Lindstedt series and Hamilton–Jacobi equation for hyperbolic tori in three time scales problems

G. Gallavotti, G. Gentile, V. Mastropietro

Università di Roma 1,2,3

27 Novembre 1998

ABSTRACT. *Interacting systems consisting of two rotators and a pendulum are considered, in a case in which the uncoupled systems have three very different characteristic time scales. The abundance of unstable quasi periodic motions in phase space is studied via Lindstedt series. The result is a strong improvement, compared to our previous results, on the domain of validity of bounds that imply existence of invariant tori, large homoclinic angles, long heteroclinic chains and drift–diffusion in phase space.*

§1. The Hamiltonian system.

1.1. Let $(\varphi,\alpha_1,\alpha_2)=(\varphi,\boldsymbol{\alpha})\in\mathbb{T}^3$ be three angles (*i.e.* positions on circles); let $(I,A_1,A_2)=(I,\mathbf{A})\in\mathbb{R}^3$ be their conjugate momenta (or "*actions*"). We consider the Hamiltonian function, depending on two parameters $\varepsilon,\eta>0$, defined by

$$\mathcal{H}=\eta^{\frac12}\Omega_1A_1+\eta\frac{A_1^2}{2J}+\eta^{-\frac12}\Omega_2A_2+\frac{I^2}{2J_0}+J_0g_0^2(\cos\varphi-1)+\varepsilon f(\varphi,\alpha_1,\alpha_2)\;, \tag{1.1}$$

with f an *even* trigonometric polynomial of degree N,N_0 in $\boldsymbol{\alpha},\varphi$ respectively; $\Omega_1,\Omega_2,J,J_0,g_0$ are positive constants.

This system describes two rotators (one anisochronous, labeled #1, and one isochronous, labeled #2) interacting with a pendulum which has its free (*i.e.* with $\varepsilon=0$) unstable equilibrium position at $I=0,\varphi=0$ and the stable one at $I=0,\varphi=\pi$. The scale of frequency of the pendulum is $O(1)$ in η; at the same time the two rotators rotate at constant speed $O(\eta^{\frac12})$ and $O(\eta^{-\frac12})$ respectively. Hence the system has three time scales: we assume $\eta<1$ so that the *slow* rotator is the #1 rotator.

The free motion admits invariant tori of dimension 2 (namely parameterized by $\mathbf{A}$, a constant vector, by $\boldsymbol{\alpha}$ arbitrary, and with $I=0$, $\varphi=0$) which are unstable and possess stable (labeled $+$) and unstable (labeled $-$) 3–dimensional manifolds (parameterized by $\mathbf{A}$, the same constant vector, by $\boldsymbol{\alpha},\varphi$ arbitrary, and with $I=\pm J_0g_0\sqrt{2(1-\cos\varphi)}$).

We shall study properties that eventually hold when $\eta\to0$. It is well known ([HMS,CG] for instance) that if ε is small most of the unperturbed tori and their manifolds still exist, just a little deformed. This means that (under the condition stated below) there exist functions $\mathbf{U}^{\pm}_{\mathbf{A}'}(\varphi,\boldsymbol{\alpha})$ and $V^{\pm}_{\mathbf{A}'}(\varphi,\boldsymbol{\alpha})$ which are divisible by ε and analytic in $\boldsymbol{\alpha},\varphi,\varepsilon$, for $\boldsymbol{\alpha}\in\mathbb{T}^2$, $|\varphi|<2\pi$, $|\varepsilon|<\varepsilon_0$, with ε_0 small enough, such that an initial datum starting on the (3–dimensional) surfaces W^σ_ε, $\sigma=\pm$, defined as

$$\mathbf{A}^\sigma(\varphi,\boldsymbol{\alpha})=\mathbf{A}'+\mathbf{U}^\sigma_{\mathbf{A}'}(\varphi,\boldsymbol{\alpha})\;,\qquad I^\sigma(\varphi,\boldsymbol{\alpha})=\pm J_0g_0\sqrt{2(1-\cos\varphi)}+V^\sigma_{\mathbf{A}'}(\varphi,\boldsymbol{\alpha})\;, \tag{1.2}$$

evolves, when the time $t\to\pm\infty$, tending to be confused with a quasi periodic motion on a invariant torus $\mathcal{T}(\mathbf{A}')$, with rotation vector

$$\boldsymbol{\omega}'=(\omega_1',\omega_2')\;,\qquad \omega_1'\stackrel{def}{=}\eta^{\frac12}\Omega_1+\eta J^{-1}A_1'\;,\qquad \omega_2'\stackrel{def}{=}\eta^{-\frac12}\Omega_2\;, \tag{1.3}$$

and furthermore such asymptotic motion takes place with $\mathbf{A}$ moving quasi periodically *with average* $\mathbf{A}'$.
All this holds if $\boldsymbol{\omega}'$ *verifies the Diophantine condition*

$$|\boldsymbol{\omega}' \cdot \boldsymbol{\nu}| > C|\boldsymbol{\nu}|^{-\tau} \,, \qquad \forall \boldsymbol{\nu} \in \mathbb{Z}^2 \setminus \{\mathbf{0}\} \,, \tag{1.4}$$

for $C, \tau > 0$ (which may depend also on η). The values of ε for which we shall be able to prove the above will be so small that the part of the stable and unstable manifolds with $|\varphi| < \frac{3}{2}\pi$ *can be represented as a graph of* $\mathbf{A}, I$ *over* $\boldsymbol{\alpha}, \varphi$.[1] Hence we look, since the beginning, for invariant tori which have the latter property.
The approach to the invariant tori, of the points that lie on their stable manifolds, will be exponential in the sense that their distances $d(t)$ to the tori will be such that

$$\lim_{t\to+\infty} t^{-1} \log d(t)^{-1} = \overline{g}_0 \,, \qquad \overline{g}_0 \equiv \overline{g}_0(\varepsilon) \stackrel{def}{=} (1 + \Gamma(\varepsilon, g_0))\, g_0 \,, \tag{1.5}$$

for a suitable analytic function $\Gamma(\varepsilon, g_0)$, divisible by ε. We shall call $\overline{g}_0$ the *Lyapunov exponent* of the torus (it will depend on ε as well as on the considered torus, *i.e.* on $\boldsymbol{\omega}'$ and on η). The exponent relative to the approach to the same torus along its unstable manifold (as $t \to -\infty$) will be the same, by time reversal symmetry defined below.
We fix throughout the paper τ ($\tau \geq 1$) and *we shall mainly study the dependence of* ε_0*, i.e. our estimate for the analyticity radius, as a function of* η*:* $\varepsilon_0 = \varepsilon_0(C, \eta)$.

1.2. Remark. The relation (1.3) between the value of the average action and the rotation vector is non trivial and it has been named in [G1,G2] (where it was pointed out) by saying that the tori of (1.1) are "torsion free" or "twistless". It is a remarkable symmetry property of (1.1), see [G1,Ge2,GGM3].

1.3. If $\varepsilon = 0$ the stable and unstable manifolds coincide (because the pendulum separatrix is degenerate); it is a degeneracy that is lost when $\varepsilon \neq 0$ and generically the manifolds will have only pairwise isolated trajectories in common, called *homoclinic trajectories.*
Nevertheless time reversal symmetry and parity symmetry[2] hold for (1.1). If S^t denotes the time evolution and the involution map i (composition of parity and time reversal) is defined by $i(\varphi, \boldsymbol{\alpha}, I, \mathbf{A}) = (2\pi - \varphi, -\boldsymbol{\alpha}, I, \mathbf{A})$, then $iS^t = S^{-t}i$ and there are relations between the stable and unstable manifolds that are preserved even for $\varepsilon \neq 0$. Namely

$$\begin{aligned} \mathbf{U}^+_{\mathbf{A}'}(\varphi, \boldsymbol{\alpha}) &= \mathbf{U}^-_{\mathbf{A}'}(2\pi - \varphi, -\boldsymbol{\alpha}) \,, \\ V^+_{\mathbf{A}'}(\varphi, \boldsymbol{\alpha}) &= V^-_{\mathbf{A}'}(2\pi - \varphi, -\boldsymbol{\alpha}) \,, \end{aligned} \tag{1.6}$$

where care must be exercised because the manifolds contain *two* points over each $\boldsymbol{\alpha}, \varphi$.[3] Hence if $\varphi \simeq \pi$ the relations in (1.6) concern points that lie on different connected manifolds; to understand what happens one should try a drawing taking into account that the above representations are considered only for $|\varphi| < \frac{3}{2}\pi$.
Looking at the manifolds at $\varphi = \pi$, *assuming their existence and that they are graphs above* $\boldsymbol{\alpha}, \varphi$ *for* $|\varphi| < \frac{3}{2}\pi$, equations (1.6) imply that, fixed $\mathbf{A}'$,

$$\mathbf{Q}(\boldsymbol{\alpha}) \stackrel{def}{=} \mathbf{U}^+_{\mathbf{A}'}(\pi, \boldsymbol{\alpha}) - \mathbf{U}^-_{\mathbf{A}'}(\pi, \boldsymbol{\alpha}) = -\mathbf{Q}(-\boldsymbol{\alpha}) \,, \tag{1.7}$$

so that $\mathbf{Q}(\mathbf{0}) = \mathbf{0}$; but, in general, $\mathbf{Q}(\boldsymbol{\alpha}) \neq \mathbf{0}$ for $\boldsymbol{\alpha} \neq \mathbf{0}$.
The function $\mathbf{Q}(\boldsymbol{\alpha})$ is called the *homoclinic splitting* (or simply *splitting*) *vector* at $\varphi = \pi$, and the determinant of the matrix with entries $\partial_{\alpha_i} Q_j(\mathbf{0})$ (splitting matrix) is called the *splitting.* One can more generally

[1] Note that if $\varepsilon = 0$ they are graphs over $\boldsymbol{\alpha}, \varphi$ for $|\varphi|$ smaller than *any* prefixed quantity $< 2\pi$.
[2] The latter symmetry is due to the assumption of evenness of f.
[3] This is in fact already so for $\varepsilon = 0$.

consider $\underline{Z} \equiv \underline{Z}(\varphi, \boldsymbol{\alpha}) = (\mathbf{U}^+_{\mathbf{A}'}(\varphi, \boldsymbol{\alpha}) - \mathbf{U}^-_{\mathbf{A}'}(\varphi, \boldsymbol{\alpha}), V^+_{\mathbf{A}'}(\varphi, \boldsymbol{\alpha}) - V^-_{\mathbf{A}'}(\varphi, \boldsymbol{\alpha}))$ which would be the splitting vector at φ. Here and henceforth the vectors in $\mathbb{R}^\ell$ will be denoted with an underlined letter (while the boldface is used for vectors in $\mathbb{R}^{\ell-1}$); so far $\ell = 3$, but shortly we shall consider $\ell \geq 3$. The function $\underline{Z}$ can be written as the gradient of a generating function Φ, *i.e.* $\underline{Z} = (\partial_\varphi \Phi, \partial_{\boldsymbol{\alpha}} \Phi)$. This is a result due to Eliasson who points out that it follows immediately from the Lagrangian nature of the stable and unstable manifolds. It is a further symmetry property.[4]

The symmetry of (1.1) (hence the consequent oddness of $\mathbf{Q}(\boldsymbol{\alpha})$) implies that there is one trajectory which swings through $\varphi = \pi$ when $\boldsymbol{\alpha}$ is exactly $\mathbf{0}$: it tends to the same invariant torus as $t \to \pm\infty$, provided the torus exists and its stable and unstable manifolds are graphs over $\boldsymbol{\alpha}, \varphi$ over an interval of φ greater than $|\varphi| < \pi$.

In this paper we prove the following result.

1.4. THEOREM. *Given the Hamiltonian (1.1), given constants $s, \Omega > 0$ and given $\eta > 0$ small enough, the following assertions hold.*

- *There are invariant tori with rotation vectors $\boldsymbol{\omega}'$ for all $\boldsymbol{\omega}'$ verifying the Diophantine condition (1.4) with constant $C = C(\eta) = \Omega e^{-s\eta^{-1/2}}$ and $|\omega'_1| \in [\frac{1}{2}\Omega_1 \eta^{\frac{1}{2}}, 2\Omega_1 \eta^{\frac{1}{2}}]$.*
- *Such tori exist for $|\varepsilon| < \varepsilon_0 = O(\eta^2)$ and for η small enough.*
- *They can be parameterized by their average actions $\mathbf{A}'$; the angular velocity is then given by the rotation vector $\boldsymbol{\omega}' \equiv (\Omega_1 \eta^{\frac{1}{2}} + \eta J_1^{-1} A'_1, \Omega_2 \eta^{-\frac{1}{2}})$ (i.e. they are "twistless") and the Lyapunov exponents have the form $\overline{g}_0 = (1 + \Gamma(\varepsilon, g_0))\, g_0$, with $\Gamma(\varepsilon, g_0)$ analytic in ε and divisible by ε.*
- *The parametric equations for such tori and for their stable and unstable manifolds ("whiskers") can be computed by a convergent perturbation series in powers of ε around the unperturbed tori with the same rotation vector and their corresponding stable and unstable manifolds.*
- *At the homoclinic intersection with $\varphi = \pi$ (existing by symmetry), between the stable manifold and the unstable manifold of each torus, the splitting is generically given by the Mel'nikov integral which is of order $O(\varepsilon^2 \eta^{-\beta} e^{-\frac{\pi}{2} g_0^{-1} \eta^{-\frac{1}{2}}})$, for ε small enough, with β depending on the degree N_0 in φ of the perturbation f: one can take $\beta = 2N_0 - 1$ and the asymptotic formula holds if $|\varepsilon| < \eta^\zeta$, $\zeta = 2(N_0 + 3)$ and η is small enough.*

1.5. REMARK. The novelty of the theorem is the "sharp" bound $\varepsilon_0 = O(\eta^2)$. If we "only" require $\varepsilon_0 = O(\eta^{\frac{9}{2}+})$ where $\frac{9}{2}+$ is any prefixed positive number $> \frac{9}{2}$ the result is proved in [GGM3] (see also [CG] or [Ge2]). The improvement is made possible by the *totally different technique* used (with respect to [GGM3]): a technique that has interest in its own right and, we think, beyond the result itself. In fact the proof of the last assertion of the theorem is the content of [GGM2], and the values of the constants β and ζ are taken from Appendix A2 of [GGM2].

1.6. Theorem 1.4 will be proved by a further extension of Eliasson's method, [E,G2,Ge1,Ge2], for the KAM theorem. The following discussion will show the correctness of the intuition that "new" small divisors appear in the perturbation expansion *at orders spaced by $O(\eta^{-1})$*. So that the coupling constant is effectively $O(\varepsilon^{\eta^{-1}})$ and the analyticity condition is expected to be $\varepsilon^{\eta^{-1}} C(\eta)^{-q}$ small (for some $q > 0$, determined as in the discussion in Remark 5.16, item (4), below). Hence the analyticity condition will be $\varepsilon C(\eta)^{-q\eta}$ small rather than the far worse $\varepsilon C(\eta)^{-q}$ small, that is implied directly from lemma 1 in [CG] (where $q = 6$ is an estimate).

In the one degree of freedom case the corresponding problem is studied in [N]: it is a problem that arises naturally in the context of Nekhoroshev theory. In our case the rotation vector is not one-dimensional, so that the cancellations between resonances typical of small divisors problems, [E,G1,Ge1,Ge2], have to be exploited in order to prove convergence of the perturbative series. The fact that the two components of the rotation vector (1.3) are so different in scale has the consequence that small divisors can appear only at

[4] It can alternatively be easily seen from the explicit expressions for the stable and unstable manifolds equations derived in [G1], which also provide a general algorithm for constructing the function Φ as a convergent series in ε for ε small; see [G3].

high orders, so that the dependence of the radius convergence on the Diophantine constant $C(\eta)$ is highly improvable with respect the "naïve" one, as explained above: the proof of such an assertion is the subject of the present paper (as, in the weaker form, already of [GGM3]).

1.7. The paper is organized as follows. In §2,3,4 the formalism is concisely illustrated and the graphic representations of the whiskers in terms of tree graphs is exhibited (for systems more general than (1.1); see (2.1) below). The analysis is brief but selfcontained, with references to [G1,Ge2] only given for further insight and details. The basic formalism is in §2, then we work out in §3 two specific examples to explain the origin of the graphical interpretation, and in §4 we set up the general Feynman rules for evaluating the equations of the whiskers (and the splitting vector as a particular case) as a sum of quantities that can be elementarily evaluated. In §5 bounds are derived, assuring the convergence of the perturbative series defining the whiskers in the more general system in (2.1) below and leading to Theorem 1.4, when restricted to the Hamiltonian (1.1).

The bounds are derived along the lines of [G1,Ge2]: the main part is the derivation of the bounds for the part of the expansion corresponding to what we call the contributions due to "trees without leaves": this is done fully and self consistently in §5 and in the related appendices. Once the bounds on the contributions from trees without leaves are established, *which is the real difficulty*, the same analysis can be applied to bound the other contributions. Since this is simply reduced, without any further technical problems, to the case of contributions from the simpler trees with no leaves we do not repeat this part of the discussion which is done in [Ge2] following the corresponding analysis done in [G1,Ge1].

To Appendix A1 we relegate some technical details, while Appendix A3 concerns the cancellation analysis of [Ge2], needed in order to treat the small divisors problems, with more details with respect to the quoted paper. An original technical part is also in Appendix A2 and deals with the improvement of the dependence of the convergence radius on the Diophantine constant $C(\eta)$.

We do not comment here on the obvious relevance of the above results for the theory of Arnol'd diffusion: see [GGM3,GGM4].

§2. Lindstedt series for whiskered tori.

We use the formalism of [Ge2]: it would be pointless to repeat here the technical work required to motivate the necessity or usefulness of the notations, and we cannot imagine that any reader may have interest in the matter that follows unless he has some experience with Eliasson's method, as exposed in [E] and complemented in [G1,G2,Ge1,Ge2]. The references to [G1,Ge2] are given only to point at places where further details on the motivations of the assertions can be found.

The following analysis innovates [Ge2] in §5 because of the extension of Siegel-Bryuno's bound described in Appendix A2 below: this section and the next two provide a *self contained* description of the graphical algorithm exploited in §5 and Appendix A2.

2.1. In the following we shall consider a Hamiltonian ("Thirring model") more general than the one in (1.1), *i.e.* a Hamiltonian which couples a pendulum with $\ell-1$ rotators via a perturbation f_1 which is always an *even trigonometric polynomial*,

$$\mathcal{H}=\boldsymbol{\omega}\cdot\mathbf{A}+\frac{1}{2J}\mathbf{A}\cdot\mathbf{A}+\frac{I^2}{2J_0}+J_0g_0^2\,f_0(\varphi)+\varepsilon J_0g_0^2\,f_1(\varphi,\boldsymbol{\alpha})+J_0g_0^2\,\gamma(\varepsilon,g_0)\,f_0(\varphi)\;, \tag{2.1}$$

where $(\boldsymbol{\alpha},\mathbf{A})\in\mathbb{T}^{\ell-1}\times\mathbb{R}^{\ell-1}$, $(\varphi,I)\in\mathbb{T}^1\times\mathbb{R}^1$, $J_0>0$, J is a diagonal matrix, with $0<\det J\le+\infty$, and

$$f_1(\varphi,\boldsymbol{\alpha})=\sum_{\substack{n\in\mathbb{Z}\\|n|\le N_0}}\sum_{\substack{\nu\in\mathbb{Z}^{\ell-1}\\|\nu|\le N}}f^1_{\underline{\nu}}\,e^{i(\boldsymbol{\nu}\cdot\boldsymbol{\alpha}+n\varphi)}\;,\qquad f^1_{\underline{\nu}}=f^1_{-\underline{\nu}}\;, \tag{2.2}$$

$$f_0(\varphi) = (\cos\varphi - 1) = \sum_{\substack{|n|=1\\ \boldsymbol{\nu}=\mathbf{0}}} f^0_{\underline{\nu}} e^{i(\boldsymbol{\nu}\cdot\boldsymbol{\alpha}+n\varphi)} \;,$$

with $\underline{\nu} = (\nu_0, \boldsymbol{\nu}) \equiv (n, \boldsymbol{\nu}) \in \mathbb{Z} \times \mathbb{Z}^{\ell-1}$ and $|\boldsymbol{\nu}| = \sum_{j=1}^{\ell-1} |\nu_j|$; we prefer to consider the Hamiltonian (2.1) with ℓ arbitrary because the Lindstedt series analysis holds for any $\ell \ge 1$. So that the value $\ell = 3$ and the existence of three time scales will be used only to obtain the second bound in (5.13) below.

The last term in (2.1) could be put together with the free pendulum potential $J_0 g_0^2\,(\cos\varphi - 1)$ thus modifying the "gravity acceleration" g_0^2 into $(1+\gamma(\varepsilon, g_0))\, g_0^2$: the term with $\gamma(\varepsilon, g_0) = \sum_{k=1}^{\infty} \gamma_k(g_0)\varepsilon^k$ is added because we follow here the approach in [Ge2]. We show that, given $s, \Omega, \eta, \boldsymbol{\omega}'$, with η small enough and $\boldsymbol{\omega}'$ verifying the Diophantine condition in Theorem 1.4, then one can fix $\gamma(\varepsilon, g_0)$ so that, for $|\varepsilon| < O(\eta^2)$, there is an invariant torus with average (over time) action $\mathbf{A}'$, with the properties in Theorem 1.4 and with Lyapunov exponent g_0 and rotation $\boldsymbol{\omega}' \equiv \boldsymbol{\omega} + J^{-1}\mathbf{A}'$. In other words by adding a *counterterm* to the Hamiltonian (1.1) one gets a new Hamiltonian system, (2.1), with an invariant torus with rotation $\boldsymbol{\omega}'$ *and Lyapunov exponent exactly equal to the prefixed* g_0 (see also §2.7 below).

We further show that, fixed $\boldsymbol{\omega}'$, $\gamma(\varepsilon, g_0)$ is jointly analytic in ε, g_0, if g_0 varies near a prefixed $\overline{g}_0 > 0$. Going back to the original Hamiltonian (1.1) we *therefore* set $g_0^2 = \overline{g}_0^2\,(1 + \gamma(\varepsilon, \overline{g}_0))$ and we can invert the latter relation as $\overline{g}_0^2 = (1 + \Gamma(\varepsilon, g_0))\, g_0^2$ for ε small enough (this will mean: for $|\varepsilon| < O(\eta^2)$). Hence by interpreting g_0^2 in (1.1) as $\overline{g}_0^2\,(1 + \gamma(\varepsilon, \overline{g}_0))$, so that (1.5) holds, we obtain Theorem 1.4 as a corollary of the above statements.

Of course a similar proof could be done without first fixing the Lyapunov exponent $\overline{g}_0$ and then inverting the relation between the "dressed exponent" $\overline{g}_0$ and the "bare" one g_0. But it is well known, from the analogous problem in renormalization theory, that it is wiser technically and conceptually to work, in perturbation theory, with prefixed physical quantities (*i.e.* dressed ones). The idea that perturbation theory would be simpler, in the technical estimates, is the key idea beyond [G1, Ge1] that is introduced in [Ge2].

2.2. From now on let us denote by $\boldsymbol{\alpha}$ the initial value of the rotators angles (*i.e.* at time $t = 0$). We define by $X_j^\sigma(t; \boldsymbol{\alpha})$, $j = 0, \ldots, 2\ell - 1$, the values of the variables at time t that are reached from initial data $X^\sigma(0; \boldsymbol{\alpha}) = (\pi, \boldsymbol{\alpha}, I^\sigma(0; \mathbf{0}), \mathbf{A}^\sigma(0; \mathbf{0}))$, with the given $\boldsymbol{\alpha}$, with $\varphi = \pi$ and with $I, \mathbf{A}$ such that $X^\sigma(0; \boldsymbol{\alpha})$ is on the stable ($\sigma = +$) or unstable ($\sigma = -$) manifolds of the invariant torus that we are searching for; the convention on the labels of X is that

$$\begin{aligned} &X_0^\sigma = \varphi^\sigma \;; \qquad X_j^\sigma = \alpha_j^\sigma \;, \qquad \text{for } 1 < j < \ell \;;\\ &X_\ell^\sigma = I^\sigma \;; \qquad X_j^\sigma = A_j^\sigma \;, \qquad \text{for } \ell < j < 2\ell \;. \end{aligned} \tag{2.3}$$

All functions in (2.3) depend on t and $\boldsymbol{\alpha}$ (the symbols $I^\sigma(t; \boldsymbol{\alpha})$ and $\mathbf{A}^\sigma(t; \boldsymbol{\alpha})$ should not be confused with $I^\sigma(\varphi, \boldsymbol{\alpha})$ and $\mathbf{A}^\sigma(\varphi, \boldsymbol{\alpha})$ defined in (1.2): in the following no ambiguity can arise as the quantities I^σ and $\mathbf{A}^\sigma$ *will be used always with the meaning in* (2.3) and as functions of $t, \boldsymbol{\alpha}$).

This is a parameterization of the stable and unstable manifolds in terms of $\boldsymbol{\alpha}, t$ where $\boldsymbol{\alpha}$ is the value of the angular coordinates at the moment in which $\varphi = \pi$, and t is the time elapsed since. The parameterization is different from the one in terms of $\boldsymbol{\alpha}, \varphi$ in (1.2) unless, of course, it is $\varphi = \pi$ and correspondingly $t = 0$. Hence the splitting vector (1.7) at $\varphi = \pi$ can also be written $Q_j(\boldsymbol{\alpha}) = X_j^+(0; \boldsymbol{\alpha}) - X_j^-(0, \boldsymbol{\alpha})$, $j = \ell + 1, \ldots, 2\ell - 1$. Note that we do not need to consider explicitly the splitting in the I–coordinates because, by energy conservation, they are functions of $\varphi, \boldsymbol{\alpha}, \mathbf{A}^\pm$.

Let $\mathbf{A}'$ be given and let $\boldsymbol{\omega}'$ in (1.3) be Diophantine with constants $C = C(\eta), \tau > 0$; see (1.4). We look for an invariant torus and for its stable and unstable manifolds with the property that the quasi periodic rotation on the torus takes place at velocity $\boldsymbol{\omega}'$ and, *at the same time*, the action variables oscillate with an average position $\mathbf{A}'$.

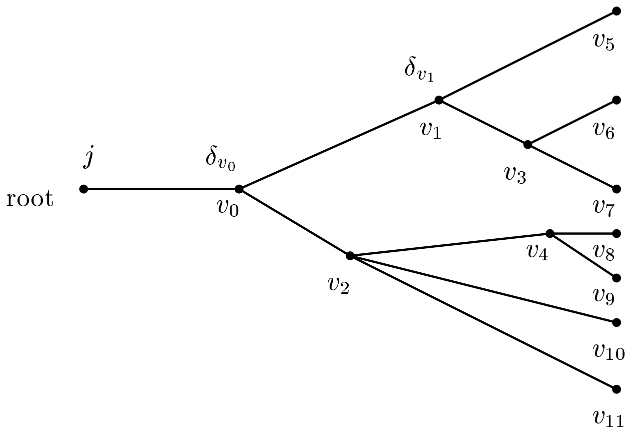
• Before proceeding we remark that the above *two* requirements may seem contradictory as there may seem to be no reason for being able to prescribe simultaneously the "spectrum" $\boldsymbol{\omega}'$ and the "average action" $\mathbf{A}'$

of the invariant tori. In fact this property of "*twistless*" motion on the tori or of "*absence of torsion*" is very remarkable (see the Remark 1.2 and [G1]): it will appear as due to the special symmetries of the system (2.1) and to the separation of the energy into a quadratic part involving actions only and an angular part involving only the angles.

Note also that we could confine ourselves to study the torus with average position $\mathbf{A}'=\mathbf{0}$, as in [G1,Ge2], because any torus can be reduced to that one through a trivial canonical transformation (a translation in the action variables). This explains why in the quoted papers only the torus covered with rotation vector $\boldsymbol{\omega}$ is explicitly considered: however in the following we consider also $\mathbf{A}'\neq\mathbf{0}$, as we are interested in showing the abundance of such tori in phase space (see the Remark 1.5).

The quantity $X_j^\sigma(t;\boldsymbol{\alpha})$ can be graphically represented as sum of *values* which can be associated with tree graphs, that we shall call "Feynman graphs" or "trees" *tout court*, see Fig.(2.4) below. The trees are partially ordered sets of points, called *nodes*, connected by unit lines, called *branches*, and they are "oriented" towards a point called *root*, which is reached by a single branch of the tree. Given two nodes v and w of a tree, we say that w precedes v ($w\le v$) if there is a path connecting w to v, oriented from w to v. With an abuse of notations we shall sometimes consider a tree as the collection of its nodes, sometimes as the collection of its branches and sometimes as the collection of both nodes and branches. The root *will not* be considered a node.

A typical tree considered below can be drawn as in Fig.(2.4): the labels meaning and the caption of such a drawing (which has to be interpreted as a mathematical formula) will be elucidated in the coming sections.



(2.4)

A tree ϑ with $m=12$, and some labels. The line numbers, distinguishing the lines, and their orientation pointing at the root, are not shown. The lines length should be the same but it is drawn of arbitrary size. The nodes labels δ_v are indicated only for two nodes.

The branch starting at the node v and linking it to the uniquely determined next node (or to the root), which we call v', will be denoted by λ_v: there is a unique correspondence between nodes and branches starting at them. We shall say that λ_v exits from v and enters v'; given a node v we shall say that a branch λ *pertains* to v if either λ enters v or λ exits from v; *e.g.* in Fig.(2.4) the line $v_1v_0\equiv\lambda_{v_1}$ "exits" v_1 and "enters" v_0, hence it pertains to both.

In [G1] two expansions are considered for the functions $X_j^\sigma(t;\boldsymbol{\alpha})$ representing the stable and unstable manifolds: one of them is used to exhibit cancellations taking place at all orders in the sums that express the coefficients of the power series in ε of the splitting vector, [G1,BCG,GGM2]; it is somewhat more involved than the other one that is convenient to just discuss convergence of the perturbation series for the splitting vector and that we shall use here. This is the reason why (as in [Ge2]) we shall not have trees whose lowest nodes carry a graphical decoration called *form factor*, or *fruit* in [G1,GGM2]. Nevertheless some of the nodes will still have a particular structure: to characterize them we introduce, below as in [Ge2], the notion of "*leaf*", which is related to the notion of fruit in [G1], from which it differs (and it, even, differs slightly from the similar notion of leaf in [Ge2]), see below for the motivation of the name.

2.3. As mentioned the drawing Fig.(2.4) has to be regarded as a mathematical formula expressing a function of the labels and of the topological structure of the trees. We now prepare the notation for the definition of "value" of a tree (following [Ge2]) (see [G1] for a simpler case): the derivation is not difficult but somewhat long and unusual for the subject (the breakthrough work [E] still does not seem to be well known in its technical aspects!). We discuss it in detail not only for completeness but in the attempt to clarify a construction that has generated quite a few new results starting from the work of [E], see [G1,GGM2,BGGM].

Let us consider the unperturbed motion $X^0(t) \equiv (\varphi^0(t), \boldsymbol{\alpha} + \boldsymbol{\omega}'t, I^0(t), \mathbf{A}')$, where $(\varphi^0(t), I^0(t))$ is the separatrix motion, generated by the pendulum in (2.1) starting at $t = 0$ in $\varphi = \pi$, $\mathbf{A} = \mathbf{A}', I = -2J_0 g_0$, so that $\varphi^0(t) = 4 \arctan e^{-g_0 t}$. Let $X^\sigma(t;\alpha)$, $\sigma = \operatorname{sign} t = \pm$, be the evolution, under the flow generated by (1.1), of the point on W^σ_ε which at time $t = 0$ is $(\pi, \boldsymbol{\alpha}, I^\sigma(\boldsymbol{\alpha}, \pi), \mathbf{A}^\sigma(\boldsymbol{\alpha}, \pi))$, see (1.2); let

$$X^\sigma(t) \equiv X^\sigma(t;\boldsymbol{\alpha}) \equiv \sum_{h\ge 0} X^{h\sigma}(t;\boldsymbol{\alpha})\varepsilon^h = \sum_{h\ge 0} X^{h\sigma}(t)\varepsilon^h, \qquad \sigma = \pm\ , \tag{2.5}$$

be the power series in ε of X^σ, (which we want to show to be convergent for ε small); note that $X^{0\sigma} \equiv X^0$ is the unperturbed whisker. We shall often omit writing explicitly the $\boldsymbol{\alpha}$ variable among the arguments of various $\boldsymbol{\alpha}$-dependent functions, to simplify the notations, and we shall regard the two functions $X^{h\sigma}(t)$, as forming a single function $X^h(t)$, which is $X^{h+}(t)$ if $\sigma = +$, $t > 0$, and $X^{h-}(t)$ if $\sigma = -$, $t < 0$.

Components of X will be labeled j, $j = 0, \ldots, 2\ell - 1$, consistently with (2.3), with the convention that $X_0 \stackrel{def}{=} X_-$ describes the coordinate φ, $(X_j)_{j=1,\ldots,\ell-1} \stackrel{def}{=} \mathbf{X}_\downarrow$ describes the $\boldsymbol{\alpha}$ coordinates, $X_\ell \stackrel{def}{=} X_+$ describes the I coordinate and $(X_j)_{j=\ell+1,\ldots,2\ell-1} \stackrel{def}{=} \mathbf{X}_\uparrow$ describes the $\mathbf{A}$ coordinates,

$$X \stackrel{def}{=} (X_j)_{j=0,\ldots,2\ell-1} \stackrel{def}{=} (X_-, \mathbf{X}_\downarrow, X_+, \mathbf{X}_\uparrow)\ , \tag{2.6}$$

i.e. we write first the angle and then the action components, first the pendulum and then the rotators. The $\uparrow$ ("*up*") and $\downarrow$ ("*down*") labels recall that the components with labels $\downarrow$ $(0 < j < \ell)$ have "lower" index than the variables with labels $\uparrow$ $(\ell < j)$, which have a "higher" index (a mnemonically useful fact, on first reading at least).

Inserting (2.5) into the Hamilton equation associated with (2.1) we get that the coefficients $X^{h\sigma}(t)$, $h \ge 1$, satisfy the hierarchy of linear equations

$$\frac{d}{dt} X^{h\sigma}(t) = L(t) X^{h\sigma}(t) + F^{h\sigma}(t)\ , \tag{2.7}$$

with $F^{h\sigma}(t)$ a 2ℓ–vector and the $2\ell \times 2\ell$–matrix $L(t)$ is

$$L(t) = \begin{pmatrix} 0 & \mathbf{0} & J_0^{-1} & \mathbf{0} \\ \mathbf{0} & 0 & \mathbf{0} & J^{-1} \\ g_0^2 J_0 \cos\varphi^0(t) & \mathbf{0} & 0 & \mathbf{0} \\ \mathbf{0} & 0 & \mathbf{0} & 0 \end{pmatrix} \tag{2.8}$$

For instance, $F^{1\sigma}(t)$ is a 2ℓ–vector with the first $0, \ldots, \ell - 1$ components vanishing (a consequence of the assumption that the perturbation only depends on the angular variables), with the ℓ–th component equal to $-J_0 g_0^2 \partial_\varphi f_1(\varphi^0(t), \boldsymbol{\alpha} + \boldsymbol{\omega}'t) + J_0 g_0^2 \gamma_1(g_0)\sin(\varphi^0(t))$ and with the remaining components equal to $-J_0 g_0^2 \partial_{\boldsymbol{\alpha}} f_1(\varphi^0(t), \boldsymbol{\alpha} + \boldsymbol{\omega}'t)$.

In general $F^{h\sigma}$ depends upon $X^0, \ldots, X^{h-1\sigma}$ *but not on* $X^{h\sigma}$. The entries of the $(2\ell \times 2\ell)$ matrix L have different meaning according to their position: the $\mathbf{0}$'s in the first and third row are $(\ell - 1)$–(row)–vectors, the $\mathbf{0}$'s in the first and third column are $(\ell-1)$–(column)–vectors, and the 0's and J^{-1} in the second and fourth column are $(\ell-1)\times(\ell-1)$–matrices, while the 0's in the first and third columns are scalars (as J_0^{-1} is). The perturbed motions will be described by *dimensionless* quantities Ξ, Φ:

$$\begin{aligned} X_j^{h\sigma} &= \Xi_j^{h\sigma}, \qquad 0\le j\le \ell-1, \qquad & X_j^{h\sigma} &= J_0 g_0\, \Xi_j^{h\sigma}, \qquad \ell\le j\le 2\ell-1\,,\\ \mathbf{F}_\uparrow^{h\sigma} &= J_0 g_0^2\, \boldsymbol{\Phi}_\uparrow^{h\sigma}, & F_+^{h\sigma} &= J_0 g_0^2\, \Phi_+^{h\sigma}\,, \end{aligned} \tag{2.9}$$

The simple form of the Hamiltonian equations for $\varphi, \boldsymbol{\alpha}$, namely $\dot\varphi = J_0^{-1} I$, $\dot{\boldsymbol{\alpha}} = \boldsymbol{\omega} + J^{-1}\mathbf{A}$ implies that $\Phi_j^{h\sigma} = F_j^{h\sigma} \equiv 0$, for $j=0,\ldots,\ell-1$. For instance

$$\Phi^{1\sigma} = \big(0, \mathbf{0}, -\partial_\varphi f_1(\varphi^0(t), \boldsymbol{\alpha}+\boldsymbol{\omega}' t) + \gamma_1(g_0)\sin(\varphi_0(t)), -\partial_{\boldsymbol{\alpha}} f_1(\varphi^0(t), \boldsymbol{\alpha}+\boldsymbol{\omega}' t)\big)\ . \tag{2.10}$$

Given the form of $L(t)$ and the vanishing of the first ℓ components $F_-^{h\sigma}$, $\mathbf{F}_\downarrow^{h\sigma}$ of $F^{h\sigma}$, for $h\ge 1$, the above hierarchy of equations (determining the stable and unstable manifolds) takes the form

$$\begin{aligned} \frac{1}{g_0}\frac{d}{dt}\Xi_+^{h\sigma} &= \cos\varphi^0\, \Xi_-^{h\sigma} + \Phi_+^{h\sigma}\,, \qquad & \frac{1}{g_0}\frac{d}{dt}\boldsymbol{\Xi}_\uparrow^{h\sigma} &= \boldsymbol{\Phi}_\uparrow^{h\sigma}\,,\\ \frac{1}{g_0}\frac{d}{dt}\Xi_-^{h\sigma} &= \Xi_+^{h\sigma}\,, & \frac{1}{g_0}\frac{d}{dt}\boldsymbol{\Xi}_\downarrow^{h\sigma} &= J_0 J^{-1}\boldsymbol{\Xi}_\uparrow^{h\sigma}\,. \end{aligned} \tag{2.11}$$

And, for all $h\ge 1$, we can easily write (via Taylor expansion and order matching) the following formula for $\Phi^{h\sigma}$ in terms of the coefficients $\Xi^0, \ldots, \Xi^{h-1\sigma}$ and of the derivatives of f_0 and $f_1\equiv f$, see (2.2). The first ℓ components of $\Phi^{h\sigma}$ vanish, as said above, $\Phi_-^{h\sigma}\equiv 0$, $\boldsymbol{\Phi}_\downarrow^{h\sigma}\equiv \mathbf{0}$, and

$$\begin{aligned} \boldsymbol{\Phi}_\uparrow^{h\sigma} &= -\sum_{|\underline{m}|\ge 0} (\partial_{\boldsymbol{\alpha}} f_1)_{\underline{m}}(\varphi^0, \boldsymbol{\alpha}+\boldsymbol{\omega}' t) \sum_{(h_j^i)_{\underline{m},h-1}} \prod_{i=0}^{\ell-1}\prod_{j=1}^{m_i} \Xi_i^{h_j^i\sigma}\,,\\ \Phi_+^{h\sigma} &\equiv -\sum_{|\underline{m}|\ge 2} (\partial_\varphi f_0(\varphi))_{\underline{m}}(\varphi^0) \sum_{(h_j^0)_{\underline{m},h}} \prod_{j=1}^{m_0} \Xi_-^{h_j^0\sigma} +\\ &\quad -\sum_{p=1}^{h}\sum_{|\underline{m}|\ge 0} \gamma_p(g_0)\,(\partial_\varphi f_0(\varphi))_{\underline{m}}(\varphi^0) \sum_{(h_j^0)_{\underline{m},h-p}} \prod_{j=1}^{m_0} \Xi_-^{h_j^0\sigma} +\\ &\quad -\sum_{|\underline{m}|\ge 0} (\partial_\varphi f_1)_{\underline{m}}(\varphi^0, \boldsymbol{\alpha}+\boldsymbol{\omega}' t) \sum_{(h_j^i)_{\underline{m},h-1}} \prod_{i=0}^{\ell-1}\prod_{j=1}^{m_i} \Xi_i^{h_j^i\sigma}\,, \end{aligned} \tag{2.12}$$

where $(G)_{\underline{m}}(\cdot)$, with $G\in\{\partial_\varphi f_0, \partial_{\boldsymbol{\alpha}} f_1, \partial_\varphi f_1\}$, and $(h_j^i)_{\underline{m},q}$, with $h_j^i\ge 1$, are defined as

$$\begin{aligned} &(G)_{\underline{m}}(\cdot) \equiv \Big(\frac{\partial_\varphi^{m_0}\partial_{\alpha_1}^{m_1}\ldots\partial_{\alpha_{\ell-1}}^{m_{\ell-1}} G}{m_0!\, m_1!\,\ldots\, m_{\ell-1}!}\Big)(\cdot)\,,\\ &(h_j^i)_{\underline{m},q} \equiv (h_1^0, \ldots, h_{m_0}^0, h_1^1, \ldots, h_{m_1}^1, \ldots, h_1^{\ell-1}, \ldots, h_{m_{\ell-1}}^{\ell-1}) \qquad \text{with} \qquad \sum_{i=0}^{\ell-1}\sum_{j=1}^{m_i} h_j^i = q\,, \end{aligned} \tag{2.13}$$

and $m_i\ge 0$, $\underline{m} = (m_0,\ldots,m_{\ell-1})$, $|\underline{m}| = \sum_{i=0}^{\ell} m_i$. Note that the first two sums in the expression for $\Phi_+^{h\sigma}$ can only involve vectors $\underline{m}$ with $m_j = 0$ if $j\ge 1$ (so that $|\underline{m}| = m_0$), because the function f_0, see (2.2), depends only on φ and not on $\boldsymbol{\alpha}$. The evolution of Ξ^h is determined by integrating (2.8), if the initial data are known. The $h=1$ case requires a suitable interpretation of the symbols, given explicitly by (2.10).

Elementary quadrability of the free pendulum equations on the separatrix leads to the following expression for the "*Wronskian matrix*" $W(t)$ of the separatrix motion for the pendulum appearing in (2.1), with initial data at $t=0$ given by $\varphi=\pi, I=-2g_0J_0$, *i.e.* $\Xi^0_+=-2$. The matrix

$$W(t)=\begin{pmatrix} w_{00}(t) & w_{0\ell}(t)\\ w_{\ell 0}(t) & w_{\ell\ell}(t)\end{pmatrix} \tag{2.14}$$

is defined to be the solution of the linearization of the free pendulum equation around the separatrix solution, with data $W(0)=1$ and with $J_0=1$ (*because we use dimensionless solutions* Ξ, see (2.11)):

$$W(t)=\begin{pmatrix} \frac{1}{\cosh g_0t} & \frac{\bar w(t)}{4}\\ -\frac{\sinh g_0t}{\cosh^2 g_0t} & \Big(1-\frac{\bar w(t)}{4}\frac{\sinh g_0t}{\cosh^2 g_0t}\Big)\cosh g_0t\end{pmatrix},\qquad \bar w(t)\equiv\frac{2g_0t+\sinh 2g_0t}{\cosh g_0t}\ . \tag{2.15}$$

The evolution of the I,φ components, *i.e.* $\Xi^{h\sigma}_j$ with $j=0,\ell$ (also identified with the components with subscripts $\pm$, see (2.6)) can be determined from $W(t)$, by integrating (2.7) for the 0 and ℓ components, to be

$$\begin{pmatrix}\Xi^{h\sigma}_-\\ \Xi^{h\sigma}_+\end{pmatrix}=W(t)\begin{pmatrix}0\\ \Xi^{h\sigma}_+(0)\end{pmatrix}+W(t)\int_0^{g_0t}W^{-1}(\tau)\begin{pmatrix}0\\ \Phi^{h\sigma}_+(\tau)\end{pmatrix}\,d\,g_0\tau\ . \tag{2.16}$$

Thus, denoting by w_{ij} $(i,j=0,\ell)$ the entries of $W(t)$, (2.16) becomes, for $h\ge1$,

$$\begin{aligned}\Xi^{h\sigma}_-(t)&=w_{0\ell}(t)\Big(\Xi^{h\sigma}_+(0)+\int_0^{g_0t}w_{00}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g_0\tau\Big)-w_{00}(t)\int_0^{g_0t}w_{0\ell}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g_0\tau\ ,\\ \Xi^{h\sigma}_+(t)&=w_{\ell\ell}(t)\Big(\Xi^{h\sigma}_+(0)+\int_0^{g_0t}w_{00}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g_0\tau\Big)-w_{\ell0}(t)\int_0^{g_0t}w_{0\ell}(\tau)\Phi^{h\sigma}_+(\tau)\,d\,g_0\tau\ ,\end{aligned} \tag{2.17}$$

having used that $\Xi^{h,\sigma}_-(0)=0$ because the initial datum for φ is fixed and ε–independent. Likewise integration of the equations (2.11) for the $\uparrow,\downarrow$ components yields, for $h\ge1$,

$$\begin{aligned}\boldsymbol{\Xi}^{h\sigma}_\downarrow(t)&=J^{-1}J_0\Big[g_0t\Big(\boldsymbol{\Xi}^{h\sigma}_\uparrow(0)+\int_0^{g_0t}\boldsymbol{\Phi}^{h\sigma}_\uparrow(\tau)\,d\,g_0\tau\Big)-\int_0^{g_0t}g_0\tau\,\boldsymbol{\Phi}^{h\sigma}_\uparrow(\tau)\,d\,g_0\tau\Big]\ ,\\ \boldsymbol{\Xi}^{h\sigma}_\uparrow(t)&=\Big(\boldsymbol{\Xi}^{h\sigma}_\uparrow(0)+\int_0^{g_0t}\boldsymbol{\Phi}^{h\sigma}_\uparrow(\tau)\,d\,g_0\tau\Big)\ ,\end{aligned} \tag{2.18}$$

having used that the $\boldsymbol{\Xi}^{h\sigma}_\downarrow(0)\equiv\mathbf{0}$ because the initial datum for $\boldsymbol{\alpha}$ is fixed and ε–independent. The equations (2.17), (2.18) can be used to find a reasonably simple algorithm to represent the whiskers equations to all orders $h\ge1$ of the perturbation expansion.

2.4. The initial data in (2.17), (2.18) have to be *determined by imposing that the solutions (to all orders) become quasi periodic* as $t\to\sigma\infty$. This is quite easy and (as to be expected) this condition is simply that $\Xi^{h\sigma}_+(0),\boldsymbol{\Xi}^{h\sigma}_\uparrow(0)$ are determined by imposing that the integrals in parentheses become integrals between $\sigma\infty$ and t, *i.e.* $\Xi^{h\sigma}_+(0)=\int_{\sigma\infty}^0\ldots$ and $\boldsymbol{\Xi}^{h\sigma}_\uparrow(0)=\int_{\sigma\infty}^0\ldots$; see below.

However the latter integrals are no longer necessarily convergent properly (a few examples suffice to see this); hence one has to go carefully through the process of imposing the correct asymptotic behavior in order to see what is the meaning to be given to such integrals $\int_{\sigma\infty}^t$. The analysis can be found in [G1, Ge2]. The result is that all expressions under integral sign can be written as sums of functions that are rather special, namely

$$M(t)=\sigma^\chi\frac{(\sigma g_0t)^j}{j!}e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}t-pg_0\sigma t}\ , \tag{2.19}$$

with $\chi, j, \boldsymbol{\nu}, p$ integers and $p \ge -1$ (see below), so that one has

$$\Xi^{h\sigma}(t) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^{\ell-1}} \sum_{p=-1}^{\infty} \tilde\Xi^{h\sigma}(\boldsymbol{\nu},p)\, e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}t - pg_0\sigma t} \;, \qquad \Phi^{h\sigma}(t) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^{\ell-1}} \sum_{p=-1}^{\infty} \tilde\Phi^{h\sigma}(\boldsymbol{\nu},p)\, e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}t - pg_0\sigma t} \;, \tag{2.20}$$

where we explicitly write down only the dependence on $\boldsymbol{\nu}$ and p (clearly also the fixed constants like $J, J_0, g_0, \ldots$ enter).

The series turn out to be convergent for $\sigma t > 0$; however their sums have *no singularity* at $t = 0$ and can be anaytically continued for $\sigma t < 0$ (*i.e.* $x \ge 1$). More precisely the functions that one has to integrate are contained in an *algebra* $\hat{\mathcal{M}}$ on which the integration operations that we need can be given a meaning.

DEFINITION ([G1]). *Let $\hat{\mathcal{M}}$ be the space of the functions of t which can be represented, for some $k \ge 0$, as*

$$M(t) = \sum_{j=0}^{k} \frac{(\sigma t g_0)^j}{j!} M_j^\sigma(x, \boldsymbol{\omega} t) \;, \quad x \equiv e^{-\sigma g_0 t} \;, \quad \sigma = \operatorname{sign} t \;, \tag{2.21}$$

with $M_j^\sigma(x,\boldsymbol{\psi})$ a trigonometric polynomial in $\boldsymbol{\psi}$ with coefficients holomorphic in the x-plane in the annulus $0 < |x| < 1$, with 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 $< \frac{\pi}{2}$, and possible polar singularities at $x = 0$. The smallest cone containing the singularities will be called the singularity cone of M.

The proper interpretation of the improper integrals $\int_{\sigma\infty}^{g_0 t} M(\tau) dg_0\tau$, which henceforth will be denoted by $\fint_{\sigma\infty}^{g_0 t} M(\tau) dg_0\tau$, is simply the *residuum* at $R = 0$ of the analytic function

$$\mathcal{I}_R M \stackrel{def}{=} \int_{\sigma\infty+i\theta}^{g_0 t} e^{-Rg_0\sigma z} M(z)\, d\, g_0 z \;, \tag{2.22}$$

(where θ is arbitrarily prefixed) which is defined and holomorphic for $\operatorname{Re} R > 0$ and large enough, *i.e.*

$$\mathcal{I}M(t) \equiv \fint_{\sigma\infty}^{g_0 t} dg_0\tau\, M(\tau) \stackrel{def}{=} \oint \frac{dR}{2\pi i R} \mathcal{I}_R M(t) \;. \tag{2.23}$$

By linear extension this defines the integration of function in $\hat{\mathcal{M}}$ for $|x| < 1$. The analyticity in x around $x = \pm 1$ and the remarks that $\frac{d}{dg_0 t}\mathcal{I}M(t) \equiv M(t)$, *i.e.* $\mathcal{I}M(t) \equiv \mathcal{I}M(t') + \int_{g_0 t'}^{g_0 t} d\, g_0\tau\, M(\tau)$, so that $\mathcal{I}M(t)$ is a special primitive of $M(t)$ (at fixed σ), allow us to analytically continue the result of the integration to a function in $\hat{\mathcal{M}}$. The operator $\mathcal{I}$ maps the algebra $\hat{\mathcal{M}}$ into itself because one checks that on the monomial (2.19) one has

$$\mathcal{I}M(t) = \begin{cases} -g_0^{-1}\sigma^{\chi+1} e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}t - pg_0\sigma t} \sum_{h=0}^{j} (g_0\sigma t)^{j-h} \frac{1}{(j-h)!} \frac{1}{(p - i\sigma g_0^{-1}\boldsymbol{\omega}'\cdot\boldsymbol{\nu})^{h+1}} \;, & \text{if } |p| + |\boldsymbol{\nu}| > 0 \;, \\ g_0^{-1}\sigma^{\chi+1} \frac{(\sigma g_0 t)^{j+1}}{(j+1)!} \;, & \text{otherwise} \;, \end{cases} \tag{2.24}$$

showing, in particular, that the radius of convergence in x of $\mathcal{I}M$, for a general M, is the same as that of M. But in general the singularities will not be polar, even when those of the M_j^σ's were such.

We shall see that the cases $|p| + |\boldsymbol{\nu}| = 0$ do not enter in the discussion (a feature of the method of [Ge2]). The complete expression of $X^{h\sigma}(t)$ becomes

$$\begin{aligned} &\Xi_-^{h\sigma}(t) = w_{0\ell}(t)\mathcal{I}(w_{00}\Phi_+^{h\sigma})(t) - w_{00}(t)\big(\mathcal{I}(w_{0\ell}\Phi_+^{h\sigma})(t) - \mathcal{I}(w_{0\ell}\Phi_+^{h\sigma})(0^\sigma)\big) \stackrel{def}{=} \mathcal{O}(\Phi_+^{h\sigma})(t) \;, \\ &\boldsymbol{\Xi}_\downarrow^{h\sigma}(t) = J^{-1}J_0\Big(\mathcal{I}^2(\boldsymbol{\Phi}_\uparrow^{h\sigma})(t) - \mathcal{I}^2(\boldsymbol{\Phi}_\uparrow^{h\sigma})(0^\sigma)\Big) \stackrel{def}{=} \overline{\mathcal{I}}^2(\boldsymbol{\Phi}_\uparrow^{h\sigma}(t)) \;, \\ &\Xi_+^{h\sigma}(t) = w_{\ell\ell}(t)\mathcal{I}(w_{00}\Phi_+^{h\sigma})(t) - w_{\ell 0}(t)\Big(\mathcal{I}(w_{0\ell}\Phi_+^{h\sigma})(t) - \mathcal{I}(w_{0\ell}\Phi_+^{h\sigma})(0^\sigma)\Big) \stackrel{def}{=} \mathcal{O}_+(\Phi_+^{h\sigma})(t) \;, \\ &\boldsymbol{\Xi}_\uparrow^{h\sigma}(t) = \mathcal{I}(\boldsymbol{\Phi}_\uparrow^{h\sigma})(t) \;, \end{aligned} \tag{2.25}$$

where $\mathcal{O}, \mathcal{O}_+, \overline{\mathcal{I}}^2$ are implicitly defined here (and $\mathcal{I}^2$ is $\mathcal{I}$ applied twice); and $\Xi^{h\sigma}, \Phi^{h\sigma} \equiv (0, \mathbf{0}, \Phi_+^{h\sigma}, \boldsymbol{\Phi}_\uparrow^{h\sigma})$ are introduced in (2.9). While $\Xi^{h\sigma}$ has non zero components over both the *angle* ($j = 0, \ldots, \ell - 1$) and over the *action* ($j = \ell, \ldots, 2\ell - 1$) components, the $\Phi^{h\sigma}$ has, as already noted, only the action directions non zero; the notation 0^σ means the limit as $t \to 0$ from the left ($\sigma = -$) or from the right ($\sigma = +$), but below we shall drop the superscript on 0 (always clear from he context because it is the same as the superscript σ of the functions $\Xi^{h\sigma}$). Furthermore, with the definitions (2.20) of $\tilde{\boldsymbol{\Phi}}_\uparrow^{h\sigma}(\boldsymbol{\nu}, p)$ one finds also the property (with the notations in (2.1))

$$\tilde{\boldsymbol{\Phi}}_\uparrow^{h\sigma}(\mathbf{0}, 0) = \mathbf{0} \;, \tag{2.26}$$

for all $h \geq 1$.

We shall repeatedly use that in order to compute $\Xi_j^{h\sigma}$ we only need $\Xi_{j'}^{h'\sigma}$ with $0 \leq j' < \ell$ (*i.e.* only $\Xi_+^{h'\sigma}, \Xi_\uparrow^{h'\sigma}$) and $h' < h$. This follows from (2.25) and (2.12): whether we want to compute an "*action component*" ($\Xi_j^{h\sigma}$, $j \geq \ell$) or an "*angle component*" ($\Xi_j^{h\sigma}$, $j < \ell$) of $\Xi^{h\sigma}$, we only need the angle components of lower orders, *i.e.* $\Xi_{j'}^{h'\sigma}$ with $h' < h$ and $j' < \ell$.

2.5. The linearity of the last of (2.25), together with (2.26) and the t–dependence of $\boldsymbol{\Phi}^{h\sigma}(t)$ in (2.20), implies that the prefixed value $\mathbf{A}'$ *has the interpretation of average action* of the quasi periodic motion on the invariant torus to which the trajectories that we study asymptote; see the third statement in Theorem 1.4. This corresponds to the identity of [CG] (see, in the latter reference, the first of (6.34) and its proof in Appendix A12) that follows from the symplectic structure of the equations of motion, according to a well known argument going back to Poincaré, [P], discussed also in [E,CZ]. It is a property that generated the qualification of "*twistless tori*" given in [G1] to such tori: the "dispersion relation" linking the frequencies to the average actions *does not change* or *is not twisted* when the perturbation is switched on. This is a property, established in the present context in (33) of [Ge2], that can be ultimately traced back to the fact that in the above models the twist condition is not needed for establishing a KAM theorem.

2.6. By combining (2.25) and (2.12), (2.13) (and recalling (2.9)) the representation in terms of trees is immediate; the integrals in (2.25) and the lower order X^h in (2.12) become *recursively* multiple (improper) integrals over dummy "time" variables.

In this operation each function $(-\partial_{\boldsymbol{\alpha}} f_1(\varphi^0(t), \boldsymbol{\alpha} + \boldsymbol{\omega}' t))_{\underline{m}}$ and $(-\partial_\varphi f_0(\varphi^0(t)))_{\underline{m}}$ is expanded as a linear combination of monomials $M(t)$ having the form $\sigma^\chi (\sigma g_0 t)^j (j!)^{-1} x^n e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu} t}$ with $x = e^{-g_0 \sigma t}$; see (2.19).

The form of (2.12) shows that the integrations occur in a hierarchical order: *hence one can describe them by a tree.* The integrands can be identified by attaching to each node of the tree suitably many labels. We shall first illustrate the construction of the trees via two examples (in §3 below): this can be useful in order to understand the general case (see also [G1,Ge2]).

2.7. We shall establish, also recursively, that $\Xi^{h\sigma}$ will be expanded in monomial like (2.19) with $j = 0$ and $p \geq 0$, so that at $t \to \pm\infty$ the quantities $\Xi^{h\sigma}$ will approach exponentially fast quasi periodic functions describing the motion on the invariant torus. The approach will be proportional to $e^{-g_0|t|}$ or to a higher power of this quantity. This, together with the remark that at order 0 (*i.e.* on the unperturbed motion) the approach is precisely proportional to $e^{-g_0|t|}$ (in the I, φ coordinates), will imply that at least for $j = 0, \ell$ (and "generically" also for the other coordinates)

$$\lim_{t\to\infty} \frac{1}{\sigma t} \log |\Xi_j^\sigma(t)|^{-1} = g_0 \;, \tag{2.27}$$

i.e. that the Lyapunov exponents of the torus are $\pm g_0$.

Before stating the general graphical rules to represent (2.25) in terms of explicitly performed integrals, we discuss in detail two examples: understanding them facilitates enormously, we think, the understanding of the general cases which will be exposed referring to the examples to make it more concrete.

§3. Two examples of the trees construction.

3.1. We discuss how to make more explicit (2.25) by performing two "third order" examples. The first order reduces trivially to the first order formulae (Mel'nikov integral); the second order is also a bit too simple and is left to the reader: the first two orders will be, of course, implicitly done below, because to compute the third order one needs the first and second too.

To third order the last line in (2.25) gives $\boldsymbol{\Xi}^{3\sigma}_{\uparrow}(t) = \mathcal{I}(\boldsymbol{\Phi}^{3\sigma}_{\uparrow})(t)$, where $\boldsymbol{\Phi}^{3\sigma}_{\uparrow}$ can be expressed through the first equation in (2.12), so that, for $j = \ell+1, \ldots, 2\ell-1$, one has

$$\begin{aligned}\Phi^{3\sigma}_j = &-\frac{1}{2}\partial_{\alpha_j}\partial^2_{\varphi} f_1\, \Xi^{1\sigma}_{-}\Xi^{1\sigma}_{-} - \partial_{\alpha_j}\partial_{\varphi} f_1\, \Xi^{2\sigma}_{-} - \sum_{p=1}^{\ell-1}\partial_{\alpha_j}\partial_{\alpha_p}\partial_{\varphi} f_1\, \Xi^{1\sigma}_{p}\Xi^{1\sigma}_{-} + \\ &-\frac{1}{2}\sum_{p,q=1}^{\ell-1}\partial_{\alpha_j}\partial_{\alpha_p}\partial_{\alpha_q} f_1\, \Xi^{1\sigma}_{p}\Xi^{1\sigma}_{q} - \sum_{p=1}^{\ell-1}\partial_{\alpha_j}\partial_{\alpha_p} f_1\, \Xi^{2\sigma}_{p}\ ,\end{aligned} \tag{3.1}$$

where $\Xi^{1\sigma}$ and $\Xi^{2\sigma}$ can be written by using once more (2.25) (the first two lines only as per the general remark in the last paragraph of §2.4).

We consider explicitly two contributions to $\boldsymbol{\Xi}^{3\sigma}_{\uparrow}(t)$. Recalling that $\sigma=+$ corresponds to the stable manifold and $\sigma=-$ to the unstable one, the first will be

$$\frac{1}{2}\int_{\sigma\infty}^{g_0 t} dg_0\tau_{v_0}(-\partial_{\alpha_j\alpha_p\alpha_q} f_1(\varphi^0(\tau_{v_0}), \boldsymbol{\alpha}+\boldsymbol{\omega}'\tau_{v_0})\ \Xi^{1\sigma}_{p}(\tau_{v_0})\,\Xi^{1\sigma}_{q}(\tau_{v_0})\ , \tag{3.2}$$

arising from the fourth contribution in the r.h.s. of (3.1). The contribution (3.2) can be written more explicitly, by using again the expression for $\boldsymbol{\Xi}^{1\sigma}_{\downarrow}$ in (2.25), as

$$\frac{1}{2}\int_{\sigma\infty}^{g_0 t} dg_0\tau_{v_0}\ \big(-\partial_{\alpha_j\alpha_p\alpha_q} f_1\big)\,(\tau_{v_0})\ \overline{\mathcal{I}}^2(-\partial_{\alpha_p} f_1(\tau_{v_1}))(\tau_{v_0})\ \overline{\mathcal{I}}^2(-\partial_{\alpha_p} f_1(\tau_{v_2}))(\tau_{v_0})\ , \tag{3.3}$$

where the $\overline{\mathcal{I}}^2$ operations involve, see (2.25), integrations over variables that we can call τ_{v_1}, τ_{v_2} and the derivatives of f_1 are evaluated at $(\varphi^0(\tau_{v_n}), \boldsymbol{\alpha}+\boldsymbol{\omega}'\tau_{v_n})$, $n=0,1,2$. Such variables have been indicated explicitly using the abbreviated notation (τ_{v_n}) and with a obvious abuses of notation (they should not appear at all, except τ_{v_0}, being dummy).

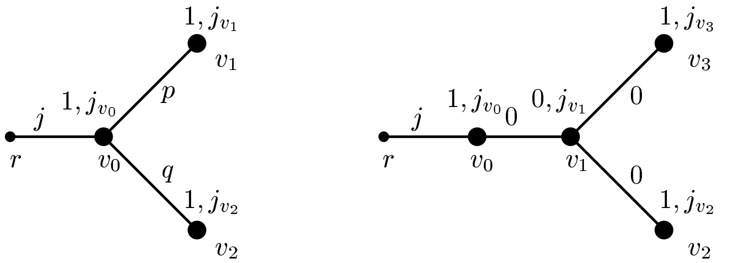
The second example is obtained by considering the contribution with $h^0_2 = 2$ from the first line of (2.12), *i.e.* the second contribution in the r.h.s. of (3.1),

$$\int_{\sigma\infty}^{g_0 t} \big(-\partial_{\alpha_j\varphi} f_1\big)\,(\tau_{v_0})\ \Xi^{2\sigma}_{-}(\tau_{v_0})\,d\tau_{v_0}\ , \tag{3.4}$$

still imagining the derivatives of f_1 evaluated at $(\varphi^0(\tau_{v_0}), \boldsymbol{\alpha}+\boldsymbol{\omega}'\tau_{v_0})$. This contribution will be the sum of several terms, because $\Xi^{2\sigma}_{-}(\tau_{v_0})$ has to be expressed by using (2.25) and (2.12). One of the (many) contributions will be

$$\frac{1}{2}\int_{\sigma\infty}^{g_0 t} dg_0\tau_{v_0}\ \big(-\partial_{\alpha_j\varphi} f_1\big)\,(\tau_{v_0})\,\mathcal{O}\Big(-\partial^3_{\varphi} f_0(\tau_{v_1})\,\mathcal{O}\big(-\partial_{\varphi} f_1(\tau_{v_2})\big)(\tau_{v_1})\,\mathcal{O}\big(-\partial_{\varphi} f_1(\tau_{v_3})\big)(\tau_{v_1})\Big)(\tau_{v_0})\ , \tag{3.5}$$

where the $\mathcal{O}$ operations involve, see (2.25), integrations over variables that we can call $\tau_{v_1}, \tau_{v_2}, \tau_{v_3}$ and the derivatives of f_0, f_1 are evaluated at $(\varphi^0(\tau_{v_n}), \boldsymbol{\alpha} + \boldsymbol{\omega}'\tau_{v_n})$, $n = 0, 1, 2, 3$. Such variables have been indicated explicitly with the same abuse of notation as above; and the dependence on τ of the derivatives of f_0, f_1 has again been simply denoted by adding the symbol (τ_{v_n}) instead of the full argument $(\varphi^0(\tau_{v_n}), \boldsymbol{\alpha} + \boldsymbol{\omega}'\tau_{v_n})$.
A complete representation of the above two contributions to $\Xi_j^{3\sigma}(t)$ is given, *with enormous notational simplification*, by the following trees:



(3.6)

where the labels on the nodes v are denoted δ_v, j_v and those on the lines λ_v are denoted j_{λ_v}.
The label $\delta_v = 0, 1$ on the node v indicates selection of f_{δ_v}, *i.e.* of f_0 or f_1, the label j_v denotes a derivative with respect to φ if $j_v = \ell$ or with respect to α_{j_v} if $j_v = \ell + 1, \ldots, 2\ell - 1$. For the label j_{λ_v} associated with the branch λ_v following v, one has $j_{\lambda_v} = j_v - \ell$ for all v except for the highest node v_0, for which one has $j_{\lambda_{v_0}} = j_{v_0}$. In the examples above, (3.3) and (3.5) correspond, respectively, to the first figure in (3.6) with $j_{v_0} = j, j_{v_1} = p + \ell, j_{v_2} = q + \ell$ and to the second with $j_{v_1} = j_{v_2} = j_{v_3} = \ell, j_{v_0} = j$, (hence $j_{\lambda_{v_1}} = j_{\lambda_{v_2}} = j_{\lambda_{v_3}} = 0$, $j_{v_0} = j$). In the examples the labels p, q correspond to $\partial_{\alpha_p}, \partial_{\alpha_q}$ in (3.3).

3.2. Remark. The exception for the meaning of $j_{\lambda_{v_0}}$ is convenient, in the above cases, as the integration over τ_{v_0} differs from the others: the inner ones evaluate $\Xi_j^{h\sigma}$ for $j = 0, \ldots, \ell - 1$, because the functions f_0, f_1 only depend on the angle variables (see the last paragraph in §2.4); the last integral, however, evaluates in the examples a component of $\boldsymbol{\Xi}_{\uparrow}^{h\sigma}$ (which is labeled $j = \ell + 1, \ldots, 2\ell - 1$), but, in general, j can be any value $j = 0, \ldots, 2\ell - 1$. Note that this is not so for the inner labels j_λ which must be angle labels $j_\lambda = 0, \ldots, \ell - 1$. So, in general, we shall have that the value of a tree with $j_{\lambda_{v_0}} = j$ contributes to $\Xi_j^{h\sigma}$.

§4. Trees and Feynman graphs approach to whiskers construction: the general case.

We now proceed to describe the general case.

4.1. To compute the splitting vector we only need to consider the variable t equal to 0. However we shall be also interested in $\Xi^{h\sigma}(\boldsymbol{\alpha}, t)$ for $\sigma t > 0$, for instance in order to study how fast the invariant torus is approached by the motions on its stable and unstable manifolds (to obtain its Lyapunov exponent). Hence it will be natural to attribute the label t to the root: this will also remind that the integral over τ_{v_0} has to be performed between $\sigma\infty$ and t, (the value $\sigma = -$ corresponds to the unstable manifold and the value $\sigma = +$ corresponds to the stable one). Since we shall *never* consider the stable manifold for $t > 0$ or the unstable for $t < 0$ the value of σ will be the same as that of the sign of t.
We shall be interested in computing not only $\mathbf{X}_{\uparrow}^{\sigma}(0; \boldsymbol{\alpha}) - \mathbf{A}'$ (or $\mathbf{X}_{\uparrow}^{\sigma}(t; \boldsymbol{\alpha}) - \mathbf{A}'$), as in [GGM2], but, more generally, $X^\sigma(t; \boldsymbol{\alpha}) - X^0(t; \boldsymbol{\alpha})$, with $\sigma = \operatorname{sign} t$, (here X^0 denotes the unperturbed motion).
In general the rules to express $X^\sigma(t; \boldsymbol{\alpha}) - X^0(t; \boldsymbol{\alpha})$ as sum of "values" associated with trees will be described now, assuming that the reader follows us by applying and checking them to the special cases (3.3), (3.5), illustrated in (3.6).

The reader might be helped in following the construction of the algorithm to express the stable and unstable manifolds below, by keeping in mind that we simply decompose the (quite involved and recursively defined by (2.25),(2.12)) expressions for the whiskers, so far obtained, *further.*

The purpose being of reducing their evaluation to *very elementary algebraic operations*: ultimately just products of simple factors associated with the nodes (and their labels) of a tree, that we shall call "coupling constants", and of factors associated with the branches (and their labels), that we shall call "propagators", each of which can be trivially evaluated and trivially bounded.

The reader familiar with Quantum Field Theory will realize the striking analogy between the algorithms discussed below and the *Feynman graphs*: in fact a "tree" will turn out as an analog to a (loopless) Feynman graph and *very likely it is* a Feynman graph of a suitable (non trivial) field theory. Our analysis amounts to a renormalization group analysis of it and it partially extends, to the case of the theory of the stable and unstable manifolds of hyperbolic tori in nearly integrale systems, the field theoretic interpretation already discussed in detail in previous works, see [GGM1] and appended references, in the study of KAM tori.

• To each node we attach an *order label* $\delta_v = 0, 1$, see Fig.(3.6), and a corresponding function f_{δ_v}: if a node v bears a label $\delta_v = 1$ the associated function is f_1 and if it bears a label $\delta_v = 0$ it is f_0.

• To each node v of a tree ϑ, see the Figure (2.4) above, we associate an integration *time variable* τ_v and an *integration operation*, which corresponds to $\overline{\mathcal{I}}^2$ or $\mathcal{O}$ if the node *is not the highest node* v_0 and to $\overline{\mathcal{I}}^2$ or $\mathcal{O}$ or $\mathcal{I}$ or $\mathcal{O}_+$ if the node v *is the highest, i.e.* $v = v_0$. This is so because in the first case (a "lower node") one must use the first two equations in (2.25) because in (2.12) only angle components of $X^{h\sigma}$ appear, while in the second case (that of the highest node) one can use all of (2.25) since we can evaluate either an angle coordinate $\Xi_j^{h\sigma}(\boldsymbol{\alpha}, t)$, $j < \ell$, or an action coordinate, $j \ge \ell$.

When $v < v_0$ the choice between the two possibilities will be marked by an *action label* j_v associated with each node: if $j_v = \ell$, $v < v_0$, then we choose $\mathcal{O}$, if $j_v = \ell + 1, \ldots, 2\ell - 1$, $v < v_0$, we choose $\overline{\mathcal{I}}^2$.

• When v is the highest node v_0, there are therefore more possibilities: to distinguish between them we use the *action label* j_{v_0} and the *branch label* $j_{\lambda_{v_0}}$, which can be equal either to j_{v_0} or to $j_{v_0} - \ell$. So when $v = v_0$ and $j_{v_0} = \ell$, we choose $\mathcal{O}$ if $j_{\lambda_{v_0}} = j_{v_0} - \ell = 0$ and $\mathcal{O}_+$ if $j_{\lambda_{v_0}} = j_{v_0} = \ell$, see (2.25), while when $v = v_0$ and $j_{v_0} > \ell$, we choose $\overline{\mathcal{I}}^2$ if $j_{\lambda_{v_0}} = j_{v_0} - \ell$ and $\mathcal{I}$ if $j_{\lambda_{v_0}} = j_{v_0}$, see (2.25).

As said in Remark 3.2, the meaning of the branch label is that a tree with $j_{\lambda_{v_0}} = j$ is a graphic representation of a "contribution" to $\Xi_j^{h\sigma}$. Therefore if $j_{\lambda_{v_0}} \ge \ell$ we call the branch an *action branch* and if $j_{\lambda_{v_0}} < \ell$ we call it an *angle branch.*

In the first of the figures in (3.6) integrals with respect to the nodes v_1, v_2 are of the type $\overline{\mathcal{I}}^2$. In the second the integrals over the τ_{v_n}, $n = 1, 2, 3$, are all of the type $\mathcal{O}$. In both cases the integrals over τ_{v_0} are of the form $\mathcal{I}$ because we fixed $j_{\lambda_{v_0}} = j > \ell$ to be an action label. We can associate a branch label j_{λ_v} also to the inner branches with $v < v_0$: however in this case one has necessarily $j_{\lambda_v} = j_v - \ell$ because the inner branches necessarily represent angle components $\Xi_j^{h\sigma}$ with $j < \ell$, see (2.12). Hence no information is carried by such labels that we define only for uniformity of notation. The latter labels appear in (3.6) as j, p, q in the first tree and as $j, 0, 0, 0$ in the second one. The labels j_λ corresponding to the lines pertaining to a node v determine, as in the examples of §3, which derivatives have to be taken of the function f_{δ_v} which is associated with v: each line λ_v with label j_{λ_v} corresponds to a derivative of f_{δ_v} with respect to φ if $j_{\lambda_v} = 0$ or to $\alpha_{j_{\lambda_v}}$ if $0 < j_{\lambda_v} < \ell$.

• The integrations over the node times τ_v must be thought of as improper integrals, in the above sense, from $\sigma\infty$ to either $\tau_{v'}$ or 0 because (2.25) contains various integrals between such extremes. *It will be convenient to distinguish between such terms.*

This can be easily done by adding, on each node, a new label ρ_v also equal to $0, 1$: if $\rho_v = 1$ this means, naturally, that in the evaluation of the integration operations relative to the node v we select the terms that correspond to integrations between $\sigma\infty$ and $\tau_{v'}$ while if $\rho_v = 0$ we select the integrations between $\sigma\infty$ and 0. We shall imagine that also the highest node carries a label ρ_{v_0} which is 1 necessarily if we consider only

$\Xi_{\uparrow}(t)$ (because this implies that the function associated with the highest node must appear differentiated with respect to a $\boldsymbol{\alpha}$–component, see above and (2.12)), but which could be 0 for the other components of $\Xi(t)$. Recall also that in this case $\tau_{v'} \equiv t$, see (2.25).

• We remark that the hierarchical structure of the integrations implies that if $\rho_v = 1$ and if v' is the node immediately following (in the direction of the root) v along the tree then one has $\tau_v > \tau_{v'}$ if $\sigma = +$ and $\tau_v < \tau_{v'}$ if $\sigma = -$, while $\tau_v, \tau_{v'}$ have the same sign but are otherwise unrelated if $\rho_v = 0$; see (2.25) and check this in the examples.
Besides the labels already introduced also the labels $\rho_v = 0, 1$, just described but not shown in (3.6), should be imagined carried by each node.

• Given a tree labeled as above we pick up the nodes v with $\rho_v = 0$ which are closest to the root, and consider the subtrees having such nodes as highest nodes. We call each such subtree, *i.e.* each such node *together with the subtrees ending in it* (and its labels), a *leaf*.[5] The name is natural if one imagines to enclose the part of the tree including the node v itself and half of the line λ_v into a circle (or, more pictorially, into a leaf shaped contour): hence, to whom tries the drawing, it will look like the *venations* of a leaf and the half line outside it will look like its *stalk*.

• All nodes which do not belong to any leaves will be called *free nodes*; they carry, by construction, a label $\rho_v = 1$, so that the corresponding time variables are hierarchically ordered from the lowest nodes up to the root: *i.e.* if $w < v$ then $\tau_w < \tau_v$ if $\sigma = +$ and $\tau_w > \tau_v$ if $\sigma = -$. Given a tree ϑ let us call ϑ_f the set of free nodes in ϑ, and call Θ_L the set of highest nodes of the leaves.

• Each f_{δ_v} function, associated with the node v with order label δ_v, can be decomposed into its Fourier harmonics. This can be done graphically by adding to each node v a *mode* label $\underline{\nu}_v = (\nu_{0v}, \boldsymbol{\nu}_v) \equiv (n_v, \boldsymbol{\nu}_v) \in \mathbb{Z}^{\ell}$, with $|\boldsymbol{\nu}_v| \le N$ and $|n_v| \le N_0$, that denotes the particular harmonic selected for the node v. If $(j_v, \delta_v, \rho_v, \underline{\nu}_v)$ are the labels of v we will associate with v the quantity $f^{\delta_v}_{\underline{\nu}_v} e^{i(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_v\tau_v + n_v\varphi^0(\tau_v))}$ multiplied by appropriate products of factors in_v (one per φ–derivative) and $i\nu_{vj}$ (one per α_j–derivative, $j = j_{\lambda_v}$).
If the mode labels $\underline{\nu}_v$ are specified for each v we shall define the *momentum* $\boldsymbol{\nu}(v)$ "flowing" on a branch λ_v as the sum of all the angle mode components $\boldsymbol{\nu}_w$ of the nodes w *preceding* the branch, with v included,

$$\boldsymbol{\nu}(v) \stackrel{def}{=} \sum_{w\in\vartheta,\ w\le v} \boldsymbol{\nu}_w \; ; \tag{4.1}$$

the momentum $\boldsymbol{\nu}(v_0)$ flowing through the root branch will be called the *total momentum* (of the tree).
We shall define also the *total free momentum* of the tree as the sum of the mode labels of the free nodes: more generally, for any free node v we can define the *free momentum* flowing through the branch λ_v as

$$\boldsymbol{\nu}_0(v) = \sum_{w\in\vartheta_f,\ w\le v} \boldsymbol{\nu}_w \; . \tag{4.2}$$

For instance in the above examples the two contributions (3.3), (3.5) (represented by figure (3.6)) are decomposed into sums of several distinct contributions once the ρ_v and the mode labels $\underline{\nu}_v$ are specified.
Likewise we can look at a leaf as a tree: the momentum $\boldsymbol{\nu}'$ flowing through its stalk will then be called the internal *leaf momentum*. Note that its value gives *no contribution* to the total free momentum of the tree to which the leaf belongs.

• The free momenta will turn out to describe the harmonics of the time dependent quasi periodic motion around the invariant tori, while the Fourier expansion modes of $X^{h\sigma}(t;\boldsymbol{\alpha})$ as a function of $\boldsymbol{\alpha}$ are related to

[5] This definition is slightly different from the one given in [Ge2], where the leaf represents a collection of trees and, as explained below, is related to a resummation operation (see also comments in §4.2, item (v), below, and (4.27)), that we do not consider here.

the sum of the free momenta *and* of all the internal leaf momenta. This is an important difference: it is a property stressed in [G1] where it is referred as "quasi flatness", source of the main difficulties and interest in the theory of homoclinic splitting, see [G1,GGM2,GGM3,G3].

4.2. The trees contributions of the examples of §3 will be sums over the various labels of "values" of trees decorated by more labels:

$$\frac{1}{2}\int_{\sigma\infty}^{g_0 t} dg_0\tau_{v_0}(-i\nu_{v_0 j})(i\nu_{v_0 p})(i\nu_{v_0 q})\, f^1_{\underline{\nu}_{v_0}}\, e^{i(\boldsymbol{\nu}_{v_0}\cdot\boldsymbol{\omega}'\tau_{v_0}+n_{v_0}\varphi^0(\tau_{v_0}))}\cdot \tag{4.3}$$
$$\cdot\,\overline{\mathcal{I}}^2\big((-i\nu_{v_1 p})\, f^1_{\underline{\nu}_{v_1}}\, e^{i(\boldsymbol{\nu}_{v_1}\cdot\boldsymbol{\omega}'\tau_{v_1}+n_{v_1}\varphi^0(\tau_{v_1}))}\big)(\rho_{v_1}\tau_{v_0})\,\overline{\mathcal{I}}^2\big((-i\nu_{v_2 q})\, f^1_{\underline{\nu}_{v_2}}\, e^{i(\boldsymbol{\nu}_{v_2}\cdot\boldsymbol{\omega}'\tau_{v_2}+n_{v_2}\varphi^0(\tau_{v_2}))}\big)(\rho_{v_2}\tau_{v_0})\ ,$$

for (3.3) and

$$\frac{1}{2}\int_{\sigma\infty}^{g_0 t} dg_0\tau_{v_0}(-i\nu_{v_0 j})(in_{v_0})\, f^1_{\underline{\nu}_{v_0}}\, e^{i(\boldsymbol{\nu}_{v_0}\cdot\boldsymbol{\omega}'\tau_{v_0}+n_{v_0}\varphi^0(\tau_{v_0}))}\,\mathcal{O}\Big((-in_{v_1})\, f^0_{\underline{\nu}_{v_1}}\, e^{in_{v_1}\varphi^0(\tau_{v_1})} \tag{4.4}$$
$$\mathcal{O}\big((-in_{v_2})\, f^1_{\underline{\nu}_{v_2}}\, e^{i(\boldsymbol{\nu}_{v_2}\cdot\boldsymbol{\omega}'\tau_{v_2}+n_{v_2}\varphi^0(\tau_{v_2}))}\big)(\rho_{v_2}\tau_{v_1})\,\mathcal{O}\big((-in_{v_3})\, f^1_{\underline{\nu}_{v_3}}\, e^{i(\boldsymbol{\nu}_{v_3}\cdot\boldsymbol{\omega}'\tau_{v_3}+n_{v_3}\varphi^0(\tau_{v_3}))}\big)(\rho_{v_3}\tau_{v_1})\Big)\ ,$$

for (3.5), with the conventions following (3.3) about the dummy integration variables.

• The integration operations are still fairly involved, as it can be seen from (2.25) and from the expressions for $\overline{\mathcal{I}}^2$ and $\mathcal{O}$. With the above conventions for the dummy variables and noting that, for any function F in $\hat{\mathcal{M}}$,

$$\overline{\mathcal{I}}^2\big(F(\tau)\big)(t) = J^{-1}J_0\left(\mathcal{I}(g_0(t-\tau)F(\tau))(t) - \mathcal{I}(g_0(t-\tau)F(\tau))(0)\right)\ , \tag{4.5}$$

we see that the integration over the τ_v has (by (2.25)) one of the two forms, when $\rho_v = 1$ and v' is not the root (so that $j_{\lambda_v} = j_v - \ell$),

$$\begin{aligned}
&(1)\quad \mathcal{I}\big((w_{0\ell}(\tau_{v'})w_{00}(\tau_v) - w_{00}(\tau_{v'})w_{0\ell}(\tau_v))e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(\tau_{v'}), && j_{\lambda_v} = 0\ ,\\
&(2)\quad \mathcal{I}\big(g_0(\tau_{v'}-\tau_v)\, e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(\tau_{v'}), && 0 < j_{\lambda_v} < \ell\ ,
\end{aligned} \tag{4.6}$$

where $G_v(\tau_v)$ is a function that depends on the structure of the tree formed by the nodes preceding v and by the labels attached to the nodes. If $\rho_v = 0$ it has one of the two forms

$$\begin{aligned}
&(1)\quad w_{00}(\tau_{v'})\mathcal{I}\big(w_{0\ell}(\tau_v)e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(0)\,, && j_{\lambda_v} = 0\ ,\\
&(2)\quad \mathcal{I}\big(g_0\tau_v\, e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(0), && 0 < j_{\lambda_v} < \ell\ .
\end{aligned} \tag{4.7}$$

• When v' is the root the operations involved in the evaluation of the τ_v–integral are slightly different *if* $j_{\lambda_{v_0}} = j_{v_0}$, *i.e.* if we are considering contributions to the action coordinates, (if $j_{\lambda_{v_0}} = j_{v_0} - \ell$ we still have integrations of the form (4.6) or (4.7)). If $j_{\lambda_{v_0}} = j_{v_0}$ the integrations are particularly simple if we are interested in the evaluation of the splitting vector (1.7), that is $j_{v_0} > \ell$ and $t = 0$; in such a case the last two of (2.25) are relevant and setting $v = v_0$ the integration over τ_{v_0} is the value for $\tau_{v'}$ of

$$\begin{aligned}
&(1)\quad \mathcal{I}\left(w_{00}(\tau_v)e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\right)(0)\ , && j_{\lambda_v} = \ell\ ,\\
&(2)\quad \mathcal{I}\left(e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\right)(0)\ , && j_{\lambda_v} > \ell\ .
\end{aligned} \tag{4.8}$$

because, if $\tau_{v'}=0$, one has $w_{\ell\ell}(0)=1$ and $w_{\ell 0}(0)=0$; see (2.15) and the last two of (2.25).
More generally, if $\tau_{v'_0}=t\neq 0$, setting $v=v_0$ and $r=v'_0$, one defines for $\rho_{v_0}=1$

$$\begin{array}{lll}
(1) & \mathcal{I}\big((w_{0\ell}(\tau_{v'})w_{00}(\tau_v)-w_{00}(\tau_{v'})w_{0\ell}(\tau_v))e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(\tau_{v'}), & j_{\lambda_v}=0\,,\\
(2) & \mathcal{I}\big(g_0(\tau_{v'}-\tau_v)\,e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(\tau_{v'}), & 0<j_{\lambda_v}<\ell\,,\\
(3) & \mathcal{I}\big((w_{\ell\ell}(\tau_{v'})w_{00}(\tau_v)-w_{\ell 0}(\tau_{v'})w_{0\ell}(\tau_v))e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(\tau_{v'})\,, & j_{\lambda_v}=\ell\,,\\
(4) & \mathcal{I}\big(e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(\tau_{v'})\,, & j_{\lambda_v}>\ell\,,
\end{array}\tag{4.9}$$

(see the last two relations in (2.25)) and for $\rho_{v_0}=0$

$$\begin{array}{lll}
(1) & w_{00}(\tau_{v'})\mathcal{I}\big(w_{0\ell}(\tau_v)e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(0)\,, & j_{\lambda_v}=0\,,\\
(2) & \mathcal{I}\big(g_0\tau_v\,e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(0), & 0<j_{\lambda_v}<\ell\,.\\
(3) & w_{\ell 0}(\tau_{v'})\mathcal{I}\big(w_{0\ell}(\tau_v)e^{i(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v+n_v\varphi^0(\tau_v))}G_v(\tau_v)\big)(0)\,, & j_{\lambda_v}=\ell\,,\\
(4) & 0\,, & j_{\lambda_v}>0\,;
\end{array}\tag{4.10}$$

note that, for $\tau_{v'}=t=0$ and $j_{\lambda_{v_0}}\geq\ell$, (4.9) and (4.10), summed together, give (4.8).

• Hence each node still describes a rather complicated set of operations: it is, therefore, convenient to consider separately the terms that appear in (4.6)÷(4.10). This can be done by simply adding further labels at each node. To this end, looking at the integrals in (4.7) and (4.10), at $\rho_v=0$, and in (4.6) and (4.9), at $\rho_v=1$, we see that the following kernels are involved in the integrals

$$\begin{aligned}
w^0_{j_{\lambda_v}}(\tau_{v'},\tau_v) &= \begin{cases} w_{00}(\tau_{v'})w_{0\ell}(\tau_v), & v>v_0\,,j_v=\ell\,\to j_{\lambda_v}=0\,,\\ g_0\tau_v, & v>v_0\,,j_v>\ell\,\to 0<j_{\lambda_v}<\ell\,,\end{cases}\\
w^0_{j_{\lambda_{v_0}}}(t,\tau_{v_0}) &= \begin{cases} w_{00}(t)w_{0\ell}(\tau_{v_0}), & j_{v_0}=\ell,\quad j_{\lambda_{v_0}}=0\,,\\ g_0\tau_{v_0}, & j_{v_0}>\ell,\quad 0<j_{\lambda_{v_0}}<\ell\,,\\ w_{\ell 0}(t)w_{0\ell}(\tau_{v_0}), & j_{v_0}=\ell,\quad j_{\lambda_{v_0}}=\ell\\ 0, & j_{v_0}>\ell,\quad j_{\lambda_{v_0}}>\ell\,,\end{cases}\\
w^1_{j_{\lambda_v}}(\tau_{v'},\tau_v) &= \begin{cases} w_{0\ell}(\tau_{v'})w_{00}(\tau_v)-w_{00}(\tau_{v'})w_{0\ell}(\tau_v), & v>v_0\,,j_v=\ell\,\to j_{\lambda_v}=0\,,\\ g_0(\tau_{v'}-\tau_v), & v>v_0\,,j_v>\ell\,\to 0<j_{\lambda_v}<\ell\,,\end{cases}\\
w^1_{j_{\lambda_{v_0}}}(t,\tau_{v_0}) &= \begin{cases} w_{0\ell}(t)w_{00}(\tau_{v_0})-w_{00}(t)w_{0\ell}(\tau_{v_0}), & j_{v_0}=\ell,\quad j_{\lambda_{v_0}}=0\,,\\ g_0(t-\tau_{v_0}), & j_{v_0}>\ell,\quad 0<j_{\lambda_{v_0}}<\ell\,,\\ w_{\ell\ell}(t)w_{00}(\tau_{v_0})-w_{\ell 0}(t)w_{0\ell}(\tau_{v_0}), & j_{v_0}=\ell,\,j_{\lambda_{v_0}}=\ell,\\ 1, & j_{v_0}>\ell,\,j_{\lambda_{v_0}}>\ell,\end{cases}
\end{aligned}\tag{4.11}$$

respectively appearing in (4.7) and (4.10), at $\rho_v=0$, and in (4.6) and (4.9), at $\rho_v=1$.
The function in (4.11) involving the Wronskian matrix elements can be computed from (2.15) and one finds, for instance, that the function in the seventh row on the r.h.s. is

$$w_{0\ell}(\tau_{v'})w_{00}(\tau_v)-w_{00}(\tau_{v'})w_{0\ell}(\tau_v)=\frac{1}{2}\left\{\frac{g_0(\tau_{v'}-\tau_v)}{\cosh g_0\tau_{v'}\,\cosh g_0\tau_v}+\frac{\sinh g_0\tau_{v'}}{\cosh g_0\tau_v}-\frac{\sinh g_0\tau_v}{\cosh g_0\tau_{v'}}\right\}\;;\tag{4.12}$$

hence if we consider (4.6)÷(4.10) we note that the integrals over τ_v involve functions that can be written, for $\rho=\rho_v,\tau=\tau_v,\tau'=\tau_{v'}$ and for suitable coefficients $c_j(\rho,\alpha,v)$, ($\rho=1$ if we consider (4.6), (4.9) and $\rho=0$ if we consider (4.7), (4.10)),

$$\sum_{\alpha=-1}^{2}T^{(\alpha)}_\rho(\rho\tau',\tau)\,Y^{(\alpha)}(\tau',\tau)\,c_j(\rho,\alpha,v)\,,\tag{4.13}$$

where $Y^{(\alpha)}(\tau',\tau)$ are given, if $x=e^{-\sigma g_0\tau}$ and $x'=e^{-\sigma g_0\tau'}$, by

$$\begin{aligned}
Y^{(-1)}(\tau',\tau) =&\frac{1}{2}\frac{\sinh g_0\tau}{\cosh g_0\tau'}\,\exp[in\varphi^0(\tau)] = \sum_{k'=1}^{\infty}\sum_{k=-1}^{\infty} y_n^{(-1)}(k',k){x'}^{k'}x^k\ , && k'\ \text{odd}\,,\\
Y^{(0)}(\tau',\tau) =&\frac{1}{2}\frac{\exp[in\varphi^0(\tau)]}{\cosh g_0\tau'\cosh g_0\tau} = \sum_{k'=1}^{\infty}\sum_{k=1}^{\infty} y_n^{(0)}(k',k){x'}^{k'}x^k\ , && k'\ \text{odd}\,,\\
Y^{(1)}(\tau',\tau) =&\frac{1}{2}\frac{\sinh g_0\tau'}{\cosh g_0\tau}\,\exp[in\varphi^0(\tau)] = \sum_{k'=-1}^{\infty}\sum_{k=1}^{\infty} y_n^{(1)}(k',k){x'}^{k'}x^k\ , && k'\ \text{odd}\,,\\
Y^{(2)}(\tau',\tau) =&\exp[in\varphi^0(\tau)] = \sum_{k=0}^{\infty}\tilde y_n^{(2)}(0,k)x^k\ , && k'\equiv 0
\end{aligned}\tag{4.14}$$

which define the coefficients $y_n^{(\alpha)}(k',k)$ for $\alpha=-1,0,1,2$ (it is easily checked that k' is *odd* in the first three relations) and we set, for $\alpha=-1,0,1,2$,

$$T_\rho^{(\alpha)}(\rho\tau',\tau)=\begin{cases} g_0(\tau'-\tau) & \text{if } \alpha \text{ is either 0 or 2 and } \rho=1\ ,\\ g_0\tau & \text{if } \alpha \text{ is either 0 or 2 and } \rho=0\ ,\\ 1 & \text{if } \alpha \text{ is either } -1 \text{ or } 1\ .\end{cases}\tag{4.15}$$

Likewise we shall set, defining the coefficients $\tilde y_n^{(\alpha)}(k',k)$, for $\alpha=-1,0,1$, and $\overline{y}_n^{(-1)}(k',k)$,

$$\begin{aligned}
&\tilde Y^{(\alpha)}(\tau',\tau) = -\tanh g_0\tau'\, Y^{(\alpha)}(\tau',\tau) \stackrel{def}{=} \sum_{k'=-\alpha}^{\infty}\sum_{k=\alpha}^{\infty}\tilde y_n^{(\alpha)}(k',k){x'}^{k'}x^k\ , \qquad \alpha=\pm1, k'=\text{odd}\ ,\\
&\tilde Y^{(0)}(\tau',\tau) = -\tanh g_0\tau'\, Y^{(0)}(\tau',\tau) \stackrel{def}{=} \sum_{k'=1}^{\infty}\sum_{k=1}^{\infty}\tilde y_n^{(0)}(k',k){x'}^{k'}x^k\ , \qquad k'\ \text{odd}\\
&\tilde Y^{(2)}(\tau',\tau) = Y^{(2}(\tau',\tau) \stackrel{def}{=} \sum_{k=1}^{\infty}\tilde y_n^{(2)}(0,k)x^k\ ,\\
&\overline{Y}^{(1)}(\tau',\tau) = \frac{\cosh g_0\tau'}{\cosh g_0\tau}\exp[in\varphi^0(\tau)] \stackrel{def}{=} \sum_{k'=-1}^{\infty}\sum_{k=1}^{\infty}\overline{y}_n^{(1)}(k',k){x'}^{k'}x^k, \qquad k'\ \text{odd}\,,\\
&\tilde T_1^{(0)}(\tau',\tau)=g_0(\tau'-\tau),\qquad \tilde T_1^{(2)}(\tau',\tau)\equiv 1,\qquad \tilde T_0^{(0)}(0,\tau)=T_0^{(0)}\,;
\end{aligned}\tag{4.16}$$

in all other cases the $T,\tilde T,\overline{T}$–functions will be defined 1 (no matter which is the value of the labels that we attribute to them: this is done to uniformize the notation.

The label k will be called the *incoming hyperbolic mode* and k' the *outgoing hyperbolic mode* for reasons that become clear by contemplating (4.19) below.

In terms of (4.14)÷(4.16) the functions (4.11) multiplied by $\exp[in_v\varphi^0(\tau_v)]$ can be expressed as in (4.13), thus defining implicitly the coefficients $c_j(\rho,\alpha,v)$ in (4.13):

$$\begin{aligned}
&w^0_{j_{\lambda_v}}(\tau_{v'},\tau_v)\,\exp[in_v\varphi^0(\tau_v)]=\begin{cases} T_0^{(0)}(0,\tau_v)\,Y^{(0)}(\tau_{v'},\tau_v)+Y^{(-1)}(\tau_{v'},\tau_v), & j_{\lambda_v}=j_v-\ell=0,\\ T_0^{(2)}(0,\tau_v)\,Y^{(2)}(\tau_{v'},\tau_v), & 0<j_{\lambda_v}=j_v-\ell<\ell,\end{cases}\\
&w^0_{j_{\lambda_{v_0}}}(t,\tau_{v_0})\,\exp[in_{v_0}\varphi^0(\tau_{v_0})]=\begin{cases} T_0^{(0)}(0,\tau_{v_0})\,Y^{(0)}(t,\tau_{v_0})+Y^{(-1)}(t,\tau_{v_0}), & j_{\lambda_{v_0}}=j_{v_0}-\ell=0,\\ T_0^{(2)}(0,\tau_{v_0})\,Y^{(2)}(t,\tau_{v_0}), & 0<j_{\lambda_{v_0}}=j_{v_0}-\ell<\ell\\ \tilde T_0^{(0)}(0,\tau_{v_0})\,\tilde Y^{(0)}(t,\tau_{v_0})+\tilde Y^{(-1)}(t,\tau_{v_0}), & j_{v_0}=j_{\lambda_{v_0}}=\ell,\\ 0, & j_{v_0}=j_{\lambda_{v_0}}>\ell,\end{cases}
\end{aligned}\tag{4.17}$$

$$w^1_{j_{\lambda_v}}(\tau_{v'},\tau_v)\,\exp[in_v\varphi^0(\tau_v)]=\begin{cases} T_1^{(0)}(\tau_{v'},\tau_v)\,Y^{(0)}(\tau_{v'},\tau_v)+Y^{(1)}(\tau_{v'},\tau_v)+ & \\ \quad -Y^{(-1)}(\tau_{v'},\tau_v), & j_{\lambda_v}=j_v-\ell=0,\\ T_1^{(2)}(\tau_{v'},\tau_v)\,Y^{(2)}(\tau_{v'},\tau_v), & 0<j_{\lambda_v}=j_v-\ell<\ell,\end{cases}$$

$$w^1_{j_{\lambda_{v_0}}}(t,\tau_{v_0})\,\exp[in_{v_0}\varphi^0(\tau_{v_0})]=\begin{cases} T_1^{(0)}(t,\tau_{v_0})\,Y^{(0)}(t,\tau_{v_0})+Y^{(1)}(t,\tau_{v_0})+ & \\ \quad -Y^{(-1)}(t,\tau_{v_0}), & j_{\lambda_{v_0}}=j_{v_0}-\ell=0,\\ T_1^{(2)}(t,\tau_{v_0})\,Y^{(2)}(t,\tau_{v_0}), & 0<j_{\lambda_{v_0}}=j_{v_0}-\ell<\ell,\\ \tilde T_1^{(0)}(t,\tau_{v_0})\,\tilde Y^{(0)}(t,\tau_{v_0})+\tilde Y^{(1)}(t,\tau_{v_0})+ & \\ \quad -\tilde Y^{(-1)}(t,\tau_{v_0}))+\overline{Y}^{(1)}(t,\tau_v), & j_{v_0}=j_{\lambda_{v_0}}=\ell,\\ \tilde T_1^{(2)}(t,\tau_{v_0})\tilde Y^{(2)}(t,\tau_{v_0}), & j_{v_0}=j_{\lambda_{v_0}}>\ell.\end{cases}$$

One could avoid introducing the $\tilde T$ functions as they are simply related to the T functions or are just identically 1: however it is convenient to introduce them to make the above formulae more symmetric and therefore easier to keep in mind while working with.

Finally we define the coefficients $\xi_j(k',0)$ by the power series expansion

$$\begin{aligned} \frac{1}{\cosh g_0\tau'}&=\sum_{k'=1}^{\infty}\xi_\ell(k',0)x'^{k'}\ , \qquad k'\ge 1\,,\ \text{odd}\ ,\\ 1&=\xi_j(0,0)\ ,\qquad j>\ell\ ,\end{aligned}\tag{4.18}$$

where $x'=e^{-\sigma g_0\tau'}$ and k' is odd, which occurs as coefficient $w_{00}(\tau')$ in (4.7) (when $\rho_v=0$, *i.e.* $v\in\Theta_L$).

The above definitions (taken from (42) and (45) in [Ge2]) suffice to discuss the whiskers (and therefore the splitting in the action variables).

• The (4.13) allow us to introduce a "relatively simple notation": we can add to each node a *badge* label $\alpha_v=(-1,0,1,2)$ that will distinguish which choice we make between the possibilities in (4.14) and (4.16) and two *hyperbolic mode* labels k'_v,k_v which select which particular term we choose in the sums in (4.14) and (4.16); they are integer numbers ≥ -1. We shall not have to introduce labels to distinguish terms coming from the expansions of $Y^{(\alpha)},\tilde Y^{(\alpha)},\overline{Y}^{(\alpha)}$ bearing the same badge α because one can check that the labels α_v together with j_v and v itself uniquely determine which choice has to be made.

In terms of the latter labels we can define a *hyperbolic momentum* of a line λ_v as a label $p(v)\in\mathbb{Z}$ which will be the sum of all the hyperbolic modes of the nodes that precede v *plus* the incoming hyperbolic mode of the node v itself: this is the sum of the labels k_w associated with all *free* nodes $w\le v$, with v included, and of the labels k'_w associated with all the *free* nodes $w<v$ or *highest* nodes of the leaves $w<v$, with v *not* included,

$$p(v)=k_v+\sum_{\substack{w\in\vartheta_f\\ w<v}}(k_w+k'_w)+\sum_{\substack{w\in\Theta_L\\ w<v}}k'_w\ .\tag{4.19}$$

A very important property is that $k_w+k'_w\ge 0$, by (4.14) and (4.16), and $k'_w\ge 0$ if $w\in\Theta_L$, by (4.18), so that $p(v)\ge -1$. Furthermore if $p(v)=0$ then *either* $k_v=-1$ and $k_w+k'_w=0$ for all $w<v$ except one single node $\tilde w<v$ (which is either in ϑ_f or Θ_L) for which $k_{\tilde w}+k'_{\tilde w}=1$, *or* $k_v=0$ and $k_w+k'_w=0$ for all $w<v$. If $p(v)=-1$ then $k_v=-1$ and $k_w+k'_w=0$ for all $w<v$.

• In the above analysis we have not taken explicitly into account the possibility of contributions to $\Phi^{h\sigma}_+$ coming from the third line in (2.12), *i.e.* counterterm contributions. They are, of course, possible and they can be taken immediately into account in the graphical representation by considering the nodes with a label $\delta_v=0$ and adding to them a *counterterm label* κ_v, a non negative integer. If $\kappa_v=0$ this will mean that the node represents a contribution from the first line of the definition of $\Phi^{h\sigma}_+$, *i.e.* a contribution that is unrelated to the counterterms, while if $\kappa_v \ge 1$ the node represents a contribution from the term with $p = \kappa_v$ in the second line contribution to $\Phi_+^{h\sigma}$ in (2.12).

4.3. The trees carry, at this point, quite a few decorating labels and each tree together with all its labels will represent a "very simple" contribution to the value of the h–th order coefficient in the Taylor expansion in ε (at fixed η of course) of the $\Xi^{h\sigma}$ vector. Very simple means that the improper integrals that correspond to each term are very easy to evaluate explicitly and lead to a result that can be expressed as a product of factors determined by the labels of the tree and associated with the nodes and with the lines, see (4.30), below. We list here the set of labels that have been introduced:

j_v	action labels
j_{λ_v}	branch labels
δ_v	order labels
ρ_v	leaf labels
$\underline{\nu}_v$	mode labels
$\boldsymbol{\nu}(v)$	momentum in the branch λ_v following v
$\boldsymbol{\nu}_0(v)$	free momentum in the branch λ_v following v
α_v	badge labels
(k_v', k_v)	hyperbolic mode labels
$p(v)$	hyperbolic momentum in the branch λ_v following v
κ_v	counterterm labels

There are some constraints between the labels, which follow from the rules stated in §4.1 and §4.2 and from the choice of the counterterms (the latter will be discussed in §4.5 below):

• one has $j_{\lambda_v} = j_v - \ell$ if $v < v_0$ and $j_{\lambda_v} = j_v$ or $j_{\lambda_v} = j_v - \ell$ if $v = v_0$ (see the third item in §4.1);
• if $\rho_v = 0$ then $\alpha_v \ne 1$, (see (4.17));
• if $j_{\lambda_v} \ne 0, \ell$, then $\alpha_v = 2$, otherwise if $j_{\lambda_v} = 0, \ell$, then α_v can be $-1, 0, 1$, (see (4.17));
• $\delta_v = 0$ implies $j_v = \ell$ (by the $\boldsymbol{\alpha}$–independence of f_0);
• $k_v, k_v', p(v) \ge -1$, (see (4.14), (4.16) and comment following (4.19));
• $(p(v), \boldsymbol{\nu}_0(v)) \ne (0, \mathbf{0})$, see Remark 4.6 below.

In terms of such labels, given a decorated tree ϑ_0 with m_0 nodes and with highest node v_0, we can define the *value* of a *subtree* ϑ with m free nodes, highest node w (preceding the highest node v_0 of ϑ_0: $w \le v_0$) and label $j_{\lambda_w} = 0, \ldots, \ell - 1$, $\rho_w = 0, 1$. It will be given by the expression

$$\mathrm{Val}(\vartheta) = \Big[\prod_{\substack{v\in\vartheta_f \\ v\le w}} \int\!\!\!\!\!\!-_{\sigma\infty}^{\rho_v g_0\tau_{v'}} dg_0\tau_v\ \mathcal{V}_v(\vartheta)\Big]\Big[\prod_{v\in\Theta_L}\mathcal{L}_v(\vartheta)\Big]\Big[\prod_{\substack{v\in\vartheta_f \\ \delta_v=0}}\gamma_{\kappa_v}(g_0)\Big] , \tag{4.20}$$

where the integration is the improper integration $\mathcal{I}$ (in the sense of §2.4), the tree ϑ consists of a "free" m–nodes tree ϑ_f with leaves attached to a (possibly empty) subset of the nodes of ϑ_f, and the following notation has been used.

(i) The coefficients $\mathcal{V}_v(\vartheta)$ and $\mathcal{L}_v(\vartheta)$ are described by the collection of labels enumerated above. They can be written, respectively, as

$$\mathcal{V}_v(\vartheta) = \bar F_{\underline{\nu}_v} \hat T_{\rho_v}^{(\alpha_v)}(\rho_v\tau_{v'},\tau_v)\, e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_v\tau_v} x_v^{k_v} \prod_{\substack{w\in\vartheta \\ w'=v}} x_v^{k_w'}\, (-1)^{\delta_{\alpha_v,-1}}\, \hat y_{n_v}^{(\alpha_v)}(k_v', k_v) , \tag{4.21}$$

and

$$\mathcal{L}_v(\vartheta) = \xi_{j_v}(k_v', 0) L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}(\vartheta) , \tag{4.22}$$

where $x_v = \exp[-\sigma g_0 \tau_v]$ and $\hat y^{(\alpha_v)}_{n_v}, \hat T^{(\alpha_v)}_{\rho_v}$ are (see (4.17)) either
- $y^{(\alpha_v)}_{n_v}$, $T^{(\alpha_v)}_{\rho_v}$, if either $v < v_0$ or $v = v_0$ and $j_{\lambda_{v_0}} = j_{v_0} - \ell$, or
- $\tilde y^{(\alpha_v)}_{n_v}$ or $\overline{y}^{(1)}_{n_v}$ and $\tilde T^{(\alpha_v)}_{\rho_v}$ or $\overline{T}^{(\alpha_v)}_{\rho_v}$, if $v = v_0$ and $j_{\lambda_{v_0}} = j_{v_0}$.

Furthermore $\rho_v = 1$ if $v < w$, while ρ_w can be either 0 or 1; j_{λ_w} can any value $0, \ldots, 2\ell - 1$ if $w = v_0$, in any other case $j_{\lambda_v} = 0, \ldots, \ell - 1$ (see above). In (4.21)

$$\bar F_{\underline{\nu}_v} = \Big(\frac{J_0}{J}\Big)^{(1-\delta_{j_v,\ell})(1-\delta_{j_v, j_{\lambda_v}})} f^{\delta_v}_{\underline{\nu}_v} \Big[(-i\nu_v)_{j_v-\ell} \prod_{\substack{w\in\vartheta \\ w'=v}} (i\nu_v)_{j_w-\ell}\Big] \tag{4.23}$$

depends on the labels $(\delta, \underline{\nu}, j)$ of the node v and of its predecessors w's (recall that by (2.2) $\underline{\nu}_v = (n_v, \boldsymbol{\nu}_v)$); in (4.22) the quantity $L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}(\vartheta)$ is called the "*value of the leaf*" v of order h_v (see item (v) below for its definition). The matrix J is not, in general, a multiple of the identity and $J_0 J^{-1}$ will be interpreted as acting on the rotator components of $\underline{\nu}_v$ (and it will be 1 when raised to the power 0).

(ii) For the purposes of the cancellations analysis performed in Appendix A3, the exact form of a few coefficients among the $y^{(\alpha_v)}_{n_v}(k'_v, k_v)$'s turns out to be essential, so that we list them here:

$$\begin{array}{ccc} y^{(-1)}_{n_v}(2,-1) = 0 \ , & y^{(-1)}_{n_v}(1,-1) = \sigma/2 \ , & y^{(-1)}_{n_v}(1,0) = 2in_v \ , \\ y^{(1)}_{n_v}(0,1) = 0 \ , & y^{(1)}_{n_v}(-1,1) = \sigma/2 \ , & y^{(1)}_{n_v}(-1,2) = 2in_v \ , \\ y^{(2)}_{n_v}(1,0) = 0 \ , & y^{(2)}_{n_v}(0,0) = 1 \ , & y^{(2)}_{n_v}(0,1) = 4in_v\sigma \ . \end{array} \tag{4.24}$$

The coefficients $\tilde y^{(-1)}(1,-1)$, $\tilde y^{(-1)}(1,0)$, $\tilde y^{(1)}(-1,1)$ and $\overline{y}^{(1)}(-1,2)$ are equal to the corresponding (*i.e.* with the same values of the labels k', k) $y^{(\alpha)}(k',k)$ coefficients.

(iii) The value of a leaf with highest node v in (4.22) is *not* the same as the value $L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}$ in [Ge2]: this is because of the above mentioned change in notation (see the sixth item in §4.1)). In [Ge2] leaf values are defined as sums of the values of all leaves (in the sense we use now) with fixed order, action label and total momentum. Then the leaf value considered here, $L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}(\vartheta)$, is a single contribution to the $L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}$ of [Ge2], and depends *only* on the part of the tree ϑ consisting of the nodes $w \le v$; if we call ϑ_v such a subtree, we can write (*temporarily, just for the purposes of comparison*) the present definition of leaf value as $L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}(\vartheta)$ $\equiv \bar L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}(\vartheta_v)$ (as it depends only on the labels of the subtree ϑ_v). In order to make a link between the different notations note that $L^{h\sigma}_{j\boldsymbol{\nu}}$ in [Ge2] would be, with our present notations, just the sum

$$L^{h\sigma}_{j\boldsymbol{\nu}} \stackrel{def}{=} \sum_{\substack{\vartheta_{v_0}\in\mathcal{T}_{\nu,h} \\ j_{v_0}=j}} \bar L^{h\sigma}_{j\boldsymbol{\nu}}(\vartheta_{v_0}) \ , \tag{4.25}$$

where ϑ_{v_0} is the part of the tree ϑ on which the leaf value really depends.

Coming back to our notations we define the *leaf value* $L^{h\sigma}_{j\boldsymbol{\nu}}(\vartheta)$, with $j = j_{v_0}$, (where the the first and third of (4.11) should be used), to be the value of a tree ϑ with $j_{\lambda_{v_0}} = j_{v_0} - \ell$ and $\rho_{v_0} = 0$.

(iv) By construction (see (2.12) and corresponding comments), and if Θ_L is the set of highest nodes in the leaves, the total perturbation order k of ϑ is

$$k = \sum_{\substack{v\in\vartheta_f \\ \delta_v=0}} \kappa_v + \sum_{v\in\vartheta_f} \delta_v + \sum_{v\in\Theta_L} h_v = \sum_{\substack{v\in\vartheta \\ \delta_v=0}} \kappa_v + \sum_{v\in\vartheta} \delta_v \ , \qquad m < 2k \ . \tag{4.26}$$

(v) Both the counterterms and the leaf values of a given perturbation order are recursively defined in terms of the same quantities with lower orders. In fact $\gamma_\kappa(g_0)$ admits a graphical representation as sum of tree values defined as in (4.20) with the difference that the integration operation corresponding to the highest node of the tree has to be suitably modified (see (4.32) below).

If $\mathrm{Val}\,(\vartheta)$ is defined as in (4.20) (and in item (iv) above) then, by construction, one has

$$\Xi_j^{h\sigma}(t;\boldsymbol{\alpha}) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^{\ell-1}} \Xi_{j\boldsymbol{\nu}}^{h\sigma}(t)\, e^{i\boldsymbol{\nu}\cdot\boldsymbol{\alpha}} \;, \qquad \Xi_{j\boldsymbol{\nu}}^{h\sigma}(t) = \frac{1}{m_0!} \sum_{\substack{\vartheta_0\in\mathcal{T}_{\boldsymbol{\nu},h}\\ j_{\lambda_{v_0}}=j}} \mathrm{Val}\,(\vartheta_0)\;, \tag{4.27}$$

where $\mathcal{T}_{\boldsymbol{\nu},h}$ is the collection of all trees with total momentum $\boldsymbol{\nu}$ and order h. In (4.30) $m_0!^{-1}$ is a combinatorial factor, which depends on the way we count trees: the simplest is to think that the tree branches of ϑ are pairwise distinct and are distinguished by a label $1,2,\ldots,m_0$, if m_0 is the number of nodes in the tree ϑ. In the latter case, which corresponds to our choice, the factor is simply $m_0!^{-1}$, see [G1,Ge2], provided we regard as identical two trees that can be overlapped by pivoting the branches entering a node (rigidly, together with the subtree attached to them) around any node: as in [G1] we shall call *numbered trees* the trees so counted.

4.4. REMARK. Since the value of any leaf with highest node v depends only on the labels of the nodes $w\le v$, the equation (4.20) factorizes into a product of leaf values times a product of counterterms (whose value, so far arbitrary, has still to be specified and it will be, in the analysis between (4.31) and (4.32) when intervening compatibility requirements will dictate its value) times a factor

$$\Big[\prod_{\substack{v\in\vartheta_f\\ v\le w}} \;\;\int_{\sigma\infty}^{\rho_v g_0\tau_{v'}}\!\!\!\!\!\!\!\!\!- \;\; dg_0\tau_v\;\mathcal{V}_v(\vartheta)\Big]\Big[\prod_{v\in\Theta_L}\xi_v(k_v',0)\Big]\;, \tag{4.28}$$

which does not depend on the leaves.

4.5. The extra effort with respect to the approach without counterterms developed in [G1,Ge1], gives here (as in [Ge2]) a reward: few combinations of powers of the "times" τ_v appear in the integrand in (4.21). The time variables, by (4.21) and (4.14)$\div$(4.17), appear only via exponentials like

$$e^{-\sigma(\tau_v-\tau_{v'})a} \qquad \text{or} \qquad (\tau_{v'}-\tau_v)\, e^{-\sigma(\tau_v-\tau_{v'})a}\;, \tag{4.29}$$

for some complex $a=(g_0p(v)-i\sigma\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v))$, yielding respectively, upon integration, a^{-1} or a^{-2}. Note also that by the hierarchical structure of the trees one has $\sigma\,(\tau_{v'}-\tau_v)\ge 0$.

This greatly simplifies the actual performance of the integration operations which, once one gets familiarity with the formalism, are trivial. One can say that the absence of high powers of the τ_v's is due to having *a priori* fixed the Lyapunov exponent g_0 by means of the counterterms (by contrast in [G1,Ge1] arbitrary powers of τ_v appeared because g_0 is *not* fixed *a priori*).

Of course the triviality of the integrations is entirely due to the above *very fine* decomposition, into terms identified by labeled trees, of the more compact (but "difficult" to integrate) integrands appearing in (4.6)$\div$(4.12) and in the middle terms in (4.14).

Once all the integration operations will have been performed, the tree value in (4.20) will become a product of *"factors"*, in complete analogy with what one is accustomed to find when defining *Feynman graphs* in Field Theory. The factors are associated with the nodes v and with the branches λ_v. The value of a tree ϑ will then be *defined* as

$$\begin{aligned}\mathrm{Val}(\vartheta) = & e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v_0)t-\sigma g_0[k'_{v_0}+p(v_0)]t}\Big[\prod_{\substack{\lambda_v\in\vartheta\\ v\in\vartheta_f}}\Big(-\frac{\sigma g_0}{g_0p(v)-i\sigma\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)}\Big)^{r_v}\Big]\;,\\ & \Big[\prod_{v\in\vartheta_f}\bar F_{\nu_v}\,(-1)^{\delta_{\alpha_v,-1}}\,y_{n_v}^{(\alpha_v)}(k_v',k_v)\Big]\Big[\prod_{v\in\Theta_L}\xi_{j_v}(k_v',0)L_{j_v\boldsymbol{\nu}(v)}^{h_v\sigma}(\vartheta)\Big]\Big[\prod_{\substack{v\in\vartheta_f\\ \delta_v=0}}\gamma_{\kappa_v}(g_0)\Big]\end{aligned} \tag{4.30}$$

where r_v is either 1 or 2, and the case $(\boldsymbol{\nu}_0(v), p(v)) = (\mathbf{0}, 0)$ has to be *excluded* for any node $v \in \vartheta_f$. This is not to claim that no trees with $\boldsymbol{\nu}_0(v) = \mathbf{0}, p(v) = 0$ can be drawn by following the above rules: this is a *further rule* to impose on the labels in order that the analysis does not become contradictory requires fixing the function $\gamma(\varepsilon, g_0)$ conveniently: the consistence criterion determines $\gamma(\varepsilon, g_0)$ uniquely. This rule is a natural extension of the corresponding rule holding in perturbation theory of KAM tori, which was discussed by Lindstedt and Newcomb for the lowest orders of the perturbation expansions and which was proved to hold at all orders by Poincaré, [P]; see the last paragraph in §2.4 above and the Remark 4.6, (c), below.

The values of the numbers r_v arise from the time variables integrals via the mechanism just illustrated above (whereby one either gets a^{-1} or a^{-2} from the integration of the functions (4.29)).

The factors $-(\sigma g_0)^{r_v}[g_0 p(v) - i\boldsymbol{\omega}' \cdot \boldsymbol{\nu}_0(v)]^{-r_v}$, associated with the branches, will be called *propagators* or *small divisors*. The first name arises from the possible interpretation of the trees as Feynman graphs of a suitable field theory, see [GGM1]; the second name corresponds to the usual name given in Mechanics to such expressions generated by perturbation expansions.

It can be useful to write, if v_0 is the highest node of ϑ and $j_{\lambda_{v_0}} = 0$,

$$\mathrm{Val}(\vartheta) = e^{i\boldsymbol{\omega}' \cdot \boldsymbol{\nu}_0(v_0)t - \sigma g_0[k'_{v_0} + p(v_0)]t} \Big(- \frac{\sigma g_0}{g_0 p(v_0) - i\sigma\boldsymbol{\omega}' \cdot \boldsymbol{\nu}_0(v_0)} \Big)^{r_{v_0}} \, \overline{\mathrm{Val}}(\vartheta) \;, \tag{4.31}$$

so defining the quantity $\overline{\mathrm{Val}}(\vartheta)$ (this is a well known kind of operation on Feynman graphs, which associates with a graph another value gruesomely called the value of the *amputated* graph – amputated tree in our case). Moreover we can define $\overline{\mathrm{Val}}(\vartheta)$ also for $(\boldsymbol{\nu}_0(v_0), p(v_0)) = (\mathbf{0}, 0)$ as no vanishing denominator appears in its expression. It is however clear that nodes with $(\boldsymbol{\nu}_0(v), p(v_0)) = (\mathbf{0}, 0)$ must not appear at all in the trees that we consider, for (4.30) to make sense as it is written. This implies, not surprisingly, a consistence problem: namely one has to check that the sum of all the $\overline{\mathrm{Val}}(\vartheta)$ over trees of a given order and with $(\boldsymbol{\nu}_0(v_0), p(v_0)) = (\mathbf{0}, 0)$ cancel so that lines λ_v with $(\boldsymbol{\nu}_0(v), p(v)) = (\mathbf{0}, 0)$ never appear, neither for $v = v_0$ nor for $v < v_0$.

The cancellation is made possible because we still have freedom to fix the counterterms and *their choice is in fact uniquely determined by the conditions that they be such that the needed cancellation takes place.* The quantities $\overline{\mathrm{Val}}(\vartheta)$ are convenient in order to find and to express the counterterms and also the "resonance values" introduced later (see Appendix A3). One checks, see Appendix A1, that the counterterms can be explicitly written, if $\mathcal{T}_{\boldsymbol{\nu},h}$ is the collection of all trees with total momentum $\boldsymbol{\nu}$ and order h, as

$$\gamma_\kappa(g_0) = -\frac{1}{2} \sum^{*}_{\substack{\vartheta \in \mathcal{T}_{\mathbf{0},\kappa}, \alpha_{v_0} = -1 \\ p(v_0) = 0, \boldsymbol{\nu}_0(v_0) = \mathbf{0}, k'_{v_0} = 1}} \overline{\mathrm{Val}}(\vartheta) \;, \tag{4.32}$$

and the * means that the sum is further restricted so that the tree contains no leaves. This choice being simply imposed by the requirement that no contribution with $(\boldsymbol{\nu}_0(v_0), p(v_0)) = (\mathbf{0}, 0)$ can arise for $\Xi^{h\sigma}_{-}(t)$; see Appendix A1.

4.6. Remarks. (a) The presence of the counterterms will manifest itself not only through the elimination of the trees whose value would be meaningless if evaluated via (4.30) but also, and mainly, in the fact that the elements of the algebra $\hat{\mathcal{M}}$ met in the successive integrations have a special form (namely always like one of the (4.29)) which implies that the result of the improper integrals is *not* as complicated as one could fear from (2.24). This leads to the simple expression (4.32) (see Appendix A1 of [G1] for what would otherwise happen without counterterms).

(b) From (4.31) one sees that if $j_{\lambda_{v_0}} > \ell$ it is natural to collect together the terms with $p(v_0) = 0$: for them, since $j_{\lambda_{v_0}} > \ell$, in (4.30) one must have $k'_{v_0} + p(v_0) = 0$ by the last of (4.14). Note also that by (2.26) the case $(\boldsymbol{\nu}_0(v), p(v_0)) = (\mathbf{0}, 0)$ is excluded. If $j_{\lambda_{v_0}} \le \ell$ we, likewise, collect the terms with $k'_{v_0} + p(v_0) = 0$

and, for similar reasons the term with $k'_{v_0}+p(v_0)=-1$ cannot be present (see again (4.14) and (4.16), and use $p(v_0)\ge -1$ supplemented by the relations between the labels $p(v_0)$ and k'_{v_0} which will be exhibited in §5.1).

Hence the cases with $p(v_0)+k'_{v_0}=-1$ are excluded by construction[6] and we see that the sum of the values of the trees with $p(v_0)+k_{v'_0}=0$ give us the equations for the actions and the angles of the invariant torus to which the whiskers considered are asymptotic: the terms with $p(v_0)+k_{v'_0}=0$ asymptote to quasi periodic functions of $\boldsymbol{\omega} t$ so that replacing $\boldsymbol{\omega} t$ by $\boldsymbol{\psi}\in T^{\ell-1}$ one gets a parameterization of the points on the tori in terms of a point $\boldsymbol{\psi}\in T^{\ell-1}$ on a "standard torus".

And the terms with $k'_{v_0}+p(v_0)=1$ provide the leading corrections. Since such terms are present already to order 0 (as one sees from the expression of the pendulum separatrix) the distance between a point moving on the stable manifold of the torus and the torus itself will be proportional to $x=e^{-g_0\sigma t}$ as $\sigma t\to\infty$ so that g_0 has the interpretation of Lyapunov exponent of the invariant torus; see (2.27) in §2.27.

(c) Summarizing: *the case* $(\boldsymbol{\nu}_0(v),p(v))=(\mathbf{0},0)$ *has to be ruled out as a consequence of (2.26) and of (4.32), respectively for the contributions to* $\boldsymbol{\Xi}^{h\sigma}_{\uparrow}$ *and to* $\Xi^{h\sigma}_{+}$ *(see the last constraint listed at the beginning of §4.3). All cases with* $k'_{v_0}+p(v_0)=-1$ *are also excluded.*

§5. Bounds.

5.1. We now discuss how to bound the value of a tree or of a sum of a small number of trees which we take for simplicity without leaves and without counterterms. The more general case will be eventually reduced, see below, to the one we consider here.

We shall discuss first how to bound values of trees without leaves and counterterms such that $p(v_0)=0$ if v_0 is the highest node; hence we shall consider trees, always without leaves and counterterms, with $p(v_0)=0$. At the end we shall see how the presence of leaves and counterterms modifies the analysis.

The following discussion is "locally" simple, but "globally" delicate and repeats that in [Ge2], §4: the conclusions are also summarized in the table 0,1,2,3 below.

From (4.19) it follows that the hyperbolic momentum $p(v)$ is $p(v)\ge -1$ and, as remarked after (4.19), $p(v)=0$ can occur only in special cases: more precisely if $p(v)=0$, then k_v is either -1 or 0, and
(1) if $k_v=0$, all free nodes w preceding v (whether immediately or not) have $k'_w+k_w=0$, while
(2) if $k_v=-1$, all free nodes w preceding v have $k'_w+k_w=0$, *except* for a single node $\tilde w<v$ such that $k'_{\tilde w}+k_{\tilde w}=1$.
In the latter case we call $\mathcal{P}$ the path of nodes (*i.e.* the totally ordered set of nodes) which connect v to $\tilde w$, both extremes included (see also [Ge2], §4).

• Supposing $p(v)=-1,0$ and recalling that $\alpha_v<2$ implies k'_v odd, the expansions (4.14) impose that there are *very few possible choices* of the values of the hyperbolic modes at $w\le v$:
(1) if there is no path or there is a path linking v to $\tilde w$ but $w\ne\tilde w$ and $w\notin\mathcal{P}$, then $k_w+k_{w'}=0$ and the cases $\alpha_w=1$, $\alpha_w=-1$ and $\alpha_w=2$ require, respectively, $(k'_w,k_w)=(-1,1)$, $(k'_w,k_w)=(1,-1)$ and $(k'_w,k_w)=(0,0)$: correspondingly $p(w)=1$, $p(w)=-1$ and $p(w)=0$. While, for $w\in\mathcal{P}$, the value of $p(w)$ "increases by one unit", *i.e.* $p(w)=2$, $p(w)=0$ and $p(w)=1$ for $w\in\mathcal{P}$;
(2) if $w=\tilde w$ then $k_w+k_{w'}=1$ and the cases $\alpha_{\tilde w}=1$, $\alpha_{\tilde w}=-1$ and $\alpha_{\tilde w}=2$ require, respectively, $(k'_{\tilde w},k_{\tilde w})=(-1,2)$, $(k'_{\tilde w},k_{\tilde w})=(1,0)$ and $(k'_{\tilde w},k_{\tilde w})=(0,1)$ (correspondingly $p(\tilde w)=2$, $p(\tilde w)=0$ and $p(\tilde w)=1$).

Note that in both cases $\alpha_w=0$ is not possible.

The above analysis covers both cases $v<v_0$ and $v=v_0$, as the functions in (4.14) and in (4.16) have the same dependence on τ (hence on k).

---

[6] The initial data $\Xi^{h\sigma}(0,\boldsymbol{\alpha})$ were determined precisely by imposing boundedness at $\sigma t=+\infty$, *i.e.* by imposing the absence of divergent terms in the expansion in powers of $x=e^{-g_0\sigma t}$ which would correspond to the terms with $p(v_0)=-1$.

• The latter properties have to be considered as a further restriction to impose on the tree labels, and play an essential rôle for the discussion of the cancellations, [Ge2].

• Moreover if $p(v) = 0$ then:
(1) if $k_v = 0$ then v can be preceded only by leaves with the highest nodes w having $j_w > \ell$, because k'_w must be 0 in such a case, so that the second of (4.18) applies;
(2) if $k_v = -1$, then all the leaves again must have the highest node w with $j_w > \ell$, except at most one leaf with highest node $\tilde{w}$ with $j_{\tilde{w}} = \ell$ and $k'_{\tilde{w}} = 1$.

(Vertical) Table 0. Possible cases when $p(v) = 0, -1$.

$p(v)$	k_v	k'_v	α_v	j_v
-1	-1	odd ≥ 1	-1	ℓ
0	-1	odd ≥ 1	-1	ℓ
0	0	≥ 1	-1	ℓ
0	0	≥ 0	2	$> \ell$

(Horizontal) Table 1. Cases $p(v) = 0$, $w \notin \mathcal{P}$.

α_w	-1	0	1	2
(k'_w, k_w)	$(1, -1)$	impossible	$(-1, 1)$	$(0, 0)$
$p(w)$	-1	impossible	1	0
j_w	ℓ	impossible	ℓ	$> \ell$

(Horizontal) Table 2. Cases $p(v) = 0$, $w \in \mathcal{P}, w > \tilde{w}$.

α_w	-1	0	1	2
(k'_w, k_w)	$(1, -1)$	impossible	$(-1, 1)$	$(0, 0)$
$p(w)$	0	impossible	2	1
j_w	ℓ	impossible	ℓ	$> \ell$

(Horizontal) Table 3. Cases $p(v) = 0$, $w = \tilde{w}$.

α_w	-1	0	1	2
(k'_w, k_w)	$(1, 0)$	impossible	$(-1, 2)$	$(0, 1)$
$p(w)$	0	impossible	2	1
j_w	ℓ	impossible	ℓ	$> \ell$

• We extend the definition of path also to the case $p(v) = 0, k_v = 0$, by setting $\mathcal{P} \stackrel{def}{=} \emptyset$ if $j_v > \ell$ and $\mathcal{P} \stackrel{def}{=} v$ if $j_v = \ell$, only for purposes of notational convenience (see (5.3) below). This is consistent with the above tables and does not change them.

5.2. Remark. Note that, if a tree (or subtree) ϑ_0 with highest node v_0 has total hyperbolic momentum $p(v_0) = 0$, then there is one and only one path $\mathcal{P}$, and, if $\mathcal{P} \neq \emptyset$, then $\mathcal{P}$ connects the node v_0 to some node $\tilde{w} < v_0$. In fact, if there is a path $\mathcal{P} \neq \emptyset$, the node $\tilde{w}$ is so defined that $k_{\tilde{w}} + k'_{\tilde{w}} = 1$: then if $k_{v_0} = 0$ there cannot be any of such nodes (and $\mathcal{P} = v_0$ in such a case), while if $k_{v_0} = -1$ there must be one and only one such node. This simply follows from the analysis in §5.1, by noting that all nodes $w < v_0$ except $\tilde{w}$ must have $k_w + k'_w = 0$.

5.3. The small divisors can be really "small" only when $p(v) = 0$: if $p(v) \neq 0$, they are bounded by $g_0^{-r_v}$, *i.e.* by a quantity of order $O(1)$. So one can consider all free nodes in the trees, among the ones having $p(v) = 0$, which are closest to the root. All free nodes v between them and the root have propagators which are not small because $|p(v)| \geq 1$ (see also [Ge2], page 298).

Given a tree ϑ_0 with m free nodes, from each subtree ϑ ending in a node v_0 with $p(v_0)=0$ (here v_0 is some node of ϑ_0: it becomes the highest node of ϑ), one obtains contributions which can be naturally collected together (recall Remark 4.4) into a contribution to the tree value (see (4.30)) consisting in a factor

$$\prod_{v\le v_0} \bar F_{\underline{\nu}_v} G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)]\, y'_v \ , \tag{5.1}$$

(here the product is over all free nodes preceding v_0) times a product of counterterms $\gamma_{\kappa_v}(g_0)$ for $v\in\vartheta_f$ with $\delta_v=0$, times a products of factors $\xi_{j_v}(k'_v,0)L^{h_v\sigma}_{j_v\boldsymbol{\nu}(v)}$, for each $v\in\Theta_L$; see (4.30). The vector $\boldsymbol{\nu}_0(v)$ is the free momentum (defined above; see (4.2)) flowing through the branch λ_v, the coefficients y'_v are related to the expansions (4.14) via

$$y'_v=\begin{cases}\frac12\left[y^{(1)}_{n_v}(-1,1)+y^{(-1)}_{n_v}(1,-1)\right]=\sigma/2 & \text{if } v\in\mathcal{P},\ \alpha_v\in\{-1,1\}\ ,\\ (-1)^{\delta_{\alpha_v,-1}}\, y^{(\alpha_v)}_{n_v}(k_{v'},k_v) & \text{otherwise }; \end{cases} \tag{5.2}$$

where $\mathcal{P}$ denotes the path in ϑ (there is always one such path, possibly the empty set, because we suppose $p(v_0)=0$); and $G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)]$ is related to the propagator of the branch λ_v, and it will have the form

$$G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)]=\begin{cases} g_0^2\left[i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)\right]^{-2} & \text{if } v\notin\mathcal{P},\ j_v>\ell,\\ -\sigma g_0^2\left[g_0^2+(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v))^2\right]^{-1} & \text{if } v\notin\mathcal{P},\ j_v=\ell,\\ g_0^{2-\delta_{j_v,\ell}}\left[-\sigma\left(g_0p(v)-i\sigma\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)\right)\right]^{-(2-\delta_{j_v,\ell})} & \text{if } v\in\mathcal{P},\ \alpha_v\neq-1,\to p(v)\neq0\\ g_0\left[i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)\right]^{-1} & \text{if } v\in\mathcal{P},\ \alpha_v=-1, \end{cases} \tag{5.3}$$

because:
(a) The first line is such because if $j_v>\ell$ one has necessarily $\alpha_v=2$, see Tables 0,1,2,3 and we have to integrate a function $g_0(\tau_{v'}-\tau_v)e^{in_v\varphi^0(\tau_v)}$ so that $k_v\ge0$: hence $k_v=p(v)=0$ and we have the second function in (4.29) to integrate.
(b) The second line is such because if $v\notin\mathcal{P}$, $j_v=\ell$ we have $w^1_\ell(\tau_{v'},\tau_v)e^{in_v\varphi^0(\tau_v)}$ which is a sum of three terms (see the third of (4.17)): the first has $k_v+k_{v'}\ge2$ so is excluded (recall that $p(v_0)=0$ and $v\le v_0$); while the second only sees the contribution to $Y^{(1)}$ with $k'_v=-1,k_v=1$, see (4.14), and the third only contributes by the term with $k'_v=1,k_v=-1$ in $Y^{(-1)}$. In the two cases one has $p(v)=1$ or $p(v)=-1$ respectively; adding up together the latter two contributions and using the first of (5.2) to compute the sum of the coefficients we get

$$\frac{-\sigma g_0 y^{(1)}_{n_v}(-1,1)}{g_0-i\sigma\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)}-\frac{-\sigma g_0 y^{(-1)}_{n_v}(1,-1)}{-g_0-i\sigma\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)}=\frac{\sigma}{2}\frac{-2\sigma g_0^2}{g_0^2+(\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v))^2}\ , \tag{5.4}$$

as it can be read from the coefficients in the intermediate column of (4.24) and from $p(v)=\alpha_v=\pm1$.
(c) The third line of (5.3) is obtained by noting that, if $v\in\mathcal{P}$, $v>\tilde w$ one has $p(v)=1+k_v$, so that, if $j_v=\ell$ and $\alpha_v\neq-1$, then $p(v)>0$, see Table 2; if $v=\tilde w$ and $\alpha_v\neq-1$, one has $p(v)\neq0$, see Table 3 (note that $\alpha_v\neq2,0$ so that we have to consider the first integrand in (4.29)).
If $j_v>\ell$ then $\alpha_v=2$, and, by the Tables 2,3, one has $k_{v'}=0,k_v\ge0$ and $k_v+k_{v'}=1$, so that $k_v=1$ and $p(v)=2$; while, if $v=\tilde w$, then $k_{v'}=0,k_v\ge0$ and $k'_v+k_v=1$ imply $k_v=1$, so that $p(v)=1$. So in both cases $p(v)\ge1$.
(d) The fourth line is found by looking at the Tables 2,3 as follows: if $\alpha_v=-1,v\in\mathcal{P},v>\tilde w$, one has $k_v+k'_v=0$, hence $k_v=-1,k_{v'}=1$ and $p(v)=0$; this happens only if $j_v=\ell$ so that we have to consider the first integral in (4.29) and we get the fourth relation.

This shows that the only trees that do not have a value tending to 0 as $t\to\sigma\infty$, *i.e.* are those with $p(v_0)+k'_{v_0}=0$ (all the others tend to 0 as a power of $x=e^{-g_0\sigma t}$), have propagators that are even functions of the momenta flowing in them. In fact the observation on the absence of paths preceding v_0 implies that only the first two propagators in (5.3) appear in such trees. Since, as already remarked, the trees with $p(v_0)+k'_{v_0}=0$ give the equations of the tori this is an interesting check that the tori equations so obtained at $t=+\infty$ and $t=-\infty$ do *coincide*. A similar analysis, and check, holds for the cases $j_{\lambda_{v_0}}\le\ell$.

5.4. Remark. Collecting together the contributions from $\alpha_v=-1$ and $\alpha_v=1$, for $v\notin\mathcal{P}$, is a convenient operation and has nothing to do with the deeper resummations that imply the cancellations necessary for convergence estimates: the systematic use of this operation should be described by adding a label to the trees on the nodes $v\notin\mathcal{P}$ and replacing on the branches which give rise to one of the two propagators in (5.4) the α_v label by the new label (*e.g.* a $*$ label which indicates that we consider the sum of the values of a tree with $\alpha_v=1$ and one with $\alpha_v=-1$). We shall do this without explicitly mentioning the new label, to simplify the notation. Moreover we can no more associate a label $p(v)$ to a node of this kind, as two factors with different $p(v)$ label ($p(v)=\pm1$ for $\alpha_v=\pm1$) have been considered together; nevertheless we shall modify slightly the definition of $p(v)$ by setting $p(v)\stackrel{def}{=}1$ in such a case (and letting it unchanged in all the other cases).

We shall continue to call $G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)]$ a *propagator* as, for the purposes of the following analysis, only such modified version of the original propagators appearing in (4.30) plays a rôle.

5.5. Furthermore we define the *degree* D of a propagator to be $D=2$ if either $v\notin\mathcal{P}$ or $v\in\mathcal{P}, j_v>\ell$ (hence $\alpha_v\neq-1$), and $D=1$ otherwise (the constraint, see (4.3), $1\le r_v\le2$ implies that the power to which the divisors appear raised is either 1 or 2); by extension we shall say that a branch λ has degree $D_\lambda=D$ if the corresponding propagator has degree D.

The coefficients $\bar F_{\underline{\nu}_v}$ and y'_v in (5.1) satisfy the bounds

$$|y'_v|\le 4N\ ,\qquad \prod_{v\le v_0}|\bar F_{\underline{\nu}_v}|\le(\mathcal{C}N^2)^m\ ,\tag{5.5}$$

for some constant $\mathcal{C}$ depending on the perturbation f_1 in (2.1); see (2.6), (2.13) and (2.18). For instance one can take

$$\mathcal{C}=\max\{|J^{-1}|J_0,1\}\max_{|n|\le N_0,\,|\boldsymbol{\nu}|\le N}|f_{\underline{\nu}}|\ ;\tag{5.6}$$

see (4.23), where $|J^{-1}|$ is the maximum of the matrix elements of the (diagonal) matrix J^{-1}.

To bound the product in (5.1), we shall consider simultaneously the cases $k_{v_0}=0,-1$; if $k_{v_0}=0$ the path $\mathcal{P}$ is supposed to be reduced to a single node, v_0, or to the empty set, $\emptyset$, depending on the value of j_{v_0}, (respectively $j_{v_0}=\ell$, and $j_{v_0}>\ell$, see above).

What follows below and in Appendix A2 really goes beyond [Ge2], although it constitutes a natural extension of it. From now now let us consider the case $\ell=3$ and the Hamiltonian (1.1).

We shall assume first a condition on the rotation vectors stronger than the Diophantine one, as done in [G1,GG,Ge2], *i.e.* we suppose that they satisfy a *strong Diophantine condition*

$$\begin{aligned}&(1)\quad C_0|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|\ge|\boldsymbol{\nu}|^{-\tau}\ ,\qquad \mathbf{0}\neq\boldsymbol{\nu}\in\mathbb{Z}^2,\qquad C_0^{-1}=\eta^{-\frac12}C(\eta)\ ,\\ &(2)\quad \min_{0\ge p\ge n}\Big|C_0|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|-2^p\Big|\ge 2^{n+1}\ ,\quad\text{if}\quad n\le0,\ \ 0<|\boldsymbol{\nu}|\le(2^{n+3})^{-1/\tau},\end{aligned}\tag{5.7}$$

where $n,p\in\mathbb{Z}$, $n\le0$, and

$$\boldsymbol{\omega}_0\equiv\eta^{-1/2}(\Omega_1+\eta^{\frac12}J^{-1}A_1)^{-1}\boldsymbol{\omega}'=\big(1,\eta^{-1}(\Omega_1+\eta^{\frac12}J^{-1}A_1)^{-1}\Omega_2\big)\ ,\tag{5.8}$$

so that $\boldsymbol{\omega}'\cdot\boldsymbol{\nu}=\eta^{1/2}(\Omega_1+\eta^{\frac12}J^{-1}A_1)\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}$. We suppose also that $A_1\in[-\eta^{-\frac12}R,\eta^{-\frac12}R]$, with $R\le J\Omega_1/2$, so that $\eta^{1/2}(\Omega_1+\eta^{\frac12}J^{-1}A_1)\ge\eta^{1/2}\Omega_1/2$.

If we write $\boldsymbol{\omega}' = (\eta^{\frac{1}{2}}\Omega_1 + \eta J^{-1}A_1', \eta^{-\frac{1}{2}}\Omega_2)$ then the measure of the set of A_1''s such that $\boldsymbol{\omega}_0$ verifies the strong Diophantine condition (5.7) has measure of size $O(C_0^{-1}\eta^{-3/2})$.

By reasoning as in [GG], once the case of strong Diophantine vectors has been understood, it can be extended to cover also the case of the usual (weaker) Diophantine condition (expressed by (1) in (5.7) above). Alternatively one could follow the approach in [GM] avoiding completely considering condition (2) in (5.7) and assuming only the "usual" condition (1) in (5.7). We shall not perform such an analysis (which can be easily adapted from the quoted papers), and we shall confine ourselves to the case of strongly Diophantine vectors.[7]

Keeping in mind that $C_0 = \eta^{-\frac{1}{2}} e^{+s\eta^{-\frac{1}{2}}}$ is enormous we shall say that

$$\begin{array}{lll} (1)\ \ G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)] & \text{is on scale 1, if } C_0|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}_0(v)| > C_0/4, \text{ or if } p(v)\neq 0; \\ (2)\ \ G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)] & \text{is on scale 0, if } 1/2 < C_0|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}_0(v)| \le C_0/4; \\ (3)\ \ G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)] & \text{is on scale } n\le -1, \text{ if } 2^{n-1} < C_0|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}_0(v)| \le 2^n. \end{array} \tag{5.9}$$

5.6. REMARK. Note that in the above definition of scale the second and third cases can arise only if $p(v) = 0$. The propagators on scale 1 can be bounded by 4^2 if $p(v) = 0$ and by 1 if $p(v) \neq 0$. Note also that the definition of the scales $n = 0$ and $n = 1$ is different from [G1,Ge2]: *this is an important modification*, exploited in Appendix A2, useful in order to take advantage from the existence of different scale times.

5.7. As it is well known, (5.1) cannot usefully be bounded by just taking the absolute value of each factor and bounding the denominators by using the Diophantine condition. This is true not only if one wants to get the improved bounds that we are studying, but also if one, more modestly, wants to show convergence for ε small enough: this is a problem usually referred to as a "*small divisors problem*".

Useful bounds are nevertheless possible, as shown first in similar cases in [E], because one can collect the contributions from the various trees into pairwise disjoint ("*non overlapping*") classes whose values add up to a quantity that verifies much better bounds than the individual elements of the same class. Each class $\mathcal{F}(\vartheta)$ will be determined by one of its elements ϑ called a *representative*. This means that there are important cancellations within each class.

The classes can be constructed by collecting trees which have the same *resonance structures*. The key notion of resonance is recalled below and the description of the classes will follow it.

DEFINITION. *A "cluster" T of scale n_T will be a maximal connected set of branches with scales $n \ge n_T$ and with at least one branch of scale n_T.*

A free node v will be defined to be internal to T, $v \in T$, if at least one of the branches leading to it or coming from it, *i.e. pertaining to* v (as defined in §4), belongs to T; a leaf with highest node v' will be defined to be internal to the cluster T if $v' \in T$.

A branch λ_v is called *external* to T if it does not belong to T but it pertains to a node v internal to T, and it said to be entering T if the node v' following it is in T, exiting from T if $v \in T$, (note that an external branch of T is not any branch outside T). We define the *degree* D_T of a cluster T to be the degree of its exiting branch, and the *order* k_T of T to be given by the same formula as (4.26), with the extra constraint that the nodes are internal to T.

DEFINITION. *A "resonance" V will be a cluster with only two external branches λ_{v_0} and λ_{v_1} carrying the*

[7] Basically the argument is the following: the analysis that we present does not change if $2^p, 2^n$ are replaced by exponentials in another base q (larger than 1) or even if they are replaced by $\gamma(p), \gamma(n)$, where $\gamma(p)/q^p \xrightarrow[p\to-\infty]{} 1$, and if in the second of (5.7) we substitute $2^p, 2^{n+1}, 2^{n+3}$ by, respectively, $\gamma(p), \gamma(n+1), \gamma(n+3)$. One then proves a simple arithmetic lemma (see [GG]), whereby it follows that, if the first of (5.7) is verified and if $\gamma(p)$ is suitably chosen, then the second holds with $\gamma(p), \gamma(n+1), \gamma(n+3)$ replacing $2^p, 2^{n+1}, 2^{n+3}$.

same free momentum, $\boldsymbol{\nu}_0(v_0)=\boldsymbol{\nu}_0(v_1)$ *and with order "not too high", i.e.*

$$k_V<\max\{N^{-1}2^{-(n_{\lambda_{v_0}}+3)/\tau},(\gamma N\eta)^{-1}\},\qquad \gamma\stackrel{def}{=}\frac{4\Omega_1}{\Omega_2}\,. \tag{5.10}$$

The branch exiting from a resonance will be called a resonant branch, and the scale $n_{\lambda_{v_0}}$ *of the two branches entering and exiting the resonance will be called the resonance-scale. The degree of the propagator of the exiting branch will be called the degree* D_V *of the resonance.*

Even though a node v either with $\delta_v=0$, $\kappa_v\ge 1$ or with $\delta_v=1$, $\boldsymbol{\nu}_v=\mathbf{0}$ is not a cluster in the above sense (because it does not consist of branches) we shall nevertheless regard it as a cluster when there are only one incoming branch and one exiting branch of equal scale. *Therefore we shall also regard it as a resonance*, if $\kappa_v<\max\{N^{-1}2^{-(n_{\lambda_{v_0}}+3)/\tau},(\gamma N\eta)^{-1}\}$ when $\delta_v=0,k_v\ge 1$ and if $1<\max\{N^{-1}2^{-(n_{\lambda_{v_0}}+3)/\tau},(\gamma N\eta)^{-1}\}$ when $\delta_v=1,\boldsymbol{\nu}_v=\mathbf{0}$; note that the restriction that if $\delta_v=0=\kappa_v$ there are at least two branches entering v implies that no node with $\delta_v=0=\kappa_v$ can be a resonance.

DEFINITION. *A resonance will be called "strong" if* $p(v_0)=p(v_1)=0$.

All resonances on scale ≤ 0 are strong (see Remark 5.8).

5.8. It is important to note that a strong resonance of degree 2 is necessarily such that *also* the degree of the entering branch *must* be 2. No branch inside it can be of order 1 and no path can precede v_0. This is so because $D_{\lambda_{v_0}}=2$ implies $j_{v_0}>\ell$ (see the first of (5.3)) and $p(v_0)=0$ implies that $\alpha_{v_0}=2,k_{v_0}=0,k'_{v_0}=0$ so any path preceding v_0 would necessarily imply the contradiction $p(v_0)=1$. Also if $D_{\lambda_{v_1}}=1$, $D_{\lambda_{v_0}}=2$ one must have $j_{v_1}=\ell$ hence $k_{v_1}=0$ (otherwise $p(v_0)>0$) so that $k'_{v_1}=0$: *but* $\alpha_{v_1}<2$ and k'_{v_1} must be odd. The cases $D_{\lambda_{v_0}}=1$, $D_{\lambda_{v_1}}=1,2$ are both allowed.

5.9. Given a tree ϑ, let V be a resonance (if there are any) with entering branch λ_{v_1} of degree $D_{\lambda_{v_1}}=2$. Then consider the family of all trees which can be obtained from ϑ by detaching the part of the tree having λ_{v_1} as root branch and reattaching it to all the remaining nodes *internal to* V *but external to the resonances contained inside the cluster* V (if any); to the just defined set of trees we add all the trees obtained by reversing simultaneously the signs of the latter modes of the nodes (this can be done as the sum of the mode vectors $\boldsymbol{\nu}_w$ of such nodes, $w\in V$, vanishes). The set of all the so obtained trees will be denoted $\mathcal{F}_V(\vartheta)$.

The definition of resonance and the strong Diophantine condition insures that all the trees so constructed have a well defined value (*i.e.* no division by zero occurs in evaluating it with the above rules); see the Remark 5.10, (1), below.

If the entering branch λ_{v_1} of the resonance has degree $D_{\lambda_{v_1}}=1$ then also the exiting branch λ_{v_0} has degree $D_{\lambda_{v_0}}=1$, and we collect together with the considered tree also the tree which is obtained from ϑ through the following operation. Replace the resonance V with *a single node* v carrying labels $\delta_v=0$ and $\kappa_v=k_V$, if k_V is the order of the resonance. The set of all the so obtained trees will be denoted by $\mathcal{F}_V(\vartheta)$: the definition of the class $\mathcal{F}_V(\vartheta)$ will therefore depend on the degree of the branch entering V.

Then repeat the above operations for all resonances in ϑ. Thus a class $\mathcal{F}(\vartheta)$ has been constructed and the number of elements of $\mathcal{F}(\vartheta)$ is bounded by the product $\prod_V 2\mathcal{N}_V$ of the numbers $\mathcal{N}_V$ of branches in each resonance V *which are not* contained inside inner resonances. The latter product is bounded by $\exp\sum_V 2\mathcal{N}_V\le\exp 2m$; the $\mathcal{F}(\vartheta)$ can be obtained starting from any of its elements (which therefore we shall call *representatives* of the class): this is again a *consequence of the strong Diophantine condition*, see [Ge2].

5.10. REMARKS. (1) The strong Diophantine condition plays a rôle here that should be stressed. In fact one checks that because of it the scale of a branch inside a resonance *cannot* change too much, as one considers the different members of a given family. Not enough to change the sets of branches that belong to a given resonance and insures that the different families of trees *do not overlap*: for this reason the strong

Diophantine condition leads to a simplification of the analysis (the simplification in the simpler case of the KAM theory). A simplification that is however not major (as explained informally in [G1] and as shown in [GG], see footnote 7 above).
(2) To see how the above difficulty is bypassed by using the alternative approach of [GM1,GM2], we refer to the conclusive comments in [GM1], §3.

5.11. Consider trees with $p(v_0)=0$, if v_0 is the highest node of the tree; then the expression of each tree value contains a product like (5.1). As mentioned in the introduction *we consider only trees without leaves.*
Since the leaf values factorize with respect the product (5.1), they can be dealt with separately, and no overlap arises with the cancellation mechanisms acting on the product (5.1): so that leaves can be easily taken into account; see §A3.3 in Appendix A3 (see also [G1,Ge1,Ge2]).
The counterterms can also be explicitly expanded in terms of tree values, according to (3.2), which again we can imagine to have no leaves, (see however the comments in §A3.3 below).
The cancellation mechanisms described in [Ge1,Ge2] (and recalled in Appendix A3) lead to the bound (on a given family $\mathcal{F}(\vartheta)$ described above, in §5.9), see (5.1), (4.30), (4.23)

$$\left(\frac{1}{\eta^{1/2}}\right)^{2m}\Big[(4N^3\mathcal{C}')^m 2^{4m} e^{2m}\prod_{n\le 0}\big(C_0^{2N_n^2}2^{-2nN_n^2}\big)\big(C_0^{N_n^1}2^{-nN_n^1}\big)\Big]\cdot \qquad (5.11)$$
$$\cdot\Big[\prod_{n\le 0}\prod_{T,\,n_T=n}\prod_{i=1}^{m_T^1(n)}2^{(n-n_i+3)}\prod_{i=1}^{m_T^2(n)}2^{2(n-n_i+3)}\Big]\,,$$

where
- $\mathcal{C}'=\max\{(2g_0/\Omega_1)^2,4^2\}\mathcal{C}$, with $\mathcal{C}$ the dimensionless constant defined in (5.6);
- m is the number of nodes $v\ge v_0$;
- N_n^j is the number of propagators on scale n and of degree j in ϑ, which can be written as

$$N_n^j=\bar N_n^j+\sum_{\substack{T\\ n_T=n,\,D_T=j}}(-1)+\sum_{\substack{T\\ n_T=n}}m_T^j(n)\,, \qquad (5.12)$$

where $m_T^j(n)$ is the number of resonances on scale n and degree j (*i.e.* with entering branch having a propagator of degree j) contained inside the cluster T;
- the terms $\bar N_n^j$, $j=1,2$, which count the number of propagators *which do not correspond to resonant branches* plus the number of clusters on scale n and of degree j in ϑ, satisfy the bounds

$$\sum_{j=1}^{2}\bar N_n^j\le 4mN2^{(n+3)/\tau},\qquad \sum_{n=-\infty}^{0}\sum_{j=1}^{2}\bar N_n^j\le 4m\gamma N\eta\,, \qquad (5.13)$$

(with $\gamma=4\Omega_1/\Omega_2$) which are proven in Appendix A2;
- the first square bracket in (5.11) is the bound on the product of individual elements in the family $\mathcal{F}(\vartheta)$ times the bound on their number $\prod_V 2\mathcal{N}_V<e^{2m}$, see above.
- the second square bracket term is the part coming from the maximum principle, (in the form of Schwarz's lemma), applied to bound the sums of the tree values ("*resummations*") over the classes $\mathcal{F}(\vartheta)$ introduced above: this is a *non trivial product of small factors* that arise from the cancellations associated with the resummations, see Appendix A3. In (5.11) n_i is the scale of the cluster V_i which is the i–th resonance inside T, as in [Ge2];
- the $\eta^{-m/2}$ arises as a lower bound on the small divisors of the form $\boldsymbol{\omega}'\cdot\boldsymbol{\nu}$ on scale $n=1$ (for $n=1$ we use the better bound $|\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}|\ge 2^2\eta^{\frac12}$).

5.12. REMARK. The first bound (5.13) holds for all n and for all Hamiltonians of the form (2.1). On the contrary the second bound in (5.13) will follow from the fact that the rotation vector $\boldsymbol{\omega}_0$ has the form (5.8), with η small, and will be used to control the (huge) factors C_0 in (5.11).

5.13. Hence by substituting (5.12) and the first of (5.13) into (5.11) we see that, for $j = 1, 2$, the $m_T^j(n)$ is taken away by the first factor in $2^{jn}2^{-jn_i}$, while the remaining 2^{-jn_i} are compensated by the -1 before the $+m_T^j(n)$ in (5.11) taken from the factors with $T = V_i$ (note that there are always enough -1's), and therefore (5.11) is bounded by

$$\left(\frac{2}{\eta^{1/2}}\right)^{2m} (4N^3\mathcal{C}')^m e^m 2^{4m} 2^{8m} C_0^{8m\gamma N\eta} \prod_{n=-\infty}^{0} 2^{-8mNn2^{(n+3)/\tau}} , \tag{5.14}$$

because the product of the factors C_0 in (5.11) can be bounded by using the second of (5.13), since the product does not contain the $n = 1$ factor). The last product in (5.14) is bounded by

$$\prod_{n=-\infty}^{0} 2^{-8mNn2^{(n+3)/\tau}} \le \exp\left[8mN2^{3/\tau}\log 2\sum_{p=1}^{\infty} p2^{-p/\tau}\right] , \tag{5.15}$$

hence, by adding the remark that the perturbation degree k and the number of tree nodes m are related by $m < 2k$, a bound on the sum over all the subtrees of order k with $p(v_0) = 0$, $\boldsymbol{\nu}(v_0) = \boldsymbol{\nu}$ (recalling that the number of trees with m nodes is $< 4^m m!$) is

$$\Delta_k \stackrel{def}{=} \left|\frac{1}{|\mathcal{F}(\vartheta)|}\sum_{\vartheta'\in\mathcal{F}(\vartheta)}\prod_{v\in\vartheta'} \bar{F}_{\underline{\nu}_v} G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)]\,\tilde{y}_v\right| \le B_0^{2k}\eta^{-2k} , \tag{5.16}$$

for some positive constant B_0. The normalization constant $|\mathcal{F}(\vartheta)|$ is introduced in order to avoid overcountings: in fact if $\vartheta' \in \mathcal{F}(\vartheta)$ then $\vartheta \in \mathcal{F}(\vartheta')$, so that, without dividing by $|\boldsymbol{\Phi}(\vartheta)|$ in (5.16), each tree would be counted $|\mathcal{F}(\vartheta)|$ times.

If $C_0^{-1} = \eta^{-\frac{1}{2}}C(\eta)$ is chosen as in the statement of Theorem 1.4, an explicit calculation gives the bound on (5.11) of the form $(\eta^{-\frac{1}{2}})^{4k}B_0^k$, $k \ge 1$, and

$$B_0 = 2^{18}(4N^3\mathcal{C}')\,\exp\left[2 + 4\gamma N\eta\log\eta + 8s\gamma N\eta^{\frac{1}{2}} + 8N2^{3/\tau}\log 2\sum_{p=1}^{\infty} p2^{-p/\tau}\right] , \tag{5.17}$$

which is bounded uniformly in η (for $\eta \le 1$).

5.14. In the previous section trees with $p(v_0) = 0$ have been considered; in particular only the contributions (5.1) arising from the value (4.30), once the corresponding tree has been deprived of leaves and counterterms, have been bounded and the bound (5.16) has been obtained through a suitable resummation operation. In such a case the sum over the labels (k_v', k_v) is trivial because the condition $p(v_0) = 0$ imposes that only a few values (up to three per node) can be assumed by the hyperbolic mode labels; also the sum over the mode labels $\underline{\nu}_v$ cannot create any problems. In fact for any node v one has $|\underline{\nu}_v| \le N$ and $|n_v| \le N_0$ (see the eighth item in §4.1).

The cases $p(v_0) \ne 0$ as well as those involving graphs containing leaves or counterterms can be treated in the same manner as already done in [G1,Ge2]. We provide, in Appendix A4, a quick description of the construction of the analyticity bound $\varepsilon_0 = D^{-1}$ with

$$e_0^{-1} = D = \left[B2^6\ell(2N+1)^{2\ell-1}(2N_0+1)\right]^2 , \qquad B = \max(B_0\eta^{-1}, B_1) \tag{5.18}$$

and B_1 is a suitable numerical constant.

The part of Theorem 1.4 not concerning the connection between the average action $\mathbf{A}'$ and the rotation vector $\boldsymbol{\omega}'$ nor the splitting size follows.

5.15. Determining the exact splitting size (*i.e.* the leading behavior asymptotically as $\eta\to 0$ with $\varepsilon<B\eta^2$) is *not* trivial because of the existence of major cancellations in the evaluation of the determinant of the splitting matrix; however the analysis in [GGM2] dealt with this question in detail: in the latter paper remarkable cancellations are exhibited and an exact formula for the splitting angles is derived (see (7.19) of [GGM2]).

One gets the results in the last item of Theorem 1.4 simply if [GGM2] and the first part of Theorem 1.4 (to estimate the remainders) are used: then the claimed bounds on the splitting follow immediately (see Remark 1.5). In [GGM3] an improvement of lemma 1 and lemma 1' of [CG] was used instead to control the density of tori in phase space (the lemmata in [CG] were, as such, useless already in the case in [GGM2] because they would require that ε be far smaller than the ε_0 of Theorem 1.4); see [GGM3], where this is discussed in detail and differs from our case only because it relied on a theorem weaker than Theorem 1.4 above (as the radius of convergence estimate there is proportional to η to the power $\frac{9}{2}+$ rather than our 2).

5.16. Remarks. (1) The bound (5.16) and the discussion in §5.14 imply the convergence of the perturbative expansions for the parametric equations of the invariant tori (for the Hamiltonian (1.1)), if $|\varepsilon|<\varepsilon_0=O(\eta^2)$. This bound on the convergence radius should be compared with the value given by [GGM3], which, for $\mathcal{N}=O(\eta^{-\frac{1}{2}})$, gives $\varepsilon_0=O(\eta^{\frac{9}{2}}/\log^2\eta^{-1})$. *As usual the Lindstedt series gives a much better estimate than the classical method* (*i.e.* an exponent 2 versus ~ 4.5). We do not see immediately how to improve substantially the classical estimate without important changes in the architecture of the proof of [GGM3], although this should be possible; on the other hand, from the above analysis, $\varepsilon_0=O(\eta^2)$ might be close to an optimal result. If so it should be no surprise that our analysis is so delicate.

(2) In the Hamiltonian (1.1),(2.1) the polynomial dependence of the interaction on the rotators angles has very likely a purely technical motivation (as it simplifies the analysis) and could probably be relaxed into a more general analytical dependence, as in [BCG]. On the contrary the hypothesis that the perturbation is a trigonometric polynomial of degree N_0 in φ is fundamental to get the correct asymptotic behavior, in order to apply the results in [GGM2], where the dominance of Mel'nikov integral is proven *provided the perturbation is polynomially small in a power of* η^{N_0} (so that the results of [GGM2] become meaningless for $N_0\to\infty$).

(3) A bound of the form (5.16) holds under the weaker condition that $C(\eta)\le e^{-s\eta^{-a}}$, with $a\le 1$ (see (5.17)).

(4) If q is defined as in §1.6 so that $|\varepsilon|C(\eta)^{q\eta}<1$ implies analyticity in ε, the above analysis gives that q can be taken $q=8\gamma N$.

In general all the bounds found so far are not uniform in N; in order to deal with the analytical case in the frame of the exploited formalism one should bound the small divisors by using the results of [GM2] or Eliasson-Siegel's bound (see for instance [BGGM]), and use explicitly as in [BCG] the exponential decay in $\boldsymbol{\nu}$ of the Fourier coefficients $f^1_{\underline{\boldsymbol{\nu}}}$.

(5) Note that we have convergence for $|\varepsilon|<O(\eta^2)$, while the asymptoticity of the splitting estimate follows only for $|\varepsilon|<O(\eta^\zeta)$ (with $\zeta=2(N_0+3)$), which *much smaller* for $\eta\to 0$. Then the question is: what will be the asymptotics for ε small enough to be in the convergence domain but too large to be in the domain of the asymptotic result? There is some evidence that there is a critical value T_c such that, if $\varepsilon=\eta^T$, then for $T>T_c$ the asymptotic formula that we can prove only for $T>\zeta$ holds, but for $T=T_c$ it is modified remaining qualitatively of the same size $O(e^{-\frac{1}{2}\eta^{-\frac{1}{2}}})$ and for $T<T_c$ it becomes qualitatively different. The analogy with the critical point scaling phenomena seems to be substantial and, keeping in mind that the above theory can be interpreted as a field theory, see [GGM1], one would say that the region $T>T_c$ is described by a trivial fixed point; a non trivial fixed point describes the case $T=T_c$ and another "low

temperature" fixed point describes the cases $T < T_c$. Evidence in this direction comes also from the theory of the standard map and its developments, [La,Gel2,Gel1].

Acknowledgments.

This work is part of the research program of the European Network on "Stability and Universality in Classical Mechanics", # ERBCHRXCT940460. Supported in part by CNR-GNFM and Rutgers University.

Appendix A1. Counterterms

A1.1 To order h one can write, by using (2.12), (2.25) and (4.11),

$$\begin{aligned}\Xi_-^{h\sigma}(t) = &\int_{\sigma\infty}^{t} d\tau\, w_0^1(t,\tau)\left(\Phi_+^{h\sigma}(\tau;\gamma_1(g_0),\ldots,\gamma_{h-1}(g_0)) + \gamma_h(g_0)\sin\varphi^0(\tau)\right) \\ &+ \int_{\sigma\infty}^{0} d\tau\, w_0^0(t,\tau)\left(\Phi_+^{h\sigma}(\tau;\gamma_1(g_0),\ldots,\gamma_{h-1}(g_0)) + \gamma_h(g_0)\sin\varphi^0(\tau)\right)\;,\end{aligned} \tag{A1.1}$$

where $\Phi_+^{h\sigma}(t;\gamma_1(g_0),\ldots,\gamma_{h-1}(g_0))$ takes into account all the contributions to $\Phi_+^{h\sigma} = (J_0 g_0^2)^{-1} F_+^{h\sigma}$ except the only one explicitly depending on $\gamma_h(g_0)$, which is given by $\gamma_h(g_0)\sin\varphi^0(\tau)$.

We shall impose, recursively, that the contributions to the integrands in (A1.1) arising from amputated trees (see comments after (4.31)) with $(\boldsymbol{\nu}_0(v_0), p(v_0)) = (\mathbf{0}, 0)$, *without leaves and without end nodes bearing a counterterm label* (see Remark A1.2 below), compensate exactly the contributions due to the trees with a single node representing a counterterm of order h (*i.e.* the terms with $\gamma_h(g_0)$ in (A1.1)). The first described type of contributions can be written

$$\mathcal{W}_1 = w_\ell^1(t,\tau)\,\Phi_+^{h\sigma}(\tau;\gamma_1(g_0),\ldots,\gamma_{h-1}(g_0)) \equiv \sum_{\boldsymbol{\nu}\in\mathbb{Z}^{\ell-1}}\sum_{p=-1}^{\infty} \tilde{\mathcal{W}}_1(\boldsymbol{\nu},p)\, e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}t - pg_0\sigma t} \tag{A1.2}$$

in the first integral in (A1.1) and

$$\mathcal{W}_0 = w_\ell^0(t,\tau)\,\Phi_+^{h\sigma}(\tau;\gamma_1(g_0),\ldots,\gamma_{h-1}(g_0)) \equiv \sum_{p=-1}^{\infty}\sum_{\boldsymbol{\nu}\in\mathbb{Z}^{\ell-1}} \tilde{\mathcal{W}}_0(\boldsymbol{\nu},p)\, e^{i\boldsymbol{\omega}'\cdot\boldsymbol{\nu}t - pg_0\sigma t} \tag{A1.3}$$

in the second one. Imposing that such contributions are canceled by the contributions with $p = 0$ arising from $w_\ell^1(t,\tau)\gamma_h(g_0)\sin\varphi^0(\tau)$ and $w_\ell^0(t,\tau)\gamma_h(g_0)\sin\varphi^0(\tau)$ gives our prescription on how to fix $\gamma_h(g_0)$.

Since two integrals are involved (one for $\rho_{v_0} = 1$ and one for $\rho_{v_0} = 0$), the first time dependent and the second time independent, two conditions may seem to be required: however note that $\sin\varphi^0(t) = 2\sinh g_0 t/\cosh^2 g_0 t$ is expanded in odd powers of $x = e^{-\sigma g_0 t}$, hence the only terms appearing in (A1.2) and (A1.3) which can contribute to $p = 0$ are, in both cases, those involving $y^{(-1)}(k'_{v_0}, k_{v_0})$, with $k_{v_0} = -1$. This means that the contributions with $p = 0$ arising from (A1.2) and (A1.3) are equal, so that no compatibility problem arises.

Note that an expression analogous to (A1.1) is obtained also for $\Xi_+^{h\sigma}(t)$; however the terms with $p = 0$ have the same form as in (A1.1) (see (4.17) for $j_{\lambda_{v_0}} = \ell$ and for $j_{\lambda_{v_0}} = 0$), so that if no term with $(\boldsymbol{\nu}_0(v_0), p(v_0) = (\mathbf{0}, 0))$ contributes to $\Xi_-^{h\sigma}(t)$, then also no term with $(\boldsymbol{\nu}_0(v_0), p(v_0)) = (\mathbf{0}, 0)$ contributes to $\Xi_+^{h\sigma}(t)$.

A1.2. REMARK. It may seem strange that we exclude from the definition of the counterterms trees with leaves: in fact one can imagine to realize trees with $(\boldsymbol{\nu}(v_0), p(v_0)) = (\mathbf{0}, 0)$ also by trees which contain leaves with a stalk bearing a label $j_w > \ell$ or $j_w = \ell$ and internal momenta $(\boldsymbol{\nu}', p')$. Such terms would give rise to $\boldsymbol{\alpha}$–dependent counterterms which of course are not allowed: however it turns out that the sum over all contributions to tree values of trees with $(\boldsymbol{\nu}(v_0), p(v_0)) = (\mathbf{0}, 0)$ from such trees cancel *exactly*: this is explained, together with the other cancellations built in our algorithm, in Appendix A3 (see §A3.5 in particular).

A1.3. Let us consider the first integral in (A1.1). Corresponding to the node v_0 of each tree whose value contributes to $\Xi_{-}^{h\sigma}(t)$ there is a coefficient $\tilde{y}_{n_{v_0}}(k'_{v_0}, k_{v_0})$, see (4.14), (4.17). Then from (A1.1) and the just formulated condition to impose we obtain

$$\sum_{\substack{\vartheta\in\mathcal{T}_{\mathbf{0},h},\alpha_{v_0}=-1\\ p(v_0)=0,k'_{v_0}=1}} \overline{\mathrm{Val}}(\vartheta) + \gamma_h(g_0)\, w^1_\ell(t,\tau)\, \sin\varphi_0(\tau)\big|_{k'=1,p=0} = 0 \;, \tag{A1.4}$$

where the sum is over the set $\mathcal{T}_{\mathbf{0},h}$ of all trees of order h and momentum $\boldsymbol{\nu}(v_0) = \mathbf{0}$, with $\boldsymbol{\nu}_0(v_0) = \mathbf{0}$ (see Remark A1.2), $p(v_0) = 0$, $j_{v_0} = \ell$ and $k'_{v_0} = 1$; hence if $p(v_0) = 0, j_{v_0} = \ell$, one must have $k_{v_0} = -1$, hence $\alpha_{v_0} = -1$ and $k'_{v_0} = 1$ which is a possible case indeed.

A trivial calculation (just take into account that $y^{(-1)}_{n_v}(1,-1) = \sigma/2$ and $\sin\varphi^0(\tau) = 4\sigma x + O(x^3)$) gives

$$w^1_\ell(t,\tau)\, \sin\varphi_0(\tau)\big|_{k'=1,p=0} = 2 \;, \tag{A1.5}$$

so that (4.32) follows; the above follows [Ge2], page 287.

Appendix A2. (Improved) resonant Siegel-Bryuno's bound

A2.1. We follow the idea of Pöschel, [Pö] (see also [G1, GG,Ge2]). In the discussion, we focus on the scale labels, so that it is quite irrelevant which value the $p(v)$'s, $v \in \vartheta$, assume, and therefore which resonances are strong and which are not.

Calling $N^*_n(\vartheta)$ the number of non resonant branches carrying a scale label $\le n$, in a tree ϑ with m nodes, we shall prove first that

$$N^*_n(\vartheta) \le 2mE_n - 1 \;, \qquad E_n \stackrel{def}{=} N2^{(3+n)/\tau}, \qquad n \le 1 \;, \tag{A2.1}$$

provided that $N^*_n(\vartheta) > 0$, and

$$N^*_0(\vartheta) \le 2m\gamma N\eta - 1 \;, \qquad \gamma \stackrel{def}{=} 4\Omega_1/\Omega_2 \;, \tag{A2.2}$$

if $N^*_0(\vartheta) > 0$.

Define, as in §5.7, $\boldsymbol{\omega}_0 = (1, \eta^{-1}(\Omega_1 + \eta^{\frac{1}{2}} J^{-1} A_1)^{-1}\Omega_2)$. Then $C_0|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}| > |\boldsymbol{\nu}|^{-\tau}$ for all $\mathbf{0} \ne \boldsymbol{\nu} \in \mathbb{Z}^{\ell-1}$; see (5.7). Assume also η so small that $C_0 \ge 2$, (this is not restrictive as we are interested in $\eta \to 0$).

Set $E_n \equiv N2^{(n+3)/\tau}$ as in (A2.1). Note that if $m \le E_n^{-1}$ one has $N^*_n(\vartheta) = 0$. In fact $m \le E_n^{-1}$ implies that, for all $v \in \vartheta$, $|\boldsymbol{\nu}_0(v)| \le NE_n^{-1}$, *i.e.* $C_0|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}_0(v)| \ge (N^{-1}E_n)^\tau = 2^{n+3}$, so that there are no clusters T with $n_T = n$. Note also that if $m \le (\gamma N\eta)^{-1}$, with $\gamma = 4\Omega_1/\Omega_2$, then $N^*_0 = 0$, as $|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}_0(v)| \ge 1$ for all $v \in \vartheta$ in such a case.

A2.2. Let us prove first the inequality (A2.1). If ϑ has the root branch either with scale $> n$, or with scale $\le n$ and resonant, then calling $\vartheta_1, \vartheta_2, \ldots, \vartheta_k$ the subtrees of ϑ ending into the highest node v_0 of ϑ and with

$m_j > E_n^{-1}$ nodes, $j = 1, \ldots, k$, one has $N_n^*(\vartheta) = N_n^*(\vartheta_1) + \ldots + N_n^*(\vartheta_k)$ and the statement is inductively implied from its validity for $m' < m$ and from the just proved fact that $N_n^*(\vartheta) = 0$ if $m \le E_n^{-1}$.

If the root branch is on scale $\le n$ and non resonant, one has $N_n^*(\vartheta) \le 1 + \sum_{i=1}^k N_n^*(\vartheta_i)$: if $k = 0$ the statement is trivial, if $k \ge 2$ the statement is again inductively implied by its validity for $m' < m$. If $k = 1$ one has $N_n^*(\vartheta) \le 1 + 2m_1 E_n - 1$, hence we once more have a trivial case unless the order m_1 of ϑ_1 is $m_1 > m - (2E_n)^{-1}$: but in the latter case we shall show that the root branch of ϑ_1 has scale $> n$.

Accepting the last statement (which will be proved below), one will obtain $N_n^*(\vartheta) = 1 + N_n^*(\vartheta_1) = 1 + N_n^*(\vartheta'_1) + \ldots + N_n^*(\vartheta'_{k'})$, where ϑ'_j's are the k' subtrees ending into the highest node of ϑ'_1 with orders $m'_j > E_n^{-1}$, $j = 1, \ldots, k'$. Going once more through the analysis the only non trivial case is if $k' = 1$ with the root branch of ϑ'_1 non resonant; and in such case $N_n^*(\vartheta'_1) = N_n^*(\vartheta''_1) + \ldots + N_n(\vartheta''_{k''})$, *etc.*, until we reach a trivial case or a tree of order $\le m - (2E_n)^{-1}$.

It remains to check that if $m - m_1 < (2E_n)^{-1}$ then the root branch of ϑ_1 has scale $> n$. Let us proceed by *reductio ad absurdum.* Suppose that the root branch of ϑ_1 is on scale $\le n$. Then $C_0|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}_0(v_0)| \le 2^n$ and $C_0|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}_0(v_1)| \le 2^n$, if v_1 is the highest node of ϑ_1. Hence $C_0|\boldsymbol{\omega}_0 \cdot (\boldsymbol{\nu}_0(v) - \boldsymbol{\nu}_0(v_1))| < 2^{n+1}$ (equality would imply violation of the strong Diophantine property, (5.7)), and the Diophantine condition implies that

$$|\boldsymbol{\nu}_0(v_0) - \boldsymbol{\nu}_0(v_1)| > 2^{-(n+1)/\tau} \equiv \delta \,, \tag{A2.3}$$

because $\boldsymbol{\nu}_0(v_0) \ne \boldsymbol{\nu}_0(v_1)$ (the root branch of ϑ being supposed non resonant). But $m - m_1 < (2E_n)^{-1}$, so that $|\boldsymbol{\nu}_0(v_0) - \boldsymbol{\nu}_0(v_1)| < (2E_n)^{-1}N < 2^{-1}2^{-(n+3)/\tau} = 2^{-(1+2/\tau)}\delta < \delta$, which contradicts inequality (A2.3).

A2.3. Let us prove now (A2.2). If ϑ has the root branch either with scale 1, or with scale ≤ 0 and resonant, then calling $\vartheta_1, \vartheta_2, \ldots, \vartheta_k$ the subtrees of ϑ ending into the highest node v_0 of ϑ and with $m_j > (\gamma N \eta)^{-1}$ nodes, $j = 1, \ldots, k$, one has $N_0^*(\vartheta) = N_0^*(\vartheta_1) + \ldots + N_0^*(\vartheta_k)$ and the statement is inductively implied from its validity for $m' < m$ and from the fact that $N_0^*(\vartheta) = 0$ if $m \le (\gamma N \eta)^{-1}$.

If the root branch is on scale ≤ 0 and non resonant, one has $N_0^*(\vartheta) \le 1 + \sum_{i=1}^k N_0^*(\vartheta_i)$: if $k = 0$ the statement is trivial, if $k \ge 2$ the statement is again inductively implied by its validity for any $m' < m$. If $k = 1$ we once more have a trivial case unless the order m_1 of ϑ_1 is $m_1 > m - (2\gamma N\eta)^{-1}$: but in the latter case the root branch of ϑ_1 has scale 1.

Accepting the last statement (which will be proved below), one will obtain $N_0^*(\vartheta) = 1 + N_0^*(\vartheta_1) = 1 + N_0^*(\vartheta'_1) + \ldots + N_0^*(\vartheta'_{k'})$, where ϑ'_j's are the k' subtrees ending into the highest node of ϑ'_1 with orders $m'_j > (2\gamma N\eta)^{-1}$. Going once more through the analysis the only non trivial case is if $k' = 1$ and in that case $N_0^*(\vartheta'_1) = N_0^*(\vartheta''_1) + \ldots + N_0(\vartheta''_{k''})$, *etc.*, until we reach a trivial case or a tree of order $\le m - (2\gamma N\eta)^{-1}$.

It remains to check that, if $m - m_1 < (2\gamma N\eta)^{-1}$, then the root branch of ϑ_1 has scale 1. Suppose that the root branch of ϑ_1 is on scale ≤ 0. Then $p(v_1) \ne 0$ and $|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}_0(v_0)| \le 1/4$, $|\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}_0(v_1)| \le 1/4$, if v_1 is the highest node of ϑ_1, *i.e.*

$$|\boldsymbol{\omega}_0 \cdot (\boldsymbol{\nu}_0(v_0) - \boldsymbol{\nu}_0(v_1))| \le 1/2 \,. \tag{A2.4}$$

As the root branch of ϑ is supposed non resonant, then $m - m_1 < (2\gamma N\eta)^{-1}$ implies that $0 < |\boldsymbol{\nu}_0(v_0) - \boldsymbol{\nu}_0(v_1)| < (2\gamma N\eta)^{-1}N = (2\gamma\eta)^{-1}$, so that one would have $|\boldsymbol{\omega}_0 \cdot (\boldsymbol{\nu}(v_0) - \boldsymbol{\nu}(v_1))| \ge 1$, which is contradictory with the inequality (A2.4).

A2.4. A similar induction can be used to prove that if the number of branches on scale n is $N_n(\vartheta) > 0$ then the number $p_n(\vartheta)$ of clusters of scale n verifies the bound

$$p_n(\vartheta) \le 2mN2^{(n+3)/\tau} - 1 \,. \tag{A2.5}$$

In fact this is true for $m \le E_n^{-1}$, if E_n is defined as in §A2.1. Otherwise, if the highest tree node v_0 is not in a cluster on scale n, one calls $\vartheta_1, \ldots, \vartheta_k$ the subtrees ending into v_0, and one has $p_n(\vartheta) = p_n(\vartheta_1) + \ldots + p_n(\vartheta_k)$, so that the statement follows by induction. If v_0 is in a cluster V of scale n, and $\vartheta_1, \ldots, \vartheta_k$ are the subtrees entering the cluster containing v_0 and with orders $m_j > E_n^{-1}$, one will find $p_n(\vartheta) = 1 + p_n(\vartheta_1) + \ldots + p_n(\vartheta_k)$.

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 and it will have a propagator of scale $\le n-1$. Therefore the cluster V must contain at least E_n^{-1} nodes. This means that $m_1 \le m-(2E_n)^{-1}$.
Finally, the bound

$$\sum_{n=-\infty}^{0} p_n(\vartheta) \le 2m\gamma N\eta - 1 \tag{A2.6}$$

is a trivial consequence of (A2.2).

A2.5. Let $\bar N_n^* \le N_n^*$ be the number of non resonant branches on scale n. Then if $\bar N_n$ is the number of non resonant branches *plus* the number of clusters on scale n, $\bar N_n^*$ verifies the bounds

$$\bar N_n^* = \left(\bar N_n^* + p_n\right) - p_n \equiv \bar N_n - p_n \le 4mN2^{(n+3)/\tau} - 2 - \sum_{\substack{T\\ n_T=n}} (1) \le 4mN2^{(n+3)/\tau} + \sum_{\substack{T\\ n_T=n}} (-1) \,. \tag{A2.7}$$

This proves that (A2.1) and (A2.5) imply an inequality analogous to the first of (5.13); likewise one derives an inequality similar to the second of (5.13) by combining (A2.2) and (A2.6).

Appendix A3. Cancellations between resonances

In this appendix we recall briefly the cancellation mechanisms of [Ge2]. We provide this as a guide to the reader and as a tune up of a fine points of the analysis of [Ge2] (the analysis in A3.2 is given here in full details while in [Ge2] it was left out).

A3.1. Consider a tree ϑ with a strong resonance V of order k_V. Let λ_{v_0} and λ_{v_1} be, respectively, the exiting and entering branches of V. There are two possibilities: either the degree of the propagator corresponding to the branch exiting from V is $D_{\lambda_{v_0}}=2$ or it is $D_{\lambda_{v_0}}=1$ (equivalently the degree of the resonance is either $D_V=2$ or $D_V=1$).
Let us discuss first the case in which the degree D_V of the resonance is $D_V=2$. Then $j_{v_0}>\ell$ (see (5.3) and the comments after the definition of strong resonance in §5.9) and, by following the notations of §5.1, we shall say that $\mathcal{P}=\emptyset$, *i.e.* there is no path $\mathcal{P}$ ending into v_0. It follows, from the properties of $\mathcal{P}$ discussed at the beginning of §5.1 above, that $p(v_1)=0$ implies $j\equiv j_{v_1}>\ell$ and $D_{\lambda_{v_1}}=2$ (see again (5.3)).
Consider all the trees belonging to the class $\mathcal{F}(\vartheta)$ which are obtained from ϑ by detaching the subtree having as branch root the entering branch λ_{v_1} of the resonance and attaching it to all the remaining nodes of V (see the definition of the class $\mathcal{F}_V(\vartheta)$ in §5.9). As a consequence of such an operation
• some of the branches internal to the resonance have changed the free momentum by an amount $\boldsymbol{\nu}_0(v_1)$, and
• if w is the node inside V to which the branch λ_{v_1} is attached and $j_{v_1}-\ell>0$, then $\bar F_{\underline{\nu}_w}$ (see (4.23)), has the form of an even function of $\underline{\nu}_w$ times a factor $(i\nu_{wj})$.
We shall call *resonance value* $\mathcal{R}_V$ the product of factors appearing in the definition of tree value and relative only to the nodes and branches internal to the resonance V:

$$\mathcal{R}_V = \bar F_{\underline{\nu}_{v_0}} y'_{v_0} \prod_{\substack{v\in V\\ v<v_0}} \bar F_{\underline{\nu}_v} y'_v G_v[\boldsymbol{\omega}'\cdot\boldsymbol{\nu}_0(v)] \,, \tag{A3.1}$$

and shall consider the resonance value as a function of the quantity $\mu\equiv\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}_0(v_1)$.
Then for $\mu=0$ a quantity proportional to $\sum_{w\in V}\boldsymbol{\nu}_{wj}$ is constructed, but such a quantity is vanishing by definition of resonance, as $j=j_{v_1}-\ell>0$.

If we sum also on an overall change of signs of the mode labels of the nodes internal to the resonance (by following the definition of the class $\mathcal{F}(\vartheta)$ given in §5.9), we obtain a zero contribution also to first order in μ (here the even parity of the perturbation f is essential, see [G1,Ge2]).

This can be seen by using the explicit form of the functions in (4.21), *i.e.* the coefficients listed in (4.24). Noting that in the present case *there cannot be any* $\mathcal{P}$ *inside* V the only propagators we can associate with the branches internal to V have the form of the two first terms of (5.3), so that, for $\mu=0$, *they are even functions of the mode labels.* Moreover in such a case the analysis in §5.1 shows that $\alpha_v=-1$, $\alpha_v=1$ and $\alpha_v=2$ imply, respectively, $k'_v=-k_v=1$, $k'_v=-k_v=-1$ and $k'_v=-k_v=0$ (the case $\alpha_v=0$ is not possible here): then no n_v labels appear in the coefficients $y_{n_v}^{(\alpha_v)}(k'_v,k_v)$ corresponding to the nodes $v\le v_0$ (see the list of coefficients in (4.24)). Therefore all the dependence on the n_v labels is through the factors $\bar F_{\underline{\nu}_v}$ in (4.23). This yields that there is an even number of the n_v (if there are any) corresponding to the nodes $v\in V$: two for each branch λ_v with $j_v=\ell$, by taking into account that $j_{v_0},j_{v_1}>\ell$, so that no change is produced by the sign reversal (since, by the parity properties of the Hamiltonians (1.1) and (2.1), one has also $f_{\underline{\nu}_v}^{\delta_v}=f_{-\underline{\nu}_v}^{\delta_v}$). This means that the resonance value is an even function of μ.

A3.2. Let us now consider the case in which the strong resonance is of degree $D_V=1$ and the tree ϑ has no leaves inside V. In such a case $\alpha_{v_0}=-1$ and $j_{v_0}=\ell$, hence $D_{\lambda_{v_0}}=1$ (see (5.3)): then a first order zero in μ will be enough. Moreover there is a $\mathcal{P}$ inside the resonance: we shall distinguish between the cases $v_1\notin\mathcal{P}$ and $v_1\in\mathcal{P}$.

Let us consider first the case $v_1\notin\mathcal{P}$ (in particular this is the case when $\mathcal{P}=v_0$, $k_{v_0}=0$, provided $k_V\ge 2$). In such a case $j_{v_1}>\ell$ and we can reason as above to obtain a first order zero. Note that in such a case there would be no cancellations between tree values of trees obtained by the sign reversal operation.

On the contrary, if $v_1\in\mathcal{P}$, then $k_{v_0}=-1$, and one has also $\alpha_{v_1}=-1$ and $j_{v_1}=\ell$. In this case consider together with the tree ϑ also the tree ϑ' obtained from ϑ by performing the following operation (recall the definition of $\mathcal{F}_V(\vartheta)$): replace the resonance V with a single node v carrying labels $\delta_v=0$ and $\kappa_v=k_V$ (if k_V is the order of the resonance V), then express the counterterm $\gamma_{\kappa_v}(g_0)$ associated with the node v in terms of trees. If ϑ_1 is the subtree having λ_{v_1} as root branch, then the values of the two considered trees ϑ and ϑ' can be written, respectively, as $\mathrm{Val}(\vartheta)=A(\vartheta)\mathcal{R}_V\mathrm{Val}(\vartheta_1)$ and $\mathrm{Val}(\vartheta')=A(\vartheta)[\gamma_{k_V}(g_0)\sigma/2]\mathrm{Val}(\vartheta_1)$, where $\sigma/2=y_v^{(-1)}(1,-1)$ and $A(\vartheta)$ takes into account the factors corresponding to all nodes *not* preceding v_0, and has the same value for both ϑ and ϑ'.

The resonance value $\mathcal{R}_V$, for $\mu=0$, can be written as

$$\mathcal{R}_V=\overline{\mathrm{Val}}(\vartheta_0)\,in_{v'_1}\ ,\qquad \text{for some } \vartheta_0\in\mathcal{T}_{\mathbf{0},k} \text{ with } p(v_0)=0,\ k_{v_0}=-1\ ,\tag{A3.2}$$

see the definitions (4.30) and (4.31) of tree value and the definition (A3.1) of resonance value: remember that we are considering resonances V with degree $D_V=1$, so that $k_{v_0}=-1$ and, as a consequence, $k'_{v_0}\ge 1$; see (4.14). The counterterm $\gamma_\kappa(g_0)$ can be represented in terms of trees as in (4.32); note that, if the tree contributing to $\gamma_\kappa(g_0)$ has $k_{v_0}=-1$, the condition $\alpha_{v_0}=-1$ implies that such a tree has a node $w>v_0$ with $k_w+k'_w=1$, while all the other nodes $v\ne w$ have $k_v+k'_v=0$.

Among the contributions in (4.32) to $\gamma_{k_V}(g_0)$ there will be a quantity $\overline{\mathrm{Val}}(\vartheta_2)$, where ϑ_2 will have the same topological form of ϑ_0 in (A3.2) with the node w such that $k_w+k'_w=1$ corresponding to the node $v'_1\in V$; then we denote both nodes by w.

Then $\overline{\mathrm{Val}}(\vartheta_0)$ will be related to $\overline{\mathrm{Val}}(\vartheta_2)$ by

$$\overline{\mathrm{Val}}(\vartheta_2)=-\left[\frac{y_{n_v}^{(\alpha_v)}(k'_w,k_w)|_{k'_w+k_w=1}}{y_{n_v}^{(\alpha_v)}(k'_w,k_w)|_{k'_w+k_w=0}}\right]\overline{\mathrm{Val}}(\vartheta_0)\ ,\tag{A3.3}$$

so that a look at the coefficients listed in and after (4.24) shows that the factor in square brackets in (A3.3) (when it is not vanishing) is equal to $4in_w\sigma$. The quantity $\overline{\mathrm{Val}}(\vartheta_2)$, in order to contribute to $\gamma_{k_V}(g_0)\sigma/2$,

has to be multiplied by a factor -4σ extra with respect to $\overline{\rm Val}(\vartheta_0)$, which, on the other hand, has to be multiplied by in_w in order to contribute to the resonance value $\mathcal{R}_V$ (see (4.32).
Then, for $\mu=0$, by summing the values of the two considered contributions one obtain

$$A(\vartheta)\Big[\overline{\rm Val}(\vartheta_0)\, in_w - \frac{1}{4\sigma}\overline{\rm Val}(\vartheta_2)\Big]{\rm Val}(\vartheta_1)\ , \tag{A3.4}$$

which is zero by (A3.3), so that a first order zero is obtained.

A3.3. If there are leaves, nothing changes in the discussion of §A3.1, as $k_{v_0}=0$ implies that only leaves w with $j_w>\ell$ are possible, and $\xi_w(k'_w,0)\equiv 1$ in such a case (see (4.18)).
In §A3.2, when discussing the case $v_1\in\mathcal{P}$, one has to take care of the case in which there is a leaf with highest node $\tilde w$ with $k'_{\tilde w}=1$ (such a leaf will be at the end of the path $\mathcal{P}$). In fact the resonances having as entering branch a branch of the path $\mathcal{P}$ cannot have any leaves with $k'_w=1$, while when considering the graphical representation for $\gamma_\kappa(g_0)$, there will be also contributions arising from trees containing a leaf: such contributions will be either of the form (4.32) with $\overline{\rm Val}(\vartheta_2)=\overline{\rm Val}(\vartheta_2) in_{v'_1}\xi_{v_1}(1,0)L^{h_1\sigma}_{\ell\boldsymbol{\nu}(v_1)}(\vartheta_2)$, or of the form $\gamma_\kappa(g_0)=\gamma_{\kappa-h_1}(g_0)\, in_{v'_1}\xi_{v_1}(1,0)L^{h_1\sigma}_{\ell\boldsymbol{\nu}(v_1)}(\vartheta_2)$, where $h_1\ge 1$, and ϑ'_2 is a suitable tree of order $k-h_1$. Then one realizes that the two contributions cancel exactly, so that no new case has to be discussed with respect to the analysis of §A3.2.

A3.4. The above discussion completes the proof of approximate cancellations of resonance values (*i.e.* of cancellations to first and second order, according to the degree of the resonant branch). The existence of cancellations, approximate to the first or second order, is all is needed to obtain the bound (5.16): the analysis continues exactly as in [Ge2] and is based on simple analyticity arguments that allow us to exploit, via the maximum principle, the fact that in a resonance with momentum $\boldsymbol{\nu}$ the functions of $\mu=\boldsymbol{\omega}_0\cdot\boldsymbol{\nu}$ that have been considered above have a zero in μ of order 1 or 2.
A complete analysis showing that the higher orders contributions (*i.e.* the part which does not cancel) can be performed as in [Ge2], Appendix B, and the final result is given by the bound (5.16) in §5.11.

A3.5. We shall show now that all contributions with $(\boldsymbol{\nu}_0(v_0),p(v_0))=(\mathbf{0},0)$ involved in the definition of the counterterms (see Appendix A1) must have automatically also $\boldsymbol{\nu}(v_0)=\mathbf{0}$. The analysis performed in §5.1 shows that in order to have $p(v_0)=0$ (for $j_{v_0}=\ell$), there can be any number of leaves with highest nodes w such that $j_w>\ell$ and only one leaf w with $j_w=\ell$ (contributing, respectively, $k'_w=0$ and $k'_w=1$ to $p(v_0)$).
Each time a leaf with $j_w>\ell$ appears, if we sum together the values of all trees obtained by detaching the leaf with its stalk, then reattaching it to all the other nodes of ϑ_f, we obtain a vanishing contribution: simply by the cancellation mechanism described in §A3.1 (assuring there the first order zero), *which, now, is an exact cancellation as the leaf does not contribute to the free momenta of the branches of ϑ_f, so that it does not modify the propagators.* So we can suppose that no leaf with $j_w>\ell$ is possible in trees involved in the determination of the counterterms.
In the same way, if we have a tree ϑ having a leaf with $j_w=\ell$ and $h_w=h-h_1$ (for some h_1), we can reason as in §A3.2 and consider, together with ϑ, also the tree formed by only one free node, carrying a counterterm label h_1 and bearing the same leaf as ϑ. The same cancellation mechanism described in §A3.2 apply now: again the only difference is that now the cancellation is exact (by the same reason as before).
This shows that no tree with leaves can contribute to $(\boldsymbol{\nu}_0(v_0),p(v_0))=(\mathbf{0},0)$, so that for such trees one has $\boldsymbol{\nu}(v_0)=\boldsymbol{\nu}_0(v_0)=\mathbf{0}$. This, together with the analysis in §A1.1, proves (A1.4) in §A1.2.

Appendix A4. Graphs with non zero total hyperbolic momentum, with leaves or with counterterms

A4.1 Consider first the cases $p(v_0) \neq 0$. In this case we consider the nodes $w < v_0$ with $p(w) = 0$ and which are the nearest to v_0: by construction all nodes z between v_0 and the just singled out nodes have $p(z) \neq 0$. Let us denote by $\tilde{\vartheta}$ the set of such nodes z and k_1 the sum of their order labels. The subtrees having as root branches the branches exiting from the nodes w can be considered as trees of the kind of the previous sections (*i.e.* with $p(w) = 0$), so that the integrations corresponding to their nodes can be performed and discussed as before and a bound $B_0^{2k_0}\eta^{-2k_0}$ follows, if $k_0 = k - k_1$.

All the other nodes (in $\tilde{\vartheta}$) have $p(z) \neq 0$, so that no real small divisor appears (*i.e.* the propagators are trivially bounded by 1). The *only delicate point to discuss* concerns the sum over the hyperbolic mode labels, but this can be done as in [Ge2], p. 292, (or in [G1], item 7 in Appendix A1), to which we refer for details beyond the summary that follows.

By noting that the Laurent expansion of each function (of x and x') appearing in (4.14) and (4.16) starts from $k \geq -1$ and $k' \geq= 1$, we can denote by M_1 the maximum of all such functions (multiplied by $1/x$ and $1/x'$ respectively when $k = -1$ and $k' = -1$) in a disk of radius $\lambda = 1/2$. If M_2 is the coefficient of the term with $k_v = -1$ or $k'_v = -1$, set $M = \max\{M_1, M_2\}$.

Consider the tree value (4.20). If $\sigma t < 1$, the first integral (corresponding to the highest node v_0) can be split into the sum of two integral, the first one from $\sigma\infty$ to $\sigma 1$ (here the value 1 is an arbitrarily chosen positive number) and the second one from $\sigma 1$ to t. Let us denote by $I_m(\vartheta)$ the first integral and $J_m(\vartheta)$ the second one, if m is the number of nodes in ϑ.

For the nodes $v \in \tilde{\vartheta}$, one has

(1) for each node the associated propagator is bounded by 1 (as $p(v) \neq 0$),

(2) $\prod_{v\in\tilde{\vartheta}} |\bar{F}_{\nu_v}| \leq (CN^2)^{2k_1}$,

(3) for each node v one has $|y_{n_v}^{(\alpha_v)}(k_v, k'_v)| \leq M\, 2^{k_v + k'_v}$,

(4) the last integration (on τ_{v_0}) produces a factor $\exp[-k'_{v_0} + p(v_0)] = \exp[-\sum_{v\in\tilde{\vartheta}}(k_v + k'_v)]$.

Then the sum over the hyperbolic mode labels can be performed and gives, for each node $v \in \tilde{\vartheta}$, a factor A^2, where

$$A = \sum_{k=-1}^{\infty} \left(\frac{2}{e}\right)^k = \frac{e^2}{2(e-2)} \,. \tag{A4.1}$$

The contribution to $I_m(\vartheta)$ arising from $\tilde{\vartheta}$ a bound $B_1^{2k_1}$ is obtained, with $B_1 = A^2 \mathcal{C} N^2 M$, so that, for $I_m(\vartheta)$ a bound

$$B^{2k}, \qquad B \geq \max\{B_0\eta^{-1}, B_1\} \tag{A4.2}$$

is obtained (see (5.6) for the meaning of $\mathcal{C}$).

By taking into account the integral $J_m(\vartheta)$, one can perform a splitting of the integration domain for the integrals corresponding to the nodes immediately preceding v_0 (now for each such nodes the second integral is from $\sigma 1$ to τ_{v_0}), and, iterating such a splitting, one finds that $\mathrm{Val}\,(\vartheta)$ can be written as sum of at most 2^m terms each of which has the form (for some integer p)

$$\left[\prod_{v\in\vartheta^*} \int_{\sigma 1}^{\tau_{v'}} d\tau_v\,(\ldots)\right] I_{m_1}(\vartheta_1)\ldots I_{m_p}(\vartheta_p) \,, \tag{A4.3}$$

where $\vartheta_1, \ldots, \vartheta_p$ are disjoint subtress of ϑ and ϑ^* is the set of the m_0 nodes in ϑ not belonging to any such subtrees and $m_1 + \ldots + m_p + m_0 = m$. The dots between the parentheses denote the product of the functions in

$$\prod_{w\in\vartheta} Y^{(\alpha_w)}(\tau'_w, \tau_w) \tag{A4.4}$$

which depend on τ_v, for a given $v \in \vartheta^*$, and therefore is a quantity bounded by 1 (see (4.14) and (4.16)). Note that the functions Y have *no singularity* as functions of their arguments $x = e^{\sigma g \tau}, x' = e^{\sigma g \tau'}$, when $x, x' = 1$ (or $\tau, \tau' = 0$), even though the values $x, x' = 1$ lie on the convergence circle (the singularities being at $x, x' = \pm i$). Furthermore each integration from $\sigma 1$ to $\tau_{v'}$, once the integrand has been bounded, gives 1, while the integrals $I_{m_j}(\vartheta_j)$, $j = 1, \ldots, p$ can be bounded as before.

Of course, if $\sigma t > 1$, the discussion is easier as no splitting of the integration domains is needed.

So we can conclude that a final bound $(2B)^{2k}$ is obtained for $\mathrm{Val}\,(\vartheta)$; so far neither leaves nor counterterms have been considered.

A4.2 Introducing the leaves and the counterterms, one sees (recall Remark 4.4) that the value of any tree ϑ can be always be written as the product of a factor like (4.28) times the product of the counterterms and of the leaf values; each counterterm can be decomposed in turn as sum of values of amputated trees (see (4.32)). As each leaf and each amputated tree can contain other leaves and counterterms we can iterate such a decomposition procedure, until, at the end, the value of the tree ϑ, with highest node v_0, turns out to be given by the product of factorizing terms which

(1) either are of the form (4.28), with $\rho_w = 0$ for any subtree with highest node $w < v_0$ and with $\rho_w = 0, 1$ if $w = v_0$,

(2) or differ from (4.28) simply because no integration is performed corresponding to the highest node.

The terms as in item (1) correspond to subtrees contributing to leaf values (for $w < v_0$) and to $\mathrm{Val}\,(\vartheta)$ itself (for $w = v_0$), while the terms as in item (2) correspond to amputated trees contributing to counterterms. Then a natural decomposition of the tree ϑ into subtrees $\tilde\vartheta$ (amputated or not) follows: each of such subtree contains neither counterterms nor leaves (by construction). Furthermore the subtrees contributing to leaves are linked to nodes of some other subtrees through their stalks, while the amputated subtrees are not linked to any node (as there is no branch exiting from the highest node). To keep memory of the node to which the counterterm label is attached we can draw a hatched line connecting the amputated subtree to such a node.

So for each subtree $\tilde\vartheta$ (amputated or not) one can reason as above and a bound B^{m_0} is obtained, if B is the same constant as before and if $m_0 \le 2k_0$ is the number of free nodes of the subtree. For all of them the resummation described in §5.9 has to be performed, to bound the values of the subtrees ϑ_v, with $p(v) = 0$, $v \in \tilde\vartheta$: such a resummation is taken into account by the constant B. By collecting together all bounds one obtains, for the (normalized) sum of the values of the all trees ϑ' generated by the resummations corresponding to the families $\mathcal{F}(\vartheta_v)$, a bound B^{2k}, if k is the order of ϑ.

Therefore we are left with the sum of all possible ways to arrange leaves and counterterms. The choice of the leaves is uniquely determined by the assignments of the labels ρ_v, $v \in \vartheta$, so that it gives a factor 2^m (recall that the number of nodes m is such that $m < 2k$).

In the same way one can deal with the counterterms: simply one has to distinguish between solid and hatched lines, so that another factor 2^m is produced.

A4.3 Then the sum over all the other labels can be performed, in the same way as for the contributions without leaves and without counterterms (see the beginning of this subsection). The sum over the hyperbolic modes has been taken into account by the constant B (see (A4.2)); moreover

- the sum over the mode labels is bounded by $(2N+1)^{m(\ell-1)}(2N_0+1)^m$,
- the sum over the angle labels is bounded by ℓ^m,
- the sum over the order labels is bounded by 2^m,
- the sum over the badge labels is bounded by 2^m.

Therefore, by taking into account that the momenta and the hyperbolic momenta are uniquely determined by the mode labels and, respectively, the hyperbolic mode labels and that the sums over the leaf labels and counterterms labels have been already considered, we are left with the sum of (unlabeled) numbered trees (see comments after (4.27)): but these are no more than $2^{2m} m!$, so that, both for $X^{k\sigma}_{j\nu}(t)$ and $\gamma_k(g_0)$, a final bound D^k is obtained, for some constant D: in terms of B the constant D is given by $D =$

$B\, 2^6 \ell\, (2N+1)^{(\ell-1)}(2N_0+1)$, *i.e.* (5.18). In particular one has that D is proportional to η^{-2}, as B is so.

Note that, as a matter of fact, we have bounded $\Xi^{h\sigma}_{j\boldsymbol{\nu}}(t)$ in (4.27) by neglecting the constraint on j and $\boldsymbol{\nu}$. Therefore, by making use of the fact that the Fourier coefficients with $|\boldsymbol{\nu}| > h(N+N_0)$ vanish at order h as a consequence of the trigonometric assumption on the perturbation f_1, see (2.2), the bound $(2B)^{2k}$ is a bound both for the Fourier coefficients of $X^\sigma(t;\boldsymbol{\alpha})$ and for the function $X^\sigma(t;\boldsymbol{\alpha})$ itself.

References

[BCG] Benettin, G., Carati, A., Gallavotti, G.: *A rigorous implementation of the Jeans–Landau–Teller approximation for adiabatic invariants*, Nonlinearity **10** (1997), 479–507.

[BGGM] Bonetto, F., Gallavotti, G., Gentile, G., Mastropietro, V.: *Lindstedt series, ultraviolet divergences and Moser's Theorem*, Annali della Scuola Normale Superiore di Pisa **26**, 545–593 (1998).

[CG] Chierchia, L., Gallavotti, G.: *Drift and diffusion in phase space*, Annales de l'Institut Henri Poincaré, B **60** (1994), 1–144; see also the *Erratum*, B **68** (1998), 135.

[CZ] Chierchia, L., Zehnder, E.,: *Asymptotic expansions of quasi-periodic motions*, Annali della Scuola Normale Superiore di Pisa Cl. Sci. (4) **16** (1989), 245–258.

[E] Eliasson, L.H.: *Absolutely convergent series expansions for quasi-periodic motions*, Mathematical Physics Electronic Journal **2** (1996), http:// mpej.unige.ch.

[G1] Gallavotti, G.: *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review*, Reviews on Mathematical Physics **6** (1994), 343–411.

[G2] Gallavotti, G.: *Twistless KAM tori*, Communications in Mathematical Physics **164** (1994), 145–156.

[G3] Gallavotti, G.: *Reminiscences on science at I.H.E.S. A problem on homoclinic theory and a brief review*, Publications Scientifiques de l'IHES, Special Volume for the 40-th anniversary (1998), ???–???, [preprint in chao-dyn 9804043].

[GG] Gallavotti, G. , Gentile, G.: *Majorant series convergence for twistless KAM tori*, Ergodic Theory and Dynamical Systems **15** (1995), 1–69.

[GGM1] Gallavotti, G. , Gentile, G., Mastropietro, V.: *Field theory and KAM tori*, p. 1–9, Mathematical Physics Electronic Journal **1** (1995), http:// mpej.unige.ch.

[GGM2] Gallavotti, G. , Gentile, G., Mastropietro, V.: *Separatrix splitting for systems with three degrees of freedom*, to appear in Communications in Mathematical Physics, 1999, [preprint with title *Pendullum: separatrix splitting* in chao-dyn 9709004].

[GGM3] Gallavotti, G., Gentile, G., Mastropietro, V.: *Hamilton-Jacobi equation, heteroclinic chains and Arnol'd diffusion in three time scales systems*, submitted for publication [Preprint in chao-dyn #9801004].

[GGM4] Gallavotti, G., Gentile, G., Mastropietro, V.: *Melnikov's approximation dominance. Some examples*, to appear in Reviews in Mathematical Physics, 1999, [preprint in chao-dyn #9804043].

[Ge1] Gentile, G: *A proof of existence of whiskered tori with quasi flat homoclinic intersections in a class of almost integrable hamiltonian systems*, Forum Mathematicum **7** (1995), 709–753.

[Ge2] Gentile, G.: *Whiskered tori with prefixed frequencies and Lyapunov spectrum*, Dynamics and Stability of Systems **10** (1995), 269–308.

[Gel1] Gelfreich, V.: *Melnikov method and exponentially small splitting of separatrices*, Physica D **101** (1997), 227–248.

[Gel2] Gelfreich, V.: *Reference systems for splitting of separatrices*, Nonlinearity **10** (1997), 175–193.

[GM1] Gentile, G., Mastropietro, V.: *KAM theorem revisited*, Physica D **90** (1996), 225–234.

[GM2] Gentile, G., Mastropietro, V.: *Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in Classical mechanics. A review with some applications*, Reviews in Mathematical Physics **8** (1996), 393–444.

[N] Neishtadt, A.I.: *The separation of motions in systems with rapidly rotating phase*, Journal of Applied Mathematics and Mechanics **48** (1984), 133–139.

[HMS] Holmes, P., Marsden, J., Scheurle, J: *Exponentially small splittings of separatrices with applications to KAM theory and degenerate bifurcations*, Contemporary Mathematics **81** (1989), 213–244.

[La] Lazutkin, V.F.,, *Splitting of separatrices for the Chirikov's standard map*, mp_arc@ math.utexas.edu, #98–421, (annotated english translation of a 1984 paper).

[P] Poincaré, H.: *Les méthode nouvelles de la Mécanique Céleste*, Vol. III, Gauthier–Villard, Paris, 1899.

[Pö] Pöschel, J.: *Invariant manifolds of complex analytic mappings*, Les Houches, XLIII (1984), Vol. II, p. 949–964, Eds. K. Osterwalder & R. Stora, North Holland, 1986.

Internet: Authors' preprints at: `http://ipparco.roma1.infn.it`
e-mail: users: giovanni, gentile, vieri, address: `@ipparco.roma1.infn.it`