A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable hamiltonian systems *

Guido Gentile[1]

Abstract: *Rotators interacting with a pendulum via small, velocity independent, potentials are considered: the invariant tori with diophantine rotation numbers are unstable and have stable and unstable manifolds ("whiskers"), whose intersections define a set of homoclinic points. The homoclinic splitting can be introduced as a measure of the splitting of the stable and unstable manifolds near to any homoclinic point. In a previous paper, [G1], cancellation mechanisms in the perturbative series of the homoclinic splitting have been investigated. This led to the result that, under suitable conditions, if the frequencies of the quasi periodic motion on the tori are large, the homoclinic splitting is smaller than any power in the frequency of the forcing ("quasi flat homoclinic intersections"). In the case $l = 2$ the result was uniform in the twist size: for $l > 2$ the discussion relied on a recursive proof, of KAM type, of the whiskers existence, (so loosing the uniformity in the twist size). Here we extend the non recursive proof of existence of whiskered tori to the more than two dimensional cases, by developing some ideas illustrated in the quoted reference.*

Key words: *KAM, homoclinic points, cancellations, perturbation theory, classical mechanics, renormalization*

1. Introduction

The existence of whiskered tori is known from the works of Melnikov, [Me], Moser, [Mo], Graff [Gr]; a general theory can be found in [LW]. In this paper we discuss the existence of whiskered tori in a special class of almost integrable hamiltonian systems. As in [G1], we consider a model consisting of a family of rotators, say $l-1$ in number, interacting with a pendulum via a conservative force (the model can be called, as in [G1], *rotator–pendulum model*, or *simple resonance model*, or *Arnold model*).[2] 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. The model can be described by the l degrees of freedom hamiltonian $H_\mu \equiv H_0 + \mu f$ given by

$$\vec\omega \cdot \vec A + \frac{1}{2} J^{-1} \vec A \cdot \vec A + \frac{I^2}{2J_0} + g^2 J_0 (\cos\varphi - 1) + \mu \sum_{\substack{|\nu| \le N \\ \vec\nu \ne \vec 0}} f_\nu \cos(\vec\alpha \cdot \vec\nu + n\varphi) \tag{1.1}$$

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, and f_ν are fixed constants. We suppose $f_{n,\vec 0} \equiv 0$, for all n: this will be clearly not restrictive. A natural *energy scale* for the model will be $E = J_0 g^2$.

We suppose *a priori* that:

---

* 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 [LW] the generation of whiskered tori is studied for systems whose hamiltonian can be expressed, in terms of action-angle variables $(\vec A, \vec\alpha)$, as $H(\vec A, \vec\alpha) = H_0(\vec A) + \mu f(\vec A, \vec\alpha)$, so that there is no hyperbolicity in the unperturbed problem. Then, under the hypothesis that the non degeneration condition $\|\partial_{A_i} \partial_{A_j} H_0\| \ge c > 0$ is fulfilled, invariant whiskered tori are constructed near perturbed periodic orbits. A case in which the above condition does not hold is studied in [CG], in connection to a celestial mechanics problem (D'Alembert procession).

Hypothesis $\mathbf{H}_1$: *the parameters $\vec{\omega}, \mu$ verify, in general:*

$$\vec{\omega} = \frac{\vec{\omega}_0}{\sqrt{\eta}}, \qquad |\mu| \le b\eta^Q, \qquad \eta \le 1 \tag{1.2}$$

with Q and b^{-1} which will be restricted to be large enough in the course of the analysis.

and:

Hypothesis $\mathbf{H}_2$: *$\vec{\omega}_0$ is a diophantine vector, i.e. :*

$$C_0 |\vec{\omega}_0 \cdot \vec{\nu}| \ge |\vec{\nu}|^{-\tau} \ , \qquad \textit{for all } \vec{0} \ne \vec{\nu} \in \mathbf{Z}^{l-1} \tag{1.3}$$

for some diophantine constant C_0 and some diophantine exponent $\tau > 0$.

The $l = 2$ and $J = +\infty$ case will *not* be excluded and corresponds to the "pendulum in a periodic force field". 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:

$$\mathcal{T}_0 \equiv \{I = 0 = \varphi\} \times \{\vec{A} \equiv \vec{A}^0 \ , \vec{\alpha} \in \mathbf{T}^{l-1}\} \tag{1.4}$$

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

$$W_0^{\pm} \equiv W_0 \equiv \{\tfrac{I^2}{2J_0} + g^2 J_0 (\cos\varphi - 1) = 0\} \times \{\vec{A} \equiv \vec{A}^0 \ , \vec{\alpha} \in \mathbf{T}^{l-1}\} \tag{1.5}$$

Then, it follows "from KAM theory", [Me],[E],[CG], that "many" unperturbed tori around the torus $\vec{A}^0 = \vec{0}$ (including the one $\vec{A}^0 = \vec{0}$ itself) can be *continued analytically* (in μ), togheter with their whiskers, into invariant tori with the same $\vec{\omega}$, for all $|\mu| < b\eta^Q$ (if b is a suitable constant, explicitly computable in terms of a few parameters associated with H_0, f in (1.1), see for instance[CG]) and for Q large enough; we shall call such tori *persistent*. The determination of b, Q requires going through an analysis very similar to that of the classical KAM theorem: hence we say that such tori and whiskers are "obtained by KAM analytic continuation".

We shall denote by $W_\mu^{\pm}$ the stable and unstable whiskers, and $\mathcal{T}_\mu$ the whiskered tori obtained by the KAM analytic continuation (in μ). The stable and unstable whiskers $W_\mu^{\pm}$ are characterized by the sets of initial data $X_\mu^{\pm}$ such that the distance$(S_\mu^t X_\mu^{\pm}, \mathcal{T}_\mu) \to 0$ exponentially fast as $t \to \pm\infty$; here S_μ^t is the hamiltonian flow generated by (1.1). The flow on the persistent whiskered tori can be described, in suitable coordinates and for all $|\mu|$ small (*i.e.* $|\mu| \le b\eta^Q$), by $\vec{\psi} \to \vec{\psi} + \vec{\omega} t$.

A legacy of the unperturbed situation (see (1.5)) is that the persistent whiskers $W_\mu^{\pm}$ arising from (1.5) are, for $|\mu|$ small enough, *graphs over the angles $\vec{\alpha} \in \mathbf{T}^{l-1}$* and φ, at least if $|\varphi| < 2\pi - \delta$ for any prefixed $\delta > 0$; hence they can be written as:

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

for suitable real–analytic (in $(\vec{\alpha}, \varphi)$ *and* μ) functions $\vec{A}^{\pm}, I^{\pm}$. In this context, it is natural to *measure the splitting between W_μ^+ and W_μ^-* at $\varphi = \pi$ and $\vec{\alpha} = \vec{0}$ by the quantity: $\delta(\vec{\alpha}) \equiv \det \partial_{\vec{\alpha}} (\vec{A}_\mu^+ - \vec{A}_\mu^-)|_{\varphi=\pi}$ at $\vec{\alpha} = 0$, and its $\vec{\alpha}$-derivatives at $\vec{\alpha} = \vec{0}$.

In [CG] an algorithm for the computation of the μ-expansion coefficients of the functions $\vec{A}^{\pm}$, $\vec{I}^{\pm}$ is introduced, and in [G1] it is used in deriving the result that the homoclinic splitting is smaller than any power in η, as $\eta \to 0$.[3] Such a result is obtained by checking several cancellation mechanisms, operating to all orders of perturbation theory. However the problem is solved selfconstistently only in the case $l = 2$, the solution of the cases $l > 2$ relying on results inherited from the KAM theory approach of [CG]: in such cases one looses the uniformity in the twist size, defined as $t_w = \min_{j=0,\ldots,l-1} J_j^{-1}$, see [G2]. It would, therefore, be nice to have a proof completely freed from KAM-type results. In [G1] the conjecture that this can be done is advanced (and motivated): in this paper we extend the selfconsistent treatment to the more general

[3] In [CF] the persistence of quasiperiodic solutions for nearly integrable hamiltonian systems, described by hamiltonians of the form $\frac{1}{2}\vec{A} \cdot \vec{A} - \mu f(\vec{\alpha})$, is proven with similar tools.

case $l\ge 2$, by using some extra cancellations which can be seen as an extension of those exposed in [G1], [G2], [GG], so obtaining a theory fully independent on KAM-type results. To be more precise, we prove the existence of the whiskered tori in a selfconsistent way, and assuring the uniformity in the twist size. The steps through which the proof proceeds are the following: 1) starting from the unpertubed motion on the separatrix, one perturbatively finds the equations of the motion on two l-dimension manifolds, one stable and the other unstable, expressed by a formal series expansion in the perturbation parameter; 2) under the hypothesis that the series converges, the motions become asymptotic to motions on $(l-1)$-dimension invariant tori; 3) one checks the series convergence.

The paper is selfcontained: §2÷§5 have a definitory nature, however, and they are almost literally taken from the review article [G1], with some abstraction effort, while the original work is in §6÷§8 (and in the appendices) and, as it has been said, it develops the ideas of [G1], [G2], [GG]. The above illustrated steps are approached in §4 and §8, where they receive a more mathematical statement.

Propositions 4.1 and 4.2 at the end of §4 provide formal statements of the above results. In §6 besides quoting our key estimate (the first of (6.2)) we briefly discuss the connection of this work with the theory of the homoclinic splitting.

Acknowledgements: I am indebted to G. Gallavotti for having originally proposed this work, and for encouragement and many clarifying discussions and suggestions all along during its draft.

2. Recursive formulae

In this section 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).

The unperturbed motion is simply:

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

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

$$X^{\sigma}_{\mu}(t)\equiv X^{\sigma}_{\mu}(t;\vec{\alpha})\equiv\sum_{k\ge0}X^{k\sigma}(t;\vec{\alpha})\mu^k=\sum_{k\ge0}X^{k\sigma}(t)\mu^k,\qquad \sigma=\pm \tag{2.2}$$

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 $X^{k\sigma}(t)$, as forming a single function $X^k(t)$, which is $X^{k+}(t)$ if $\sigma=\operatorname{sign}t=+$, and $X^{k-}(t)$ if $\sigma=\operatorname{sign}t=-$.

Inserting (2.2) into the Hamilton equation associated with (1.1), we see that the coefficients $X^{k\sigma}(t)$ satisfy the hierarchy of equations:

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

where:

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

and where $F^{k\sigma}$ depends upon $X^0,...,X^{k-1\sigma}$ but not on $X^{k\sigma}$; 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 $\vec{0}$'s in the first and third row are $(l-1)$ (row) vectors, the $\vec{0}$'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.

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

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

($i.e.$ 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:

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

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

$$F_-^{k\sigma} \equiv 0 \ , \qquad \vec{F}_\downarrow^{k\sigma} \equiv \vec{0} \ , \qquad \vec{F}_\uparrow^{k\sigma} = -\sum_{|\vec{m}|\geq 0} (\partial_{\vec{\alpha}} f)_{\vec{m}}(\varphi^0, \vec{\alpha}+\vec{\omega}t) \sum_{(k_j^i)_{\vec{m},k-1}} \prod_{i=0}^{l-1}\prod_{j=1}^{m_i} X_i^{k_j^i\sigma} \tag{2.7}$$

$$F_+^{k\sigma} \equiv \sum_{|\vec{m}|\geq 2} (g^2 J_0 \sin\varphi)_{\vec{m}}(\varphi^0) \sum_{(k_j)_{\vec{m},k}} \prod_{j=1}^{m} X_-^{k_j\sigma} - \sum_{|\vec{m}|\geq 0} (\partial_\varphi f)_{\vec{m}}(\varphi^0, \vec{\alpha}+\vec{\omega}t) \sum_{(k_j^i)_{\vec{m},k-1}} \prod_{i=0}^{l-1}\prod_{j=1}^{m_i} X_i^{k_j^i\sigma}$$

where $(G)_{\vec{m}}(\cdot)$, with $G = \partial_{\vec{\alpha}} f, \partial_\varphi f, g^2 J_0 \sin\varphi$, and $(k_j^i)_{\vec{m},p}$, with $k_j^i \geq 1$, $m_i \geq 0$, $p = k, k-1$, are defined as:

$$\begin{aligned} (G)_{\vec{m}}(\cdot) &\equiv \Big(\frac{\partial_\varphi^{m_0}\partial_{\alpha_1}^{m_1}\ldots\partial_{\alpha_{l-1}}^{m_{l-1}}\partial_I^{m_l}\partial_{A_1}^{m_{l+1}}\ldots\partial_{A_{l-1}}^{m_{2l-1}} G}{m_0!\, m_1!\, \ldots\, m_{l-1}!\, m_l!\, m_{l+1}!\, \ldots\, m_{2l-1}!}\Big)(\cdot) \\ (k_j^i)_{\vec{m},p} &\equiv (k_1^0, \ldots, k_{m_0}^0, k_1^1, \ldots, k_{m_1}^1, \ldots, k_1^{2l-1}, \ldots, k_{m_{2l-1}}^{2l-1}) \qquad \text{s.t.} \sum k_j^i = p \end{aligned} \tag{2.8}$$

Note that the first sum in the expression for $\vec{F}_+^h$ can only involve vectors $\vec{m}$ with $m_j = 0$ if $j \geq 1$, because the function $J_0 g^2 \cos\varphi$ depends only on φ and not on $\vec{\alpha}$, (hence also $k_j^i = 0$ if $i > 0$). We use here the above notation to uniformize the notations.

The evolution of X^k is determined by integrating (2.6), if the initial data are known. The $k = 1$ case requires a suitable interpretation of the symbols, in according to equation (2.4).

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 $\varphi = \pi, I = 2gJ_0$:

$$W(t) = \begin{pmatrix} \frac{1}{\cosh gt} & \frac{\bar{w}}{4J_0 g} \\ -J_0 g \frac{\sinh gt}{\cosh^2 gt} & (1 - \frac{\bar{w}}{4}\frac{\sinh gt}{\cosh^2 gt})\cosh gt \end{pmatrix}, \qquad \bar{w} \equiv \frac{2gt + \sinh 2gt}{\cosh gt} \tag{2.9}$$

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

$$\begin{pmatrix} X_-^{k\sigma} \\ X_+^{k\sigma} \end{pmatrix} = W(t)\begin{pmatrix} 0 \\ X_+^{k\sigma}(0) \end{pmatrix} + W(t)\int_0^t W^{-1}(\tau)\begin{pmatrix} 0 \\ F_+^{k\sigma}(\tau) \end{pmatrix} d\tau \tag{2.10}$$

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

$$\begin{aligned} X_+^{k\sigma}(t) &= w_{ll}(t)X_+^{k\sigma}(0) + w_{ll}(t)\int_0^t w_{00}(\tau)F_+^{k\sigma}(\tau)d\tau - w_{l0}(t)\int_0^t w_{0l}(\tau)F_+^{k\sigma}(\tau)\,d\tau \\ X_-^{k\sigma}(t) &= w_{0l}(t)X_+^{k\sigma}(0) + w_{0l}(t)\int_0^t w_{00}(\tau)F_+^{k\sigma}(\tau)d\tau - w_{00}(t)\int_0^t w_{0l}(\tau)F_+^{k\sigma}(\tau)\,d\tau \end{aligned} \tag{2.11}$$

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

$$\begin{aligned} \vec{X}_\uparrow^{k\sigma}(t) =& \vec{X}_\uparrow^{k\sigma}(0) + \int_0^t \vec{F}_\uparrow^{k\sigma}(\tau)d\tau \\ \vec{X}_\downarrow^{k\sigma}(t) =& J^{-1}\Big(t\vec{X}_\uparrow^{k\sigma}(0) + \int_0^t d\tau\,(t-\tau)\vec{F}_\uparrow^{k\sigma}(\tau)\Big) \end{aligned} \tag{2.12}$$

having used that the $\vec{X}_{\downarrow}^{k\sigma}(0) \equiv \vec{0}$ 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 $k \geq 1$ of the perturbation expansion.

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

We introduce some integrations operations that can be performed on the functions introduced in §2. The operation is simply the integration over t from $\sigma\infty$ to t, $\sigma = \operatorname{sign} t$. In general such 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 $X^k(t)$, 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 3.1: *Let $\hat{\mathcal{M}}$ be the space of the functions of t which can be represented, for some $k \geq 0$, as:*

$$M(t) = \sum_{j=0}^{k} \frac{(\sigma t g)^j}{j!} M_j^\sigma(x, \vec{\omega} t) \ , \quad x \equiv e^{-\sigma g t} \ , \quad \sigma = \operatorname{sign} t \tag{3.1}$$

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) $M_k^\sigma \neq 0$. The number k will be called the t-degree of M. The smallest cone containing the singularities will be called the singularity cone *of M.*

Definition 3.2: *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 3.3: *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 3.4: *Let $\hat{\mathcal{M}}^k, \hat{\mathcal{M}}_0^k, \mathcal{M}^k, \mathcal{M}_0^k$ denote the subspaces of $\hat{\mathcal{M}}, \hat{\mathcal{M}}_0, \mathcal{M}, \mathcal{M}_0$, respectively, containing the functions of t-degree $\leq k$.*

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 (3.1), 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:

$$\sigma^\chi \frac{(\sigma t g)^j}{j!} x^h e^{i\vec{\omega}\cdot\vec{\nu} t} \tag{3.2}$$

where $\chi = 0, 1$ (*i.e.* the (3.2) 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 $\{\sigma \operatorname{Re} t > 0, |\operatorname{Im} gt| < \pi/2 - d \equiv \xi\}$). 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) 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 3.5: *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:*

$$\mathcal{I}_R M(t) \equiv \int_{\sigma\infty + i\theta}^{t} e^{-Rg\sigma z} M(z)\, dz \tag{3.3}$$

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 (3.3) 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:

$$\mathcal{I} M(t) \equiv \oint \frac{dR}{2\pi i R} \mathcal{I}_R M(t) \tag{3.4}$$

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 (3.2) 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, $\mathcal{I} : \hat{\mathcal{M}}^k \to \hat{\mathcal{M}}^{k+1}$; 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:

$$\mathop{{\int\!\!\!\!\!\!-}}\nolimits_{(\sigma)}^{t} M(\tau) d\tau \equiv \mathcal{I} M(t) \;, \qquad M \in \hat{\mathcal{M}},\ \sigma = \mathrm{sign}\, \mathrm{Re}\, t \tag{3.5}$$

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:

$$\mathop{{\int\!\!\!\!\!\!-}}\nolimits_{\sigma\infty}^{t} M(\tau) d\tau \equiv \mathcal{I} M(0^\sigma) + \int_0^t M(\tau) d\tau \quad . \tag{3.6}$$

4. Analytic expressions of the expansion coefficients for the whiskers

We will show that the X^k'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 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 $X^h \in \mathcal{M}^{2h-1}$, $h < k$, and $F^h \in \mathcal{M}^{2(h-1)}$, $\vec F^h_\uparrow \in \mathcal{M}_0^{2(h-1)}$, $h \leq k$, and, furthermore, that the singularity cone consists of just the imaginary axis (*i.e.* the singularities of the functions defining X^k, F^k are on the segments on the imaginary axis $(-i\infty, -i]$ and $[+i, +i\infty)$).

This means, in particular, that F^h, X^h can be represented as:

$$\begin{aligned} F^h(x,\vec\psi,t) &= \sum_{j=0}^{2(h-1)} \frac{(\sigma t g)^j}{j!} F_j^{h\sigma}(x,\vec\psi), \qquad h=1,\ldots,k \\ X^h(x,\vec\psi,t) &= \sum_{j=0}^{2h-1} \frac{(\sigma t g)^j}{j!} X_j^{h\sigma}(x,\vec\psi), \qquad h=1,\ldots,k-1 \end{aligned} \tag{4.1}$$

by setting $\vec\psi = \vec\omega t$, $\sigma = \operatorname{sign} t$, $x = e^{-g\sigma t}$, with $F_j^{k\sigma}, X_j^{k\sigma}$ 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 F_j^h, X_j^h 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 $F^h(\vec\psi,\sigma\infty)$ dropping the subscript 0 as there is no ambiguity.

Furthermore the inductive hypothesis is enriched by:

$$\vec F^{h\sigma}_{\uparrow\vec 0}(\sigma\infty) = \vec 0, \qquad \text{for all } h\le k \tag{4.2}$$

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: $X^{h\sigma}_{j\vec\nu}(t)$ and $F^{h\sigma}_{j\vec\nu}(t)$ are the Fourier transforms in $\vec\psi$ of $X^{h\sigma}_j(t,\vec\psi)$ and $F^{h\sigma}_j(t,\vec\psi)$, respectively.

Let us suppose, just as an assumption for the time being, that $X^{h\sigma}(t)$ and, from (2.7), hence also $F^{h\sigma}(t)$ are bounded as $t\to\sigma\infty$ for all h, so that $X^{h\sigma}_j(0,\vec\psi) = 0$ if $j\ge 1$: we show then that the latter information is very strong and permits us to determine X^k. This does not imply the convergence of the series: however in §8 we prove such a result, so justifying the boundedness hypothesis and completing the research of bounded motions.

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

To find expressions for $X^k_\uparrow$ we start from the equations (2.6) with initial time at some instant T. And we use that $\mathcal{I}F(t)$ is a primitive of the function $F(t)$, see comment preceding (3.5), so that:

$$\vec X^{k\sigma}_\uparrow(t) = \vec X^{k\sigma}_\uparrow(T) + \mathcal{I}\vec F^{k\sigma}_\uparrow(t) - \mathcal{I}\vec F^{k\sigma}_\uparrow(T) \tag{4.3}$$

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

The function $\vec X^{k\sigma}_\uparrow(T)$ tends to become quasi periodic with exponential speed as $T\to\sigma\infty$: in fact it becomes asymptotic to the $j=0$ component, see (4.1), at $x=0$: $\vec X^{k\sigma}_{0\uparrow}(0,\vec\omega T)$, (in the sense that the difference tends to 0, bounded proportionally to $(g|T|)^{2k-1}e^{-g|T|}$). The function $\mathcal{I}\vec F^{k\sigma}_\uparrow(T)$ also becomes asymptotically quasi periodic with exponential speed *and* $\vec 0$ *average*, because $\vec F^{k\sigma}_\uparrow \in \mathcal{M}^{2(k-1)}_0$ and by the definition of $\mathcal{I}$: therefore the two quasi periodic functions of T must cancel modulo a constant equal to $\langle \vec X^{k\sigma}_{0\uparrow}(0,\cdot)\rangle \equiv \vec X^{k\sigma}_{\uparrow\vec 0}(\sigma\infty)$.

Hence it follows that:

$$\vec X^{k\sigma}_\uparrow(t) = \vec X^{k\sigma}_{\uparrow\vec 0}(\sigma\infty) + \mathcal{I}\vec F^{k\sigma}_\uparrow(t) \tag{4.4}$$

and, by inserting (4.4) into the second of (2.12), (considering also that $\int_0^t \tau \vec F^{k\sigma}_\uparrow(\tau)\,d\tau = t\mathcal{I}\vec F^{k\sigma}_\uparrow(t) +$ a t-bounded function), we see that the $\vec X^{k\sigma}_\downarrow(t)$ can be bounded only if:

$$\vec X^{k\sigma}_{\uparrow\vec 0}(\sigma\infty) = \vec 0, \qquad \text{hence:} \qquad \vec X^{k\sigma}_\uparrow(t) = \mathcal{I}\vec F^{k\sigma}_\uparrow(t) \tag{4.5}$$

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

$$\vec X^{k\sigma}_\downarrow(t) = J^{-1}\big(\mathcal{I}^2\vec F^{k\sigma}_\uparrow(t) - \mathcal{I}^2\vec F^{k\sigma}_\uparrow(0^\sigma)\big) \equiv J^{-1}\overline{\mathcal{I}}^2\vec F^{k\sigma}_\uparrow(t) \tag{4.6}$$

which gives a simple form to the time evolution of the $\vec\alpha$ (*i.e.* $\downarrow$) component of X^k in terms of the operator $\overline{\mathcal{I}}^2$ defined by the r.h.s. of (4.6).

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

$$X_{+}^{k\sigma}(0) = -\fint_{0}^{\sigma\infty} w_{00}(\tau) F_{+}^{k\sigma}(\tau)\, d\tau \tag{4.7}$$

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

$$\begin{aligned} X_{+}^{k\sigma}(t) &= w_{ll}(t) \fint_{(\sigma)}^{t} w_{00}(\tau) F_{+}^{k\sigma}(\tau) d\tau - w_{l0}(t) \int_0^t w_{0l}(\tau) F_{+}^{k\sigma}(\tau) d\tau \equiv \mathcal{O}_+ F_{+}^{k\sigma}(t) \\ X_{-}^{k\sigma}(t) &= w_{0l}(t) \fint_{(\sigma)}^{t} w_{00}(\tau) F_{+}^{k\sigma}(\tau) d\tau - w_{00}(t) \int_0^t w_{0l}(\tau) F_{+}^{k\sigma}(\tau) d\tau \equiv \mathcal{O} F_{+}^{k\sigma}(t) \end{aligned} \tag{4.8}$$

The (4.5), equ(4.6), (4.8) and the boundedness request imply (4.1) for $h = k + 1$, as we can show by reasoning as in [G1]. As already remarked before (4.3) we note again that, since $F_{\uparrow\vec{0}}^{h\sigma}(\sigma\infty) = \vec{0}$ for $h \le k$, the $\vec{X}_{\uparrow}^{k}, \vec{X}_{\downarrow}^{k}$ functions are in fact in $\mathcal{M}^{2(k-1)}$, as the $\mathcal{I}$ operation, on such $\vec{F}_{\uparrow}^{k}$ functions, does not increase the degree. Also, if one looks carefully at the $X_{\pm}^{h\sigma}$–evaluation in terms of $F_{+}^{h\sigma}$, 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 (4.1) is proved for X^k, and it remains to check it for F^{k+1}. This follows from the expression of F^{k+1}, see (2.7), in terms of the X^h with $h \le k$: 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 X's contains at least two factors (which therefore have order labels smaller than k and verify the inductive hypothesis); and, furthermore, the coefficients $(\partial_{\vec{\alpha}} f)_{\vec{m}}(\varphi_0, \vec{\omega} t)$ or $g^2 J_0 \sin\varphi_0$ or $g^2 J_0 \cos\varphi_0$ do not contain any terms that can possibly increase the degree. Hence $F^{k+1} \in \mathcal{M}^{2k}$. To see that $\vec{F}_{\uparrow}^{(k+1)\sigma} \in \mathcal{M}_0^{2k}$, *i.e.* $\vec{F}_{\uparrow\vec{0}}^{(k+1)\sigma}(\sigma\infty) = \vec{0}$, 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:

$$\vec{F}_{\uparrow\vec{0}}^{k\sigma}(\sigma\infty) \equiv \int_{\mathbf{T}^{l-1}} \vec{F}_{\uparrow}^{k\sigma}(\vec{\psi}, \sigma\infty) \frac{d\vec{\psi}}{(2\pi)^{l-1}} \equiv \langle \vec{F}_{\uparrow}^{k\sigma}(\cdot, \sigma\infty) \rangle = \vec{0} \tag{4.9}$$

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

$$\begin{aligned} X_{-}^{h}(t) =& w_{0l}(t) \mathcal{I}(w_{00} F_{+}^{h})(t) - w_{00}(t) \big(\mathcal{I}(w_{0l} F_{+}^{h})(t) - \mathcal{I}(w_{0l} F_{+}^{h})(0^\sigma)\big) \equiv \mathcal{O}(F_{+}^{h})(t) \\ \vec{X}_{\downarrow}^{h}(t) =& J^{-1} \left(\mathcal{I}^2(\vec{F}_{\uparrow}^{h})(t) - \mathcal{I}^2(\vec{F}_{\uparrow}^{h})(0^\sigma)\right) \equiv J^{-1} \overline{\mathcal{I}}^2(\vec{F}_{\uparrow}^{h}(t)) \\ X_{+}^{h}(t) =& w_{ll}(t) \mathcal{I}(w_{00} F_{+}^{h})(t) - w_{l0}(t) \left(\mathcal{I}(w_{0l} F_{+}^{h})(t) - \mathcal{I}(w_{0l} F_{+}^{h})(0^\sigma)\right) \equiv \mathcal{O}_+(F_{+}^{h})(t) \\ \vec{X}_{\uparrow}^{h}(t) =& \mathcal{I}(\vec{F}_{\uparrow}^{h})(t) \end{aligned} \tag{4.10}$$

where $\mathcal{O}, \mathcal{O}_+, \overline{\mathcal{I}}^2, \mathcal{I}$ are defined here and in §3; and $X^h \equiv (X_-, \vec{X}_\downarrow, X_+, \vec{X}_\uparrow) = (X_j^h)$, $j = 0, \ldots 2l - 1$, $F^h = (0, \vec{0}, F_{+}^{h}, \vec{F}_{\uparrow}^{h})$. Note that while X^h has non zero components over both the *angle* ($j = 0, \ldots, l - 1$) components and over the *action* ($j = l, \ldots, 2l - 1$) the F^h has only the action components non zero. *Furthermore the above functions describe a motion on the whisker W^σ with initial data at some $\vec{\alpha}$ and $\varphi = \pi$.*

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

Proposition 4.1: *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 $W_\mu^{\pm}$, (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 scattering and homoclinic phase shifts, see [CG], [G1].

Proposition 4.2: *If $(I, \vec{A}, \varphi, \vec{\alpha}) \in W_\mu^\pm$, i.e. if $(I, \vec{A}, \varphi, \vec{\alpha}) = X_\mu^\sigma$, then the evolution $S_t(I, \vec{A}, \varphi, \vec{\alpha})$ converges to a quasiperiodic motion on the torus $\mathcal{T}$ of Proposition 4.1. And in fact the convergence is exponential in the sense that for $\sigma t \geq 0$:*

$$\left| X^\sigma(x, \vec{\psi} + \vec{\omega} t, t) - X^\sigma(0, \vec{\psi}, \sigma\infty) \right| \leq C e^{-\frac{1}{2} g \sigma t} \tag{4.11}$$

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 §8.

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 the analysis in this paper should also allow us to find the expansion of $g(\mu)$ in a convergent power series in μ: however we do not discuss this further.

5. Tree formalism: part I

In this section we review the graphical formalism developed in [G1], §5, in order to represent, via equations (4.10) and (2.7), the generic h-th order contribution to the homoclinic splitting.

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

$$\begin{aligned} & f^\delta(\varphi, \vec{\alpha}) \equiv \sum_{\nu = (n, \vec{\nu})} \frac{f_\nu^\delta}{2} e^{i(n\varphi + \vec{\nu}\cdot\vec{\alpha})}, \qquad \delta = 0, 1 \\ & f^0(\varphi, \vec{\alpha}) \equiv J_0 g^2 \cos\varphi = \sum_{\substack{\nu,\, \vec{\nu} = \vec{0} \\ n = \pm 1}} \frac{f_\nu^0}{2} e^{in\varphi}, \qquad f^1(\varphi, \vec{\alpha}) \equiv f(\varphi, \vec{\alpha}) = \sum_\nu \frac{f_\nu^1}{2} e^{i(n\varphi + \vec{\nu}\cdot\vec{\alpha})} \end{aligned} \tag{5.1}$$

(the introduction of 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 will 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. 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 $\nu_v \in \mathbf{Z}^l$, $|\nu_v| \leq N$, $j_v = l, \ldots, 2l-1$, $\delta_v = 0, 1$. 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* $j_{\lambda_v} \geq l$, 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 $h = \sum_{v \geq v_0} \delta_v$, *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 $m = 1 + \sum_{v \geq v_0} m_v$, 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 $m_v \geq 2$, it is $h \leq m < 2h$. In order to dispose of a label counting the number of nodes of a subtree, we 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

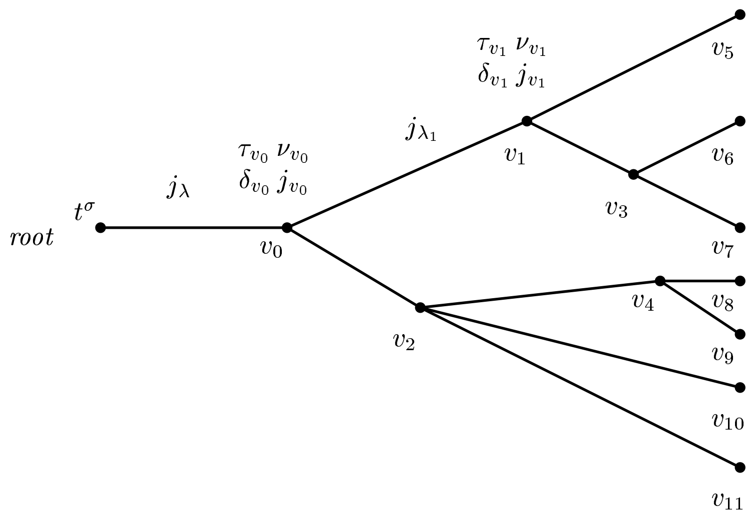

Fig.5.1: A tree θ with $m_{v_0} = 2, m_{v_1} = 2, m_{v_2} = 3, m_{v_3} = 2, m_{v_4} = 2$ and $m = 12$; the root branch label is defined to be $j_\lambda = j$.

be such that all the decorations of the tree match.[4] 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 $\vec\nu(v) = \sum_{w\ge v} \vec\nu_w$, if $\nu_v = (n_v, \vec\nu_v)$ is the *mode label* of v. The *total momentum* is $\vec\nu(v_0) = \sum_{v\ge v_0} \vec\nu_v$; we say also that $\vec\nu_v$ is the momentum "emitted" by the node v.

Then to each node v there corresponds a factor:

$$\frac{1}{2}(-i\nu_v)_{j_v - l} f^{\delta_v}_{\nu_v}\, e^{i(n_v\varphi^0(\tau_v)+(\vec\alpha+\vec\omega\tau_v)\cdot\vec\nu_v)} \prod_{s=0}^{l-1} (i\nu_{vs})^{m_s} \tag{5.2}$$

(the last product is missing if no nodes follow v) which is univoquely determined by the sets of labels attached to v, and to each branch λ we associate an improper integration operation with upper limit t, denoted $\mathcal{O}$, $J^{-1}\overline{\mathcal{I}}^2$, $\mathcal{O}_+$, $\mathcal{I}$ in (4.10), and the branch label will be $j_\lambda = 0$ when representing $\mathcal{O}$, $j_\lambda = 1, \ldots, l-1$ for $J^{-1}\overline{\mathcal{I}}^2$, $j_\lambda = l$ 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 (5.2).
(3) Furthermore out of $(_v$ and next to it we write a symbol $\mathcal{E}^T_v$ which we interpret differently, depending on the label j_{λ_v} on λ_v:

$$\mathcal{E}^T_v\Big(_v \cdot \Big)_v \equiv \begin{cases} \mathcal{O}\Big(_v \cdot \Big)_v(\tau_{v'}) , & \text{if } v \ge v_0 \ , \quad j_{\lambda_v} = 0 \ , \\ J^{-1}\overline{\mathcal{I}}^2\Big(_v \cdot \Big)_v(\tau_{v'}) & \text{if } v \ge v_0 \ , \quad 1 \le j_{\lambda_v} \le l-1 \ , \\ \mathcal{O}_+\Big(_v \cdot \Big)_v(\tau_{v'}) , & \text{if } v = v_0 \ , \quad j_{\lambda_v} = l; \\ \mathcal{I}\Big(_v \cdot \Big)_v(\tau_{v'}) , & \text{if } v = v_0 \ , \quad l+1 \le j_{\lambda_v} \le 2l-1 \end{cases} \tag{5.3}$$

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 (5.3), if $v = v_0$, has to be interpreted as the limit as $t \to 0^\sigma$.

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

Then it follows that $X_j^h(t)$ can be written as:

$$X_j^h(t) = \sum_{\theta\in trees} \frac{1}{m(\theta)!} \sum_{labels;\, \sum_v \delta_v = h} \tilde V_j(t;\theta) \tag{5.4}$$

where $m(\theta)$ = number of branches of θ = number of nodes of θ.

Remark : If we do not perform the operation $\mathcal{E}^T$ 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 $F^h_{j_{v_0}}(t)$. In particular, from (4.10), we deduce that the whiskers splitting $\Delta_j^h(\vec\alpha) = X_j^{h+}(0;\vec\alpha) - X_j^{h-}(0;\vec\alpha)$ is given by:

$$\Delta^h_-(\vec\alpha) \equiv 0 \ , \quad \vec\Delta^h_\downarrow(\vec\alpha) \equiv \vec 0 \ , \quad \Delta^h_+(\vec\alpha) = -\int_{-\infty}^{+\infty} d\tau\ w_{00}(\tau) F^{h\sigma}_+(\tau) \ , \quad \vec\Delta^h_\uparrow(\vec\alpha) = -\int_{-\infty}^{+\infty} d\tau\ \vec F^{h\sigma}_\uparrow(\tau) \tag{5.5}$$

where $F_j^{h\sigma}$ is defined as above prescribed. Note that if $\vec\alpha = \vec 0$ then we are at a homoclinic point, because the hamiltonian (1.1) is even: so that (5.5) is identically vanishing also for the components $j = l, \ldots 2l-1$.

6. Theory of the homoclinic splitting: results

As a consequence of the above analysis and the analysis in [G1], we get that, in general, the angles of homoclinic splitting, (or $\delta(\alpha)$, introduced in §1), are smaller than any power in η. Let us denote $\Delta^h_{\vec\nu}$ 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 splitting $(\mu,\vec\alpha) \to \Delta(\vec\alpha) \equiv X^+_\mu(0;\vec\alpha) - X^-_\mu(0;\vec\alpha)$; then the property of smallness is an immediate consequence of the following bounds.

Let $d \in (0, \frac{\pi}{2})$, and let:

$$\varepsilon_h \equiv \varepsilon_h(d) \equiv \sup_{0<|\vec\nu_0|\le Nh} e^{-|\vec\omega\cdot\vec\nu_0| g^{-1}(\frac{\pi}{2}-d)}, \qquad \beta = 4(N_0+1), \qquad p = 4\tau \tag{6.1}$$

where N_0 is the maximal φ–harmonic of the perturbation f in (1.1). Note that, if $l=2$, it is $\varepsilon_h \equiv \varepsilon_1$. Then, for $j \ge l$ and for all $J \in [J_0, +\infty)$ and $h \ge 1$:

$$|\Delta^h_{j\vec\nu}| \le gJ_0 D B^{h-1} \quad , \qquad |\Delta^h_{j\vec\nu}| \le gJ_0 D d^{-\beta} (B d^{-\beta})^{h-1} (h-1)!^{p+2} \varepsilon_h \tag{6.2}$$

where D and B are suitable dimensionless constants depending on the various parameters describing (1.1), *but not on the perturbation parameters* η,μ. Note also that since we always suppose that f is a trigonometric polynomial of degree N, it is actually $\Delta^h_{j\vec\nu} = 0$ if $|\vec\nu| > Nh$. Both bounds in (6.2) are uniform in $J \ge J_0$ and one can take $J \to +\infty$. The second equation in (6.2) has been proven in [G1], §8 and Appendix A1, by using some cancellation mechanisms operating to all orders in the perturbative series of the homoclinic splitting. To the first one the following section is devoted, as it represents the original result with respect to [CG], [G1].

In this section we confine ourselves to show that, by reasoning as in [G1], the bounds (6.2) imply that the splitting is smaller than any power, so justifying the expression "quasi flat homoclinic interesections".

By (6.2), the splitting can be bounded, for any multiindex $\vec a$, by:

$$|\partial^{\vec a}_{\vec\alpha} \vec\Delta_\uparrow(\vec 0)| \le gJ_0 D \sum_{h=1}^\infty \sum_{0<|\vec\nu|\le Nh} |\mu|^h |\vec\nu|^{|\vec a|} \min\{B^{h-1}, B_h \varepsilon_h(d)\} \tag{6.3}$$

having denoted $B_h = B^{h-1} d^{-\beta h} (h-1)!^{p+2}$. Note that, if N is the trigonometric degree of the polynomial f in (1.1), the sums over $\vec\nu$ can be suppressed by multiplying the h-th term by the mode counting factor $\bar C^h \equiv (2N+1)^{h(l-1)+h|\vec a|}$ (where $\bar C$ is the maximum number of non zero Fourier components times the maximum of $|\vec\nu|^{\vec a}$).

¿From this bound it follows that $|\partial^{\vec a}_{\vec \alpha}\vec\Delta_\uparrow|$ is smaller than any power in η (see (1.2)). In fact we can split the sum over h in (6.3) into a finite sum, $\sum_{1\le h\le h_0}(\cdot)$ and a "remainder", $\sum_{h>h_0}(\cdot)$; then, if η is small enough, and η, Q in (1.2) are such that $b\eta^Q=\mu_0$, $\mu_0^{-1}>B\bar C$, and $|\mu|<\mu_0/2$, we find

$$\sum_{h>h_0}(\cdot)\le \frac{gJ_0D}{B}\sum_{h>h_0}(|\mu|B\bar C)^h\le \frac{2gJ_0D}{B}\Big(\frac{|\mu|}{\mu_0}\Big)^{h_0} \tag{6.4}$$

and:

$$\sum_{h=1}^{h_0}(\cdot)\le gJ_0D\,h_0|\mu|\,\bar C^{h_0}d^{-\beta h_0}B^{h_0-1}(h_0-1)!^{p+2}\varepsilon_{h_0}(d) \tag{6.5}$$

Thus if $\mu=\eta^{Q+s}$, $d=\sqrt{\eta}$, and $s\ge1$ we see that fixing $h_0=r/s$, for any $r>1$, the $|\partial^{\vec a}_{\vec\alpha}\vec\Delta_\uparrow|$ is bounded by a (r-dependent) constant times η^r (as in such a case (6.5) is just a remainder, exponentially small in $\eta^{-1/2}$).

7. Tree formalism: part II

We introduce the dimensionless quantities related to the homoclinic splitting by:

$$\Delta^h_{l\vec\nu}=J_0g\,\bar\Delta^h_{l\vec\nu}\ ,\qquad \Delta^h_{j\vec\nu}=Jg\,\bar\Delta^h_{j\vec\nu}\ ,\quad (l+1\le j\le 2l-1) \tag{7.1}$$

and denote $\Xi^{h\sigma}_j(t)$, $\sigma=\pm$, $0\le j<2l$, the dimensionless quantities corresponding to the perturbed motions $X^{h\sigma}_j(t)$: obviously $\bar\Delta^h_{j\vec\nu}=\Xi^{h+}_{j\vec\nu}(0)-\Xi^{h-}_{j\vec\nu}(0)$, $j\ge l$.

Given a tree θ, with $m(\theta)=m$, we can write its contribution to $\Xi^{h\sigma_{v_0}}_{j\vec\nu}(t)$, $j\ge l$, as:

$$\begin{aligned}\frac{1}{m!}\,\tilde V_j(t;\theta)=&\frac{1}{m!}\prod_{v_0\le v\in\theta}\oint\frac{dR_v}{2\pi iR_v}\sum_{\rho_v=0,1}\int_{\sigma_{v'}\infty}^{\rho_v g\tau_{v'}}d\,g\tau_v\,e^{-\sigma_v gR_v\tau_v}\,w^{\rho_v}_{j_v}(\tau_{v'},\tau_v)\\&\cdot\Big[\frac{(-i\nu_v)_{j_v-l}}{2}\,c_{\nu_v}\,e^{i(n_v\varphi^0(\tau_v)+\vec\nu_v\cdot\vec\omega\tau_v)}\prod_{s=0}^{l-1}(i\nu_{vs})^{m^s_v}\Big]\end{aligned} \tag{7.2}$$

where $\tau'_{v_0}=t$, $j_{v_0}=j$, and we have defined the dimensionless coefficients c_{ν_v} as:

$$c_{\nu_v}\equiv[(J_0g^2)^{-1}\delta_{j_v,l}+(Jg^2)^{-1}\big(1-\delta_{j_v,l}\big)\delta_v]f^{\delta_v}_{\nu_v}\ ,$$

where $\delta_{j_v,l}$ is 1 if $j_v=l$, and 0 otherwise (*i.e.* $j_v=l+1,\ldots,2l-1$), and used (4.10), by setting:

$$\begin{aligned}
w^0_{j_v}(\tau_{v'},\tau_v)&=\begin{cases}w_{00}(\tau_{v'})\bar w_{0l}(\tau_v), & v>v_0\ ,j_v=l\\ g\tau_v, & v>v_0\ ,j_v>l\end{cases}\\
w^0_{j_{v_0}}(t,\tau_{v_0})&=\begin{cases}\bar w_{l0}(t)\bar w_{0l}(\tau_{v_0}), & j_v=l\\ 0, & j_v>l\end{cases}\\
w^1_{j_v}(\tau_{v'},\tau_v)&=\begin{cases}\bar w_{0l}(\tau_{v'})w_{00}(\tau_v)-w_{00}(\tau_{v'})\bar w_{0l}(\tau_v), & v>v_0\ ,j_v=l\\ g(\tau_{v'}-\tau_v), & v>v_0\ ,j_v>l\end{cases}\\
w^1_{j_{v_0}}(t,\tau_{v_0})&=\begin{cases}w_{ll}(t)w_{00}(\tau_{v_0})-\bar w_{l0}(t)\bar w_{0l}(\tau_{v_0}), & j_{v_0}=l\\ 1, & j_{v_0}>l\end{cases}
\end{aligned} \tag{7.3}$$

with the dimensionless matrix elements $\bar w_{0l}$, $\bar w_{l0}$ given, respectively, by $\bar w_{0l}=(J_0g)^{-1}w_{0l}=\bar w/4$, $\bar w_{l0}=-(Jg)^{-1}w_{l0}$, and m is the total number of branches (root branch included) and the integers m^s_v decompose m_v and count the number of branches emerging from v and carrying the labels $s=0,\ldots,l-1$. If $j<l$, then $w^{\rho_{v_0}}_{j_{v_0}}$ is defined as $w^{\rho_v}_{j_v}$, $v>v_0$.

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:

$$\begin{aligned}
w^0_{j_v}(\tau_{v'},\tau_v)=&\frac12\Big\{\frac{g\tau_v}{\cosh g\tau_{v'}\ \cosh g\tau_v}+\frac{\sinh g\tau_v}{\cosh g\tau_{v'}}\Big\}\\
w^1_{j_v}(\tau_{v'},\tau_v)=&\frac12\Big\{\frac{g(\tau_{v'}-\tau_v)}{\cosh g\tau_{v'}\ \cosh g\tau_v}+\frac{\sinh g\tau_{v'}}{\cosh g\tau_v}-\frac{\sinh g\tau_v}{\cosh g\tau_{v'}}\Big\}
\end{aligned} \tag{7.4}$$

Then we can write:

$$
\begin{aligned}
w^0_{j_v}(\tau_{v'},\tau_v)\,e^{in_v\varphi^0(\tau_v)} &= \begin{cases} g\tau_v\, y_v^{(0)}(\tau_{v'},\tau_v)+y_v^{(-1)}(\tau_{v'},\tau_v)\,, & \text{if } j_v=l \\ g\tau_v\, y_v^{(2)}(\tau_v)\,, & \text{if } j_v>l \end{cases} \\
w^1_{j_v}(\tau_{v'},\tau_v)\,e^{in_v\varphi^0(\tau_v)} &= \begin{cases} g(\tau_{v'}-\tau_v)\, y_v^{(0)}(\tau_{v'},\tau_v)+y_v^{(1)}(\tau_{v'},\tau_v)-y_v^{(-1)}(\tau_{v'},\tau_v)\,, & \text{if } j_v=l \\ g(\tau_{v'}-\tau_v)\, y_v^{(2)}(\tau_v)\,, & \text{if } j_v>l \end{cases}
\end{aligned}
\tag{7.5}
$$

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

$$
\begin{aligned}
&y_v^{(-1)}(\tau_{v'},\tau_v)=\frac{1}{2}\frac{\sinh g\tau_v}{\cosh g\tau_{v'}}\;e^{in_v\varphi^0(\tau_v)} &&y_v^{(1)}(\tau_{v'},\tau_v)=\frac{1}{2}\frac{\sinh g\tau_{v'}}{\cosh g\tau_v}\;e^{in_v\varphi^0(\tau_v)}\\
&y_v^{(0)}(\tau_{v'},\tau_v)=\frac{1}{2}\frac{1}{\cosh g\tau_v\cosh g\tau_{v'}}\;e^{in_v\varphi^0(\tau_v)} &&y_v^{(2)}(\tau_{v'},\tau_v)=e^{in_v\varphi^0(\tau_v)}
\end{aligned}
\tag{7.6}
$$

and admit the following Laurent expansion:

$$
\begin{aligned}
&y_v^{(-1)}(\tau_{v'},\tau_v)=\sum_{k'_v=1}^{\infty}\sum_{k_v=-1}^{\infty}y_v^{(-1)}(k'_v,k_v)x_{v'}^{k'_v}x_v^{k_v} &&y_v^{(1)}(\tau_{v'},\tau_v)=\sum_{k'_v=-1}^{\infty}\sum_{k_v=1}^{\infty}y_v^{(1)}(k'_v,k_v)x_{v'}^{k'_v}x_v^{k_v}\\
&y_v^{(0)}(\tau_{v'},\tau_v)=\sum_{k'_v=1}^{\infty}\sum_{k_v=1}^{\infty}y_v^{(0)}(k'_v,k_v)x_{v'}^{k'_v}x_v^{k_v} &&y_v^{(2)}(\tau_v)=\sum_{k_v=0}^{\infty}y_v^{(2)}(0,k_v)x_v^{k_v}
\end{aligned}
\tag{7.7}
$$

with $x_v=\exp[-\sigma_v g\tau_v]$, $\sigma_v=\mathrm{sign}\,\tau_v$, and $x_{v'}=\exp[-\sigma_{v'}g\tau_{v'}]$, $\sigma_{v'}=\mathrm{sign}\,\tau_{v'}$. 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: $y_v^{(-1)}(1,-1)=\sigma_v/2$, $y_v^{(-1)}(1,0)=2in_v$, $y_v^{(-1)}(1,1)=-\sigma_v/2$, $y_v^{(0)}(1,1)=2$, $y_v^{(0)}(1,2)=8in_v\sigma_v$, $y_v^{(1)}(-1,1)=\sigma_{v'}/2$, $y_v^{(1)}(0,1)=0$, $y_v^{(1)}(1,1)=-\sigma_{v'}/2$, $y_v^{(2)}(0,0)=1$, $y_v^{(2)}(0,1)=4in_v\sigma_v$. We define the sets Λ_α, $\alpha=-1,0,1,2$, as: $\Lambda_\alpha=\{v\in\theta\,:\alpha_v=\alpha\}$.

Then, for each tree node, we have four more labels, k_v,k'_v,ρ_v,α_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 $\tilde V_j(t;\theta)$ introduced in §5 can be replaced with a new tree value, $V_j(t;\theta)$, taking into account also the new labels, and (5.4) holds still provided $\tilde V_j(t;\theta)$ is replaced with $V_j(t;\theta)$. The generic contribution $(1/m!)\;V_j(t;\theta)$ to (7.2), corresponding to a given tree θ, with $m(\theta)=m$, is:

$$
\frac{1}{m!}\;V_j(t;\theta)=\frac{1}{m!}\prod_{v_0\le v\in\theta}\oint\frac{dR_v}{2\pi iR_v}\int_{\sigma_{v'}\infty}^{\rho_v g\tau_{v'}}d\,g\tau_v\;\mathcal V_v(\theta)
\tag{7.8}
$$

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

$$
\mathcal V_v(\theta)\equiv F_{\nu_v}\;T_v(g\tau_{v'},g\tau_v)\;e^{-\sigma_v R_v g\tau_v}\;e^{i\omega_v\tau_v}\;x_v^{k_v}\prod_{j=1}^{m_v}x_v^{k'_{v_j}}\;,
\tag{7.9}
$$

$\omega_v=\vec\omega\cdot\vec\nu_v$, 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 is missing if v is a top node), and $T_v(g\tau_{v'},g\tau_v)$ is defined as:

$$
T_v(g\tau_{v'},g\tau_v)=(\delta_{\alpha_v,2}+\delta_{\alpha_v,0})\,[(1-\rho_v)g\tau_v+\rho_v g(\tau_{v'}-\tau_v)]+(\delta_{\alpha_v,-1}+\delta_{\alpha_v,1})
\tag{7.10}
$$

(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 $\rho_v=1$). We have set:

$$
F_{\nu_v}=\frac{(-i\nu_v)_{j_v-l}}{2}\;c_{\nu_v}\;\Big[\prod_{s=0}^{l-1}(i\nu_{vs})^{m_v^s}\Big]\,(-1)^{\delta_{\alpha_v,-1}\delta_{\rho_v,1}}\;y_v^{(\alpha_v)}(k'_v,k_v)\equiv\Phi_{\nu_v}\;(-1)^{\delta_{\alpha_v,-1}\delta_{\rho_v,1}}\;y_v^{(\alpha_v)}(k'_v,k_v)
\tag{7.11}
$$

where the coefficients Φ_{ν_v} satisfy the following bound:

$$\Big|\prod_{v\ge v_0}\Phi_{\nu_v}\Big|\le\Big(\frac{N}{2}F_0N\Big)^m\equiv\mathcal{C}^m \tag{7.12}$$

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:

$$\left|\prod_{v\ge v_0}y_v^{(\alpha_v)}(k'_v,k_v)\right|\le M^{2m}\prod_{v\ge v_0}\lambda^{k_v+k'_v} \tag{7.13}$$

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 $y_v^{(v)}$'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 equal to a constant $M_3=M_2\lambda^{-1}=1$, so that we can set $M=\max\{M_1,M_2\}$.[5]

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.

Now we give three definitions about trees which perhaps do not deserve really a their own name, since they do not correspond to any object admitting a natural interpretation, (expecially the second and third ones), but they will appear in the following discussion, and therefore it will be useful to have a name to label them.

Definition 7.1 : *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 attached to the node v, (in particular with the same h_v), and we sum their values: the so obtained quantity $\bar L_{j_v}^{h_v\sigma_v}(\tau_{v'})$ will be associated to 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 can associate to a reduced tree $\bar\theta$ a value $V_j(t;\bar\theta)$, where, corresponding to each free node v, there is a factor $\mathcal{V}_v(\bar\theta)\equiv\mathcal{V}_v(\theta)$ as in (7.9), and, corresponding to each leaf v, there is factor $\bar L_{j_v}^{h_v\sigma_v}(\tau_{v'})$.*

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 contribution to $\Xi^{h_v\sigma_v}_{j_{\lambda_v}\vec\nu(v)}(\tau_{v'})$, $j_{\lambda_v}=j_v-l$, ($\vec\nu(v)$ is the momentum of the node v, as it is defined in §5), whose dependence on $\tau_{v'}$ reveals itself only through the factor, (see the third line in (7.3)):

$$\xi_v(\tau_{v'})=[w_{00}(\tau_{v'})\delta_{j_v,l}+(1-\delta_{j_v,l})] \tag{7.14}$$,

so that we can write $\bar L_{j_v}^{h_v\sigma_v}(\tau_{v'})=\xi_v(\tau_{v'})\ \bar L_{j_v}^{h_v\sigma_v}(0)$. We define $\bar L_{j_v}^{h_v\sigma_v}(0)$ as the *value of the leaf* v of the reduced tree. Also the factor (7.14) admits a series expansion like the functions $y_v^{(\alpha_v)}$'s in (7.7):

$$\xi_v(\tau_{v'})=\sum_{k'_v=1}^{\infty}\xi_v(k'_v,0)x_{v'}^{k'_v} \tag{7.15}$$

[5] The request that *all* the x satisfy the property $|x|<\lambda$ is not so strong: in the cases 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 Lemma 8.3 below).

We can use explicitly the order of the integration variables, so defining:

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

and writing:

$$\begin{aligned}
\prod_{\bar{\theta}_f \ni v \geq v_0} e^{-k_v g\sigma\tau_v} &= e^{-k(v_0)g\sigma\tau_{v_0}} \cdot \prod_{\bar{\theta}_f \ni v > v_0} e^{-k(v)g\sigma(\tau_v - \tau_{v'})} \\
\prod_{\bar{\theta} \ni v \geq v_0} e^{-k'_v g\sigma\tau_v} &= e^{-[k'(v_0)+k'_{v_0}]g\sigma\tau_{v_0}} \cdot \prod_{\bar{\theta} \ni v > v_0} e^{-k'(v)g\sigma(\tau_v - \tau_{v'})} \\
\prod_{\bar{\theta}_f \ni v \geq v_0} e^{-R_v g\sigma\tau_v} &= e^{-\sum_{w \geq v_0} R_w g\sigma\tau_{v_0}} \cdot \prod_{\bar{\theta}_f \ni v > v_0} e^{-\sum_{w \geq v} R_w g\sigma(\tau_v - \tau_{v'})} \\
\prod_{\bar{\theta}_f \ni v \geq v_0} e^{i\omega_v\tau_v} &= e^{i\omega(v_0)\tau_{v_0}} \cdot \prod_{\bar{\theta}_f \ni v > v_0} e^{i\omega(v)(\tau_v - \tau_{v'})}
\end{aligned} \tag{7.16}$$

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 $\bar{L}_{j_v}^{h_v\sigma_v}(0)$, which is independent on $\tau_{v'}$, times a factor (7.14), which one has to take into account in the computation of $p(\tilde{v})$, for each $\tilde{v} < v$. Note that only the free nodes contribute to $k(v)$ and $\omega(v)$; we can write $\omega(v) = \vec{\omega} \cdot \vec{\nu}_0(v)$, where $\vec{\nu}_0(v)$ is the "free momentum" of the reduced tree. Note also that the leaves with $j_v = l$ are such that, in (7.16), $k'_v \geq 1$, see (7.15), (7.7), while, if $j_v > l$, it is $k'_v = 0$; in both cases we can define k_v to be identically vanishing, so attaching such a label, for convenience, also to the leaves.

Definition 7.2 : *Given a tree θ, we set $\mathcal{L}_{-1} \equiv \{v \in \theta : v \in \Lambda_{-1},\ and\ p(v) = 0\}$. We define the* generalized reduced tree $\bar{\theta}^G$ *in the following way. Let us draw a bubble encircling each node $v > v_0$, $v \notin \mathcal{L}_{-1}$, with $\rho_v = 0$, and the entire subtree emerging from it, and let us delete all the so obtained bubbles, except 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, we consider all the possible trees with the same labels attached to the node v, (in particular with the same h_v), and we sum their values: the so obtained quantity $L_{j_v}^{h_v\sigma_v}(\tau_{v'})$ will be associated to a fat point, replacing the original bubble, which will be called a* leaf *(of the generalized reduced tree). We still call leaves the fat points, and* free nodes *the generalized 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 generalized 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 to the fat point representing it. We can associate to a generalized reduced tree $\bar{\theta}^G$ a value $V_j(t;\bar{\theta}^G)$, where, corresponding to each free node v, there is a factor $\mathcal{V}_v(\bar{\theta}^G) \equiv \mathcal{V}_v(\theta)$ as in (7.9), and, corresponding to each leaf v, there is factor $L_{j_v}^{h_v\sigma_v}(\tau_{v'})$.*

Remark 1 : The Definition 7.1 is only a preliminary definition preluding to Definition 7.2, which is more involved, but a useful one. The generalized reduced trees are different from the reduced trees as to the resummation procedure of the leaves, (for instance, a tree contributing to a generalized reduced tree with $\rho_v = 0$, for one $v \in \mathcal{L}_{-1}$, can be counted also among the trees contributing to the reduced tree in which v is a leaf). So the leaves of the reduced trees are different from the leaves of the generalized reduced trees, (that's why we have used different symbols to label their values). The more natural notion is the first one, since it allows us to order the time variables; but this is not sufficient to prove our result, and so the introduction of the generalized reduced trees is necessary to become aware of some cancellation mechanisms which can be implemented only by considering together the nodes $v \in \theta$ in $\mathcal{L}_{-1}$, with $\rho_v = 0, 1$. This will be explicitly exploited in the proof of Lemma 8.2.
Remark 2 : The reduced degree is so defined that the degree of a generalized 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 3 : Note that, unlike what happened in §5, now only to the free nodes an integration time variable is associated. This could be found a little misleading as to the notion of node, with respect with the usual terminology, (see [G1], [G2], [GG]); 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 of the proof of Lemma 8.1 below).

We remark also that it is still possible write

$$L_{j_v}^{h_v\sigma_v}(\tau_{v'}) = \xi_v(\tau_{v'})\, L_{j_v}^{h_v\sigma_v}(0)\ , \tag{7.17}$$

being $\xi_v(\tau_{v'})$ defined in (7.14). Again we call $L_{j_v}^{h_v\sigma_v}(0)$ the *value of the leaf* v of the generalized reduced tree. Eventually we define the *free momentum* of the generalized reduced tree with first node v_0 as $\vec{\nu}_0(v_0) = \sum_{\bar\theta_f^G \ni w \geq v_0} \vec{\nu}_w$. Note that, if $(1/m!)V_j(t;\bar\theta^G)$ is a contribution to $\Xi_{j\vec{\nu}}^{h\sigma_{v_0}}(t)$, $\vec{\nu} \equiv \vec{\nu}(v_0)$, then it is $\vec{\nu}_0(v_0) \neq \vec{\nu}$, since $\vec{\nu}_0(v_0)$ takes into account only the free nodes of $\bar\theta^G$, while $\vec{\nu}$ depends also on the momentum labels affixed to the leaves.

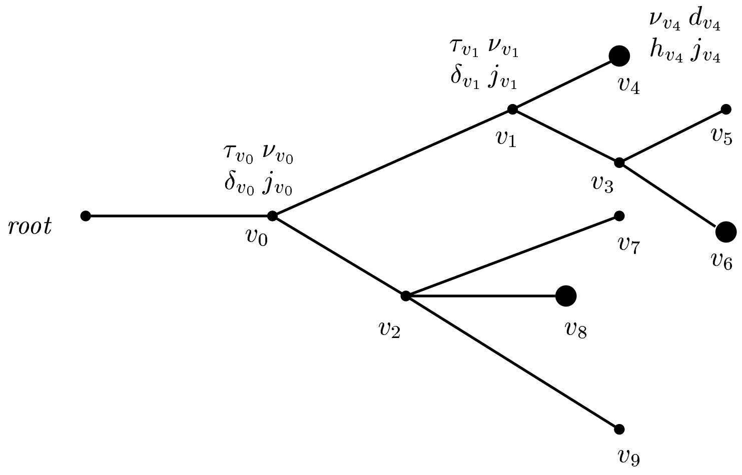

Fig.7.1: A generalized reduced tree $\bar\theta^G$ with $\mathcal{N}_L = 3$ leaves, $m_{v_0} = 2, m_{v_1} = 2, m_{v_2} = 3, m_{v_3} = 2$, and reduced degree $d_{v_0} = 7$; the branch label is defined to be $j_\lambda = j$. Each fat point represents a leaf.

As done in the case of the reduced trees, we can define also for a generalized reduced tree $\bar\theta^G$ the sets $\bar\theta_f^G \equiv \{v \in \bar\theta^G : v \text{ is a free node }\}$ and $\bar\theta_L^G \equiv \{v \in \bar\theta^G : v \text{ is a leaf }\}$, verifying the properties $\bar\theta^G = \bar\theta_f^G \cup \bar\theta_L^G$ and $\bar\theta_f^G \cap \bar\theta_L^G = \emptyset$.

We note that the equalities (7.16) cannot be used for generalized reduced trees, since the time variables are no longer ordered. Nevertheless it is still possible to exploit them partially. In fact, let us consider a generalized reduced tree, and let us single out the nodes v's in $\mathcal{L}_{-1}$: for each such node v we introduce a label $D(v)$, the *depth label*, counting the maximum number of nodes in $\mathcal{L}_{-1}$ we can meet moving forward along any path connecting v to the top nodes. Let us start from the nodes $v^{(0)}$'s in $\mathcal{L}_{-1}$ with $D(v^{(0)}) = 0$: all the following free nodes v's have $\rho_v = 1$, so that their time variables are ordered, and we can use the relations (7.16) from $v^{(0)}$ to the top nodes following it. Then we sum the two contributions with $\rho_{v^{(0)}} = 0$ and $\rho_{v^{(0)}} = 1$, and we obtain a function of $\tau_{v^{(0)\prime}}$. (Note that the sum over such two contributions corresponds to perform an integration from 0 to $\tau_{v^{(0)\prime}}$, instead of two improper integrations, since the functions which we integrate are equal up to the sign, see the $y_v^{(-1)}$ term in (7.4) and (7.6)). As second step, we consider the nodes $v^{(1)}$'s in $\mathcal{L}_{-1}$ with $D(v^{(1)}) = 1$: all the following nodes have $\rho_v = 1$, since the nodes with depth zero have disappeared, (*i.e.* we have integrated already over them), and so the relations (7.16) can be exploited again. Then we sum over the two contributions $\rho_{v^{(1)}} = 0$ and $\rho_{v^{(1)}} = 1$, and we obtain a function of $\tau_{v^{(1)\prime}}$. And so on: we iterate the procedure until the first node of the generalized reduced tree is reached.

The result of the whole procedure will be found inductively when explaining the proof of Lemma 8.2.

Definition 7.3 : *Given a generalized reduced tree $\bar\theta^G$, we define the* stripped value *of the generalized reduced tree $V_j^S(t;\bar\theta^G)$ as the value we obtain by associating to each free node a factor $\mathcal{V}_v(\bar\theta^G) \equiv \mathcal{V}_v(\theta)$ as in (7.9), but retaining for each leaf only the factor $\xi_v(\tau_{v'})$ in (7.17). Note that the discarded contribution of the leaf v is nothing else but its value, $L_{j_v}^{h_v\sigma_v}(0)$, as it is defined after (7.17).*

Remark : The just given definition may appear too involved. Perhaps it is so, but it turns out to be notationally useful, as will become clear along the proof of Lemma 8.1, see in particular (8.7) below. In particular we note that the contribution of a leaf $v \in \bar\theta^G$ to a stripped value $V_j^S(t;\bar\theta^G)$ 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.

8. Analyticity of the homoclinic splitting

It can be useful to elucidate the problems arising in the treatment and to sketch the strategy followed in order to solve them. If all the nodes v had $p(v) \neq 0$, then all the integrals would trivially factorize, (there would be no need to distinguish between reduced trees and generalized reduced trees), 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 $\tau_{v_0}^p$, $p \le 2m$, 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 $n_v \ge 1$, 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 checking that each time a power t^p appears, it comes together with a factor $1/p!$, (and it is $p \le m$, since the case $p(v) = \omega(v) = 0$ is not possible when $T_v(g\tau_{v'}, g\tau_v) \neq 1$, see below); the second is treated in part by exploiting some new cancellations related to the particular structure of the kernels (7.4), which can be very easily visualized in terms of the generalized reduced trees introduced in Definition 7.2. Other cancellation mechanisms will be used in Appendices A1 and A2, and are essentially taken from [G1].

To do explicitly what has been said, it will be necessary to single out the cases in which such problems can really arise. Therefore, in order to study the contributions to $\Xi_{j\vec\nu}^{h\sigma}(t)$, it will turn out to be useful to distinguish between several cases, according to the value of the labels $p(v_0)$ and k_{v_0}. For each considered case we obtain a lemma giving us a convergence result: as a consequence of such lemmata, Theorem 8.1 below will follow.

The idea is the following. 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 (7.3), $p(v)$ can never be "too negative", and, in fact, it is always $p(v) \ge -1$; 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 or to 1, (according to some rules which will appear clearer along the below discussion). If $p(v) = 0$, then, as we shall see, k_v can assume only the values either $k_v = 0$ or $k_v = -1$. If $k_v = 0$, the integrals over the τ_w's, $w \ge v$, can be bounded by using the theory of the twistless KAM tori and the Eliasson's cancellations, (see Lemma 8.2); while, if $k_v = -1$, the integrals over the τ_w's, $w \ge v$, can be inductively studied, by exploiting also the previous result, (see Lemma 8.2).[6] It remains to study the cases $p(v) \neq 0$, but they follow quite easily, if we use the two above results, by explicit calculations, (see Lemma 8.3). 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.

Lemma 8.1: *Let us consider the contribution to $\Xi_{j\vec\nu}^{h\sigma}(t)$, $\vec\nu \in \mathbf{Z}^{l-1}$, $\sigma = \pm$, $j < l$, corresponding to the stripped value $V_j^S(t;\bar\theta^G)$ of a generalized reduced tree $\bar\theta^G$ with labels $p(v_0) = 0$, $k_{v_0} = 0$; 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) If $0 < j \le l-1$, such a sum can be written as:

$$e^{i\vec\omega\cdot\vec\nu_0(v_0)\rho_{v_0}t} \prod_{\bar\theta_f^G \ni v \ge v_0} \Phi_{\nu_v} G_v[\omega(v)] \tag{8.1}$$

where Φ_v is defined in (7.11), $0 < |\vec\nu_0(v_0)| \le m_0 N$, m_0 being the number of free nodes in $\bar\theta^G$, and $G_v[\omega(v)]$ is defined to be:

$$G_v[\omega(v)] = \begin{cases} [ig^{-1}\omega(v)]^{-2} & \text{if } j_v > l \\ [1 + g^{-2}\omega^2(v)]^{-1} & \text{if } j_v = l \end{cases} \tag{8.2}$$

[6] It is important to stress that a subtree with first node v represents a contribution to $\Xi_{j_{\lambda_v}\vec\nu(v)}^{h_v\sigma_v}(\tau_{v'})$, so that it is possible to express $\Xi_{j_{\lambda_{v_0}}\vec\nu(v_0)}^{h_{v_0}\sigma_{v_0}}(t)$ in terms of analogous functions of lower order, with $j_{\lambda_v} < l$. This allows us to look for an inductive proof about the structure of a tree with $p(v_0) = 0$, $k_{v_0} = -1$, since the case in which there is no node $v > v_0$ with $p(v) = 0$, $k_v = -1$, is easy, (if the assertion about the case $p(v) = 0$, $k_v = 0$, is accepted).

with $j_{v_0} > l$.
2) If $j = 0$, *the sum over* $\rho_{v_0} = 0, 1$ *gives:*

$$(-in_{v_0})\frac{gt}{\cosh gt}\,\Phi_{\nu_{v_0}}\Big(\prod_{\bar{\theta}^G_f \ni v > v_0} \Phi_{\nu_v} G_v[\omega(v)]\Big)\int_0^1 ds e^{is\vec{\omega}\cdot\vec{\nu}_0(v_0)t} \tag{8.3}$$

where $|\vec{\nu}_0(v_0)| \le m_0 N$, *and the function* $G[\omega(v)]$ *is defined in (8.2).*
3) The sum over all the generalized reduced trees with labels $p(v_0)$ *and* k_{v_0} *fixed to be zero, of the expressions (8.1) or (8.3), admits the bound* $D_0 C_0^{m_0-1}$ *for some constants* $C_0, D_0 > 0$, *if* m_0 *is the number of free nodes,* $m_0 < 2h_0$, *with* $h_0 \le h$ *being the reduced order of* $\bar{\theta}^G$.

Remark 1 : Note that the first two statements are easy consequences of the definitions, while the third one is rather deep, being essentially equal to the KAM theorem, as it appears from the proof, (see also [G2] and [GG]).
Remark 2: We note in advance that, as will be shown along the proof of the lemma, when contributions with $\alpha_v = 1$ and $\alpha_v = -1$ are summed together, the corresponding nodes v turn out to have, in the respective cases, $p(v) = 1$ and $p(v) = -1$, so that $p(v) \neq 0$, *i.e.* $v \notin \mathcal{L}_{-1}$. Therefore, since the cancellation implemented in Lemma 8.2 below occurs between contributions with a different label ρ_v affixed to a node $v \in \mathcal{L}_{-1}$, no cancellations overlapping can arise.

Lemma 8.2: *The contribution to* $\Xi^{h\sigma}_{j\vec{\nu}}(t)$, $\vec{\nu} \in \mathbf{Z}^{l-1}$, $\sigma = \pm 1$, $j = 0$, *arising from the sum of the stripped values of all the generalized reduced trees of reduced degree* m_0, *with labels* $p(v_0) = 0$ *and* $k_{v_0} = -1$, *can be written as:*

$$\sum_{r=1}^{m_0} Q^r_{v_0}(x)\frac{(gt)^r}{r!}\;E(m_0 - r)\int \mu_r(ds) e^{i\vec{\omega}(s)\cdot\vec{\nu}_0(v_0)t} A_{v_0}(\vec{\omega}\cdot\vec{\nu}_0(v_0), r, s) \tag{8.4}$$

where $|\vec{\nu}_0(v_0)| \le m_0 N$, r *is the number of nodes in* $\mathcal{L}_{-1}$, $s = \{s_1, \ldots, s_r\}$, *with* $s_i \in [0,1]$, $i = 1, \ldots, r$, *being "interpolation parameters", and* $\mu_r(ds)$ *is a suitable normalized positive measure:*

$$\mu_r(ds) = ds_1 ds_2 \ldots ds_{r-1} ds_r\;[r\, s_1^{r-1}]\,[(r-1)\, s_2^{r-2}] \ldots [s_{r-1}]\;,$$

and the nodes in $\mathcal{L}_{-1}$ *are totally ordered so that* $w_i < w_j$ *for any* $i < j$, *with* $w_1 = v_0$, $i = 1, \ldots, r$.[7] *The function* $\vec{\omega}(s)\cdot\vec{\nu}_0(v_0) \equiv \vec{\omega}(v_0, s)$ *is defined in the following way. Let us call* $\theta(w_i)$ *the (generalized reduced) tree with first node* w_i, *and* $\theta(w_i)\backslash\theta(w_{i+1})$ *the tree obtained from* $\theta(w_i)$ *by deleting the entire subtree emerging from* w_{i+1} *(recall that* $w_{i+1} > w_i$*), the node* w_{i+1} *included. Then:*

$$\omega(v_0, s) = \sum_{i=1}^{r} s_1 \ldots s_i \sum_{w \in \theta(w_i)\backslash\theta(w_{i+1})} \omega_w \tag{8.5}$$

Note that $\omega(v_0, s)$ *satisfies the property that* $0 \le |\omega(v_0, s)| \le m_0 N$, *as* $\omega(v_0)$ *did. The functions* $Q^r_{v_0}(x)$ *are defined as:* $Q^r_{v_0}(x) = \sum_{k\ge 1}^{\infty} Q^r_{v_0}(k)x^k$, $x = \exp[-\sigma g t]$, *and the functions* $E(m_0 - r)$ *and* $A_{v_0}(\vec{\omega}\cdot\vec{\nu}_0(v_0), r, s)$ *verify the bounds:* $E(p) \le e^{2p}$, *and* $|A_{v_0}(\vec{\omega}\cdot\vec{\nu}_0(v_0), r, s)| \le D_1 C_1^{m_0-1}$ *for some constants* $C_1, D_1 > 0$.

Let us consider a generalized reduced tree with given shape and collection of indices, and let us consider the $p(v)$ labels. Let us single out the nodes v's, with $p(v) = 0$: then each such node will be enclosed, together with all the generalized reduced subtree emerging from it, inside a bubble β_v which will be wiggly if $j_v > 0$, and smooth if $j_v = l$. 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 generalized 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 *withered flowers* the wiggly bubbles, and *fresh flowers* the smooth ones; unlike the leaves, the flowers will not be considered nodes. A generalized 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 either: i) the sum over all the generalized reduced trees

[7] That is the nodes $w_1, \ldots, w_r$ belong to a connected monotone path.

of the stripped values (8.1), times the product of the leaf values, (if the flower is withered), or: ii) the sum over all the generalized reduced trees of the stripped values (8.3), times the product of the leaf values, (if the flower is fresh, and $k_v = 0$), or: iii) an expression differing from (8.4) inasmuch it lumps together also the leaf values, (if the flower is fresh, and $k_v = -1$). We shall see later that, in order to obtain the latter expression, it will be enough to substitute the function $A_{v_0}(\vec{\omega}\cdot\vec{\nu}_0(v_0), r, s)$ in (8.4) with a function which admits the same bound, being m_0 replaced with m, (see also note 9).

The degree of a generalized reduced tree is given by the number of its free nodes plus the sum of the degrees of its withered and fresh flowers, and of its leaves; analogously, the order of a generalized reduced tree is given by the sum of the order labels of its nodes plus the sum of the orders of its flowers.

All the withered flowers give a contribution to the stripped value of the generalized reduced tree of the form (8.1), (by Lemma 8.1), and the dependence on the time variable reveals itself only through the exponential factor $\exp[i\vec{\omega}\cdot\vec{\nu}(v)\tau_v]$. As to the fresh flowers, they contribute to the stripped value a factor (8.4), (we can imagine to rewrite (8.3) in the same form, with the constraints $Q^1_v(x) = -in_v(\cosh gt)^{-1}$ and $Q^r_v(x)(x) = 0$ if $r \geq 2$). Obviously in both cases we have to take into account the leaf values too.

Lemma 8.3: *The contribution to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$, $\vec{\nu} \in \mathbf{Z}^{l-1}$, $\sigma = \pm 1$, $2l > j \geq 0$, arising from the sum of the values of all the generalized reduced trees of degree m, with labels $p(v_0) \neq 0$, can be written as:*

$$\sum_{r_0=0}^{m-1}\sum_{r=0}^{m-1} Q^r_{v_0}(x)\,\frac{(gtr_0)^r}{r!}\,E(m-1-r)\int \mu_r(ds)e^{i\vec{\omega}(s)\cdot\vec{\nu}_0(v_0)t}B_{v_0}(\vec{\omega}\cdot\vec{\nu}_0(v_0), r, s) \tag{8.6}$$

where $|\vec{\nu}_0(v_0)| \leq mN$, r_0 is the number of fresh flowers, r is the sum of the powers of the time variables the fresh flowers contribute, $\mu_r(ds)$ and $\vec{\omega}(s)\cdot\vec{\nu}_0(v_0)$ are defined as in Lemma 8.2, r_0^r is meant as 1 when $r = r_0 = 0$, and $Q^r_{v_0}(x) = \sum_{k\geq r_0}^{\infty} Q^r_{v_0}(k)x^k$, $x = \exp[-\sigma gt]$; the function $E(m-1-r)$ admits the same bound as the homonymous one in Lemma 8.2, and $|B_{v_0}(\vec{\omega}\cdot\vec{\nu}_0(v_0), r, s)| \leq D_2 C_2^{m-1}$ for some constants $D_2, C_2 > 0$.

Proof of Lemma 8.1 : Let us consider a generalized reduced tree $\bar{\theta}^G$; if $p(v_0) = 0$, $k_{v_0} = 0$, the root branch can be $j_{v_0} = l$, or $j_{v_0} > l$. If v_0 is the only tree node (*i.e.* if $\bar{\theta}^G$ is the *trivial tree*), the result is obvious, by direct check. Otherwise, for each $\bar{v} \geq v_0$, $\bar{v} \in \bar{\theta}^G$, it is $p(\bar{v}) = k_{\bar{v}} + \sum_{\bar{\theta}^G \ni w > \bar{v}}(k_w + k'_w)$, see (7.16), where $k_w + k'_w \geq 0$, for each w, see (7.7), and $k_w \equiv 0$ if w is a leaf, see (7.15). 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 $k_{\tilde{w}} + k'_{\tilde{w}} = 1$. Under the hypothesis of the lemma, only the first case must be considered here. 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 $k'_w = 0$. We note that the case $p(\bar{v}) = 0$ and $\alpha_{\bar{v}} = -1$ is not possible: *this means that, in the case we are studying, as far as the free nodes are concerned, the generalized reduced trees behave in the same way as the reduced trees, and, in particular, the time variables are ordered and (7.16) can be directly applied, (in particular we can set $\sigma_w = \sigma_{v_0} \equiv \sigma$, $\forall\, w \geq v_0$, $w \in \bar{\theta}^G_f$).* Then we can write:

$$\sum_{\theta} V_j(t;\theta) = \sum_{\bar{\theta}^G} V^S_j(t;\bar{\theta}^G)\prod_{i=1}^{\mathcal{N}_L} L^{h_{v_i}\sigma_{v_i}}_{j_i}(0) \tag{8.7}$$

where $\mathcal{N}_L$ is the number of leaves of the generalized reduced tree $\bar{\theta}^G$, and $j_i \equiv j_{\lambda_{v_i}}$, where v_i is the i-th leaf. Note that (8.7) is the product of factorizing terms, which can be treated separately, being independent on each other; each $L^{h_{v_i}\sigma_{v_i}}_{j_i}(0)$, $i > 0$, corresponds to a leaf and has as first node a node v_i with $\rho_{v_i} = 0$, while $V^S_j(t;\bar{\theta}^G)$ can have either ρ_{v_0} or $\rho_{v_0} = 1$. Moreover each $L^{h_{v_i}\sigma_{v_i}}_{j_i}(0)$, $i > 0$, 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 $V^S_j(t;\bar{\theta}^G)$, 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: see Lemma 8.3). Note that $V^S_j(t;\bar{\theta}^G)$ corresponds to the stripped value of a generalized reduced tree, so that the hypothesis of Lemma 8.1 applies to it.

As indicated in the statement of the lemma, if $j_w = l$ we consider together the cases $w \in \Lambda_{-1}$ and $w \in \Lambda_1$: they give a contribution to (7.8), containing, as far as the w node is concerned, a factor $\Phi_{\nu_w}\exp[i\omega(w)(\tau_w - \tau'_w)]$ times $e^{-g\sigma(\tau_w-\tau_{w'})}y^{(1)}_w(-1,1) - e^{g\sigma(\tau_w-\tau_{w'})}y^{(-1)}_w(1,-1) = (\sigma/2)[e^{-g\sigma(\tau_w-\tau_{w'})} - e^{g\sigma(\tau_w-\tau_{w'})}]$. ¿From (7.8) and (7.16) we can obtain a sequence of factorizing integrals; then, for the top nodes different from the leaves (top free nodes), we have

$$\oint \frac{dR_v}{2\pi i R_v} \int_{\sigma\infty}^{0} d\, g\tau_v \;\; T_v(-g\tau_v)\, e^{-gR_v \sum_{w\le v} \sigma\tau_w}\, e^{i\tau_v\omega_v}\, e^{-gk_v\sigma\tau_v} \tag{8.8}$$

where $T_v(-g\tau_v) = (-g\tau_v)^{1-\delta_{j_v,l}}$, see (7.10). The time integration is trivial and yields:

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

where $k_v = k(v) = p(v)$ and $\omega_v = \omega(v)$. 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 (4.9) requires in such a case $\omega(v) \ne 0$. If $j_v = l$, we have to sum together the two contributions $k_v = \pm 1$; if $j_v > l$, we have a factor $y_v^{(2)}(0,0) = 1$. Therefore the residue at $R_v = 0$ is

$$\begin{cases} \left[ig^{-1}\omega(v)\right]^{-2} & \text{if } j_v > l \\ \left[1 + g^{-2}\omega^2(v)\right]^{-1} & \text{if } j_v = l \end{cases} \tag{8.9}$$

(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 generalized reduced tree obtained from $\bar\theta^G$ by deleting the original top free nodes, and we have again to consider an expression like (8.8), so that all the integrations can be performed in the same way, for each $v \ne v_0$, if only we take 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 (8.9).

In the end, only the node v_0 is left. Since $k_{v_0} = 0$, if $j_{v_0} > l$, we have a coefficient $y^{(2)}(0,0) = 1$: 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 $\exp[i\omega(v_0)\tau_{v_0}]$, and we obtain (8.1), if $G_v[\omega(v)]$ is defined as in (8.2). Otherwise, if $j_{v_0} = l$, then $k_{v_0} = 0$ requires $v_0 \in \Lambda_{-1}$, and we have a coefficient (see (7.6)):

$$(-1)^{\rho_{v_0}} \sum_{k'_{v_0}=1}^{\infty} y_{v_0}^{(-1)}(k'_{v_0}, 0) x^{k_{v'_0}} = \frac{(-1)^{\rho_{v_0}}}{2} \frac{2in_{v_0}}{\cosh gt} \tag{8.10}$$

and, if we integrate in τ_{v_0} and sum together the contributions $\rho_{v_0} = 0, 1$, we obtain (8.3). So Lemma 8.1 is proven if we show that the bound $D_0 C_0^{m_0 - 1}$, in the statement 3) of Lemma 8.1, holds. This will be done in Appendices A1, A2 and A3. ■

Proof of Lemma 8.2 : The expression (8.4) can be checked by induction. The case $p(v_0) = 0$ and $k_{v_0} = -1$ is the case put aside in the above discussion, (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 $w > v_0$, $w \ne \tilde w$), 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 $\mathcal{P}$.

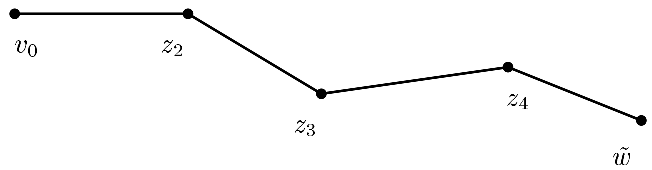

Fig.8.1: A path $\mathcal{P}$ connecting the first node v_0 of the generalized reduced tree $\bar\theta^G$, (single path tree), with the node $\tilde w$, (defined as the node verifying the condition $k_{\tilde w} + k'_{\tilde w} = 1$), with $m_{\mathcal{P}} = 5$, $z_1 = v_0$ and $z_5 = \tilde w$.

Given a generalized reduced tree $\bar\theta^G$ with $p(v_0) = 0$, and $k_{v_0} = -1$, then it *will* have a path $\mathcal{P}$: so we call it a *single path tree*. For each z_i, 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 case: $z_i \in \Lambda_{-1}$, $z_i \in \Lambda_2$, $z_i \in \Lambda_1$.

Note that $\mathcal{L}_{-1} \cap [\bar{\theta}^G \setminus \mathcal{P}] = \emptyset$, as can be seen by *reductio ad absurdum*: in fact, if $w \in \mathcal{L}_{-1}$ 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. In particular this shows that the nodes in $\mathcal{L}_{-1}$ are totally ordered as it is said in the statement of the lemma.

As a consequence of what has been said, we see that, in order to obtain the contribution to $\Xi^{h\sigma}_{j\bar{\nu}}(t)$, with $p(v_0) = 0$, $k_{v_0} = -1$, we have to consider the sum of products of several factorizing terms, as in proof of Lemma 8.1, (8.7), which are of the same type of before, up to the first factor, which is given by the stripped value of a generalized reduced tree with a fixed shape, and labels $p(v_0) = 0$, $k_{v_0} = -1$. Therefore we have to study only this term.

For each z_i we consider separately the generalized reduced subtree with root equal to z_i and first node z_{i+1}, and the remaining $m_{z_i} - 1$ generalized reduced subtrees $\bar{\theta}^G_{ij}$, with root z_i, and first node v_{ij}, $j = 1, \ldots, m_{z_i} - 1$, if $\{v_{ij}\}$ is the set of nodes immediately following z_i, different from z_{i+1}.

We treat in a different way the case in which there is no node with $p(z_i) = 0$, and the case in which there is at least one such node. In the first case, if $\tilde{w}$ is not a leaf, since the *a priori* possible situations are either $k_{\tilde{w}} = 1$ and $k'_{\tilde{w}} = 0$, or $k_{\tilde{w}} = 0$ and $k'_{\tilde{w}} = 1$, it must be $k_{\tilde{w}} = 0$ and $k'_{\tilde{w}} = 1$, because $y^{(1)}_v(0, 1) = 0$; if $\tilde{w}$ is a leaf, then again $k_{\tilde{w}} \equiv 0$ and $k'_{\tilde{w}} = 1$. Therefore the node $\tilde{w}$ can be treated as in the proof of Lemma 8.1, and so we can study the generalized reduced subtrees $\bar{\theta}^G_{ij}$, $\forall\, z_i$, so obtaining from each of them a contribution of the form either $\exp[i \sum_j \omega(v_{ij}) \tau_{z_i}]$ times $\prod_{w \in \cup_j \bar{\theta}^G_{ij}} G_w[\omega(w)]$, if v_{ij} is a free node, or $L^{h_{v_{ij}} \sigma_{v_{ij}}}_{j \lambda_{v_{ij}}}(0)$, if v_{ij} is a leaf. Therefore we are left with the integrations along the path $\mathcal{P}$: but it is always $p(z_i) \neq 0$, so that we can factorize the integrations and obtain a product of terms $(p(z_i) - i\sigma g^{-1} \omega(z_i))$ to some negative power (1 or 2), which can be bounded by 1.

Otherwise, if there are nodes $z_i \in \mathcal{P}$ with $p(z_i) = 0$, (8.4) can be verified by induction: this is done in Appendix A4, so that the proof of Lemma 8.2 lacks only the control of the sums over all the generalized reduced trees. But the number of addends is trivially bounded, if m_0 is the reduced degree of the generalized reduced tree, 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, ($\leq (3N)^{l m_0} \cdot 2^{m_0} \cdot 3^{m_0} \cdot 3^{m_0}$). ■

Proof of Lemma 8.3 : 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 8.1, 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 fresh flower a contribution (8.4) arises, and, if v is the end point of the flower stem, we can decompose the powers of $\tau_{v'}$ as in the proof of Lemma 8.2, so constructing several paths along the generalized reduced tree, (which will be called a *multiple paths tree*), where the paths are uniquely determined by the request that they connect the first node v_0 to the fresh flowers stems. Then we can explicitly perform the integrations over the time variables of the nodes belonging to the paths, and it can be checked that no factorials arise, by reasoning as in the proof of Lemma 8.2, (the details can be found in Appendix A5).

Nevertheless we must be careful, because we still have to sum over the labels $p(v)$, (the sum over the other labels can be treated as in the previous cases). 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}:

$$\int_{\sigma\infty}^{gt} d\, g\tau_{v_0}\, (\ldots) = \int_{\sigma\infty}^{\sigma} d\, g\tau_{v_0}\, (\ldots) + \int_{\sigma}^{gt} d\, g\tau_{v_0}\, (\ldots) \equiv I_m + \int_{\sigma}^{gt} d\, g\tau_{v_0}\, (\ldots) \tag{8.11}$$

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 (7.13) 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 (7.7)), and the convergence follows. We have left the term in (8.11) 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:

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

(all the free nodes v's have $p(v) \neq 0$, so that $\rho_v = 1$) where $\tilde{\theta}^G$ is a subtree of $\bar{\theta}^G$ 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 $I_m \le C^m$, 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 $\exp[-p(v_0)g\sigma t] \le \exp[-p(v_0)]$ 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 generalized 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 generalized 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.[8] Therefore the bound (8.6), in the statement of Lemma 8.3, can be inductively checked, exploiting the results of Lemmata 8.1 and 8.2 too, as far as the leaves with label $p(v) = 0$ are concerned.[9] This completes the proof of Lemma 8.3. ■

We can now state the fundamental result giving the convergence property of the series defining the whiskered tori, and so completing the proof of Proposition 4.1.

Theorem 8.1: *Let us denote by $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ the dimensionless perturbed motion, $0 \le j < 2l$. We can always write it in the form:*

$$\Xi^{h\sigma}_{j\vec{\nu}}(t) = \sum_{r=0}^{2h-1} \tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(x, \vec{\omega}t; r) \, \frac{(gt)^r}{r!} \tag{8.12}$$

where $|\vec{\nu}| \le (2h-1)N$, and $\tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(x, \vec{\omega}t; r)$ is an analytic function in x, $\tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(x, \vec{\omega}t; r) = \sum_{p=0}^{\infty} \tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(p, \vec{\omega}t; r) x^p$, with $|(gt)^r \, \tilde{\Xi}^{h\sigma}_{j\vec{\nu}}(p, \vec{\omega}t; r)| \le \bar{D}\bar{C}^{2h-1} \, r!$, for some constant $\bar{C}, \bar{D} > 0$, and for all $r \ge 0$, $p \ge 0$, for any $\sigma t \ge 0$.

The result stated in Theorem 8.1 follows directly from Lemma 8.3, as far the contribution $|p(v_0)| \ge 1$ is concerned, if we take into account the inequalities $x^p e^{-px} \le 1$, $x^p e^{-x} \le p!$, $\forall \; p \ge 0$, $x \ge 0$, and we explicitly bound the sum over r_0 in (8.6).

For the contributions $p(v_0) = 0$, it follows from Lemma 8.1 and Lemma 8.2, or better from their proof, as we have to estimate also the contribution to $\Xi^{h\pm}_{j\vec{\nu}}(t)$, with $j \ge l$: it is easily seen that the discussion can be repeated essentially unchanged and leads to the same convergence result. The leaves can be treated as in the proof of Lemma 8.3, so that the writing (8.12) is proven. ■

Obviously, if we want to find a bound on the homoclinic splitting, we can write $\bar{\Delta}^h_{j\vec{\nu}} = \Xi^{h+}_{j\vec{\nu}}(0) - \Xi^{h-}_{j\vec{\nu}}(0)$, so obtaining the same bound of Theorem 8.1, up to a factor 2. This proves the first of (6.2), which therefore can be considered a corollary of Theorem 8.1.

Appendix A1: Proof of the convergence bound in Lemma 8.1

[8] If we recall the proof of the convergence bound of Lemma 8.1, (as it is carried out in Appendices A1, A2, A3), we can note that it was obtained by exploiting some cancellations we could implemented by summing together different generalized reduced trees, (inside the same family $\mathcal{F}(\theta)$, see Appendix A2); 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 generalized 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 generalized reduced tree $\bar{\theta}^G$, with all its leaves of fixed orders; then we sum over all the terms of the family $\mathcal{F}(\theta)$, in which $\bar{\theta}^G$ is contained, so that the cancellation mechanism is implemented.

[9] Note that the leaves can have $p(v) = 0$, so that, if this is the case, the bounds of Lemma 8.1 and Lemma 8.2 have to be implemented. A leaf v with $p(v) = 0$ contributing, *e.g.* , to the generalized reduced tree value (8.7) through the factor $L_j^{h_{v_j}\sigma_{v_j}}(0)$ admits a representation analogous to the same (8.7), and can be expressed as a sum of terms, which are given by the product of the stripped value of the generalized 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 (8.6) can be assumed to hold, and an inductive proof can be performed.

As we have seen in §8, from the case $p(v_0) = k_{v_0} = 0$ we obtain a contribution to $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ containing a factor:

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

and we want to find a bound on the sum of (A1.1) over all the generalized reduced trees with $p(v_0)$ and k_{v_0} fixed to the above values.

Given a generalized reduced tree $\bar{\theta}^G$, it will be characterized by its shape and by a collections of labels, as shown in §5 and §7. Let us proceed as in [G2], and let us suppose a condition over the rotation vectors stronger than the hypothesis H_2, *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 [GG] to eliminate such an unneeded hypothesis. Therefore we shall make the assumption that the rotation vectors $\vec{\omega}$'s satisfy the *strong diophantine condition*:

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

where $n, p \in \mathbf{Z}$, $n \leq 0$. 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:

$$G \equiv G_v[\omega(v)] = \begin{cases} -(gC_0)^2[\vec{\omega}_0 \cdot \vec{\nu}_0(v)]^{-2} & \text{if } j_v > l \\ -(gC_0)^2 \left[(gC_0)^2[1 + (\vec{\omega}_0 \cdot \vec{\nu}_0(v))^2]\right]^{-1} & \text{if } j_v = l \end{cases} \tag{A1.3}$$

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 $n \leq 0$, if $2^{n-1} < |\vec{\omega}_0 \cdot \vec{\nu}_0(v)| \leq 2^n$.

Note that, if $j_v > l$, then, if G is on scale $n \leq 0$, 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 $j_v = l$, then $|G| \leq 1$. Such a definition, despite its asymmetry, turns out to be useful in the following estimates, and allows us to use, nearly without changes, the results of [G1]; we can get rid of the new factor $(2gC_0)^2$, by defining $C_1 = \max\{1, (2gC_0)^2\}$, and introducing a coefficient C_1^m in the bound (7.12). This implies a simple redefinition of the constant $\mathcal{C}$ in (7.12), and we can say that, if G is on scale n, then $|G| < 2^{-2n}$, $\forall\, n \leq 0$.

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 "generalized reduced tree", and the symbol θ instead of $\bar{\theta}^G$; however it is always in the meaning of the latter that the first one has to be interpretated. In Appendix A4 we will come back to the complete name. Moreover we call momentum tout court *the free momentum $\vec{\nu}_0(v_0)$.*

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 §5), 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 incoming 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 $\leq E\, 2^{-n\varepsilon}$, 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 entering 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). 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 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).

The key remark is that the resonant trees 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 total momentum by the amount $\vec{\nu}_0(v)$: this means that some of the propagators $\big[i\omega_0(w)\big]^{-2}$ have become $\big[i(\omega_0(w)+\varepsilon)\big]^{-2}$, and some of the propagators $\big[-(gC_0)^2(1+\omega_0^2(w))\big]^{-1}$ have become $\big[-(gC_0)^2(1+(\omega_0(w)+\varepsilon)^2)\big]^{-1}$, if $\varepsilon \equiv \omega_0(v)$, and:
2) there is one of the node factors which changes by taking successively the values ν_{wj}, j being the branch label of the branch leading to v, and $w \in V$ is the node to which such 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, if we sum as well on a overall change of sign of the $\vec{\nu}_w$ values (which 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.

This 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 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 some minor changes), and so we relegate it to Appendix A2 below. Here we confine ourselves to state the final result.

Let us denote $m_T(n)$ the number of resonances of scale n contained in a cluster T, and define the tree family $\mathcal{F}(\theta)$ as follows (such definitions will become more clear in Appendix A2, as that of resonance given above, which has not been completely used so far). Given a resonance V of θ we detach the part of θ above λ_V (λ_V included) 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, and 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:

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

where:
1) N_n is the number of propagators of scale n in θ ($n=1$ does not appear as $|G| \ge 1$ in such cases), satisfying the inequality:

$$N_n \le \frac{4m}{E\,2^{-\varepsilon n}} + \sum_{T,\,n_T=n}(-1+m_T(n)) \tag{A1.5}$$

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.

Hence substituting (A1.5) into (A1.4) we see that the $m_T(n)$ is taken away by the first factor in $2^{2n}2^{-2n_i}$, while the remaining 2^{-2n_i} are compensated by the -1 before the $+m_T(n)$ 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:

$$\frac{1}{m!} e^{2m} 2^{4m} 2^{6m} \prod_n 2^{-8nmE^{-1}2^{\varepsilon n}} \le \frac{1}{m!} B_0^m \tag{1.6}$$

for $B_0 = e^2 2^{10} \exp\left[\left(2^{3+3\tau^{-1}} \log 2\right) \sum_{p=1}^{\infty} p 2^{-p\tau^{-1}}\right]$. Note that the propagators with $j_v = l$ are bounded by 1, independently on the scale label n: in fact the above described algorithm produces a gain only for the strong resonances.

To sum over the trees we note that, fixed θ, the collection of clusters is fixed. Therefore we only have to multiply (1.6) by the number of tree shapes for θ, ($\le 2^{2m} m!$), and by the number of ways of attaching the mode labels, ($\le (3N)^{lm}$), and the $p(v), \alpha_v$ labels, ($\le 3^m$), so that we can bound $\Xi^{h\sigma}_{j\vec{\nu}}(t)$ by an exponential of m and the bound of Lemma 8.1, 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, let $m_T(n)$ be the number of resonances of scale n contained in T, (*i.e.* with incoming branches of scale n), we have the inequality (A1.5), which is an adaptation presented in [G1] of the version of *Brjuno's lemma* as it is exposed in [P]: a proof is in Appendix A3.

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 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 $\tilde{V}$. To the just defined set of trees we add the trees obtained by reversing simoultaneously the signs of the node modes $\vec{\nu}_w$, for $w \in \tilde{V}$: the change of sign is performed independently for the various resonant clusters. This defines a family of $\prod 2M_V$ 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, as it is explained in note 8.

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 $\prod 2M_V \le e^{2m}$ products of propagators of the members of the family $\mathcal{F}(\theta^1)$ can be bounded above by $\prod_\lambda 2^{-2(n_\lambda-2)} = 2^{4m}\prod_\lambda 2^{-2n_\lambda}$, 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 second order 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 (A1.4); the result is explained as follows:
i) 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 for $|\varepsilon_i| < 2^{n_i-3}$ if $n_i \equiv n_{V_i}$, provided $n_i > n+3$.
ii) the resummation says that the dependence on the ε_i's has a second order zero in each. 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 $(|\varepsilon_i|\,2^{-n_i+3})^2$ 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) > 1, we can leave them in it to simplify the notation.

The main point here (and the main difference with respect to the otherwise identical discussion of [G1]) is that, for $n \le 0$, not all the resonances are strong resonances, so that $m_T(n)$ is a bound on the number of strong resonances, to which all the cancellations exploited in Appendix A1 apply.

Appendix A3: Resonant Siegel-Brjuno bound

In the following discussion, which is taken from [G1], 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.

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 with scale $> n$ then calling $\theta_1, \theta_2, \ldots, \theta_k$ the subtrees of θ emerging from the first vertex of θ and with $m_j > E\,2^{-\varepsilon n}$ lines, 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 (A1.5). 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 $N_n^* = 0$.

In the other case it is $N_n^* \le 1 + \sum_{i=1}^k N_n^*(\theta_i)$, and if $k = 0$ the statement is trivial, or 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 $m_1 > m - \frac12 E\,2^{-n\varepsilon}$. 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 is either a resonance or it 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 $\le m - \frac12 E\,2^{-n\varepsilon}$.

It remains to check that if $m_1 > m - \frac12 E\,2^{-n\varepsilon}$ then the root line of θ_1 has scale $> n$, unless it is entering a resonance.

Suppose that the root line of θ_1 is not entering a resonance. Note that $|\vec\omega\cdot\vec\nu_0(v_0)| \le 2^n, |\vec\omega\cdot\vec\nu_0(v)| \le 2^n$, if v_0, v_1 are the first vertices 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 $m - m_1 < \frac12 E\,2^{-n\varepsilon}$ (and we are not considering the resonances), it follows that $m - m_1 < \frac12 E\,2^{-n\varepsilon}$ is inconsistent: it would in fact imply that $\vec\nu_0(v_0) - \vec\nu_0(v_1)$ is a sum of $m - m_1$ vertex modes and therefore $|\vec\nu_0(v_0) - \vec\nu_0(v_1)| < \frac12 NE\,2^{-n\varepsilon}$ 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

$m_1 \le m - E2^{-\varepsilon n}$. Then (A1.5) is proved.

Appendix A4: Bound on the single path trees

Let us consider a single path tree, and let us denote by $z > v_0$ the first node with $p(z) = 0$, (the case in which such a node does not exist has been considered already in §8). Let us suppose inductively that (8.4) holds (for $m = 1$ it can be checked easily to be valid, for some constant C_1). Then the generalized reduced subtree $\bar\theta_1^G$ with root z has $m_1 \le m - q$ nodes, $q \ge 1$ being the number of nodes of $\mathcal{P}$ preceding z, and (8.4) is supposed to hold for it by the inductive hypothesis. We treat as in the case considerd in §8 all the generalized reduced subtrees $\bar\theta_{ij}^G$, $z_i < z$, and we are left with the integrations over the nodes $z_i \in \mathcal{P}$, $z_i < z$:

$$\int_0^{g\tau_{v_0}} dg\tau_{v_0}\, e^{i\omega(v_0)\tau_{v_0}}\, y_{v_0}^{(-1)}(k'_{v_0},-1)\prod_{i=2}^{q}\ \int_{\sigma\infty}^{g\tau_{z_{i-1}}} dg\tau_{z_i}\, e^{[-g\sigma p(z_i)+i\omega(z_i,s)](\tau_{z_i}-\tau_{z_{i-1}})}\cdot$$
$$\cdot\,[g(\tau_{z_{i-i}}-\tau_{z_i})]^{1-\delta_{j_{z_i},l}}y_{z_i}^{(\alpha_{z_i})}(k'_{z_i},k_{z_i})\sum_{r_1=1}^{m_1}\frac{(g\tau_z)^{r_1}}{r_1!}E(m_1-r_1)$$

where $\omega(z_i,s)\equiv\vec\omega(s)\cdot\vec\nu_0(z_i)$, $z_{q+1}=z$, $p(z_i)\ne 0$, if $q\ge i>1$, and $\omega(z_i,s)$ depends on s through the addend $\omega(z,s)$. The node $v_0\equiv z_1$ has $p(v_0)=0$, and $\rho_{v_0}=0,1$, so that we sum over the two possible values of the latter label. We decompose:

$$\begin{aligned}&\sum_{r_1=0}^{m_1}\frac{(g\tau_z)^{r_1}}{r_1!}E(m_1-r_1)\prod_{1<i\le q}[g(\tau_{z_{i-1}}-\tau_{z_i})]^{1-\delta_{j_{z_i},l}}=\sum_{r_1=1}^{m_1}\frac{1}{r_1!}\sum_{n_q=0}^{r_1}\binom{r_1}{n_q}\cdot\\&\cdot[g(\tau_{z_{q-1}}-\tau_{z_q})]^{r_1-n_q+1-\delta_{j_{z_q},l}}\sum_{n_{q-1}=0}^{n_q}\binom{n_q}{n_{q-1}}[g(\tau_{z_{q-2}}-\tau_{z_{q-1}})]^{n_q-n_{q-1}+1-\delta_{j_{z_{q-1}},l}}\\&\ldots\sum_{n_2=0}^{n_3}\binom{n_3}{n_2}[g(\tau_{z_2}-\tau_{z_1})]^{n_3-n_2+1-\delta_{j_{z_2},l}}[g\tau_{z_1}]^{n_2}E(m_1-r_1)\prod_{i=1}^{q}(-1)^{1-\delta_{j_{z_i},l}}\end{aligned}\tag{A4.1}$$

(it is $j_{v_0}=l$) and all the integrations over the nodes $z_i\in\mathcal{P}$ give factors $p(z_i)-i\sigma g^{-1}\omega(z_i)$ to some negative power $n_i-n_{i-1}+1-\delta_{j_{z_i},l}$, (which we bound again by 1), times a factorial of the same power $[n_i-n_{i-1}+1-\delta_{j_{z_i},l}]!$, so that we are left with:

$$\begin{aligned}&\sum_{r_1=1}^{m_1}\frac{1}{r_1!}\sum_{n_q=0}^{r_1}\frac{r_1!}{n_q!}(r_1-n_q+1)^{1-\delta_{j_{z_q},l}}\sum_{n_{q-1}=0}^{n_q}\frac{n_q!}{n_{q-1}!}(n_q-n_{q-1}+1)^{1-\delta_{j_{z_{q-1}},l}}\\&\qquad\ldots\sum_{n_2=0}^{n_3}\frac{n_3!}{n_2!}(n_3-n_2+1)^{1-\delta_{j_{z_2},l}}(g\tau_{v_0})^{n_2}\,E(m_1-r_1)=\sum_{r=0}^{m_1}\frac{(g\tau_{v_0})^r}{r!}\tilde E(m_1-r)\,c(q)\end{aligned}$$

where $c(q)\le\exp[2(q-1)]$, and $\tilde E(m_1-r)$ is so defined to satisfy the same bound as $E(m_1-r)$, as can be easily checked:

$$\begin{aligned}&\sum_{r_1=1}^{m_1}\sum_{n_2=0}^{r_1}\frac{(g\tau_{v_0})^{n_2}}{n_2!}\sum_{n_3=n_2}^{r_1}\cdots\sum_{n_q=n_{q-1}}^{r_1}e^{2(m_1-r_1)}(r_1-n_q+1)^{1-\delta_{j_{z_q},l}}\ldots(n_3-n_2+1)^{1-\delta_{j_{z_2},l}}\\&\le\sum_{r_1=1}^{m_1}\sum_{n_2=0}^{r_1}\frac{(g\tau_{v_0})^{n_2}}{n_2!}\sum_{n_3=0}^{r_1-n_2}\cdots\sum_{n_q=0}^{r_1-\sum_{i=2}^{q-1}n_i}e^{2(m_1-r_1)+(r_1-n_2)+(q-1)}\\&\le\sum_{n_2=0}^{m_1}\frac{(g\tau_{v_0})^{n_2}}{n_2!}e^{q-1}\,e^{2(m_1-n_2)}\sum_{r_1=n_2}^{m_1}e^{n_2-r_1}\frac{(r_1-n_2)^{q-1}}{(q-1)!}\\&\le\sum_{n_2=0}^{m_1}\frac{(g\tau_{v_0})^{n_2}}{n_2!}e^{q-1}\,e^{2(m_1-n_2)}\sum_{p_2=0}^{m_1-n_2}e^{-p_2}\frac{(p_2)^{q-1}}{(q-1)!}\\&\le\sum_{r=0}^{m_1}\frac{(g\tau_{v_0})^r}{r!}e^{q-1}\,e^{2(m_1-r)}\frac{e^{q-1}}{(q-1)!}\sum_{p=0}^{\infty}p^{q-1}\,e^{-p}\end{aligned}\tag{A4.2}$$

where the last sum is bounded by $(q-1)!$, so that we obtain a factor bounded by $e^{2(m_1-r)}$, times a factor bounded by $e^{2(q-1)}$, which will be taken into account by the term $A_{v_0}(\vec{\omega}\cdot\vec{\nu}_0(v_0),r,s)$. It remains to perform the last integration; we have trivially:

$$\int_0^{gt} dg\tau_{v_0} e^{i\omega(v_0)\tau_{v_0}} \frac{(g\tau_{v_0})^r}{r!} = \frac{(gt)^{r+1}}{r!}\int_0^1 ds_1\, e^{is_1\omega(v_0)t}\, s_1^r = \frac{(gt)^{r+1}}{(r+1)!}\int_0^1 ds_1\, e^{is_1\omega(v_0)t}\, [(r+1)s_1^r]$$

so that the power of the time variable increases by 1 for all $r=0,\ldots,m_1$, and it is:

$$\omega(v_0,s) = s_1\left[\sum_{w\in\theta(v_0)\setminus\theta(z)} \omega_w + \omega(z,s)\right]$$

according to (8.5), and:

$$\mu_m(ds) = ds_1 ds_2 \ldots ds_{m-1} ds_m\, [m\, s_1^{m-1}]\,[(m-1)\, s_2^{m-2}]\ldots[s_{m-1}], \qquad \int \mu_m(ds) = 1 \tag{A4.3}$$

is inductively proven to be consistent.

Since $m_1+1\le m$, the above discussion and the convergence bound of Lemma 8.1 complete the result stated in Lemma 8.2.

Appendix A5: Bound on the multiple paths trees

Let us consider a multiple paths tree $\bar{\theta}^G$, with m_0 nodes, w_0 withered flowers, and r_0 fresh flowers; the latter ones are of degree, respectively, $m_1,\ldots,m_{r_0}$, $m_1+\ldots+m_{r_0}\equiv\mathcal{M}\le m-1$, and individuate a set of r_0 paths $\mathcal{P}_i$, $i=1,\ldots,r_0$, connecting their stems to the first node v_0. Let us remark that: 1) each fresh flower $\beta_{\bar{v}}$ is characterized to have $k'_{\bar{v}}\ge 1$, so that it contributes at least 1 to each $p(v)$, $v\le\bar{v}$; 2) each v crossed by r'_0 paths will have a label $p(v)\ge\max\{r'_0-1,1\}$, so that the case $p(v)=-1$ is not possible since it would require $k_v\le -2$.

The first step is to decompose the power of $\tau_{v'}$, for all the fresh flowers, as in (A4.1), but now including the first node v_0 too, because it behaves as all the others nodes, (remind that $p(v_0)\neq 0$), so that, since all the time variables are ordered and all the time dependence of the functions is through the differences $g(\tau_v-\tau_{v'})$'s, the integrals factorize. Note that in particular, along the paths $\mathcal{P}_i$, $i=1,\ldots,r_0$ the integrals are convergent, (in other words they are not improper integrals). The case $r_0=0$ is trivial, and gives a contribution to (8.6) with $r=r_0=0$; therefore in the following we suppose $r_0\ge 1$.

Then, to each free node $v\ge v_0$ in θ, there corresponds a factor:

$$[g(\tau_{v'}-\tau_v)]^{1-\delta_{j_v,l}}\, e^{(-g\sigma p(v)+i\omega(v))(\tau_v-\tau_{v'})} \prod_{i\,:\,v\in\mathcal{P}_i}\left[\binom{n_v^i}{n_{v'}^i}[g(\tau_v-\tau_{v'})]^{n_v^i-n_{v'}^i}\right] \tag{A5.1}$$

which, once we have integrated on the time difference variable, gives:

$$\left[\prod_{i\,:\,v\in\mathcal{P}_i}\binom{n_v^i}{n_{v'}^i}\right]\frac{\left[\sum_{i\,:\,v\in\mathcal{P}_i}\left(n_v^i-n_{v'}^i\right)+1-\delta_{j_v,l}\right]!}{(p(v)+ig^{-1}\omega(v))^{\sum_{i\,:\,v\in\mathcal{P}_i}\left(n_v^i-n_{v'}^i\right)+1-\delta_{j_v,l}}}(-1)^{1-\delta_{j_v,l}} \tag{A5.2}$$

where it is $p(v)\neq 0$, for any v, and $p(v)\ge\max\{1,\left(\sum_{i\,:\,v\in\mathcal{P}_i}1\right)-1\}$, if v is crossed by at least one $\mathcal{P}_i$, for some i, (because of the above remarks), so that the denominator can be always be bounded by 1.

Moreover we have a factor:

$$\left[\prod_{i=1}^{r_0}\frac{E(m_i-r_i)}{r_i!}\right](gt)^{n_{v_0}^i}\, e^{[-g\sigma(p(v_0)+k_{v_0})+i\omega(v_0,s)]t} \tag{A5.3}$$

where, obviously, the product is on the fresh flowers, and $t=\tau'_{v_0}$, and $\omega(v_0,s)$ depends on s only through the terms $\omega(v_i,s)$ arising from the fresh flowers β_{v_i}, $i=1,\ldots,r_0$.

We have to sum the product of the above factors over the collection of indices $\{r_i, n^i_v\}$, with the constraint $m_i \ge r_i \ge 1$, $r_i \ge n^i_v \ge n^i_{v'} \ge 0$, for all $i = 1, \ldots, r_0$, and $v \ge v_0$. We can write:

$$\Big[\sum_{i\,:\,v\in\mathcal{P}_i}(n^i_v - n^i_{v'}) + 1 - \delta_{j_v,l}\Big]! \le \Big[\sum_{i\,:\,v\in\mathcal{P}_i}(n^i_v - n^i_{v'}) + 1\Big]^{1-\delta_{j_v,l}}\Big[\sum_{i\,:\,v\in\mathcal{P}_i} 1\Big]^{\sum_{i\,:\,v\in\mathcal{P}_i}(n^i_v - n^i_{v'})}\prod_{i\,:\,v\in\mathcal{P}_i}(n^i_v - n^i_{v'})!$$

so that the the exponential simplifies the denominator in (A5.2), up to a factor ≤ 2, so giving an overall factor $\le 2^m$; moreover:

$$\prod_{v\ge v_0}\Big[\sum_{i\,:\,v\in\mathcal{P}_i}(n^i_v - n^i_{v'}) + 1\Big]^{1-\delta_{j_v,l}}\prod_{i=1}^{r_0} E(m_i - r_i) \le e^m \prod_{i=1}^{r_0} e^{2(m_i - r_i) + (r_i - n^i_{v_0})}$$

where m in e^m is a bound on the number of nodes of $\cup_{i=1}^{r_0}\mathcal{P}_i$. We can perform the sums as in Appendix A4, (A4.2), and, for each path $\mathcal{P}_i$, $i = 1, \ldots, r_0$, we obtain a contribution:

$$\sum_{r_i=0}^{m_i}\frac{(gt)^{r_i}}{r_i!}\,e^{2(m_i - r_i)}$$

so that we can rewrite:

$$\sum_{r_1=0}^{m_1}\cdots\sum_{r_{r_0}=0}^{m_{r_0}}\prod_{i=1}^{r_0}\frac{(gt)^{r_i}}{r_i!}\,e^{2(m_i-r_i)} = \sum_{r=0}^{\mathcal{M}} e^{2(\mathcal{M}-r)}\frac{(g\tau_{v_0})^r}{r!}\sum_{\substack{\{0\le r_i\le m_i\}_{i=1}^{r_0}\\ \sum_{i=1}^{r_0} r_i = r}}\frac{r}{r_1!\ldots r_{r_0}!}$$

where the last sum is bounded by r_0^r. Since $\mathcal{M} \le m-1$, and $p(v_0) \ge \max\{1, r_0 - 1\}$, $p(v_0) + k'_{v_0} \ge \max\{r_0, 1\}$, we obtain (8.6).

References

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

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

[BfG] 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.

[GG] Gallavotti, G., Gentile, G.: *Non recursive proof of the KAM theorem*, archived in `mp_arc@math.utexas.edu`, #93-229.

[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.

[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.

[Mo] 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.