Commun. Math. Phys. 316, 489–529 (2012)
Digital Object Identifier (DOI) 10.1007/s00220-012-1548-2

Communications in
**Mathematical
Physics**

Oscillator Synchronisation under Arbitrary Quasi-periodic Forcing

Livia Corsi, Guido Gentile

Dipartimento di Matematica, Università di Roma Tre, Roma 00146, Italy.
E-mail: lcorsi@mat.uniroma3.it; gentile@mat.uniroma3.it

Received: 30 September 2011 / Accepted: 22 February 2012
Published online: 4 September 2012 –

Abstract: We study the problem of existence of response solutions for a real-analytic one-dimensional system, consisting of a rotator subject to a small quasi-periodic forcing with Bryuno frequency vector. We prove that at least one response solution always exists, without any assumption on the forcing besides smallness and analyticity. This strengthens the results available in the literature, where generic non-degeneracy conditions are assumed. The proof is based on a diagrammatic formalism and relies on renormalisation group techniques, which exploit the formal analogy with problems of quantum field theory; a crucial role is played by remarkable identities between classes of diagrams.

1. Introduction

Synchronisation and resonance phenomena are of the greatest relevance in the theory of dynamical systems, and have been extensively investigated since the earlier days of physics and applied mathematics; see [3] for an overview. Recently, the search for synchronisation has been extended also to chaotic systems (see for instance [4,32]). The very idea of synchronisation suggests that two or more systems adjust given properties of their motion to a common behaviour. For instance, when subjected to a periodic forcing, one-dimensional dynamical systems typically develop periodic solutions with the same frequency as the forcing term (response solutions); in the perturbation regime, the solution can be seen as a deformation of some unperturbed trajectory, which is fixed among all possible ones by a suitable locking of the phase (synchronisation).

In the chaotic case, the situation becomes much more subtle. Moreover, the presence or absence of dissipation may lead to very different behaviours when perturbing a chaotic system: very non-intuitive phenomena may take place in volume preserving systems (one can think of the pathological foliations discussed in [34,35]), while, for the same model in the presence of dissipation, a periodic forcing gives rise to strange attractors which can be quite smooth [19]. On the contrary, when considering periodic perturbations of integrable systems, too pathological situations are not expected to occur. For both Hamiltonian and dissipative one-dimensional systems, a response solution is a natural outcome (in the dissipative case, such a solution may become an attractor and hence plays a fundamental role in the understanding of the dynamics) and at worst it lacks analyticity in the perturbation parameter. Even taking a quasi-periodic perturbation, one can find solutions with the same frequency (vector) as the forcing, independently of the presence of dissipation. From a technical point of view, dealing with such solutions requires solving small divisor problems, similar to those appearing in the context of KAM theory. However, the problem cannot be considered quite solved for quasi-periodic perturbations, since a complete analysis is available in the literature only under some non-degeneracy assumptions on the perturbation. Such assumptions are generically satisfied, but still do not allow us to draw general conclusions. The aim of this paper is to investigate such non-generic situations, at least in a simple class of models. More precisely the problem can be described as follows.

Consider the one-dimensional system

$$\ddot{\beta} = -\varepsilon F(\boldsymbol{\omega} t, \beta), \qquad F(\boldsymbol{\omega} t, \beta) := \partial_\beta f(\boldsymbol{\omega} t, \beta), \tag{1.1}$$

where $\beta \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, $f : \mathbb{T}^{d+1} \to \mathbb{R}$ is a real-analytic function, $\boldsymbol{\omega} \in \mathbb{R}^d$ and ε is a real number, called the *perturbation parameter*; hence the *forcing function* (or perturbation) F is quasi-periodic in t, with *frequency vector* $\boldsymbol{\omega}$.

It is well known that, for $d = 1$ (periodic forcing) and ε small enough, there exist periodic solutions to (1.1) with the same period as the forcing. In fact the existence of periodic solutions to (1.1), or to the more general equation

$$\ddot{\beta} = -\partial_\beta V(\beta) - \varepsilon F(\boldsymbol{\omega} t, \beta), \tag{1.2}$$

with $V : \mathbb{R} \to \mathbb{R}$ real-analytic, can be discussed by relying on the *Melnikov method* [9,28]. A possible approach consists in splitting the equations of motion into two separate equations, the so-called *range equation* and *bifurcation equation*. One can solve the first equation in terms of a free parameter and then fix the latter by solving the second equation (which represents an implicit function problem). This is usually done by assuming some *non-degeneracy condition* involving the perturbation: the time average of the perturbation over the unperturbed motion is assumed to have a simple zero. As a byproduct, this entails analyticity of the solution. If no such condition is assumed, a result of the same kind still holds [1,10,38], but the scenario appears slightly more complicated. For instance the existing periodic solutions might no longer be analytic in the perturbation parameter.

If the forcing is quasi-periodic and $\boldsymbol{\omega}$ satisfies some Diophantine condition, one can still study the problem of existence of quasi-periodic solutions with the same frequency vector $\boldsymbol{\omega}$ as the forcing, for ε small enough. The analysis becomes much more involved, because of the small divisor problem. However, under the same generic non-degeneracy condition as above, the analysis can be carried out in a similar way and the bifurcation scenario can be described in a rather detailed way; see for instance [6]. On the contrary, if no assumption at all is made on the perturbation, the small divisor problem and the implicit function problem become inevitably tangled together and new difficulties arise. In this paper we focus on this situation: we study (1.1) without making any assumption on the forcing function besides analyticity and study the problem of existence of *response solutions*, i.e. a quasi-periodic solution with frequency vector $\boldsymbol{\omega}$. Of course, we also make some assumption of strong irrationality on the frequency vector $\boldsymbol{\omega}$, say we assume some mild Diophantine condition, such as the Bryuno condition (see below). We shall prove the existence of at least one response solution to (1.1) for ε small enough.

Note that (1.1) can be seen as the Hamilton equations for the system described by the Hamiltonian function

$$H(\boldsymbol{\alpha}, \beta, \boldsymbol{A}, B) = \boldsymbol{\omega} \cdot \boldsymbol{A} + \frac{1}{2} B^2 + \varepsilon f(\boldsymbol{\alpha}, \beta), \tag{1.3}$$

where $\boldsymbol{\omega} \in \mathbb{R}^d$ is fixed, $(\boldsymbol{\alpha}, \beta) \in \mathbb{T}^d \times \mathbb{T}$ and $(\boldsymbol{A}, B) \in \mathbb{R}^d \times \mathbb{R}$ are conjugate variables, f is an analytic periodic function of $(\boldsymbol{\alpha}, \beta)$ and $\cdot$ denotes the standard scalar product in $\mathbb{R}^d$. Indeed, the corresponding Hamilton equations for the angle variables are closed, and are given by

$$\dot{\boldsymbol{\alpha}} = \boldsymbol{\omega}, \qquad \ddot{\beta} = -\varepsilon \partial_\beta f(\boldsymbol{\alpha}, \beta), \tag{1.4}$$

that we can rewrite as (1.1). Therefore the problem of existence of response solutions can be seen as a problem of persistence of lower-dimensional (or resonant) tori, more precisely of d-dimensional tori for a particular system with $d + 1$ degrees of freedom. As usual in the context of KAM theory, by 'persisting torus' we mean an invariant torus for the perturbed system which is close to an unperturbed invariant torus and reduces to it when the perturbation is switched off. In the case of (1.3), the d-dimensional invariant tori $\mathcal{T}_{\boldsymbol{\omega},\beta_0} = \{(\boldsymbol{\alpha}, \beta, \boldsymbol{A}, B) : \boldsymbol{A} = \boldsymbol{0}, B = 0, \boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{\omega} t, \beta = \beta_0, \boldsymbol{\alpha}_0 \in \mathbb{T}^d\}$ of the unperturbed system, i.e. of (1.3) for $\varepsilon = 0$, foliate the $(d + 1)$-dimensional invariant torus $\mathcal{T}_{\boldsymbol{\omega}} = \cup_{\beta_0 \in \mathbb{T}} \mathcal{T}_{\boldsymbol{\omega},\beta_0}$ (note that each $\mathcal{T}_{\boldsymbol{\omega},\beta_0}$ is a submanifold of $\mathcal{T}_{\boldsymbol{\omega}}$); then we say that a d-dimensional invariant torus persists for the perturbed system ($\varepsilon \neq 0$) if there is a d-dimensional invariant manifold which is close to an unperturbed invariant torus $\mathcal{T}_{\boldsymbol{\omega},\beta_0}$ of $\mathcal{T}_{\boldsymbol{\omega}}$ for some β_0 (i.e. depends continuously on ε and reduces to $\mathcal{T}_{\boldsymbol{\omega},\beta_0}$ as $\varepsilon \to 0$) and is traversed quasi-periodically with the same frequency vector $\boldsymbol{\omega}$ as the unperturbed one.

The existence of d-dimensional tori in systems with $d + 1$ degrees of freedom, without imposing any non-degeneracy condition on the perturbation except analyticity, was first proved by Cheng [8], for convex unperturbed Hamiltonians and for $\boldsymbol{\omega}$ satisfying the standard Diophantine condition $|\boldsymbol{\omega} \cdot \boldsymbol{\nu}| \geq \gamma |\boldsymbol{\nu}|^{-\tau}$ for all $\boldsymbol{\nu} \in \mathbb{Z}^d \setminus \{\boldsymbol{0}\}$ and for some $\gamma > 0$ and $\tau > d - 1$ (here and henceforth $|\boldsymbol{\nu}| = |\boldsymbol{\nu}|_1 = |\nu_1| + \cdots + |\nu_d|$, if ν_i are the components of $\boldsymbol{\nu}$). In the case (1.3) the unperturbed Hamiltonian is isochronous in all but one angle variables (and we assume on $\boldsymbol{\omega}$ the weaker Bryuno condition), so that Cheng's paper does not cover our result; we defer further comments to Sect. 7.

The method we shall use is based on the analysis and resummation of the perturbation series through renormalisation group techniques [17,18,21–23,27] and not on an iteration scheme *à la* KAM. The main problems arise from the accumulation of small divisors: from a technical point of view this is reflected in the leading part of the self-energies (see Sect. 3 for details). In the non-degenerate case, the leading part is non-zero, but 'it has the right sign', so that the resummation can be checked to be well-defined. In the 'weakly non-degenerate' case (i.e. when the time average of the perturbation over the unperturbed motion has an odd order zero – see Hypothesis 2), the leading part can be arbitrarily small but still non-zero and it is much harder to keep control of its size. This will be obtained by introducing suitable cut-offs and eventually showing that the cut-offs can be removed. As a byproduct we find that, if the leading part is formally zero (that is if its formal power series expansion in terms of the perturbation parameter – which is always well defined order by order; see Appendix H – vanishes), then the full resonant torus persists and is analytical in the perturbation parameter (see Remark 6.7). If on the one hand one can argue that this could be expected, on the other hand a rigorous proof is not immediate and requires some work. In general, the response solution can only be proved to be continuous in the perturbation parameter; for further comments see Sect. 7.

A crucial role in our proof will be played by remarkable identities between classes of diagrams. By exploiting the analogy of the method with the techniques of quantum field theory, one can see the solution as the one-point Schwinger function of a suitable Euclidean field theory – this has been explicitly shown in the case of KAM tori [20]. In the case of classical KAM theorem, identities between diagrams analogous to those we prove and use follow from the translation invariance of the whole KAM torus; see [5]. We conjecture that our identities reflect some suitable Ward identity of the field theory symmetries also in the present case.

2. Results

Consider Eq. (1.1) and let the solution for the unperturbed system be given by $\beta(t) = \beta_0$. We want to study whether for some value of β_0 such a solution can be continued under perturbation. Define the *Bryuno function* [7] as

$$\mathcal{B}(\boldsymbol{\omega}) := \sum_{m=0}^{\infty} \frac{1}{2^m} \log \frac{1}{\alpha_m(\boldsymbol{\omega})}, \qquad \alpha_m(\boldsymbol{\omega}) := \inf_{0<|\boldsymbol{\nu}|\le 2^m} |\boldsymbol{\omega}\cdot\boldsymbol{\nu}|. \tag{2.1}$$

Hypothesis 1. *$\boldsymbol{\omega}$ satisfies the Bryuno condition $\mathcal{B}(\boldsymbol{\omega}) < \infty$.*

If $\boldsymbol{\omega}$ satisfies the standard Diophantine condition $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}| \ge \gamma|\boldsymbol{\nu}|^{-\tau}$ for all $\boldsymbol{\nu} \in \mathbb{Z}^d\setminus\{\mathbf{0}\}$ then it also satisfies the Bryuno condition, since $\alpha_m(\boldsymbol{\omega}) \ge \gamma\, 2^{-m\tau}$.

Write

$$f(\boldsymbol{\alpha},\beta) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d} f_{\boldsymbol{\nu}}(\beta) e^{i\boldsymbol{\nu}\cdot\boldsymbol{\alpha}}, \qquad F(\boldsymbol{\alpha},\beta) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d} F_{\boldsymbol{\nu}}(\beta) e^{i\boldsymbol{\nu}\cdot\boldsymbol{\alpha}}. \tag{2.2}$$

Hypothesis 2. *β_0^* is a zero of order $\mathfrak{n}$ for $F_{\mathbf{0}}(\beta)$, with $\mathfrak{n}$ odd, and $\varepsilon\partial_\beta^{\mathfrak{n}} F_{\mathbf{0}}(\beta_0^*) < 0$.*

Eventually we shall want to get rid of Hypothesis 2: however, we shall first assume it to simplify the analysis, and at the end we shall show how to remove it.

We look for a solution to (1.1) of the form $\beta(t) = \beta_0 + b(t)$, with

$$b(t) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d_*} e^{i\boldsymbol{\nu}\cdot\boldsymbol{\omega}t} b_{\boldsymbol{\nu}}, \tag{2.3}$$

where $\mathbb{Z}^d_* := \mathbb{Z}^d\setminus\{\mathbf{0}\}$. In Fourier space (1.1) becomes

$$(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2 b_{\boldsymbol{\nu}} = \varepsilon[F(\boldsymbol{\omega}t,\beta)]_{\boldsymbol{\nu}}, \qquad \boldsymbol{\nu} \ne \mathbf{0}, \tag{2.4a}$$

$$0 = [F(\boldsymbol{\omega}t,\beta)]_{\mathbf{0}}, \tag{2.4b}$$

where

$$[F(\boldsymbol{\psi},\beta)]_{\boldsymbol{\nu}} = \sum_{r\ge 0} \sum_{\substack{\boldsymbol{\nu}_0+\dots+\boldsymbol{\nu}_r=\boldsymbol{\nu} \\ \boldsymbol{\nu}_0\in\mathbb{Z}^d \\ \boldsymbol{\nu}_i\in\mathbb{Z}^d_*,\, i=1,\dots,r}} \frac{1}{r!}\partial_\beta^r F_{\boldsymbol{\nu}_0}(\beta_0) \prod_{i=1}^r b_{\boldsymbol{\nu}_i}.$$

Our first result will be the following.

Theorem 2.1. *Consider Eq. (1.1) and assume Hypotheses 1 and 2. If ε is small enough, there exists at least one quasi-periodic solution $\beta(t)$ to (1.1) with frequency vector $\boldsymbol{\omega}$, such that $\beta(t) \to \beta_0^*$ as $\varepsilon \to 0$.*

The proof will be carried out through Sects. 3 to 5. First, after introducing the basic notations in Sect. 3, we shall show in Sect. 4 that, under the assumption that further conditions are satisfied (see below), for ε small enough and arbitrary β_0 there exists a solution

$$\beta(t) = \beta_0 + b(t; \varepsilon, \beta_0), \tag{2.5}$$

to (2.4a), depending on ε, β_0, with $b(t) = b(t; \varepsilon, \beta_0)$ a zero-average function. For such a solution define $G(\varepsilon, \beta_0) := [F(\boldsymbol{\omega} t, \beta(t))]_{\mathbf{0}}$, and consider the implicit function equation

$$G(\varepsilon, \beta_0) = 0. \tag{2.6}$$

Then we shall prove in Sect. 5 that one can fix $\beta_0 = \beta_0(\varepsilon)$ in such a way that (2.6) holds and the conditions mentioned above are also satisfied. Hence for such $\beta_0(\varepsilon)$ the function (2.5) is a solution of the whole system (2.4).

The conditions which are first assumed to be satisfied and then a posteriori checked can be illustrated as follows. The resummation procedure turns out to be well-defined if the small divisors of the resummed series can be bounded proportionally to the small divisors of the formal series. However, it is not obvious at all that this is possible, because the latter are of the form $(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^{-2}$, with $\boldsymbol{\nu} \in \mathbb{Z}^d_*$, whereas the small divisors of the resummed series are $((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2 - \mathcal{M}_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu}; \varepsilon, \beta_0))^{-1}$, for suitable functions $\mathcal{M}_n$ (see Sect. 3). In the non-degenerate case one can fix β_0 in such a way that $\mathcal{M}_n(x; \varepsilon, \beta_0) = a\,\varepsilon + O(\varepsilon^2)$, with $a\varepsilon < 0$, so that $((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2 - \mathcal{M}_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu}; \varepsilon, \beta_0))^{-1} \sim (\boldsymbol{\omega}\cdot\boldsymbol{\nu})^{-2}$. If the non-degeneracy assumption is removed, the latter property is much more difficult to check. So we introduce some cut-offs in order to make the property automatically satisfied: essentially we replace $\mathcal{M}_n(x; \varepsilon, \beta_0)$ with $\mathcal{M}_n(x; \varepsilon, \beta_0)\,\xi_n(\mathcal{M}(0; \varepsilon, \beta_0))$, for suitable 'cut-off functions' ξ_n. However, the introduction of the cut-offs changes the series, in such a way that if on the one hand the modified series can be proved to be convergent, on the other hand in principle it no longer solves the equations of motion. This turns out to be the case only if one can prove that the cut-offs can be removed (that is if one can replace the cut-off functions ξ_n with 1). So, the last part of the proof consists in showing that, by suitably choosing the value β_0, this occurs and hence the series with the cut-offs reduces to the solution.

Next, we shall see how to remove Hypothesis 2 in order to prove the existence of a response solution without any assumption on the forcing function, so as to obtain the following result, which is the main result of the paper.

Theorem 2.2. *Consider Eq. (1.1) and assume Hypothesis 1. There exists $\varepsilon_0 > 0$ such that for all ε with $|\varepsilon| < \varepsilon_0$ there is at least one quasi-periodic solution to (1.1) with frequency vector $\boldsymbol{\omega}$. Such a solution depends continuously on ε.*

If $F_{\mathbf{0}}(\beta)$ does not identically vanish, then Theorem 2.2 follows immediately from Theorem 2.1. Indeed, the function $f_{\mathbf{0}}(\beta)$ is analytic and periodic, hence, if it is not identically constant, it has at least one maximum point β_0' and one minimum point β_0'', where $\partial_\beta^{\mathfrak{n}'+1} f_{\mathbf{0}}(\beta_0') < 0$ and $\partial_\beta^{\mathfrak{n}''+1} f_{\mathbf{0}}(\beta_0'') > 0$, for some $\mathfrak{n}'$ and $\mathfrak{n}''$ both odd. Let ε be fixed small enough, say $|\varepsilon| < \varepsilon_0$ for a suitable ε_0: choose $\beta_0^* = \beta_0'$ if $\varepsilon > 0$ and $\beta_0^* = \beta_0''$ if $\varepsilon < 0$. Then Hypothesis 2 is satisfied, and we can apply Theorem 2.1 to deduce the existence of a quasi-periodic solution with frequency vector $\boldsymbol{\omega}$. However, the function $f_{\mathbf{0}}(\beta)$ can be identically constant and hence $F_{\mathbf{0}}(\beta)$ can vanish identically, so that some further work will be needed to prove Theorem 2.2: this will be performed in Sect. 6.

Note that, except for the very special case discussed in Sect. 6, the response solution is not expected to be analytic in ε. In some cases smoothness in ε or some fractional power of ε can be obtained (see Sect. 7), but in general the solution is only proved to be continuous in ε. Moreover, generically β_0 has to be chosen suitably as a function of ε, whereas it remains arbitrary in the "completely degenerate case" in Sect. 6.

3. Diagrammatic Rules and Multiscale Analysis

We want to study whether it is possible to express the function $b(t;\varepsilon,\beta_0)$ appearing in (2.5) as a convergent series. Let us start by writing formally

$$b(t;\varepsilon,\beta_0)=\sum_{k\geq 1}\varepsilon^k b^{(k)}(t;\beta_0)=\sum_{k\geq 1}\varepsilon^k\sum_{\boldsymbol{\nu}\in\mathbb{Z}^d_*}e^{\mathrm{i}\boldsymbol{\nu}\cdot\boldsymbol{\omega}t}b^{(k)}_{\boldsymbol{\nu}}(\beta_0). \tag{3.1}$$

If we define recursively for $k\geq 1$ and $\boldsymbol{\nu}\neq\mathbf{0}$,

$$b^{(k)}_{\boldsymbol{\nu}}(\beta_0)=\frac{1}{(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2}[F(\boldsymbol{\omega}t,\beta)]^{(k-1)}_{\boldsymbol{\nu}}, \tag{3.2}$$

where $[F(\boldsymbol{\omega}t,\beta)]^{(0)}_{\boldsymbol{\nu}}=F_{\boldsymbol{\nu}}(\beta_0)$ and, for $k\geq 1$ and $\boldsymbol{\nu}\in\mathbb{Z}^d$,

$$[F(\boldsymbol{\omega}t,\beta)]^{(k)}_{\boldsymbol{\nu}}=\sum_{s\geq 1}\sum_{\substack{\boldsymbol{\nu}_0+\dots+\boldsymbol{\nu}_s=\boldsymbol{\nu}\\ \boldsymbol{\nu}_0\in\mathbb{Z}^d\\ \boldsymbol{\nu}_i\in\mathbb{Z}^d_*,\,i=1,\dots,s}}\frac{1}{s!}\partial^s_\beta F_{\boldsymbol{\nu}_0}(\beta_0)\sum_{\substack{k_1+\dots+k_s=k,\\ k_i\geq 1}}\prod_{i=1}^{s}b^{(k_i)}_{\boldsymbol{\nu}_i}(\beta_0), \tag{3.3}$$

the series (3.1) turns out to be a formal solution of (2.4a) only: the coefficients $b^{(k)}_{\boldsymbol{\nu}}(\beta_0)$ are well defined for all $k\geq 1$ and all $\boldsymbol{\nu}\in\mathbb{Z}^d_*$ – see Appendix H – and solve (2.4a) order by order – as is straightforward to check.

Write also (recall the definition of $G(\varepsilon,\beta_0)$ before (2.6)), again formally,

$$G(\varepsilon,\beta_0)=\sum_{k\geq 0}\varepsilon^k G^{(k)}(\beta_0), \tag{3.4}$$

with $G^{(0)}(\beta_0)=F_{\mathbf{0}}(\beta_0)$ and, for $k\geq 1$,

$$G^{(k)}(\beta_0)=\sum_{s\geq 1}\sum_{\substack{\boldsymbol{\nu}_0+\dots+\boldsymbol{\nu}_s=\mathbf{0}\\ \boldsymbol{\nu}_0\in\mathbb{Z}^d\\ \boldsymbol{\nu}_i\in\mathbb{Z}^d_*,\,i=1,\dots,s}}\frac{1}{s!}\partial^s_\beta F_{\boldsymbol{\nu}_0}(\beta_0)\sum_{\substack{k_1+\dots+k_s=k,\\ k_i\geq 1}}\prod_{i=1}^{s}b^{(k_i)}_{\boldsymbol{\nu}_i}(\beta_0). \tag{3.5}$$

Of course, Hypothesis 1 yields that the formal series (3.4) is well-defined too.

Unfortunately the power series (3.1) and (3.4) may be not convergent (as far as we know). However we shall see how to construct two series (convergent if β_0 is suitably chosen) whose formal expansions coincide with (3.1) and (3.4). As we shall see, this leads to express the response solution as a series of contributions each of which can be graphically represented as a suitable diagram.

A graph is a set of points and lines connecting them. A *tree* θ is a graph with no cycle, such that all the lines are oriented toward a single point (*root*) which has only one incident line ℓ_θ (*root line*). All the points in a tree except the root are called *nodes*. The orientation of the lines in a tree induces a partial ordering relation ($\preceq$) between the nodes and the lines: we can imagine that each line carries an arrow pointing toward the root. Given two nodes v and w, we shall write $w \prec v$ every time v is along the path (of lines) which connects w to the root.

We denote by $N(\theta)$ and $L(\theta)$ the sets of nodes and lines in θ respectively. Since a line $\ell \in L(\theta)$ is uniquely identified by the node v which it leaves, we may write $\ell = \ell_v$. We write $\ell_w \prec \ell_v$ if $w \prec v$, and $w \prec \ell = \ell_v$ if $w \preceq v$; if ℓ and ℓ' are two comparable lines, i.e. $\ell' \prec \ell$, we denote by $\mathcal{P}(\ell, \ell')$ the (unique) path of lines connecting ℓ' to ℓ, with ℓ and ℓ' not included (in particular $\mathcal{P}(\ell, \ell') = \emptyset$ if ℓ' enters the node ℓ exits).

With each node $v \in N(\theta)$ we associate a *mode* label $\boldsymbol{\nu}_v \in \mathbb{Z}^d$ and we denote by s_v the number of lines entering v. With each line ℓ we associate a *momentum* $\boldsymbol{\nu}_\ell \in \mathbb{Z}^d_*$, except for the root line which can have either zero momentum or not, i.e. $\boldsymbol{\nu}_{\ell_\theta} \in \mathbb{Z}^d$. Finally, we associate with each line ℓ also a *scale label* such that $n_\ell = -1$ if $\boldsymbol{\nu}_\ell = \mathbf{0}$, while $n_\ell \in \mathbb{Z}_+$ if $\boldsymbol{\nu}_\ell \neq \mathbf{0}$ (so far there is no relation between non-zero momenta and scale labels: a constraint will appear later on, see Remark 3.7). Note that one can have $n_\ell = -1$ only if ℓ is the root line of θ. We force the following *conservation law*:

$$\boldsymbol{\nu}_\ell = \sum_{\substack{w \in N(\theta) \\ w \prec \ell}} \boldsymbol{\nu}_w. \tag{3.6}$$

In the following we shall call trees tout court the trees with labels, and we shall use the term *unlabelled tree* for the trees without labels. We shall say that two trees are *equivalent* if they can be transformed into each other by continuously deforming the lines in such a way that these do not cross each other and also labels match. This provides an equivalence relation on the set of the trees. From now on we shall call trees such equivalence classes.

Given a tree θ we call the *order* of θ the number $k(\theta) = |N(\theta)| = |L(\theta)|$ (for any finite set S we denote by $|S|$ its cardinality) and the *total momentum* of θ the momentum associated with ℓ_θ. We shall denote by $\Theta_{k,\boldsymbol{\nu}}$ the set of trees with order k and total momentum $\boldsymbol{\nu}$. A subset $T \subset \theta$ is a *subgraph* of θ if it is formed by a set of nodes $N(T) \subseteq N(\theta)$ and lines $L(T) \subseteq L(\theta)$ connecting them (possibly including the root line, and in such a case we say that the root is included in T) in such a way that $N(T) \cup L(T)$ is connected. If T is a subgraph of θ we call the *order* of T the number $k(T) = |N(T)|$. We say that a line enters T if it connects a node $v \notin N(T)$ to a node $w \in N(T)$, and we say that a line exits T if it connects a node $v \in N(T)$ to a node $w \notin N(T)$ or to the root (which is not included in T in this case). Of course, if a line ℓ enters or exits T, then $\ell \notin L(T)$.

Remark 3.1. One has $\sum_{v \in N(\theta)} s_v = k(\theta) - 1$.

A *cluster* T on scale n is a maximal subgraph of a tree θ such that all the lines have scales $n' \leq n$ and there is at least a line with scale n. The lines entering the cluster T and the line coming out from it (unique if existing at all) are called the *external* lines of T.

A *self-energy cluster* is a cluster T such that (i) T has only one entering line ℓ'_T and one exiting line ℓ_T, (ii) one has $\boldsymbol{\nu}_{\ell_T} = \boldsymbol{\nu}_{\ell'_T}$ and hence $\sum_{v \in N(T)} \boldsymbol{\nu}_v = \mathbf{0}$.

For any self-energy cluster T, set $\mathcal{P}_T = \mathcal{P}(\ell_T, \ell'_T)$. More generally, if T is a subgraph of θ with only one entering line ℓ' and one exiting line ℓ, we can set $\mathcal{P}_T = \mathcal{P}(\ell, \ell')$. We shall say that a self-energy cluster is on scale -1, if $N(T) = \{v\}$ with of course $\boldsymbol{\nu}_v = \mathbf{0}$ (so that $\mathcal{P}_T = \emptyset$).

A *left-fake cluster* T on scale n is a connected subgraph of a tree θ with only one entering line ℓ'_T and one exiting line ℓ_T such that (i) all the lines in T have scale $\le n$ and there is in T at least a line on scale n, (ii) ℓ'_T is on scale $n+1$ and ℓ_T is on scale n, and (iii) one has $\boldsymbol{\nu}_{\ell_T} = \boldsymbol{\nu}_{\ell'_T}$. Analogously a *right-fake cluster* T on scale n is a connected subgraph of a tree θ with only one entering line ℓ'_T and one exiting line ℓ_T such that (i) all the lines in T have scale $\le n$ and there is in T at least a line on scale n, (ii) ℓ'_T is on scale n and ℓ_T is on scale $n+1$, and (iii) one has $\boldsymbol{\nu}_{\ell_T} = \boldsymbol{\nu}_{\ell'_T}$. Roughly speaking, a left-fake (respectively right-fake) cluster T fails to be a self-energy cluster only because the exiting (respectively the entering) line is on scale equal to the scale of T; note that left- and right-fake clusters are not even clusters.

Remark 3.2. Given a self-energy cluster T, the momenta of the lines in $\mathcal{P}_T$ depend on $\boldsymbol{\nu}_{\ell'_T}$ because of the conservation law (3.6). More precisely, for all $\ell \in \mathcal{P}_T$ one has $\boldsymbol{\nu}_\ell = \boldsymbol{\nu}^0_\ell + \boldsymbol{\nu}_{\ell'_T}$ with $\boldsymbol{\nu}^0_\ell = \sum_{w \in N(T), w \prec \ell} \boldsymbol{\nu}_w$, while all the other labels in T do not depend on $\boldsymbol{\nu}_{\ell'_T}$. Clearly, this holds also for left-fake and right-fake clusters.

We say that two self-energy clusters T_1, T_2 have the same *structure* if setting $\boldsymbol{\nu}_{\ell'_{T_1}} = \boldsymbol{\nu}_{\ell'_{T_2}} = \mathbf{0}$ one has $T_1 = T_2$. Of course this provides an equivalence relation on the set of all self-energy clusters. The same considerations apply for left-fake and right-fake clusters. From now on we shall call self-energy, left-fake and right-fake clusters tout court such equivalence classes.

A *renormalised tree* is a tree in which no self-energy clusters appear; analogously a *renormalised subgraph* is a subgraph of a tree θ which does not contains any self-energy cluster. Denote by $\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$ the set of renormalised trees with order k and total momentum $\boldsymbol{\nu}$, by $\mathfrak{R}_n$ the set of renormalised self-energy clusters on scale n, and by $\mathfrak{LF}_n$ and $\mathfrak{RF}_n$ the sets of (renormalised) left-fake and right-fake clusters on scale n respectively.

For any $\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$ we associate with each node $v \in N(\theta)$ a *node factor*

$$\mathcal{F}_v(\beta_0) := \frac{1}{s_v!} \partial_\beta^{s_v} F_{\boldsymbol{\nu}_v}(\beta_0). \tag{3.7}$$

We associate with each line $\ell \in L(\theta)$ with $n_\ell \ge 0$, a *dressed propagator* $\mathcal{G}_{n_\ell}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell; \varepsilon, \beta_0)$ (propagator tout court in the following) defined recursively as follows.

Introduce first a partition of unity. The idea is to have at disposal some labels (that we called scales) characterising the sizes of the small divisors $(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2$. Roughly we would like to associate with a line ℓ a scale label n if $\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell \approx \alpha_n(\boldsymbol{\omega})$; to be more precise, since the sequence $\{\alpha_n(\boldsymbol{\omega})\}_{n\ge0}$ is only non-increasing, for the scales to be uniquely defined one should take a decreasing subsequence $\{\alpha_{m_n}(\boldsymbol{\omega})\}_{n\ge0}$ and say that ℓ has scale n if $\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell$ is of order $\alpha_{m_n}(\boldsymbol{\omega})$. It would be tempting to use a sharp partition through step functions with supports $[\alpha_{m_n}(\boldsymbol{\omega}), \alpha_{m_{n-1}}(\boldsymbol{\omega}))$, in order to associate with a line ℓ a scale n if $|\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell| \in [\alpha_{m_n}(\boldsymbol{\omega}), \alpha_{m_{n-1}}(\boldsymbol{\omega}))$. However, it turns out to be more convenient using a smooth partition through compact support functions Ψ_n (because we have to take derivatives of quantities involving such functions). Therefore we shall proceed as follows. Given a decreasing sequence ρ_n, $n = 0, 1, \ldots$, of positive numbers with $\rho_{n+1} \le \rho_n/2$,

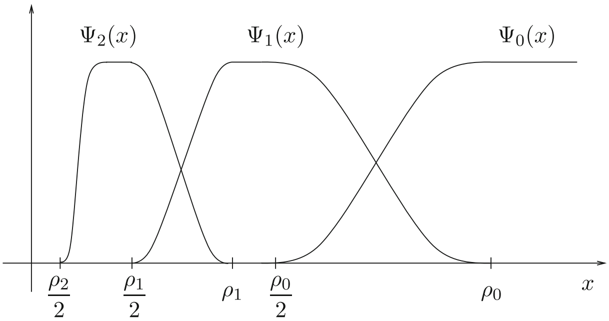

Fig. 1. Some of the C^∞ functions $\Psi_n(x)$ partitioning the unity in $\mathbb{R}\backslash\{0\}$

let $\chi : \mathbb{R} \to \mathbb{R}$ be a C^∞ even function, non-increasing for $x \geq 0$, such that

$$\chi(x) = \begin{cases} 1, & |x| \leq 1/2, \\ 0, & |x| \geq 1, \end{cases} \tag{3.8}$$

and set $\chi_n(x) = \chi(x/\rho_n)$ for $n \geq 0$ and $\chi_{-1}(x) = 1$. Set also $\Psi_n(x) = \chi_{n-1}(x) - \chi_n(x)$ for $n \geq -1$; see Fig. 1.

Remark 3.3. For all $x \neq 0$ one has $\sum_{n=0}^{\infty} \Psi_n(x) = 1$, and more generally, if $\psi_n(x) = 1 - \chi_n(x)$, for all $x \neq 0$ and all $p \geq 0$, one has $\psi_p(x) + \sum_{n \geq p+1} \Psi_n(x) = 1$.

Next, we introduce the sequences $\{m_n, p_n\}_{n\geq 0}$, with $m_0 = 0$ and, for all $n \geq 0$, $m_{n+1} = m_n + p_n + 1$, where $p_n := \max\{q \in \mathbb{Z}_+ : \alpha_{m_n}(\boldsymbol{\omega}) < 2\alpha_{m_n+q}(\boldsymbol{\omega})\}$, with $\alpha_m(\boldsymbol{\omega})$ defined in (2.1). The subsequence $\{\alpha_{m_n}(\boldsymbol{\omega})\}_{n\geq 0}$ of $\{\alpha_m(\boldsymbol{\omega})\}_{m\geq 0}$ is decreasing. A convenient partition of unity is then obtained by choosing $\rho_n = \alpha_{m_n}(\boldsymbol{\omega})/8$ (the factor 8 could be replaced by any number ≥ 8, as the proof of the forthcoming Lemma 4.1 shows; see in particular Remark 4.3 below).

Define, for $n \geq 0$,

$$\mathcal{G}_n(x; \varepsilon, \beta_0) := \Psi_n(x) \left(x^2 - \mathcal{M}_{n-1}(x; \varepsilon, \beta_0) \right)^{-1}, \tag{3.9a}$$

$$\begin{aligned} \mathcal{M}_{n-1}(x; \varepsilon, \beta_0) &:= \sum_{q=-1}^{n-1} \chi_q(x) M_q(x; \varepsilon, \beta_0), \\ M_q(x; \varepsilon, \beta_0) &:= \sum_{T \in \mathfrak{R}_q} \varepsilon^{k(T)} \mathscr{V}_T(x; \varepsilon, \beta_0), \end{aligned} \tag{3.9b}$$

$$\mathscr{V}_T(x; \varepsilon, \beta_0) := \Bigg(\prod_{v \in N(T)} \mathcal{F}_v(\beta_0) \Bigg) \Bigg(\prod_{\ell \in L(T)} \mathcal{G}_{n_\ell}(\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell; \varepsilon, \beta_0) \Bigg), \tag{3.9c}$$

where $\mathscr{V}_T(x; \varepsilon, \beta_0)$ is the *renormalised value* of T at fixed momentum $\boldsymbol{\nu}_{\ell'_T}$ such that $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}_{\ell'_T}$. Here and henceforth, the sums and the products over the empty set have to be considered as zero and 1 respectively. Note that $\mathscr{V}_T$ depends on ε because the propagators do, and on $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}_{\ell'_T}$ only through the propagators associated with the lines $\ell \in \mathcal{P}_T$ (see Remark 3.2).

Set $\mathcal{M} = \{\mathcal{M}_n(x; \varepsilon, \beta_0)\}_{n\geq -1}$. We call *self-energies* the quantities $\mathcal{M}_n(x; \varepsilon, \beta_0)$.

Remark 3.4. One has $|\mathfrak{R}_{-1}| = 1$, so that $\mathcal{M}_{-1}(x;\varepsilon,\beta_0) = M_{-1}(x;\varepsilon,\beta_0) = \varepsilon\partial_{\beta_0} F_{\mathbf{0}}(\beta_0)$. Therefore the propagators $\mathcal{G}_n$ and the self-energies $\mathcal{M}_n$ are uniquely defined for $n \ge 0$.

Remark 3.5. One has $\partial_{\beta_0}\mathcal{G}_n(x;\varepsilon,\beta_0) = \mathcal{G}_n(x;\varepsilon,\beta_0)(x^2 - \mathcal{M}_{n-1}(x;\varepsilon,\beta_0))^{-1}\partial_{\beta_0} \mathcal{M}_{n-1}(x;\varepsilon,\beta_0)$.

Set also $\mathcal{G}_{-1}(0;\varepsilon,\beta_0) = 1$ (so that we can associate a propagator also with the root line of $\theta \in \Theta^{\mathcal{R}}_{k,\mathbf{0}}$). For any subgraph S of any $\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$ define the *renormalised value* of S as

$$\mathscr{V}(S;\varepsilon,\beta_0) := \Bigg(\prod_{v\in N(S)} \mathcal{F}_v(\beta_0)\Bigg)\Bigg(\prod_{\ell\in L(S)} \mathcal{G}_{n_\ell}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\varepsilon,\beta_0)\Bigg). \tag{3.10}$$

Finally set

$$b^{[k]}_{\boldsymbol{\nu}}(\varepsilon,\beta_0) := \sum_{\theta\in\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}} \mathscr{V}(\theta;\varepsilon,\beta_0), \qquad G^{[k]}(\varepsilon,\beta_0) := \sum_{\theta\in\Theta^{\mathcal{R}}_{k+1,\mathbf{0}}} \mathscr{V}(\theta;\varepsilon,\beta_0), \tag{3.11}$$

and define formally

$$b^{\mathcal{R}}(t;\varepsilon,\beta_0) := \sum_{k\ge 1}\varepsilon^k \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d_*} e^{\mathrm{i}\boldsymbol{\nu}\cdot\boldsymbol{\omega}t} b^{[k]}_{\boldsymbol{\nu}}(\varepsilon,\beta_0), \tag{3.12a}$$

$$G^{\mathcal{R}}(\varepsilon,\beta_0) := \sum_{k\ge 0}\varepsilon^k G^{[k]}(\varepsilon,\beta_0). \tag{3.12b}$$

The series (3.12) will be called the *resummed series*. The term “resummed” comes from the fact that if we formally expand (3.12) in powers of ε, we obtain (3.1) and (3.4), as is easy to check.

Remark 3.6. If T is a renormalised left-fake (respectively right-fake) cluster, we can (and shall) write $\mathscr{V}(T;\varepsilon,\beta_0) = \mathscr{V}_T(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_{\ell'_T};\varepsilon,\beta_0)$ since the propagators of the lines in $\mathcal{P}_T$ depend on $\boldsymbol{\omega}\cdot\boldsymbol{\nu}_{\ell'_T}$. In particular one has

$$\sum_{T\in\mathfrak{LF}_n} \varepsilon^{k(T)}\,\mathscr{V}_T(x;\varepsilon,\beta_0) = \sum_{T\in\mathfrak{RF}_n} \varepsilon^{k(T)}\,\mathscr{V}_T(x;\varepsilon,\beta_0) = M_n(x;\varepsilon,\beta_0).$$

Remark 3.7. Given a renormalised tree θ such that $\mathscr{V}(\theta;\varepsilon,\beta_0) \ne 0$, for any line $\ell \in L(\theta)$ (except possibly the root line) one has $\Psi_{n_\ell}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell) \ne 0$, and hence

$$\frac{\alpha_{m_{n_\ell}}(\boldsymbol{\omega})}{16} < |\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell| < \frac{\alpha_{m_{n_\ell -1}}(\boldsymbol{\omega})}{8}, \tag{3.13}$$

where $\alpha_{m_{-1}}(\boldsymbol{\omega})$ has to be interpreted as $+\infty$. The same considerations apply to any subgraph of θ and to renormalised self-energy clusters. Moreover, by the definition of $\{\alpha_{m_n}(\boldsymbol{\omega})\}_{n\ge 0}$, the number of scales which can be associated with a line ℓ in such a way that the propagator does not vanish is at most 2; see Fig. 1.

4. Bounds and Convergence of the Resummed Series: Part 1

The key bounds on renormalised graphs are based on the idea of Siegel [36], in the version exploited by Pöschel [31] and Eliasson [13]. The first step is to obtain a bound on the number of lines on a given scale in a renormalised tree or self-energy cluster.

For $\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$, let $\mathfrak{N}_n(\theta)$ be the number of lines on scale $\geq n$ in θ, and set

$$K(\theta) := \sum_{v \in N(\theta)} |\boldsymbol{\nu}_v|. \tag{4.1}$$

More generally, for any renormalised subgraph T of any tree θ call $\mathfrak{N}_n(T)$ the number of lines on scale $\geq n$ in T, and set

$$K(T) := \sum_{v \in N(T)} |\boldsymbol{\nu}_v|. \tag{4.2}$$

Lemma 4.1. *For any $\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$ such that $\mathscr{V}(\theta; \varepsilon, \beta_0) \neq 0$ one has $\mathfrak{N}_n(\theta) \leq 2^{-(m_n-2)} K(\theta)$ for all $n \geq 0$.*

The proof is given in Appendix A.

Lemma 4.2. *For any $T \in \mathfrak{R}_n$ such that $\mathscr{V}_T(x; \varepsilon, \beta_0) \neq 0$, one has $K(T) \geq 2^{m_n-1}$ and $\mathfrak{N}_p(T) \leq 2^{-(m_p-2)} K(T)$ for all $0 \leq p \leq n$.*

The proof is given in Appendix B.

Remark 4.3. By looking carefully at the proofs of the two lemmas above, one realises that if (3.13) is replaced with

$$\frac{\alpha_{m_{n_\ell}}(\boldsymbol{\omega})}{32} < |\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell| < \frac{\alpha_{m_{n_\ell-1}}(\boldsymbol{\omega})}{4}, \tag{4.3}$$

the same bounds on $\mathfrak{N}_n(\theta)$ and $\mathfrak{N}_p(T)$ as in Lemmas 4.1 and 4.2 still hold. This will be used later – see Remark 4.9 below.

To prove that the resummed series (3.12) converges, we first make the assumption that the propagators $\mathcal{G}_{n_\ell}(x; \varepsilon, \beta_0)$ are bounded essentially as $1/x^2$: we shall see that in that case the convergence of the series can be routinely checked. Then, in Sect. 5, we shall check that the assumption is justified.

Definition 4.4. *We shall say that $\mathcal{M}$ satisfies* Property 1 *if for all $n \geq -1$ one has*

$$\Psi_{n+1}(x)|x^2 - \mathcal{M}_n(x; \varepsilon, \beta_0)| \geq \Psi_{n+1}(x)x^2/2.$$

Lemma 4.5. *Assume $\mathcal{M}$ to satisfy Property 1. Then the series (3.12), with the coefficients given in (3.11), converge for ε small enough.*

The proof is given in Appendix C.

Lemma 4.6. *Assume $\mathcal{M}$ to satisfy Property 1. Then for ε small enough the function $b^{\mathcal{R}}(t; \varepsilon, \beta_0)$ in (3.12), with the coefficients given in (3.11), solves Eq. (2.4a).*

The proof is given in Appendix D.

Definition 4.7. *We shall say that* $\mathcal{M}$ *satisfies* Property 1-p *if for* $-1 \le n < p$ *one has*

$$\Psi_{n+1}(x)|x^2 - \mathcal{M}_n(x;\varepsilon,\beta_0)| \ge \Psi_{n+1}(x)x^2/2.$$

Note that Property 1 is equivalent to assuming Property 1-p for all $p \ge 0$. The reason to introduce the last definition is that – as the following Lemma 4.8 will yield – if one assumes Property 1-p (i.e. that the inequalities $\Psi_{n+1}(x)|x^2 - \mathcal{M}_n(x;\varepsilon,\beta_0)| \ge \Psi_{n+1}(x)x^2/2$ hold for all $n < p$) then the propagators can be controlled for all scales $\le p$ and hence the self-energies $\mathcal{M}_p$ can be easily bounded by using Lemma 4.2. This will be exploited later on to prove iteratively that Property 1-p holds for all $p \ge 0$ and hence Property 1 holds too.

Lemma 4.8. *Assume* $\mathcal{M}$ *to satisfy Property 1-p. Then for any* $0 \le n \le p$ *the self-energies are well defined and one has*

$$|M_n(x;\varepsilon,\beta_0)| \le \varepsilon^2 K_1 e^{-K_2 2^{m_n}}, \tag{4.4a}$$

$$|\partial_x^j M_n(x;\varepsilon,\beta_0)| \le \varepsilon^2 C_j e^{-\overline{C}_j 2^{m_n}}, \qquad j = 1, 2, \tag{4.4b}$$

for suitable constants $K_1, K_2, C_1, C_2, \overline{C}_1$ *and* $\overline{C}_2$.

The proof is in Appendix E.

Remark 4.9. One can write

$$\mathcal{M}_n(x;\varepsilon,\beta_0) = \mathcal{M}_n(0;\varepsilon,\beta_0) + x\,\partial_x\mathcal{M}_n(0;\varepsilon,\beta_0) + x^2\int_0^1 \mathrm{d}\tau\;(1-\tau)\,\partial_x^2\mathcal{M}_n(\tau x;\varepsilon,\beta_0).$$

Then one checks, by relying on Remark 4.3, that $\partial_x^j\mathcal{M}_n(\tau x;\varepsilon,\beta_0)$ admits the same bounds as in Lemma 4.8, for $j = 0, 1, 2$ and $\tau \in [0, 1]$. This implies that

$$|\mathcal{M}_n(x;\varepsilon,\beta_0) - \mathcal{M}_n(0;\varepsilon,\beta_0) - x\,\partial_x\mathcal{M}_n(0;\varepsilon,\beta_0)| \le C\varepsilon^2x^2.$$

Lemma 4.10. *Assume* $\mathcal{M}$ *to satisfy Property 1-p. Then one has* $\mathcal{M}_n(x;\varepsilon,\beta_0) = \mathcal{M}_n(0;\varepsilon,\beta_0) + O(\varepsilon^2x^2)$ *for all* $0 \le n \le p$.

Proof. We shall prove that $\mathcal{M}_n(x;\varepsilon,\beta_0) = \mathcal{M}_n(-x;\varepsilon,\beta_0)$, by induction on $n \ge -1$. For $n = -1$ the identity is obvious since $\mathcal{M}_{-1}$ does not depend on x. Assume now $\mathcal{M}_q(x;\varepsilon,\beta_0) = \mathcal{M}_q(-x;\varepsilon,\beta_0)$ for all $q < n$. This implies $\mathcal{G}_q(x;\varepsilon,\beta_0) = \mathcal{G}_q(-x;\varepsilon,\beta_0)$ for $q \le n$. Let $T \in \mathfrak{R}_n$ and consider the self-energy cluster T_1 obtained from T by taking ℓ_T as the entering line and ℓ'_T as the exiting line (i.e. $\ell'_{T_1} = \ell_T$ and $\ell_{T_1} = \ell'_T$) and by taking $\boldsymbol{\nu}_{\ell'_{T_1}} = -\boldsymbol{\nu}_{\ell'_T}$. Hence the momenta of the lines belonging to $\mathcal{P}_T$ change signs, while all the other momenta do not change: therefore all propagators are left unchanged. Hence $M_n(x;\varepsilon,\beta_0) = M_n(-x;\varepsilon,\beta_0)$, so that $\partial_x M_n(0;\varepsilon,\beta_0) = 0$ for all $n \le p$, and, by Lemma 4.8, this is enough to prove the assertion. □

Lemma 4.11. *Assume* $\mathcal{M}$ *to satisfy Property 1. Then the function* $G^{\mathcal{R}}(\varepsilon,\beta_0)$ *and the self-energies* $\mathcal{M}_n(x;\varepsilon,\beta_0)$ *are* C^∞ *in both* ε *and* β_0.

Proof. It follows from the explicit expressions for $G^{\mathcal{R}}(\varepsilon,\beta_0)$ and $\mathcal{M}_n(x;\varepsilon,\beta_0)$. □

Define formally

$$\mathcal{M}_\infty(x;\varepsilon,\beta_0) = \lim_{n\to\infty} \mathcal{M}_n(x;\varepsilon,\beta_0), \tag{4.5}$$

and note that if $\mathcal{M}$ satisfies Property 1, then $\mathcal{M}_\infty(x;\varepsilon,\beta_0)$ is well defined and moreover it is C^∞ in both ε and β_0. The following result plays a crucial role; the proof is deferred to Appendix F.

Lemma 4.12. *Assume $\mathcal{M}$ to satisfy Property 1. Then one has $\varepsilon\partial_{\beta_0} G^{\mathcal{R}}(\varepsilon,\beta_0) = \mathcal{M}_\infty(0;\varepsilon,\beta_0)$.*

Remark 4.13. If we take the formal power expansions of both $G^{\mathcal{R}}(\varepsilon,\beta_0)$ and $\mathcal{M}_\infty(0;\varepsilon,\beta_0)$, we obtain tree expansions where self-energy clusters are allowed; see Sect. 6 for further details. Then the identity $\varepsilon\partial_{\beta_0} G^{\mathcal{R}}(\varepsilon,\beta_0) = \mathcal{M}_\infty(0;\varepsilon,\beta_0)$ is easily found to be satisfied to any perturbation order. However, without any resummation procedure, we are no longer able to prove the convergence of the series, so that the identity becomes a meaningless "$\infty = \infty$".

Remark 4.14. The identity $\varepsilon\partial_{\beta_0} G^{\mathcal{R}}(\varepsilon,\beta_0) = \mathcal{M}_\infty(0;\varepsilon,\beta_0)$, in Lemma 4.12, can be seen as an identity between classes of diagrams. In turn, in light of a possible quantum field formulation of the problem, this can be thought as a consequence of some deep Ward identity of the corresponding field theory. Ward identities play a crucial role in quantum field theory. The analogy between KAM theory and quantum field theory has been widely stressed in the literature [20,5,12]; in particular the cancellations which assure the convergence of the perturbation series for maximal KAM tori are deeply related to a Ward identity, as shown in [5], which can be seen as a remarkable identity between classes of graphs. In the case studied in this paper, we have a similar situation, made fiddlier by the fact that we have to deal with nonconvergent series to be resummed, and it is well known that identities which are trivial on a formal level can turn out to be difficult to prove rigorously [30]. However, we expect a Ward identity to hold also in our case, so as to imply that $\varepsilon\partial_{\beta_0} G^{\mathcal{R}}(\varepsilon,\beta_0) = \mathcal{M}_\infty(0;\varepsilon,\beta_0)$. It would be interesting to confirm the expectation and to determine the Ward identity explicitly.

Given $x_0 \in \mathbb{R}$ and an interval $(a,b) \subset \mathbb{R}$ such that $x_0 \in (a,b)$, we call the half-neighbourhoods of x_0 the two intervals (a,x_0) and (x_0,b).

Lemma 4.15. *Assume $\mathcal{M}$ to satisfy Property 1. Then the implicit function equation $G^{\mathcal{R}}(\varepsilon,\beta_0) = 0$ admits a solution $\beta_0 = \beta_0(\varepsilon)$, such that $\beta_0(0) = \beta_0^*$. Moreover in a suitable half-neighbourhood of $\varepsilon = 0$, one has $\varepsilon\partial_{\beta_0} G^{\mathcal{R}}(\varepsilon,\beta_0(\varepsilon)) \le 0$.*

Proof. Property 1 allows us to write $G^{\mathcal{R}}(\varepsilon,\beta_0) = F_{\mathbf{0}}(\beta_0) + O(\varepsilon)$, so that by Hypothesis 2 one has $\partial^{\mathfrak{n}}_{\beta_0} G^{\mathcal{R}}(0,\beta_0^*) \neq 0$. Then there exist two half-neighbourhood V_-, V_+ of $\beta_0 = \beta_0^*$ such that $G^{\mathcal{R}}(0,\beta_0) > 0$ for $\beta_0 \in V_+$ and $G^{\mathcal{R}}(0,\beta_0) < 0$ for $\beta_0 \in V_-$. Hence, by continuity, for all $\beta_0 \in V_+$ there exists a neighbourhood $U_+(\beta_0)$ of $\varepsilon = 0$ such that $G^{\mathcal{R}}(\varepsilon,\beta_0) > 0$ for all $\varepsilon \in U_+(\beta_0)$ and, for the same reason, for all $\beta_0 \in V_-$ there exists a neighbourhood $U_-(\beta_0)$ of $\varepsilon = 0$ such that $G^{\mathcal{R}}(\varepsilon,\beta_0) < 0$ for all $\varepsilon \in U_-(\beta_0)$. Therefore, again by continuity, there exists a continuous curve $\beta_0 = \beta_0(\varepsilon)$ defined in a suitable neighbourhood $U = (-\bar{\varepsilon},\bar{\varepsilon})$ such that $\beta_0(0) = \beta_0^*$ and $G^{\mathcal{R}}(\varepsilon,\beta_0(\varepsilon)) \equiv 0$. Moreover, if $\partial^{\mathfrak{n}}_{\beta_0} G^{\mathcal{R}}(0,\beta_0^*) > 0$, then V_+, V_- are of the form (β_0^*, v_+) and (v_-,β_0^*) respectively, and therefore $\partial_{\beta_0} G^{\mathcal{R}}(c,\beta_0(c)) \ge 0$ for all $c \in U$. If on the contrary $\partial^{\mathfrak{n}}_{\beta_0} G^{\mathcal{R}}(0,\beta_0^*) < 0$, one has $V_+ = (v_+, \beta_0^*)$ and $V_- = (\beta_0^*, v_-)$, and then $\partial_{\beta_0} G^{\mathcal{R}}(c, \beta_0(c)) \le 0$ for all $c \in U$. Hence the assertion follows in both cases, again by Hypothesis 2. □

Remark 4.16. Note that Lemma 4.15 implies only continuity of the curve $\beta(\varepsilon)$; see also comments at the end of Sect. 7.

Remark 4.17. If $\mathcal{M}$ satisfies Property 1, one has $G^{\mathcal{R}}(\varepsilon, \beta_0) = [F(\boldsymbol{\omega} t, \beta_0 + b^{\mathcal{R}}(t; \varepsilon, \beta_0))]_{\mathbf{0}}$ and hence, if $\beta_0 = \beta_0(\varepsilon)$ is the solution referred to in Lemma 4.15, by Lemma 4.6 the function $\beta(t; \varepsilon) = \beta_0(\varepsilon) + b^{\mathcal{R}}(t; \varepsilon, \beta_0(\varepsilon))$ solves the equation of motion (1.1).

Remark 4.18. The results of this section are not sufficient to prove Theorem 2.1 because we have assumed – without proof – that Property 1 is satisfied. In Sect. 5 we shall show that Property 1 is found to be satisfied along a suitable continuous curve $\beta_0 = \overline{\beta}_0(\varepsilon)$ such that $G^{\mathcal{R}}(\varepsilon, \overline{\beta}_0(\varepsilon)) = 0$.

5. Convergence of the Resummed Series: Part 2

In this section we shall remove the assumption that the self-energies satisfy Property 1 of Definition 4.4 – see Remark 4.18. We shall proceed as follows. We sligthly modify the propagators by replacing the self-energies $\mathcal{M}_n(x; \varepsilon, \beta_0)$ with new quantities $\overline{\mathcal{M}}_n(x; \varepsilon, \beta_0)$ and we prove recursively that such quantities satisfy the symmetry properties of Lemma 4.10: this will imply that Property 1 holds. Then we shall check *a posteriori* that on a suitable curve $\beta_0 = \beta_0(\varepsilon)$ one has $\overline{\mathcal{M}}_n(x; \varepsilon, \beta_0) = \mathcal{M}_n(x; \varepsilon, \beta_0)$; moreover, thanks to the identity of Lemma 4.12, on such a curve also the bifurcation equation is satisfied, so that the function $\beta_0(\varepsilon) + b^{\mathcal{R}}(t; \varepsilon, \beta_0(\varepsilon))$, with $b^{\mathcal{R}}$ given by (3.12a), turns out to be well-defined and solve the equations of motion.

For all $n \ge 0$, define the C^∞ non-increasing functions ξ_n such that

$$\xi_n(x) = \begin{cases} 1, & x \le \alpha_{m_{n+1}}(\boldsymbol{\omega})^2/2^{12}, \\ 0, & x \ge \alpha_{m_{n+1}}(\boldsymbol{\omega})^2/2^{11}, \end{cases} \tag{5.1}$$

and set $\xi_{-1}(x) = 1$. Define recursively, for all $n \ge 0$, the propagators

$$\overline{\mathcal{G}}_n(x; \varepsilon, \beta_0) = \Psi_n(x) \left(x^2 - \overline{\mathcal{M}}_{n-1}(x; \varepsilon, \beta_0) \xi_{n-1}(\overline{\mathcal{M}}_{n-1}(0; \varepsilon, \beta_0)) \right)^{-1}, \tag{5.2}$$

with $\overline{\mathcal{M}}_{-1}(x; \varepsilon, \beta_0) = \varepsilon \partial_\beta F_{\mathbf{0}}(\beta_0)$, and for $n \ge 0$,

$$\overline{\mathcal{M}}_n(x; \varepsilon, \beta_0) = \overline{\mathcal{M}}_{n-1}(x; \varepsilon, \beta_0) + \chi_n(x) \overline{M}_n(x; \varepsilon, \beta_0), \tag{5.3}$$

where we have set

$$\overline{M}_n(x; \varepsilon, \beta_0) := \sum_{T \in \mathfrak{R}_n} \varepsilon^{k(T)} \overline{\mathscr{V}}_T(x; \varepsilon, \beta_0), \tag{5.4a}$$

$$\overline{\mathscr{V}}_T(x; \varepsilon, \beta_0) := \Bigg(\prod_{v \in N(T)} \mathcal{F}_v(\beta_0) \Bigg) \Bigg(\prod_{\ell \in L(T)} \overline{\mathcal{G}}_{n_\ell}(\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell; \varepsilon, \beta_0) \Bigg), \tag{5.4b}$$

with $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}_{\ell'_T}$. Set $\overline{\mathcal{M}} = \{\overline{\mathcal{M}}_n(x; \varepsilon, \beta_0)\}_{n \ge -1}$ and $\overline{\mathcal{M}}^\xi = \{\overline{\mathcal{M}}_n(x; \varepsilon, \beta_0) \xi_n (\overline{\mathcal{M}}_n(0; \varepsilon, \beta_0))\}_{n \ge -1}$.

Lemma 5.1. $\overline{\mathcal{M}}^{\xi}$ *satisfies Property 1.*

Proof. We shall prove that $\overline{\mathcal{M}}^{\xi}$ satisfies Property 1-p for all $p \geq 0$, by induction on p. Property 1-0 is trivially satisfied for ε small enough. Assume $\overline{\mathcal{M}}^{\xi}$ to satisfy Property 1-p. Then we can repeat – almost word by word – the proofs of Lemmas 4.8 and 4.10 (see also Remark 4.9) so as to obtain $\overline{\mathcal{M}}_p(x; \varepsilon, \beta_0) = \overline{\mathcal{M}}_p(0; \varepsilon, \beta_0) + O(\varepsilon^2 x^2)$, hence, by the definition of the function ξ_p, $\overline{\mathcal{M}}^{\xi}$ satisfies Property 1-$(p+1)$, and thence the assertion follows. □

Set

$$\overline{\mathscr{V}}(\theta; \varepsilon, \beta_0) := \left(\prod_{v \in N(\theta)} \mathcal{F}_v(\beta_0) \right) \left(\prod_{\ell \in L(\theta)} \overline{\mathcal{G}}_{n_\ell}(\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell; \varepsilon, \beta_0) \right), \tag{5.5a}$$

$$\overline{b}_{\boldsymbol{\nu}}^{[k]}(\varepsilon, \beta_0) := \sum_{\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}} \overline{\mathscr{V}}(\theta; \varepsilon, \beta_0), \tag{5.5b}$$

and define

$$\overline{b}(t, \varepsilon, \beta_0) = \sum_{k \geq 1} \varepsilon^k \overline{b}^{[k]}(\varepsilon, \beta_0) = \sum_{k \geq 1} \varepsilon^k \sum_{\boldsymbol{\nu} \in \mathbb{Z}^d_*} e^{i\boldsymbol{\nu} \cdot \boldsymbol{\omega} t} \overline{b}_{\boldsymbol{\nu}}^{[k]}(\varepsilon, \beta_0). \tag{5.6}$$

Note that, by (the proof of) Lemma 4.5 the series (5.6) converges. Define also

$$\overline{\mathcal{M}}_\infty(x; \varepsilon, \beta_0) := \lim_{n \to \infty} \overline{\mathcal{M}}_n(x; \varepsilon, \beta_0), \tag{5.7}$$

and note that, by Lemma 5.1 the limit in (5.7) is well defined and it is C^∞ in both ε and β_0. Introduce the C^∞ functions $\overline{G}(\varepsilon, \beta_0)$ such that $\overline{\mathcal{M}}_\infty(0; \varepsilon, \beta_0) = \varepsilon \partial_{\beta_0} \overline{G}(\varepsilon, \beta_0)$ and $\overline{G}(0, \beta_0^*) = 0$, and for any such function consider the implicit function equation

$$\overline{G}(\varepsilon, \beta_0) = 0. \tag{5.8}$$

Recall the definition of half-neighbourhood after Remark 4.14.

Lemma 5.2. *The implicit function equation (5.8) admits a solution $\beta_0 = \overline{\beta}_0(\varepsilon)$ such that $\overline{\beta}_0(0) = \beta_0^*$. Moreover in a suitable half-neighbourhood of $\varepsilon = 0$, one has $\varepsilon \partial_{\beta_0} \overline{G}(\varepsilon, \overline{\beta}_0(\varepsilon)) \leq 0$.*

Proof. By construction, all the functions $\overline{G}(\varepsilon, \beta_0)$ are smooth and of the form $\overline{G}(\varepsilon, \beta_0) = F_{\mathbf{0}}(\beta_0) + O(\varepsilon)$. Then the result follows straightforward from (the proof of) Lemma 4.15. □

Lemma 5.3. *Let $\beta_0 = \overline{\beta}_0(\varepsilon)$ be the solution referred to in Lemma 5.2. Then one has $\xi_n(\overline{\mathcal{M}}_n(0; \varepsilon, \overline{\beta}_0(\varepsilon))) \equiv 1$ for all $n \geq 0$, in a suitable half-neighbourhood of $\varepsilon = 0$.*

Proof. If $\beta_0 = \overline{\beta}_0(\varepsilon)$, one has $\overline{\mathcal{M}}_\infty(0; \varepsilon, \overline{\beta}_0(\varepsilon)) = \varepsilon \partial_{\beta_0} \overline{G}(\varepsilon, \overline{\beta}_0(\varepsilon)) \leq 0$, by Lemma 5.2 in a suitable half-neighbourhood of $\varepsilon = 0$. Hence, as the bound (4.4a) holds also for $\overline{M}_n(x; \varepsilon, \beta_0)$, one has

$$\begin{aligned} \overline{\mathcal{M}}_n(0; \varepsilon, \overline{\beta}_0(\varepsilon)) &\leq \overline{\mathcal{M}}_n(0; \varepsilon, \overline{\beta}_0(\varepsilon)) - \overline{\mathcal{M}}_\infty(0; \varepsilon, \overline{\beta}_0(\varepsilon)) \\ &\leq \sum_{p \geq n+1} |\overline{M}_p(0; \varepsilon, \overline{\beta}_0(\varepsilon))| \leq 2K_1 \varepsilon^2 e^{-K_2 2^{m_{n+1}}} \leq \frac{\alpha_{m_{n+1}}(\boldsymbol{\omega})^2}{2^{13}}, \end{aligned} \tag{5.9}$$

so that the assertion follows by the definition of ξ_n. □

Lemma 5.4. *One can choose $\overline{G}(\varepsilon, \beta_0)$ such that $G^{\mathcal{R}}(\varepsilon, \overline{\beta}_0(\varepsilon)) = \overline{G}(\varepsilon, \overline{\beta}_0(\varepsilon)) = 0$. In particular $\beta(t; \varepsilon) = \overline{\beta}_0(\varepsilon) + b^{\mathcal{R}}(t; \varepsilon, \overline{\beta}_0(\varepsilon))$ defined in (3.12) solves the equation of motion (1.1).*

Proof. For any $\overline{G}(\varepsilon, \beta_0)$ as above there is a curve $\overline{\beta}_0(\varepsilon)$ along which $\mathcal{M} = \overline{\mathcal{M}} = \overline{\mathcal{M}}^{\xi}$, and hence $\mathcal{M}$ satisfies Property 1 and $\overline{G}(\varepsilon, \overline{\beta}_0(\varepsilon)) = 0$. By Lemma 4.12 also $G^{\mathcal{R}}(\varepsilon, \beta_0)$ is one of such primitives and then the assertion follows. □

Remark 5.5. Note that without Lemma 4.12 we were able to prove only the existence of curves on which the solution of the range equation (2.4a) is well-defined. On the other hand Lemma 4.12 guarantees that the solution of the bifurcation equation (2.4b) is one of such curves, say $\overline{\beta}_0(\varepsilon)$, so that the function $\beta(t; \varepsilon) = \overline{\beta}_0(\varepsilon) + b^{\mathcal{R}}(t; \varepsilon, \overline{\beta}_0(\varepsilon))$ given by (3.12) is well defined and solves the equation of motion (1.1).

6. Proof of Theorem 2.2

If $F_{\mathbf{0}}(\beta_0)$ vanishes identically, let us come back to the formal expansion (3.4) of $G(\varepsilon, \beta_0)$, where $G^{(0)}(\beta_0) = F_{\mathbf{0}}(\beta_0) \equiv 0$ by hypothesis.

Assume first that there exists $k_0 \in \mathbb{N}$ such that all functions $G^{(k)}(\beta_0)$ are identically zero for $0 \leq k \leq k_0 - 1$, while $G^{(k_0)}(\beta_0)$ is not identically vanishing. Then we can write

$$G(\varepsilon, \beta_0) = \varepsilon^{k_0} \left(G^{(k_0)}(\beta_0) + G^{(>k_0)}(\varepsilon, \beta_0) \right), \tag{6.1}$$

with $G^{(>k_0)}(\varepsilon, \beta_0) = O(\varepsilon)$, and we can solve the equation of motion up to order k_0 without fixing the parameter β_0. On the other hand $G^{(k_0)}$ is the β_0-derivative of the time-average of the k_0-th order of the Lagrangian $g^{(k_0)}$ computed along a solution of the range equation, which is analytic and periodic: since it is not identically constant, it admits at least one maximum $\bar{\beta}_0'$ and one minimum $\bar{\beta}_0''$, so that one can assume the following

Hypothesis 3. *β_0^* is a zero of order $\bar{\mathfrak{n}}$ for $G^{(k_0)}(\beta_0)$ with $\bar{\mathfrak{n}}$ odd, and $\varepsilon^{k_0+1}\partial_{\beta_0}^{\bar{\mathfrak{n}}} G^{(k_0)}(\beta_0^*) < 0$.*

Indeed, if k_0 is even one can choose $\beta_0^* = \bar{\beta}_0'$ for $\varepsilon > 0$, and $\beta_0^* = \bar{\beta}_0''$ for $\varepsilon < 0$; if k_0 is odd we have to fix $\beta_0^* = \bar{\beta}_0'$: in both cases Hypothesis 3 is satisfied.

Then one can adapt the proof in the previous sections to cover this case. Namely, as the formal expansion of $G^{\mathcal{R}}$ coincide with that of G, one sets $G^{\mathcal{R}}(\varepsilon, \beta_0) =: \varepsilon^{k_0} G_*(\varepsilon, \beta_0)$ and hence, if $\mathcal{M}$ satisfies Property 1,

$$\mathcal{M}_\infty(0; \varepsilon, \beta_0) = \varepsilon^{k_0+1} \partial_{\beta_0} G_*(\varepsilon, \beta_0). \tag{6.2}$$

On the other hand, Hypothesis 3 and Lemma 4.15 guarantee the existence of a continuous curve $\beta_0(\varepsilon)$ such that $\beta_0(0) = \beta_0^*$, $G_*(\varepsilon, \beta_0(\varepsilon)) \equiv 0$ and if k_0 is even then $\varepsilon^{k_0+1}\partial_{\beta_0} G_*(\varepsilon, \beta_0(\varepsilon)) \leq 0$ in a suitable half-neighbourhood of $\varepsilon = 0$, while if k_0 is odd and β_0^* is a maximum for $g^{(k_0)}$, then $\partial_{\beta_0} G_*(\varepsilon, \beta_0(\varepsilon)) \leq 0$ in a whole neighbourhood of $\varepsilon = 0$. Then one can reason as in Sect. 5 to obtain the result.

Finally, assume $G^{(k)}(\beta_0) \equiv 0$ for all $k \geq 0$. We shall see that no resummation is necessary in that case: this situation is reminiscent of the "null-renormalisation" case considered in [26] when studying the stability problem for Hill's equation with a quasi-periodic perturbation.

We define trees and clusters according to the definitions previously done. On the other hand, we slight change the definition of self-energy clusters. Namely, a cluster T on scale $n \geq 0$ with only one entering line ℓ'_T and one exiting line ℓ_T and with $\boldsymbol{\nu}_{\ell_T} = \boldsymbol{\nu}_{\ell'_T}$, is called a self-energy cluster if one has $\boldsymbol{\nu}^0_\ell \neq \mathbf{0}$ for all $\ell \in \mathcal{P}_T$. The definition of self-energy cluster does not change for the self-energy cluster on scale -1. We denote by $\Theta_{k,\boldsymbol{\nu}}$ the set of trees with order k and momentum $\boldsymbol{\nu}$ as in Sect. 3, and by $\mathfrak{S}^k_n$ the set of (non-renormalised) self-energy clusters with order k and scale n; note that self-energy clusters are allowed both in $\Theta_{k,\boldsymbol{\nu}}$ and in $\mathfrak{S}^k_n$.

For any subgraph S of any tree $\theta \in \Theta_{k,\boldsymbol{\nu}}$ and for any $T \in \mathfrak{S}^k_n$, define the (non-renormalised) value of S and T as in (3.10) and (3.9c) respectively, but with the (undressed) propagators defined as

$$\mathcal{G}_{n_\ell}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell) := \begin{cases} \dfrac{\Psi_{n_\ell}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell)}{\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell^2}, & n_\ell \geq 0, \\ 1, & n_\ell = -1. \end{cases} \tag{6.3}$$

Note that now the values of trees and self-energy clusters do not depend on ε, and they depend on β_0 only through the node factors. From now on we do not write explicitly the dependence on β_0 to lighten the notations. For all $k \geq 1$, define

$$b^{(k)}_{\boldsymbol{\nu}} := \sum_{\theta\in\Theta_{k,\boldsymbol{\nu}}} \mathscr{V}(\theta), \qquad G^{(k-1)} := \sum_{\theta\in\Theta_{k,\mathbf{0}}} \mathscr{V}(\theta), \tag{6.4a}$$

$$M^{(k)}_n(x) := \sum_{T\in\mathfrak{S}^k_n} \mathscr{V}_T(x), \qquad \mathcal{M}^{(k)}_n(x) := \sum_{p=0}^{n} M^{(k)}_p(x), \qquad n \geq -1, \tag{6.4b}$$

$$\mathcal{M}^{(k)}_\infty(x) := \lim_{n\to\infty} \mathcal{M}^{(k)}_n(x). \tag{6.4c}$$

The coefficients (6.4a) coincide with (3.2) and (3.5), as is easy to check; in particular, for all $k \geq 1$ one has $\sum_{\theta\in\Theta_{k,\mathbf{0}}} \mathscr{V}(\theta) \equiv 0$ by assumption.

Remark 6.1. One has $\mathfrak{S}^k_{-1} = \mathfrak{S}^1_n = \emptyset$ for $k \geq 2$ and $n \geq 0$. On the other hand $|\mathfrak{S}^1_{-1}| = 1$ and $\mathscr{V}_T(x) = \partial_{\beta_0} F_{\mathbf{0}} \equiv 0$ if T is the self-energy cluster in $\mathfrak{S}^1_{-1}$; see Remark 3.4. Hence $M^{(1)}_n(x) = \mathcal{M}^{(1)}_n(x) = \mathcal{M}^{(1)}_\infty(x) = M^{(k)}_{-1} = \mathcal{M}^{(k)}_{-1} \equiv 0$ for all $n \geq -1,\ k \geq 1$.

Given a tree θ with $\mathscr{V}(\theta) \neq 0$, we shall say that a line $\ell \in L(\theta)$ is *resonant* if it is the exiting line of a self-energy cluster T, otherwise we shall say that ℓ is *non-resonant*. For any subgraph T of any tree $\theta \in \Theta_{k,\boldsymbol{\nu}}$, denote by $\mathfrak{N}^*_n(T)$ the number of non-resonant lines on scale $\geq n$ in T, and set $K(T)$ as in (4.2). Define also, for any line $\ell \in L(T)$, $\zeta_\ell := \min\{n \in \mathbb{Z}_+ : \Psi_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell) \neq 0\}$ and denote by $\mathfrak{N}^\bullet_n(T)$ the number of non-resonant lines $\ell \in L(T)$ such that $\zeta_\ell \geq n$. Then we can prove the analogue of Lemmas 4.1 and 4.2, namely the following results.

Lemma 6.2. *For any $\theta \in \Theta_{k,\boldsymbol{\nu}}$ such that $\mathscr{V}(\theta) \neq 0$ one has $\mathfrak{N}^\bullet_n(\theta) \leq 2^{-(m_n-2)} K(\theta)$, for all $n \geq 0$.*

Lemma 6.3. *For any $T \in \mathfrak{S}^k_n$ such that $\mathscr{V}_T(x) \neq 0$ one has $K(T) \geq 2^{m_n-1}$ and $\mathfrak{N}^\bullet_p(T) \leq 2^{-(m_p-2)} K(T)$, for all $0 \leq p \leq n$.*

We omit the proofs of the two results above as it would be essentially a repetition of those for Lemmas 4.1 and 4.2, respectively. Note that, since self-energy clusters are now allowed, for the proof of Lemma 6.3 one needs that the momenta of the lines in $\mathcal{P}_T$ are different from those of the external lines: this explains the new definition of self-energy clusters.

In light of Lemmas 6.2 and 6.3, although one has the 'good bound' $1/x^2$ for the propagators, one cannot prove the convergence of the power series (3.1) as done in Lemma 4.5, because we do not have any bound for the number of resonant lines, which in principle can accumulate 'too much'. In fact, we need a gain factor proportional to $(\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell)^2$ for each resonant line ℓ.

Lemma 6.4. *For all $n \geq 0$ and for all $k \geq 2$ one has $\partial_x M_n^{(k)}(0) = 0$ and hence $\partial_x \mathcal{M}_n^{(k)}(0) = 0$.*

Proof. As the propagators are trivially even in the momenta, one can repeat (almost word by word) the proof of Lemma 4.10 so as to obtain the result. □

Lemma 6.5. *One has $\mathcal{M}_\infty^{(k)}(0) \equiv 0$ for all $k \geq 2$.*

Proof. One has (see also Remark 4.13) $\partial_{\beta_0} G^{(k-1)} \equiv \mathcal{M}_\infty^{(k)}(0)$ so that the assertion follows. □

Lemma 6.6. *For all $k \geq 1$ one has $|\mathcal{M}_n^{(k)}(x)|\Psi_{n+1}(x) \leq C^k x^2 \Psi_{n+1}(x)$ for some $C > 0$.*

The latter result, proved in Appendix G, implies the convergence of the series (3.1). Indeed, for any tree θ, consider the set $\mathcal{T}_1(\theta)$ of its maximal self-energy clusters and sum together the values of the trees obtained by replacing each $T \in \mathcal{T}_1(\theta)$ with any self-energy cluster with the same order and scale $< \min\{n_{\ell_T}, n_{\ell'_T}\}$. Then the product of the propagators of the non-resonant lines outside $\mathcal{T}_1(\theta)$ is bounded thanks to Lemma 6.3, while the product of the propagators of the resonant lines exiting any self-energy cluster $T \in \mathcal{T}_1(\theta)$ times the product of the corresponding self-energy values is bounded through Lemma 6.6.

Remark 6.7. We have obtained the convergence of the power series (3.1) and (3.4) for any β_0 and any ε small enough. Hence, in this case, the response solution turns out to be analytic in both ε and β_0.

Remark 6.8. Note that the problem under study has analogies with the problem considered in [24]. In that case, the resummation adds to the small divisor $\mathrm{i}\boldsymbol{\omega} \cdot \boldsymbol{\nu}$ a quantity $-\varepsilon(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2 + \mathcal{M}_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu}; \varepsilon)$, and one can prove that $\mathcal{M}_n(x, \varepsilon)$ is smooth in x and it is real at $x = 0$, so that the dressed propagator is proportional to $1/(\mathrm{i}\boldsymbol{\omega}\cdot\boldsymbol{\nu} - \varepsilon(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2 + \mathcal{M}_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu}; \varepsilon))$, and hence can be bounded essentially as the undressed one. In the present case, both the small divisor $(\boldsymbol{\omega} \cdot \boldsymbol{\nu})^2$ and the correction are real, but, up to negligible corrections, they turn out to have the same sign (for a suitable choice of β_0^*), so that once more the dressed propagator can be bounded as the undressed one.

7. Conclusions

In this paper we proved the existence of response solutions to (1.1) for ε small enough and $\boldsymbol{\omega}$ satisfying the Bryuno condition, with no other assumption on the perturbation than analyticity. As we said in the Introduction, the result can be interpreted as a result on persistence of lower-dimensional tori in quasi-integrable systems. As far as we know, the only other result in the literature on the persistence of lower-dimensional tori with no assumption on the perturbation is due to Cheng [8]. He proved that, for convex unperturbed Hamiltonians, given any $(d+1)$-dimensional unperturbed resonant torus on which the flow is quasi-periodic with frequency vector $\boldsymbol{\omega} \in \mathbb{R}^d$ satisfying the standard Diophantine condition, there exists at least one d-dimensional submanifold of the resonant torus persisting under small perturbations and still carrying a quasi-periodic flow with the same frequency vector. We have proved a result of the same kind for Eq. (1.1), that is the existence of at least one response solution for ε small enough – see Theorem 2.2 in Sect. 2.

Of course, if the one hand we can look at the problem of existence of response solutions to (1.1) as a problem of persistence of d-dimensional tori in a system with $d + 1$ degress of freedom, on the other hand d degrees of freedom trivially evolve according to the first equation in (1.4): in fact (1.1) is actually a 1-degree of freedom system with an arbitrary quasi-periodic perturbation. However, even though the Hamiltonian (1.3) can be seen as a simplified model for the problem of lower-dimensional tori, we think that our result can be of interest by its own for the following reasons.

First of all, Cheng's result does not directly apply, since the convexity property he requires is obviously not satisfied by the Hamiltonian (1.3). Very likely Cheng's method could be extended to the case (1.3): indeed the anisochrony condition is expected to be removable (and in fact it is, as our result yields) when the perturbation depends only on the angle variables – in the case of maximal tori this has been explicitly shown; see for instance [15,29]. However, at least, a proof would require some adaptation from Cheng's paper. Moreover, just because of its simplicity, the model is particularly suited to point out the main issues of the proof, avoiding all aspects that would add only technical intricacies without shedding further light on the problem of persistence of resonant tori. Finally, our method is completely different from Cheng's: in fact one of the main motivations for us was to provide an alternative approach to the problem.

We also mention that we allow a weaker Diophantine condition on the frequency vector, i.e. the Bryuno condition. Recently the Bryuno condition has been widely studied in the theory of small divisors problems. Its relevance is also related to the possibility of describing properties of the analyticity domain, such as the radius of convergence, of the solutions in terms of the Bryuno function; this has been explictly shown in some simple cases, such as the Siegel problem [37], the semi-standard map [11] and the standard map [2]. It is generally believed that any analytical KAM-type problem that can be solved under the standard Diophantine condition, can also be solved using the Bryuno condition.

Note also that, in contrast to the case of periodic perturbations, the quasi-periodic solution to (1.1) is not expected to be analytic in ε nor is some fractional power of ε. Already in the non-degenerate case the solution has been proved only to be C^∞ smooth in ε [17], and analyticity is very unlikely. In the degenerate case, under some further assumptions on the forcing one obtains smoothness in some fractional power of ε [18]. However, in general no more than continuity in ε can be proved. This is ultimately related to the implicit function problem (2.6): the best we can do is to show that there exists a continuous solution $\beta_0(\varepsilon)$ to (2.6). Furthermore the argument is not constructive (a different situation arises in the case of the forced strongly dissipative systems studied in [24,25], where the proof of existence of response solutions has been made fully constructive and C^∞ smoothness follows).

This not surprising: the same non-constructive feature of the proof appears in Cheng's approach.

As emerges from the proof, very likely, the only case in which the quasi-periodic solution is analytic is when the infinitely many conditions $G^{(k)}(\beta_0) \equiv 0$ are satisfied. Of course this is a highly non-generic situation: moreover in that case, the entire resonant torus persists – see Remark 6.7.

A. Proof of Lemma 4.1

First of all we note that if $\mathfrak{N}_n(\theta) \geq 1$, then $K(\theta) \geq 2^{m_n-1}$. Indeed, if a line ℓ has scale $n_\ell \geq n$, then

$$|\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell| \leq \frac{1}{8}\alpha_{m_{n-1}}(\boldsymbol{\omega}) < \frac{1}{4}\alpha_{m_{n-1}+p_{n-1}}(\boldsymbol{\omega}) = \frac{1}{4}\alpha_{m_n-1}(\boldsymbol{\omega}) < \alpha_{m_n-1}(\boldsymbol{\omega}),$$

and hence, by definition of $\alpha_m(\boldsymbol{\omega})$, one has $K(\theta) \geq |\boldsymbol{\nu}_\ell| \geq 2^{m_n-1}$. Now we prove the bound $\mathfrak{N}_n(\theta) \leq \max\{2^{-(m_n-2)}K(\theta) - 1, 0\}$ by induction on the order.

If the root line of θ has scale $n_{\ell_\theta} < n$ then the bound follows by the inductive hypothesis. If $n_{\ell_\theta} \geq n$, call $\ell_1, \ldots, \ell_r$ the lines with scale $\geq n$ closest to ℓ_θ (that is such that $n_{\ell'} < n$ for all lines $\ell' \in \mathcal{P}(\ell_\theta, \ell_i)$, $i = 1, \ldots, r$); see Fig. 2. If $r = 0$ then $\mathfrak{N}_n(\theta) = 1$ and $|\boldsymbol{\nu}| \geq 2^{m_n-1}$, so that the bound follows. If $r \geq 2$ the bound follows once more by the inductive hypothesis. If $r = 1$, then ℓ_1 is the only entering line of a cluster T which is not a self-energy cluster as $\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$, and hence $\boldsymbol{\nu}_{\ell_1} \neq \boldsymbol{\nu}$. But then

$$|\boldsymbol{\omega} \cdot (\boldsymbol{\nu} - \boldsymbol{\nu}_{\ell_1})| \leq |\boldsymbol{\omega} \cdot \boldsymbol{\nu}| + |\boldsymbol{\omega} \cdot \boldsymbol{\nu}_{\ell_1}| \leq \frac{1}{4}\alpha_{m_{n-1}}(\boldsymbol{\omega}) < \frac{1}{2}\alpha_{m_{n-1}+p_{n-1}}(\boldsymbol{\omega}) = \frac{1}{2}\alpha_{m_n-1}(\boldsymbol{\omega}),$$

as both ℓ_θ and ℓ_1 are on scale $\geq n$, so that one has $K(T) \geq |\boldsymbol{\nu} - \boldsymbol{\nu}_{\ell_1}| \geq 2^{m_n-1}$. Now, call θ_1 the subtree of θ with root line ℓ_1. Then one has $\mathfrak{N}_n(\theta) = 1 + \mathfrak{N}_n(\theta_1) \leq 1 + \max\{2^{-(m_n-2)}K(\theta_1) - 1, 0\}$, so that $\mathfrak{N}_n(\theta) \leq 2^{-(m_n-2)}(K(\theta) - K(T)) \leq 2^{-(m_n-2)}K(\theta) - 1$, again by induction.

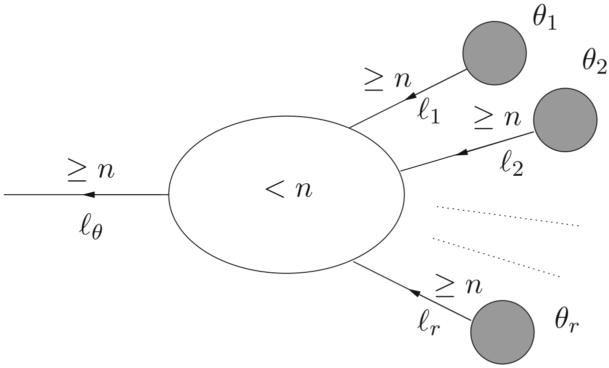

Fig. 2. Construction used in the proof of Lemma 4.1 when $n_{\ell_\theta} \geq n$

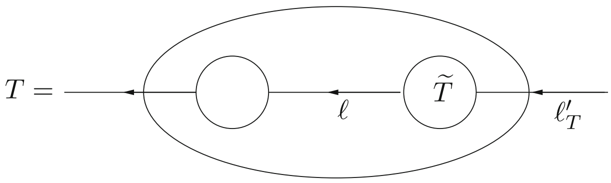

Fig. 3. Construction used to prove $K(T) \geq 2^{m_n-1}$ when there is a line $\ell \in \mathcal{P}_T$ on scale n

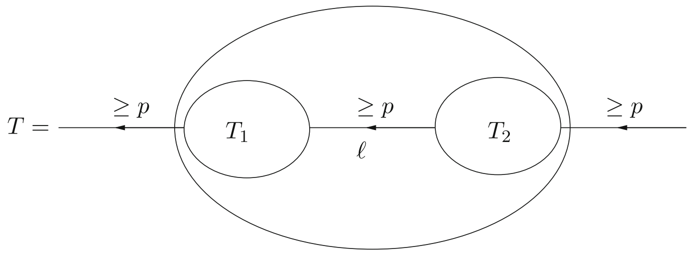

Fig. 4. Construction used to prove Lemma 4.2

B. Proof of Lemma 4.2

We first prove that for all $n \geq 0$ and all $T \in \mathfrak{R}_n$, one has $K(T) \geq 2^{m_n-1}$. In fact if $T \in \mathfrak{R}_n$ then T contains at least a line on scale n. If there is $\ell \in L(T) \setminus \mathcal{P}_T$ with $n_\ell = n$, then

$$|\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell| < \frac{1}{8} \alpha_{m_{n-1}}(\boldsymbol{\omega}) < \alpha_{m_n-1}(\boldsymbol{\omega}),$$

and hence $K(T) \geq |\boldsymbol{\nu}_\ell| > 2^{m_n-1}$. Otherwise, let $\ell \in \mathcal{P}_T$ be the line on scale n which is closest to ℓ'_T. Call $\widetilde{T}$ the subgraph (actually the cluster) consisting of all lines and nodes of T preceding ℓ; see Fig. 3. Then $\boldsymbol{\nu}_\ell \neq \boldsymbol{\nu}_{\ell'_T}$, otherwise $\widetilde{T}$ would be a self-energy cluster. Therefore $K(T) > |\boldsymbol{\nu}_\ell - \boldsymbol{\nu}_{\ell'_T}| > 2^{m_n-1}$ as both ℓ, ℓ'_T are on scale $\geq n$.

Given a tree θ, call $\mathcal{C}(n, p)$ the set of renormalised subgraphs T of θ with only one entering line ℓ'_T and one exiting line ℓ_T both on scale $\geq p$, such that $L(T) \neq \emptyset$ and $n_\ell \leq n$ for any $\ell \in L(T)$. Note that $\mathfrak{R}_n \subset \mathcal{C}(n, p)$ for all $n, p \geq 0$. We prove that $\mathfrak{N}_p(T) \leq \max\{K(T)2^{-(m_p-2)} - 1, 0\}$ for all $0 \leq p \leq n$ and all $T \in \mathcal{C}(n, p)$. The proof is by induction on the order. Call $N(\mathcal{P}_T)$ the set of nodes in T connected by lines in $\mathcal{P}_T$. If all lines in $\mathcal{P}_T$ are on scale $< p$, then $\mathfrak{N}_p(T) = \mathfrak{N}_p(\theta_1) + \cdots + \mathfrak{N}_p(\theta_r)$ if $\theta_1, \ldots, \theta_r$ are the subtrees with root line entering a node in $N(\mathcal{P}_T)$, and hence the bound follows from (the proof of) Lemma 4.1. If there exists a line $\ell \in \mathcal{P}_T$ on scale $\geq p$, call T_1 and T_2 the subgraphs of T such that $L(T) = \{\ell\} \cup L(T_1) \cup L(T_2)$, and note that if $L(T_1), L(T_2) \neq \emptyset$, then $T_1, T_2 \in \mathcal{C}(n, p)$; see Fig. 4.

Hence, by the inductive hypothesis one has $\mathfrak{N}_p(T) = 1 + \mathfrak{N}_p(T_1) + \mathfrak{N}_p(T_2) \leq 1 + \max\{2^{-(m_p-2)}K(T_1) - 1, 0\} + \max\{2^{-(m_p-2)}K(T_2) - 1, 0\}$. If both $\mathfrak{N}_p(T_1), \mathfrak{N}_p(T_2)$ are zero the bound trivially follows as $K(T) \geq 2^{m_p-1}$, while if both are non-zero one has $\mathfrak{N}_p(T) \leq 2^{-(m_p-2)}(K(T_1) + K(T_2)) - 1 = 2^{-(m_p-2)}K(T) - 1$. Finally if only one is zero, say $\mathfrak{N}_p(T_1) \neq 0$ and $\mathfrak{N}_p(T_2) = 0$, then T_2 is a cluster and hence $\boldsymbol{\nu}_\ell \neq \boldsymbol{\nu}_{\ell'_T}$, which implies $K(T_2) \geq 2^{m_p-1}$, so that $\mathfrak{N}_p(T) \leq 2^{-(m_p-2)}K(T_1) = 2^{-(m_p-2)}K(T) - 2^{-(m_p-2)}K(T_2) \leq 2^{-(m_p-2)}K(T) - 1$. On the other hand, either $T_2 \in \mathcal{C}(n, p)$ or it is constituted by only one node v with $\boldsymbol{\nu}_v \neq \mathbf{0}$, so that $K(T_2) > 2^{m_p-1}$ in both cases. The same argument can be used in the case $\mathfrak{N}_p(T_1) = 0$ and $\mathfrak{N}_p(T_2) \neq 0$.

C. Proof of Lemma 4.5

Let $\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$. The analyticity of f, hence of F, implies that there exist positive constants F_1, F_2, ξ such that for all $v \in N(\theta)$ one has

$$|\mathcal{F}_v(\beta_0)| = \frac{1}{s_v!}|\partial_\beta^{s_v} F_{\boldsymbol{\nu}_v}(\beta_0)| \le F_1 F_2^{s_v} e^{-\xi|\boldsymbol{\nu}_v|}. \tag{C.1}$$

Moreover Property 1 implies $|\mathcal{G}_n(x;\varepsilon,\beta_0)| \le c_0 \alpha_{m_n}(\boldsymbol{\omega})^{-2}$ for all $n \ge 0$ and for some positive constant c_0, and hence by Lemma 4.1 one can bound

$$\begin{aligned}
\prod_{\ell \in L(\theta)} |\mathcal{G}_{n_\ell}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\varepsilon,\beta_0)| &\le \prod_{n\ge 0}\left(\frac{c_0}{\alpha_{m_n}(\boldsymbol{\omega})^2}\right)^{\mathfrak{N}_n(\theta)} \\
&\le \left(\frac{c_0}{\alpha_{m_{n_0}}(\boldsymbol{\omega})^2}\right)^{k} \prod_{n\ge n_0+1}\left(\frac{c_0}{\alpha_{m_n}(\boldsymbol{\omega})^2}\right)^{\mathfrak{N}_n(\theta)} \\
&\le \left(\frac{c_0}{\alpha_{m_{n_0}}(\boldsymbol{\omega})^2}\right)^{k} \exp\left(8K(\theta)\sum_{n\ge n_0+1}\frac{1}{2^{m_n}}\log\frac{c_0^{1/2}}{\alpha_{m_n}(\boldsymbol{\omega})}\right) \le D^k(n_0)\exp(\xi(n_0)K(\theta)),
\end{aligned}$$

with

$$D(n_0) = \frac{c_0}{\alpha_{m_{n_0}}(\boldsymbol{\omega})^2}, \qquad \xi(n_0) = 8\sum_{n\ge n_0+1}\frac{1}{2^{m_n}}\log\frac{c_0^{1/2}}{\alpha_{m_n}(\boldsymbol{\omega})}.$$

Then, by Hypothesis 1, one can choose n_0 such that $\xi(n_0) \le \xi/2$. The sum over the other labels is bounded by a constant to the power k, and hence one can bound

$$\sum_{\theta\in\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}} |\mathscr{V}(\theta;\varepsilon,\beta_0)| \le C_0 C_1^k e^{-\xi|\boldsymbol{\nu}|/2},$$

for some constants C_0, C_1, and this is enough to prove the assertion.

D. Proof of Lemma 4.6

We shall prove that, the function $b^{\mathcal{R}}$ defined in (3.12) satisfies the equation of motion (2.4a), i.e. we shall check that $b^{\mathcal{R}} = \varepsilon g F(\boldsymbol{\omega}t, \beta_0 + b^{\mathcal{R}})$, where g is the pseudo-differential operator with kernel $g(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) = 1/(\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2$. We can write the Fourier coefficients of $b^{\mathcal{R}}$ as

$$b^{\mathcal{R}}_{\boldsymbol{\nu}} = \sum_{n\ge 0} b^{[n]}_{\boldsymbol{\nu}}, \qquad b^{[n]}_{\boldsymbol{\nu}} = \sum_{k\ge 1}\varepsilon^k \sum_{\theta\in\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}(n)} \mathscr{V}(\theta;\varepsilon,\beta_0), \tag{D.1}$$

where $\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}(n)$ is the subset of $\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}$ such that $n_{\ell_\theta} = n$.

Using Remark 3.3 and Lemma 4.5, in Fourier space one can write

$$
\begin{aligned}
g(\boldsymbol{\omega}\cdot\boldsymbol{\nu})[\varepsilon F(\boldsymbol{\omega}t,\beta_0+b^{\mathcal{R}})]_{\boldsymbol{\nu}} &= g(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\sum_{n\geq 0}\Psi_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu})[\varepsilon F(\boldsymbol{\omega}t,\beta_0+b^{\mathcal{R}})]_{\boldsymbol{\nu}}\\
&= g(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\sum_{n\geq 0}\Psi_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu})(\mathcal{G}_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0))^{-1}\mathcal{G}_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)[\varepsilon F(\boldsymbol{\omega}t,\beta_0+b^{\mathcal{R}})]_{\boldsymbol{\nu}}\\
&= g(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\sum_{n\geq 0}\Big((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2-\mathcal{M}_{n-1}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\Big)\mathcal{G}_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)[\varepsilon F(\boldsymbol{\omega}t,\beta_0+b^{\mathcal{R}})]_{\boldsymbol{\nu}}\\
&= g(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\sum_{n\geq 0}\Big((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2-\mathcal{M}_{n-1}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\Big)\sum_{k\geq 1}\varepsilon^k\sum_{\theta\in\overline{\Theta}^{\mathcal{R}}_{k,\boldsymbol{\nu}}(n)}\mathscr{V}(\theta;\varepsilon,\beta_0),
\end{aligned}
$$

where $\overline{\Theta}^{\mathcal{R}}_{k,\boldsymbol{\nu}}(n)$ differs from $\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}}(n)$ as it contains also trees θ which have one self-energy cluster with exiting line ℓ_θ. If we separate the trees containing such self-energy cluster from the others, we obtain

$$
\begin{aligned}
[\varepsilon F(\boldsymbol{\omega}t,\beta_0+b^{\mathcal{R}})]_{\boldsymbol{\nu}} &= \sum_{n\geq 0}\Big((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2-\mathcal{M}_{n-1}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\Big)b^{[n]}_{\boldsymbol{\nu}}+\sum_{n\geq 0}\Psi_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\\
&\quad\times\sum_{p\geq n}\sum_{q=-1}^{n-1}M_q(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)b^{[p]}_{\boldsymbol{\nu}}+\sum_{n\geq 1}\Psi_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\sum_{p=0}^{n-1}\sum_{q=-1}^{p-1}M_q(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)b^{[p]}_{\boldsymbol{\nu}}\\
&= \sum_{n\geq 0}\Big((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2-\mathcal{M}_{n-1}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\Big)b^{[n]}_{\boldsymbol{\nu}}\\
&\quad+\sum_{p\geq 0}\left(\sum_{q=-1}^{p-1}M_q(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\sum_{n\geq q+1}\Psi_n(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\right)b^{[p]}_{\boldsymbol{\nu}}\\
&= \sum_{n\geq 0}\Big((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2-\mathcal{M}_{n-1}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\Big)b^{[n]}_{\boldsymbol{\nu}}\\
&\quad+\sum_{n\geq 0}\left(\sum_{q=-1}^{n-1}M_q(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\chi_q(\boldsymbol{\omega}\cdot\boldsymbol{\nu})\right)b^{[n]}_{\boldsymbol{\nu}}\\
&= \sum_{n\geq 0}\Big((\boldsymbol{\omega}\cdot\boldsymbol{\nu})^2-\mathcal{M}_{n-1}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)\Big)b^{[n]}_{\boldsymbol{\nu}}+\sum_{n\geq 0}\mathcal{M}_{n-1}(\boldsymbol{\omega}\cdot\boldsymbol{\nu};\varepsilon,\beta_0)b^{[n]}_{\boldsymbol{\nu}},
\end{aligned}
$$

and hence

$$
g(\boldsymbol{\omega}\cdot\boldsymbol{\nu})[\varepsilon F(\boldsymbol{\omega}t,\beta_0+b^{\mathcal{R}})]_{\boldsymbol{\nu}}=\sum_{n\geq 0}b^{[n]}_{\boldsymbol{\nu}}=b^{\mathcal{R}}_{\boldsymbol{\nu}},
$$

so that the proof is complete.

E. Proof of Lemma 4.8

Property 1-p implies $|\mathcal{G}_n(x;\varepsilon,\beta_0)|\leq c_0\alpha_{m_n}(\boldsymbol{\omega})^{-2}$ for all $0\leq n\leq p$. Then, using also Lemma 4.2 and the fact that any self-energy cluster in $\mathfrak{R}_n$ has at least two nodes for any $n\geq 0$, we obtain

$$|M_n(x;\varepsilon,\beta_0)| \le \sum_{T\in\mathfrak{R}_n} |\varepsilon|^{k(T)} |\mathscr{V}_T(x;\varepsilon,\beta_0)| \le \sum_{k\ge 2} |\varepsilon|^k C^k e^{-K_2 2^{m_n}},$$

so that (4.4a) is proved for ε small enough. Now we prove (4.4b) by induction on n. For $n=0$ the bound is obvious. Assume then (4.4b) to hold for all $n'<n$. For any $T\in\mathfrak{R}_n$ such that $\mathscr{V}_T(x;\varepsilon,\beta_0)\neq 0$ one has

$$\partial_x \mathscr{V}_T(x;\varepsilon,\beta_0) = \sum_{\ell\in\mathcal{P}_T}\left(\prod_{v\in N(T)}\mathcal{F}_v(\beta_0)\right)\left(\partial_x\mathcal{G}_{n_\ell}(x_\ell;\varepsilon,\beta_0)\prod_{\ell'\in L(T)\setminus\{\ell\}}\mathcal{G}_{n_{\ell'}}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_{\ell'};\varepsilon,\beta_0)\right),$$

where $x_\ell=\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell = x+\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell^0$ and

$$\partial_x\mathcal{G}_{n_\ell}(x_\ell;\varepsilon,\beta_0) = \frac{\partial_x\Psi_{n_\ell}(x_\ell)}{x_\ell^2-\mathcal{M}_{n_\ell-1}(x_\ell;\varepsilon,\beta_0)} - \frac{\Psi_{n_\ell}(x_\ell)\left(2x_\ell-\partial_x\mathcal{M}_{n_\ell-1}(x_\ell;\varepsilon,\beta_0)\right)}{\left(x_\ell^2-\mathcal{M}_{n_\ell-1}(x_\ell;\varepsilon,\beta_0)\right)^2}.$$

One has

$$|\partial_x\Psi_{n_\ell}(x_\ell)| \le |\partial_x\chi_{n_\ell-1}(x_\ell)| + |\partial_x\psi_{n_\ell}(x_\ell)| \le \frac{B_1}{\alpha_{m_{n_\ell}}(\boldsymbol{\omega})},$$

for some constant B_1 and, by (4.4a), the inductive hypothesis and Hypothesis 1,

$$\begin{aligned}|\partial_x\mathcal{M}_{n_\ell-1}(x_\ell;\varepsilon,\beta_0)| &\le \sum_{q=0}^{n_\ell-1}|(\partial_x\chi_q(x_\ell))M_q(x_\ell;\varepsilon,\beta_0)| + \sum_{q=0}^{n_\ell-1}|\partial_x M_q(x_\ell;\varepsilon,\beta_0)| \\ &\le \varepsilon^2 B_1K_1\sum_{q\ge0}\frac{1}{\alpha_{m_q}(\boldsymbol{\omega})}e^{-K_2 2^{m_q}} + \varepsilon^2 C_1\sum_{q\ge0}e^{-\overline{C}_1 2^{m_q}} \le \varepsilon^2 B_2,\end{aligned}$$

for some constant B_2. Hence, at the cost of replacing the bound for the propagators with $\widetilde{C}\alpha_{m_{n_\ell}}(\boldsymbol{\omega})^{-4}$ for some constant $\widetilde{C}$, one can rely upon Lemma 4.2 to obtain (4.4b) for $j=1$. For $j=2$ one can reason analogously.

F. Proof of Lemma 4.12

First of all, for any renormalised tree θ set

$$\partial_v\mathscr{V}(\theta;\varepsilon,\beta_0) := \partial_{\beta_0}\mathcal{F}_v(\beta_0)\left(\prod_{w\in N(\theta)\setminus\{v\}}\mathcal{F}_w(\beta_0)\right)\left(\prod_{\ell\in L(\theta)}\mathcal{G}_{n_\ell}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell;\varepsilon,\beta_0)\right), \tag{F.1}$$

and

$$\begin{aligned}\partial_\ell\mathscr{V}(\theta;\varepsilon,\beta_0) &:= \partial_{\beta_0}\mathcal{G}_{n_\ell}(x_\ell;\varepsilon,\beta_0)\left(\prod_{v\in N(\theta)}\mathcal{F}_v(\beta_0)\right)\left(\prod_{\lambda\in L(\theta)\setminus\{\ell\}}\mathcal{G}_{n_\lambda}(x_\lambda;\varepsilon,\beta_0)\right)\\ &= \mathcal{A}_\ell(\theta,x_\ell;\varepsilon,\beta_0)\,\partial_{\beta_0}\mathcal{G}_{n_\ell}(x_\ell;\varepsilon,\beta_0)\,\mathcal{B}_\ell(\theta;\varepsilon,\beta_0),\end{aligned} \tag{F.2}$$

where $x_\ell := \boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell$, $\partial_{\beta_0}\mathcal{G}_{n_\ell}(x_\ell;\varepsilon,\beta_0)$ is written according to Remark 3.5,

$$\mathcal{A}_\ell(\theta,x_\ell;\varepsilon,\beta_0) := \Bigg(\prod_{\substack{v\in N(\theta)\\ v\not\prec\ell}}\mathcal{F}_v(\beta_0)\Bigg)\Bigg(\prod_{\substack{\ell'\in L(\theta)\\ \ell'\not\preceq\ell}}\mathcal{G}_{n_{\ell'}}(x_{\ell'};\varepsilon,\beta_0)\Bigg), \tag{F.3a}$$

$$\mathcal{B}_\ell(\theta;\varepsilon,\beta_0) := \Bigg(\prod_{\substack{v\in N(\theta)\\ v\prec\ell}}\mathcal{F}_v(\beta_0)\Bigg)\Bigg(\prod_{\substack{\ell'\in L(\theta)\\ \ell'\prec\ell}}\mathcal{G}_{n_{\ell'}}(x_{\ell'};\varepsilon,\beta_0)\Bigg). \tag{F.3b}$$

Let us define in the analogous way $\partial_v\mathscr{V}_T(x;\varepsilon,\beta_0)$ and $\partial_\ell\mathscr{V}_T(x;\varepsilon,\beta_0)$ for any self-energy cluster T, and let us write

$$\partial_{\beta_0}\mathscr{V}(\theta;\varepsilon,\beta_0) = \partial_N\mathscr{V}(\theta;\varepsilon,\beta_0) + \partial_L\mathscr{V}(\theta;\varepsilon,\beta_0), \tag{F.4}$$

where

$$\partial_N\mathscr{V}(\theta;\varepsilon,\beta_0) := \sum_{v\in N(\theta)}\partial_v\mathscr{V}(\theta;\varepsilon,\beta_0), \qquad \partial_L\mathscr{V}(\theta;\varepsilon,\beta_0) := \sum_{\ell\in L(\theta)}\partial_\ell\mathscr{V}(\theta;\varepsilon,\beta_0). \tag{F.5}$$

Let us also write

$$\partial_{\beta_0}\mathscr{V}_T(x;\varepsilon,\beta_0) = \partial_N\mathscr{V}_T(x;\varepsilon,\beta_0) + \partial_L\mathscr{V}_T(x;\varepsilon,\beta_0), \tag{F.6}$$

for any $T\in\mathfrak{R}_n$, $n\geq 0$, where the derivatives ∂_N and ∂_L are defined as in the previous cases (F.5), with $N(T)$ and $L(T)$ replacing $N(\theta)$ and $L(\theta)$, respectively, so that we can split

$$\begin{aligned}\partial_{\beta_0}M_n(x;\varepsilon,\beta_0) &= \partial_N M_n(x;\varepsilon,\beta_0) + \partial_L M_n(x;\varepsilon,\beta_0),\\ \partial_{\beta_0}\mathcal{M}_n(x;\varepsilon,\beta_0) &= \partial_N\mathcal{M}_n(x;\varepsilon,\beta_0) + \partial_L\mathcal{M}_n(x;\varepsilon,\beta_0),\end{aligned} \tag{F.7}$$

again with obvious meaning of the symbols.

Remark F.1. We can interpret the derivative ∂_v as all the possible ways to attach an extra line (carrying a momentum $\mathbf{0}$) to the node v, so that $\sum_{k\geq 0}\varepsilon^{k+1}\sum_{\theta\in\Theta^{\mathcal{R}}_{k+1,\mathbf{0}}}\partial_N\mathscr{V}(\theta;\varepsilon,\beta_0)$ produces contributions to $\mathcal{M}_\infty(0;\varepsilon,\beta_0)$.

Given any $\theta\in\Theta^{\mathcal{R}}_{k,\mathbf{0}}$ we have to study the derivative (F.4). The terms (F.1) produce immediately contributions to $\mathcal{M}_\infty(0;\varepsilon,\beta_0)$ by Remark F.1. Thus, we have to study the derivatives $\partial_\ell\mathscr{V}(\theta;\varepsilon,\beta_0)$ appearing in the second sum in (F.5). Here and henceforth, we shall not write any longer explicitly the dependence on ε and β_0 of both propagators and self-energies, in order not to overwhelm the notation.

For any $\theta\in\Theta^{\mathcal{R}}_{k,\mathbf{0}}$ such that $\mathscr{V}(\theta;\varepsilon,\beta_0)\neq 0$ and for any line $\ell\in L(\theta)$, either there is only one scale n such that $\Psi_n(x_\ell)\neq 0$ (and in that case $\Psi_n(x_\ell)=1$ and $\Psi_{n'}(x_\ell)=0$ for all $n'\neq n$) or there exists only one $n\geq 0$ such that $\Psi_n(x_\ell)\Psi_{n+1}(x_\ell)\neq 0$.

1. If $\Psi_n(x_\ell) = 1$ one has

$$\begin{aligned}\partial_\ell \mathscr{V}(\theta;\varepsilon,\beta_0) &= \mathcal{A}_\ell(\theta,x_\ell)\,\frac{\Psi_n(x_\ell)}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\frac{1}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}\,\mathcal{B}_\ell(\theta)\\ &= \mathcal{A}_\ell(\theta,x_\ell)\,\frac{\Psi_n(x_\ell)}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\frac{\Psi_n(x_\ell)}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}\,\mathcal{B}_\ell(\theta)\\ &= \mathcal{A}_\ell(\theta,x_\ell)\,\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_n(x_\ell)\,\mathcal{B}_\ell(\theta),\end{aligned}\tag{F.8}$$

where (here and henceforth) we shorten $\mathcal{A}_\ell(\theta,x_\ell) = \mathcal{A}_\ell(\theta,x_\ell;\varepsilon,\beta_0)$ and $\mathcal{B}_\ell(\theta) = \mathcal{B}_\ell(\theta;\varepsilon,\beta_0)$.

Remark F.2. Note that if we split $\partial_{\beta_0} = \partial_N + \partial_L$ in (F.8), the term with $\partial_N\mathcal{M}_{n-1}(x_\ell)$ is a contribution to $\mathcal{M}_\infty(0)$.

If there is only one $n \ge 0$ such that $\Psi_n(x_\ell)\Psi_{n+1}(x_\ell) \ne 0$, then $\Psi_n(x_\ell)+\Psi_{n+1}(x_\ell) = 1$ and $\chi_q(x_\ell) = 1$ for all $q = -1,\ldots,n-1$, so that $\psi_{n+1}(x_\ell) = 1$, and hence $\Psi_{n+1}(x_\ell) = \chi_n(x_\ell)$. Moreover it can happen only (see Remark 3.7) $n_\ell = n$ or $n_\ell = n+1$.

2. Consider first the case $n_\ell = n+1$. One has

$$\begin{aligned}\partial_\ell \mathscr{V}(\theta;\varepsilon,\beta_0) &= \mathcal{A}_\ell(\theta,x_\ell)\,\mathcal{G}_{n+1}(x_\ell)\partial_{\beta_0}\mathcal{M}_n(x_\ell)\frac{1}{x_\ell^2-\mathcal{M}_n(x_\ell)}\,\mathcal{B}_\ell(\theta)\\ &= \mathcal{A}_\ell(\theta,x_\ell)\,\mathcal{G}_{n+1}(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\frac{\Psi_n(x_\ell)+\Psi_{n+1}(x_\ell)}{x_\ell^2-\mathcal{M}_n(x_\ell)}\,\mathcal{B}_\ell(\theta)\\ &\quad+\mathcal{A}_\ell(\theta,x_\ell)\,\mathcal{G}_{n+1}(x_\ell)\partial_{\beta_0}M_n(x_\ell)\frac{\chi_n(x_\ell)}{x_\ell^2-\mathcal{M}_n(x_\ell)}\,\mathcal{B}_\ell(\theta)\\ &= \mathcal{A}_\ell(\theta,x_\ell)\,\mathcal{G}_{n+1}(x_\ell)\left(\sum_{q=-1}^{n}\partial_{\beta_0}M_q(x_\ell)\right)\mathcal{G}_{n+1}(x_\ell)\,\mathcal{B}_\ell(\theta)\\ &\quad+\mathcal{A}_\ell(\theta,x_\ell)\,\mathcal{G}_{n+1}(x_\ell)\left(\sum_{q=-1}^{n-1}\partial_{\beta_0}M_q(x_\ell)\right)\mathcal{G}_n(x_\ell)\,\mathcal{B}_\ell(\theta)\\ &\quad+\mathcal{A}_\ell(\theta,x_\ell)\,\mathcal{G}_{n+1}(x_\ell)\left(\sum_{q=-1}^{n-1}\partial_{\beta_0}M_q(x_\ell)\right)\\ &\quad\times\mathcal{G}_n(x_\ell)M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell)\,\mathcal{B}_\ell(\theta).\end{aligned}\tag{F.9}$$

We can represent graphically the three contributions in (F.9) as in Fig. 5: we represent the derivative ∂_{β_0} as an arrow pointing toward the graphical representation of the differentiated quantity; see also Figs. 7, 10 and 12.

Remark F.3. Note that the $M_n(x_\ell)$ appearing in the latter line of (F.9) has to be interpreted (see Remark 3.6) as

$$\sum_{T\in\mathfrak{L}\mathfrak{F}_n}\varepsilon^{k(T)}\,\mathscr{V}_T(x_\ell;\varepsilon,\beta_0).$$

Note also that, again, if we split $\partial_{\beta_0} = \partial_N + \partial_L$ in (F.9), all the terms with $\partial_N M_q(x_\ell)$ are contributions to $\mathcal{M}_\infty(0)$.

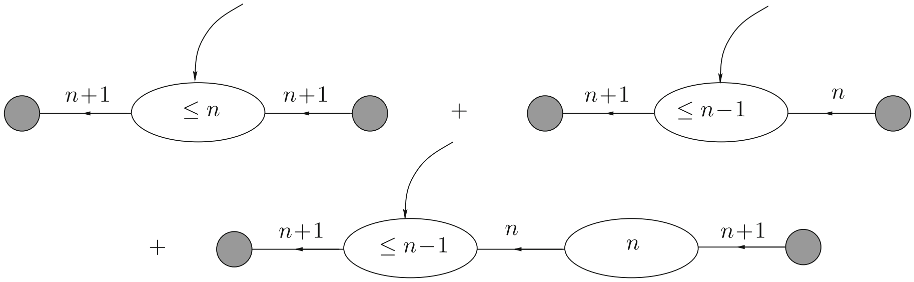

Fig. 5. Graphical representation of the derivative $\partial_\ell \mathscr{V}(\theta;\varepsilon,\beta_0)$ according to (F.9)

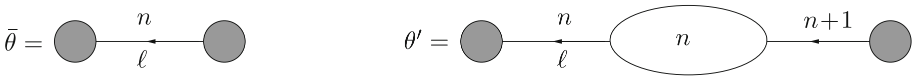

Fig. 6. The renormalised tree $\bar\theta$ and the renormalised trees θ' of the set $\tau_1(\bar\theta,\ell)$ associated with $\bar\theta$

Now consider the case $n_\ell = n$.

3. If ℓ is not the exiting line of a left-fake cluster, set $\bar\theta = \theta$; otherwise, if ℓ is the exiting line of a left-fake cluster T, define – if possible – $\bar\theta$ as the renormalised tree obtained from θ by removing T and ℓ'_T. In both cases, define – if possible – $\tau_1(\bar\theta,\ell)$ as the set constituted by all the renormalised trees θ' obtained from $\bar\theta$ by inserting a left-fake cluster, together with its entering line, between ℓ and the node v which ℓ exits; see Fig. 6. Here and henceforth, if S is a subgraph with only one entering line $\ell'_S = \ell_v$ and one exiting line ℓ_S and we "remove" S together with ℓ'_S, we mean that we also reattach the line ℓ_S to the node v.

Remark F.4. The construction of the set $\tau_1(\bar\theta,\ell)$ could be impossible if the removal or the insertion of a left-fake cluster T, together with its entering line ℓ'_T, produce a self-energy cluster. We shall see later how to deal with these cases.

Then one has

$$\begin{aligned}\partial_\ell \mathscr{V}(\bar\theta;\varepsilon,\beta_0) + \partial_\ell \sum_{\theta'\in\tau_1(\bar\theta,\ell)} \mathscr{V}(\theta';\varepsilon,\beta_0)\\ = \mathcal{A}_\ell(\bar\theta,x_\ell)\,\partial_{\beta_0}\mathcal{G}_n(x_\ell)\,(1+M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell))\;\mathcal{B}_\ell(\bar\theta),\end{aligned} \tag{F.10}$$

where

$$\begin{aligned}&\partial_{\beta_0}\mathcal{G}_n(x_\ell)\,(1+M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell))\\ &\quad= \mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_n(x_\ell) + \mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\frac{\Psi_{n+1}(x_\ell)}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}\\ &\qquad+\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_n(x_\ell)M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell)\\ &\qquad+\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\frac{\Psi_{n+1}(x_\ell)}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell)\\ &\quad= \mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_n(x_\ell) + \mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_{n+1}(x_\ell)\\ &\qquad-\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\frac{\chi_n(x_\ell)}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell)\end{aligned}$$

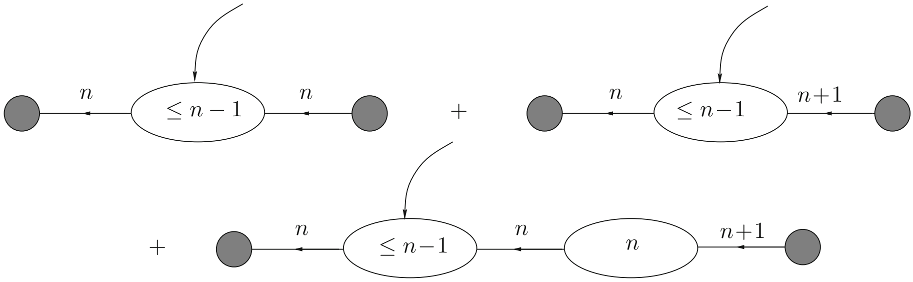

Fig. 7. Graphical representation of the three contributions in the last two lines of (F.11)

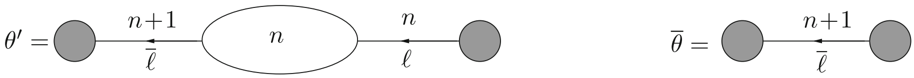

Fig. 8. The trees θ' of the set $\tau_2(\overline{\theta},\overline{\ell})$ obtained from $\overline{\theta}$ when $\ell \in L(\theta)$ enters a right-fake cluster

$$
\begin{aligned}
&+\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_n(x_\ell)M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell)\\
&+\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\frac{\Psi_{n+1}(x_\ell)}{x_\ell^2-\mathcal{M}_{n-1}(x_\ell)}M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell)\\
=\;&\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_n(x_\ell)+\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_{n+1}(x_\ell)\\
&+\mathcal{G}_n(x_\ell)\partial_{\beta_0}\mathcal{M}_{n-1}(x_\ell)\mathcal{G}_n(x_\ell)M_n(x_\ell)\mathcal{G}_{n+1}(x_\ell),
\end{aligned}
\tag{F.11}
$$

so that also in this case, if we split $\partial_{\beta_0}=\partial_N+\partial_L$, all the terms with $\partial_N\mathcal{M}_{n-1}$ are contributions to $\mathcal{M}_\infty(0)$ – see Remark F.2. Again, we can represent graphically the three contributions obtained inserting (F.11) in (F.10): see Fig. 7.

4. Assume now that ℓ is not the exiting line of a left-fake cluster, and the insertion of a left-fake cluster, together with its entering line, produces a self-energy cluster. Note that this can happen only if ℓ is the entering line of a renormalised right-fake cluster T. Let $\overline{\ell}$ be the exiting line (on scale $n+1$) of the renormalised right-fake cluster T, call $\overline{\theta}$ the renormalised tree obtained from θ by removing T and ℓ and call $\tau_2(\overline{\theta},\overline{\ell})$ the set of renormalised trees θ' obtained from $\overline{\theta}$ by inserting a right-fake cluster, together with its entering line, before $\overline{\ell}$; see Fig. 8.

By construction one has

$$
\begin{aligned}
\mathscr{V}(\overline{\theta};\varepsilon,\beta_0)&=\mathcal{A}_{\overline{\ell}}(\overline{\theta},x_\ell)\,\mathcal{G}_{n+1}(x_{\overline{\ell}})\,\mathcal{B}_{\overline{\ell}}(\overline{\theta}),\\
\sum_{\theta'\in\tau_2(\overline{\theta},\overline{\ell})}\mathscr{V}(\theta';\varepsilon,\beta_0)&=\mathcal{A}_{\overline{\ell}}(\overline{\theta},x_\ell)\,\mathcal{G}_{n+1}(x_{\overline{\ell}})\,M_n(x_{\overline{\ell}})\,\mathcal{G}_n(x_{\overline{\ell}})\,\mathcal{B}_{\overline{\ell}}(\overline{\theta}),
\end{aligned}
$$

where we have used that $x_\ell = x_{\overline{\ell}}$.

Consider the contribution to $\partial_{\overline{\ell}}\mathscr{V}(\overline{\theta};\varepsilon,\beta_0)$ – see (F.9) – given by

$$
\mathcal{A}_{\overline{\ell}}(\overline{\theta},x_{\overline{\ell}})\mathcal{G}_{n+1}(x_{\overline{\ell}})\partial_L M_n(x_{\overline{\ell}})\mathcal{G}_{n+1}(x_{\overline{\ell}})\mathcal{B}_{\overline{\ell}}(\overline{\theta}).
\tag{F.12}
$$

Call $\mathfrak{R}_n(T)$ the subset of $\mathfrak{R}_n$ such that if $T'\in\mathfrak{R}_n(T)$ the exiting line $\ell_{T'}$ exits also the renormalised right-fake cluster T; note that the entering line ℓ of T must be also the exiting line of some renormalised left-fake cluster T'' contained in T'; see Fig. 9.

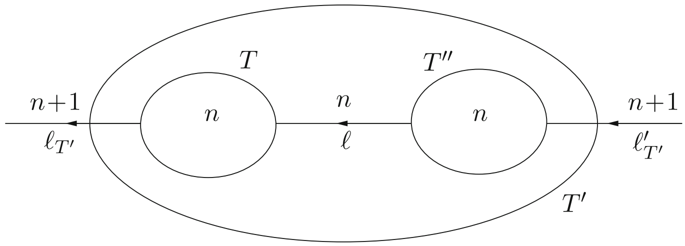

Fig. 9. A self-energy cluster $T' \in \mathfrak{R}_n(T)$

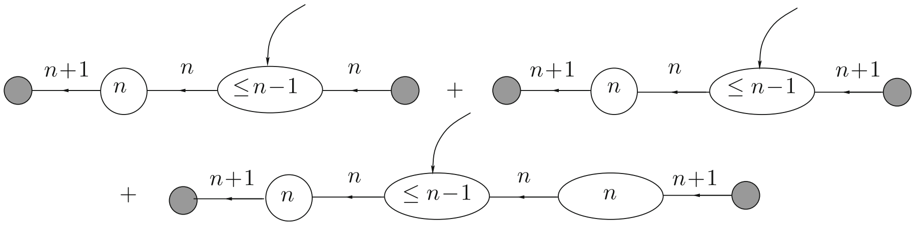

Fig. 10. Graphical representation of the three contributions arising from (F.14)

Define

$$M_n(T, x_{\overline{\ell}}; \varepsilon, \beta_o) = \sum_{T' \in \mathfrak{R}_n(T)} \varepsilon^{k(T')} \mathscr{V}_{T'}(x_{\overline{\ell}}; \varepsilon, \beta_0). \tag{F.13}$$

Hence one has

$$\begin{aligned} &\partial_\ell \sum_{\theta' \in \tau_2(\overline{\theta}, \overline{\ell})} \mathscr{V}(\theta'; \varepsilon, \beta_0) + \mathcal{A}_{\overline{\ell}}(\overline{\theta}, x_\ell)\, \mathcal{G}_{n+1}(x_\ell)\, \partial_\ell \sum_{T \in \mathfrak{R}\mathfrak{F}_n} M_n(T, x_{\overline{\ell}}) \mathcal{G}_{n+1}(x_\ell)\, \mathcal{B}_{\overline{\ell}}(\overline{\theta}) \\ &\quad = \mathcal{A}_{\overline{\ell}}(\overline{\theta}, x_\ell)\, \mathcal{G}_{n+1}(x_\ell) M_n(x_\ell) \partial_{\beta_0} \mathcal{G}_n(x_\ell)\, (1 + M_n(x_\ell) \mathcal{G}_{n+1}(x_\ell))\, \mathcal{B}_{\overline{\ell}}(\overline{\theta}), \end{aligned} \tag{F.14}$$

where we have used again that $x_\ell = x_{\overline{\ell}}$. Thus, one can reason as in (F.11), so as to obtain the sum of three contributions, as represented in Fig. 10.

5. Finally, consider the case in which ℓ is the exiting line of a renormalised left-fake cluster, T_0 and the removal of T_0 and ℓ'_{T_0} creates a self-energy cluster.

Set (for a reason that will become clear later) $\theta_0 = \theta$ and $\ell_0 = \ell$. Then there is a maximal $m \geq 1$ such that there are $2m$ lines $\ell_1, \ldots, \ell_m$ and $\ell'_1, \ldots \ell'_m$, with the following properties:

(i) $\ell_i \in \mathcal{P}(\ell_{\theta_0}, \ell_{i-1})$, for $i = 1, \ldots, m$,
(ii) $n_{\ell_i} = n + i < \max\{p : \Psi_p(x_{\ell_i}) \neq 0\} = n + i + 1$, for $i = 0, \ldots, m-1$, while $n_m := n_{\ell_m} = n + m + \sigma$, with $\sigma \in \{0, 1\}$,
(iii) $\boldsymbol{\nu}_{\ell_i} \neq \boldsymbol{\nu}_{\ell_{i-1}}$ and the lines preceding ℓ_i but not ℓ_{i-1} are on scale $\leq n + i - 1$, for $i = 1, \ldots, m$,
(iv) $\boldsymbol{\nu}_{\ell'_i} = \boldsymbol{\nu}_{\ell_i}$, for $i = 1, \ldots, m$,
(v) if $m \geq 2$, ℓ'_i is the exiting line of a left-fake cluster T_i, for $i = 1, \ldots, m-1$,
(vi) $\ell'_i \prec \ell'_{T_{i-1}}$ and all the lines preceding $\ell'_{T_{i-1}}$ but not ℓ'_i are on scale $\leq n + i - 1$, for $i = 1, \ldots, m$,
(vii) $n'_m := n_{\ell'_m} = n + m + \sigma'$ with $\sigma' \in \{0, 1\}$.

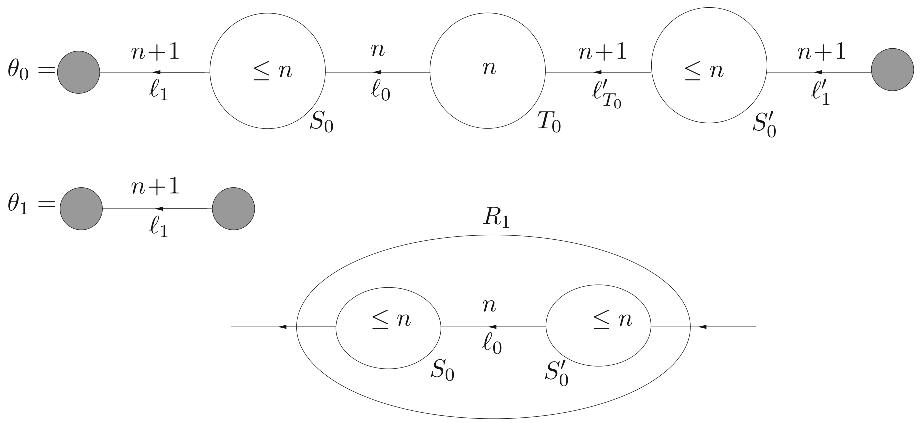

Fig. 11. The renormalised trees θ_0 and θ_1 and the self-energy cluster R_1 in Case 5 with $m = 1$ and $\sigma = \sigma' = 0$. Note that the set S'_0 is a cluster, but not a self-energy cluster

Note that one cannot have $\sigma = \sigma' = 1$, otherwise the subgraph between ℓ_m and ℓ'_m would be a self-energy cluster. Note also that (ii), (iv) and (v) imply $n_{\ell'_i} = n + i$ for $i = 1, \dots, m - 1$ if $m \geq 2$. Call S_i the subgraph between ℓ_{i+1} and ℓ_i, and S'_i the cluster between ℓ'_{T_i} and ℓ'_{i+1} for all $i = 0, \dots, m-1$. For $i = 1, \dots, m$, call θ_i the renormalised tree obtained from θ_0 by removing everything between ℓ_i and the part of θ_0 preceding ℓ'_i, and note that if $m \geq 2$, Properties (i)–(vii) hold for θ_i but with $m - i$ instead of m, for all $i = 1, \dots, m - 1$.

For $i = 1, \dots, m$, call R_i the self-energy cluster obtained from the subgraph of θ_{i-1} between ℓ_i and ℓ'_i, by removing the left-fake cluster T_{i-1} together with ℓ'_{T_i}. Note that $L(R_i) = L(S_{i-1}) \cup \{\ell_{i-1}\} \cup L(S'_{i-1})$ and $N(R_i) = N(S_{i-1}) \cup N(S'_{i-1})$; see Fig. 11.

For $i = 0, \dots, m-1$, given $\ell', \ell \in L(\theta_i)$, with $\ell' \prec \ell$, call $\mathcal{P}^{(i)}(\ell, \ell')$ the path of lines in θ_i connecting ℓ' to ℓ (hence $\mathcal{P}^{(i)}(\ell, \ell') = \mathcal{P}(\ell, \ell') \cap L(\theta_i)$). For any $i = 0, \dots, m-1$ and any $\ell \in \mathcal{P}^{(i)}(\ell_i, \ell'_m)$, let $\tau_3(\theta_i, \ell)$ be the set of all renormalised trees which can be obtained from θ_i by replacing each left-fake cluster preceding ℓ but not ℓ'_m with all possible left-fake clusters. Set also $\tau_3(\theta_{m-1}, \ell'_m) = \theta_{m-1}$.

Note that

$$\begin{aligned} \mathcal{A}_{\ell_m}(\theta_m, x_{\ell_m}) \mathcal{G}_{n_m}(x_{\ell_m}) \mathscr{V}(S_{m-1}) &= \mathcal{A}_{\ell_{m-1}}(\theta_{m-1}, x_{\ell_{m-1}}), \\ \mathscr{V}(S'_{m-1}) \mathcal{G}_{n'_m}(x_{\ell_m}) \mathcal{B}_{\ell_m}(\theta_m) &= \mathcal{B}_{\ell'_{T_{m-1}}}(\theta_{m-1}), \end{aligned} \tag{F.15}$$

and one among Cases **1**–**4** holds for $\ell_m \in L(\theta_m)$ so that we can consider the contribution to $\partial_{\ell_m} \mathscr{V}(\theta_m; \varepsilon, \beta_0)$ (together with other contributions as in **3** and **4** if necessary) given by – see (F.8), (F.9) and (F.11) –

$$\mathcal{A}_{\ell_m}(\theta_m, x_{\ell_m}) \mathcal{G}_{n_m}(x_{\ell_m}) \partial_{\ell_{m-1}} \mathscr{V}_{R_m}(x_{\ell_m}) \mathcal{G}_{n'_m}(x_{\ell_m}) \mathcal{B}_{\ell_m}(\theta_m).$$

Then one has

$$\begin{aligned} &\mathcal{A}_{\ell_m}(\theta_m, x_{\ell_m}) \mathcal{G}_{n_m}(x_{\ell_m}) \partial_{\ell_{m-1}} \mathscr{V}_{R_m}(x_{\ell_m}) \mathcal{G}_{n'_m}(x_{\ell_m}) \mathcal{B}_{\ell_m}(\theta_m) \\ &\quad + \partial_{\ell_{m-1}} \sum_{\theta' \in \tau_3(\theta_{m-1}, \ell_{m-1})} \mathscr{V}(\theta'; \varepsilon, \beta_0) \end{aligned}$$

$$= \mathcal{A}_{\ell_{m-1}}(\theta_{m-1}, x_{\ell_{m-1}}) \partial_{\beta_0} \mathcal{G}_{n+m-1}(x_{\ell_{m-1}}) \\ \times \left(1 + M_{n+m-1}(x_{\ell_{m-1}}) \mathcal{G}_{n+m}(x_{\ell_{m-1}})\right) \mathcal{B}_{\ell'_{T_{m-1}}}(\theta_{m-1}), \tag{F.16}$$

and hence we obtain, reasoning as in (F.11),

$$\mathcal{A}_{\ell_{m-1}}(\theta_{m-1}, x_{\ell_{m-1}}) \mathcal{G}_{n+m-1}(x_{\ell_{m-1}}) \partial_{\beta_0} \mathcal{M}_{n+m-2}(x_{\ell_{m-1}}) \mathcal{G}_{n+m-1}(x_{\ell_{m-1}}) \mathcal{B}_{\ell'_{T_{m-1}}}(\theta_{m-1}) \\ + \mathcal{A}_{\ell_{m-1}}(\theta_{m-1}, x_{\ell_{m-1}}) \mathcal{G}_{n+m-1}(x_{\ell_{m-1}}) \partial_{\beta_0} \mathcal{M}_{n+m-2}(x_{\ell_{m-1}}) \mathcal{G}_{n+m}(x_{\ell_{m-1}}) \mathcal{B}_{\ell'_{T_{m-1}}}(\theta_{m-1}) \\ + \mathcal{A}_{\ell_{m-1}}(\theta_{m-1}, x_{\ell_{m-1}}) \mathcal{G}_{n+m-1}(x_{\ell_{m-1}}) \partial_{\beta_0} \mathcal{M}_{n+m-2}(x_{\ell_{m-1}}) \mathcal{G}_{n+m-1}(x_{\ell_{m-1}}) \\ \times M_{n+m-1}(x_{\ell_{m-1}}) \mathcal{G}_{n+m}(x_{\ell_{m-1}}) \mathcal{B}_{\ell'_{T_{m-1}}}(\theta_{m-1}). \tag{F.17}$$

Then, for $i = m-1, \dots, 1$ we recursively reason as follows. Set

$$\mathcal{B}_{\ell'_{T_i}}(\tau_3(\theta_i, \ell'_{i+1})) := \sum_{\theta' \in \tau_3(\theta_i, \ell'_{i+1})} \mathcal{B}_{\ell'_{T_i}}(\theta'),$$

and note that

$$\mathcal{A}_{\ell_i}(\theta_i, x_{\ell_i}) \mathcal{G}_{n+i}(x_{\ell_i}) \mathscr{V}(S_{i-1}) = \mathcal{A}_{\ell_{i-1}}(\theta_{i-1}, x_{\ell_{i-1}}), \\ \mathscr{V}(S'_{i-1}) \mathcal{G}_{n+i}(x_{\ell_i}) M_{n+i}(x_{\ell_i}) \mathcal{G}_{n+i+1}(x_{\ell_i}) \mathcal{B}_{\ell'_{T_i}}(\tau_3(\theta_i, \ell'_{i+1})) = \mathcal{B}_{\ell'_{T_{i-1}}}(\tau_3(\theta_{i-1}, \ell'_i)). \tag{F.18}$$

Consider the contribution

$$\mathcal{A}_{\ell_i}(\theta_i, x_{\ell_i}) \mathcal{G}_{n+i}(x_{\ell_i}) \partial_{\ell_{i-1}} \mathscr{V}_{R_i}(x_{\ell_i}) \mathcal{G}_{n+i}(x_{\ell_i}) M_{n+i}(x_{\ell_i}) \mathcal{G}_{n+i+1}(x_{\ell_i}) \mathcal{B}_{\ell'_{T_i}}(\tau_3(\theta_i, \ell'_{i+1})), \tag{F.19}$$

obtained at the $(i+1)^{\text{th}}$ step of the recursion. By (F.18) one has (see Fig. 12)

$$\mathcal{A}_{\ell_i}(\theta_i, x_{\ell_i}) \mathcal{G}_{n+i}(x_{\ell_i}) \partial_{\ell_{i-1}} \mathscr{V}_{R_i}(x_{\ell_i}) \mathcal{G}_{n+i}(x_{\ell_i}) M_{n+i}(x_{\ell_i}) \mathcal{G}_{n+i+1}(x_{\ell_i}) \mathcal{B}_{\ell'_{T_i}}(\tau_3(\theta_i, \ell'_{i+1})) \\ + \partial_{\ell_{i-1}} \sum_{\theta' \in \tau_3(\theta_{i-1}, \ell_{i-1})} \mathscr{V}(\theta'; \varepsilon, \beta_0) = \mathcal{A}_{\ell_{i-1}}(\theta_{i-1}, x_{\ell_{i-1}}) \partial_{\beta_0} \mathcal{G}_{n+i-1}(x_{\ell_{i-1}}) \\ \times \left(1 + M_{n+i-1}(x_{\ell_{i-1}}) \mathcal{G}_{n+i}(x_{\ell_{i-1}})\right) \mathcal{B}_{\ell'_{T_{i-1}}}(\tau_3(\theta_{i-1}, \ell'_i)), \tag{F.20}$$

which produces, as in (F.17), the contribution

$$\mathcal{A}_{\ell_{i-1}}(\theta_{i-1}, x_{\ell_{i-1}}) \mathcal{G}_{n+i-1}(x_{\ell_{i-1}}) \partial_{\ell_{i-2}} \mathscr{V}_{R_{i-1}}(x_{\ell_{i-1}}) \mathcal{G}_{n+i-1}(x_{\ell_{i-1}}) \\ \times M_{n+i-1}(x_{\ell_{i-1}}) \mathcal{G}_{n+i}(x_{\ell_{i-1}}) \mathcal{B}_{\ell'_{T_{i-1}}}(\tau_3(\theta_{i-1}, \ell'_i)). \tag{F.21}$$

Hence we can proceed recursively from θ_m up to θ_0, until we obtain

$$\mathcal{A}_{\ell_0}(\theta_0, x_{\ell_0}) \mathcal{G}_n(x_{\ell_0}) \partial_{\beta_0} \mathcal{M}_{n-1}(x_{\ell_0}) \mathcal{G}_n(x_{\ell_0}) \mathcal{B}_{\ell'_{T_0}}(\tau_3(\theta_0, \ell'_1)) \\ + \mathcal{A}_{\ell_0}(\theta_0, x_{\ell_0}) \mathcal{G}_n(x_{\ell_0}) \partial_{\beta_0} \mathcal{M}_{n-1}(x_{\ell-0}) \mathcal{G}_{n+1}(x_{\ell_0}) \mathcal{B}_{\ell'_{T_0}}(\tau_3(\theta_0, \ell'_1)) \\ + \mathcal{A}_{\ell_0}(\theta_0, x_{\ell_0}) \mathcal{G}_n(x_{\ell_0}) \partial_{\beta_0} \mathcal{M}_{n-1}(x_{\ell_0}) \mathcal{G}_n(x_{\ell_0}) M_n(x_{\ell_0}) \mathcal{G}_{n+1}(x_{\ell_0}) \mathcal{B}_{\ell'_{T_0}}(\tau_3(\theta_0, \ell'_1)). \tag{F.22}$$

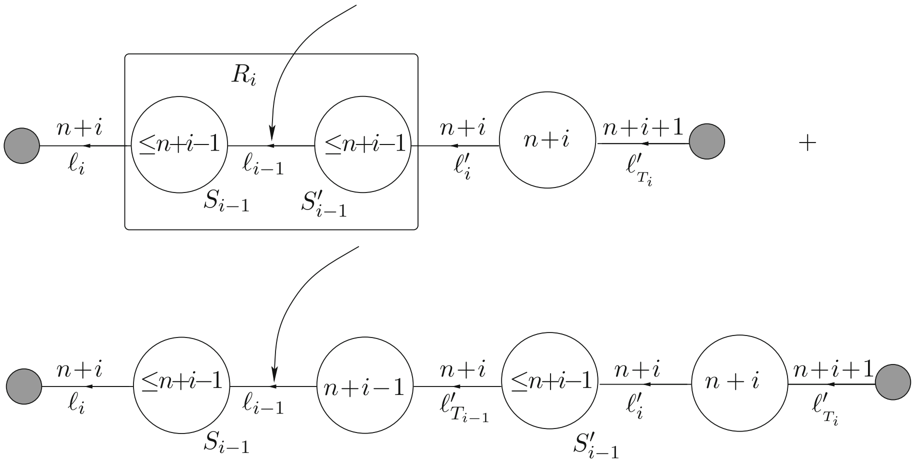

Fig. 12. Graphical representation of the left-hand side of (F.20)

Once again, if we split $\partial_{\beta_0} = \partial_N + \partial_L$, all the terms with $\partial_N \mathcal{M}_{n-1}$ are contributions to $\mathcal{M}_\infty(0)$.

6. We are left with the derivatives $\partial_L M_q(x; \varepsilon, \beta_0)$, $q \leq n$, when the differentiated propagator is not one of those used along the Cases **4** or **5**; see for instance (F.14), (F.16) and (F.20). One can reason as in the case $\partial_L \mathscr{V}(\theta; \varepsilon, \beta_0)$, by studying the derivatives $\partial_\ell \mathscr{V}_T(x_\ell; \varepsilon, \beta_0)$ and proceed iteratively along the lines of Cases **1** to **5** above, until only lines on scales 0 are left. In that case the derivatives $\partial_{\beta_0} \mathcal{G}_0(x_\ell; \varepsilon, \beta_0)$ produce derivatives $\partial_{\beta_0} M_{-1}(x; \varepsilon, \beta_0) = \varepsilon \partial^2_{\beta_0} F_{\mathbf{0}}(\beta_0)$ (see Remarks 3.4 and 3.5). Therefore, for $n = -1$, in the splitting (F.7), there are no terms with the derivatives ∂_ℓ, and the derivatives ∂_v can be interpreted as said in Remark F.1. It is also easy to realise that, by construction, each contribution to $\mathcal{M}_\infty(0; \varepsilon, \beta_0)$ appears as one term among those considered in the discussion above. Hence the assertion follows.

Remark F.5. If we used a sharp scale decomposition instead of the C^∞ one, the proof above would be much easier. More precisely, if we defined the (discontinuous) function

$$\chi(x) := \begin{cases} 1, & |x| \leq 1, \\ 0, & |x| > 1, \end{cases}$$

and consequently changed the definitions of ψ, and χ_n, ψ_n and Ψ_n for $n \geq 0$, we could reduce the proof of Lemma 4.12 to (iterations of) Case **1**. Moreover in such a case, defining $G_n^{\mathcal{R}}(\varepsilon, \beta_0) = \sum_{k \geq 0} \varepsilon^k G_n^{[k]}(\varepsilon, \beta_0)$ and $G_n^{[k]}(\varepsilon, \beta_0) = \sum_{\theta \in \Theta^{\mathcal{R}}_{k+1,\mathbf{0},n}} \mathscr{V}(\theta, \varepsilon, \beta_0)$, with $\Theta^{\mathcal{R}}_{k,\boldsymbol{\nu},n} = \{\theta \in \Theta^{\mathcal{R}}_{k,\boldsymbol{\nu}} \, : \, n_\ell \leq n \text{ for all } \ell \in L(\theta)\}$, we would obtain the stronger identity $\mathcal{M}_n(0; \varepsilon, \beta_0) = \varepsilon \partial_{\beta_0} G_n^{\mathcal{R}}(\varepsilon, \beta_0)$ for all $n \geq -1$. On the other hand, the bound (4.4b) in Lemma 4.8 would be no longer true because of the derivative $\partial_x \Psi_n$, so that further work would be however needed; see for instance [16] where a sharp scale decomposition is used for the standard KAM theorem and $\boldsymbol{\omega}$ satisfying the standard Diophantine condition. Analogously, using a scale decomposition depending on the whole sequence $\{\alpha_n(\boldsymbol{\omega})\}_{n \geq 0}$, i.e. defining the function χ as in (3.8) but setting $\chi_n(x) = \chi(8x/\alpha_n(\boldsymbol{\omega}))$ and consequently changing the definitions of ψ_n and Ψ_n for $n \geq 0$, Remark 3.7 does not hold anymore so that the analysis of Cases **2–5** becomes more complicated.

G. Proof of Lemma 6.6

For all $n \geq 0$ one can write

$$\mathcal{M}_n^{(k)}(x) = -\mathcal{M}_{>n}^{(k)}(0) + x^2 \int_0^1 \mathrm{d}\tau_0 \ (1-\tau_0)\, \partial_x^2 \mathcal{M}_n^{(k)}(\tau_0 x), \tag{G.1}$$

where we have used Lemmas 6.4 and 6.5 and we have defined

$$\mathcal{M}_{>n}^{(k)}(0) = \sum_{p=n+1}^{\infty} M_p^{(k)}(0). \tag{G.2}$$

Call $\mathfrak{S}_{>n}^k$ and $\mathfrak{S}_{\leq n}^k$ the sets of self-energy clusters with order k and scale $p > n$ and $p \leq n$, respectively, and denote by $\mathfrak{S}^k$ the set of all self-energy clusters of order k. Then one can write

$$\mathcal{M}_n^{(k)}(x) = -\sum_{T \in \mathfrak{S}_{>n}^k} \mathscr{V}_T(0) + x^2 \sum_{T \in \mathfrak{S}_{\leq n}^k} \int_0^1 \mathrm{d}\tau_0 \ (1-\tau_0)\, \partial_x^2\, \mathscr{V}_T(\tau_0 x). \tag{G.3}$$

Given a self-energy cluster T, we define $\mathcal{T}(T)$ the set of self-energy clusters in T and by $\mathcal{T}_1(T)$ the set of maximal self-energy clusters strictly contained in T. Given a line $\ell \in L(T)$ there exist $p = p(\ell) \geq 1$ self-energy clusters $T_0, \ldots, T_{p-1}$ such that $T = T_0 \supset T_1 \supset \ldots \supset T_{p-1}$, $T_j \in \mathcal{T}_1(T_{j-1})$ for $j = 1, \ldots, p-1$ and T_{p-1} is the minimal self-energy cluster containing ℓ: we call $\mathcal{C}_\ell(T) := \{T_j\}_{j=0}^{p-1}$ the *cloud* of ℓ and $\{T_j\}_{j=1}^{p-1}$ the *internal cloud* of ℓ. If $p = 1$ then the internal cloud of ℓ is the empty set.

With each T contributing to $\mathcal{M}_n^{(k)}(x)$ through (G.3) we associate a label $\delta_T \in \{\mathcal{L}, \mathcal{R}\}$, by setting $\delta_T = \mathcal{L}$ if the scale of T is $> n$ and $\delta_T = \mathcal{R}$ if the scale of T is $\leq n$: If we define

$$\mathcal{L}_T = -\mathscr{V}_T(0), \qquad \mathcal{R}_T(x) = x^2 \int_0^1 \mathrm{d}\tau_0 \ (1-\tau_0)\, \partial_x^2\, \mathscr{V}_T(\tau_0 x), \tag{G.4}$$

then we associate with T the value $\mathcal{L}_T$ if $\delta_T = \mathcal{L}$ and the value $\mathcal{R}_T(x)$ if $\delta_T = \mathcal{R}$.

Consider first a contribution $\mathcal{L}_T$ in (G.3). Call $\mathfrak{F}(T)$ the set of all self-energy clusters $T' \in \mathfrak{S}_{>n}^k$, where each $T_i \in \mathcal{T}_1(T)$ is replaced by any self-energy cluster with the same order k_i as T_i and scale $\leq n_i$, if $n_i + 1 = n_{T_i}$; here and henceforth, given a self-energy cluster T, we define $n_T = \min\{n_{\ell_T}, n_{\ell'_T}\}$. Call T^* the set of nodes and lines obtained from T by removing all nodes and lines belonging to the self-energy clusters $T' \in \mathcal{T}_1(T)$; one has $(T')^* = T^*$ for all $T' \in \mathfrak{F}(T)$. For all $i = 1, \ldots, |\mathcal{T}_1(T)|$ we sum together

$$\sum_{T_i \in \mathfrak{S}_{\leq n_i}^{k_i}} \mathscr{V}_{T_i}(x_{\ell'_{T_i}}(\tau_0)) = \mathcal{M}_{n_i}^{(k_i)}(x_{\ell'_{T_i}}(\tau_0)), \tag{G.5}$$

where $x_{\ell'_{T_i}}(\tau_0) = x^0_{\ell'_{T_i}} := \boldsymbol{\omega} \cdot \boldsymbol{\nu}^0_{\ell'_{T_i}}$; the choice of such a notation will be clear later (see after (G.12)). By using (G.3) for $k = k_i$ and $n = n_i$, we decompose each $\mathcal{M}_{n_i}^{(k_i)}(x_i(\tau_0))$ into a sum over self-energy clusters, that we still denote by T_i, and associate with each of them a label $\delta_{T_i} \in \{\mathcal{L}, \mathcal{R}\}$, where $\delta_{T_i} = \mathcal{L}$ for $T_i \in \mathfrak{S}^{k_i}_{>n_i}$ and $\delta_{T_i} = \mathcal{R}$ for $T_i \in \mathfrak{S}^{k_i}_{\leq n_i}$. Then we can write

$$\sum_{T' \in \mathfrak{F}(T)} \mathcal{L}_{T'} = - \sum_{\substack{T_i \in \mathfrak{S}^{k_i} \\ i=1,\dots,|\mathcal{T}_1(T)|}} \mathscr{V}_{T^*}(0) \Bigg(\prod_{\substack{i=1 \\ \delta_{T_i} = \mathcal{R}}}^{|\mathcal{T}_1(T)|} \mathcal{R}_{T_i}(x^0_{\ell'_{T_i}}) \Bigg) \Bigg(\prod_{\substack{i=1 \\ \delta_{T_i} = \mathcal{L}}}^{|\mathcal{T}_1(T)|} \mathcal{L}_{T_i} \Bigg), \tag{G.6}$$

where

$$\mathscr{V}_{T^*}(0) = \Bigg(\prod_{v \in N(T^*)} \mathcal{F}_v \Bigg) \Bigg(\prod_{\ell \in L(T^*)} \mathcal{G}_\ell(x^0_\ell) \Bigg). \tag{G.7}$$

Define $\widetilde{T}$ as the set of nodes and lines such that $N(\widetilde{T}) = N(T^*) \cup N(T_1) \cup \ldots \cup N(T_{|\mathcal{T}_1(T)|})$ and $L(\widetilde{T}) = L(T^*) \cup L(T_1) \cup \ldots \cup L(T_{|\mathcal{T}_1(T)|})$ and set $\mathcal{T}_1(\widetilde{T}) = \{T_i\}_{i=1}^{|\mathcal{T}_1(T)|}$. Note that if $\delta_{T_i} = \mathcal{L}$ then T_i contains at least one line with scale $\geq n_i + 1$, so that T_i cannot be considered as a cluster of T' for any $T' \in \mathfrak{F}(T)$; on the contrary if $\delta_{T_i} = \mathcal{R}$ then T_i is a cluster of T' for some $T' \in \mathfrak{F}(T)$. In both cases the external lines ℓ_{T_i} and ℓ'_{T_i} of T_i belong to $L(T^*)$; in particular $x_{\ell'_{T_i}}$ is fixed once and for all, independent of T_i. We call a *replacement* the operation which, given T, generates all sets $\widetilde{T}$ which are summed over in (G.6): each $\widetilde{T}$ can be imagined as obtained from a self-energy cluster $T' \in \mathfrak{F}(T)$ by (1) replacing some self-energy clusters $T_i \in \mathcal{T}_1(T)$ with new sets (still called T_i) which are no longer clusters of T' and (2) substituing all $\mathscr{V}_{T_i}(x_{\ell'_{T_i}}(\tau_0))$ with $\mathcal{L}_{T_i}$, when T_i is not a cluster of T', and with $\mathcal{R}_{T_i}(x_{\ell'_{T_i}}(\tau_0))$, when T_i is still a cluster of T'. By passing from the sum over $T' \in \mathfrak{F}(T)$ to the sum over the sets $\widetilde{T}$ in (G.6), we group together various contributions to exploit the cancellations and at the end everything is decomposed again into a sum over single sets $\widetilde{T}$.

Next, we consider a contribution

$$\mathcal{R}_T(x) = x^2 \int_0^1 \mathrm{d}\tau_0 \; (1 - \tau_0)\, \partial_x^2 \mathscr{V}_T(\tau_0 x)$$

to (G.3). We write

$$\begin{aligned} \partial_x^2 \mathscr{V}_T(\tau_0 x) = & \sum_{\ell_1 \neq \ell_2 \in L(T)} \Big(\partial_x \mathcal{G}_{n_{\ell_1}}(x_{\ell_1}(\tau_0)) \Big) \Big(\partial_x \mathcal{G}_{n_{\ell_2}}(x_{\ell_2}(\tau_0)) \Big) \\ & \times \Bigg(\prod_{\ell \in L(T) \setminus \{\ell_1, \ell_2\}} \mathcal{G}_{n_\ell}(x_\ell(\tau_0)) \Bigg) \Bigg(\prod_{v \in N(T)} \mathcal{F}_v \Bigg) \\ & + \sum_{\ell_1 \in L(T)} \Big(\partial_x^2 \mathcal{G}_{n_{\ell_1}}(x_{\ell_1}(\tau_0)) \Big) \Bigg(\prod_{\ell \in L(T) \setminus \{\ell_1\}} \mathcal{G}_{n_\ell}(x_\ell(\tau_0)) \Bigg) \Bigg(\prod_{v \in N(T)} \mathcal{F}_v \Bigg), \end{aligned}$$

where $x_\ell(\tau_0) = x^0_\ell + \tau_0 x$ if $\ell \in \mathcal{P}_T$ and $x_\ell = x^0_\ell$ otherwise, with $x^0_\ell := \boldsymbol{\omega} \cdot \boldsymbol{\nu}^0_\ell$. To simplify the notations we associate with each line $\ell \in L(T)$ a label $d_\ell \in \{0, 1, 2\}$, which denotes the number of derivatives acting on the corresponding propagator, and set

$$\overline{\mathcal{G}}_\ell(x) = \partial_x^{d_\ell} \mathcal{G}_{n_\ell}(x); \tag{G.8}$$

then we rewrite $\partial_x^2 \mathscr{V}_T(\tau_0 x)$ as

$$\partial_x^2 \mathscr{V}_T(\tau_0 x) = \sum_{\ell_1,\ell_2 \in L(T)} \Bigg(\prod_{\ell \in L(T)} \overline{\mathcal{G}}_\ell(x_\ell(\tau_0)) \Bigg) \Bigg(\prod_{v \in N(T)} \mathcal{F}_v \Bigg), \tag{G.9}$$

where the constraint $\sum_{\ell \in L(T)} d_\ell = 2$ is understood.

For $i = 1, 2$ there exist p_i self-energy clusters $T_0^{(i)}, T_1^{(i)}, \ldots, T_{p_i-1}^{(i)}$ such that, setting $T_0^{(1)} = T_0^{(2)} = T$, $\{T_j^{(i)}\}_{j=0}^{p_i-1}$ is the cloud of the line ℓ_i (note that the two lines ℓ_1 and ℓ_2 may coincide: in such a case one has only one cloud). With each self-energy cluster T' of the internal clouds of ℓ_1 or ℓ_2 we associate a label $\delta_{T'} = 0$: we denote by $\mathcal{T}_0^*(T)$ the set of such self-energy clusters, i.e. $\mathcal{T}_0^*(T) = \{T_j^{(i)} : i = 1, 2,\ j = 1, \ldots, p_i - 1\}$. Set $\mathcal{T}^*(T) = \mathcal{T}(T) \backslash \mathcal{T}_0^*(T)$ and denote by $\mathcal{T}_1^*(T)$ the set of maximal self-energy clusters in $\mathcal{T}^*(T)$: one can think of $\mathcal{T}_1^*(T)$ as the set of self-energy clusters which become maximal in T when ignoring the internal clouds of ℓ_1 and ℓ_2. Finally denote by T^* the set (of nodes and lines) obtained from the self-energy cluster T by removing all nodes and lines belonging to the self-energy clusters $T' \in \mathcal{T}_1^*(T)$. Note that both $\mathcal{T}_1^*(T)$ and T^* depend on ℓ_1 and ℓ_2. Call $\mathfrak{F}(T)$ the set of all self-energy clusters $T' \in \mathfrak{S}_{\le n}^k$ obtained from T by replacing each $T_i \in \mathcal{T}_1^*(T)$ with any self-energy cluster with the same order k_i as T_i and scale $\le n_i$, with $n_i + 1 = n_{T_i}$. Of course $\mathfrak{F}(T)$ too depends on ℓ_1 and ℓ_2; on the other hand one has $T^* = (T')^*$ for all $T' \in \mathfrak{F}(T)$. Then, if we sum together all contributions we obtain by choosing the lines $\ell_1, \ell_2 \in L(T)$ and the self-energy clusters T' in the corresponding set $\mathfrak{F}(T)$, we have

$$\begin{aligned} &\sum_{\ell_1,\ell_2 \in L(T)} \sum_{T' \in \mathfrak{F}(T)} x^2 \int_0^1 \mathrm{d}\tau_0\, (1-\tau_0) \Bigg(\prod_{v \in N(T')} \mathcal{F}_v \Bigg) \Bigg(\prod_{\ell \in L(T')} \overline{\mathcal{G}}_\ell(x_\ell(\tau_0)) \Bigg) \\ &= \sum_{\ell_1,\ell_2 \in L(T)} \mathscr{V}_{T^*}(x) \Bigg(\prod_{i=1}^{|\mathcal{T}_1^*(T)|} \mathcal{M}_{n_i}^{(k_i)}(x_{\ell'_{T_i}}(\tau_0)) \Bigg), \end{aligned} \tag{G.10}$$

where

$$\mathscr{V}_{T^*}(x) = x^2 \int_0^1 \mathrm{d}\tau_0\, (1-\tau_0) \Bigg(\prod_{v \in N(T^*)} \mathcal{F}_v \Bigg) \Bigg(\prod_{\ell \in L(T^*)} \overline{\mathcal{G}}_\ell(x_\ell(\tau_0)) \Bigg). \tag{G.11}$$

Again, by using (G.3), we decompose each $\mathcal{M}_{n_i}^{(k_i)}(x_{\ell'_{T_i}}(\tau_0))$ into a sum over self-energy clusters, still denoted by T_i, and we associate with each of them a label $\delta_{T_i} \in \{\mathcal{L}, \mathcal{R}\}$. This leads to

$$\begin{aligned} &\sum_{\ell_1,\ell_2 \in L(T)} \sum_{T' \in \mathfrak{F}(T)} x^2 \int_0^1 \mathrm{d}\tau_0\, (1-\tau_0) \Bigg(\prod_{v \in N(T')} \mathcal{F}_v \Bigg) \Bigg(\prod_{\ell \in L(T')} \overline{\mathcal{G}}_\ell(x_\ell(\tau_0)) \Bigg) \\ &= \sum_{\ell_1,\ell_2 \in L(T)} \sum_{\substack{T_i \in \mathfrak{S}^{k_i} \\ i=1,\ldots,|\mathcal{T}_1^*(T)|}} \mathscr{V}_{T^*}(x) \Bigg(\prod_{\substack{i=1 \\ \delta_{T_i}=\mathcal{R}}}^{|\mathcal{T}_1^*(T)|} \mathcal{R}_{T_i}(x_{\ell'_{T_i}}(\tau_0)) \Bigg) \Bigg(\prod_{\substack{i=1 \\ \delta_{T_i}=\mathcal{L}}}^{|\mathcal{T}_1^*(T)|} \mathcal{L}_{T_i} \Bigg), \end{aligned} \tag{G.12}$$

where $\mathscr{V}_{T^*}(x)$ is defined in (G.11). Both (G.7) and (G.11) contain also the product of the propagators of the resonant lines in $L(T^*)$. We define the sets $\widetilde{T}$ and $\mathcal{T}_1(\widetilde{T})$ as after (G.7), with $\mathcal{T}_1^*(T)$ replacing $\mathcal{T}_1(T)$.

To make the notations uniform, for $\delta_T = \mathcal{L}$ we set $\mathcal{T}_0^*(T) = \emptyset$ and $\tau_0 = 0$, so that $\mathcal{T}_1^*(T) = \mathcal{T}_1(T)$ and $x_{\ell'_{T_i}}(\tau_0) = x^0_{\ell'_{T_i}}$ in such a case. For given $\widetilde{T}$, define $\widehat{T}$ as the self-energy cluster obtained from $\widetilde{T}$ by the following *pruning* operation: we remove each self-energy cluster $T_i \in \mathcal{T}_1(\widetilde{T})$ with $\delta_{T_i} = \mathcal{L}$ and replace it with a node v_i with labels $\boldsymbol{\nu}_{v_i} = \mathbf{0}$, $s_{v_i} = 1$ and node factor $\mathcal{F}_{v_i} = 1$; it turns out to be convenient to associate with such a node v_i two further labels $n_{v_i} = n_{T_i}$ and $k_{v_i} = k(T_i)$.

An important fact is that the values $\mathcal{L}_{T_i}$ factorise in both (G.6) and (G.12). For each self-energy cluster T_i with $\delta_{T_i} = \mathcal{R}$ we apply once more (G.12), where now the roles of T, $\mathcal{T}_1^*(T)$ and T_i are played by, respectively, T_i, $\mathcal{T}_1^*(T_i)$ and $T_{i,j}$, with $j = 1, \ldots, |\mathcal{T}_1^*(T_i)|$. For each $i = 1, \ldots, |\mathcal{T}_1^*(T)|$ such that $\delta_{T_i} = \mathcal{R}$ we introduce the sets $\widetilde{T}_i$ and construct the sets $\widehat{T}_i$ by pruning $\widetilde{T}_i$. Again, the values $\mathcal{L}_{T_{i,j}}$ corresponding to the self-energy clusters $T_{i,j}$ with $\delta_{T_{i,j}} = \mathcal{L}$ factorise, while the values corresponding to the self-energy clusters $T_{i,j}$ with $\delta_{T_{i,j}} = \mathcal{R}$ can be dealt with by relying again on (G.12): once more we apply the replacement and pruning operations, i.e. we introduce the sets $\widetilde{T}_{i,j}$ as before and construct the sets $\widehat{T}_{i,j}$ by pruning $\widetilde{T}_{i,j}$. And so on: at each step we first apply the replacement operation whenever there is a self-energy cluster T' with $\delta_{T'} = \mathcal{R}$, so obtaining a new set $\widetilde{T}'$, then we apply the pruning operation to $\widetilde{T}'$ by replacing all self-energy clusters $T'' \in \mathcal{T}_1(\widetilde{T}')$ with nodes v', with the labels as described above, and so obtaining a self-energy cluster $\widehat{T}'$. The factorising values $\mathcal{L}_{T'}$ can be treated in the same way, by applying iteratively the replacement and pruning operations.

Furthermore, at each step the order of the self-energy clusters has decreased, so that eventually the procedure stops. Therefore we end up with a sum of terms, each of which is given by the product of factors with the following structure. Each factor is the value a self-energy cluster $S^\star$ obtained by successive replacement and pruning operations starting from a self-energy cluster S, with $\delta_S \in \{\mathcal{L}, \mathcal{R}\}$ if $S = T$ and $\delta_S = \mathcal{L}$ otherwise. By construction, all self-energy clusters $T' \in \mathcal{T}(S^\star)\backslash\{S^\star\}$ carry a label $\delta_{T'} \in \{\mathcal{R}, 0\}$. Moreover each $S^\star$ can contain nodes v' with labels $\boldsymbol{\nu}_{v'} = \mathbf{0}$, $s_{v'} = 1$, $k_{v'} \in \mathbb{N}$ and $n_{v'} \geq 0$ and node factor $\mathcal{F}_{v'} = 1$; the label $n_{v'}$ is such that $n_{v'} + 1$ is the minimum between the scales of the lines entering and exiting v'. We call $N_{\mathcal{L}}(S^\star)$ the set of such nodes and $L_{\mathcal{L}}(S^\star)$ the set of lines exiting one such node. Note that the nodes in $N_{\mathcal{L}}(S^\star)$ can be regarded as self-energy clusters on scale -1 so that the lines in $L_{\mathcal{L}}(S^\star)$, which were resonant as lines in $L(S)$, are resonant as lines in $S^\star$ as well. Each node $v' \in N_{\mathcal{L}}(S^\star)$ has been obtained by pruning a self-energy cluster T' with $\delta_{T'} = \mathcal{L}$, $k(T') = k_v$ and scale $\geq n_{v'}$; note that also T', through successive replacement and pruning operations, produces self-energy clusters $(T')^\star$ which can be dealt with as $S^\star$. We define the *depth* $D(S^\star)$ recursively by setting $D(T^\star) = 0$ and if there exists $v' \in N_{\mathcal{L}}(S^\star)$ which has been obtained by pruning a self-energy cluster S' then $D((S')^\star) = D(S^\star) + 1$. Another important remark is that the propagator of each line $\ell \in L(S^\star)$ is differentiated at most twice.

Then the value of each $S^\star$ such that $\delta_S = \mathcal{L}$ is

$$-\mathscr{V}_{S^\star}(0) = -\prod_{\substack{T' \in \mathcal{T}(S^\star) \\ \delta_{S'} = \mathcal{R}}} x^2_{\ell'_{T'}}(\underline{\tau}) \int_0^1 \mathrm{d}\tau_{T'}\,(1 - \tau_{T'}) \Bigg(\prod_{v \in N(S^\star)} \mathcal{F}_v\Bigg)\Bigg(\prod_{\ell \in L(S^\star)} \overline{\mathcal{G}}_\ell(x_\ell(\underline{\tau}))\Bigg), \tag{G.13}$$

where the set of *interpolation parameters* $\underline{\tau} = \{\tau_{T'} : T' \in \mathcal{T}(S^\star)\}$ and the set of *interpolated arguments* $\{x_\ell(\underline{\tau})\}_{\ell\in L(S^\star)}$ are defined as follows. If $\delta_{T'} = 0$ then $\tau_{T'} = 1$, if $\delta_{T'} = \mathcal{L}$ then $\tau_{T'} = 0$ and if $\delta_{T'} = \mathcal{R}$ then $\tau_{T'} \in [0, 1]$. Given a line $\ell \in L(S^\star)$ consider its cloud $\mathcal{C}_\ell(S^\star)$ and define $\mathcal{C}^*_\ell(S^\star) = \{T' \in \mathcal{C}_\ell(S^\star) : \ell \in \mathcal{P}_{T'}\}$. If T' is the minimal self-energy cluster containing ℓ such that $\ell \in \mathcal{P}_{T'}$, we set

$$\overline{\boldsymbol{\nu}}^0_\ell = \sum_{\substack{w\in N(T')\\ w\prec\ell}} \boldsymbol{\nu}_w$$

so that $\boldsymbol{\nu}_\ell = \overline{\boldsymbol{\nu}}^0_\ell + \boldsymbol{\nu}_{\ell'_{T'}}$ (in general only if $T' = S^\star$ one has $\overline{\boldsymbol{\nu}}^0_\ell = \boldsymbol{\nu}^0_\ell$, with $\boldsymbol{\nu}^0_\ell$ defined in Remark 3.2). Then $x_\ell(\underline{\tau})$ depends only on the parameters $\tau_{T'}$ with $T' \in \mathcal{C}^*_\ell(S^\star)$ and, if we set $\mathcal{C}^*_\ell(S^\star) = \{T_0, T_1, \dots, T_p\}$,

$$x_\ell(\underline{\tau}) = \overline{x}^0_\ell + \tau_p\left(\overline{x}^0_{\ell_p} + \tau_{p-1}\left(\overline{x}^0_{\ell_{p-1}} + \tau_{p-2}\left(\dots + \tau_0\overline{x}_{\ell_0}\right)\right)\right), \tag{G.14}$$

where we have shortened $\tau_i = \tau_{T_i}$ and $\ell_i = \ell'_{T_i}$ and set $\overline{x}^0_{\ell_i} = \boldsymbol{\omega}\cdot\overline{\boldsymbol{\nu}}^0_{\ell_i}$.

For each set $S^\star$ one finds the bound

$$\left|\mathscr{V}_{S^\star}(0)\right| \le C_1^{k(S^\star)} e^{-(\xi/2)K(S^\star)}\left(\prod_{\ell\in L_{\mathcal{L}}(S^\star)} \alpha_{m_{n_\ell}}(\boldsymbol{\omega})^{-2}\right), \tag{G.15}$$

for some positive constant C_1. To obtain (G.15) we can reason as follows.

Each label $\delta_{T'} = \mathcal{R}$ means that either there are two lines $\ell_1, \ell_2 \in L((T')^*)$ with $d_{\ell_1} = d_{\ell_2} = 1$ or one line $\ell_1 \in L((T')^*)$ with $d_{\ell_1} = 2$. Then, when bounding the product of propagators, with respect to the bound

$$\prod_{\ell\in L(S)} c_0\alpha_{m_{n_\ell}}(\boldsymbol{\omega})^{-2},$$

with $c_0 = 32^2$, for each T' with $\delta_{T'} = \mathcal{R}$ we have an extra factor

$$c_1\,\alpha_{m_{n_{\ell_1}}}(\boldsymbol{\omega})^{-1}\alpha_{m_{n_{\ell_2}}}(\boldsymbol{\omega})^{-1}|x_{\ell_{T'}}|^2, \tag{G.16}$$

where $\alpha_{m_{n_{\ell_i}}}(\boldsymbol{\omega})^{-1}$ is due to the derivative acting on the line ℓ_i and c_1 is a suitable constant (ℓ_1 and ℓ_2 may coincide). On the contrary we have no gain factors corresponding to lines exiting the self-energy clusters T' with $\delta_{T'} = 0$; moreover all resonant lines can be differentiated once or twice. In order to deal with all such lines we need some preliminary results.

Lemma G.1. *Given a self-energy cluster T such that $\mathscr{V}_T(x) \neq 0$, if $\ell \in L(T)$ is a resonant line, let T' be the minimal self-energy cluster containing ℓ. Then there is at least one non-resonant line $\ell' \in L(T')$ with $\boldsymbol{\nu}_{\ell'} = \boldsymbol{\nu}_\ell$ and hence $\zeta_{\ell'} = \zeta_\ell$.*

Proof. If ℓ is a resonant line there are $p \ge 1$ self-energy clusters $T_1, \dots, T_p$ with $\ell'_{T_i} = \ell_{T_{i+1}}$ for $i = 1, \dots, p-1$, such that $\ell = \ell_{T_j}$ for some $j = 1, \dots, p$ and ℓ'_{T_p} is non-resonant. Then $\ell'_{T_i} \in L(T')$ for $i = j, \dots, p$: otherwise there would be $j' \in \{j, \dots, p\}$ such that $\ell_{T_{j'}} \in L(T')$ and $\ell'_{T_{j'}} \notin L(T')$, and hence $\boldsymbol{\nu}^0_{\ell_{T_{j'}}} = \mathbf{0}$, so that T' would not be a self-energy cluster. In particular $\ell'_{T_p} \in L(T')$. Obviously $\boldsymbol{\nu}_{\ell_{T'_p}} = \boldsymbol{\nu}_\ell$ and $\zeta_{\ell'_{T_p}} = \zeta_\ell$, so that the assertion follows. □

Given a resonant line ℓ call $\lambda(\ell)$ the non-resonant line which is associated with ℓ by Lemma G.1. Of course the application $\ell \to \lambda(\ell)$ is not necessarily injective. Denote by $L_R(S^\star)$ the set of resonant lines in $S^\star$ and set $L_{NR}(S^\star) = L(S^\star) \backslash L_R(S^\star)$. Set also $L_D(S^\star) = \{\ell \in L_R(S^\star) : d_\ell > 0\}$, $L_0(S^\star) = \{\ell \in L_R(S^\star) : \ell = \ell_{T'}$ for some $T' \in \mathcal{T}(S^\star)$ with $\delta_{T'} = 0\}$ and $L_R^*(S^\star) = L_0(S^\star) \cup L_D(S^\star)$.

Lemma G.2. *Let $\ell_1, \ldots, \ell_p \in L_R^*(S^\star)$ and, for all $i = 1, \ldots, p$ denote by T_i' the minimal self-energy cluster containing ℓ_i. If $T_1' = \ldots = T_p' =: T'$ then $p \le 2$ and $d_{\ell_1} + \cdots + d_{\ell_p} \le 2$.*

Proof. Let $T'' \supseteq T'$ the minimal self-energy cluster containing the lines $\ell_1, \ldots, \ell_p$ and such that $\delta_{T''} = \mathcal{R}$. Then there are at most either two lines $\ell, \ell' \in L(T')$ such that $d_\ell = d_{\ell'} = 1$ or one line $\ell \in L(T')$ such that $d_\ell = 2$. □

Define the multiplicity (function) of a non-injective map as the cardinality of its pre-image sets [14,33].

Lemma G.3. *There exists an application $\Lambda : L_R^*(S^\star) \to L_{NR}(S^\star)$ with multiplicity at most 2 such that $\zeta_\ell = \zeta_{\Lambda(\ell)}$.*

Proof. If $\ell \in L_R^*(S^\star)$ and T' is the minimal self-energy cluster containing ℓ, by Lemma G.1, there is at least one line ℓ' such that $\ell' = \lambda(\ell)$, i.e. such that $\ell' \in L_{NR}(S^\star)$, $\ell' \in L(T')$ and $\zeta_{\ell'} = \zeta_\ell$. By Lemma G.2 there can be at most two lines $\ell_1, \ell_2 \in L_R^*(S^\star)$ such that, if T_1' and T_2' denote the minimal self-energy clusters which contain ℓ_1 and ℓ_2, respectively, then $T_1' = T_2'$. Therefore there are at most two resonant lines $\ell_1, \ell_2 \in L_R^*(S^\star)$ such that $\lambda(\ell_1) = \lambda(\ell_2)$. □

Let Λ be as in Lemma G.3. Define $L_{NR}^*(S^\star) = \Lambda(L_R^*(S^\star))$. By Lemma G.3 for $\ell \in L_{NR}^*(S^\star)$ the set $\Lambda^{-1}(\ell)$ contains at most two elements. Finally for $\ell \in L_R^*(S^\star)$ define $\sigma_\ell^* = 0$ if ℓ exits a self-energy cluster T' with $\delta_{T'} = 0$ and $\sigma_\ell^* = 2$ if ℓ exits a self-energy cluster T' with $\delta_{T'} = \mathcal{R}$.

Lemma G.4. *For all $\ell \in L_{NR}^*(S^\star)$ one has*

$$|\overline{\mathcal{G}}_\ell| \prod_{\ell' \in \Lambda^{-1}(\ell)} (\boldsymbol{\omega} \cdot \boldsymbol{\nu}_{\ell'})^{\sigma_{\ell'}^*} |\overline{\mathcal{G}}_{\ell'}| \le c_2\, \alpha_{m_{n_\ell}}(\boldsymbol{\omega})^{-(2+a)}, \tag{G.17}$$

with $a = 4$ and $c_2 > 0$.

Proof. If $\Lambda^{-1}(\ell)$ contains only one line ℓ', then, if $\ell' \in L_0(S^\star)$ one has $d_{\ell'} \le 1$ and $d_\ell + d_{\ell'} \le 1$, so that (G.17) follows with $a = 3$. If $\ell' \in L_D(S^\star)$ exits a self-energy cluster T' with $\delta_{T'} = \mathcal{R}$, then $d_{\ell'} \le 2$ and $d_\ell + d_{\ell'} \le 2$: hence (G.17) follows with $a = 2$.

If $\Lambda^{-1}(\ell)$ contains two distinct lines ℓ_1' and ℓ_2', we distinguish between the following cases: if $\ell_1', \ell_2' \in L_0(S^\star)$ then $d_\ell = d_{\ell_1'} = d_{\ell_2'} = 0$, and hence the bound follows with $a = 4$; if both lines ℓ_1' and ℓ_2' exit self-energy clusters with label $\mathcal{R}$ then $d_{\ell_1'} + d_{\ell_2'} \le 2$, so that one finds the bound (G.17) with $a = 2$; if $\ell_1' \in L_0(S^\star)$ while $\ell_2' \in L_D(S^\star)$ exits a self-energy cluster T' with $\delta_{T'} = \mathcal{R}$, then $d_{\ell_2'} \le 1$, $d_{\ell_1'} + d_{\ell_2'} \le 1$ and $d_\ell + d_{\ell_1'} + d_{\ell_2'} \le 1$, so that the bound follows once more with $a = 3$. □

Then (G.16) and Lemma G.4 imply

$$\Bigg(\prod_{\ell\in L_{NR}(S^\star)}|\overline{\mathcal{G}}_\ell|\Bigg)\Bigg(\prod_{\substack{T'\in\mathcal{T}(S^\star)\\ \delta_{T'}=\mathcal{R}}}|x_{\ell'_{T'}}|^2\Bigg)\Bigg(\prod_{\ell\in L_R^*(S^\star)}|\overline{\mathcal{G}}_\ell|\Bigg)\le\prod_{\ell\in L_{NR}(S^\star)}\tilde{C}\,\alpha_{m_{n_\ell}}(\boldsymbol{\omega})^{-6},$$

for some positive constant $\tilde{C}$. The conclusion is that, at the price of replacing the bound of the propagator of each line $\ell\in L_{NR}(S^\star)$ with $\tilde{C}\alpha_{m_{n_\ell}}(\boldsymbol{\omega})^{-6}$ to take into account the extra derivatives and the self-energy clusters T' with $\delta_{T'}=0$, we can bound the product of all propagators in terms of the product of propagators of the non-resonant lines times the product of the propagators of the lines in $L_{\mathcal{L}}(S^\star)$. Then, by using Lemma 6.3, the bound (G.15) follows.

To deal with the factors $\alpha_{m_{n_\ell}}(\boldsymbol{\omega})^{-2}$ of the lines $\ell\in L_{\mathcal{L}}(S^\star)$ in (G.15), we proceed iteratively by starting from the self-energy clusters $(S')^\star$ with label $\delta_{S'}=\mathcal{L}$ which have highest depth, say D. If $(S')^\star$ is one of such sets then $L_{\mathcal{L}}((S')^\star)=\emptyset$ and hence the bound (G.15) follows with the last product replaced by 1:

$$\big|\mathscr{V}_{(S')^\star}(0)\big|\le C_1^{k((S')^\star)}e^{-(\xi/2)K((S')^\star)}. \tag{G.18}$$

If $(S')^\star$ has depth $D-1$, then each node $v''\in N_{\mathcal{L}}((S')^\star)$ has been obtained by pruning a self-energy cluster S'' such that $N_{\mathcal{L}}((S'')^\star)=\emptyset$, then we apply the bound (G.18) to each such $(S'')^\star$. In particular, we can extract a factor $e^{-(\xi/4)K((S'')^\star)}\le e^{-(\xi/8)2^{m_{n_{S''}}}}$ from each of them and, by exploiting that $n_{\ell'_{S''}}\ge n_{S''}+1$, use it to compensate the corresponding factor $\alpha_{m_{\ell'_{S''}}}(\boldsymbol{\omega})^{-2}$. And so on, iteratively, up to the self-energy cluster $S^\star$ itself.

We have still to sum over all the possible self-energy clusters T. To take into account the sum over the scale labels n_ℓ, $\ell\in L(T)$, simply recall that for each momentum $\boldsymbol{\nu}_\ell$ only two scale labels are allowed (see Remark 3.7). To sum over the mode labels $\boldsymbol{\nu}_v$, $v\in N(T)$, we can neglect all constraints and use a factor $e^{-(\xi/4)|\boldsymbol{\nu}_v|}$ for each node $v\in N(T)$. Finally, we have to sum over all possible 'shapes' of self-energy clusters of order k, that is over all possible unlabelled self-energy clusters of order k: this is bounded as C^k for some constant C. In conclusion, for all $p>n$ we have $\big|M_p^{(k)}(0)\big|\le D_1^k e^{-D_2 2^{m_p}}$, for some positive constants D_1 and D_2. Hence the assertion follows from (G.2), as far as in (G.1) the contribution $-\mathcal{M}_{>n}^{(k)}(0)$ is concerned. To take into account the other contributions we reason in the same way, with the only difference that a label $\delta_T=\mathcal{R}$ is associated with T, so that the corresponding value $\mathscr{V}_{T^*}(x)$ is expressed as in (G.13), but with $S=T$ and the second product over all $T'\in\mathcal{T}(S^\star)$. This produces once more the desired bound, so that the result follows.

H. Existence of the Formal Power Series

Here we shall prove that the formal power series (3.1) is well defined for all $k\ge 1$ and all $\boldsymbol{\nu}\in\mathbb{Z}_*^d$. Set

$$\varepsilon_n=\varepsilon_n(\boldsymbol{\omega}):=\frac{1}{2^n}\log\frac{1}{\alpha_n(\boldsymbol{\omega})}, \tag{H.1}$$

and note that by Hypothesis 1 $\varepsilon_n\to 0$ as $n\to\infty$.

To study the formal power series we can use the tree expansion introduced in Sect. 6, simply without imposing the conditions $G^{(k)}(\beta_0) = 0$. So we write the coefficients $b_\nu^{(k)}$ as in (6.4a). For all $\theta \in \Theta_{k,\boldsymbol{\nu}}$ one has

$$
\begin{aligned}
|\mathscr{V}(\theta)| &\le \left(\prod_{v\in N(\theta)} e^{-\xi|\boldsymbol{\nu}_v|}\right)\left(\prod_{\ell\in L(\theta)} \frac{1}{|\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell|^2}\right) = e^{-\xi K(\theta)}\left(\prod_{\ell\in L(\theta)} \frac{1}{|\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell|^2}\right) \\
&= e^{-\xi K(\theta)/2}\left(e^{-\xi K(\theta)/2k}\right)^k \left(\prod_{\ell\in L(\theta)} \frac{1}{|\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell|^2}\right) \\
&\le e^{-\xi|\boldsymbol{\nu}|/2} \prod_{\ell\in L(\theta)} e^{-\xi|\boldsymbol{\nu}_\ell|/2k} \frac{1}{|\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell|^2} \\
&\le e^{-\xi|\boldsymbol{\nu}|/2} \prod_{\ell\in L(\theta)} e^{-\xi 2^{\overline{n}_\ell}/4k} \frac{1}{\alpha_{\overline{n}_\ell}(\boldsymbol{\omega})^2} = e^{-\xi|\boldsymbol{\nu}|/2} \prod_{\ell\in L(\theta)} e^{(-\xi/4k+2\varepsilon_{\overline{n}_\ell})2^{\overline{n}_\ell}}, \qquad \text{(H.2)}
\end{aligned}
$$

where we have set $\overline{n}_\ell = n(\boldsymbol{\nu}_\ell) := \inf\{n \ge 0 \,:\, |\boldsymbol{\nu}_\ell| \le 2^n\}$. But then, since the sum over all the shapes and all the labels except the mode labels is bounded by a constant to the power k, one has

$$
\sum_{\theta\in\Theta_{k,\boldsymbol{\nu}}} |\mathscr{V}(\theta)| \le e^{-\xi|\boldsymbol{\nu}|/2} C^k \left(\sum_{n\ge 0} e^{(-\xi/4k+2\varepsilon_n)2^n}\right)^k \le e^{-\xi|\boldsymbol{\nu}|/2} C^k B(k)^k, \qquad \text{(H.3)}
$$

where C is a suitable constant and $B(k)$ is a constant depending on k. Therefore the assertion follows.

Remark H.1. Of course $B(k)$ grows with k (for instance if $\boldsymbol{\omega}$ is Diophantine one has $B(k) \approx k$) and hence the bound (H.3) is not enough to obtain the convergence of the power series.

References

1. Andronov, A.A., Leontovich, E.A., Gordon, I.I., Maĭer, A.G.: *Theory of bifurcations of dynamic systems on a plane*. Jerusalem-London: Halsted Press/Israel Program for Scientific Translations, 1973
2. Berretti, A., Gentile, G.: Bryuno function and the standard map. Commun. Math. Phys. **220**(3), 623–656 (2001)
3. Blekman, I.I.: *Synchronization in Science and Technology*. New York: ASME Press, 1988
4. Boccaletti, S., Kurths, J., Osipov, G., Valladares, D.L., Zhou, C.S.: The synchronization of chaotic systems. Phys. Rep. **366**, 1–101 (2002)
5. Bricmont, J., Gawędzki, K., Kupiainen, A.: KAM theorem and quantum field theory. Commun. Math. Phys. **201**(3), 699–727 (1999)
6. Broer, H.W., Hanssmann, H., You, J.: Bifurcations of normally parabolic tori in Hamiltonian systems. Nonlinearity **18**(4), 1735–1769 (2005)
7. Bryuno, A.D.: *Analytic form of differential equations. I, II*. Trudy Moskov. Mat. Obšč. **25**, 119–262 (1971), ibid. **26**, 199–239 (1972). English translations: Trans. Moscow Math. Soc. **25**(1971), 131–288 (1973); ibid. **26**(1972), 199–239 (1974)
8. Cheng, Ch.-Q.: Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems. Commun. Math. Phys. **177**(3), 529–559 (1996)
9. Chow, S.-N., Hale, J.K.: *Methods of bifurcation theory*. Grundlehren der Mathematischen Wissenschaften **251**, New York-Berlin: Springer-Verlag, 1982
10. Corsi, L., Gentile, G.: Melnikov theory to all orders and Puiseux series for subharmonic solutions. J. Math. Phys. **49**(11), 112701 (2008)

11. Davie, A.M.: The critical function for the semistandard map. Nonlinearity **7**(1), 219–229 (1994)
12. De Simone, E., Kupiainen, A.: The KAM theorem and renormalization group. Erg. Th. Dyna. Sys. **29**(2), 419–431 (2009)
13. Eliasson, L.H.: *Absolutely convergent series expansions for quasi periodic motions*. Math. Phys. Electron. J. **2**, Paper 4, 33 pp. (1996) (electronic)
14. Franz, A.: Hausdorff dimension estimates for non-injective maps using the cardinality of the pre-image sets. Nonlinearity **13**(5), 1425–1438 (2000)
15. Gallavotti, G.: Twistless KAM tori. Commun. Math. Phys. **164**(1), 145–156 (1994)
16. Gallavotti, G., Bonetto, F., Gentile, G.: *Aspects of ergodic, qualitative and statistical theory of motion*. Texts and Monographs in Physics, Berlin: Springer-Verlag, 2004
17. Gallavotti, G., Gentile, G.: Hyperbolic low-dimensional invariant tori and summation of divergent series. Commun. Math. Phys. **227**(3), 421–460 (2002)
18. Gallavotti, G., Gentile, G., Giuliani, A.: Fractional Lindstedt series. J. Math. Phys. **47**(1), 012702, 33 pp. (2006)
19. Gallavotti, G., Gentile, G., Giuliani, A.: Resonances within chaos. Chaos **22**(2), 026108 (2012)
20. Gallavotti, G., Gentile, G., Mastropietro, V.: Field theory and KAM tori. Math. Phys. Electron. J. **1**, Paper 5, 13 pp. (1995) (electronic)
21. Gentile, G.: Quasi-periodic solutions for two level systems. Commun. Math. Phys. **242**(1-2), 221–250 (2003)
22. Gentile, G.: Resummation of perturbation series and reducibility for Bryuno skew-product flows. J. Stat. Phys. **125**(2), 321–361 (2006)
23. Gentile, G.: Degenerate lower-dimensional tori under the Bryuno condition. Erg. Th. Dyna. Sys. **27**(2), 427–457 (2007)
24. Gentile, G.: Quasi-periodic motions in strongly dissipative forced systems. Erg. Th. Dyna. Sys. **30**(5), 1457–1469 (2010)
25. Gentile, G.: Construction of quasi-periodic response solutions in forced strongly dissipative systems. Forum Math. **24**(4), 791–808 (2012)
26. Gentile, G., Cortes, D.A., Barata, J.C.A.: Stability for quasi-periodically perturbed Hill's equations. Commun. Math. Phys. **260**(2), 403–443 (2005)
27. Gentile, G., Gallavotti, G.: Degenerate elliptic resonances. Commun. Math. Phys. **257**(2), 319–362 (2005)
28. Guckenheimer, J., Holmes, Ph.: *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Applied Mathematical Sciences **42**, New York: Springer-Verlag, 1990
29. Lochak, P.: Tores invariants à torsion évanescente dans les systèmes hamiltoniens proches de l'intégrable. C. R. Acad. Sci. Paris Sér. I Math. **327**(9), 833–836 (1998)
30. Mastropietro, V.: *Non-perturbative renormalization*. Hackensack, NJ: World Scientific Publishing, 2008
31. Pöschel, J.: On invariant manifolds of complex analytic mappings near fixed points. Exposition. Math. **4**(2), 97–109 (1986)
32. Pikovsky, A., Rosenblum, M., Kurths, J.: *Synchronization. A universal concept in nonlinear sciences*. Cambridge Nonlinear Science Series **12**, Cambridge: Cambridge University Press, 2001
33. Reitmann, V.: *Dimension estimates for invariant sets of dynamical systems*. In: Ergodic theory, analysis, and efficient simulation of dynamical systems, Ed. B. Fiedler, Berlin: Springer, 2001, pp. 585–615
34. Ruelle, D., Wilkinson, A.: Absolutely singular dynamical foliations. Commun. Math. Phys. **219**, 481–487 (2001)
35. Shub, M., Wilkinson, A.: Pathological foliations and removable zero exponents. Invent. Math. **139**, 495–508 (2000)
36. Siegel, C.L.: Iteration of analytic functions. Ann. Math. **43**, 607–612 (1942)
37. Yoccoz, J.-C.: Théorème de Siegel, nombres de Bruno et polinômes quadratiques. Astérisque **231**, 3–88 (1995)
38. Zhang, Zh.F., Li, B.Y.: High order Melnikov functions and the problem of uniformity in global bifurcation. Ann. Mat. Pura Appl. (4) **161**, 181–212 (1992)

Communicated by G. Gallavotti