Fractional Lindstedt series

G. Gallavotti[1], **G. Gentile**[2] and **A. Giuliani**[3]

[1] Dipartimento di Fisica, Università di Roma Uno, I-00185

[2] Dipartimento di Matematica, Università di Roma Tre, Roma, I-00146

[3] Dipartimento di Matematica, Università di Roma Due, Roma, I-00133

ABSTRACT. *The parametric equations of the surfaces on which highly resonant quasi-periodic motions develop (lower-dimensional tori) cannot be analytically continued, in general, in the perturbation parameter ε, i.e. they are not analytic functions of ε. However rather generally quasi-periodic motions whose frequencies satisfy only one rational relation ("resonances of order 1") admit formal perturbation expansions in terms of a fractional power of ε depending on the degeneration of the resonance. We find conditions for this to happen, and in such a case we prove that the formal expansion is convergent after suitable resummation.*

1. Introduction

Resonances play an important role in the theory of dynamical systems. A possible application is provided by problems of celestial mechanics, such as the phenomenon of resonance locking between rotation and orbital periods of the satellites [GP]. In fact, the presence of friction can select resonant motions which remain stable when friction (on astronomical time scales) becomes negligible. In such a case maximal KAM tori can be really observed only approximately and on very short time scale, whereas, on very large time scales one expects that only periodic motions survive. On intermediate time scale one can imagine that quasi-periodic motions, involving a number of frequencies less than the total number of degrees of freedom (and decreasing with time), describe most of the observed dynamics. This makes interesting and important to study quasi-periodic motions occurring on lower-dimensional tori for nearly-integrable Hamiltonian systems. These quasi-periodic motions are characterized by frequencies satisfying s rational relations, with $r = N - s$ ranging between 1 (periodic motions) and the number N of degrees of freedom (KAM tori). The number s equals the number of normal frequencies appearing in the perturbed motions, whereas r is the number of independent components of the rotation vector.

The analysis of such motions is simpler under some (generic) non-degeneracy assumptions on the perturbations. On the contrary the situation becomes immediately very complicated if no restriction at all is made on the perturbation. Situations of this kind arise also in similar contexts: we can mention the conservation of KAM tori under perturbations for systems of N harmonic oscillators, proved for $N = 2$ but conjectured to hold in general [H], [R2], and the stability of Hill's equation under quasi-periodic perturbations [GBC]. In the case of lower-dimensional tori, the non-degeneracy assumption is that the normal frequencies become different from zero when the perturbation is switched on. If such an assumption is removed, then only partial results hold, and only in the case $s = 1$, that is only in the case of one normal frequency [Ch1], [Ch2].

Starting from the work of Eliasson [E2], a new approach to KAM theory of quasi–integrable Hamiltonian systems arose, based on the analysis of cancellations in the "Lindstedt series" for the functions mapping unperturbed motions (uniform rotations, in suitable coordinates) into corresponding perturbed ones.

A convenient way to exploit cancellations to resolve apparent divergences in the Lindstedt series is through methods inspired by quantum field theory, consisting in graphical expansions, summation of classes of diverging subdiagrams, iterative study of the flow of the effective constants and the possible introduction of counterterms. With these techniques and ideas, a number of known results have been reproduced and new ones have been obtained; see [G1], [Ge2], [Ge3] and [GBG] for some reviews.

In this paper we follow the latter approach to investigate the conservation of $(N-1)$-dimensional tori for systems with N degrees of freedom. More precisely we consider N degrees of freedom systems described by analytic Hamiltonians of the form

$$H(I,\varphi) = H_0(I) + \varepsilon f(I,\varphi), \tag{1.1}$$

with $(I,\varphi) \in \mathcal{D}\times\mathbb{T}^N$, where $\mathbb{T} = \mathbb{R}\backslash 2\pi\mathbb{Z}$ is the standard torus, $\mathcal{D}$ is an open subset of $\mathbb{R}^N$, ε is a real parameter, and the free Hamiltonian $H_0(I)$ is assumed to be uniformly convex: $\partial_I^2 H_0(I) \ge C > 0$ for all $I \in \mathcal{D}$.

Definition 1 (Simple resonance). *A* simple Diophantine resonance *for the unperturbed system is a motion taking place on the torus* $\{I_0\}\times\mathbb{T}^N$ *with* $\omega_0 \stackrel{def}{=} \partial_I H_0(I_0)$ *satisfying a rational relation* $\omega_0\cdot\nu_0 = 0$ *for some* $\nu_0 \in \mathbb{Z}^N$ *and* $|\omega_0\cdot\nu| > C|\nu|^{-\tau}$ *for suitable* $C,\tau > 0$ *and for all* $0 \ne \nu \in \mathbb{Z}^N$ *not parallel to* ν_0.

Thus, if $\varepsilon = 0$, the motions with rotation velocity ω_0 will foliate the torus $\{I_0\}\times\mathbb{T}^N$ into a one parameter family of invariant tori of dimension $N-1$.

For $\varepsilon \ne 0$ invariant tori with dimension $N-1$ run by quasi-periodic motions with spectrum ω_0 will, in general, only continue to exist "close" to some of the unperturbed tori $\{I_0\}\times\mathbb{T}^N$. The problem is simpler under non-degeneracy assumptions on the average $\langle f\rangle$ of the perturbing function f on the torus $\{I_0\}\times\mathbb{T}^N$; it has been studied in [GG1] and [GG2] with techniques employed, under different assumptions on the perturbation, in this paper. Define the average of f on $\{I_0\}\times\mathbb{T}^N$ as

$$\langle f(\varphi)\rangle \stackrel{def}{=} \lim_{T\to\infty}\frac{1}{T}\int_0^T f(\varphi+\omega_0 t)dt. \tag{1.2}$$

It depends nontrivially on φ in the sense that in general it is a non-constant periodic function of φ. This is more easily visualized in coordinates *adapted* to the resonance: they are defined by a linear canonical transformation $(I,\varphi) \stackrel{def}{=} \big(I_0 + S^{-1}(\mathbf{A},B), S^T(\boldsymbol{\alpha},\beta)\big)$, with S a non-singular integer components $N\times N$ matrix with determinant $\det S = 1$ such that $\omega_0 \equiv S^T(\boldsymbol{\omega},0)$ with $\boldsymbol{\omega}\in\mathbb{R}^{N-1}$, $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}| > C_0|\boldsymbol{\nu}|^{-\tau_0}$ for *all* $\mathbf{0} \ne \boldsymbol{\nu} \in \mathbb{Z}^{N-1}$ and $\mathbf{A}\in\mathbb{R}^{N-1}, B\in\mathbb{R}, \boldsymbol{\alpha}\in\mathbb{T}^{N-1}, \beta\in\mathbb{T}$.

In the coordinates $(\mathbf{A}, B, \boldsymbol{\alpha}, \beta)$ the Hamiltonian becomes an analytic function of $(\mathbf{A}, B)$, in the domain obtained from $\mathcal{D}$ under the transformation S, and $(\boldsymbol{\alpha}, \beta) \in \mathbb{T}^{N-1} \times \mathbb{T}$. It will be of the form $H(\mathbf{A}, B, \boldsymbol{\alpha}, \beta) = H_0'(\mathbf{A}, B) + \varepsilon f(\boldsymbol{\alpha}, \beta)$, where $H_0'(\mathbf{A}, B) = \boldsymbol{\omega} \cdot \mathbf{A} + H_0(\mathbf{A}, B)$, with H_0 vanishing to second order at $\mathbf{A} = \mathbf{0}, B = 0$ and uniformly convex in the domain where it is defined. Of course the functions H, H_0, f have a different meaning with respect to those appearing in (1.1), but we prefer to use the same notation for simplicity. For the same reason we still shall use the notation (I, φ) to denote the new action-angle variables, by setting $I = (\mathbf{A}, B)$ and $\varphi = (\boldsymbol{\alpha}, \beta)$.

So in the new coordinates the Hamiltonian H, the rotation vector $\boldsymbol{\omega} \in \mathbb{R}^{N-1}$ and the average $f_{\mathbf{0}}(\mathbf{A}, B, \beta)$ of the perturbing energy can be supposed to be such that

$$
\begin{aligned}
& H = \boldsymbol{\omega} \cdot \mathbf{A} + H_0(\mathbf{A}, B) + \varepsilon f(\mathbf{A}, B, \boldsymbol{\alpha}, \beta), \qquad |\boldsymbol{\omega} \cdot \boldsymbol{\nu}| \geq \frac{C_0}{|\boldsymbol{\nu}|^{\tau_0}} \qquad \forall\, \mathbf{0} \neq \boldsymbol{\nu} \in \mathbb{Z}^{N-1}, \\
& f_{\mathbf{0}}(\mathbf{A}, B, \beta) \stackrel{def}{=} \int \frac{d\boldsymbol{\alpha}}{(2\pi)^{N-1}} f(\mathbf{A}, B, \boldsymbol{\alpha}, \beta), \quad \text{if} \quad f(\mathbf{A}, B, \boldsymbol{\alpha}, \beta) = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^{N-1}} e^{i\boldsymbol{\nu} \cdot \boldsymbol{\alpha}} f_{\boldsymbol{\nu}}(\mathbf{A}, B, \beta),
\end{aligned}
\tag{1.3}
$$

near the unperturbed resonance and *without any loss of generality.* The corresponding equations of motion for $X(t) \equiv (\mathbf{A}(t), B(t), \boldsymbol{\alpha}(t), \beta(t))$ are

$$
\begin{aligned}
\dot{\mathbf{A}} &= -\,\varepsilon \partial_{\boldsymbol{\alpha}} f, \\
\dot{B} &= -\,\varepsilon \partial_{\beta} f, \\
\dot{\boldsymbol{\alpha}} &= \boldsymbol{\omega} + \partial_{\mathbf{A}} H_0 + \varepsilon \partial_{\mathbf{A}} f, \\
\dot{\beta} &= \partial_B H_0 + \varepsilon \partial_B f,
\end{aligned}
\qquad \text{or} \qquad \dot{X} = E \partial_X H(X), \qquad \text{with} \qquad E = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},
\tag{1.4}
$$

where E is the standard $2N \times 2N$ symplectic matrix. We study the existence of motions which can be described by a constant β_0 and by functions $\mathbf{A}(\boldsymbol{\psi}), B(\boldsymbol{\psi}), \mathbf{a}(\boldsymbol{\psi}), b(\boldsymbol{\psi})$ of $\boldsymbol{\psi} \in \mathbb{T}^{N-1}$, which tend to 0 as $\varepsilon \to 0$, such that, by posing

$$
\begin{aligned}
\mathbf{A}(t) &= \mathbf{A}(\boldsymbol{\psi} + \boldsymbol{\omega}\, t), & B(t) &= B(\boldsymbol{\psi} + \boldsymbol{\omega}\, t), \\
\boldsymbol{\alpha}(t) &= \boldsymbol{\psi} + \boldsymbol{\omega}\, t + \mathbf{a}(\boldsymbol{\psi} + \boldsymbol{\omega}\, t), & \beta(t) &= \beta_0 + \mathbf{b}(\boldsymbol{\psi} + \boldsymbol{\omega}\, t),
\end{aligned}
\tag{1.5}
$$

one obtains, for all $\boldsymbol{\psi} \in \mathbb{T}^{N-1}$, solutions to the equations of motion (1.4). For brevity we shall sometimes write $X(t)$, instead of $X(\boldsymbol{\psi} + \boldsymbol{\omega} t)$, to indicate solutions of (1.4) of the form (1.5).

Note that, for $\varepsilon = 0$, the motions (1.5) reduce to

$$
\mathbf{A}^{(0)}(t) = \mathbf{0}, \qquad B^{(0)}(t) = 0, \qquad \boldsymbol{\alpha}^{(0)}(t) = \boldsymbol{\psi} + \boldsymbol{\omega} t, \qquad \beta^{(0)}(t) = \beta_0,
\tag{1.6}
$$

where $\boldsymbol{\psi} \in \mathbb{T}^{N-1}$ are arbitrary. The motions $X^{(0)}(t) = (\mathbf{A}^{(0)}(t), B^{(0)}(t), \boldsymbol{\alpha}^{(0)}(t), \beta^{(0)}(t))$ represent the unperturbed resonant motions filling a one parameter family of $N-1$ dimensional invariant tori (parameterized by β_0).

Given any function $G(\boldsymbol{\psi}, \varepsilon)$ we shall denote by $G_{\boldsymbol{\nu}}$ the $\boldsymbol{\nu}$-th Fourier component of its Fourier expansion in $\boldsymbol{\psi}$ and we shall call $[G]^k$ the k-th order term obtained by expanding in ε the function G. Furthermore we shall denote by $[G]^k_{\boldsymbol{\nu}}$ the $\boldsymbol{\nu}$-th Fourier component of the Fourier expansion of $[G]^k$ and by $G_{\neq \mathbf{0}}$ the function $G - G_{\mathbf{0}}$.

A formal solution of (1.4), (1.5), as a power series in ε, $X(\boldsymbol{\psi}) = X^{(0)}(\boldsymbol{\psi}) + \varepsilon X^{(1)}(\boldsymbol{\psi}) + \ldots$, is well known to exist if $\beta_0 \in \mathbb{T}$ is a stationarity point for the average $f_{\mathbf{0}}(\mathbf{0}, 0, \beta)$ (i.e. a point with such that $\partial_\beta f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0) = 0$) which is not degenerate (i.e. $\partial^2_\beta f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0) \neq 0$); cf. for instance [P], [JLZ] and [GG1]. *Here we consider explicitly the case in which the non-degeneracy condition on* $\partial^2_\beta f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0)$ *can fail to hold. This is a case in which in general no formal solution in powers of* ε *can be constructed.*

If β_0 is such that $\partial_\beta f_{\mathbf{0}}(\beta_0) = 0$, then there exists a function X' such that $X^{(0)} + \varepsilon X'$ solves (1.4) up to terms of order ε^2 (excluded). Namely $X'(\boldsymbol{\psi}) = (\mathbf{A}'(\boldsymbol{\psi}), B'(\boldsymbol{\psi}), \mathbf{a}'(\boldsymbol{\psi}), b'(\boldsymbol{\psi}))$ is obtained by applying the operator $(\boldsymbol{\omega} \cdot \boldsymbol{\partial}_{\boldsymbol{\psi}})^{-1}$ to the vector

$$\Big(- \partial_\varphi f(\mathbf{0}, 0, \boldsymbol{\psi}, \beta_0)\ ,\ \partial_I f(\mathbf{0}, 0, \boldsymbol{\psi}, \beta_0) + \partial^2_I H_0(\mathbf{0}, 0)\big(\mathbf{A}'(\boldsymbol{\psi}), B'(\boldsymbol{\psi})\big)\Big), \tag{1.7}$$

where $\partial_\varphi, \partial_I$ denote respectively the derivatives with respect to the angle and action variables. This means that one first determines $\mathbf{A}'(\boldsymbol{\psi})$ and $B'(\boldsymbol{\psi})$ by solving $(\boldsymbol{\omega} \cdot \partial_{\boldsymbol{\psi}})\big(\mathbf{A}'(\boldsymbol{\psi}), B'(\boldsymbol{\psi})\big) = -\partial_\varphi f(\mathbf{0}, 0, \boldsymbol{\psi}, \beta_0)$; the (otherwise arbitrary) averages $\mathbf{A}'_{\mathbf{0}}, B'_{\mathbf{0}}$ are fixed by requiring that the operator $(\boldsymbol{\omega} \cdot \partial)^{-1}$ can be applied to the vector formed by the last N components of (1.7), i.e. $\partial_I f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0) + \partial^2_I H_0(\mathbf{0}, 0)\big(\mathbf{A}'_{\mathbf{0}}, B'_{\mathbf{0}}\big) = \mathbf{0}$. In this way also $\mathbf{a}'(\boldsymbol{\psi})$ and $b'(\boldsymbol{\psi})$ can be obtained. The averages of the angle variables will be chosen $\mathbf{a}'_{\mathbf{0}} = \mathbf{0}$ while we leave $b'_{\mathbf{0}}$ as a free parameter (to be suitably fixed at higher orders to make the equations (1.4) formally solvable).

There are, however other solutions which are correct up to order ε^2 (excluded). If $\partial^j_\beta f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0) = 0$ for all $j \leq k_0$ and $\partial^{k_0+1}_\beta f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0) \neq 0$, one can imagine to add to the free parameter $b'_{\mathbf{0}}$ any polynomial in powers of $\eta = |\varepsilon|^{\frac{1}{k_0}}$ of degree $< k_0$. It can be checked (and it will be explicitly shown in next sections) that, for any choice of this polynomial, a solution of the equation of motions correct up to terms of order ε^2 (excluded) exists. This is, ultimately, the reason why a consistent expansions in powers of ε cannot in general be continued beyond first order: the consistency condition of the equations of motion necessary to improve (to second order included) the solution fixes one coefficient of such (a priori arbitrary) polynomials in η and, in general, this really forces the expansion to be an expansion in powers of η different from $\eta^{k_0} = \varepsilon$.

The analysis below shows that an expansion in η is actually possible at all orders if the following assumptions are satisfied.

Assumptions.

(a) *The constant* $a \stackrel{def}{=} [\partial_\beta f(X^{(0)} + \varepsilon\, X'_0)]^1_{\mathbf{0}}$ *is* $a \neq 0$.

(b) *The matrix* $\partial^2_I H_0(\mathbf{0}, 0)$ *is positive definite.*

(c) *There is* $k_0 > 0$ *such that* $\partial^j_\beta f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0) = 0$ *for all* $j \leq k_0$ *and* $c \stackrel{def}{=} \frac{1}{k_0!}\partial^{k_0+1}_\beta f_{\mathbf{0}}(\mathbf{0}, 0, \beta_0) \neq 0$.

The constant a appearing in assumption (a), written explicitly, is

$$\begin{aligned} a \stackrel{def}{=} \Big[&\partial_{\beta,\varphi} f(X^{(0)}(\boldsymbol{\psi})) \cdot (\boldsymbol{\omega} \cdot \partial_{\boldsymbol{\psi}})^{-1}\Big(\partial_I f(X^{(0)}(\boldsymbol{\psi})) + \partial^2_I H_0(\mathbf{0}, 0)\big(\mathbf{A}'(\boldsymbol{\psi}), B'(\boldsymbol{\psi})\big)\Big)_{\neq \mathbf{0}} - \\ &- \partial_{\beta,I} f(X^{(0)}(\boldsymbol{\psi})) \cdot (\boldsymbol{\omega} \cdot \partial_{\boldsymbol{\psi}})^{-1}\partial_\varphi f(X^{(0)}(\boldsymbol{\psi}))_{\neq \mathbf{0}}\Big]_{\mathbf{0}}. \end{aligned} \tag{1.8}$$

For instance if f depends only on the angle variables φ and $\partial^2_I H(\mathbf{0}, 0) = \mathbb{1}$ the constant a is

$$a = \frac{1}{2}\partial_\beta \sum_{\boldsymbol{\nu}} \frac{|\boldsymbol{\nu}|^2 |f_{\boldsymbol{\nu}}(\beta)|^2 + |\partial_\beta f_{\boldsymbol{\nu}}(\beta)|^2}{(\boldsymbol{\omega} \cdot \boldsymbol{\nu})^2}\Bigg|_{\beta = \beta_0}. \tag{1.9}$$

The number k_0 appearing in assumption (c) is a measure of the "degeneration" of the resonance, while the *order* of the resonance is the number of rational relations between the unperturbed frequencies, which is 1 in our case as the unperturbed motions have $N-1$ independent frequencies. The case $k_0 = 1$ was considered in [GG2] and we will not consider it again here. We shall focus on the case $k_0 \geq 2$, in which case a formal power solution in ε to (1.4) does not exist. In fact, one checks that, as a consequence of assumptions (a) and (c), the average over $\boldsymbol{\psi}$ of the r.h.s. of the second equation in (1.4) is different from 0 at order ε^2, for any possible choice of the free parameter $b'_{\mathbf{0}}$ introduced after (1.7).

However, under the assumptions above, it is possible to find a formal solution to (1.4) such that the average of b is $b_{\mathbf{0}} = O(|\varepsilon|^{1/k_0})$: the average of b will be fixed in terms of the constants a and c in such a way that the average over $\boldsymbol{\psi}$ of the r.h.s. of the second equation in (1.4) is 0 at all orders in $|\varepsilon|^{1/k_0}$.

The necessity of a fractional powers expansion can be seen by a heuristic argument sketched in Appendix A1: the argument also suggests, as a conjecture, the forthcoming Theorem 1 and motivates the assumptions (a) *to* (c).

In Appendix A1 it is in fact shown that, in the simple case $H_0(\mathbf{A}, B) = \frac{1}{2}(\mathbf{A}^2 + B^2)$ and f depending only on the angles, a canonical transformation (explicitly constructed in Appendix A1), defined in a neighborhood of $\{\mathbf{0}\} \times \{0\} \times \mathbb{T}^{N-1} \times \{\beta_0\}$, maps $\boldsymbol{\omega} \cdot \mathbf{A} + H_0(\mathbf{A}, B) + \varepsilon f(\boldsymbol{\alpha}, \beta)$ into

$$\begin{aligned} &\boldsymbol{\omega} \cdot \mathbf{A} + H_0(\mathbf{A}, B) + \frac{\varepsilon\, c}{k_0 + 1}(\beta - \beta_0)^{k_0+1} + \varepsilon^2\, a\,(\beta - \beta_0) + \\ &\qquad + O(\varepsilon(\beta - \beta_0)^{k_0+2}) + O(\varepsilon^2(\beta - \beta_0)^2) + O(\varepsilon I^2) + O(\varepsilon^3). \end{aligned} \tag{1.10}$$

As discussed in Appendix A1, the Hamiltonian equations corresponding to (1.10) can be consistently solved to order ε^2. In particular, for some choices of the signs of ε, a, c, the angle β admits an approximate *quadratic* equilibrium point $O(|\varepsilon|^{1/k_0})$, whose stability depends again on the relative signs of ε, a, c. This second order computation suggests the conjecture that the unperturbed motion $X^{(0)}(t)$ can be continued at $\varepsilon \neq 0$, provided the average of β is chosen $O(|\varepsilon|^{1/k_0})$. The perturbed motion, if existing beyond second order, will take place on a torus (a small perturbation of the free one) which we shall call elliptic or hyperbolic, depending on the stability of the behavior of the linearization of the motion of β in the vicinity of its equilibrium point: if the corresponding pair of nonzero Lyapunov exponents is imaginary then the torus will be said *elliptic*, if it is real it will be said *hyperbolic*.

In the following a *sparse Cantor set dense at* 0 will mean a set $\mathcal{E}$ contained in an interval $I = [-\varepsilon_0, 0]$ or $I = [0, \varepsilon_0]$ with an open dense complement in I and with 0 as a density point in the sense of Lebesgue integration. In particular $\mathcal{E}$ will have positive measure.

Theorem 1. *Consider the system described by the Hamiltonian (1.3), under the assumptions* (a), (b) *and* (c). *There exists* $\varepsilon_0 > 0$ *such that for* $|\varepsilon| < \varepsilon_0$ *the following holds.*

(i) If k_0 *is odd and* $c\,\varepsilon < 0$ *there is at least one hyperbolic invariant torus of dimension* $N-1$ *with rotation vector* $\boldsymbol{\omega}$. *If* $c\,\varepsilon > 0$ *there is a sparse Cantor set* $\mathcal{E} \subset [-\varepsilon_0, 0]$ *with the property that for* $\varepsilon \in \mathcal{E}$ *there is at least one elliptic torus of dimension* $(N-1)$ *with rotation vector* $\boldsymbol{\omega}$.

(ii) If k_0 *is even and* $\varepsilon\, c$ *has the same sign of* $-a$, *there is at least one hyperbolic invariant torus of dimension* $N-1$ *with rotation vector* $\boldsymbol{\omega}$. *Moreover, there is a sparse Cantor set* $\mathcal{E}$ *dense at* 0

such that, if $\varepsilon \in \mathcal{E}$ and $\varepsilon\, c$ has the same sign of $-a$, there is at least one elliptic invariant torus of dimension $N-1$ with rotation vector $\boldsymbol{\omega}$.

Remarks. (1) If assumption (c) is violated, since f is analytic in β, then $f_{\mathbf{0}}(\mathbf{0},0,\beta) \equiv 0$ as a function of β. In this case we can perform a canonical transformation removing the perturbation at order ε and casting the Hamiltonian into the form $H' = H_0' + \varepsilon^2 f'$, for some new analytic functions H_0' and f'. If f' satisfies the assumptions above we can apply Theorem 1 to H'. If H' satisfies assumption (c) and violates assumption (a) we cannot say much. If H' violates assumption (c) we can again remove the perturbation at lowest order through a new canonical transformation and cast the Hamiltonian into the new form $H'' = H_0'' + \varepsilon^4 f''$ and hope to be able to apply Theorem 1 to H''. And so on.
(2) Assumption (a) is essential and, if it does not hold, our expansion may fail to be convergent. In fact, in some cases (k_0 even), it is easy to show that if assumption (a) fails there cannot be perturbed motions of the form (1.5), see Appendix A2 for an example. If k_0 is odd, [Ch1] proved existence of hyperbolic tori even if assumption (a) fails; in these cases we expect that a new perturbation parameter must be identified. The heuristic analysis in Appendix A1 concretely suggests plausible results to be expected if $a = 0$ (under alternative assumptions).
(3) We expect that assumption (b) is not essential and that it could be weakened into the request that $\partial_I^2 H_0(\mathbf{0},0)$ is non-degenerate and that $\partial_{BB}^2 H_0(\mathbf{0},0) \neq 0$. Certainly the convexity assumption simplifies some of the estimates (see Appendix A6) and we did not attempt to eliminate it.
(4) The only known result on the problems considered here is in [Ch1], where conservation of $(N-1)$-dimensional "hyperbolic tori" is proved under weaker assumptions, although we study conservation of both hyperbolic and elliptic $(N-1)$–dimensional tori under assumptions (a) to (c) above.
(5) In principle one could proceed in a different way rather than following our approach. One could first perform the canonical transformation described in Appendix A1 and leading to the Hamiltonian (1.10), then study the system so obtained with other techniques, such as those in [E1], [Po], [R1] or [JLZ]. To this aim one should use that in the new coordinates the unperturbed Hamiltonian contains terms of order η^{2k_0-1}, while the perturbation is of order η^{2k_0}, hence has a further η.
(6) The analysis via Lindstedt series and summations of classes of diagrams is the main aspect of this work. Also the analyticity properties in η at $\boldsymbol{\omega}$ fixed are, to our knowledge, new.
(7) Technically the present work is strongly inspired by [GG1], [Ge1] and [GG2]. The proofs that can be taken literally from [GG2] will not be repeated here: hence familiarity with the latter reference is essential.
(8) The resummed series has manifest holomorphy properties which show that the set $\mathcal{E}$ is in the boundary of a complex domain where the functions $X(\boldsymbol{\psi})$ are analytic in ε: we do not discuss the details (see [GG1],[GG2]).

The paper is organized as follows. In Sections 2 and 3 we prove formal solvability of the equation of motion in power series in $|\varepsilon|^{1/k_0}$ and we describe a graphical representation of the terms appearing in the formal power series. Such a representation involves labeled rooted trees, and is very similar to diagrammatic representations through Feynman diagrams arising in quantum field theory. In Sections 4 to 6 we describe an iterative resummation scheme which eliminates some classes of divergent subdiagrams and iteratively changes the power expansion, and we prove convergence of the resummed series. Some details of the proofs are deferred to the Appendices.

2. Lindstedt series

We define $\varepsilon = \sigma\eta^{k_0}$, with $\sigma \in \{\pm 1\}$, $\eta > 0$, and look for a family of formal solutions of the equations of motion in powers of η parameterized by $\boldsymbol{\psi} \in \mathbb{T}^{N-1}$, in the special form $X(t) = (\mathbf{A}(t), B(t), \boldsymbol{\alpha}(t), \beta(t))$:

$$
\begin{aligned}
\mathbf{A}(t) &= \sum_{k=k_0}^{\infty} \eta^k \sum_{\boldsymbol{\nu}\in\mathbb{Z}^{N-1}} e^{i\boldsymbol{\nu}\cdot(\boldsymbol{\psi}+\boldsymbol{\omega}t)} \mathbf{A}_{\boldsymbol{\nu}}^{(k)}, \\
B(t) &= \sum_{k=k_0}^{\infty} \eta^k \sum_{\boldsymbol{\nu}\in\mathbb{Z}^{N-1}} e^{i\boldsymbol{\nu}\cdot(\boldsymbol{\psi}+\boldsymbol{\omega}t)} B_{\boldsymbol{\nu}}^{(k)}, \\
\boldsymbol{\alpha}(t) &= \boldsymbol{\psi} + \boldsymbol{\omega}t + \mathbf{a}(t), \qquad \mathbf{a}(t) = \sum_{k=k_0}^{\infty} \eta^k \sum_{\boldsymbol{\nu}\neq\mathbf{0}} e^{i\boldsymbol{\nu}\cdot(\boldsymbol{\psi}+\boldsymbol{\omega}t)} \mathbf{a}_{\boldsymbol{\nu}}^{(k)}, \\
\beta(t) &= \beta_0 + b(t), \qquad b(t) = \sum_{k=k_0}^{\infty} \eta^k \sum_{\boldsymbol{\nu}\neq\mathbf{0}} e^{i\boldsymbol{\nu}\cdot(\boldsymbol{\psi}+\boldsymbol{\omega}t)} b_{\boldsymbol{\nu}}^{(k)} + \sum_{k=1}^{\infty} \eta^k b_{\mathbf{0}}^{(k)},
\end{aligned}
\tag{2.1}
$$

where all the involved functions have also a dependence on β_0 which has not been made explicit. The formal series (2.1) will be called *Lindstedt series* as it extends the corresponding notions already used in the theory of quasi-periodic motions on maximal tori.

In the following the average $b_{\mathbf{0}}^{(k)}$ *will be abbreviated as* β_k.

Remarks. (1) The functions $\mathbf{A}, B, \mathbf{a}$ and $b - b_{\mathbf{0}}$ have been chosen as power series in η starting with the k_0–th order as the first non trivial order; $\mathbf{a}$ has been chosen with zero average (this just corresponds to a redefinition of the origin of $\mathbb{T}^{N-1}$), while the average of b has been chosen as a series in η starting with the first order; the coefficients $\beta_k \stackrel{def}{=} b_{\mathbf{0}}^{(k)}$ will be chosen in such a way that the Lindstedt series admits a formal solution.

(2) With the choices in (2.1), equations (1.4) are identically solved for any order $k < k_0$ in η. The parameters β_k, $k < k_0$, are left as free parameters, to be explicitly chosen below.

To write the generic k–th order of (1.4) we introduce the following definitions: given any function $\boldsymbol{\psi} \to F(X(\boldsymbol{\psi}))$, let $[F]^{(k)}$ be the k–th order in the Taylor expansion of F in η and let $[F]_{\boldsymbol{\nu}}^{(k)}$ be the $\boldsymbol{\nu}$-th Fourier component of the Fourier expansion of $[F]^{(k)}$. Note that this notation concerns expansions in η and is different from the one introduced after (1.6) which dealt with expansions in ε: with the new notation the quantity denoted $a = [\partial_\beta f(X^{(0)}) + \varepsilon X')]_{\mathbf{0}}^{1}$ before (1.8) has to be written as $a = \sigma\,[\partial_\beta f(X^{(0)} + \sigma\eta^{k_0} X')]_{\mathbf{0}}^{(k_0)}$ as the superscript ^{k} denotes k-th order in ε while $^{(k)}$

denotes k-th order in η. Then, if $\boldsymbol{\nu}\neq\mathbf{0}$ and $k\geq k_0$, the equations (1.4) become

$$\begin{aligned}(i\boldsymbol{\omega}\cdot\boldsymbol{\nu})\,\mathbf{A}^{(k)}_{\boldsymbol{\nu}} &= -\sigma[\partial_{\boldsymbol{\alpha}} f]^{(k-k_0)}_{\boldsymbol{\nu}},\\ (i\boldsymbol{\omega}\cdot\boldsymbol{\nu})\,B^{(k)}_{\boldsymbol{\nu}} &= -\sigma[\partial_{\beta} f]^{(k-k_0)}_{\boldsymbol{\nu}},\\ (i\boldsymbol{\omega}\cdot\boldsymbol{\nu})\,\mathbf{a}^{(k)}_{\boldsymbol{\nu}} &= [\partial_{\mathbf{A}} H_0]^{(k)}_{\boldsymbol{\nu}}+\sigma[\partial_{\mathbf{A}} f]^{(k-k_0)}_{\boldsymbol{\nu}},\\ (i\boldsymbol{\omega}\cdot\boldsymbol{\nu})\,b^{(k)}_{\boldsymbol{\nu}} &= [\partial_{B} H_0]^{(k)}_{\boldsymbol{\nu}}+\sigma[\partial_{B} f]^{(k-k_0)}_{\boldsymbol{\nu}},\end{aligned}\tag{2.2}$$

and, since $\boldsymbol{\omega}$ satisfies the Diophantine property, they can be solved *provided*

$$\begin{aligned}\mathbf{0} &=[\partial_{\boldsymbol{\alpha}} f]^{(k-k_0)}_{\mathbf{0}},\\ 0 &=[\partial_{\beta} f]^{(k-k_0)}_{\mathbf{0}},\\ \mathbf{0} &=[\partial_{\mathbf{A}} H_0]^{(k)}_{\mathbf{0}}+\sigma[\partial_{\mathbf{A}} f]^{(k-k_0)}_{\mathbf{0}},\\ 0 &=[\partial_{B} H_0]^{(k)}_{\mathbf{0}}+\sigma[\partial_{B} f]^{(k-k_0)}_{\mathbf{0}}.\end{aligned}\tag{2.3}$$

If assumptions (a), (b) and (c) in Section 1 are satisfied, such a formal solution can be shown to exist provided the formal series for the average $b_{\mathbf{0}}$ is suitably chosen.

Lemma 1. *Under the assumptions* (a), (b) *and* (c) *if* k_0 *is odd, a formal solution of (1.4) in the form (2.1) always exists; if* k_0 *is even, a formal solution of (1.4) in the form (2.1) exists if* $\sigma ac<0$. *When a formal solution exists,* $\mathbf{A}^{(k)},B^{(k)},\mathbf{a}^{(k)},b^{(k)}$ *are uniquely fixed in the case* k_0 *odd. If* k_0 *is even, there are two possible such sequences corresponding to the choices* $\beta_1\equiv b^{(1)}_{\mathbf{0}}=\pm(-\sigma a/c)^{1/k_0}$.

Proof. Let $X^{(0)}(\boldsymbol{\psi})=(\mathbf{0},0,\boldsymbol{\psi},\beta_0)$, $\varepsilon=\sigma\,\eta^{k_0}$ with $\sigma=\mathrm{sign}\,(\varepsilon)$ and $\eta>0$, and look for a formal solution $t\to(\mathbf{A}(t),B(t),\boldsymbol{\alpha}(t),\beta(t))$ obtained by setting $\boldsymbol{\psi}=\boldsymbol{\psi}+\boldsymbol{\omega}t$ in

$$X(\boldsymbol{\psi})=X^{(0)}(\boldsymbol{\psi})+\sum_{k=k_0}^{\infty}\eta^k X^{(k)}(\boldsymbol{\psi})+\sum_{k=1}^{\infty}\eta^k\xi^{(k)},\tag{2.4}$$

where $\xi^{(k)}=(\mathbf{0},0,\mathbf{0},\beta_k)$. Set $X^{(h)}\equiv 0$ for $0<h<k_0$ and $\xi^{(0)}=0$.

Suppose inductively that $X^{(h)}$ has been determined for $h=0,1,\ldots,k-1$, for $k\geq 1$. Consider a generic polynomial $Y(\boldsymbol{\psi})\stackrel{def}{=}\sum_{h=0}^{k-1}\eta^h Y^{(h)}(\boldsymbol{\psi})$ and suppose, inductively, that if $Y^{(h)}=X^{(h)}+\xi^{(h)}$ the function $Y(\boldsymbol{\omega}t)$ solves the equation of motion $\dot X=E\partial_X H(X)$ up to order $k-1$. This is true for $k=1$. Let $\Delta\stackrel{def}{=}(\boldsymbol{\omega}\cdot\partial_{\boldsymbol{\psi}})$ and remark that the identities $\int\sum_{j=1}^{2N}\partial_{\psi_i}Y_j\cdot\partial_{Y_j}H(Y)\,d\boldsymbol{\psi}=0$ and $\int\sum_{j=1}^{2N}\partial_{\psi_i}Y_j\cdot(E\Delta Y)_j\,d\boldsymbol{\psi}=0$ are identities *for all periodic functions* $Y(\boldsymbol{\psi})$.

Then $0\equiv\int\sum_j\partial_{\psi_i}Y_j\cdot((E\,\Delta Y)_j+\partial_{Y_j}H(Y))\,d\boldsymbol{\psi}$ and $\dot Y=E\partial_Y H(Y)+O(\eta^k)$ imply (note that Y has degree $<k$ in η) $\int\partial_{\psi_i}Y^{(0)}_j\cdot\,[\partial_{Y_j}H(Y)]^{(k)}\;d\boldsymbol{\psi}=\mathbf{0}$, i.e. $\int[\partial_{\boldsymbol{\alpha}}H(Y)]^{(k)}\,d\boldsymbol{\psi}=\mathbf{0}$ or

$$\big[\partial_{\boldsymbol{\alpha}} f\big(\sum_{h=0}^{k-1}\eta^h(X^{(h)}+\xi^{(h)})\big)\big]^{(k-k_0)}_{\mathbf{0}}=\mathbf{0},\tag{2.5}$$

which is the first of the compatibility conditions (2.3). This leaves $\mathbf{a}^{(k)}_{\mathbf{0}}$ as an arbitrary parameter: we can set $\mathbf{a}^{(k)}_{\mathbf{0}}=\mathbf{0}$. We also remark that, if $I\equiv(\mathbf{A},B)$, the last two of the four equations in (2.3) have the form

$$(E\partial)_{\varphi}\partial_I H_0(\mathbf{0},0)\,I^{(k)}_{\mathbf{0}}+F^{(k)}(\{X^{(h)},\xi^{(h)}\}_{0\leq h\leq k-1})=0.\tag{2.6}$$

By assumption (b) in Section 1, $\partial_I^2 H_0(\mathbf{0},0) \equiv (E\partial)_\varphi \partial_I H_0(\mathbf{0},0)$ is invertible; hence to impose the last two compatibility conditions in (2.3), it suffices to fix $I_{\mathbf{0}}^{(k)} = -[\partial_I^2 H_0(\mathbf{0},0)]^{-1} F^{(k)}(\{X^{(h)}, \xi^{(h)}\}_{0\le h\le k-1})$, which is a known function of $\{X^{(h)},\xi^{(h)}\}_{0\le h\le k-1}$, for any possible choice of the functions $\{\xi^{(h)}\}_{0\le h\le k-1}$.

So, to fulfill the compatibility conditions (2.3) and to continue the inductive construction of $X^{(k)},\xi^{(k)}$, we are left with imposing the second of (2.3), which can only hold if β_{k-2k_0+1} satisfies certain compatibility properties. In fact if $k<2k_0$ there is no requirement because the necessary condition that $[\partial_\beta f]_{\mathbf{0}}^{(k-k_0)}=0$ is automatically satisfied by assumption (c).

For $k=2k_0$ the condition $[\partial_\beta f]_{\mathbf{0}}^{(k_0)}=0$ can be expressed as follows. Let $c \stackrel{def}{=} \frac{1}{k_0!}\partial_\beta^{k_0+1} f_{\mathbf{0}}(\mathbf{0},0,\beta_0)$ and remark that with the notations leading to (1.8) it is $X^{(k_0)}\equiv \sigma X'$: then the second of (2.3) becomes

$$\Big[\partial_{\beta X} f(X^{(0)})\cdot X^{(k_0)} + \frac{1}{k_0!}\partial_\beta^{k_0+1} f(X^{(0)})\,(\beta_1)^{k_0}\Big]_{\mathbf{0}} \equiv \sigma\, a + c\,\beta_1^{k_0} \;=\; 0, \tag{2.7}$$

which means

$$\beta_1 \;=\; \begin{cases} (-a\sigma/c)^{1/k_0} & \text{if } k_0 \text{ is odd,}\\ \pm(-a\sigma/c)^{1/k_0} & \text{if } k_0 \text{ is even and } a\sigma c<0, \end{cases} \tag{2.8}$$

which is the compatibility condition to which β_1 must be subject.

For $k>2k_0$ the second of (2.3) only involves a sum of quantities depending on $X^{(h)}$ with $h\le k-k_0$ and on $\xi^{(h)}$ with $h\le k-2k_0$ *with the exception of a single term*, proportional to $ck_0\beta_1^{k_0-1}\beta_{k-2k_0+1}$, involving $\xi^{(k-2k_0+1)}$. Therefore the second compatibility condition in (2.3) can be fulfilled by properly fixing β_{k-2k_0+1} in terms of $X^{(h)}$ with $h\le k-k_0$ and of $\xi^{(h)}$ with $h\le k-2k_0$, *provided* β_1 *exists and* $\beta_1\ne 0$, i.e. *provided* $a\ne 0$ as assumed here.

This means that if k_0 is odd all β_k are uniquely determined while if k_0 is even there are two possible sequences β_k depending on the two choices for β_1 in (2.8). ■

Remarks. (1) For k_0 even the choice of β_k will be possible only if ε has sign σ such that $-\sigma\, a/c>0$.
(2) With the notation (2.1) one has $X_{\mathbf{0}}^{(k)} = (\mathbf{A}_{\mathbf{0}}^{(k)}, B_{\mathbf{0}}^{(k)},0,0)$ as $\mathbf{a}_{\mathbf{0}}^{(k)}=\mathbf{0}$ and $b_{\mathbf{0}}^{(k)}=\xi^{(k)}$.

3. Tree formalism and formal series

Given that the formal Lindstedt series is well defined, by proceeding as in [G2] a graphical representation of the contributions to the perturbative series will be introduced. The idea to get the rules explained below is to start by representing the *r.h.s.* of (2.2) as a power series in η by simply expanding the functions $\partial_\varphi f, \partial_I H$ in their arguments around $X^{(0)}(\boldsymbol{\psi})$ and then each argument again in powers until one obtains a power series in η.

Since the arguments of f depend on the $X^{(h)}$ with $h<k$ we obtain recursively a natural representation in terms of trees. Consider $(X_{\boldsymbol{\nu}}^{(k)})_\gamma$ where γ ranges in the symbols list

$$\mathcal{A} \stackrel{def}{=} \{A_1,\ldots,A_{N-1},B,\alpha_1,\ldots,\alpha_{N-1},\beta\}\,. \tag{3.1}$$

This notation is more convenient than the alternative one which would simply label the components with a label $\gamma=1,2,\ldots,N,N+1,\ldots,2N$ because the first $N-1$ components play a very different role than the N-th or the remaining ones. It can be disturbing as the labels in (3.1) have the meaning of canonical variables when they do not appear as labels: hence the reader should keep this in mind in what follows. Occasionally, only in cases of possible confusion, we shall use the labels $1,\ldots,2N$ instead of (3.1).

It will be convenient to call the first N labels *action components* and the last N *angle components*; the first $N-1$ of the two groups will be called, respectively, *fast actions* and *fast angles* component labels while the last will be called a *slow action* or, respectively, *slow angle* component label. When, occasionally, it will turn out to be useful, we shall denote the N action labels simply by I and the angle labels by φ. If $\gamma\in\mathcal{A}$ and γ is one of the first N labels (i.e. it is an action component), then $\gamma+N$ will indicate the corresponding label of the angle components and vice versa if $\gamma\in\mathcal{A}$ is an angle component then $\gamma-N$ will denote the corresponding label in the group of action components. With a slight abuse of notation, given $\gamma\in\mathcal{A}$, we shall write $\gamma\in I$ if γ is an action component and $\gamma\in\varphi$ if it is an angle component. If $\gamma,\gamma'\in\mathcal{A}$ then $\delta_{\gamma\gamma'}$ denotes the Kronecker delta; we shall also use the notation

$$\delta_{\gamma\varphi}=\begin{cases}1, & \text{if } \gamma\in\varphi\ ,\\ 0, & \text{if } \gamma\in I\ ,\end{cases}\qquad \delta_{\gamma I}=\begin{cases}1, & \text{if } \gamma\in I\ ,\\ 0, & \text{if } \gamma\in\varphi\ .\end{cases} \tag{3.2}$$

Looking at (2.2) we see that, for $\boldsymbol{\nu}\neq\mathbf{0}$ and $k\geq k_0$, we can write it as

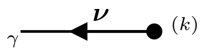
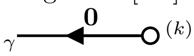
$$\begin{aligned}(X^{(k)}_{\boldsymbol{\nu}})_\gamma=\frac{\delta_{\gamma\gamma'}}{i\boldsymbol{\omega}\cdot\boldsymbol{\nu}}\Big[\sigma\sum_{p\geq1}\sum_{\substack{\boldsymbol{\nu}_0+\boldsymbol{\nu}_1+\ldots+\boldsymbol{\nu}_p=\boldsymbol{\nu}\\ k_1+\ldots k_p=k-k_0}}\frac{1}{p!}(E\partial)_{\gamma'}\partial_{\gamma_1\ldots\gamma_p}f_{\boldsymbol{\nu}_0}(\mathbf{0},0,\beta_0)\prod_{n=1}^p(X^{(k_n)}_{\boldsymbol{\nu}_n}+\xi^{(k_n)}\delta_{\boldsymbol{\nu}_n\mathbf{0}})_{\gamma_n}+\\ +\sum_{p\geq1}\sum_{\substack{\boldsymbol{\nu}_1+\ldots+\boldsymbol{\nu}_p=\boldsymbol{\nu}\\ k_1+\ldots k_p=k}}\frac{1}{p!}(E\partial)_{\gamma'}\partial_{\gamma_1\ldots\gamma_p}H_0(\mathbf{0},0)\prod_{n=1}^p(X^{(k_n)}_{\boldsymbol{\nu}_n})_{\gamma_n}\Big].\end{aligned} \tag{3.3}$$

Here the symbols ∂_γ have to be interpreted as derivatives of $f_{\boldsymbol{\nu}}$ with respect to action arguments of f if γ is an action component, as derivatives with respect to β if $\gamma=\beta$ and as multiplications by $i\boldsymbol{\nu}_i$ if $\gamma=\alpha_i$, $i=1,\ldots,N-1$. *A summation convention is conveniently adopted for pairs of repeated component labels.* In the second line no $\xi^{(k_n)}$ appears because $\xi^{(h)}$ has only angle components but H_0 does not depend on the angles.

For $\boldsymbol{\nu}=\mathbf{0}$ and $\gamma\in I$ (i.e. γ an action component; cf. Remark (2) at the end of Section 2), we use (2.6) and write

$$\begin{aligned}(X^{(k)}_{\mathbf{0}})_\gamma=-\big(\partial_I^2H_0(\mathbf{0},0)^{-1}\big)_{\gamma,\gamma'-N}\,\delta_{\gamma'\varphi}\Big[\sum_{p=2}^\infty\sum_{\substack{k_1+\ldots k_p=k\\ \boldsymbol{\nu}_1+\ldots+\boldsymbol{\nu}_p=\mathbf{0}}}\frac{1}{p!}(E\partial)_{\gamma'}\partial_{\gamma_1\ldots\gamma_p}H_0(\mathbf{0},0)\prod_{n=1}^p(X^{(k_n)}_{\boldsymbol{\nu}_n})_{\gamma_n}+\\ +\sigma\sum_{p=1}^\infty\sum_{\substack{k_1+\ldots k_p=k-k_0\\ \boldsymbol{\nu}_0+\boldsymbol{\nu}_1+\ldots+\boldsymbol{\nu}_p=\mathbf{0}}}\frac{1}{p!}(E\partial)_{\gamma'}\partial_{\gamma_1\ldots\gamma_p}f_{\boldsymbol{\nu}_0}(\mathbf{0},0,\beta_0)\prod_{n=1}^p(X^{(k_n)}_{\boldsymbol{\nu}_n}+\xi^{(k_n)}\delta_{\boldsymbol{\nu}_n\mathbf{0}})_{\gamma_n}\Big].\end{aligned} \tag{3.4}$$

For $\boldsymbol{\nu}=\mathbf{0}$ and γ an angle label the (2.2) are identically satisfied if the $\xi^{(h)}$ are determined as prescribed in Section 2. The equation that fixes $\xi^{(h)}$ is again a relation of the type of (3.3) and (3.4) as we shall discuss in detail later.

Proceeding as in [G2] and [GG1] we represent $(X^{(h)}_{\boldsymbol{\nu}})_\gamma$ as  and $\xi^{(h)} \equiv (\mathbf{0}, 0, \mathbf{0}, \beta^{(k)})$ as  and we realize that (3.3) and (3.4) can be very conveniently (as it turns out) represented by graphs of the type represented in Figure 1.

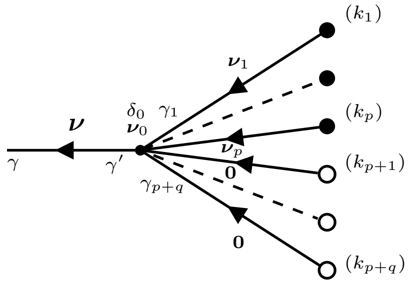
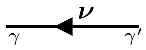


FIGURE 1. Graphical representation of equations (3.3) and (3.4). For $\boldsymbol{\nu}\neq\mathbf{0}$ the graph represents (3.3), while for $\boldsymbol{\nu}=\mathbf{0}$ it represents (3.4).

The label $\delta_0 = 0, 1$, that we call the *degree label* of the node to which it is associated, identifies the terms in (3.3) or (3.4) which contain inside the square brackets H_0 (then $\delta_0 = 0$) or f (then $\delta_0 = 1$). The other label attached to the first node indicates the harmonic $\boldsymbol{\nu}_0$ (mode label) selected in the terms with f (i.e. with $\delta_0 = 1$) and we take $\boldsymbol{\nu}_0 = \mathbf{0}$ if $\delta_0 = 0$ (because H_0 does not depend on the angles). A further component label γ' is attached to the right extreme of the line exiting the central node in Figure 1 and it indicates the derivative $(E\partial)_{\gamma'}$ in (3.3) and (3.4). We call a component label a "right" or "left" component label if it is attached at the beginning or at the end of the (oriented) line.

The component labels attached to the first node determine the components of the tensors defined by the derivatives of $f_{\boldsymbol{\nu}_0}$ or of H_0. The labels attached to the left extreme of each endline (to the right of the bifurcation point in the figure) determine which component of $(X^{(k_n)}_{\boldsymbol{\nu}_n})_{\gamma_n}$ is taken in the products of X's in (3.3) or (3.4). Finally the root line symbolizes the factor that is outside the square brackets in the expression (3.3) or in the expression (3.4): it will be called *propagator* of the line  . Hence the propagator of the line will be the matrix

$$\widetilde{g}_{\gamma\gamma'}(\boldsymbol{\nu}) \;= \begin{cases} (i\boldsymbol{\omega}\cdot\boldsymbol{\nu})^{-1}\,\delta_{\gamma\gamma'}, & \text{if } \boldsymbol{\nu}\neq\mathbf{0}, \\ -\left(\partial_I^2 H_0(\mathbf{0},0)^{-1}\right)_{\gamma,\gamma'-N}\delta_{\gamma I}\delta_{\gamma'\varphi}, & \text{if } \boldsymbol{\nu}=\mathbf{0}, \end{cases} \tag{3.5}$$

and $\boldsymbol{\nu}$ will be called the *momentum* of the line.

Finally the endpoints are divided into endpoints representing $X^{(h)}$ for some h, marked by bullets, and others representing $\xi^{(h)}$ marked by white disks that we shall call *leaves*.

It follows from the analysis of Section 2 that, setting β_1 equal to the constant in (2.8) and replacing $\frac{1}{(k_0-1)!}\partial_\beta^{k_0+1} f_{\mathbf{0}}(\mathbf{0}, 0.\beta_0)$ with $k_0\, c$ by the definition of c before (2.7), the coefficients β_h with $h \geq 2$ can be derived in terms of β_1 from the relation

$$\beta_{k-2k_0+1} = \frac{1}{k_0\, c\, \beta_1^{k_0-1}} \sum_{s\geq 1} \sideset{}{^*}\sum_{\substack{k_1+\ldots+k_s=k-k_0 \\ \boldsymbol{\nu}_0+\boldsymbol{\nu}_1+\ldots+\boldsymbol{\nu}_s=\mathbf{0}}} \frac{1}{s!}(E\partial)_B \partial_{\gamma_1\ldots\gamma_s} f_{\boldsymbol{\nu}_0} \prod_{m=1}^{s} (X^{(k_m)}_{\boldsymbol{\nu}_m} + \xi^{(k_m)}\delta_{\boldsymbol{\nu}_m\mathbf{0}})_{\gamma_m}, \tag{3.6}$$

where the derivatives of $f_{\boldsymbol{\nu}_0}$ are, as above, evaluated at $(\mathbf{0},0,\beta_0)$ and the $*$ on the sum recalls that we are excluding from the sum the contribution equal to $-k_0\,c\,\beta_1^{k_0-1}\beta_{k-2k_0+1}$.

At this point we can repeat the construction and replace each endpoint of Figure 1 ending in a node or in a leaf with a node into which merge several lines each of which comes out of a node or a leaf with some labels $(h),\boldsymbol{\nu},\gamma$ representing $(X^{(h)}_{\boldsymbol{\nu}})_\gamma$ for some $h,\boldsymbol{\nu},\gamma$ or, in the case of leaves, $(\xi^{(h)})_\gamma$.

Using also the relations (3.4) and (3.6), the construction can be iterated until we are left with a tree θ whose endpoints either carry a degree label 1 or are leaves representing β_1: we denote by $L(\theta)$ the set of endpoints representing such leaves. The constraint (3.6) can be automatically implied by imagining that the propagator of a line ℓ with momentum $\boldsymbol{\nu}$ is

$$g_{\gamma\gamma'}(\boldsymbol{\nu})=\begin{cases}(i\boldsymbol{\omega}\cdot\boldsymbol{\nu})^{-1}\,\delta_{\gamma\gamma'}, & \text{if } \boldsymbol{\nu}\neq\mathbf{0},\\ \delta_{\gamma\gamma'}, & \text{if } \boldsymbol{\nu}=\mathbf{0},\ \mathfrak{v}\in L(\theta),\\ -\big(\partial_I^2H_0(\mathbf{0},0)^{-1}\big)_{\gamma,\gamma'-N}\,\delta_{\gamma I}\delta_{\gamma'\varphi}+\dfrac{\delta_{\gamma\beta}\delta_{\gamma' B}}{\sigma\eta^{2k_0-1}k_0\,c\,\beta_1^{k_0-1}}, & \text{if } \boldsymbol{\nu}=\mathbf{0},\ \mathfrak{v}\notin L(\theta),\end{cases}\tag{3.7}$$

where $\mathfrak{v}$ is the node preceding ℓ on τ. We can write $(X_{\boldsymbol{\nu}})_\gamma$ as a formal power series in η, whose terms can be computed in terms of tree values. Let a tree θ be a tree diagram with nodes which look like the node drawn in Figure 1 and let $V(\theta)$, $L(\theta)$ and $\Lambda(\theta)$ denote the sets of nodes, leaves and lines of θ, respectively. The *tree value* $\mathrm{Val}(\theta)$ will be a monomial in η obtained by multiplying
(1) a factor

$$\begin{aligned}[F(\boldsymbol{\nu}_\mathfrak{v})]_{\gamma'_{\ell_\mathfrak{v}},\boldsymbol{\gamma}(\mathfrak{v})} &=\sigma\eta^{k_0}(E\partial)_{\gamma'_{\ell_\mathfrak{v}}}\partial_{\gamma_{\mathfrak{v}1}\ldots\gamma_{\mathfrak{v}p_\mathfrak{v}}}f_{\boldsymbol{\nu}_\mathfrak{v}}(\mathbf{0},0,\beta_0), &&\text{if } \delta_\mathfrak{v}=1,\\ [F(\boldsymbol{\nu}_\mathfrak{v})]_{\gamma'_{\ell_\mathfrak{v}},\boldsymbol{\gamma}(\mathfrak{v})} &=(E\partial)_{\gamma'_{\ell_\mathfrak{v}}}\partial_{\gamma_{\mathfrak{v}1}\ldots\gamma_{\mathfrak{v}p_\mathfrak{v}}}H_0(\mathbf{0},0), &&\text{if } \delta_\mathfrak{v}=0,\end{aligned}\tag{3.8}$$

per each node $\mathfrak{v}\in V(\theta)$ into which merge $p_\mathfrak{v}\ge 1$ lines carrying component labels $\boldsymbol{\gamma}(\mathfrak{v})=(\gamma_{\mathfrak{v}1}\ldots\gamma_{\mathfrak{v}p_\mathfrak{v}})$ and emerges a line $\ell_\mathfrak{v}$ carrying a label $\gamma'_{\ell_\mathfrak{v}}$;
(2) a factor $\sigma\eta^{k_0}(E\partial)_{\gamma'_{\ell_\mathfrak{v}}}f_{\boldsymbol{\nu}_\mathfrak{v}}(\mathbf{0},0,\beta_0)$ per each endnode $\mathfrak{v}\notin L(\theta)$ (note that necessarily $\delta_\mathfrak{v}=1$);
(3) a factor $\eta\beta_1$ per each leaf $\mathfrak{v}\in L(\theta)$ (note that necessarily $\boldsymbol{\nu}_\mathfrak{v}=\mathbf{0}$ and $\gamma_{\ell_\mathfrak{v}}=B$);
(4) a factor $g_{\gamma_\ell\gamma'_\ell}(\boldsymbol{\nu})$ per line $\ell\in\Lambda(\theta)$, given by (3.7).

Note that the construction described above and in Section 2 forbids the presence in the tree diagrams of some configurations of nodes. More precisely, calling *trivial* the nodes $\mathfrak{v}$ with $p_\mathfrak{v}=1$ and $\boldsymbol{\nu}_\mathfrak{v}=\mathbf{0}$ and *b–trivial* the nodes $\mathfrak{v}$ with $p_\mathfrak{v}=k_0$, $\boldsymbol{\nu}_\mathfrak{v}=\mathbf{0}$, $\gamma'_{\ell_\mathfrak{v}}=B$ and immediately preceded by at least k_0-1 leaves, the following configurations of nodes:
(i) *trivial nodes with* $\delta_\mathfrak{v}=0$ *and the entering line with* $\mathbf{0}$ *momentum* and
(ii) *b–trivial nodes with the exiting line with* $\mathbf{0}$ *momentum and all the entering lines with left component labels equal to* β
are not allowed, in the sense that $\mathrm{Val}(\theta)=0$ if θ contains such configurations of nodes. The reason why the trees containing such configurations of nodes are forbidden is a consequence of the use of the relations (3.4) and (3.6). Trees with no forbidden node will be called *allowed trees.*

The result is that the perturbed motion runs on a $(n-1)$-dimensional torus whose equations are formally written as a sum of values of (allowed) tree diagrams, computable by using rules very similar to those listed in [GG1], with the value of an allowed tree θ defined as

$$\mathrm{Val}(\theta)=\frac{1}{|\Lambda(\theta)|!}\Big(\prod_{\mathfrak{v}\in V(\theta)}\varepsilon^{\delta_\mathfrak{v}}\Big)(\eta\beta_1)^{|L(\theta)|}\Big(\prod_{\mathfrak{v}\in V(\theta)}F_\mathfrak{v}\Big)\Big(\prod_{\ell\in\Lambda(\theta)}G_\ell\Big),\tag{3.9}$$

where $G_\ell \stackrel{def}{=} g_{\gamma_\ell\,\gamma'_\ell}(\boldsymbol{\nu}_\ell)$ with $\boldsymbol{\nu}_\ell = \boldsymbol{\nu}_{\ell_\mathfrak{v}} = \sum_{\mathfrak{w}\le\mathfrak{v}} \boldsymbol{\nu}_\mathfrak{v}$, while $F_\mathfrak{v}$ is the node factor (a tensor) defined in items (1) and (2) above. In the product all the labels are summed over, except for the root label γ_{ℓ_0}, where ℓ_0 is the line entering the root. We call $\Theta^o_{k,\boldsymbol{\nu},\gamma}$ the set of trees with *degree* $|L(\theta)| + k_0 \sum_{\mathfrak{v}\in V(\theta)} \delta_\mathfrak{v} - (2k_0-1)\sum_{\ell\in\Lambda(\theta)} \delta_{\boldsymbol{\nu}_\ell \mathbf{0}}\delta_{\gamma_\ell\beta}\delta_{\gamma'_\ell B} = k$, $\boldsymbol{\nu}_{\ell_0} = \boldsymbol{\nu}$ and $\gamma_{\ell_0} = \gamma$ (these are trees whose value is proportional to η^k). As in [GG2], we denote by $\Theta_{k,\boldsymbol{\nu},\gamma}$ the set of trees with k nodes, and with labels $\boldsymbol{\nu}_{\ell_0} = \boldsymbol{\nu}$ and $\gamma_{\ell_0} = \gamma$ associated with the root line.

The definitions above are given so that the formal series for X_γ is given by the sum

$$X^{(0)}_\gamma + \sum_{\boldsymbol{\nu}\in\mathbb{Z}^{N-1}} \mathrm{e}^{i\boldsymbol{\nu}\cdot(\boldsymbol{\psi}+\boldsymbol{\omega}t)} \sum_{k=1}^{\infty} \sum_{\theta\in\Theta_{k,\boldsymbol{\nu},\gamma}} \mathrm{Val}(\theta). \tag{3.10}$$

The above rules give, order by order in powers of η, the solution of the perturbed equations of motion. In particular one can check that the trees contributing a monomial in η of degree $k\ge 1$ to the conjugating function have a number of lines that is bounded above and below proportionally to k. This is a property extensively used in the convergence analysis, for instance to show that the number of non-numbered trees of degree k is bounded by a constant to the power k; cf. [GG2] for details. An explicit bound is $\ge k/k_0$ and $\le 3k_0k$, see Appendix A3.

4. Elimination of the trivial nodes

The problem of proving convergence of the series just defined is very similar to that treated in [GG2]. As in [GG2] the difficulty is that even exploiting the cancellations analogous to those of the maximal tori case, we are left with tree graphs containing chains of subdiagrams with one entering and one exiting lines, carrying the same momentum, that we call "self–energy (sub)diagrams". Naively such subdiagrams are the source of bad bounds on the k–th the contribution proportional to η^k to the series. Proceeding as in in [GG2] we iteratively resum such chains into "renormalized propagators" and change step by step the structure of the perturbation series. At each step we define different rules to compute the tree values: we assign at each line ℓ a scale label $[n_\ell]$, with $n_\ell \ge -1$, depending on the size of its propagators, and, at the n–th step ($n = 0, 1, \ldots$), we will not allow trees containing chains of self-energy diagrams on scale $\ge n-1$; at the same time we will assign to each line a propagator different from that in (3.7), depending on the value of its scale label, as explained below.

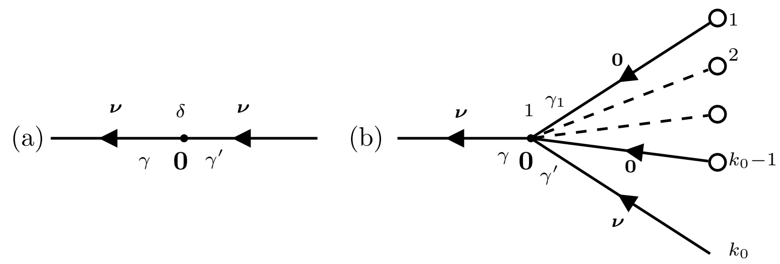


FIGURE 2. Examples of lowest order self-energy clusters on scale $[-1]$. In (a) one has $\gamma = B$ and $\gamma' \in I$ or vice versa, while in (b) one has $\gamma = B$ and $\gamma' = \beta$. The order label δ in the first graph can be $\delta = 0, 1$. The leaves in the second graph represent factors β_1 so that they contribute to the order of the self-energy clustser a power η^{k_0-1}.

The first step is the removal of the trivial nodes and of a class of few other subgraphs that can be present in *allowed trees*, see Section 3, whose value is not 0. Assign, in a way similar to that described in [GG2], a scale label $[-1]$ to the $\mathbf{0}$–momentum propagator and let $\mathcal{S}^{\mathcal{R}}_{k,-1}$ be the set of self-energy clusters T on scale $[-1]$ and degree k (i.e. such that $k = |L(T)| + \sum_{\mathfrak{v}\in V(T)} \delta_{\mathfrak{v}}\, k_0 - (2k_0-1)\sum_{\ell\in\Lambda(T)} \delta_{\boldsymbol{\nu}_\ell \mathbf{0}} \delta_{\gamma_\ell \beta} \delta_{\gamma'_\ell B}$), consistently with the notion of degree defined after (3.9).

A *cluster on scale* $[-1]$ is either a single node or a maximal connected set of nodes and lines, such that all the lines are on scale $[-1]$; the lines of non-zero momentum that enter or exit the nodes of the cluster are called *external lines*. A *self-energy cluster on scale* $[-1]$ is defined as a cluster on scale $[-1]$ such that there are only two external lines (one exiting and one entering), and they are connected to the same node of the cluster. This means that a self-energy cluster is formed by a node $\mathfrak{v}$ with $\mathbf{0}$-mode together with all lines (if any) of momentum $\mathbf{0}$, hence on scale $[-1]$, which are linked to $\mathfrak{v}$ by a path consisting of lines of momentum $\mathbf{0}$; see Figure 2. *A cluster on scale* $[-1]$ *with only one exiting and one entering line, which are not connected to the same node, will not be considered a self-energy cluster on scale* $[-1]$.

The self-energy value, i.e. the value $\mathcal{V}_T(\eta)$ of the self-energy cluster T on scale $[-1]$, is defined, similarly to (3.9), as

$$\mathcal{V}_T(\eta) = \frac{1}{|\Lambda(T)|!} \Big(\prod_{\mathfrak{v}\in V(T)} \varepsilon^{\delta_{\mathfrak{v}}} \Big) (\eta\beta_1)^{|L(T)|} \Big(\prod_{\mathfrak{v}\in V(T)} F_{\mathfrak{v}} \Big) \Big(\prod_{\mathfrak{v}\in\Lambda(T)} G_\ell \Big), \tag{4.1}$$

where $V(T)$, $L(T)$ and $\Lambda(T)$ are the set of nodes, leaves and lines, respectively, contained inside T. Then we can define

$$\mathcal{M}^{[0]}(\eta) = \sum_{k=k_0}^{\infty} \sum_{T\in\mathcal{S}^{\mathcal{R}}_{k,-1}} \mathcal{V}_T(\eta), \tag{4.2}$$

where the sum runs over all the self-energy clusters on scale $[-1]$.

Remarks. (1) The clusters T in (4.1) and (4.2) with $k = k_0$ correspond to the trivial nodes appearing in the tree expansion of previous section. In the next section the general notion of

cluster will be introduced.
(2) The clusters on scale $[-1]$ must contain only lines on scale $[-1]$ and there are infinitely many of them and (4.2) can be illustrated by the two clusters in Figure 2. The self-energy values $\mathcal{V}_T(\eta)$ corresponding to the graphs in Figure 2 give the lowest order(in powers of η) contributions to different entries of the matrix $\mathcal{M}^{[0]}$, as discussed in the caption of Figure 2.
(3) Unlike the case in [GG2], the equation (4.1), defining the matrix $\mathcal{M}^{[0]}(\eta)$, is really an infinite series; but it is still convergent, if η is sufficiently small. Convergence of the series defining $\mathcal{M}^{[0]}$ is a straightforward consequence of the fact that a self–energy cluster of degree k has a number of lines bounded by constant times k, see Appendix A3, and of the fact that the propagators of the lines with $\mathbf{0}$-momentum can be bounded by an $O(1)$ constant if the line is preceded by a leaf or if $\gamma = I_i, \gamma' = \varphi_j$ and can be bounded by an $O(1)$ constant times η^{-2k_0+1} if the line is not preceded by a leaf and $\gamma = \beta, \gamma' = B$.

By construction, $\mathcal{M}^{[0]}(\eta)$ is real and has the following special structure:

$$\mathcal{M}^{[0]}(\eta) = \sum_{k=k_0}^{\infty} \sum_{T \in \mathcal{S}^{\mathcal{R}}_{k,-1}} \mathcal{V}_T = \begin{pmatrix} Q & R \\ P & -Q^{\dagger} \end{pmatrix} \tag{4.3}$$

where the superscript † denotes Hermitian conjugation and P, Q, R are $N \times N$ matrices which to lowest order in η have the form (with a natural meaning of the symbols, in agreement with our convention on the component labels (3.1))

$$P = \partial_I^2 H_0 + \varepsilon \partial_I^2 f_{\mathbf{0}}, \qquad Q = \begin{pmatrix} 0 \\ -\varepsilon \partial_{\beta I} f_{\mathbf{0}} \end{pmatrix}, \qquad R = \begin{pmatrix} 0 & 0 \\ 0 & -\frac{\varepsilon \eta^{k_0-1}}{(k_0-1)!} \partial_\beta^{k_0+1} f_{\mathbf{0}}\, \beta_1^{k_0-1} \end{pmatrix}, \tag{4.4}$$

where $f_{\mathbf{0}}$ has $(\mathbf{0}, 0, \beta_0)$ as arguments. The complete expression, including the higher orders in η, simply replaces the *non-zero* terms in (4.4) by convergent series that we can write

$$P = \partial_I^2 H_0 + \varepsilon \overline{M}_{II}, \qquad Q = \begin{pmatrix} 0 \\ -\varepsilon \overline{M}_{\beta I} \end{pmatrix}, \qquad R = \begin{pmatrix} 0 & 0 \\ 0 & -\varepsilon \eta^{k_0-1} \overline{M}_{\beta\beta} \end{pmatrix}, \tag{4.5}$$

where the entries $\overline{M}_{\gamma\gamma'}(\eta)$ are bounded uniformly in η, for η small enough. They can be computed order by order using the rules of the last section. *The vanishing entries of the matrix $\mathcal{M}^{[0]}(\eta)$ remain zero to all orders*: a property which simply follows from the definition of value of a self–energy cluster and from the remark that a derivative ∂_γ with $\gamma = \boldsymbol{\alpha}$ acting on $f_{\mathbf{0}}$ must be interpreted as a multiplication by $\mathbf{0}$. Note that, since P, Q, R are real and P, R are symmetric, $\mathcal{M}^{[0]}(\eta)$ satisfies the following symmetry properties:

$$E \mathcal{M}^{[0]}(\eta) E = \left[\mathcal{M}^{[0]}(\eta)\right]^T, \qquad \left[\mathcal{M}^{[0]}(\eta)\right]^* = \mathcal{M}^{[0]}(\eta), \tag{4.6}$$

where $*$ denotes complex conjugation and T transposition. A consequence of (4.6) is that $\mathcal{M}^{[0]} E$ is Hermitian.

Now, we formally resum the chains of self–energy clusters on scale $[-1]$ into the new propagator

$$g^{[\geq 0]}(x;\eta) = \frac{1}{ix - \mathcal{M}^{[0]}(\eta)}, \tag{4.7}$$

where $x \neq 0$. This means that we modify the tree expansion described in previous section: the new expansion will involve only trees not containing self–energy graphs on scale $[-1]$ (and in particular containing neither trivial nor b–trivial nodes) and with the propagator of the lines with non-zero momentum replaced by (4.7). It is not clear at all that the new resummed series is well defined: on the contrary it will become clear that it is affected by divergence problems similar to those of the original series. In some sense, we just eliminated a few of the possible source of problems (that are actually an infinite class of divergent sub–diagrams). However, the idea is to begin by eliminating this first few sources of problems and then, step by step, iteratively eliminate one after the other all possibile sources of problems (first the less "dangerous" and then the more and more dangerous ones).

Certainly we must at least suppose that $ix - \mathcal{M}^{[0]}(\eta)$ can be inverted: otherwise the values of the trees representing the new series might even be meaningless! To give a meaning to $(ix - \mathcal{M}^{[0]}(\eta))^{-1}$ it is sufficient to impose $\det(ix - \mathcal{M}^{[0]}(\eta)) \neq 0$ for $x \neq 0$, by eliminating a denumerable dense set of values of η. One can compute the determinant of $(ix - \mathcal{M}^{[0]}(\eta))$, finding $\det(ix - \mathcal{M}^{[0]}(\eta))) = -(ix)^{2(N-1)} \cdot \Big[x^2 + \varepsilon\eta^{k_0-1}\overline{M}_{\beta\beta}\big(\partial^2_{BB}H_0 + \varepsilon\overline{M}_{BB}\big) + \varepsilon^2|\overline{M}_{\beta B}|^2\Big]$, so that the condition of invertibility of $(ix - \mathcal{M}^{[0]}(\eta))$ for $x \neq 0$ becomes

$$x^2 + \varepsilon\eta^{k_0-1}\overline{M}_{\beta\beta}\big(\partial^2_{BB}H_0 + \varepsilon\overline{M}_{BB}\big) + \varepsilon^2|\overline{M}_{\beta B}|^2 \neq 0 \tag{4.8}$$

for all $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}$, $\boldsymbol{\nu} \in \mathbb{Z}^{N-1}$. In computing $\det(ix - \mathcal{M}^{[0]}(\eta))$ the property $\overline{M}_{\beta B} = -\overline{M}_{B\beta} \in \mathbb{R}$ has been used.

If η is chosen according to (4.8), we have that the norm of $g^{[\geq 0]}(x;\eta)$ is equal to the eigenvalue of $\big[ix - \mathcal{M}^{[0]}(\eta)\big]E$ with smallest absolute value: this is because $\|ix - \mathcal{M}^{[0]}(\eta)\| = \|(ix - \mathcal{M}^{[0]}(\eta))E\|$ and, as remarked after (4.6), $(ix - \mathcal{M}^{[0]}(\eta))E$ is Hermitian. Here and in the following we use the uniform norm: given a $2N \times 2N$ complex matrix g, we define $\|g\| = \sup_{z\in\mathbb{C}^{2N},|z|=1} |g\,z|$, where $|z| = \sum_{i=1}^{2N} |z_i|$.

An approximate computation of the eigenvalues of $(ix - \mathcal{M}^{[0]}(\eta))E$, see Appendix A4, implies the following result.

Lemma 2. *There exists an $O(1)$ constant $\rho > 0$ such that if $x^2 > \rho\eta^{2k_0-1}$ then the resummed propagator $g^{[\geq 0]}(x;\eta)$ in (4.7) can be bounded as:*

$$\Big|g^{[\geq 0]}(x;\eta)\Big| \leq \max\Big\{\frac{2\mu_N^{(0)}}{x^2}, \frac{2}{\mu_1^{(0)}}\Big\}, \tag{4.9}$$

where $\mu_1^{(0)}$ and $\mu_N^{(0)}$ are, respectively, the minimum and maximum eigenvalues of P, see (4.3) and (4.5).

From now on we shall proceed following [GG2]. First of all we assume that $|\varepsilon|$ is in an interval $(\overline{\varepsilon}/4, \overline{\varepsilon}]$ such that, by setting $\overline{\varepsilon} = \overline{\eta}^{k_0}$, we can define the integer n_0 through

$$C_0^2 2^{-2(n_0+1)} < \rho\overline{\varepsilon}\,\overline{\eta}^{k_0-1} \leq C_0^2 2^{-2n_0}, \tag{4.10}$$

with ρ as in Lemma 2.

In the first range of scales (in which $x^2 \ge 2C_0^2 2^{-2n_0}$) the small denominators can be bounded by the "classical" small divisor x^2 and we proceed as described in the section "Non–resonant resummations" of [GG2].

For smaller scales we shall see below that the small divisor will bounded below by an $O(1)$ constant times $\min\{x^2, |x^2+\varepsilon\eta^{k_0-1}\lambda^{[n]}(x;\varepsilon)|\}$, with $\lambda^{[n]}(x;\varepsilon)$ a suitable $O(1)$ function. However here the distinction between the *hyperbolic* case ($\varepsilon\eta^{k_0-1}\lambda^{[n]}(x;\varepsilon) > 0$) and the *elliptic* one ($\varepsilon\eta^{k_0-1}\lambda^{[n]}(x;\varepsilon) < 0$) has to be made: in the hyperbolic case the small divisors will be always bounded by an $O(1)$ constant times x^2, even for $x^2 \le O(\varepsilon\eta^{k_0-1})$, while in the elliptic case we will have to proceed differently, essentially as described in the section "Infrared resummations" of [GG2].

5. Multiscale analysis and non resonant resummation

The resummations will be defined via trees with no self–energy clusters on scale $[-1]$ and with lines bearing further labels. Moreover the definition of propagator will be changed, hence the values of the trees will be different from the ones in Section 3: they are constructed recursively as follows.

We introduce a *multiscale decomposition* (see [GG2]): we call $\psi(D)$ a C^∞ non-decreasing compact support function defined for $D \ge 0$,

$$\psi(D) = 1, \quad \text{for} \quad D \ge C_0^2, \qquad \psi(D) = 0, \quad \text{for} \quad D \le C_0^2/4, \tag{5.1}$$

where C_0 is the Diophantine constant of $\boldsymbol{\omega}$, and let $\chi(D) = 1-\psi(D)$. Define also $\psi_n(D) = \psi(2^{2n}D)$ and $\chi_n(D) = \chi(2^{2n}D)$ for all $n \ge 0$. Hence $\psi_0 = \psi$, $\chi_0 = \chi$ and

$$1 \equiv \psi_n(D(x)) + \chi_n(D(x)), \qquad \text{for all } n \ge 0, \tag{5.2}$$

for all choices of the function $D(x) \ge 0$: in particular for $D(x) = x^2$, that we shall now use.

A simple way to represent the value of a tree as sum of many terms is to make use of the identity in (5.2). The resummed propagator $g^{[\ge 0]}(x;\eta) \stackrel{def}{=} (ix - \mathcal{M}^{[0]}(\eta))^{-1}$ of each line with non-zero momentum (hence with $x \neq 0$) is written as

$$g^{[\ge 0]}(x;\eta) = \psi_0(x^2)\, g^{[\ge 0]}(x;\eta) + \chi_0(x^2)\, g^{[\ge 0]}(x;\eta) \stackrel{def}{=} g^{[0]}(x;\eta) + g^{\{\ge 1\}}(x;\eta), \tag{5.3}$$

and we note that $g^{[0]}(x;\eta)$ vanishes if x^2 is smaller than $(C_0/2)^2$.

If we replace each $g^{[\ge 0]}(x;\eta)$ with the sum in (5.3) then the value of each tree with k nodes is split as a sum of up to 2^k terms which can be identified by affixing on each line with momentum $\boldsymbol{\nu} \neq \mathbf{0}$ a label $[0]$ or $\{\ge 1\}$. Further splittings of the tree values can be achieved as follows.

Definition 2 (Propagators). *Let n_0 be an integer, and for $1 \le p < n_0$, give $2N \times 2N$ matrices $\mathcal{M}^{[p]}(x;\eta)$ satisfying the symmetry properties*

$$E\mathcal{M}^{[p]}(x;\eta)E = \left[\mathcal{M}^{[p]}(-x;\eta)\right]^T, \qquad \left[\mathcal{M}^{[p]}(x;\eta)\right]^* = \mathcal{M}^{[p]}(-x;\eta). \tag{5.4}$$

Let $\mathcal{M}^{[0]}(x;\eta) \equiv \mathcal{M}^{[0]}(\eta)$ *and* $\mathcal{M}^{[\le n]}(x;\eta) = \sum_{p=0}^{n} \mathcal{M}^{[p]}(x;\eta)$*. Define for* $n_0 \ge n \ge 1$ *the* propagators

$$\begin{aligned} g^{[n-1]}(x;\eta) &\stackrel{def}{=} \frac{\psi_n(x^2)\prod_{m=0}^{n-1}\chi_m(x^2)}{ix-\mathcal{M}^{[\le n-1]}(x;\eta)}, \\ g^{\{\ge n\}}(x;\eta) &\stackrel{def}{=} \frac{\prod_{m=0}^{n-1}\chi_m(x^2)}{ix-\mathcal{M}^{[\le n-1]}(x;\eta)}, \\ g^{[\ge n]}(x;\eta) &\stackrel{def}{=} \frac{\prod_{m=0}^{n-1}\chi_m(x^2)}{ix-\mathcal{M}^{[\le n]}(x;\eta)}, \end{aligned} \tag{5.5}$$

and $g^{[0]}(x;\eta) = \psi_0(x^2)\,(ix - \mathcal{M}^{[0]}(\eta))^{-1}$*. We call the labels* $[n], \{\ge n\}, [\ge n]$ *scale labels.*

Remarks. (1) The matrices $\mathcal{M}^{[p]}(x;\eta)$ will be defined recursively under the requirement that the functions $(\mathbf{A}, B, \mathbf{a}, b)$ defining the parametric equations (2.1) of the invariant torus will be expressed in terms of trees whose lines carry scale labels indicating that their values are computed with the propagators in (5.5).
(2) To have the propagators in (5.5) well defined, we have to eliminate for each value of p a denumerable set of ε's, by imposing that $(ix - \mathcal{M}^{[\le p]}(x;\eta))$ be invertible, in analogy with (4.8).
(3) So far n_0 can be any integer number. It will be fixed as prescribed after Lemma 2, in such a way that a bound like (4.9) will hold for all propagators $g^{[p]}$ with $p < n_0$.

To define recursively the matrices we introduce the notions of clusters and of self-energy clusters of a tree whose lines and nodes carry the same labels introduced so far and *in addition* each line carries a scale label which can be either $[-1]$, if the momentum of the line is zero, or $[p]$, with $0 \le p < n_0$, or $[\ge n_0]$, with n_0 the same integer appearing in the statement of Definition 2 (still to be suitably fixed). Given a tree θ decorated in this way we give the following definition, for $n < n_0$.

Definition 3 (Clusters). *(i) A* cluster T *on scale* $[n]$*, with* $0 \le n$*, is a maximal set of nodes and lines connecting them with propagators on scales* $[p]$*,* $p \le n$*, one of which, at least, on scale exactly* $[n]$*. We denote with* $V(T)$*,* $L(T)$ *and* $\Lambda(T)$ *the set of nodes, the set of leaves and the set of lines, respectively, contained in* T*. The number of nodes in* T *will be denoted by* k_T*.*
(ii) The $m_T \ge 0$ *lines entering the cluster* T *and the possible line coming out of it (unique if existing at all) are called the* external lines *of the cluster* T*.*
(iii) Given a cluster T *on scale* $[n]$*, we shall call* $n_T = n$ *its scale.*

Remarks. (1) The clusters on scale $[-1]$ were defined before (4.1): they can contain either only lines with scale $[-1]$ or no line at all (i.e. they can contain just a single node).
(2) Here $n < n_0$. However the definition above is given in such a way that it will extend unchanged when also scales equal to or larger than n_0 will be introduced.
(3) The clusters of a tree can be regarded as sets of lines hierarchically ordered by inclusion and have hierarchically ordered scales.
(4) A cluster T is not a tree (in our sense); however we can uniquely associate a tree with it by adding the entering and the exiting lines and by imagining that the lower extreme of the exiting line is the root and that the highest extremes of the entering lines are nodes carrying a mode label

equal to the momentum flowing into them (cf. [GG2], Figure 3).

Definition 4 (Self-energy clusters). *A* self-energy cluster *on scale* $[n]$*, with* $n \geq 0$*, of a tree* θ *will be any cluster* T *on scale* $[n]$ *with the following properties:*
(i) T *has only one entering line* ℓ_T^2 *and one exiting line* ℓ_T^1*;*
(ii) $\sum_{\mathfrak{v}\in V(T)} \boldsymbol{\nu}_{\mathfrak{v}} = \mathbf{0}$*;*
(iii) there is no line on scale $[-1]$ *along the path connecting* ℓ_T^2 *to* ℓ_T^1*.*
We call k_T *the number of nodes in* $V(T)$*.*

Remarks. (1) The essential property of a self-energy cluster is that it has necessarily just one entering line and one exiting line, and both have *equal momentum* (because $\sum_{\mathfrak{v}\in V(T)} \boldsymbol{\nu}_{\mathfrak{v}} = \mathbf{0}$). Note that both scales of the external lines of a self-energy cluster T are strictly larger than the scale of T regarded as a cluster, but they can be different from each other by just one unit.
(2) The self-energy clusters on scale $[-1]$ were defined before (4.1).
(3) For $n \geq 0$, the number of nodes of any self-energy cluster on scale $[n]$ is ≥ 2, and the corresponding degree is $\geq 2k_0$. This can be seen as follows. Call T_0 the connected subset of T containing no line on scale $[-1]$ and containing the two nodes to which the external lines are attached. Then T_0 must have at least two nodes with $\delta = 1$, and $T \setminus T_0$ is the union of subtrees with positive degree.
(4) The clusters which satisfy properties (i) and (ii), but not (iii) are not considered self-energy clusters. The same happened for the self-energy clusters on scale $[-1]$. The reason for such a definition is that the cancellation mechanisms that will imply the bounds needed on the derivatives of the self-energy values (see next definition and Lemma 3 below) can be derived only under such an extra condition. On the other hand the clusters which verify only the first two properties, but contain lines on scale $[-1]$ along the path connecting the external lines, require no resummation, and can be dealt with in the same way as the other clusters which are not self-energy clusters. We note that in [GG2] property (iii) was explicitly required only for self-energy clusters on scale $[-1]$ but it was not mentioned any more for the others: the proofs in [GG2], however, implicitly used property (iii) also for the self–energy clusters on scale > -1.

Definition 5 (Renormalized trees). *Let* $\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu},\gamma}$ *be the set of trees with degree* k *(see comments after (3.9)), root line momentum* $\boldsymbol{\nu}$ *and root label* γ *which contain no self-energy clusters. Such trees will be called* renormalized trees.

Definition 6 (Self–energy matrices). *(i) We denote with* $\mathcal{S}^{\mathcal{R}}_{k,n}$ *the set of self-energy clusters with degree* k *and scale* $[n]$ *which do not contain other self-energy clusters; we call them* renormalized self-energy clusters *on scale* n*.*
(ii) Given a self-energy cluster $T \in \mathcal{S}^{\mathcal{R}}_{k,n}$ *we shall define the* self-energy value *of* T *as the matrix*[1]

$$\mathcal{V}_T(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\eta) = \frac{1}{|\Lambda(T)|!}\Big(\prod_{\mathfrak{v}\in V(T)} \varepsilon^{\delta_{\mathfrak{v}}}\Big)(\eta\beta_1)^{|L(T)|}\Big(\prod_{\mathfrak{v}\in V(T)} F_{\mathfrak{v}}\Big)\Big(\prod_{\ell\in\Lambda(T)} g_\ell^{[n_\ell]}\Big) \tag{5.6}$$

[1] This is a matrix because the self-energy cluster inherits the labels γ, γ' attached to the endnode of the entering line and to the initial node of the exiting line.

where $g_\ell^{[n_\ell]} = g^{[n_\ell]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\eta)$*. Note that, necessarily,* $n_\ell \le n$*. The* $k_T - 1$ *lines of the self-energy cluster* T *will be imagined as distinct and to carry a number label ranging in* $\{1,\ldots,k_T-1\}$*.*
(iii) The self-energy matrices $\mathcal{M}^{[n]}(x;\eta)$, $n \ge 1$, *will be defined recursively as*

$$\mathcal{M}^{[n]}(x;\eta) = \Big(\prod_{p=0}^{n-1}\chi_p(x^2)\Big)\sum_{k=2k_0}^{\infty}\sum_{T\in\mathcal{S}^{\mathcal{R}}_{k,n-1}}\mathcal{V}_T(x;\eta) \stackrel{def}{=} \Big(\prod_{p=0}^{n-1}\chi_p(x^2)\Big)M^{[n]}(x;\eta), \tag{5.7}$$

where the self-energy values are evaluated by means of the propagators on scales $[p]$*, with* $p = -1,0,\ldots,n$*.*

The definition (5.7) makes sense because we have already defined the propagators on scale [0] and the matrices $\mathcal{M}^{[0]}(x;\eta)\equiv\mathcal{M}^{[0]}(\eta)$ (cf. Definition 2). Of course we have still to check that the series converges.

With the above new definitions let $h_{\boldsymbol{\nu},\gamma}$ with $\gamma = \mathbf{A}, B, \boldsymbol{\alpha}, \beta$ be the values of $\mathbf{A}_{\boldsymbol{\nu}}, B_{\boldsymbol{\nu}}, \mathbf{a}_{\boldsymbol{\nu}}, b_{\boldsymbol{\nu}}$. We have the formal identities

$$h_{\boldsymbol{\nu},\gamma} = \sum_{k=1}^{\infty}\sum_{\theta\in\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu},\gamma}}\mathrm{Val}(\theta), \tag{5.8}$$

where we have redefined the *value* of a tree $\theta\in\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu},\gamma}$ as

$$\mathrm{Val}(\theta) = \frac{1}{|\Lambda(\theta)|!}\Big(\prod_{\mathfrak{v}\in V(\theta)}\varepsilon^{\delta_{\mathfrak{v}}}\Big)(\eta\beta_1)^{|L(T)|}\Big(\prod_{\ell\in\Lambda(\theta)}g^{[n_\ell]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\eta)\Big)\Big(\prod_{\mathfrak{v}\in V(\theta)}F_{\mathfrak{v}}\Big), \tag{5.9}$$

with $[n_\ell] = [-1],[0],\ldots,[n_0-1],[\ge n_0]$. Note that (5.8) is not a power series in η.

The statement in (5.8) is checked to be an identity between formal series (as in the corresponding check in [GG2]).

As a first step to bypass the formal level, the series (5.7) defining $M^{[n]}(x;\eta)$ has to be shown to be really convergent. This will be true because in the evaluation of $M^{[n]}(x;\eta)$ *the only involved propagators have scales* $[p]$ *with* $p\le n-1$ so that, see the factors $\psi_n(x^2),\chi_n(x^2)$ in (5.5), their denominators not only do not vanish but have controlled sizes that can be bounded below proportionally to x^2 by (4.9), i.e. simply by a constant times $C_0^2|\boldsymbol{\nu}|^{-2\tau_0}$. Using this fact one can actually show that the matrices $M^{[n]}(x;\eta)$ are well defined and satisfy symmetry properties similar to (5.4).

Furthermore *cancellations* similar to the maximal KAM tori cancellations hold even in this case so that $M^{[n]}(x;\eta)$ has a special structure, as described in the following Lemma (see Appendices A5 and A6 for a proof).

Lemma 3. *Let* $\overline{\varepsilon} < \varepsilon_0$ *with* ε_0 *small enough, and* $\varepsilon\in I(\overline{\varepsilon}) = (\overline{\varepsilon}/4,\overline{\varepsilon}]$*. Define also* $\overline{\eta}$ *through* $\overline{\varepsilon} = \overline{\eta}^{k_0}$*. If* $1\le n<n_0$*, with* n_0 *defined in (4.10)), the following properties hold.*
(i) The series defining the matrices $\mathcal{M}^{[\le n]}(x;\eta)$*,* $x = \boldsymbol{\omega}\cdot\boldsymbol{\nu}$*, converge and the matrices satisfy the same symmetry properties noted for* $\mathcal{M}^{(0)}$

$$E\mathcal{M}^{[\le n]}(x;\eta)E = \big[\mathcal{M}^{[\le n]}(-x;\eta)\big]^T,\qquad \big[\mathcal{M}^{[\le n]}(x;\eta)\big]^* = \mathcal{M}^{[\le n]}(-x;\eta), \tag{5.10}$$

where E is the $2N\times 2N$ symplectic matrix, (1.4). Hence, by (5.10), $(ix-\mathcal{M}^{[\le n]})E$ is Hermitian.
(ii) The matrix $\mathcal{M}^{[\le n]}(x;\eta)$ can be written in the form

$$\begin{pmatrix} Q^{[\le n]}(x;\eta) & R^{[\le n]}(x;\eta) \\ P^{[\le n]}(x;\eta) & -Q^{[\le n]\dagger}(x;\eta) \end{pmatrix}, \tag{5.11}$$

where, if $\chi_{n-1}(x^2)\neq 0$,

$$P^{[\le n]}(x;\eta)=\begin{pmatrix} \partial^2_{\mathbf{A}\mathbf{A}}H_0+\varepsilon\overline{M}^{[\le n]}_{\mathbf{A}\mathbf{A}}(x;\eta) & \partial^2_{\mathbf{A}B}H_0+\varepsilon\overline{M}^{[\le n]}_{\mathbf{A}B}(x;\eta) \\ \partial^2_{B\mathbf{A}}H_0+\varepsilon\overline{M}^{[\le n]}_{B\mathbf{A}}(x;\eta) & \partial^2_{BB}H_0+\varepsilon\overline{M}^{[\le n]}_{BB}(x;\eta) \end{pmatrix}=[P^{[\le n]}(x;\eta)]^\dagger,$$

$$Q^{[\le n]}(x;\varepsilon)=\begin{pmatrix} ix\varepsilon^2\overline{M}^{[\le n]}_{\boldsymbol{\alpha}\mathbf{A}}(x;\eta) & ix\varepsilon^2\overline{M}^{[\le n]}_{\boldsymbol{\alpha}B}(x;\eta) \\ \varepsilon\overline{M}^{[\le n]}_{\beta\mathbf{A}}(x;\eta) & \varepsilon\overline{M}^{[\le n]}_{\beta B}(x;\eta) \end{pmatrix}, \tag{5.12}$$

$$R^{[\le n]}(x;\varepsilon)=\begin{pmatrix} x^2\varepsilon^2\overline{M}^{[\le n]}_{\boldsymbol{\alpha}\boldsymbol{\alpha}}(x;\eta) & ix\varepsilon^2\overline{M}^{[\le n]}_{\boldsymbol{\alpha}\beta}(x;\eta) \\ -ix\varepsilon^2\overline{M}^{[\le n]}_{\beta\boldsymbol{\alpha}}(x;\eta) & \varepsilon\eta^{k_0-1}\overline{M}^{[\le n]}_{\beta\beta}(x;\eta) \end{pmatrix}=[R^{[\le n]}(x;\eta)]^\dagger,$$

(iii) The entries $\overline{M}^{[\le n]}_{\gamma,\gamma'}$ are such that the corrections $\overline{M}^{[n]}_{\gamma,\gamma'}(x;\eta)=\overline{M}^{[\le n]}_{\gamma,\gamma'}(x;\eta)-\overline{M}^{[\le n-1]}_{\gamma,\gamma'}(x;\eta)$ are bounded, uniformly in x and ε for ε small enough, by $B\,e^{-\kappa_1 2^{n/\tau}}$ for $n<n_0$ and for suitable n_0-independent constants $B,\kappa_1,\tau>0$; one can take $\tau=\tau_0$
(iv) One has

$$\|\partial_x\mathcal{M}^{[\le n]}(x,\eta)\|\le B\varepsilon^2,\qquad \|\partial_\varepsilon\big[\mathcal{M}^{[\le n]}(x,\eta)-\mathcal{M}^{[0]}(x,\eta)\big]\|\le B\,\varepsilon, \tag{5.13}$$

where the derivatives must be interpreted in the sense of Whitney and the constants $B,\kappa_1,\tau>0$ can be taken the same as in item (iii).

Remarks. (1) The symmetry property (5.10) is proved in Appendix A5. Note that at the first step $\mathcal{M}^{[0]}(x;\eta)$ satisfies it, see (4.6).
(2) The key property in (5.12) is that some entries of $Q^{[\le n]}(x;\eta)$ and $R^{[\le n]}(x;\eta)$ are proportional to x or x^2: this is proved by exploiting cancellations among families of self–energy clusters, as described in detail in Appendix A6; note that the single self–energy clusters contributing to $\mathcal{M}^{[\le n]}(x;\eta)$ do not have in general the structure in (5.12) and only their sum has.

A crucial technical point in the proof of Lemma 3 is the fact that, if the scale n_ℓ of a line ℓ is smaller than n_0, as defined in (4.10), then the corresponding propagator admits a bound that is qualitatively the same as (4.9). More precisely the following result holds.

Lemma 4. *Let $\varepsilon\in(\overline{\varepsilon}/4,\overline{\varepsilon}]$. The propagator $g^{[0]}(x;\eta)$ admits the same bound as $g^{[\le 0]}(x;\eta)$ in (4.9). For $1\le n<n_0$, with n_0 given by (4.10), the propagator on scale $[n]$ can be bounded by*

$$\left|g^{[n]}(x;\eta)\right|\le\frac{C}{x^2}, \tag{5.14}$$

for some positive constant C.

The proof of Lemma 4 proceeds as that of Lemma 2, in Appendix A4, and we do not repeat it here. Once the bound (5.14) is established, the proof of convergence is the same as the one discussed in Appendix A3 of [GG2] and we do not repeat it here. Item (ii) of Lemma 3 simply follows from convergence and from the remark that $\mathcal{M}^{[\le n]}(x;\eta) - \mathcal{M}^{[0]}(x;\eta)$ is of order ε^2 (cf. Remark (3) after Definition 4). The bounds in items (iii) and (iv) of Lemma 3 also follow from the proof of convergence, see Appendix A3 of [GG2].

We have therefore constructed a new representation of the formal series for the parametric equations for the invariant torus: in it only trees with lines carrying a scale label $[-1], [0], \ldots, [n_0-1]$ or $[\ge n_0]$ and *no self–energy clusters* are present (note that, so far, self–energy clusters may have only scales $[n]$ with $n < n_0$). The above lemma will be the starting block of the construction that follows.

6. Renormalization: the infrared resummation

From the proof of Lemma 2 it is clear that for $x^2 \le \rho\eta^{2k_0-1}$ and $n \ge n_0$ it will not be possible to bound $g^{[n]}(x;\eta)$ by a constant times x^{-2}.

So, the first problem to face when reaching scales $n \ge n_0$ is the computation of the eigenvalues of $(ix - \mathcal{M}^{[\le n]}(x;\eta))E$, in terms of which an estimate of the size of the propagator $g^{[n]}$ can be deduced.

If $\mathcal{M}^{[\le n]}(x;\eta)$ does have the structure in (5.12) then an approximate computation of the eigenvalues of $(ix - \mathcal{M}^{[\le n]})E$ leads to the following Lemma, proved in Appendix A7.

Lemma 5. *Let $n \ge n_0$ and let us assign a matrix $\mathcal{M}^{[\le n]}(x;\eta)$ satisfying the properties in (5.12) admitting right and left derivatives with respect to x and ε, bounded as the derivatives in (5.13), and having the structure described by (5.12), with the entries $\overline{M}^{[\le n]}_{\gamma,\gamma'}(x;\eta)$ such that the corrections $\overline{M}^{[n]}_{\gamma,\gamma'}(x;\eta) = \overline{M}^{[\le n]}_{\gamma,\gamma'}(x;\eta) - \overline{M}^{[\le n-1]}_{\gamma,\gamma'}(x;\eta)$ can be bounded (as in item (iii) of Lemma 3) by $|\overline{M}^{[n]}_{\gamma,\gamma'}(x;\eta)| \le Be^{-\kappa_1 2^{n/\tau_1}}$, for suitable constants B, κ_1 and τ_1. Then the uniform norm of $(ix - \mathcal{M}^{[\le n]}(x;\eta))$ can be bounded below by*

$$\|ix - \mathcal{M}^{[\le n]}(x;\eta)\| \ge \frac{1}{4\mu}\min\{x^2, |x^2 - \lambda^{[n]}(x;\eta)|\}, \tag{6.1}$$

where $\mu = \mu_N^{(0)}$ is the largest eigenvalue of $P^{[0]}$ and $\lambda^{[n]}(x;\eta) = \ell^{[n]}(x;\eta)\varepsilon\eta^{k_0-1}$ is a real function with $\ell^{[n]}(x;\eta) = ck_0\partial_B^2 H_0 \beta_1^{k_0-1}\big(1+O(\eta)\big)$. Furthermore $\lambda^{[n]}(x;\eta)$ is right and left differentiable in ε and x and the derivatives satisfy the following dimensional bounds:

$$C^{-1}\eta^{k_0-1} \le |\partial_\varepsilon^\pm \lambda^{[n]}(x;\eta)| \le C\eta^{k_0-1}, \qquad |\partial_x^\pm \lambda^{[n]}(x;\eta)| \le C\eta^{k_0}, \tag{6.2}$$

for some positive constant C.

Remarks. (1) The bound (6.1) suggests to replace the classical small divisor x^2 used in previous sections by the (n–dependent) quantity $\min\{x^2, |x^2 - \lambda^{[n]}(x;\eta)|\}$, that is essentially what we shall do below. In particular we shall replace the argument x^2 of the support functions ψ, χ by a quantity

$\Delta^{[n]}(x;\eta)$, behaving like $\min\{x^2, |x^2 - \lambda^{[n]}(x;\eta)|\}$, which will be the measure of the strength of the resonance for scales larger than n_0. With this choice we shall in general introduce a singularity in the definition of the propagators and self–energy matrices on scales $n \geq n_0$: this is due to the presence of a minimum and of an absolute value in the definition of $\Delta^{[n]}(x;\eta)$. This is ultimately why in (6.2) the right and left derivatives appear, rather than the plain derivatives, as in (5.13) above. This could be avoided by using a smoothed version of the quantity $\Delta^{[n]}(x;\eta)$ introduced below, but we shall not discuss it here.
(2) The bound on the derivatives of $\lambda^{[n]}(x;\eta)$ with respect to ε follows from the expression of $\mathcal{M}^{[0]}(\eta)$, in (4.3) to (4.5), and from the bounds (5.13), which allow to control the corrections. On the contrary the bound on on the derivatives of $\lambda^{[n]}(x;\eta)$ with respect to x does not follow directly from (5.13), and it is explained in Appendix A7.

Depending on the sign of ε, from now on, the analysis changes qualitatively. If ε is such that $c\varepsilon\beta_1^{k_0-1} < 0$ (so that $\lambda^{[n]}(x;\eta)$ in (6.1) is negative), the minimum in (6.1) is always realized by the classical small divisor x^2. This implies that in this range of scales it is possible to proceed in the same way discussed above in Section 5. We do not repeat the details, and we concentrate on the opposite case, namely ε with the sign such that $\lambda^{[n]}(x;\eta) > 0$, which presents new difficulties.

In this case convergence problems can still arise from the propagators $g^{[\geq n_0]}(x;\eta)$, which become uncontrollably large for $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}$ close to the eigenvalue $\lambda^{[n]}(x;\eta)$ in (6.1).

We introduce a sequence of *self–energies* $\underline{\lambda}^{[n]}(\eta)$, [GG2], representing the locations of the singularities of $ix - \mathcal{M}^{[\leq n]}(x;\eta)$ and, correspondingly, we introduce a sequence of *propagator divisors* $\Delta^{[n]}(x;\eta)$, which will give a bound for the size of the propagator on scale $n \geq n_0$. Then, we modify the scale decomposition, measuring the strength of the singularity in terms of $\Delta^{[n]}(x;\eta)$ rather than, as done in Section 5, in terms of the classical small divisor x^2.

The choice of the scale decomposition will be done in such a way that the dimensional bounds for the propagators will be the same as for those with scales $n < n_0$: this will reduce the analysis of the infrared resummations to the same convergence proof discussed in Appendix A3 of [GG2].

We introduce the following definition.

Definition 7 (Self-energies and propagator divisors). *Let the function $\lambda^{[n]}(x;\eta)$ be as in Lemma 5.*
(i) The sequence of self-energies $\underline{\lambda}^{[n]}(\eta)$ *is defined for $n \geq n_0$ by*

$$\underline{\lambda}^{[n]}(\eta) \stackrel{def}{=} \lambda^{[n]}\Big(\sqrt{\underline{\lambda}^{[n-1]}(\eta)}, \eta\Big), \qquad \underline{\lambda}^{[n_0]}(\eta) \stackrel{def}{=} \lambda^{[n_0]}(0;\eta), \tag{6.3}$$

provided $\underline{\lambda}^{[n]}(\eta) \geq 0$, $n \geq n_0$.
(ii) The propagator divisors *are defined for $n \geq n_0$ by*

$$\Delta^{[n]}(x;\eta) \stackrel{def}{=} \min\{x^2, \big|x^2 - \underline{\lambda}^{[n]}(\eta)\big|\}. \tag{6.4}$$

By repeating the analysis of Section 5 we can represent the function $X(\boldsymbol{\psi})$ via sums of values of trees whose lines can carry scale labels $[-1], [0], \ldots, [n_0-1], [n_0], [n_0+1], \ldots$ and which contain no self-energy clusters (i.e. they are renormalized trees; see Definition 5 in Section 5). The new propagators will be defined by the same procedure used to eliminate the self-energy clusters on scales $[n]$ with $n \le n_0 - 1$. However for $n \ge n_0$ the scale of a line will be determined in terms of the recursively defined $\Delta^{[n]}(x;\eta)$, (6.4), rather than in terms of x^2, see (5.2).

Set $X_{n_0-1}(x) \stackrel{def}{=} \prod_{m=0}^{n_0-1} \chi_m(x^2)$, $Y_n(x;\eta) \stackrel{def}{=} \prod_{m=n_0}^{n} \chi_m(\Delta^{[m]}(x;\eta))$ for $n \ge n_0$ and $Y_{n_0-1} \equiv 1$. The definition of the new propagators will be

$$\begin{aligned} g^{[n_0]} &\stackrel{def}{=} X_{n_0-1}(x)\,\psi_{n_0}(\Delta^{[n_0]}(x;\eta))\,(x^2 - \mathcal{M}^{[\le n_0]}(x;\eta))^{-1}, \\ g^{[n_0+1]} &\stackrel{def}{=} X_{n_0-1}(x)\,\chi_{n_0}(\Delta^{[n_0]}(x;\eta))\,\psi_{n_0+1}(\Delta^{[n_0+1]}(x;\eta))\,(x^2 - \mathcal{M}^{[\le n_0+1]}(x;\eta))^{-1}, \\ &\dots \\ g^{[n]} &\stackrel{def}{=} X_{n_0-1}(x)\,Y_{n-1}(x;\eta)\,\psi_n(\Delta^{[n]}(x;\eta))\,(x^2 - \mathcal{M}^{[\le n]}(x;\eta))^{-1}, \end{aligned} \tag{6.5}$$

and so on, using indefinitely the identity $1 \equiv \psi_n(\Delta^{[n]}(x;\eta)) + \chi_n(\Delta^{[n]}(x;\eta))$ to generate the new propagators.

In this way we obtain a formal representation of $X(\boldsymbol{\psi})$ as a sum of tree values in which only renormalized trees appear and in which each line ℓ carries a *scale label* $[n_\ell]$. This means that we can formally write $X(\boldsymbol{\psi})$ as in (5.8), with $\mathrm{Val}(\theta)$ defined according to (5.9), but now the scale label $[n_\ell]$ is such that n_ℓ can assume all integer values ≥ -1, and no line carries a scale label like $[\ge n]$: *only scale labels* $[n]$ *are possible.*

We can summarize the discussion above in the following definition.

Definition 8 (Propagators and self–energy matrices). *Given a sequence of $2N \times 2N$ matrices $\mathcal{M}^{[\le m]}(x;\eta)$, $m \ge 1$, let $\mathcal{M}^{[n]}(x;\eta) = \mathcal{M}^{[\le n]}(x;\eta) - \mathcal{M}^{[\le n-1]}(x;\eta)$ with $\mathcal{M}^{[\le 0]}(x;\eta) \equiv \mathcal{M}^{[0]}(\eta)$ (see (4.2)), so that $\mathcal{M}^{[\le n]}(x;\eta) = \sum_{m=0}^{n} \mathcal{M}^{[m]}(x;\eta)$. Setting $\Delta^{[n]}(x;\eta) \equiv x^2$ if $n < n_0$, define for all $n \ge 0$*

$$g^{[n]}(x;\eta) = \frac{\psi_n(\Delta^{[n]}(x;\eta)) \prod_{m\ge 0}^{n-1} \chi_m(\Delta^{[m]}(x;\eta))}{x^2 - \mathcal{M}^{[\le n]}(x;\eta)}. \tag{6.6}$$

(for $n = 0$ this means $\psi_0(x^2)\,(ix^2 - \mathcal{M}^{[0]}(\eta))^{-1}$). We say that $g_\ell^{[n]} = g^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\varepsilon)$ is a propagator on scale $[n]$. The matrices $\mathcal{M}^{[m]}(x;\eta)$ will be defined as in Section 5 for $n < n_0$ and will be defined recursively also for $n \ge n_0$ in terms of the self-energy clusters $\mathcal{S}^{\mathcal{R}}_{k,n-1}$ introduced in Definition 4, Section 5, setting for $n > n_0$ (cf. (5.7))

$$\mathcal{M}^{[n]}(x;\eta) = \Big(\prod_{m=0}^{n-1} \chi_m(\Delta^{[m]}(x;\eta))\Big) \sum_{k=2}^{\infty} \sum_{T \in \mathcal{S}^{\mathcal{R}}_{k,n-1}} \mathcal{V}_T(x;\eta), \tag{6.7}$$

where the self-energy values $\mathcal{V}_T(x;\eta)$ are evaluated by means of propagators on scales less than $[n]$. Note that we have already defined (consistently with (6.7))) the matrices $\mathcal{M}^{[\le n]}$ with $n < n_0$ and the propagators on scale $[-1], [0], \dots, [n_0 - 1]$ (so that (6.6) defines also $g^{[n_0]}(x;\eta)$).

Of course the above definition makes sense only if the series (6.7) can be shown to be convergent for all n. For this purpose an inductive assumption on the propagators on the scales $[m]$, $0 \le m < n$ is necessary.

Inductive assumption. *Let n_0 be fixed as in Lemma 3.*
(i) For $0 \le m \le n-1$ the matrices $\mathcal{M}^{[m]}(x;\eta)$ are defined by convergent series for all $|\varepsilon| \in I(\overline{\varepsilon}) =$

$[\overline{\varepsilon}/4,\overline{\varepsilon}]$ and, for all x, they satisfy the properties (i) and (iv) of Lemma 3. Moreover they can be represented as in (5.11), with $P^{[\le n]},Q^{[\le n]},R^{[\le n]}$ as in (5.12) and the entries $\overline{M}^{[\le n]}_{\gamma'\gamma}$ bounded as described after (5.12).
(ii) There exist $K>0$ and open sets $\mathcal{E}^o_m$, $m=0,\ldots,n$, with $\mathcal{E}^o_m\subset I(\overline{\varepsilon})$, such that, defining recursively $\underline{\lambda}^{[m]}(\eta)$ in terms of $\underline{\lambda}^{[m-1]}(\eta)$ for $m=n_0,\ldots,n-1$ by (i) in Definition 7 above, while setting $\underline{\lambda}^{[m]}(\eta)\equiv 0$ for $m=0,\ldots,n_0-1$, and defining $\tau_1\stackrel{def}{=}\tau_0(1+\delta_1)+N$, with $\delta_1>0$, one has for $\varepsilon\notin\mathcal{E}^o_m$

$$\begin{aligned}&\Gamma^{[m]}(x;\eta)\stackrel{def}{=}\Big||x|-\sqrt{\underline{\lambda}^{[m]}(\eta)}\Big|\ge 2^{-\frac12 m}\frac{C_0}{|\boldsymbol{\nu}|^{\tau_1}},\\ &|\mathcal{E}^o_m|\le K2^{-\frac12 m}\overline{\varepsilon}\,\overline{\eta}^{\delta_1(k_0-1/2)},\end{aligned}\tag{6.8}$$

with $\overline{\varepsilon}=\overline{\eta}^{k_0}$, for all $m\le n-1$ and all x.

Remark. As in [GG2], a key point is checking that $\Delta^{[n]}(x;\eta)$ can be used to bound below the denominators of the non-vanishing propagators on scale $[n]$. If x has scale $[n]$, with $n\ge n_0$, one has

$$\begin{aligned}\Big|x^2-\lambda^{[n]}(x;\eta)\Big|&\ge\Big|x^2-\underline{\lambda}^{[n]}(\eta)\Big|-\Big|\underline{\lambda}^{[n]}(\eta)-\lambda^{[n]}(x;\eta)\Big|\\ &\ge\frac12\Big|x^2-\underline{\lambda}^{[n]}(\eta)\Big|+2^{-(n+2)}C_0-\Big|\lambda^{[n]}\big(\sqrt{\underline{\lambda}^{[n-1]}(\eta)},\eta\big)-\lambda^{[n]}(x;\eta)\Big|\\ &\ge\frac12\Big|x^2-\underline{\lambda}^{[n]}(\eta)\Big|\qquad\Rightarrow\qquad\|x^2-\mathcal{M}^{[n]}(x,\eta)\|\ge\frac{1}{8\mu}\,\Delta^{[n]}(x;\eta),\end{aligned}\tag{6.9}$$

having used the lower cut-off $\psi_n(\Delta^{[n]}(x;\eta))$ in the propagator (see (6.5)) to obtain the first two terms in the second line while the upper cut-off $\chi_{n-1}(\Delta^{[n-1]}(x;\eta))$ has been used to obtain positivity of the difference between the second and third terms in the second line, after applying the second inequality in (6.2), so that the last term in the second line of (6.9) can be bounded above proportionally to $\varepsilon 2^{-n}C_0$.

The inequality (6.9) allows us a complete word by word reduction of the proof of the inductive assumption above to the corresponding inductive assumption of [GG2]. The symbols here and in [GG2] have been chosen to coincide so that the analysis in Section 6 and Appendix A3 of [GG2] can be taken over and reinterpreted, with no change, as proofs of the above inductive hypothesis, apart for the check of the measure $\mathcal{E}^0_m$ in (6.8), which in the present case is slightly different from the corresponding computation in [GG2]. The estimate (6.8) of the measure of $\mathcal{E}^0_m$ is explicitly given in Appendix A8.

We can summarize the previous discussion into the following Lemma.

Lemma 6. *There is ε_0 small enough such that for $\overline{\varepsilon}<\varepsilon_0$ and $\varepsilon\in I(\overline{\varepsilon})$, if the inductive hypothesis is assumed for $0\le m\le n-1$ then it holds for $m=n$.*

The series for $X(\boldsymbol{\psi})$ is now fully "renormalized" and its terms are well defined for ε in a set $\mathcal{E}$ whose measure is large (near 0). The series is expressed as a sum of renormalized tree values of tree graphs without any self–energy graph. Therefore the series is convergent (by Siegel's lemma); see the corresponding discussion in Appendix A3 of [GG2].

In the derivation geometric series with ratio $z > 1$ have been considered and the rule $\sum_{k=0}^{\infty} z^k = (1-z)^{-1}$ has been repeatedly used: this is not mathematically rigorous and therefore an a posteriori check must be made that the function $X(\boldsymbol{\psi})$, via (1.5), actually does satisfy the equations (1.4). The check, however, is a repetition of the corresponding (essentially algebraic) check in Appendix A5 of [GG2].

Therefore the proof of Theorem 1 is complete

7. Conclusions.

We conclude by mentioning some problems that seem to us of interest.

(1) First we note that the proof of Theorem 1 also provided some informations about the analyticity properties in η, hence in ε, of the surviving lower-dimensional tori; cf. [GG1] and [GG2] for further details. As in the quoted papers, it can be interesting to further investigate such properties, and related ones, such as Borel summability.

(2) The uniqueness of the resonant tori appears to be a hard problem. Even a proof of Borel summability would not resolve the problem as there could be solutions which are not Borel summable. Just as in the maximal tori case analyticity is not sufficient to guarantee uniqueness. The very recent [BT] does not settle the question, as it proves uniqueness of the C^{∞} diffeomorphism mapping the unperturbed invariant tori which are conserved into those which are explictly constructed, but it dose not eliminate the possibility that other nearby quasi-periodic motions with the same rotation vectors exist.

(3) Another open problem concerns the possibility of removing the assumptions we made in this paper. We have been able to do this in several particular cases but we did not find a satisfactory general formulation: for instance can something be said in general in the case in which the average of the perturbation vanishes identically?

(4) Finally, it would be interesting to understand what happens in the case of n-dimensional tori, with n strictly less than $N-1$ (that is in the case of several normal frequencies), by relaxing the assumptions on the perturbing potential with respect the results existing in literature, as done here and in Cheng's paper [Ch1] and [Ch2] for $n = N-1$. This is a substantially more difficult endeavour with respect to that considered here and in Cheng's quoted papers.

Appendix A1. Heuristic analysis leading to fractional series

Consider the simple case $H(\mathbf{A}, B, \boldsymbol{\alpha}, \beta) = \boldsymbol{\omega} \cdot \mathbf{A} + \frac{1}{2}(\mathbf{A}^2 + B^2) + \varepsilon f(\boldsymbol{\alpha}, \beta)$, and let assumptions (a) and (c) in Section 1 be satisfied by f. Let us consider the generating function

$$\Gamma(\mathbf{A}', B', \boldsymbol{\alpha}, \beta) = \boldsymbol{\alpha} \cdot \mathbf{A}' + \beta B' + \varepsilon \Psi_1(\mathbf{A}', B', \boldsymbol{\alpha}, \beta) + \varepsilon^2 \Psi_2(\mathbf{A}', B', \boldsymbol{\alpha}, \beta) + \varepsilon^2 \boldsymbol{\alpha} \cdot \boldsymbol{\zeta} + \varepsilon^2 \beta \rho, \quad (A1.1)$$

where, if $\Delta \stackrel{def}{=} \boldsymbol{\omega} \cdot \partial_{\boldsymbol{\alpha}}$,

$$\Psi_1(\mathbf{A}', B', \boldsymbol{\alpha}, \beta) \stackrel{def}{=} -\Delta^{-1}\big[1 - (\mathbf{A}' \cdot \partial_{\boldsymbol{\alpha}} + B' \partial_{\beta})\, \Delta^{-1}\big] f_{\neq \mathbf{0}}(\boldsymbol{\alpha}, \beta), \qquad (7.2)$$

and, if $\Phi(\mathbf{A}', B', \boldsymbol{\alpha}, \beta) = \frac{1}{2}\big[(\partial_{\boldsymbol{\alpha}} \Psi_1)^2 + (\partial_{\beta} \Psi_1)^2\big]$,

$$\begin{aligned} &\Psi_2(\mathbf{A}', B', \boldsymbol{\alpha}, \beta) \stackrel{def}{=} -\Delta^{-1}\big[1 - (\mathbf{A}' \cdot \partial_{\boldsymbol{\alpha}} + B' \partial_{\beta})\, \Delta^{-1}\big] \Phi_{\neq \mathbf{0}}(\mathbf{A}', B', \boldsymbol{\alpha}, \beta), \\ &\boldsymbol{\zeta} = -\partial_{\mathbf{A}'} \Phi_{\mathbf{0}}(\mathbf{A}', 0, \beta_0)\big|_{\mathbf{A}'=\mathbf{0}}, \qquad \rho = -\partial_{B'} \Phi_{\mathbf{0}}(\mathbf{0}, B', \beta_0)\big|_{B'=0}. \end{aligned} \qquad (A1.3)$$

The canonical map $(\mathbf{A}, B, \boldsymbol{\alpha}, \beta) \longleftrightarrow (\mathbf{A}', B', \boldsymbol{\alpha}', \beta')$ generated by $\Gamma(\mathbf{A}', B', \boldsymbol{\alpha}, \beta)$ transforms the Hamiltonian $H(\mathbf{A}, B, \boldsymbol{\alpha}, \beta)$ into

$$H'(\mathbf{A}', B', \boldsymbol{\alpha}', \beta') = \boldsymbol{\omega} \cdot \mathbf{A}' + \frac{1}{2}(\mathbf{A}'^2 + B'^2) + \varepsilon f_{\mathbf{0}}(\beta') - \\ - \varepsilon^2 \partial_{\beta'} f_{\mathbf{0}}(\beta') \cdot \Delta^{-2} \partial_{\beta'} f_{\neq \mathbf{0}}(\boldsymbol{\alpha}', \beta') + \varepsilon^2 \Phi_{\mathbf{0}}(\mathbf{0}, 0, \beta') + O(\varepsilon I'^2) + O(\varepsilon^3) \ . \tag{A1.4}$$

Writing the Hamiltonian equations generated by (A1.4), we realize that in a small neighborhood of β_0 the evolution equation for β' takes the form:

$$\ddot{\beta}' = -c\varepsilon(\beta' - \beta_0)^{k_0} - \varepsilon^2 a + \varepsilon^2 a'(\beta - \beta_0) + O\big(\varepsilon(\beta' - \beta_0)^{k_0+1}\big) + O\big(\varepsilon^2(\beta' - \beta_0)^2\big) + O(\varepsilon^3) \ , \tag{A1.5}$$

where $a = \partial_{\beta'} \Phi_{\mathbf{0}}(\mathbf{0}, 0, \beta_0)$ is the same constant defined in (1.9) and a' is a suitable constant.

Therefore, under the assumption that both a and c are non-zero, the perturbed equilibrium of the angle β' is equal, *up to high order corrections*, to $\beta_0' = \beta_0 + \delta\beta_0$, with

$$\delta\beta_0 = \begin{cases} (-a\varepsilon/c)^{1/k_0} & \text{if } k_0 \text{ is odd,} \\ \pm(-a\varepsilon/c)^{1/k_0} & \text{if } k_0 \text{ is even and } a\varepsilon/c < 0. \end{cases} \tag{A1.6}$$

If k_0 is odd and $c\varepsilon > 0$ the approximate perturbed equilibrium point β_0' is quadratically stable, while if $c\varepsilon < 0$ β_0' is quadratically unstable: hence we shall say that the resonant invariant torus is elliptic if $c\varepsilon > 0$ and hyperbolic if $c\varepsilon < 0$. The center of oscillations is displaced by $O(\varepsilon^{1/k_0})$ so that a fractional series in powers of $\eta = \varepsilon^{1/k_0}$ has to be expected (at best).

If k_0 is even the unperturbed equilibrium point β_0 can be continued into a perturbed equilibrium $\beta_0' = \beta_0 + \delta\beta_0$ (with $\delta\beta_0$ vanishing as $\varepsilon \to 0$) *only if* $a\varepsilon c < 0$. In this case, if $c\varepsilon\delta\beta_0 > 0$ the approximate perturbed equilibrium point β_0' is quadratically stable, while if $c\varepsilon\delta\beta_0 < 0$ β_0' it is quadratically unstable: hence we shall say that the resonant invariant torus is elliptic if $c\varepsilon\delta\beta_0 > 0$ and hyperbolic if $c\varepsilon\delta\beta_0 < 0$. Given the results known for elliptic and hyperbolic resonances Theorem 1 becomes a natural conjecture.

If $c \neq 0$ and $a = 0$ the theory depends on the corrections in (A1.5): for instance if $a' \neq 0$ a natural conjecture is that the new equilibrium point for β' is approximately $\beta_0' + \delta\beta_0$, with $\delta\beta_0$ satisfying the equation $(\delta\beta_0)^{k_0-1} = \varepsilon a'/c$, whenever this equation is solvable. The stability of this equilibrium point can be again analyzed in terms of the relative signs of ε, a, c. If $a' = 0$ one expects that the higher order terms have to be studied in details.

Appendix A2. On assumption (a)

Here we want to show that assumption (a) is not a purely technical assumption; that is we want to shows that generically a Hamiltonian of the form (1.3) violating assumption (a) cannot admit quasi-periodic motions of codimension 1 continuously connected to an unperturbed motion of the form (1.6) in the limit $\varepsilon \to 0$. We show this by producing an explicit example, given by an Hamiltonian of the form (1.3), satisfying assumptions (c) and (b) and violating assumption (a), for which we can prove absence of quasi-periodic motions of codimension 1 tending to an unperturbed motion of the form (1.6) as $\varepsilon \to 0$.

The counterexample is given by the following Hamiltonian

$$H = A + \frac{1}{2}A^2 + \frac{1}{2}B^2 + \varepsilon\left(\frac{\sin^3\beta}{3} + \sin\beta\cos\alpha\right), \tag{A2.1}$$

where $A, B \in \mathbb{R}^1$ and $\alpha, \beta \in \mathbb{T}^1$. If $\varepsilon = 0$ the Hamiltonian (A2.1) is strictly convex (then it satisfies assumption (b)) and it admits the unperturbed periodic motions

$$A(t) = 0, \qquad \alpha(t) = t, \qquad B(t) = 0, \qquad \beta(t) = \beta_0, \tag{A2.2}$$

parametrized by the choice of $\beta_0 \in \mathbb{T}^1$. Choosing $\beta_0 = 0$, we see that $f_{\mathbf{0}}(0) = \partial_\beta f_{\mathbf{0}}(0) = \partial_\beta^2 f_{\mathbf{0}}(0) = 0$ and $\partial_\beta^3 f_{\mathbf{0}}(0) = 2 \neq 0$ (so that assumption (c) is satisfied with $k_0 = 2$). Using the fact that f is independent of the action variables and that $f(\alpha, \beta = 0) = \partial_\beta^2 f(\alpha, \beta = 0) \equiv 0$, we see that the quantity a defined in (1.8) is identically 0: this means that assumption (a) is violated in the case under analysis.

We now investigate the possible existence of periodic motions of the form

$$\alpha(t) = t + a(t;\varepsilon), \qquad A(t) = \dot a(t;\varepsilon), \qquad \beta(t) = \eta(\varepsilon) + b(t;\varepsilon), \qquad B(t) = \dot b(t;\varepsilon), \tag{A2.3}$$

where $\eta(\varepsilon)$ is a continuous function of ε such that $\lim_{\varepsilon\to 0}\eta(\varepsilon) = 0$ and $a(t;\varepsilon), b(t;\varepsilon)$ are 0 average periodic functions of t of period 2π. We want to show that it is impossible that a motion of the form (A2.3) satisfies the Hamiltonian equations of motion

$$\begin{aligned} &\ddot a(t;\varepsilon) = \varepsilon \sin(t + a(t;\varepsilon))\sin(\eta + b(t;\varepsilon)) \ , \\ &\ddot b(t;\varepsilon) = -\varepsilon\big[\sin^2(\eta + b(t;\varepsilon))\cos(\eta + b(t;\varepsilon)) + \cos(t + a(t;\varepsilon))\cos(\eta + b(t;\varepsilon))\big]. \end{aligned} \tag{A2.4}$$

In fact a 0 average periodic solution to (A2.4) is necessarily of the form

$$\begin{aligned} &a(t;\varepsilon) = -\varepsilon\big(\eta\sin t + \frac{\varepsilon}{4}\sin t\cos t\big) + \varepsilon^2\, O(\varepsilon^2 + \eta^2), \\ &b(t;\varepsilon) = \varepsilon\cos t + \varepsilon\, O(\varepsilon^2 + \eta^2). \end{aligned} \tag{A2.5}$$

Averaging over t the second of (A2.4) we find

$$\int_0^{2\pi} \frac{\mathrm{d}t}{2\pi}\big(\cos(t + a(t;\varepsilon))\cos(\eta + b(t;\varepsilon)) + \sin^2(\eta + b(t;\varepsilon))\cos(\eta + b(t;\varepsilon))\big) = 0. \tag{A2.6}$$

Now, using the expressions (A2.5), we see that (A2.6) is equal to $\eta^2 + \frac{\varepsilon^2}{2}$, plus terms of order at least $\eta^4 + \varepsilon^4$, and this leads to a contradiction.

Appendix A3. Counting the number of trees lines

We want to show that the number of lines of a tree contributing to $\mathbf{A}^{(k)}$, $B^{(k)}$, $\boldsymbol{\alpha}^{(k)}$ or $b^{(k)}$ can be bounded, above and below, by an $O(1)$ constant times k. A lower bound k/k_0 is an immediate consequence of the definitions.

We proceed by induction, using that:
(1) the trees contributing to $\mathbf{A}^{(k_0)}$ and $B^{(k_0)}$ have one line (see (2.2)); those contributing to $\mathbf{A}^{(k_0+1)}$ and $B^{(k_0+1)}$ have two lines;
(2) the trees contributing to $\mathbf{a}_{\boldsymbol{\nu}}^{(k_0)}$ and $b_{\boldsymbol{\nu}}^{(k_0)}$, with $\boldsymbol{\nu} \neq \mathbf{0}$, have at most two lines and those contributing to $\mathbf{a}_{\boldsymbol{\nu}}^{(k_0+1)}$ and $b_{\boldsymbol{\nu}}^{(k_0+1)}$, with $\boldsymbol{\nu} \neq \mathbf{0}$, have at most three lines;
(3) $b_{\mathbf{0}}^{(1)} \equiv \beta_1$ is represented by a trivial tree (i.e. a tree with one line) and the trees contributing to $b_{\mathbf{0}}^{(2)} \equiv \beta_2$ have at most $k_0 + 2$ lines.

We introduce the following definitions:
$B_k(I)$ is the maximum number of lines of the trees contributing to $\mathbf{A}^{(k)}, B^{(k)}$;
$B_k(\varphi)$ is the maximum number of lines of the trees contributing to $\mathbf{a}_{\boldsymbol{\nu}}^{(k)}, b_{\boldsymbol{\nu}}^{(k)}$, $\boldsymbol{\nu} \neq \mathbf{0}$;
$B_k(b_{\mathbf{0}})$ is the maximum number of lines of the trees contributing to $b_{\mathbf{0}}^{(k)} \equiv \beta_k$.

We shall make the following inductive assumption:

$$\begin{aligned} &B_k(I) \leq 3k_0(k - k_0 + 1) - 4k_0 - 2 && k \geq k_0 + 1, \\ &B_k(\varphi) \leq 3k_0(k - k_0 + 1) - 4k_0 - 1 && k \geq k_0 + 1, \\ &B_{k-k_0+1}(b_{\mathbf{0}}) \leq 3k_0(k - k_0 + 1) - 4k_0 && k \geq k_0 + 1. \end{aligned} \tag{A3.1}$$

By the remarks above and the fact that $k_0 \geq 2$ it follows that at the first step (that is $k = k_0 + 1$) the inequalities above are verified.

Assume inductively the inequalities in (A3.1) for $k_0 + 1 \leq k \leq h - 1$ and let us prove them for $k = h \geq k_0 + 2$. Let us start with the third inequality. We call $\mathfrak{v}_0$ the first node preceding the root. By the rules explained in Section 3 we have that the sum of the orders of the $s = s_{\mathfrak{v}_0}$ subtrees entering $\mathfrak{v}_0$ must be h. So, using the inductive assumptions, and calling (see Figure 3):

s_I the number of subtrees of type $I^{(k)}, k \geq k_0 + 1$, entering $\mathfrak{v}_0$;
s'_I the number of subtrees of type $I^{(k_0)}$ entering $\mathfrak{v}_0$;
s_φ the number of subtrees of type $\varphi_{\boldsymbol{\nu}}^{(k)}, k \geq k_0 + 1, \boldsymbol{\nu} \neq \mathbf{0}$, entering $\mathfrak{v}_0$;
s'_φ the number of subtrees of type $\varphi_{\boldsymbol{\nu}}^{(k_0)}, \nu \neq \mathbf{0}$, entering $\mathfrak{v}_0$;
s_0 the number of subtrees of type $b_{\mathbf{0}}^{(k)}, k \geq 2$, entering $\mathfrak{v}_0$;

s_0' the number of subtrees of type $b_{\mathbf{0}}^{(1)}$ entering $\mathfrak{v}_0$,

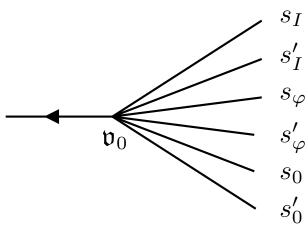


FIGURE 3. The lines entering v_0 represent *symbolically* the bundle of subtrees entering $\mathfrak{v}_0$ and with root lines of types $I^{(k)}, I^{(k_0)}, \varphi_{\boldsymbol{\nu}}^{(k)}, \varphi_{\boldsymbol{\nu}}^{(k_0)}$, with $k \geq k_0+1$, and β_k, β_1, with $k \geq 2$. The number of subtrees in each bundle is indicated by the right labels.

$B_{h-k_0+1}(b_{\mathbf{0}})$ can be bounded by the maximum over the choices of $s_I, s_I', s_\varphi, s_\varphi', s_0, s_0'$ of:

$$1+3k_0h-\Big\{(3k_0^2+k_0+2)s_I+(3k_0^2-1)s_I'+(3k_0^2+k_0+1)s_\varphi+(3k_0^2-2)s_\varphi'+4k_0s_0+(3k_0-1)s_0'\Big\}. \quad (A3.2)$$

Let us first consider the case $s_I+s_I'+s_\varphi+s_\varphi' \geq 1$.
In this case, if $s=1$, since $h \geq k_0+2$, we must have that either s_φ or s_I is $=1$, and, in both cases, we can bound (A3.2) by $1+3k_0h-(3k_0^2+k_0+1)$, that is the desired bound; if, on the contrary, $s \geq 2$, then (A3.2) can be bounded by $1+3k_0h-(3k_0^2-2)-(3k_0-1) \leq 3k_0h-3k_0^2-k_0$.
Let us now consider the case $s_I+s_I'+s_\varphi+s_\varphi'=0$, in which case $s=s_0+s_0'$ and necessarily $s \geq k_0$ (note that in this case the $s+1$ derivatives in (3.6) must be derivatives with respect to β so that $s \geq k_0$). If $s \geq k_0+1$, then (A3.2) can be bounded by $1+3k_0h-(k_0+1)(3k_0-1) \leq 3k_0h-3k_0^2-k_0$. If $s=k_0$, by the rules in Section 3 it must be $s_0' \leq k_0-2$, so that (A3.2) can be bounded by $1+3k_0h-4k_0(k_0-s_0')-(3k_0-1)s_0' \leq 1+3k_0h-4k_0^2+(k_0-2)(k_0+1)$ that implies the desired inductive bound.

Let us now consider the first of the three inequalities in (A3.1). By the rules explained in Section 3 we have that the sum of the degrees of the s subtrees entering $\mathfrak{v}_0$ is $h-k_0$ or h if, respectively, $\delta_{\mathfrak{v}_0}=1,0$. If $\delta_{\mathfrak{v}_0}=1$, $B_h(I)$ can be bounded by the maximum over the choices of $s_I, s_I', s_\varphi, s_\varphi', s_0, s_0'$ of $1+3k_0(h-k_0)-\{\cdots\}$, where the brackets include the same expression as in (A3.2), and we can bound this number by $1+3k_0h-3k_0^2-(3k_0-1)$ that is even better than the desired bound. If $\delta_{\mathfrak{v}_0}=0$, one must have $s=s_I+s_I' \geq 2$ (as $s=1$ corresponds to a not allowed tree graph) and $B_h(I)$ can be bounded by $1+3k_0h-2(3k_0^2-1)$, that is even better than the desired bound.

We finally consider the second of the three inequalities in (A3.1). By the rules explained in Section 3 we have that the sum of the orders of the s subtrees entering $\mathfrak{v}_0$ must be $h-k_0$ or h, if, respectively, $\delta_{\mathfrak{v}_0}=1,0.$. If $\delta_{\mathfrak{v}_0}=1$ the number of lines of the corresponding trees can be bounded in the same way as we proceeded above for the first inequality in (A3.1). If $\delta_{\mathfrak{v}_0}=0$, only lines of type I can enter $\mathfrak{v}_0$, and then the corresponding trees have a number of lines bounded by $1+3k_0h-\{(3k_0^2+k_0+2)s_I+(3k_0^2-1)s_I'\}$. So, if $s_I \geq 1$ the desired bound follows. If $s_I=0$, since $h \geq k_0+2$, we must have $s_I' \geq 2$, and again the bound follows.

Appendix A4. Proof of Lemma 2

To estimate the eigenvalues of $(ix-\mathcal{M}^{[0]}(\eta))E$, we start by rewriting $(ix-\mathcal{M}^{[0]}(\eta))E$ in the form

$$(ix-\mathcal{M}^{[0]}(\eta))E=\begin{pmatrix}-R & -ix+Q\\ ix+Q^{\dagger} & P\end{pmatrix}. \tag{A4.1)}$$

Note that, by the strict convexity of the Hessian of H_0 if ε is small enough, P admits N positive eigenvalues $0<\mu_1^{(0)}\le\cdots\le\mu_N^{(0)}\equiv\mu$. Let $v_\alpha(x;\eta)=\begin{pmatrix}u\\ v\end{pmatrix}$ be an eigenvector of $(ix-\mathcal{M}^{[0]})E$ with eigenvalue α (here u and v are two column vectors of dimension N):

$$\begin{pmatrix}-R & -ix+Q\\ ix+Q^{\dagger} & P\end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}=\begin{pmatrix}-Ru+(-ix+Q)v\\ (ix+Q^{\dagger})u+Pv\end{pmatrix}=\alpha\begin{pmatrix}u\\ v\end{pmatrix}. \tag{A4.2}$$

Now, either $|\alpha|\ge\mu_1^{(0)}/2$ or $\frac{P}{2}<P-\alpha<\frac{3P}{2}$. In the latter case $(P-\alpha)$ is invertible, $2/(3\mu_N^{(0)})\le(P-\alpha)^{-1}\le 2/\mu_1^{(0)}$, and we can rewrite (A4.2) as

$$v=-(P-\alpha)^{-1}(ix+Q^{\dagger})u,\qquad -Ru-(-ix+Q)(P-\alpha)^{-1}(ix+Q^{\dagger})u=\alpha u. \tag{A4.3}$$

The second equation in (A4.3) is of the form

$$\Big[-x^2(P-\alpha)^{-1}+O(\varepsilon x)+O(\varepsilon\eta^{k_0-1})\Big]u=\alpha u, \tag{A4.4}$$

so that $|\alpha|\ge\frac{2}{3\mu}x^2+O(\varepsilon x)+O(\varepsilon\eta^{k_0-1})$. If $x^2\ge\rho|\varepsilon|\eta^{k_0-1}$, for ρ large enough, the latter estimate implies $|\alpha|\ge\frac{1}{2\mu}x^2$, and Lemma 2 is proven.

Appendix A5. Symmetry properties of the self-energy matrices

In this section we discuss the symmetry properties (5.10) of $\mathcal{M}^{[\le n]}(x;\eta)$.

We begin with proving (5.10) for $\mathcal{M}^{[n]}(x;\eta)$. We inductively suppose that the same symmetry properties hold for both $\mathcal{M}^{[p]}(x;\eta)$ and $g^{[p]}(x;\eta)$, for $0\le p<n$,

$$Eg^{[p]}(x;\eta)E=[g^{[p]}(-x;\eta)]^T,\qquad [g^{[p]}(x;\eta)]^*=g^{[p]}(-x;\eta), \tag{A5.1}$$

where the $*$ denotes complex conjugation and T transposition. Note that if $p=0$, $\mathcal{M}^{[0]}(\eta)$ and $g^{[0]}(x;\eta)$ satisfy the desired symmetry properties, as discussed in Section 4.

Consider the representation of $\mathcal{M}^{[n]}(x;\eta)$ given by (5.6) and (5.7): in the following confusion between T denoting transposition and T denoting a self–energy cluster will be avoided by renaming the cluster T with a new symbol D. Note that $F^*(\boldsymbol{\nu}_{\mathfrak{v}})=F(-\boldsymbol{\nu}_{\mathfrak{v}})$, so that, using the second of (A5.1), $[\mathcal{V}_D(x;\eta)]^*$ can be written as[2]

$$\begin{aligned}[\mathcal{V}_D(x;\eta)]^* &=\frac{(\eta\beta_1)^{|L(D)|}}{|\Lambda(D)|!}\Big(\prod_{\mathfrak{v}\in V(D)}\varepsilon^{\delta_{\mathfrak{v}}}\Big)\Big(\prod_{\mathfrak{v}\in V(D)}F^*(\boldsymbol{\nu}_{\mathfrak{v}})\Big)\Big(\prod_{\ell\in\Lambda(D)}[g^{[n_\ell]}(x_\ell;\eta)]^*\Big)= \\ &=\frac{(\eta\beta_1)^{|L(D)|}}{|\Lambda(D)|!}\Big(\prod_{\mathfrak{v}\in V(D)}\varepsilon^{\delta_{\mathfrak{v}}}\Big)\Big(\prod_{\mathfrak{v}\in V(D)}F(-\boldsymbol{\nu}_{\mathfrak{v}})\Big)\Big(\prod_{\ell\in\Lambda(D)}g^{[n_\ell]}(-x_\ell;\eta)\Big)=\mathcal{V}_{D^*}(-x;\varepsilon),\end{aligned} \tag{A5.2}$$

[2] We stress that we use the convention that the internal γ indices of D are summed over, while the momentum labels are not summed over; in this way the value $\mathcal{V}_D$ explicitly depends only on: the momenta of the internal nodes, the external momentum and the two γ labels associated to the incoming and outcoming lines.

where D^* is a cluster topologically equivalent to D, with opposite mode labels (clearly the correspondence $D \longleftrightarrow D^*$ is one–to–one). Summing over the choices of D (see (5.7)) the second of (5.10) follows.

To prove the first of (5.10), it is convenient to shorten notations; see Figure 4 for a pictorial representation of the symbols. Given a self–energy cluster D, let us denote by $\mathfrak{v}$ and $\mathfrak{v}'$ the nodes such that the exiting line of D comes out from $\mathfrak{v}'$ and the entering line of D enters $\mathfrak{v}$.

(1) Let $\mathfrak{v}' \equiv \mathfrak{v}_1, \ldots, \mathfrak{v}_n \equiv \mathfrak{v}$ be the nodes on the path $\mathcal{L} \equiv \mathcal{L}(D)$ joining $\mathfrak{v}$ to $\mathfrak{v}'$, and $\boldsymbol{\nu}_j$ be their mode labels; let also $\ell_j = (\mathfrak{v}_{j+1}\mathfrak{v}_j)$ be the line joining the two successive nodes $\mathfrak{v}_{j+1}$ and $\mathfrak{v}_j$ in $\mathcal{L}$; let γ'_j, γ_j be the component labels at the beginning of the line in $\mathcal{L}$ exiting from $\mathfrak{v}_j$ and at the end of the line in $\mathcal{L}$ entering the node $\mathfrak{v}_j$ (they are, respectively, the lines ℓ_{j-1} and ℓ_j). Recall that, by definition of self–energy cluster, any propagator g_j associated to the lines $\ell_j \in \mathcal{L}$ has a scale $[n_{\ell_j}]$, with $n_{\ell_j} \geq 0$, i.e. there cannot be $\mathbf{0}$-momentum propagators along the path $\mathcal{L}$.

(2) Let $\boldsymbol{\vartheta}_j$ be the (possibly empty) family of subtrees of D with root in $\mathfrak{v}_j$ and with no line in common with $\mathcal{L}$ (not represented in Figure 4: $\boldsymbol{\vartheta}_j$ can be imagined to be a set of trees with the root line ending in $\mathfrak{v}_j$ or, possibly, $\boldsymbol{\vartheta}_j$ may consist just in $\mathfrak{v}_j$).

(3) Let N_j be the $N \times N$ symmetric matrix $\partial_{\gamma'_j\gamma_j} f_{\boldsymbol{\nu}_j}(\beta_0, \boldsymbol{\vartheta}_j)$, where $f_{\boldsymbol{\nu}_j}(\beta_0, \boldsymbol{\vartheta}_j)$ is a function depending only on the set of trees $\boldsymbol{\vartheta}_j$ and on $f_{\boldsymbol{\nu}_j}(\beta)$ (it can be read off (5.6) and (5.7)) and $\partial_{\gamma'}, \partial_\gamma$ must be interpreted as explained after (3.2).

(4) The momentum flowing in a generic line ℓ of the graph D will be the sum of the momentum $\boldsymbol{\nu}^0_\ell$ that would flow on the line if the entering line momentum $\boldsymbol{\nu}$ was $\mathbf{0}$ plus $\boldsymbol{\nu}$ if the line is on the path $\mathcal{L}$; then the propagator matrix g_j associated to the line ℓ_j of $\mathcal{L}$ is equal to $g^{[n_{\ell_j}]}(x^0_j + x; \eta)$, where x^0_j is the scalar product of $\boldsymbol{\omega} \cdot \boldsymbol{\nu}^0_{\ell_j}$ and $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}$.

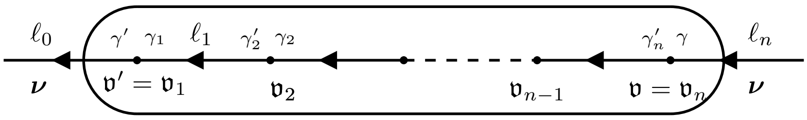


FIGURE 4. A self-energy cluster D and the path $\mathcal{L}$ connecting its external lines. The subtrees internal to D with root on $\mathcal{L}$ are not drawn.

Then the matrix $\mathcal{V}_D(x;\eta)$, see (A5.1), can be concisely written as

$$\big(\mathcal{V}_D(x;\eta)\big)_{\gamma'\gamma} = \Big(EN_1 g_1 EN_2 g_2 \ldots g_{n-2} EN_{n-1} g_{n-1} EN_n\Big)_{\gamma'\gamma}. \tag{A5.3}$$

Interchanging the external lines (i.e. having $\mathfrak{v}'$ as entering node and $\mathfrak{v}$ as exiting node) generates a new self–energy cluster D' in which the momenta flowing in the lines of the subtrees $\boldsymbol{\vartheta}_j$ with roots on the nodes $\mathfrak{v}_j \in \mathcal{L}$ are unchanged while the momentum on the line $\ell_j \in \mathcal{L}$ changes from $\boldsymbol{\nu}^0_{\ell_j} + \boldsymbol{\nu}$, with $\boldsymbol{\nu}^0_{\ell_j}$ equal to the momentum which would flow on ℓ_j if the external momentum $\boldsymbol{\nu}$ was set equal to $\mathbf{0}$, to a new value $-\boldsymbol{\nu}^0_{\ell_j} + \boldsymbol{\nu}$.

The matrix $\mathcal{V}_{D'}(x;\eta)$ can therefore be written

$$\big(\mathcal{V}_{D'}(x;\eta)\big)_{\gamma'\gamma} = \big(EN_n g'_{n-1}\, EN_{n-1} g'_{n-2} \cdots g'_2\, EN_2\, g'_1\, E\, N_1\big)_{\gamma'\gamma}, \tag{A5.4}$$

where $g'_j = g^{[n_{\ell_j}]}(-x^0_{\ell_j} + x; \eta)$. Inserting $\mathbb{1} \equiv E^T E$ after $N_n, \ldots, N_2$ and using the symmetry of $N_1, \ldots, N_n$, the inductive validity of the first of (A5.1) and the transposition rules for matrix products, we get

$$\left(\mathcal{V}^T_{D'}(x;\eta)\right)_{\gamma'\gamma} = \Big(N_1 g''_1 E N_2 g''_2 \ldots g''_{n-2} E N_{n-1} g''_{n-1} E N_n E^T\Big)_{\gamma'\gamma}, \tag{A5.5}$$

where $g''_j \equiv g^{[n_{\ell_j}]}(x^0_{\ell_j} - x;\eta)$. Hence $E\mathcal{V}^T_{D'}(x;\eta)E \equiv \mathcal{V}_D(-x;\eta)$ completing the inductive proof of (A5.1), because $\mathcal{M}^{[n]}(x;\eta)$ is defined as a sum over D of the values $\mathcal{V}_D$ (see (5.7)).

Remark. From (5.10) and (A5.1) follows that $g^{[\geq n]}(x;\eta)$ satisfies the same symmetry properties as $\mathcal{M}^{[n]}(x;\eta)$:

$$E\, g^{[\geq n]}(x;\eta)\, E = [g^{[\geq n]}(-x;\eta)]^T, \qquad [g^{[\geq n]}(x;\eta)]^* = g^{[\geq n]}(-x;\eta). \tag{A5.6}$$

Appendix A6. Cancellations

In this section we discuss the cancellations needed to show that $\mathcal{M}^{[\leq n]}(x;\eta)$ has the structure described in (5.11) and (5.12). In particular we want to show that the elements $\left(\mathcal{M}^{[n]}(x;\eta)\right)_{\gamma'\gamma}$ with either $\gamma' = A_i$ or $\gamma = \alpha_j$ are proportional to x, while the elements with $\gamma' = A_i$ and $\gamma = \alpha_j$ are proportional to x^2, for $x^2 \leq \rho\eta^{2k_0-1}$.

As in [GG2], the contributions to $\mathcal{M}^{[n]}(x;\eta)$ coming from clusters such that

$$\sum_{\mathfrak{v}\in V(T)} |\boldsymbol{\nu}_\mathfrak{v}| \geq (C_0/2^6|x|)^{1/\tau_0} \tag{A6.1}$$

require no cancellations: in fact the exponential decay as $\boldsymbol{\nu}_\mathfrak{v} \to \infty$ of the node functions implies that the contributions coming from clusters satisfying (A6.1) are smaller than Cx^2, for some constant $C > 0$.

When (A6.1) does not hold we need to exploit cancellations, because a priori there is no reason for some of the entries of $\mathcal{M}^{[n]}(x;\eta)$ to be proportional to x or x^2 for small x. Following the same strategy used in Appendix A2 of [GG2] (and of Appendix A4 of [GG1]), the necessary cancellations can be checked to occurr by collecting clusters violating condition (A6.1) into families. Such cancellations are due to the same mechanism pointed out first in [G1], [G2] and [CF]. The following analysis follows [GM].

Given a self–energy cluster D on scale $[n-1]$, let us call D_0 the connected subset of D containing no line on scale $[-1]$ and containing the nodes $\mathfrak{v}$ and $\mathfrak{v}'$ to which the entering and exiting lines are attached. By definition $\sum_{\mathfrak{v}\in D_0} \boldsymbol{\nu}_\mathfrak{v} = \mathbf{0}$. Note also that, if the path $\mathcal{L}$ connecting $\mathfrak{v}$ and $\mathfrak{v}'$ is non-trivial (i.e. $\mathfrak{v} \neq \mathfrak{v}'$), $\mathcal{L}$ is completely contained into D_0; moreover, if D does not contain lines on scale $[-1]$, then $D_0 \equiv D$. We define the family $\mathcal{F} \equiv \mathcal{F}_D$ as the set of self–energy clusters obtained from D by shifting the entering and exiting lines by reattaching them to the nodes of D_0 in all possible ways. We then consider the sum

$$\mathcal{V}_\mathcal{F}(x;\eta) = \sum_{D\in\mathcal{F}} \mathcal{V}_D(x;\eta), \tag{A6.2}$$

contributing to the r.h.s. of (5.7). When summing $\mathcal{V}_\mathcal{F}(x;\eta)$ over all distinct families of clusters on scale $[n-1]$, we recover the quantity $M^{[n]}(x;\eta)$ defined in (5.7). To prove that $\mathcal{M}^{[\leq n]}$ has the structure described by (5.11) and (5.12), we proceed in the following way. We first note that, by the same analysis of Section 6 of [GG2], $\mathcal{V}_{\mathcal{F}}(x;\eta)$ is differentiable with respect to x in the sense of Whitney; then, we show that $\big(\mathcal{V}_{\mathcal{F}}(0;\eta)\big)_{\gamma'\gamma}=0$ if $\gamma'=A_i$ or $\gamma=\alpha_j$ and $\big(\partial_x\mathcal{V}_{\mathcal{F}}(0;\eta)\big)_{\gamma'\gamma}=0$ if $\gamma'=A_i$ and $\gamma=\alpha_j$.

Let us begin with proving that $\big(\mathcal{V}_{\mathcal{F}}(0;\eta)\big)_{\gamma'\gamma}=0$ if $\gamma'=A_i$ or $\gamma=\alpha_j$, and to be definite let us assume that $\gamma=\alpha_j$ for some $j=1,\ldots,N-1$. Recall that, employing the notation of Appendix A5, we can write each matrix $\mathcal{V}_D(x;\eta)$ appearing in the sum (A6.2) in the form (A5.3). By the explicit expression of (A5.3) and of the matrices N_j, we see that, when computing $\mathcal{V}_D(x;\eta)$ in $x=0$, we find $(\mathcal{V}_D(0;\eta))_{\gamma'\alpha_j}=O(\mathcal{F},\mathfrak{v}',\eta)_{\gamma'}\,(\boldsymbol{\nu}_{\mathfrak{v}})_j$ for some function $O(\mathcal{F},\mathfrak{v}',\eta)_{\gamma'}$ *independent* of the choice of the node $\mathfrak{v}$ to which the entering line with label $\gamma=\alpha_j$ is attached to.

If we change D within $\mathcal{F}$ by attaching the entering line to all possible nodes in D_0 (keeping $\mathfrak{v}'$ fixed), we see that the global contribution from such self–energy clusters is equal to $O(\mathcal{F},\mathfrak{v}',\eta)_{\gamma'}\cdot$ $\cdot\sum_{\mathfrak{v}\in D_0}(\boldsymbol{\nu}_{\mathfrak{v}})_j$, which is 0 by the property $\sum_{\mathfrak{v}\in D_0}(\boldsymbol{\nu}_{\mathfrak{v}})_j=\mathbf{0}$. So, if $\gamma=\alpha_j$, the proof of the fact that $\big(\mathcal{V}_{\mathcal{F}}(0;\eta)\big)_{\gamma'\gamma}=0$ is complete. In the case that $\gamma'=A_i$ the proof is completely analogous and we do not repeat it here.

Let us now turn to the proof of the fact that, if $\gamma'=A_i$ and $\gamma=\alpha_j$, then $\big(\partial_x\mathcal{V}_{\mathcal{F}}(0;\eta)\big)_{\gamma'\gamma}=0$. First note that, by the explicit form (A5.3) of the value $\mathcal{V}_D(x;\eta)$ and in particular by the fact that $\mathcal{V}_D(x;\eta)$ depends on x only through the propagators along the path $\mathcal{L}$, we find that, when differentiating $\mathcal{V}_D(x;\eta)$ with respect to x, the effect of the derivative is as follows:

$$\partial_x\mathcal{V}_D(x;\eta)=\sum_{i=1}^{n-1}\Big(EN_1g_1\cdots g_{i-1}EN_i\Big)\partial_x g_i\Big(EN_{i+1}g_{i+1}\cdots g_{n-1}EN_n\Big). \tag{A6.3}$$

Given the line $\ell_i\in D_0$, we call D_1 and D_2 the two connected distinct subsets of D_0 obtained from D_0 by detaching the line ℓ_i and such that D_1 contains the node $\mathfrak{v}'$ to which the line exiting from D is attached, while D_2 contains the node $\mathfrak{v}$ to which the line entering D is attached. D_1 and D_2 are two subgraphs of D with two external lines, one coinciding with one of the external lines of D, the other coinciding with ℓ_i, see Figure 5.

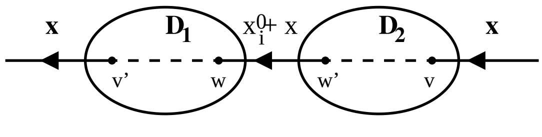


FIGURE 5. Subdiagrams D_1 and D_2 in a self-energy cluster D. The drawn lines connected to $\mathfrak{v}$ and $\mathfrak{v}'$ are the external lines of D.

If we define the values of D_1,D_2 by formulae analogous to (A5.3), we can rewrite each term under the sum in (A6.3) as $\mathcal{V}_{D_1}(x;\eta)\partial_x g^{[n_{\ell_i}]}(x_i^0+x;\eta)\mathcal{V}_{D_2}(x;\eta)$, where, with the same notations of Appendix A5, we defined $x_i^0=\boldsymbol{\omega}\cdot\boldsymbol{\nu}^0_{\ell_i}$, with $\boldsymbol{\nu}^0_{\ell_i}$ equal to the momentum which would flow on ℓ_i if the external momentum $\boldsymbol{\nu}$ was set equal to $\mathbf{0}$.

In particular let us consider the value $(\mathcal{V}_{D_1}(0;\eta)\partial_x g^{[n_{\ell_i}]}(x_i^0;\eta)\mathcal{V}_{D_2}(0;\eta))_{A_i\alpha_j}$ at $x=0$, that is one of the contributions we are interested in, and again note that by the very definition of $\mathcal{V}_{D_1}$ and $\mathcal{V}_{D_2}$, it holds that $\big(\mathcal{V}_{D_1}(0;\eta)\big)_{A_i\gamma_1}=(\boldsymbol{\nu}_{\mathfrak{v}'})_i\Omega(\mathfrak{w})_{\gamma_1}$ and $\big(\mathcal{V}_{D_2}(0;\eta)\big)_{\gamma_2\alpha_j}=\Omega(\mathfrak{w}')_{\gamma_2}(\boldsymbol{\nu}_{\mathfrak{v}})_j$, where

$\mathfrak{w},\mathfrak{w}'$ are the two nodes where ℓ_i enters into and exists from, see Figure 5, and $\Omega(\mathfrak{w})_{\gamma_1}$, $\Omega(\mathfrak{w}')_{\gamma_2}$ are two functions depending on the structure of D_1 and D_2, *but not on the choices of* $\mathfrak{v}',\mathfrak{v}$ within D_1, D_2.

This implies that if we change D within $\mathcal{F}$ by attaching the entering line to all possible nodes in $D_2 \cap D_0$ and by attaching the exiting line to all possible nodes in $D_1 \cap D_0$, keeping ℓ_i fixed, the sum of the contributions of the form $(\mathcal{V}_{D_1}(0;\eta)\partial_x g^{[n_{\ell_i}]}(x_i^0;\eta)\mathcal{V}_{D_2}(0;\eta))_{A_i\alpha_j}$ from this class of graphs is equal to

$$\Big(\sum_{\mathfrak{v}'\in D_1\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}'})_i\Big)\Omega(\mathfrak{w})_{\gamma_1}(\partial_x g^{[n_{\ell_i}]}(x_i^0;\eta))_{\gamma_1\gamma_2}\Omega(\mathfrak{w}')_{\gamma_2}\Big(\sum_{\mathfrak{v}\in D_2\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}})_j\Big). \tag{A6.4}$$

Note that, given any graph D with the structure in Figure 5, the graph D' in which the entering and exiting lines are interchanged (so that the external lines enter in $\mathfrak{v}'$ and exit from $\mathfrak{v}$, see Figure 6) the momentum through the line ℓ_i changes from $\boldsymbol{\nu}_{\ell_i^0}+\boldsymbol{\nu}$ to $-\boldsymbol{\nu}_{\ell_i^0}+\boldsymbol{\nu}$.

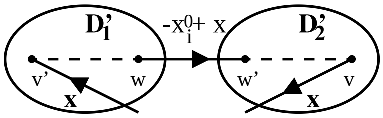


FIGURE 6. The self-energy cluster D' obtained from D by interchanging the external lines of D: the exiting line is attached to a node $\mathfrak{v}\in D_2'$ and the entering line is attached to a node $\mathfrak{v}\in D_1'$.

At $x=0$, if we change D into D', the value $\mathcal{V}_{D_1}(0;\eta)\partial_x g^{[n_{\ell_i}]}(x_i^0;\eta)\mathcal{V}_{D_2}(0;\eta)$ is changed into $\mathcal{V}_{D_2'}(0;\eta)\ \partial_x g^{[n_{\ell_i}]}(-x_i^0;\eta)\mathcal{V}_{D_1'}(0;\eta)$. In analogy with the results of Appendix A5, employing the symmetry relations (A5.1), we find $\mathcal{V}_{D_2'}(0;\eta)=E[\mathcal{V}_{D_2}(0;\eta)]^TE$ and $\mathcal{V}_{D_1'}(0;\eta)=E[\mathcal{V}_{D_1}(0;\eta)]^TE$. Using also the symmetry $E\partial_x g^{[n_{\ell_i}]}(-x_i^0;\eta)E=-\partial_x[g^{[n_{\ell_i}]}]^T(x_i^0;\eta)$, we find that the element (A_i,α_j) of the contribution corresponding to the graph in Figure 6 can be written as

$$\begin{aligned}&\Big(\mathcal{V}_{D_2'}(0;\eta)\partial_x g^{[n_{\ell_i}]}(-x_i^0;\eta)\mathcal{V}_{D_1'}(0;\eta)\Big)_{A_i\alpha_j}=-\Big(E[\mathcal{V}_{D_2}(0;\eta)]^T\partial_x[g^{[n_{\ell_i}]}]^T(x_i^0;\eta)\mathcal{V}_{D_1}^T(0;\eta)E\Big)_{A_i\alpha_j}=\\&=-\Big(E\mathcal{V}_{D_1}(0;\eta)\partial_x g^{[n_{\ell_i}]}(x_i^0;\eta)\mathcal{V}_{D_2}(0;\eta)E\Big)_{\alpha_jA_i}=-\Big(\mathcal{V}_{D_1}(0;\eta)\partial_x g^{[n_{\ell_i}]}(x_i^0;\eta)\mathcal{V}_{D_2}(0;\eta)\Big)_{A_j\alpha_i}.\end{aligned} \tag{A6.5}$$

Repeating the discussion leading to (A6.4) we see that, if we sum (A6.5) over the graphs obtained by attaching the entering line to all possible nodes $\mathfrak{v}'\in D_1'\cap D_0$ and attaching the exiting line to all possible nodes $\mathfrak{v}\in D_2'\cap D_0$, we get

$$-\Big(\sum_{\mathfrak{v}'\in D_1'\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}'})_j\Big)\Omega(\mathfrak{w})_{\gamma_1}(\partial_x g^{[n_{\ell_i}]}(x_i^0;\eta))_{\gamma_1\gamma_2}\Omega(\mathfrak{w}')_{\gamma_2}\Big(\sum_{\mathfrak{v}\in D_2'\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}})_i\Big). \tag{A6.6}$$

Now, using the fact that $\sum_{\mathfrak{v}\in D_0}\boldsymbol{\nu}_{\mathfrak{v}}=\mathbf{0}$ and that D_1', D_2' and D_1, D_2 are topologically equivalent, we also find that

$$\sum_{\mathfrak{v}'\in D_1'\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}'})_j=-\sum_{\mathfrak{v}\in D_2\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}})_j,\qquad \sum_{\mathfrak{v}\in D_2'\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}})_i=-\sum_{\mathfrak{v}'\in D_1\cap D_0}(\boldsymbol{\nu}_{\mathfrak{v}'})_i, \tag{A6.7}$$

so that (A6.4) and (A6.6) sum up to 0.

This completes the proof of the fact that

$$\sum_{D\in\mathcal{F}} \mathcal{V}_D(\varepsilon;0)_{A_i\alpha_j}=0, \qquad \partial_x \sum_{D\in\mathcal{F}} \mathcal{V}_D(\varepsilon;x)_{A_i\alpha_j}\big|_{x=0}=0. \tag{A6.8}$$

Remark. It can be checked that the symmetry (A5.1) "only" implies that, if we define $C_{ij}(x)\stackrel{def}{=}\sum_{D\in\mathcal{F}}\mathcal{V}_D(\varepsilon;x)_{A_i\alpha_j}$, it holds $C_{ij}(x)=C_{ji}(-x)=C^*_{ji}(x)$: therefore the cancellation to second order (in x), expressed by (A6.7), *is not a parity cancellation* (unless $i=j$ where by the self-adjointness of the matrix C it follows that $C_{ii}(x)$ is real and, hence, even in x). If the coefficients $f_{\boldsymbol{\nu}}(\beta)$ are assumed real then the self-adjointness property (A5.1) implies that $C_{ij}(x)$ is even in x and the second order cancellation at $x=0$ is an obvious parity cancellation and the proof above is greatly simplified, as remarked in the simple cases considered in [G1] and [G2].

Having proved that $\big(\mathcal{M}^{[n]}(0;\eta)\big)_{\gamma'\gamma}=0$ if either $\gamma'=A_i$ or $\gamma=\alpha_j$ and that $\big(\partial_x\mathcal{M}^{[n]}(0;\eta)\big)_{\gamma'\gamma}=0$ if $\gamma'=A_i$ and $\gamma=\alpha_j$, it has still to be shown that the elements $\big(\mathcal{M}^{[n]}(x;\eta)\big)_{\gamma'\gamma}$ (with the suitable choices of γ' and γ) satisfy appropriate bounds once the factors x determining the order of zero at $x=0$ are extracted. From convergence one expects that the bounds on the rests of the Taylor series in x around $x=0$ should still be proportional to ε^2.

This is in fact true, *under the condition that (A6.1) does not hold*: if (A6.1) does not hold, then changing the nodes where the external lines are attached to *does not change the scale* of the internal lines. And this implies that the bounds on the derivatives with respect to x are qualitatively the same as the bounds on $\mathcal{M}^{[n]}(x;\eta)$ and boundedness of the entries of $\overline{M}^{[\le n]}$ follows, see (5.12).

Appendix A7. Bounds on the propagator

In this section we want to get a lower bound for the norm of $(ix-\mathcal{M}^{[n]}(x;\eta))E$, by an approximate computation of its eigenvalue which is lowest in absolute value, for $x^2\le\rho\eta^{2k_0-1}$, with ρ the same constant appearing in Lemma 3. Moreover we want to prove dimensional bounds for the derivatives in x and ε of the approximate lowest eigenvalue.

Write $(ix-\mathcal{M}^{[\le n]}(x;\eta))E$ as $(ix-\mathcal{M}^{[\le n]}(x;\eta))E=\begin{pmatrix}-R^{[\le n]}(x;\eta) & -ix+Q^{[\le n]}(x;\eta)\\ ix+Q^{[\le n]\dagger}(x;\eta) & P^{[\le n]}(x;\eta)\end{pmatrix}$, where $P^{[\le n]},Q^{[\le n]},R^{[\le n]}$ are the matrices defined in (5.12).

The N eigenvalues of $P^{[\le n]}(x;\eta)$ are of order 1 and positive; call them $\mu_1^{(n)}\le\ldots\le\mu_N^{(n)}$. Since $P^{[\le n]}(x;\eta)$ differs from $P^{[0]}(x;\eta)$ by $O(\varepsilon^2)$, the eigenvalues $\mu_i^{(n)}$ can be written as $\mu_i^{(n)}=\mu_i^{(0)}(1+O(\varepsilon^2))$.

To estimate the smallest eigenvalue of $(ix-\mathcal{M}^{[\le n]}(x;\eta))E$ we can follow a strategy adapted from the proof of Lemma 2. Let $v_\alpha(x;\eta)=\begin{pmatrix}u\\ v\end{pmatrix}$ be an eigenvector of $(ix-\mathcal{M}^{[\le n]}(x;\eta))E$ with eigenvalue α (here u and v are two column vectors of dimension N):

$$\begin{aligned}&\begin{pmatrix}-R^{[\le n]}(x;\eta) & -ix+Q^{[\le n]}(x;\eta)\\ ix+Q^{[\le n]\dagger}(x;\eta) & P^{[\le n]}(x;\eta)\end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}=\\ &\quad=\begin{pmatrix}-R^{[\le n]}u+(-ix+Q^{[\le n]})v\\ (ix+Q^{[\le n]\dagger})u+P^{[\le n]}v\end{pmatrix}=\alpha\begin{pmatrix}u\\ v\end{pmatrix}.\end{aligned} \tag{A7.1}$$

To estimate the eigenvalue α with smallest absolute value, we restrict attention to the case $P^{[\le n]} - \alpha \ge P^{[\le n]}/2$, as in the opposite case $|\alpha| > \mu_1^{(n)}/2$. Then $(P^{[\le n]} - \alpha)$ is invertible, $(P^{[\le n]} - \alpha)^{-1} = [P^{[\le n]}]^{-1} + O(\alpha)$ and we can rewrite (A7.1) as

$$\begin{aligned} & v = -(P^{[\le n]} - \alpha)^{-1}(ix + Q^{[\le n]\dagger})u, \\ & - R^{[\le n]} - (-ix + Q^{[\le n]})(P^{[\le n]} - \alpha)^{-1}(ix + Q^{[\le n]\dagger})u = \alpha u. \end{aligned} \tag{A7.2}$$

From the second of (A7.2) we see that α must solve an equation of the form

$$\det\Big[N^{[\le n]} - \alpha\big(1 + O(x) + O(\varepsilon)\big)\Big] = 0, \tag{A7.3}$$

where we defined $N^{[\le n]} \stackrel{def}{=} -R^{[\le n]} - (-ix + Q^{[\le n]})[P^{[\le n]}]^{-1}(ix + Q^{[\le n]\dagger})$ and we used that $Q^{[\le n]} = O(\varepsilon)$. This means that α is an eigenvalue of a Hermitian matrix of the form $N^{[\le n]}\big(1 + O(x) + O(\varepsilon)\big)$, so that, calling $\lambda_i^{[n]}$ the eigenvalues of $N^{[\le n]}$, we have (see [Ka])

$$\alpha = \lambda_i^{[n]}\big(1 + O(x) + O(\varepsilon)\big), \tag{A7.4}$$

for some $i = 1, \ldots, N$. The conclusion of the previous discussion is that the problem of computing the smallest eigenvalue of $(ix - \mathcal{M}^{[n]}(x;\eta))E$ is essentially equivalent to the problem of computing the smallest eigenvalue of $N^{[\le n]}$.

An explicit computation of $N^{[\le n]}$ shows that it can be written in the form

$$\begin{aligned} N^{[\le n]} &= -x^2\mathcal{H}_0^{-1} + \big\{ - R^{[\le n]} + x^2\mathcal{H}_0^{-1} - (-ix + Q^{[\le n]})[P^{[\le n]}]^{-1}(ix + Q^{[\le n]\dagger})\big\} = \\ &= -x^2\mathcal{H}_0^{-1} + \begin{pmatrix} 0 & i\varepsilon x D(\eta) \\ -i\varepsilon x D^\dagger(\eta) & -\varepsilon\eta^{k_0-1}\widetilde{M}_{\beta\beta}(\eta) \end{pmatrix} + O(\varepsilon x^2), \end{aligned} \quad . \tag{A7.5}$$

where $\mathcal{H}_0 \stackrel{def}{=} \begin{pmatrix} \partial^2_{\mathbf{A}\mathbf{A}}H_0 & \partial^2_{\mathbf{A}B}H_0 \\ \partial^2_{B\mathbf{A}}H_0 & \partial^2_{BB}H_0 \end{pmatrix}$, $\widetilde{M}_{\beta\beta}(\eta) = -ck_0\beta_1^{k_0-1}(1 + O(\eta))$ and $D(\eta) = D_0 + O(\eta)$, with

$$\begin{pmatrix} D_0 \\ 0 \end{pmatrix} \stackrel{def}{=} -\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\mathcal{H}_0^{-1}\begin{pmatrix} \partial^2_{\mathbf{A}\beta}f_{\mathbf{0}}(\beta_0) \\ \partial^2_{B\beta}f_{\mathbf{0}}(\beta_0) \end{pmatrix} \tag{A7.6}$$

The off–diagonal $O(\varepsilon x)$ elements can be eliminated through a small rotation of $N^{[\le n]}$. In fact, defining the Hermitian matrix Z as:

$$Z = \begin{pmatrix} 0 & [\widetilde{M}_{\beta\beta}(\eta)]^{-1}D(\eta) \\ [\widetilde{M}_{\beta\beta}(\eta)]^{-1}D(\eta) & 0 \end{pmatrix}, \tag{A7.7}$$

a computation shows that

$$\begin{aligned} e^{ix\eta^{-k_0+1}Z}N^{[\le n]}e^{-ix\eta^{-k_0+1}Z} &\equiv -x^2\mathcal{H}_0^{-1} + \delta N^{[\le n]} = \\ &= -x^2\mathcal{H}_0^{-1} + \begin{pmatrix} 0 & 0 \\ 0 & -\varepsilon\eta^{k_0-1}\widetilde{M}_{\beta\beta}(\eta) \end{pmatrix} + O(\eta^{1/2}x^2), \end{aligned} \tag{A7.8}$$

where we used the property $x\eta^{-k_0+1} \leq (\rho\eta)^{1/2}$, if $x^2 \leq \rho\varepsilon\eta^{k_0-1}$. Multiplying the r.h.s. of (A7.8) times $\mathcal{H}_0^{1/2}$ both from the left and from the right (remark that $\mathcal{H}_0^{1/2}$ is a well–defined positive $O(1)$ operator) we see that the norm of $N^{[\leq n]}$ can be bounded above and below by an $O(1)$ constant times the norm of

$$-x^2 + \mathcal{H}_0^{1/2}\delta N^{[\leq n]}\mathcal{H}_0^{1/2}. \tag{A7.9}$$

Finally, using the explicit structure of $\delta N^{[\leq n]}$, see (A7.8), $N-1$ of the eigenvalues of (A7.9) are equal to $-x^2\big(1+O(\eta^{1/2})\big)$, while the last one is $-x^2\big(1+O(\eta^{1/2})\big) - \varepsilon\eta^{k_0-1}\widetilde{M}_{\beta\beta}(\varepsilon)\partial_B^2 H_0$.

From this computation, the bound (6.1) of Lemma 5 follows. To address the differentiability of $\lambda^{[n]}(x;\eta)$ with respect to ε and x, we note that $\mathcal{H}_0^{1/2}\delta N^{[\leq n]}\mathcal{H}_0^{1/2}$ is C^∞ in the sense of Whitney (because such is $\mathcal{M}^{[\leq n]}$, see the corresponding discussion in Appendix A3 of [GG2]) and its derivatives admit the following dimensional bounds:

$$\|\,\partial_\varepsilon\big(\mathcal{H}_0^{1/2}\delta N^{[\leq n]}\mathcal{H}_0^{1/2}\big)\| \leq C\eta^{k_0-1}, \qquad \|\,\partial_x\big(\mathcal{H}_0^{1/2}\delta N^{[\leq n]}\mathcal{H}_0^{1/2}\big)\| \leq C\eta^{k_0}. \tag{A7.10}$$

Given this and the fact that the last eigenvalue of $\mathcal{H}_0^{1/2}\delta N^{[\leq n]}\mathcal{H}_0^{1/2}$ is isolated (it is non degenerate and its distance from the others is $O(\varepsilon\eta^{k_0-1})$), we can represent $\lambda^{[n]}(x;\eta)$ in the form:

$$\lambda^{[n]}(x;\eta) = -\eta^{2k_0-1}\mathrm{Tr}\,\Big(\frac{1}{2\pi i}\oint_\gamma \frac{z\,\mathrm{d}z}{z-\eta^{1-2k_0}\mathcal{H}_0^{1/2}\delta N^{[\leq n]}\mathcal{H}_0^{1/2}}\Big), \tag{A7.11}$$

where z is a complex variable *independent of* ε, x and γ is a circle in the complex plane around $\widetilde{M}_{\beta\beta}(\eta)\partial_B^2 H_0$ not sorrounding the origin, with ε–independent radius. The derivatives of $\lambda^{[n]}(x;\eta)$ with respect to ε and x can be computed differentiating the r.h.s. of (A7.11) and, using the dimensional bounds (A7.10), we find the same dimensional bounds for the derivatives of $\lambda^{[n]}(x;\eta)$:

$$|\partial_\varepsilon\lambda^{[n]}(x;\eta)| \leq C\eta^{k_0-1}, \qquad |\partial_x\lambda^{[n]}(x;\eta)| \leq C\eta^{k_0}. \tag{A7.12}$$

The proof of Lemma 5 is concluded.

Appendix A8. Excluded values of the perturbation parameter

In this section we want to check that, by imposing the first condition in (6.8), the measure of excluded values of ε can be bounded by the second of (6.8). To estimate the measure of the excluded ε's we proceed as in Appendix A2 of [GG2]. We present a proof valid for any $n \geq n_0$. The dimensional bounds (6.2) and the property $|\lambda^{[n]}(\eta) - \lambda^{[n-1]}(\eta)| \leq Ce^{-\kappa_1 2^{n/\tau_1}}$ (see Lemma 5) implies the following dimensional bound for the derivative of $\sqrt{\lambda^{[n]}(\eta)}$:

$$|\partial_\varepsilon\sqrt{\lambda^{[n]}(\eta)}| \leq C'\eta^{-1/2}. \tag{A8.1}$$

Then the first condition in (6.8) excludes, for each $\boldsymbol{\nu}$, a subinterval of $I(\overline{\varepsilon})$ whose measure is bounded by

$$\frac{C_0 2^{-m/2}\eta^{1/2}}{C'|\boldsymbol{\nu}|^{\tau_1}} \tag{A8.2}$$

The Diophantine condition on $\boldsymbol{\omega}$ implies that if the first condition in (6.8) is invalid then $|\boldsymbol{\nu}|$ cannot be too small:

$$\frac{C_0 2^{-m/2}}{|\boldsymbol{\nu}|^{\tau_1}} + \sqrt{\lambda^{[n]}(\eta)} \ge |x| \ge C_0 |\boldsymbol{\nu}|^{-\tau_0} \; . \tag{A8.3}$$

Therefore $\frac{C_0}{2|\boldsymbol{\nu}|^{\tau_0}} \ge C'' \eta^{k_0 - 1/2}$, hence in this case we only have to consider the values of $\boldsymbol{\nu}$ such that $|\boldsymbol{\nu}|^{\tau_0} \ge (C_0/2C'')\eta^{\frac{1}{2}-k_0}$. Summing (A8.2) over the $\boldsymbol{\nu}$'s satisfying this constraint, and using the definition $\tau_1 = \tau_0(1+\delta_1) + N$ we find that the total measure of excluded ε's can be bounded by

$$\frac{C_0 2^{-m/2}\eta^{1/2}}{C'} \Big(\frac{2C''\eta^{k_0-1/2}}{C_0}\Big)^{1+\delta_1} \sum_{|\boldsymbol{\nu}|\ne 0} \frac{1}{|\boldsymbol{\nu}|^N} \equiv K C_0 2^{-m/2} |\varepsilon| \eta^{\delta_1(k_0-1/2)}, \tag{A8.4}$$

which is in fact the second inequality in (6.8).

References

[BT] H. Broer, F. Takens, *Unicity of KAM tori*, Preprint, Groningen, 2005.

[Ch1] Ch.-Q. Cheng, *Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems*, Communications in Mathematical Physics **177** (1996), no. 3, 529–559.

[Ch2] Ch.-Q. Cheng, *Lower-dimensional invariant tori in the regions of instability for nearly integrable Hamiltonian systems*, Communication in Mathematical Physics **203** (1999), no. 2, 385-419.

[CF] L. Chierchia, C. Falcolini, *A direct proof of a theorem by Kolmogorov in Hamiltonian systems*, Annali della Scuola Normale Superiore di Pisa Classe Scienze (4) **21** (1994), no. 4, 541–593.

[E1] L.H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Annali della Scuola Normale Superiore di Pisa Classe Scienze (4) **15** (1988), no. 1, 115–147.

[E2] L.H. Eliasson, *Absolutely Convergent Series Expansions for Quasi Periodic Motions*, Mathematical Physics Electronic Journal **2** (1996), no. 4, 1–33.

[H] M. Herman, *Some open problems in dynamical systems*, Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998), Documenta Mathematica 1998, Extra Vol. II, 797-808.

[JLZ] À. Jorba, R. de la Llave, M. Zou, *Lindstedt series for lower-dimensional tori*, Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995), 151–167, NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences, 533, Kluwer Acad. Publ., Dordrecht, 1999.

[G1] G. Gallavotti, *Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable hamiltonian systems. A review*, Reviews in Mathematical Physics **6** (1994), no. 3, 343–411.

[G2] G. Gallavotti, *Twistless KAM tori*, Communications in Mathematical Physics **164** (1994), no. 1, 145–156.

[GBG] G. Gallavotti, F. Bonetto, G. Gentile, *Aspects of the ergodic, qualitative and statistical theory of motion*, Springer-Verlag, Berlin, 2004.

[GG1] G. Gallavotti, G. Gentile, *Hyperbolic low-dimensional invariant tori and summations of divergent series*, Communications in Mathematical Physics **227** (2002), no. 3, 421–460.

[GG2] G. Gentile, G. Gallavotti, *Degenerate elliptic resonances*, Communications in Mathematical Physics **257** (2005), no. 2, 319–362.

[Ge1] G. Gentile, *Quasi-periodic solutions for two-level systems*, Communications in Mathematical Physics **242** (2003), no. 1–2, 221–250.

[Ge2] G. Gentile, *Diagrammatic techniques in perturbations theory, and applications*, Symmetry and perturbation theory (Rome, 1998), 59–78, World Sci. Publishing, River Edge, NJ, 1999.

[Ge3] G. Gentile, *Diagrammatic techniques in perturbation theory*, Encyclopaedia of Mathematical Physics, Eds. J.-P. Francoise, G. Naber and T. Sh. Tsun, Elsevier, Amsterdam, to appear.

[GBC] G. Gentile, D.A. Cortez, J.C.A. Barata, *Stability for quasi-periodically perturbed Hill's equations*, Communications in Mathematical Physics, to appear.

[GM] G. Gentile, V. Mastropietro, *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), no. 3, 393–444.

[GP] P. Goldreich, S. Peale, *Spin-orbit coupling in the solar system*, Astronomy Journal **71** (1966), no. 6, 425–438.

[Ka] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, Berlin-New York, 1976.

[P] H. Poincaré, *Les méthodes nouvelles de la mécanique céleste*, Tome II, Gauthier-Villars, Paris, 1893.

[Po] J. Pöschel, *On elliptic lower-dimensional tori in Hamiltonian systems*, Mathematische Zeitschrift **202** (1989), no. 4, 559–608.

[R1] H. Rüssmann, *Invariant tori in non-degenerate nearly integrable Hamiltonian systems*, Regular and Chaotic Dynamics **6** (2001), no. 2, 119–204.

[R2] H. Rüssmann, *Stability of elliptic fixed points of analytic area-preserving mappings under the Bruno condition*, Ergodic Theory and Dynamical Systems **22** (2002), no. 5, 1551-1573.