu}(v)}(0)$, it can be treated in the same way, and therefore admits the same bound.

Therefore we can inductively check, by exploiting the results of Lemma 5.1 too, (as far as the leaves with label $p(v) = 0$ are concerned), that the contribution to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$, $\vec{\nu} \in \mathbf{Z}^{l-1}$, $\sigma = \pm 1$, $2l > j \geq 0$, $|\vec{\nu}_0(v_0)| \leq mN$, arising from the sum of the values of all the reduced trees of degree m, with labels $p(v_0) \neq 0$, can be bounded by $D_2 C_2^{m-1}$ for some constants $D_2, C_2 > 0$. In fact, a leaf v with $p(v) = 0$ contributing, *e.g.* , to the reduced tree value through the factor $L^{h_{v_i}\sigma_{v_i}}_{j_i\vec{\nu}(v_i)}(0)$ admits a representation analogous to the same (5.1) and can be expressed as a sum of terms, which are given by the product of the stripped value of the reduced tree with first node v times the values of its leaves. The procedure can be iterated for all the leaves with $p(v)$ labels equal to zero, and in this way we can get rid of them and are left only with leaves having $p(v) \neq 0$. Then the bound $D_2 C_2^{m-1}$ can be assumed to hold, and an inductive proof can be performed.

It remains to study the case $p(v_0) = 0$, $j \geq l$, but one immediately see that this can be discussed as the case $p(v_0) = 0$, $j < l$, of Lemma 5.1, so that Theorem 5.1 is proven. ■

Appendix A1: Proof of the convergence bound in Lemma 5.1

As we have seen in §5.2, from the case $p(v_0) = 0$ we obtain a contribution to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ containing a product:

$$\prod_{v\geq v_0} \bar{F}_{\nu_v} G_v[\omega(v)] \tag{A1.1}$$

where $G_v[\omega(v)]$ is defined in (5.9), and we want to find a bound on the sum of (A1.1) over all the reduced trees with $p(v_0)$ fixed to the above value. We consider together the cases $k_{v_0} = 0, -1$; if $k_{v_0} = 0$ the path $\mathcal{P}$ is supposed to be reduced to a single node, v_0, or to the empty set, $\emptyset$, according to the value of j_{v_0}, (respectively $j_{v_0} = l$, and $j_{v_0} > l$; see also the first paragraph in the proof of Theorem 5.1).

Given a reduced tree $\bar{\theta}$, it will be characterized by its shape and by a collections of labels. Let us proceed as in [G1], [G2], and let us suppose a condition over the rotation vectors stronger than the hypothesis H in

§1, *i.e.* let us suppose that they satisfy a *strong diophantine condition*. This is not really necessary, but it simplifies the proof, and, once the result is obtained, we can reason as in [GGe] to eliminate such an unneeded hypothesis; as the discussion can be repeated quite unchanged with respect to [GGe], we simply refer to it. Therefore we shall make the assumption that the rotation vectors $\vec{\omega}$'s satisfy the *strong diophantine condition*:

$$\begin{array}{lll} 1) & C_0|\vec{\omega}\cdot\vec{\nu}| \ge |\vec{\nu}|^{-\tau} & \vec{0}\ne\vec{\nu}\in \mathbf{Z}^{l-1} \\ 2) & \min\limits_{0\ge p\ge n}\Big|C_0|\vec{\omega}\cdot\vec{\nu}|-2^p\Big| \ge 2^{n+1} & \text{if} \quad n\le 0,\ \ 0<|\vec{\nu}|\le (2^{n+3})^{-\tau^{-1}} \end{array} \tag{A1.2}$$

where $n, p\in\mathbf{Z}$, $n\le 0$. We fix a scaling parameter γ, which we take $\gamma=2$, and define (in analogy to quantum field theory: see, *e.g.*, [BfG], [G4]) a propagator:

$$G\equiv G_v[\omega(v)] = \begin{cases} (gC_0)^2[i\vec{\omega}_0\cdot\vec{\nu}_0(v)]^{-2} & \text{if } v\notin\mathcal{P} \text{ and } j_v>l \\ (gC_0)^2\left[(gC_0)^2+(\vec{\omega}_0\cdot\vec{\nu}_0(v))^2\right]^{-1} & \text{if } v\notin\mathcal{P} \text{ and } j_v=l \\ (\sigma)^{\delta_{j_v,l}}(gC_0)^{2-\delta_{j_v,l}}\left[(gC_0)p(v)-i\sigma\vec{\omega}_0\cdot\vec{\nu}_0(v)\right]^{-(2-\delta_{j_v,l})} & \text{if } v\in\mathcal{P} \text{ and } \alpha_v\ne -1 \\ (gC_0)\left[i\vec{\omega}_0\cdot\vec{\nu}_0(v)\right]^{-1} & \text{if } v\in\mathcal{P} \text{ and } \alpha_v=-1 \end{cases} \tag{A1.3}$$

where $\vec{\omega}_0=C_0\vec{\omega}$ is a dimensionless frequency, and we say that:
1) G is on scale 1, if $|\vec{\omega}_0\cdot\vec{\nu}_0(v)|>1$;
2) G is on scale $n\le 0$, if $2^{n-1}<|\vec{\omega}_0\cdot\vec{\nu}_0(v)|\le 2^n$.

Note that, if $v\notin\mathcal{P}$, $j_v>l$, then, if G is on scale $n\le 0$, it is $|G|<(gC_0)^2 2^{-2(n-1)}$, and, if it is on scale 1, it is $|G|<(gC_0)^2$, while, if $v\notin\mathcal{P}$, $j_v=l$, then $|G|\le 1$; if $v\in\mathcal{P}$, if $\alpha_v\ne -1$, then $|G|\le 1$, otherwise, if $\alpha_v=-1$, then $|G|<(gC_0)2^{-(n-1)}$. We can get rid of the new factor $(gC_0)^2$, by defining $C_1=\max\{1,(gC_0)^2\}$, and introducing a coefficient C_1^m in the bound (3.16). This implies a simple redefinition of the constant $\mathcal{C}$ in (3.16), and we can say that, if G is on scale n, then, $\forall\, n\le 1$, $|G|<2^{-2(n-1)}$, if $v\notin\mathcal{P}$, and $|G|<2^{-(n-1)}$, if $v\in\mathcal{P}$.

*Henceforth (and in the following two appendices), with an abuse of notation aiming to not overwhelm the discussion, let us use the term "tree" instead of the more cumbersome "reduced tree", (and the symbol θ instead of $\bar{\theta}$); however it is always in the meaning of the latter that the first one has to be interpretated. Moreover we call momentum (*tout court*) of the node v the free momentum $\vec{\nu}_0(v)$.*

Given a tree θ we can attach a *scale label* to each branch $v'v$ (v' being the node preceding v): it is equal to n if n is the scale of the branch propagator. Note that the labels thus attached to a tree are uniquely determined by the tree: they will have only the function of helping to visualize the orders of magnitude of the various tree branches.

Looking at such labels we identify the connected clusters T of nodes that are linked by a continuous path of branches with the same scale label n_T or a higher one. We shall say that *the cluster T has scale n_T*. Since the tree branches carry an arrow pointing to the root, (see §3), we can associate to each cluster a collection of incoming branches (*branches entering T*) and a collection of outgoing branches (*branches exiting from T*).

Definition A1.1: *Among the clusters we consider the ones with the property that there is only one tree branch entering them and only one exiting and both carry the same momentum. If V is one such cluster, we denote λ_V the outgoing branch, and $n=n_{\lambda_V}$ its scale label. We say that such a V is a* resonance *if the number of branches contained in V is $\le E\,2^{-n\varepsilon}$, where E,ε are defined by: $E\equiv 2^{-3\varepsilon}N^{-1}$, $\varepsilon=\tau^{-1}$. We shall say that n_{λ_V} is the* resonance scale, *and λ_V a* resonant line.

Note that if λ_V is the branch exiting from the resonance V, the branch scale n_{λ_V} is smaller than the smallest scale $n'=n_V$ of the branches inside V.

Definition A1.2: *Given a resonance V, let λ_v and $\lambda_{v'}$ be, respectively, the incoming and outgoing branches, (so that $\lambda_V\equiv\lambda_{v'}$), and v, v' the nodes which λ_v, $\lambda_{v'}$, respectively, lead to (v' is inside the resonance, and v outside).*[11] *We say that V is a* strong resonance *if it is $\vec{\nu}_0(v)=\vec{\nu}_0(v')$, (as in all resonances), and $p(v)=p(v')\equiv 0$. A tree with strong resonances will be called a* resonant tree.

Remark : We shall see in the following discussion that only the strong resonances can give problems, so that in fact they are the only "true resonances" (in the usual meaning of the word). The reason why we have

[11] Recall that the ordering is opposite to the gravity.

introduced a new name for them is simply to maintain the definition of resonance given in [G1], as it will turn out that some properties which we need follow from the very definition of resonance, and it will be not important if the considered resonances are strong or not (see, in particular, Appendix A3).

Definition A1.3: *Given a propagator G as in (A1.3), we associate to it a label which we call the degree D_G of the propagator, and which we set equal to 2, if G is given by the first or second term of the r.h.s. of (A1.3), and equal to 1, if G is given by the third or fourth term. We associate to each cluster T a label D_T, which we call the* degree *of the cluster T: we set $D_T = j$, $j = 1, 2$, if the propagator of the corresponding outgoing branch has degree j. Given a strong resonance V on scale n, let us consider the outgoing branch (resonant line); the corresponding propagator is given by either the first term of the r.h.s. of (A1.3), or the fourth one; then the degree of a strong resonance is equal to the degree of the corresponding resonant line, and the propagator G of the resonant line on scale n is such that $|G| < 2^{-D_V(n-1)}$, if D_V is the degree of the strong resonance.*

The key remark is that the resonant trees (see Definition A1.1) cancel almost exactly. We have already all is needed to see why this happens. We can reason in the following way.

Given a tree θ with a strong resonance V, we call, as before, v the node which the entering branch leads to, and v' the node which the exiting branch leads to; moreover let us call θ_2 the subtree with first node v. Imagine to detach from the tree θ the subtree θ_2, then attach it to all the remaining nodes $w \in V$. We obtain a family of trees whose contributions to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ differ because:
1) some of the branches above v' have changed momentum by the amount $\vec{\nu}_0(v)$: this means that, if $\varepsilon \equiv \omega_0(v)$ and $w \in V$, some of the propagators $\left[i\omega_0(w)\right]^{-2}$ have become $\left[i(\omega_0(w) + \varepsilon)\right]^{-2}$, some of the propagators $\left[(gC_0)^2 + (\omega_0^2(w))^2\right]^{-1}$ have become $\left[(gC_0)^2 + (\omega_0(w) + \varepsilon)^2\right]^{-1}$, some of the propagators $(\sigma)^{\delta_{j_w,l}}\left[(gC_0)p(w) -i\sigma\omega_0(w)\right]^{-(2-\delta_{j_w,l})}$ have become $(\sigma)^{\delta_{j_w,l}}\left[(gC_0)p(w) -i\sigma(\omega_0(w)+\varepsilon)\right]^{-(2-\delta_{j_w,l})}$ and some of the propagators $\left[i\omega_0(w)\right]^{-1}$ have become $\left[i\omega_0(w) + \varepsilon\right]^{-1}$, and:
2) there is one of the node factors which changes by taking successively the values ν_{wj}, $j \equiv j_{\lambda_v}$ being the branch label of the branch leading to v, and $w \in V$ is the node to which such a branch is reattached.

Hence if $\varepsilon = 0$ we would build in this resummation a quantity proportional to: $\sum_{w\in V} \nu_{wj} = \nu_{0j}(v) - \nu_{0j}(v')$, which is zero, because $\vec{\nu}_0(v') = \vec{\nu}_0(v)$ means that the sum of the $\vec{\nu}_w$'s vanishes, and $0 < j < l$, if $p(v) = 0$.

Since $\varepsilon \neq 0$, we can expect to see a sum of order ε^2 for the strong resonances of degree 2, if we sum as well on a overall change of sign of the ν_w values (whose components $\vec{\nu}_w$ sum up to $\vec{0}$, so that all the $\vec{\nu}_w$ can reverse their direction without breaking the relationship which has to exist between the modes). We use the fact that for each branch inside the resonance we have a propagator which is an even function in its argument: if the strong resonance V has $D_V = 2$, then there is no path $\mathcal{P}$ inside V, so that the only propagators we can associate to the branches internal to V are of the form of the first two terms in the r.h.s. of (A1.3).[12] Moreover, in such cases, no n_w label apppears in the $y_w^{(\alpha_v)}$'s, (see the list of coefficients after (3.10)), so that all the dependence on the n_w's is through the factors $\bar{F}_{\nu_w}$ of (3.15): therefore there is an even number of the n_w's, (if there are any), corresponding to the nodes inside the resonance (two for each branch), so that no change is produced by the sign reversal, (recall also that $f^{\delta_v}_{\nu_v} \equiv f^{\delta_v}_{-\nu_v}$). On the contrary, if the strong resonance V has degree $D_V = 1$, then only a sum of order ε can be obtained, but this is enough, since in this case the "small divisor" appears to the first power.

All that has been said can be true only if $\varepsilon \ll \omega(w)$ for any $w \in V$: but our Definition A1.1 of resonance has been set up precisely to make such property automatically verified, as it is explained in Appendix A2.

Once we have singled out the trees which need a more careful analysis, and found that they have almost the same properties of the resonant trees defined in [G1], we can proceed in the same way of the quoted reference: in fact the discussion follows quite closely [G1], Appendix A3, (with respect to which a slight change is required in order to treat the strong resonance of degree 1), and so we relegate it to Appendix A2 below. Here we confine ourselves to state the final result.

Let us denote $m^j_T(n)$ the number of strong resonances of scale n and degree j, contained in a cluster T, and define the tree family $\mathcal{F}(\theta)$ as follows.[13] Given a strong resonance V of θ we detach the part of θ above λ_v (λ_v included), being λ_v the entering branch of the resonance, (see Definition A1.2), and attach

[12] More generally, if a cluster T has degree $D_T = 2$, then all the clusters inside T have degree 2; and if the propagator corresponding to a line λ has degree 2, the propagators of all the lines in θ following λ have the same degree 2.

[13] Such definitions will become more clear in Appendix A2, as that of resonance given above, which has not been completely used so far.

it successively to the points $w \in \tilde{V}$, where $\tilde{V}$ is the set of nodes of V (including the endpoint w_1 of λ_v contained in V) outside the resonances contained in V, and, if $D_V = 2$, we add also the trees obtained by reversing simoultaneously the signs of the node modes $\vec{\nu}_w$, $w \in V$. We repeat the entire procedure for all the resonances of θ.

Then the result is that the contribution to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ we obtain from a given trees family $\mathcal{F}(\theta)$ is bounded by:

$$\frac{1}{m!}\left[2^{4m}e^{2m}\prod_{n\le 0}2^{-2nN_n^2}2^{-nN_n^1}\right]\left[\prod_{n\le 0}\prod_{T,\,n_T=n}\prod_{i=1}^{m_T^1(n)}2^{(n-n_i+3)}\prod_{i=1}^{m_T^2(n)}2^{2(n-n_i+3)}\right] \tag{A1.4}$$

where:
1) N_n^j is the number of propagators of scale n and of degree j in θ ($n = 1$ does not appear as $|G| \ge 1$ in such case), which can be written as:

$$N_n^j = \bar{N}_n^j + \sum_{\substack{T \\ n_T=n, D_T=j}} (-1) + \sum_{\substack{T \\ n_T=n}} m_T^j(n) \tag{A1.5}$$

with the terms $\bar{N}_n^j$, $j = 1, 2$, satisfying the inequality:

$$\sum_{j=1}^{2} \bar{N}_n^j \le \frac{4m}{E\,2^{-\varepsilon n}} \tag{A1.6}$$

(which is proven in Appendix A3).
2) The first square bracket is the bound on the product of individual elements in the family $\mathcal{F}(\theta)$ times the bound e^{2m} on their number (see Appendix A2).
3) The second term is the part coming from the maximum principle, (in the form of Schwarz's lemma), applied to bound the resummations, as it is explained in Appendix A2, n_i being the scale of the cluster V_i which is the i-th resonance inside T, (note that the resonance scale is n, see Definition A1.1).

Hence substituting (A1.5) and (A1.6) into (A1.4) we see that, for $j = 1, 2$, the $m_T^j(n)$ is taken away by the first factor in $2^{jn}2^{-jn_i}$, while the remaining 2^{-jn_i} are compensated by the -1 before the $+m_T^j(n)$ in (A1.4) taken from the factors with $T = V_i$ (note that there are always enough -1's), and therefore the product (A1.4) is bounded by:

$$\frac{1}{m!}e^{2m}2^{4m}2^{6m}2^m\prod_n 2^{-8nmE^{-1}2^{\varepsilon n}} \le \frac{1}{m!}B_0^m \tag{1.7}$$

for $B_0 = e^2 2^{11}\exp\left[\left(2^{3+3\tau^{-1}}\log 2\right)\sum_{p=1}^{\infty} p 2^{-p\tau^{-1}}\right]$. Note that the propagators with $j_v = l$, $v \notin \mathcal{P}$, and the propagators with $j_v = l$, $v \in \mathcal{P}$, $\alpha_v \ne -1$, are bounded by 1, independently on the scale label n: in fact the above described algorithm produces a gain only for the strong resonances. Then the bound of the second statement of Lemma 5.1, with $m = m_0$, follows. ∎

Appendix A2: Approximate cancellation of the strong resonances

Let us consider a tree θ and its clusters. We wish to estimate the number N_n of branches with scale $n \le 0$ in it, assuming $N_n > 0$. Denoting T a cluster of scale n, and $m_T^j(n)$ the number of resonances of scale n and degree j contained in T, (*i.e.* with outgoing branches of scale n and degree j), we have the relation (A1.5) supplemented by the inequality (A1.6), which is an adaptation of the version of *Brjuno's lemma* as it is exposed in [P]: a proof is in Appendix A3, and is taken from [G1], with some minor changes.

Recall that, given a tree θ^1, we define the family $\mathcal{F}(\theta^1)$ generated by θ^1 as follows. If V is a resonance of θ^1 we detach the part of θ^1 above λ_v, (recall Definition A1.2), and attach it successively to the points $w \in \tilde{V}$, where $\tilde{V}$ is the set of nodes of V (including the endpoint w_1 of λ_v contained in V) outside the resonances contained in V. We say that a branch λ is in $\tilde{V}$, if λ is contained in V and has at least one point in $\tilde{V}$; we denote by n_λ its scale. For each resonance V of θ^1 we shall call M_V the number of nodes in $\tilde{V}$. If the resonance degree is $D_V = 2$, then to the just defined set of trees we add the trees obtained by reversing simoultaneously the signs of the node modes ν_w, for $w \in \tilde{V}$: the change of sign is performed independently for the various resonant clusters. This defines a family of $\le \prod 2M_V$ trees that we call $\mathcal{F}(\theta_1)$. The number $\prod 2M_V$ will be bounded by $\exp \sum 2M_V \le e^{2m}$.

It is important to note that the definition of resonance given in Definition A1.1 is such that the above operation (of shift of the node to which the branch entering V is attached) does not change too much the scales of the tree branches inside the resonances: the reason is simply that inside a resonance of scale n the number of branches is not very large being $\le \overline{N}_n \equiv E\,2^{-n\varepsilon}$. Let λ be a branch, in a cluster T, contained inside the resonances $V = V_1 \subset V_2 \subset \ldots$ of scales $n = n_1 > n_2 > \ldots$: then the shifting of the branches λ_{V_i} can cause at most a change in the size of the propagator of λ by at most $2^{n_1} + 2^{n_2} + \ldots < 2^{n+1}$.

Since the number of branches inside V is smaller than $\overline{N}_n$ the quantity $\vec{\omega}_0 \cdot \vec{\nu}_\lambda$ of λ has the form $\vec{\omega}_0 \cdot \vec{\nu}^0_\lambda + \sigma_\lambda \vec{\omega}_0 \cdot \vec{\nu}_{\lambda_V}$ if $\vec{\nu}^0_\lambda$ is the momentum of the branch λ "inside the resonance V", *i.e.* it is the sum of all the $\vec{\nu}_v$ of the nodes v preceding λ in the sense of the branch arrows, but contained in V; and $\sigma_\lambda = 0, \pm 1$. Therefore not only $|\vec{\omega}_0 \cdot \vec{\nu}^0_\lambda| \ge 2^{n+3}$ (because $\vec{\nu}^0_\lambda$ is a sum of $\le \overline{N}_n$ node modes, so that $|\vec{\nu}^0_\lambda| \le N\overline{N}_n$), but $\vec{\omega}_0 \cdot \vec{\nu}^0_\lambda$ is "in the middle" of the diadic interval containing it and does not get out of it if we add a quantity bounded by 2^{n+1} (like $\sigma_\lambda \vec{\omega}_0 \cdot \vec{\nu}_{\lambda_V}$): this follows from the second inequality in (A1.2), *i.e.* from the strong diophantine condition hypothesis. *Hence no branch changes scale as θ varies in $\mathcal{F}(\theta^1)$, if $\vec{\omega}$ verifies a strong diophantine condition.*

Let θ^2 be a tree not in $\mathcal{F}(\theta^1)$ and construct $\mathcal{F}(\theta^2)$, *etc.* We define a collection $\{\mathcal{F}(\theta^i)\}_{i=1,2,\ldots}$ of pairwise disjoint families of trees. We shall sum all the contributions to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ coming from the individual members of each family. This is a basic feature of the summation procedure.[14]

We call ε_V the quantity $\vec{\omega}_0 \cdot \vec{\nu}_{\lambda_V}$ associated with the resonance V. If λ is a line with both extremes in $\tilde{V}$ we can imagine to write the quantity $\vec{\omega}_0 \cdot \vec{\nu}_\lambda$ as $\vec{\omega}_0 \cdot \vec{\nu}^0_\lambda + \sigma_\lambda \varepsilon_V$, with $\sigma_\lambda = 0, \pm 1$. Since $|\vec{\omega}_0 \cdot \vec{\nu}_\lambda| > 2^{n_V - 1}$ we see that the product of the propagators is holomorphic in ε_V for $|\varepsilon_V| < 2^{n_V-3}$. In fact $|\vec{\omega}_0 \cdot \vec{\nu}^0_\lambda| \ge 2^{n+3}$ because V is a resonance; therefore $|\vec{\omega}_0 \cdot \vec{\nu}_\lambda| \ge 2^{n+3} - 2^n \ge 2^{n+2}$ so that $n_V \ge n+3$. On the other hand note that $|\vec{\omega}_0 \cdot \vec{\nu}^0_\lambda| > 2^{n_V-1} - 2^n$ so that $|\vec{\omega}_0 \cdot \vec{\nu}^0_\lambda + \sigma_\lambda \varepsilon_V| \ge 2^{n_V-1} - 2^n - 2^{n_V-3} \ge 2^{n_V-1} - 2\,2^{n_V-3} \ge 2^{n_V-2}$, for $|\varepsilon_V| < 2^{n_V-3}$. While ε_V varies in such complex disk the quantity $|\vec{\omega}_0 \cdot \vec{\nu}_\lambda|$ does not become smaller than $2^{n_V-1} - 2\,2^{n_V-3} \ge 2^{n_V-2}$. Note that the quantity 2^{n_V-3} will usually be $\gg 2^{n_{\lambda_V}-1}$ which is the value ε_V actually can reach in every tree in $\mathcal{F}(\theta^1)$; this can be exploited in applying the maximum priciple, as done below.

It follows that, if V is a strong resonance, calling n_λ the scale of the branch λ in θ^1, each of the $\le \prod 2M_V \le e^{2m}$ products of propagators of the members of the family $\mathcal{F}(\theta^1)$ can be bounded above by $\prod_\lambda 2^{-D_V(n_\lambda-2)} \le 2^{4m} \prod_\lambda 2^{-D_V n_\lambda}$, if regarded as a function of the quantities $\varepsilon_V = \vec{\omega}_0 \cdot \vec{\nu}_{\lambda_V}$, for $|\varepsilon_V| \le 2^{n_V-3}$, associated with the resonant clusters V. This even holds if the ε_V are regarded as independent complex parameters.

By construction it is clear that the sum of the $\prod 2M_V \le e^{2m}$ terms, giving the contribution from the trees in $\mathcal{F}(\theta^1)$, vanishes to j-th order, $j = D_V$, in the ε_V parameters, (by the approximate cancellation discussed in Appendix A1). Hence we can apply the maximum principle to bound the contribution from the family $\mathcal{F}(\theta^1)$, so obtaining the second term in square brackets of (A1.6); the result is explained as follows:

1) the dependence on the variables $\varepsilon_{V_i} \equiv \varepsilon_i$ relative to resonances $V_i \subset T$ with scale $n_{\lambda_V} = n$ is holomorphic for $|\varepsilon_i| < 2^{n_i-3}$ if $n_i \equiv n_{V_i}$, provided $n_i > n+3$.

2) the resummation says that the dependence on the ε_i's has a first order zero in each, if the strong resonance degree is 1, and a second order zero in each, if the strong resonance degree is 2. Hence the maximum principle tells us that we can improve the bound given by the first factor in (A1.4) by the product of factors $(|\varepsilon_i|\,2^{-n_i+3})^j$, $j = D_V$, if $n_i > n+3$. If $n_i = n+3$ we cannot gain anything: but since the contribution to the bound from such terms in (A1.4) is > 1, we can leave them in it to simplify the notation. The details can be found in Appendix A4.

Appendix A3: Resonant Siegel-Brjuno bound

In the following discussion, we consider the scale labels, so that, it is quite irrelevant which value the $p(v)$'s, $v \in \theta$, assume, and therefore which resonances are strong and which are not.

---

[14] The proof of the convergence bound of Lemma 5.1 presented here is obtained by exploiting some cancellations we can implement by summing together different reduced trees, (inside the same family $\mathcal{F}(\theta)$); one could think that the leaf values give problems, since they introduce an extra difference between the terms we sum, so making us loose the cancellation mechanism. This is not the case, because the reduced trees appearing in $\mathcal{F}(\theta)$ are obtained by shifting a part of θ, *with all its leaves*, so that no further difference is introduced. To be more precise, we rearrange the sums as follows: fix a reduced tree θ, with all its leaves of fixed orders; then we sum over all the terms of the family $\mathcal{F}(\theta)$, in which θ is contained, so that the cancellation mechanism is implemented.

Calling N^*_n the number of non resonant lines carrying a scale label $\le n$. We shall prove first that $N^*_n \le 2m(E2^{-\varepsilon n})^{-1} - 1$ if $N^*_n > 0$.

If θ has the root line either with scale $> n$, or with scale n and resonant, then calling $\theta_1, \theta_2, \ldots, \theta_k$ the subtrees of θ emerging from the first node of θ and with $m_j > E\,2^{-\varepsilon n}$ lines, $j = 1, \ldots, k$, it is $N^*_n(\theta) = N^*_n(\theta_1) + \ldots + N^*_n(\theta_k)$ and the statement is inductively implied from its validity for $m' < m$ provided it is true that $N^*_n(\theta) = 0$ if $m < E2^{-\varepsilon n}$, which is is certainly the case if E is chosen as in (1.7).[15]

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

If $k = 1$ we once more have a trivial case unless the order m_1 of θ_1 is $m_1 > m - \frac{1}{2}E\,2^{-n\varepsilon}$. Finally, and this is the real problem as the analysis of a few examples shows, we claim that in the latter case the root line of θ_1 has scale $> n$.

Accepting the last statement it will be: $N^*_n(\theta) = 1 + N^*_n(\theta_1) = 1 + N^*_n(\theta'_1) + \ldots + N^*_n(\theta'_{k'})$, with θ'_j being the k' subtrees emerging from the first node of θ'_1 with orders $m'_j > E\,2^{-\varepsilon n}$: this is so because the root line of θ_1 will not contribute its unit to $N^*(\theta_1)$. Going once more through the analysis the only non trivial case is if $k' = 1$ and in that case $N^*_n(\theta'_1) = N^*_n(\theta''_1) + \ldots + N_n(\theta''_{k''})$, *etc*, until we reach a trivial case or a tree of order $\le m - \frac{1}{2}E\,2^{-n\varepsilon}$.

It remains to check that if $m_1 > m - \frac{1}{2}E\,2^{-n\varepsilon}$ then the root line of θ_1 has scale $> n$. Since the root line of θ is not a resonant line, the root line of θ_1 cannot carry the same momentum. Suppose that the root line of θ_1 is on scale n. Then $|\vec\omega \cdot \vec\nu_0(v_0)| \le 2^n, |\vec\omega \cdot \vec\nu_0(v_1)| \le 2^n$, if v_0, v_1 are the first nodes of θ and θ_1 respectively. Hence $\delta \equiv |(\vec\omega \cdot (\vec\nu_0(v_0) - \vec\nu_0(v_1))| \le 2\,2^n$ and the diophantine assumption implies that $|\vec\nu_0(v_0) - \vec\nu_0(v_1)| > (2\,2^n)^{-\tau^{-1}}$, or $\vec\nu_0(v_0) = \vec\nu_0(v_1)$. The latter case being discarded as the root line of θ is non resonant, it follows that $m - m_1 < \frac{1}{2}E\,2^{-n\varepsilon}$ is inconsistent: it would in fact imply that $\vec\nu_0(v_0) - \vec\nu_0(v_1)$ is a sum of $m - m_1$ node modes and therefore $|\vec\nu_0(v_0) - \vec\nu_0(v_1)| < \frac{1}{2}NE\,2^{-n\varepsilon}$ hence $\delta > 2^3\,2^n$ which is contradictory with the above opposite inequality.

A similar induction can be used to prove that if $N^*_n > 0$ then the number p^*_n of clusters of scale n verifies the bound $p^*_n \le 2m\,(E2^{-\varepsilon n})^{-1} - 1$. In fact this is true for $m \le E2^{-\varepsilon n}$. Let, therefore, $p(\theta)$ be the number of clusters of scale n: if the first tree node v_0 is not in a cluster of scale n it is $p(\theta) = p(\theta_1) + \ldots + p(\theta_k)$, with the above notation, and the statement follows by induction. If v_0 is in a cluster of scale n we call $\theta_1, \ldots, \theta_k$ the subdiagrams emerging from the cluster containing v_0 and with orders $m_j > E2^{-\varepsilon n}$. It will be $p(\theta) = 1 + p(\theta_1) + \ldots + p(\theta_k)$. Again we can assume that $k = 1$, the other cases being trivial. But in such case there will be only one branch entering the cluster V of scale n containing v_0 and it will have a propagator of scale $\le n - 1$. Therefore the cluster V must contain at least $E2^{-\varepsilon n}$ nodes. This means that $m_1 \le m - E2^{-\varepsilon n}$.

Therefore we add and subtract from N^j_n the quantity $\sum_T m^j_T(n)$, where the sum is over the clusters satisfying the constraint $n_T = n, D_T = j$, and exploit the inequality:

$$\sum_{j=1}^{2} \Big[N^j_n + \sum_{\substack{T \\ n_T = n, D_T = j}} m^j_T(n) \Big] \le N^*_n + p^*_n$$

so that (A1.6) is proven.

Appendix A4: Dimensional estimate of the order of zero in ε_i.

Consider a family $\mathcal{F}(\theta^1) \equiv \mathcal{F}$. Let B be the first factor in (A1.4) without the e^{2k}, *i.e.* a "naive" bound on the sum of the values of each of the trees in the family.

Between the resonances there exists an inclusion relation; let us define "first generation resonances" the innermost resonances, *i.e.* the resonances $V^1_{j_1}$, $j_1 \ge 1$, containing no other resonances, "second generation resonances" the next to innermost resonances $V^2_{j_2}$, $j_2 \ge 1$, *i.e.* the resonances which become innermost if all the original innermost ones are regarded as single nodes, and so on. Let $\varepsilon^i_{j_i} = \vec\omega \cdot \vec\nu_{\lambda_{V^i_{j_i}}}$: each $\varepsilon^i_{j_i}$ is a function of the values $\varepsilon^k_{j_k}$, corresponding to resonances following $V^i_{j_i}$ along the tree.

[15] Note that, if $m \le E\,2^{-n\varepsilon}$, it is, for all momenta $\vec\nu$ of the lines, $|\vec\nu| \le NE\,2^{-n\varepsilon}$, *i.e.* $|\vec\omega \cdot \vec\nu| \ge (NE\,2^{-n\varepsilon})^{-\tau} = 2^3\,2^n$ so that there are no clusters T with $n_T = n$ and $N^*_n = 0$.

Consider a first generation resonance V_1^1 of scale $n_{V_1^1}$: it is $|\varepsilon_1^1| < \gamma_{n_{\lambda_{V_1^1}}}$, and the values of the trees in $\mathcal{F}$ are analytic in ε_1^1 for $|\varepsilon_1^1| < \gamma_{n_{V_1^1}-3}$. Note that if the other $\varepsilon_{j_i}^i$'s vary in their analyticity domains, ε_1^1, considered as a function of them, can assume a value outside its own analyticity domain when $n_{V_1^1} = n_{\lambda_{V_1^1}} + 3$,[16] although the contribution of the factors corresponding to the tree branches to the first square bracket in (A1.4) remains as in (A1.4).

Then if we sum the values of the considered trees collecting them into families (of $\leq 2M_{V_1^1}$ terms) corresponding to the Eliasson's resummation related to the resonance V_1^1, *only*, we obtain a sum of functions each of which has a zero of second order in ε_1^1, independently on the other values $\varepsilon_{j_i}^i$'s.

Therefore the considered sums are bounded by:

$$B\left(2M_{V_1^1}\, 2^{D_{V_1^1}(n_{\lambda_{V_1^1}}-n_{V_1^1}+3)}\right) ,$$

when the "gain factor" can be left also when it is not obtained, (*i.e.* if $n_{V_1^1} = n_{\lambda_{V_1^1}} + 3$), since, in such a case, the dimensional bound, which cannot be improved, is given by B, and $n_{\lambda_{V_1^1}} - n_{V_1^1} + 3 = 0$.

We then consider another innermost resonance V_2^1 (if existent). We perform the same resummation, obtaining a bound:

$$B\prod_{i=1}^{2}\left(2M_{V_i^1}\, 2^{D_{V_i^1}(n_{\lambda_{V_i^1}}-n_{V_i^1}+3)}\right)$$

on each of the subfamilies of $\mathcal{F}$ that we consider in this way (each consisting of $\leq 2M_{V_1^1}\, 2M_{V_2^1}$ trees).

And we continue until all the N_1 innermost resonances have been considered.

Then we consider a second generation resonance V_1^2. As we perform the resummation related to the $2M_{V_1^2}$ terms associated with the new resonance, we regard the sum of the values of each of the groups of trees as a function of $\varepsilon_1^2 = \vec{\omega}\cdot\vec{\nu}_{\lambda_{V_1^2}}$ (also the values $\varepsilon_{j_1}^1$'s corresponding to the innermost resonances contained in V_1^2 are regarded as dependent on ε_1^2): for all the values of ε_1^2, with $|\varepsilon_1^2| < \gamma_{n_{V_1^2}-3}$, such a sum is analytic if $n_{V_1^1} > n_{\lambda_{V_1^1}} + 3$ for each first generation resonance, and is bounded, *in every case* by:

$$B_1 \equiv B\prod_{i=1}^{N_1}\left(2M_{V_i^1}\, 2^{D_{V_i^1}(n_{\lambda_{V_i^1}}-n_{V_i^1}+3)}\right)$$

as it can be argued analogously to the previous discussion. The further sum over the values of the $\leq 2M_{V_1^2}$ elements involved in the new resummation creates a function of ε_1^2 with a second order zero so that we can improve the bound of such a larger collection of trees by:

$$B_1\, 2M_{V_1^2}\, 2^{D_{V_i^2}(n_{\lambda_{V_i^2}}-n_{V_i^2}+3)}$$

and we can continue in this way until the second generation of resonances is exhausted, and so on until no resonances are left, and there is only a big group of terms collected in the successive resummations (containing all the values of the trees in $\mathcal{F}$) and the bound (A1.4) is consequently obtained.

References

[A1] Arnold, V.: *Proof of A.N. Kolmogorov's theorem on the preservation of quasi-periodic motions under small perturbations of the hamiltonian*, Usp. Mat. Nauk, **18**, 5, 13-40, 1963; english traslation: Russ. Math. Surv. **18**, 5, 9-36, 1963.

[A2] Arnold, V.: *Instability of dynamical systems with several degrees of freedom*, Sov. Mathematical Dokl., **5**, 581-585, 1966.

[16] When the $\varepsilon_{j_i}^i$'s are considered as variables defined in a larger analyticity domain, the scale labels can change but no more than one unity; this follows from the analysis in Appendix A2, as can be easily checked.

[BG] Benettin, G., Gallavotti, G.: *Stability of motions near resonances in quasi-integrable hamiltonian systems*, J. Statistical Physics, **44**, 293-338, 1986.

[BeG] Benfatto, G., Gallavotti, G.: *Perturbation theory of the Fermi surface in a quantum liquid. A general quasi particle formalism and one dimensional systems*, Journal of Statistical Physics, **59**, 541-664, 1990.

[CF] Chierchia, L., Falcolini, C.: *A direct proof of the theorem by Kolmogorov in hamiltonian systems*, preprint, archived in `mp_arc@math.utexas.edu`, #93-171.

[CG] Chierchia, L., Gallavotti, G.: *Drift and diffusion in phase space*, in print in Annales de l'Institut H. Poincarè, archived in `mp_arc@math.utexas.edu`, #92-92.

[E] Eliasson L. H.: *Absolutely convergent series expansions for quasi–periodic motions*, report 2–88, Dept. of Math., University of Stockholm, 1988.

[FG] Felder, G., Gallavotti, G.: *Perturbation theory and non renormalizable scalar fields*, Communications in Mathematical Physics, **102**, 549-571, 1986.

[G1] Gallavotti, G.: *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbative series of certain completely integrable systems. A review.*, in print in Reviews in Mathematical Physics, archived in `mp_arc@math.utexas.edu`, #93-164.

[G2] Gallavotti, G.: *Twistless KAM tori*, in print in Communications in Mathematical Physics, archived in `mp_arc@math.utexas.edu`, #93-172.

[G3] Gallavotti, G.: *The elements of Mechanics*, Springer, 1983.

[G4] Gallavotti, G.: *Renormalization theory and ultraviolet stability for scalar fields via renormalization group methods*, Reviews in Modern Physics, **57**, 471-572, 1985. See also, Gallavotti, G.: *Quasi integrable mechanical systems*, Les Houches, XLIII (1984), vol. II, p. 539-624, Ed. K. Osterwalder, R. Stora, North Holland, 1986.

[G5] Gallavotti, G.: *Invariant tori: a field theoretic point of view on Eliasson's work*, talk at the meeting in honor of the 60-th birthday of G. Dell'Antonio (Capri, may 1993), Marseille, CNRS-CPT, preprint.

[GGe] Gallavotti, G., Gentile, G.: *Non recursive proof of the KAM theorem*, in print in Ergodic Theory and Dynamical Systems, archived in `mp_arc@math.utexas.` edu, #93-229.

[Ge] Gentile, G.: *A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable hamiltonian systems*, submitted to Forum Mathematicum, archived in `mp_arc@math.utexas.edu`, #94-33.

[Gr] Graff, S.M.: *On the conservation for hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations, **15**, 1-69, 1974.

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

[K] Kolmogorov, A.N.: *On conservation of conditionally periodic motions under small perturbations of the Hamiltonian*, Dokl. Akad. Nauk SSSR, **98**, 4, 527-530, 1954. See also: Benettin, G., Galgani, L., Giorgilli, A., Strelcyn, J.M.: *A proof of Kolmogorov theorem on invariant tori using canonical transformations defined by the Lie method*, Nuovo Cimento, **79**, B, 201-223, 1984.

[LW] de la Llave, R., Wayne, E.: *Whiskered and low dimensional tori in nearly integrable dynamical systems*, preprint.

[Me] Melnikov, V.K.: *On the stability of the center for periodic perturbations*, Trans. Moscow Math. Soc., **12**, 1-57, 1963.

[Mo1] Moser, J.: *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Gött., II. Math.-Phys. Kl. 1962, 1-20, 1962.

[Mo2] Moser, J.: *Convergent series expansions for quasi-periodic motions*, Matematische Annalen, **169**, 136-176, 1967

[P] Pöschel, J.: *Invariant manifolds of complex analytic mappings*, Les Houches, XLIII (1984), vol. II, p. 949-964, Ed. K. Osterwalder, R. Stora, North Holland, 1986.

[S] Siegel, K.: *Iterations of analytic functions*, Annals of Mathematics, **43**, 607-612, 1943.

[T] Thirring, W.: *Course in Mathematical Physics*, vol. 1, p. 133, Springer, Wien, 1983.