Resummation of perturbation series and reducibility for Bryuno skew-product flows

Guido Gentile

Dipartimento di Matematica, Università di Roma Tre, Roma, I-00146, Italy.
E-mail: gentile@mat.unirom3.it

Abstract

We consider skew-product systems on $\mathbb{T}^d \times \mathrm{SL}(2,\mathbb{R})$ for Bryuno base flows close to constant coefficients, depending on a parameter, in any dimension d, and we prove reducibility for a large measure set of values of the parameter. The proof is based on a resummation procedure of the formal power series for the conjugation, and uses techniques of renormalisation group in quantum field theory.

1 Introduction

Consider the linear differential equation

$$\dot{x} = (\lambda A + \varepsilon f(\boldsymbol{\omega} t))\, x, \tag{1.1}$$

on $\mathrm{SL}(2,\mathbb{R})$, where $\lambda \in [a,b] \subset \mathbb{R}$, ε is a small real parameter, $\boldsymbol{\omega} \in \mathbb{R}^d$ is a vector with rationally independent components, and $A, f \in \mathfrak{sl}(2,\mathbb{R})$, with A is a constant matrix and f an analytic function periodic in its arguments. We say that f is quasi-periodic in time t.

Reducibility for (1.1) means the existence of a quasi-periodic change of variables which takes the system into a system with constant coefficients:

$$x = B(\boldsymbol{\omega} t) y, \qquad \dot{y} = A_0 y, \tag{1.2}$$

with $B \in \mathrm{SL}(2,\mathbb{R})$ analytic and $A_0 \in \mathfrak{sl}(2,\mathbb{R})$ constant. In particular if the solution $y(t)$ is periodic then the solution $x(t)$ is quasi-periodic, hence bounded for all times.

A special case of (1.1) is the one-dimensional Schrödinger equation with a weak quasi-periodic potential, or with arbitrary quasi-periodic potential for large energy. By assuming a suitable non-resonance condition on the frequency vector $\boldsymbol{\omega}$, reducibility for ε small enough and for a large measure set of values λ in $[a,b]$ (for which quasi-periodic solutions exist) was proved by Dinaburg and Sinai [5], by using KAM techniques; see also [27] for a review. Weaker non-resonance conditions were shown to be possible by Rüssmann [28], then used by Moser and Pöschel [26] to enlarge the set of values λ for which reducibility can be obtained. Reducibility almost everywhere in λ and for small ε has been obtained by Eliasson [7], for $\boldsymbol{\omega}$ a Diophantine vector.

A brief survey on the problem of reducibility for skew-product systems can be found in [8, 9]. In particular results similar to those by Eliasson, – i.e. reducibility almost everywhere for Diophantine frequency vectors, – in the case of other Lie groups, also not close to constant coefficients, have been obtained by Krikorian[20, 21]. Very recently, Avila and Krikorian [1] proved, by using renormalization techniques, that, if $\boldsymbol{\omega}$ belongs to a subset of full measure of the Diophantine vectors in $d = 2$, for all values of ε and almost everywhere in λ, quasi-periodic Schrödinger cocycles are either reducible or non-uniformly hyperbolic.

Rüssmann's non-resonance condition is weaker than the usual Diophantine one, and is expressed in terms of a suitable approximation function [28, 26]. In $d = 2$ it is equivalent to Bryuno's condition. Bryuno vectors have been explicitly considered in the case of skew products for the first time by Lopes Dias in [25], where in $d = 2$ a normal form theorem (analogous to Lemma 22 below) is proved with renormalisation group techniques. However, the non-resonance condition with the eigenvalues of λA is still assumed to be of Diophantine type. Renormalisation group techniques have been also used in [19] for any d, in the case of Diophantine vectors.

In this paper we consider Bryuno vectors in any dimension, and, for ε small enough, we prove reducibility on a large measure set of values of λ. The estimates we find for the excluded set are much better than those provided by standard KAM methods (cf. for instance [20]). The techniques we use are those of renormalisation group typical of quantum field theory, based on a diagrammatic representation of the equation in terms of trees, and are inspired to those used in [14, 15, 17]. Trees for skew-products were already introduced by Iserles and Nørsett [23, 22], but they used expansions in time, hence not suited for the study of global properties, such as reducibility and quasi-periodicity.

We can formulate our result as follows.

Theorem 1 *Let $A \in \mathfrak{sl}(2, \mathbb{R})$ be a constant matrix with imaginary eigenvalues and $f \in \mathfrak{sl}(2, \mathbb{R})$ an analytic quasi-periodic function of time. Let $\boldsymbol{\omega} \in \mathbb{R}^d$ be a Bryuno vector. Then there exists $\varepsilon_0 > 0$ and $\sigma > 0$ such that for all $|\varepsilon| < \varepsilon_0$ the set of values $\lambda \in [a, b]$ for which the system (1.1) is not reducible is of Lebesgue measure less than* $\mathrm{const}.|\varepsilon|^{\sigma}$.

The proof will proceed through the following steps. In Section 2, we reduce the study of system (1.1) to the study of a system of differential equations in $\mathbb{C}^2$, that we call here the "auxiliary system", and we see that the property for x to have $\det x = 1$ can be interpreted as the existence of a suitable first integral for the new system. Next, in Section 3 we look for a quasi-periodic solution of the auxiliary system: first, we try for solutions in the form of formal power series in ε. However, in order to define such a series, even order by order, we cannot fix λ. Instead, we write $\lambda = \lambda_0 + \mu$, with λ_0 in some interval Λ_0, and we see that for fixed λ_0 there exists a formal power series for μ such that the auxiliary system admits a formal power series solution. Hence the formal series turn out to be well-defined order by order. Moreover, they can be proved to be convergent. This result can be obtained by performing a suitable resummation leading to simpler series, that we shall call renormalised series (for a discussion of the method within the standard KAM theory we refer to [13, 11]). The renormalised series are introduced in Section 4, and in Section 5 are proved to converge to well-defined functions. The latter are analytic in ε and solve the differential equation of the auxiliary system, provided λ_0 is chosen in a subset Λ_0^* of Λ_0. Finally in Section 6 we have to control that the set of values $\lambda \in [a, b]$ for which the above procedure can be followed coincide with $[a, b]$, up to a small measure set.

We conclude with two comments.

Given the system (1.1) one could also consider another problem: fix λ and study for which values of ε (small enough) the system is reducible. This a natural question if, for instance, instead of the Schrödinger equation, one considers Hill's equation, where there is no free parameter other than ε itself. Under suitable (generic) conditions on the potential (like $f_{11,\mathbf{0}} \neq 0$) the problem is of comparable difficulty (cf. [24, 30] for Diophantine $\boldsymbol{\omega}$), and reducibility on a large measure set of values of ε can be proved. But, in general, if no condition at all is assumed on the potential (besides analyticity), further difficulties arise; cf. [16, 4, 12] for similar situations. In particular in [16] Hill's equation perturbed with a small quasi-periodic potential was studied under the standard Diophantine condition, and reducibility for a Cantor set of values of ε was proved.

More generally one can consider skew-products flows on $\mathbb{T}^d \times \mathrm{SL}(n, \mathbb{R})$, for any n (and any d). In principle our techniques apply also in such a case: of course the tree formalism becomes more involved. Also, less smooth potentials can be considered, like in [20, 19, 1], but in the case of Bryuno vectors analyticity is likely to be the optimal regularity condition on the potential.

2 Preliminary considerations

Assume $\lambda \in [a,b] \subset \mathbb{R} \setminus \{0\}$; we shall see later that the condition $0 \notin [a,b]$ can be relaxed (cf. the end of Section 6). Let $A \in \mathfrak{sl}(2,\mathbb{R})$ with imaginary eigenvalues. Possibly renaming a and b we can assume that the eigenvalues be $\pm i$. Let $f : \mathbb{T}^d \to \mathfrak{sl}(2,\mathbb{R})$ be real-analytic, $\boldsymbol{\omega} \in \mathbb{R}^d$ a real vector, and ε a real parameter.

Consider the ordinary differential equation

$$\dot{x} = (\lambda A + \varepsilon f(\boldsymbol{\omega} t))\, x, \tag{2.1}$$

on $\mathrm{SL}(2,\mathbb{R})$.

We can assume that A be of the form

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{2.2}$$

and, through a suitable change of coordinates, we obtain

$$D := MAM^{-1} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad M = \frac{1}{2}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}, \qquad M^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \tag{2.3}$$

Then, for $z = MxM^{-1}$, we find the equation

$$\dot{z} = (\lambda D + \varepsilon g(\boldsymbol{\omega} t))\, z, \tag{2.4}$$

with $g = MfM^{-1}$.

Let us introduce some notations. Given a 2×2 matrix M, we write

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \tag{2.5}$$

and we denote by $[A,B]$ the commutator of the two matrices A and B. For $z \in \mathbb{C}$ denote by z^* the complex conjugate of z. By $\delta_{i,j}$ we denote the Kronecker delta. We set $\mathbb{Z}_+ = \{n \in \mathbb{Z} : n \geq 0\} = \mathbb{N} \cup \{0\}$, and for $d \in \mathbb{N}$ and $\mathbf{0} \in \mathbb{Z}^d$, define $\mathbb{Z}^d_* = \mathbb{Z}^d \setminus \{\mathbf{0}\}$. Given any set $A \subset \mathbb{R}$, we denote by $\mathrm{meas}(A)$ the Lebesgue measure of A.

Lemma 1 *Let $g = MfM^{-1}$, with $f \in \mathfrak{sl}(2,\mathbb{R})$ and M given as in (2.3). Then $g \in \mathfrak{sl}(2,\mathbb{C})$, and one has $g_{11} = g_{22}^*$ and $g_{12} = g_{21}^*$.*

Proof. The property for g to be traceless follows from the fact that $\mathrm{tr}\,(MfM^{-1}) = \mathrm{tr}\, f = 0$. The relations between the entries of g can be checked by a direct computation:

$$\begin{aligned} 2g_{11} &= f_{11} + f_{22} + i\,(f_{12} - f_{21}), \\ 2g_{12} &= f_{11} - f_{22} - i\,(f_{12} + f_{21}), \\ 2g_{21} &= f_{11} - f_{22} + i\,(f_{12} + f_{21}), \\ 2g_{22} &= f_{11} + f_{22} - i\,(f_{12} - f_{21}), \end{aligned} \tag{2.6}$$

where all entries f_{ij} are real. ■

Define

$$\begin{aligned} \mathfrak{M} &:= \{G \in \mathrm{SL}(2,\mathbb{C}) : G_{11} = G_{22}^*, \quad G_{12} = G_{21}^*\}, \\ \mathfrak{m} &:= \{g \in \mathfrak{sl}(2,\mathbb{C}) : g_{11} = g_{22}^*, \quad g_{12} = g_{21}^*\}. \end{aligned} \tag{2.7}$$

It is easy to see that $\mathfrak{M}$ is a subgroup, and $\mathfrak{m}$ is the corresponding Lie algebra.

Lemma 2 *Consider the equation* $\dot z = Sz$*, with* $S = S(t) \in \mathfrak{m}$ *and* $z(0) \in \mathfrak{M}$*. Then* $z(t) \in \mathfrak{M}$ *for all* $t \in \mathbb{R}$ *for which the solution is defined.*

Proof. Write explicitly the equations for the entries of z:

$$\begin{aligned} \dot z_{11} &= S_{11} z_{11} + S_{12} z_{21}, \\ \dot z_{12} &= S_{11} z_{12} + S_{12} z_{22}, \\ \dot z_{21} &= S_{21} z_{11} + S_{22} z_{21} = S_{12}^* z_{11} + S_{11}^* z_{21}, \\ \dot z_{22} &= S_{21} z_{12} + S_{22} z_{22} = S_{12}^* z_{12} + S_{11}^* z_{22}, \end{aligned} \tag{2.8}$$

so that, by setting $w = (w_1, w_2)$, with $w_1 = z_{11} - z_{22}^*$ and $w_2 = z_{21} - z_{12}^*$, one obtains $\dot w = Sw$. If $z(0) \in \mathfrak{M}$ then $w(0) = 0$, so that $w(t) = 0$ for all $t \in \mathbb{R}$. Moreover, if $\delta(t) = \det z(t)$, one finds

$$\dot\delta = (S_{11} + S_{11}^*)\,(z_{11} z_{22} - z_{12} z_{21}) = (S_{11} + S_{11}^*)\,\delta, \tag{2.9}$$

where $S_{11} + S_{11}^* = S_{11} + S_{22} = \operatorname{tr} S = 0$. Hence $\delta(t) = \delta(0) = 1$. ■

Therefore it is not restrictive to consider the differential equation

$$\dot x = (\lambda A + \varepsilon f(\boldsymbol{\omega} t))\, x, \tag{2.10}$$

on $\mathfrak{M}$, with

$$A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad f \in C^\omega(\mathbb{T}^d, \mathfrak{m}), \tag{2.11}$$

and this we shall do henceforth. Write $\lambda = \lambda_0 + \mu$, and set $x = B(\boldsymbol{\omega} t) y$, with y solution of

$$\dot y = \lambda_0 A y, \qquad y(0) = \mathbb{1}, \tag{2.12}$$

that is

$$y(t) = \begin{pmatrix} \mathrm{e}^{i\lambda_0 t} & 0 \\ 0 & \mathrm{e}^{-i\lambda_0 t} \end{pmatrix}. \tag{2.13}$$

Then $B = B(\boldsymbol{\omega} t)$ must solve the differential equation

$$\dot B + \lambda_0 [B, A] = (\varepsilon f + \mu A)\, B, \tag{2.14}$$

and one has $\det B = 1$ if $\det x(0) = 1$.

Write

$$B := 1 + \beta, \qquad \beta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{2.15}$$

Then the following result holds.

Lemma 3 *With the notations in (2.15) one has* $a = d^*$*,* $b = c^*$*, and* $a + d + ad - bc$ *is constant along the flow. If* $\det B = 1$ *then*

$$\operatorname{tr}\beta + \det\beta = a + d + (ad - bc) = 0. \tag{2.16}$$

for all $t \in \mathbb{R}$*.*

Proof. Since $\mathfrak{M}$ is a group and $y \in \mathfrak{M}$, then $B \in \mathfrak{M}$ if $x \in \mathfrak{M}$. More generally, $\det B(\boldsymbol{\omega} t) = \det B(0)$, which means that $\det B = 1 + a + d + ad - bc$ is constant along the flow. By requiring $\det B = 1$ gives (2.16). ■

In terms of β, (2.14) becomes

$$\dot{\beta} + \lambda_0[\beta, A] = (\varepsilon f + \mu A)\,(1+\beta)\,, \tag{2.17}$$

which, written explicitly for the corresponding entries, gives

$$\begin{aligned}
\dot{a} &= \varepsilon f_{11} + i\mu + \varepsilon\,(f_{11}a + f_{12}c) + i\mu\, a,\\
\dot{b} - 2i\lambda_0 b &= \varepsilon f_{12} + \varepsilon\,(f_{11}b + f_{12}d) + i\mu\, b,\\
\dot{c} + 2i\lambda_0 c &= \varepsilon f_{21} + \varepsilon\,(f_{21}a + f_{22}c) - i\mu\, c,\\
\dot{d} &= \varepsilon f_{22} - i\mu + \varepsilon\,(f_{21}b + f_{22}d) - i\mu\, d.
\end{aligned} \tag{2.18}$$

If we use that $d = a^*$ and $b = c^*$, equations (2.18) reduce to two independent equations

$$\begin{aligned}
\dot{a} &= \varepsilon f_{11} + i\mu + \varepsilon\,(f_{11}a + f_{12}c) + i\mu\, a,\\
\dot{c} + 2i\lambda_0 c &= \varepsilon f_{21} + \varepsilon\,(f_{21}a + f_{22}c) - i\mu\, c,
\end{aligned} \tag{2.19}$$

which is the system the we are going to study.

We can view (2.19) as a system of ordinary differential equations on $\mathbb{C}^2$. Such a system admits a first integral, as the following result shows.

Lemma 4 *Given the system (2.19), the function*

$$H = H(a,c) := a + a^* + \left(|a|^2 - |c|^2\right) \tag{2.20}$$

is a constant of motion, that is $\dot{H} = 0$.

Proof. Just note that (2.19) is a rewriting of (2.14). Lemma 3 shows that $\det B$ is a constant of motion. In terms of a and c, this means that (2.20) is conserved along the flow. ■

3 Formal series

For any function F defined on $\mathbb{T}^d$ set, formally,

$$F(\boldsymbol{\psi}) = \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d} \mathrm{e}^{i\boldsymbol{\nu}\cdot\boldsymbol{\psi}} F_{\boldsymbol{\nu}}, \tag{3.1}$$

where $\cdot$ denotes the standard inner product in $\mathbb{R}^d$. If F is analytic the Fourier coefficients $F_{\boldsymbol{\nu}}$ decay exponentially at infinity. In particular if $f \in C^{\omega}(\mathbb{T}^d, \mathfrak{sl})$ there exists two constants F_0 and κ_0 such that $|f_{jj',\boldsymbol{\nu}}| \le F_0 \mathrm{e}^{-\kappa_0|\boldsymbol{\nu}|}$ for $j, j' = 1, 2$.

Assume that $\boldsymbol{\omega} \in \mathbb{R}^d$ is a *Bryuno vector*. This means that, by setting $\alpha_n(\boldsymbol{\omega}) = \inf_{|\boldsymbol{\nu}|\le 2^n} |\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$, one has

$$\mathcal{B}(\boldsymbol{\omega}) := \sum_{n=0}^{\infty} \frac{1}{2^n} \log\frac{1}{\alpha_n(\boldsymbol{\omega})} < \infty. \tag{3.2}$$

In terms of the Fourier coefficients $\beta_{\boldsymbol{\nu}}$, (2.19) gives for $\boldsymbol{\nu} \neq \mathbf{0}$

$$\begin{aligned}
i\boldsymbol{\omega}\cdot\boldsymbol{\nu}\, a_{\boldsymbol{\nu}} &= \varepsilon f_{11,\boldsymbol{\nu}} + \varepsilon\,(f_{11}a + f_{12}c)_{\boldsymbol{\nu}} + i\mu\, a_{\boldsymbol{\nu}},\\
i\,(\boldsymbol{\omega}\cdot\boldsymbol{\nu} + 2\lambda_0)\, c_{\boldsymbol{\nu}} &= \varepsilon f_{21,\boldsymbol{\nu}} + \varepsilon\,(f_{21}a + f_{22}c)_{\boldsymbol{\nu}} - i\mu\, c_{\boldsymbol{\nu}},
\end{aligned} \tag{3.3}$$

and for $\boldsymbol{\nu} = \mathbf{0}$

$$\begin{aligned}
0 &= \varepsilon f_{11,\mathbf{0}} + i\mu + \varepsilon\,(f_{11}a + f_{12}c)_{\mathbf{0}} + i\mu\, a_{\mathbf{0}},\\
2i\lambda_0\, c_{\mathbf{0}} &= \varepsilon f_{21,\mathbf{0}} + \varepsilon\,(f_{21}a + f_{22}c)_{\mathbf{0}} - i\mu\, c_{\mathbf{0}}.
\end{aligned} \tag{3.4}$$

3.1 Recursive equations

Assume $\lambda \neq 0$. We shall see that $\mu = O(\varepsilon)$, so that the assumption is satisfied for all $\lambda \in [a, b]$ if ε is small enough and $0 \notin [a, b]$. In fact it would be enough to require that $\min\{|a|, |b|\}$ be of order $|\varepsilon|^\sigma$; cf. the end of Section 6.

We can write a formal power series in ε for β, by setting

$$\beta = \beta(\boldsymbol{\omega} t) = \sum_{k=1}^{\infty} \varepsilon^k \beta^{(k)}(\boldsymbol{\omega} t), \qquad \beta^{(k)}(\boldsymbol{\psi}) = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^d} \mathrm{e}^{i\boldsymbol{\nu}\cdot\boldsymbol{\psi}} \beta^{(k)}_{\boldsymbol{\nu}}. \tag{3.5}$$

The properties $a = d^*$ and $b = c^*$ imply $a^*_{\boldsymbol{\nu}} = d_{-\boldsymbol{\nu}}$ and $b^*_{\boldsymbol{\nu}} = c_{-\boldsymbol{\nu}}$. In the same way $f \in \mathfrak{m}$ yields $f^*_{11,\boldsymbol{\nu}} = f_{22,-\boldsymbol{\nu}}$, hence $f_{11,\boldsymbol{\nu}} + f^*_{11,-\boldsymbol{\nu}} = 0$, and $f^*_{12,\boldsymbol{\nu}} = f_{21,-\boldsymbol{\nu}}$.

If we write also

$$\mu = \sum_{k=1}^{\infty} \varepsilon^k \mu^{(k)}, \tag{3.6}$$

and we insert (3.5) and (3.6) into (3.3) and (3.4) we find

$$\begin{aligned} a^{(1)}_{\boldsymbol{\nu}} &= -i \frac{f_{11,\boldsymbol{\nu}}}{\boldsymbol{\omega}\cdot\boldsymbol{\nu}}, \\ c^{(1)}_{\boldsymbol{\nu}} &= -i \frac{f_{21,\boldsymbol{\nu}}}{\boldsymbol{\omega}\cdot\boldsymbol{\nu} + 2\lambda_0}, \end{aligned} \tag{3.7}$$

for $k = 1$ and $\boldsymbol{\nu} \neq \mathbf{0}$,

$$\begin{aligned} \mu^{(1)} &= i\, f_{11,\mathbf{0}}, \\ c^{(1)}_{\mathbf{0}} &= -\frac{i}{2\lambda_0}\, f_{21,\mathbf{0}}, \end{aligned} \tag{3.8}$$

for $k = 1$ and $\boldsymbol{\nu} = \mathbf{0}$,

$$\begin{aligned} a^{(k)}_{\boldsymbol{\nu}} &= -i \frac{1}{\boldsymbol{\omega}\cdot\boldsymbol{\nu}} \left(\sum_{\boldsymbol{\nu}_1+\boldsymbol{\nu}_2=\boldsymbol{\nu}} \left(f_{11,\boldsymbol{\nu}_1} a^{(k-1)}_{\boldsymbol{\nu}_2} + f_{12,\boldsymbol{\nu}_1} c^{(k-1)}_{\boldsymbol{\nu}_2} \right) + i \sum_{k_1+k_2=k} \mu^{(k_1)} a^{(k_2)}_{\boldsymbol{\nu}} \right), \\ c^{(k)}_{\boldsymbol{\nu}} &= -i \frac{1}{\boldsymbol{\omega}\cdot\boldsymbol{\nu} + 2\lambda_0} \left(\sum_{\boldsymbol{\nu}_1+\boldsymbol{\nu}_2=\boldsymbol{\nu}} \left(f_{21,\boldsymbol{\nu}_1} a^{(k-1)}_{\boldsymbol{\nu}_2} + f_{22,\boldsymbol{\nu}_1} c^{(k-1)}_{\boldsymbol{\nu}_2} \right) - i \sum_{k_1+k_2=k} \mu^{(k_1)} c^{(k_2)}_{\boldsymbol{\nu}} \right), \end{aligned} \tag{3.9}$$

for $k \geq 2$ and $\boldsymbol{\nu} \neq \mathbf{0}$, and

$$\begin{aligned} \mu^{(k)} &= i \left(\sum_{\boldsymbol{\nu}_1+\boldsymbol{\nu}_2=\mathbf{0}} \left(f_{11,\boldsymbol{\nu}_1} a^{(k-1)}_{\boldsymbol{\nu}_2} + f_{12,\boldsymbol{\nu}_1} c^{(k-1)}_{\boldsymbol{\nu}_2} \right) + i \sum_{k_1+k_2=k} \mu^{(k_1)} a^{(k_2)}_{\mathbf{0}} \right), \\ c^{(k)}_{\mathbf{0}} &= -\frac{i}{2\lambda_0} \left(\sum_{\boldsymbol{\nu}_1+\boldsymbol{\nu}_2=\mathbf{0}} \left(f_{21,\boldsymbol{\nu}_1} a^{(k-1)}_{\boldsymbol{\nu}_2} + f_{22,\boldsymbol{\nu}_1} c^{(k-1)}_{\boldsymbol{\nu}_2} \right) - i \sum_{k_1+k_2=k} \mu^{(k_1)} c^{(k_2)}_{\mathbf{0}} \right), \end{aligned} \tag{3.10}$$

for $k \geq 2$ and $\boldsymbol{\nu} = \mathbf{0}$.

Lemma 5 *Let $\mu^{(k)}$ and $c^{(k)}_{\mathbf{0}}$ be fixed for all $k \geq 1$ according to (3.8) and (3.10). Then there are formal power series (3.5) and (3.6) for β and μ, respectively, recursively determined from (3.7) to (3.10), which solve order by order equations (2.19). The constants $a^{(k)}_{\mathbf{0}}$ can be arbitrarily fixed.*

We omit the easy proof, which can be obtained also as a byproduct of the forthcoming analysis. Therefore, the formal solubility of the equations (2.19) requires that $\mu^{(k)}$ and $c^{(k)}_{\mathbf{0}}$ be fixed to all orders $k \geq 1$, while all coefficients $a^{(k)}_{\mathbf{0}}$ are left undetermined. We can fix the latter by requiring (2.16).

Lemma 6 *If we fix*

$$a_{\mathbf{0}}^{(1)} = 0 \tag{3.11}$$

for $k = 1$, and

$$a_{\mathbf{0}}^{(k)} = -\frac{1}{2} \sum_{k_1+k_2=k} \sum_{\boldsymbol{\nu}\in\mathbb{R}^d} \left(a_{\boldsymbol{\nu}}^{(k_1)} a_{\boldsymbol{\nu}}^{(k_2)*} - c_{\boldsymbol{\nu}}^{(k_1)} c_{\boldsymbol{\nu}}^{(k_2)*} \right), \tag{3.12}$$

for $k \geq 2$, then

$$H^{(k)} := a^{(k)} + a^{(k)*} + \sum_{k_1+k_2=k} \left(a^{(k_1)} a^{(k_2)*} - c^{(k_1)} c^{(k_2)*} \right) = 0, \tag{3.13}$$

for all $k \in \mathbb{N}$.

Proof. By Lemma 4 to all orders $k \geq 1$ the function $H^{(k)}$ is formally a constant, so that $H_{\boldsymbol{\nu}}^{(k)} = 0$ for all $k \geq 1$ and all $\boldsymbol{\nu} \neq \mathbf{0}$, while $H_{\mathbf{0}}^{(1)} = a_{\mathbf{0}}^{(1)} + a_{\mathbf{0}}^{(1)*}$ and

$$H_{\mathbf{0}}^{(k)} = a_{\mathbf{0}}^{(k)} + a_{\mathbf{0}}^{(k)*} + \sum_{k_1+k_2=k} \left(a^{(k_1)} a^{(k_2)*} - c^{(k_1)} c^{(k_2)*} \right)_{\mathbf{0}} \tag{3.14}$$

for $k \geq 2$ are constants. If we fix $a_{\mathbf{0}}^{(k)}$ recursively according to (3.11) and (3.12), then $H_{\mathbf{0}}^{(k)} = 0$, so that (3.13) follows. ■

The recursive equations (3.7) to (3.12) can be graphically represented in terms of linear trees as follows.

Call $u = (u_1, u_2) = (a, c)$. Set also $u_{1,\boldsymbol{\nu}}^{(k)} = a_{\boldsymbol{\nu}}^{(k)}$ and $u_{2,\boldsymbol{\nu}}^{(k)} = c_{\boldsymbol{\nu}}^{(k)}$, and represent $u_{j,\boldsymbol{\nu}}^{(k)}$ as a line carrying the labels $j \in \{1, 2\}$ and $\boldsymbol{\nu} \in \mathbb{Z}^d$ exiting from a bullet carrying the label k, with $k \in \mathbb{N}$. We call $k, j, \boldsymbol{\nu}$ the *order* label, the *component* label and the *momentum* label, respectively. We colour the bullet with white if $\boldsymbol{\nu} = \mathbf{0}$ and with grey if $\boldsymbol{\nu} \neq \mathbf{0}$; in the latter case, for $k = 1$ we draw the bullet as a black bullet instead of a grey one; cf. Figure 1. We call *graph elements* the graphs which are drawn this way. We represent also $\mu^{(k)}$ by a graph element, by using the same graph for $u_{\mathbf{0}}^{(k)}$ except that $j = 3$, i.e. we set $\mu^{(k)} = u_{3,\mathbf{0}}^{(k)}$.

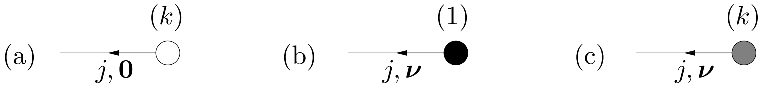

Figure 1: Graph elements representing (a) $u_{j,\mathbf{0}}^{(k)}$ for $j = 1, 2$ and $\mu^{(k)}$ for $j = 3$, (b) $u_{j,\boldsymbol{\nu}}^{(1)}$, $\boldsymbol{\nu} \neq \mathbf{0}$, and (c) $u_{j,\boldsymbol{\nu}}^{(k)}$. Only in (a) one can have $j = 3$, otherwise $j = 1, 2$. For $\boldsymbol{\nu} = 0$ the latter graph reduces to the first graph, while for $k = 1$ and $\boldsymbol{\nu} \neq \mathbf{0}$ it reduces to the second graph.

Then equations (3.9) can be represented as shown in Figure 2, provided we give some rules in order to associate with the graphs suitable numerical values.

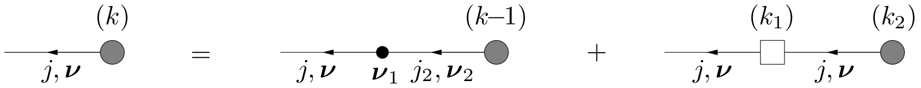

Figure 2: Graphical representation of (3.9), expressing the coefficient $u_{j,\boldsymbol{\nu}}^{(k)}$ for $k \geq 2$, $j = 1, 2$, and $\boldsymbol{\nu} \neq \mathbf{0}$ in terms of the coefficients $u_{j',\boldsymbol{\nu}'}^{(k')}$, with $k' < k$. In the first graph one has the constraint $\boldsymbol{\nu} = \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2$, while in second graph one has the constraint $k = k_1 + k_2$.

In the two graphs on the right hand side of Figure 2 there are two lines ℓ_1 and ℓ_2, with labels $(j_{\ell_1}, \boldsymbol{\nu}_{\ell_1}) = (j, \boldsymbol{\nu})$ and $(j_{\ell_2}, \boldsymbol{\nu}_{\ell_2}) = (j_2, \boldsymbol{\nu}_2)$, respectively. In the first graph we associate a *mode* label $\boldsymbol{\nu}_v = \boldsymbol{\nu}_1 \in \mathbb{Z}^d$ and a *node factor* $F_v = f_{jj_2,\boldsymbol{\nu}_1}$ with the black point v between the two lines. In the second graph we associate a *mode* label $\boldsymbol{\nu}_v = \boldsymbol{\nu}_1 = \mathbf{0}$, an *order* label $k_v = k_1$ and a *node factor* $F_v = (-1)^{j+1} i\mu^{(k_1)} \delta_{j,j_2}$ with the white square v between the two lines. In both graphs we have the constraint $\boldsymbol{\nu} = \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2$, which fixes $\boldsymbol{\nu}_2 = \boldsymbol{\nu}$ in the second graph. With the line ℓ_1 we associate a *propagator* g_{ℓ_1}, such that $g_{\ell_1} = 1/i\boldsymbol{\omega}\cdot\boldsymbol{\nu}$ if $j = 1$ and $g_{\ell_1} = 1/i(\boldsymbol{\omega}\cdot\boldsymbol{\nu} + 2\lambda_0)$ if $j = 2$, – note that in both graphs one has $\boldsymbol{\nu} \neq \mathbf{0}$. Finally the line ℓ_2 together with the grey bullet which it comes out from forms a graph element as shown in Figure 1(c), so that it represents $u^{(k_2)}_{j_2,\boldsymbol{\nu}_2}$, with $k_2 = k - 1$ in the first graph and $(j_2, \boldsymbol{\nu}_2) = (j, \boldsymbol{\nu})$ in the second one.

To obtain $u^{(k)}_{j,\boldsymbol{\nu}}$, with $\boldsymbol{\nu} \neq \mathbf{0}$, one has to sum over all labels the products of the propagator g_{ℓ_1} times the node factor F_v times the coefficient $u^{(k_2)}_{j_2,\boldsymbol{\nu}_2}$ represented by the graph element attached either to the black point or to the white square, with the constraint that the labels $j, \boldsymbol{\nu}, k$ are kept fixed. The quantity that one obtains this way is just the right hand side of equations (3.9). Of course $j = 1$ means that the corresponding graphs represent contributions to $a^{(k)}_{\boldsymbol{\nu}}$, and $j = 2$ means that they represent contributions to $c^{(k)}_{\boldsymbol{\nu}}$.

Analogously we can represent graphically (3.10) as in Figure 3. The difference with respect to Figure 2 is that now $\boldsymbol{\nu} = \mathbf{0}$, and $j \in \{2, 3\}$. For $j = 3$ we obtain a contribution to $\mu^{(k)}$, whereas for $j = 2$ we have a contribution to $c^{(k)}_{\mathbf{0}}$. The quantities to be associated with the black points, the white bullets, the white squares and the graph elements are the same as defined in the case of Figure 2. With the line ℓ_1 we associate a propagator g_{ℓ_1}, such that $g_{\ell_1} = i$ if $j = 3$ and $g_{\ell_1} = -i/2\lambda_0$ if $j = 2$.

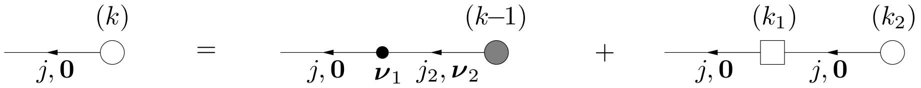

Figure 3: Graphical representation of (3.10), expressing the constants $\mu^{(k)}$ (if $j = 3$) and $c^{(k)}_{\mathbf{0}}$ (if $j = 2$) for $k \geq 2$ in terms of the coefficients $u^{(k')}_{j',\boldsymbol{\nu}'}$, with $k' < k$. In the first graph one has the constraint $\mathbf{0} = \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2$, while in second graph one has the constraint $k = k_1 + k_2$.

Finally, also $a^{(k)}_{\mathbf{0}}$ can be graphically represented from equation (3.12) in terms of the coefficients with lower order; cf. Figure 4. In such a case, in the graph on the right hand side, the line ℓ_1 which carries the labels $(j_{\ell_1}, \boldsymbol{\nu}_{\ell_1}) = (1, \mathbf{0})$ has propagator $g_{\ell_1} = 1/2$, and comes out from a white bullet v with two entering lines carrying labels $(j_{\ell_2}, \boldsymbol{\nu}_{\ell_2}) = (j_1, \boldsymbol{\nu}_1)$ and $(j_{\ell_3}, \boldsymbol{\nu}_{\ell_3}) = (j_2, \boldsymbol{\nu}_2)$, with the constraints $j_1 = j_2$ and $\boldsymbol{\nu}_1 + \boldsymbol{\nu}_2 = \mathbf{0}$. The node factor is $F_v = (-1)^j$.

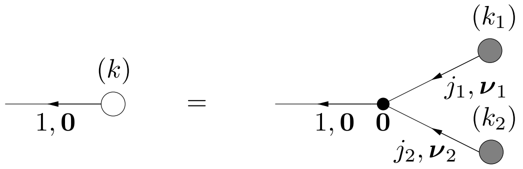

Figure 4: Graphical representation of (3.12), expressing the constant $a^{(k)}_{\mathbf{0}}$ (hence $j = 1$) for $k \geq 2$ in terms of the coefficients $u^{(k')}_{j',\boldsymbol{\nu}'}$, with $k' < k$. One has the constraints $\mathbf{0} = \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2$, $k = k_1 + k_2$, and $j_1 = j_2 \in \{1, 2\}$.

3.2 Linear trees

We can iterate the graphical construction given in Figures 2, 3 and 4 by developing further the graph elements on the right hand side according to same figures. At the end we obtain that $u^{(k)}_{j,\boldsymbol{\nu}}$, $\boldsymbol{\nu} \neq 0$, $\mu^{(k)}$ and $c^{(k)}_{\mathbf{0}}$ can all be expressed in terms of linear trees (or chains), which are constructed as follows.

A *tree* is a collection of *points* and *lines* connecting them, such that all lines are oriented toward a unique point, with the property that only one line enters such a point. The latter is called the *root* of the tree, and the line entering the root is called the *root line*. By construction any point different from the root has one and only one line coming out from it, called the *exiting line* of the point. A *linear tree* is a tree such that each point has only one line going into it, called the *entering line* of the point, except one which has no entering line at all. The latter is called the *endpoint* of the tree. All the points except the root and the endpoint are called the *nodes* of the tree.

Denote by $V(\theta)$ and $L(\theta)$ the set of nodes and the set of lines, respectively, in the tree θ. One has $|L(\theta)| = |V(\theta)| + 1$. Sometimes it can be convenient to denote by $P(\theta)$ the set of nodes plus the endpoint of θ.

We can number the lines and nodes as $\ell_1, \ldots, \ell_N$, and $v_1, \ldots, v_{N-1}$, with $N = |L(\theta)| \geq 1$, in such a way that ℓ_N connects the endpoint v_N to the node v_{N-1} (the *first node*), each line ℓ_k, $k = 2, \ldots, N-1$, connects the node v_k to the node v_{k-1}, and ℓ_1 connects the node v_1 (the *last node*) to the root.

A node v can be either a *black point* or a *white square*: in the latter case one must have $\boldsymbol{\nu}_v = \mathbf{0}$. The endpoint of the tree can be either a *white bullet* or a *black bullet*: the line ℓ coming out from the endpoint carries a momentum $\boldsymbol{\nu}_\ell = \mathbf{0}$ in the first case and a momentum $\boldsymbol{\nu}_\ell \neq \mathbf{0}$ in the second one. Examples of trees are depicted in Figure 5 and 6.

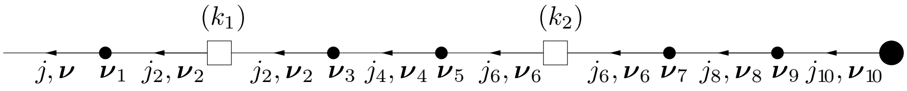

Figure 5: An example of tree of order k with 7 nodes and 8 lines, and with an endpoint which is a black bullet. One has the constraints $k = 6 + k_1 + k_2$, $\boldsymbol{\nu} = \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2$, $\boldsymbol{\nu}_2 = \boldsymbol{\nu}_3 + \boldsymbol{\nu}_4$, $\boldsymbol{\nu}_4 = \boldsymbol{\nu}_5 + \boldsymbol{\nu}_6$, $\boldsymbol{\nu}_6 = \boldsymbol{\nu}_7 + \boldsymbol{\nu}_8$, $\boldsymbol{\nu}_8 = \boldsymbol{\nu}_9 + \boldsymbol{\nu}_{10}$. The constraint that the lines connected to the white squares carry the same component and momentum labels has been taken into account explicitly. The order labels of the black points and of the black bullet are not shown, as they are necessarily 1. Also the mode label of the black bullet is not shown, as it is necessarily $\boldsymbol{\nu}_{10}$

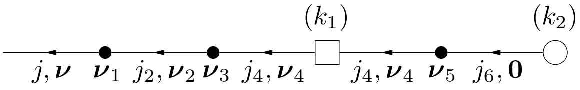

Figure 6: An example of tree of order k with 4 nodes and 5 lines, and with an endpoint which is a white bullet. One has the constraints $k = 3 + k_1 + k_2$, $\boldsymbol{\nu} = \boldsymbol{\nu}_1 + \boldsymbol{\nu}_2$, $\boldsymbol{\nu}_2 = \boldsymbol{\nu}_3 + \boldsymbol{\nu}_4$, $\boldsymbol{\nu}_4 = \boldsymbol{\nu}_5 + \boldsymbol{\nu}_6$, with $\boldsymbol{\nu}_6 = \mathbf{0}$. The constraint that the lines connected to the white square carry the same component and momentum labels has been taken into account explicitly. The order labels of the black points are not shown, as they are necessarily 1.

With each node v which is a black point we associate an *order* label $k_v = 1$ and a *mode* label $\boldsymbol{\nu}_v \in \mathbb{Z}^d$, and with each node which is a white square we associate an *order* label $k_v \in \mathbb{N}$ and a *mode* label $\boldsymbol{\nu}_v = \mathbf{0}$. If the endpoint v is a black bullet we associate with it an *order* label $k_v = 1$ and a *mode* label $\boldsymbol{\nu}_v \in \mathbb{Z}^d$, if it is a white bullet we associate with it an *order* label $k_v \in \mathbb{N}$ and a *mode* label $\boldsymbol{\nu}_v = \mathbf{0}$. With each line ℓ we associate a *component* label $j_\ell \in \{1, 2, 3\}$ and a *momentum* $\boldsymbol{\nu}_\ell \in \mathbb{Z}^d$. For each node we have the *conservation law* that the momentum of the exiting line equals the sum of the mode of the node plus the momentum of the entering line

As the tree is linear, for each node v there are only one line $\ell = \ell_v$ which comes out from it and only one line ℓ'_v which enters it. If $v = v_k$ this means that $\ell_v = \ell_{k-1}$ and $\ell'_v = \ell_k$. With these notations, the conservation law reads $\boldsymbol{\nu}_{\ell_v} = \boldsymbol{\nu}_v + \boldsymbol{\nu}_{\ell'_v}$.

Once all labels have been assigned, we associate with each node v the *node factor*

$$F_v := \begin{cases} f_{j_{\ell_v}, j_{\ell'_v}, \boldsymbol{\nu}_v} \left(1 - \delta_{j_{\ell_v},3}\right) + f_{1, j_{\ell'_v}, \boldsymbol{\nu}_v} \delta_{j_{\ell_v},3}, & v \text{ is a black point}, \\ (-1)^{j_{\ell_v}+1} i \mu^{(k_v)} \delta_{\boldsymbol{\nu}_v, \mathbf{0}} \delta_{j_{\ell_v}, j_{\ell'_v}}, & v \text{ is a white square}, \end{cases} \tag{3.15}$$

with the endpoint v the *endpoint factor*

$$F_v := \begin{cases} f_{j_{\ell_v} 1, \boldsymbol{\nu}_{\ell_v}}, & v \text{ is a black bullet}, \\ u^{(k_v)}_{j_{\ell_v}, \mathbf{0}}, & v \text{ is a white bullet}, \end{cases} \tag{3.16}$$

and with each line ℓ coming out from a node the *propagator*

$$g_\ell := \begin{cases} -i/\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell, & \boldsymbol{\nu}_\ell \neq \mathbf{0},\ j_\ell = 1, \\ -i/(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell + 2\lambda_0), & \boldsymbol{\nu}_\ell \neq \mathbf{0},\ j_\ell = 2, \\ -i/2\lambda_0, & \boldsymbol{\nu}_\ell = \mathbf{0},\ j_\ell = 2, \\ i, & \boldsymbol{\nu}_\ell = \mathbf{0},\ j_\ell = 3, \end{cases} \tag{3.17}$$

and with the line ℓ coming out from the endpoint the *propagator*

$$g_\ell := \begin{cases} -i/\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell, & \boldsymbol{\nu}_\ell \neq \mathbf{0},\ j_\ell = 1, \\ -i/(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell + 2\lambda_0), & \boldsymbol{\nu}_\ell \neq \mathbf{0},\ j_\ell = 2, \\ 1, & \boldsymbol{\nu}_\ell = \mathbf{0},\ j_\ell = 1,2, \end{cases} \tag{3.18}$$

The propagators (3.17) and (3.18) are equal as far as $\boldsymbol{\nu}_\ell \neq \mathbf{0}$, but they are different when $\boldsymbol{\nu}_\ell = \mathbf{0}$.

One has the further constraints that one can have $\boldsymbol{\nu}_\ell = \mathbf{0}$ in (3.17) only if ℓ is the root line, and $j_{\ell_v} = 3$ in (3.15) again only if ℓ_v is the root line. In particular the only lines which can have vanishing momentum are the root line and the line coming out from the endpoint, and the only line which can have component label $j = 3$ is the root line. Finally if $|P(\theta)| = 1$ then the endpoint of θ has to be a black bullet. Define $\Theta^0_{k,j,\boldsymbol{\nu}}$ the set of linear trees with labels $j, \boldsymbol{\nu}$ associated with the root line, and with $\sum_{v\in V(\theta)} k_v = k$.

Lemma 7 *Let $\boldsymbol{\omega}$ be a Bryuno vector. One has*

$$\begin{aligned} u^{(k)}_{j,\boldsymbol{\nu}} &= \sum_{\theta\in\Theta^0_{k,j,\boldsymbol{\nu}}} \mathrm{Val}(\theta), \qquad k \geq 1,\ \boldsymbol{\nu} \neq \mathbf{0},\ j = 1,2, \\ \mu^{(k)} &= \sum_{\theta\in\Theta^0_{k,3,\mathbf{0}}} \mathrm{Val}(\theta), \qquad c^{(k)}_{\mathbf{0}} = \sum_{\theta\in\Theta^0_{k,2,\mathbf{0}}} \mathrm{Val}(\theta), \qquad k \geq 1, \end{aligned} \tag{3.19}$$

where the tree value $\mathrm{Val}(\theta)$ *is given by*

$$\mathrm{Val}(\theta) = \Big(\prod_{\ell\in L(\theta)} g_\ell \Big) \Big(\prod_{v\in P(\theta)} F_v \Big), \tag{3.20}$$

with the propagators g_ℓ defined by (3.17) and (3.18), and the factors F_v defined by (3.15) and (3.16). One has $\mu^{(k)} \in \mathbb{R}$ for all $k \geq 1$.

Proof. The only non-trivial statement is that $\mu^{(k)}$ is real, – the other assertions can be easily derived from the discussion above (or can be proved by induction on k).

We prove that $\mu^{(k)} \in \mathbb{R}$ by induction. One has $\mu^{(1)} \in \mathbb{R}$ because $\mu^{(1)} = i f_{11,\mathbf{0}}$, and $f_{11,\mathbf{0}}$ is purely imaginary.

If $k \geq 2$, for each tree $\theta \in \Theta^0_{k,3,\mathbf{0}}$ we distinguish three cases: (a) the endnode of θ is a black bullet, (b) the endnode is a white bullet and the line coming out from it carries a label $j = 1$, and (c) the endnode is a white bullet and the line coming out from it carries a label $j = 2$.

We discuss first case (a). Given θ we consider the tree $\tau = \tau(\theta)$ obtained as follows. First, detach the root line from the last node and attach it to the endnode, and change the orientation of all lines; then the last node of θ becomes the endnode of τ (graphically it is transformed from a black point into a black bullet) and vice versa. Second, change the sign of all the mode labels.

Of course we can write $\mu^{(k)} = \sum_{\theta\in\Theta^0_{k,3,\mathbf{0}}} \mathrm{Val}(\tau(\theta))$. If we compare $\tau(\theta)$ with θ we see that the propagators are not changed, because the sum of all the mode labels is zero, i.e. $\sum_{v\in P(\theta)} \boldsymbol{\nu}_v = \mathbf{0}$. The node factors corresponding to white squares v are not changed (they remain $\pm i\mu^{(k_v)}$), while the node factors corresponding to black points are changed from $f_{j_{\ell_v} j'_{\ell_v},\boldsymbol{\nu}_v}$ into $f_{j'_{\ell_v} j_{\ell_v},-\boldsymbol{\nu}_v}$. The same happens to the endnode factor, which becomes $f_{1 j_{\ell_v},-\boldsymbol{\nu}_v}$. Recall that one has $f^*_{12,-\boldsymbol{\nu}} = f_{21,\boldsymbol{\nu}}$, $f^*_{21,-\boldsymbol{\nu}} = f_{12,\boldsymbol{\nu}}$, and $f^*_{11,-\boldsymbol{\nu}} = -f_{11,\boldsymbol{\nu}}$; moreover $g^*_\ell = -g_\ell$, as it follows from (3.17) and (3.18), and $\mu^{(k_v)*} = \mu^{(k_v)}$ by the inductive assumption.

Then if we compute $\mathrm{Val}^*(\tau)$ we obtain $\mathrm{Val}^*(\tau) = (-1)^{|L(\theta)|}(-1)^{|J(\theta)|}\mathrm{Val}(\theta)$, where $L(\theta)$ is the set of lines in θ, and $J(\theta)$ is the set of $v \in P(\theta)$ with $j_{\ell_v} = j_{\ell'_v}$ (we set $j'_{\ell_v} = 1$ if v is the endnode), hence including the white squares. It is immediate to realize that $|P(\theta) \setminus J(\theta)|$ is even, so that $|J(\theta)|$ has the same parity as $|P(\theta)|$. As $|P(\theta)| = |L(\theta)|$ this yields $\mathrm{Val}^*(\tau) = \mathrm{Val}(\theta)$.

In case (b) we can write $\mathrm{Val}(\theta) = \mathrm{Val}(\theta_1)\, a_{\mathbf{0}}^{(k_1)}$ for suitable θ_1 and k_1, with $a_{\mathbf{0}}^{(k_1)}$ real by (3.12). More precisely θ_1 is the tree of order $k - k_1$ obtained from θ by detaching the graph element representing $a_{\mathbf{0}}^{(k_1)}$ and replacing the first node with an endpoint. Then we can construct a tree $\tau_1 = \tau(\theta_1)$, and reason for θ_1 as done for θ in case (a). The same conclusions hold, in particular one finds $\mathrm{Val}^*(\tau_1) = \mathrm{Val}(\theta_1)$.

Finally in case (c) we can write $\mathrm{Val}(\theta) = \mathrm{Val}(\theta_1)\, c_{\mathbf{0}}^{(k_1)}$ for suitable θ_1 and k_1, and develop $c_{\mathbf{0}}^{(k_1)}$ in terms of trees (according with a procedure which will be extensively used in the following), and so on, until we reach a tree which belongs to case (a) or case (b), up to the fact it can contain lines ℓ with $\boldsymbol{\nu}_\ell = \mathbf{0}$ and $g_\ell = -i/2\lambda_0$; see (3.17) and (3.18). Therefore we can reason as in the previous cases (a) and (b).

By putting together all the cases, at the end we obtain $\mu^{(k)} = \mu^{(k)*}$. ■

Note that the set $\Theta^0_{k,1,\mathbf{0}}$ does not appear in (3.19). This is necessary as the map $\theta \to \mathrm{Val}(\theta)$ is not defined for $\theta \in \Theta^0_{k,1,\mathbf{0}}$; see (3.17). In fact, $a_{\mathbf{0}}^{(k)}$ cannot be represented as a sum of values of linear trees, but still we can write for $k \geq 2$ (and setting $a_{\mathbf{0}}^{(1)} = 0$)

$$a_{\mathbf{0}}^{(k)} = \frac{1}{2} \sum_{k_1+k_2=k} \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d} \sum_{j=1,2} (-1)^j \sideset{}{'}\sum_{\theta_1\in\Theta^0_{k_1,1,\boldsymbol{\nu}}} \mathrm{Val}(\theta_1) \sideset{}{'}\sum_{\theta_2\in\Theta^0_{k_2,1,\boldsymbol{\nu}}} \mathrm{Val}^*(\theta_2), \tag{3.21}$$

where $'$ means that we must interpret

$$\sideset{}{'}\sum_{\theta\in\Theta^0_{k,1,\mathbf{0}}} \mathrm{Val}(\theta) := a_{\mathbf{0}}^{(k)}. \tag{3.22}$$

Hence also $a_{\mathbf{0}}^{(k)}$ can be expressed in terms of linear trees.

3.3 Nonlinear trees

Each node represented by a white square can be further expanded in terms of trees as follows. First replace the white square v with a black point and attach to the latter a further graph element representing $\mu^{(k_v)}$, if k_v is the order label of v (cf. Figure 5), hence the graph element is expressed in terms of trees according to the first graph in Figure 3. With the new node v, represented by a black point, we associate a mode label $\boldsymbol{\nu}_v = \mathbf{0}$ and an order label $k_v = 0$.

In the same way also the endpoints which are drawn as white bullets can be expanded according to the second graph in Figure 3 if the exiting line carries a component label $j = 2$ and according to the graph in Figure 4 if the exiting line carries a component label $j = 1$.

Of course if we do this, then nonlinear trees appear. Nonlinear trees are partially ordered sets of points and lines connecting them, and not totally ordered sets, such as linear trees are. The advantage of this procedure, however, is that at the end, the trees have only endpoints with order 1 and all the node factors are quantities fixed (and not to be determined iteratively). The new trees can have also nodes with two entering lines. If we denote by p_v the *branching number* of the point v, that is the number of lines entering v, then $p_v = 1, 2$ if v is a node, while $p_v = 0$ is v is an endpoint.

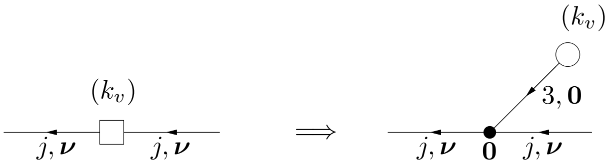

Figure 7: The quantity $\mu^{(k_v)}$ appearing in the node factor associated with v can be expressed according to (3.10). This can be interpreted graphically by replacing the white square as shown in the figure: the graph element entering the node v represent $\mu^{(k_v)}$, and it can be further developed in terms of trees according to Figure 3.

A node v with $p_v = 2$ has the following properties. Denote by ℓ_0 the exiting line of v, and by ℓ_1 and ℓ_2 the entering lines of v. Then either (i) $j_{\ell_0} = 1$, $\boldsymbol{\nu}_{\ell_0} = \mathbf{0}$ and $j_{\ell_1} = j_{\ell_2}$, $\boldsymbol{\nu}_{\ell_1} = \boldsymbol{\nu}_{\ell_2}$, or (ii) $j_{\ell_1} = 3$, $\boldsymbol{\nu}_{\ell_1} = \mathbf{0}$ and $j_{\ell_2} = j_{\ell_0}$, $\boldsymbol{\nu}_{\ell_2} = \boldsymbol{\nu}_{\ell_0} \neq \mathbf{0}$ or (iii) $j_{\ell_2} = 3$, $\boldsymbol{\nu}_{\ell_2} = \mathbf{0}$ and $j_{\ell_1} = j_{\ell_0}$, $\boldsymbol{\nu}_{\ell_1} = \boldsymbol{\nu}_{\ell_0} \neq \mathbf{0}$. Moreover in case (i) one has to take the complex conjugate of all propagators, node factors and endpoint factors of the subtree with root line ℓ_2. In all cases $k_v = 0$ and $\boldsymbol{\nu}_v = \mathbf{0}$, so that the conservation law is obeyed also in this case; cf. Figure 7. The corresponding *node factor* is

$$F_v := \begin{cases} (1/2)(-1)^j \delta_{\boldsymbol{\nu}_v,\mathbf{0}} \delta_v^\star, & p_v = 2, \text{ case (i)}, \\ (1/2)(-1)^{j+1} i \delta_{\boldsymbol{\nu}_v,\mathbf{0}} \delta_v^\star, & p_v = 2, \text{ cases (ii) and (iii)}, \end{cases} \tag{3.23}$$

where $\delta_v^\star$ recalls the constraints on the labels of the entering and exiting lines of v, which are detailed above and illustrated in Figure 8. The factor 1/2 in the second line of (3.23) aims to avoid overcountings of trees.

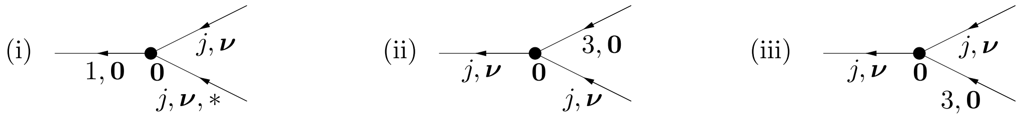

Figure 8: Nodes with branching number 2. The corresponding node factors are defined in (3.23). The entering lines are assumed to come out from other nodes or from endpoints, and the exiting line either enters another node or is the root line. The label $*$ on the lower entering line of the first graph means that one has to take the complex conjugate of the value of the subtree with that root line.

The nodes with branching number 1 can be only black points, because there are no more white squares. Hence (3.15) must be replaced with

$$F_v := f_{j_{\ell_v}, j_{\ell'_v}, \boldsymbol{\nu}_v} \left(1 - \delta_{j_{\ell_v},3}\right) + f_{1, j_{\ell'_v}, \boldsymbol{\nu}_v} \delta_{j_{\ell_v},3}, \qquad p_v = 1, \tag{3.24}$$

which represents the *node factor* of any node v with $p_v = 1$. The corresponding order label is $k_v = 1$, always. A line ℓ exiting from a node v can have also $\boldsymbol{\nu}_\ell = \mathbf{0}$ when $j_\ell = 2$.

All endpoints v have, by construction, $k_v = 1$, and are drawn as bullets coloured with black if $\boldsymbol{\nu}_{\ell_v} \neq \mathbf{0}$ and coloured with white if $\boldsymbol{\nu}_{\ell_v} = \mathbf{0}$, in the latter case one must have $j_{\ell_v} = 2$, as $a_{\mathbf{0}}^{(1)} = 0$; see (3.11). The *endpoint factor* of the endpoint v is given by

$$F_v := f_{j_{\ell_v} 1, \boldsymbol{\nu}_{\ell_v}}, \tag{3.25}$$

which replaces (3.16). If v is a white bullet then necessarily $j_{\ell_v} = 2$.

Finally, with the new rules, the propagator of any line ℓ is given by

$$g_\ell := \begin{cases} -i/\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell, & \boldsymbol{\nu}_\ell \neq \mathbf{0}, \ j_\ell = 1, \\ -i/(\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell + 2\lambda_0), & \boldsymbol{\nu}_\ell \neq \mathbf{0}, \ j_\ell = 2, \\ 1, & \boldsymbol{\nu}_\ell = \mathbf{0}, \ j_\ell = 1, \\ -i/2\lambda_0, & \boldsymbol{\nu}_\ell = \mathbf{0}, \ j_\ell = 2, \\ i, & \boldsymbol{\nu}_\ell = \mathbf{0}, \ j_\ell = 3, \end{cases} \tag{3.26}$$

which replaces both (3.17) and (3.18).

An example of tree with the new rules is given in Figure 9. The order labels are not shown, for simplicity (as the are identically 1, except for the nodes with branching number 2, which have order label 0).

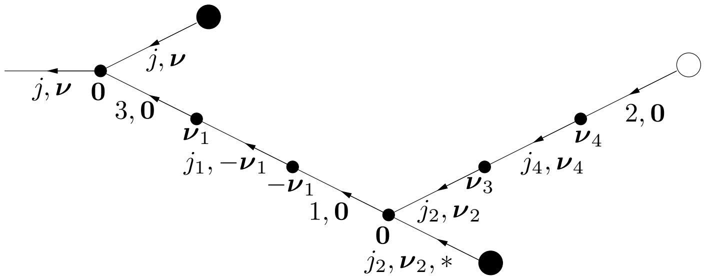

Figure 9: An example of tree of order $k = 7$ with 6 nodes, 3 endpoints and 9 lines. All endpoints and all nodes with branching number 1 have order 1, while the nodes with branching number 2 have order 0 (hence it is useless to write the orders explicitly). In principle the mode labels of the nodes with branching number 2 and the labels of the lines coming out from the endpoints which are white bullets could be omitted, as they are uniquely determined. The conservation law for the momenta has been taken into account explicitly, except for $\boldsymbol{\nu}_2 = \boldsymbol{\nu}_3 + \boldsymbol{\nu}_4$. One has $|V_1(\theta)| = 4$ and $|V_2(\theta)| = |L_0(\theta)| = 2$, so that $k_\theta = 7$.

We still denote by $V(\theta)$ and $L(\theta)$ the number of nodes and lines in θ. Define also $E(\theta)$ the number of endpoints of θ, and set $P(\theta) = V(\theta) \cup E(\theta)$. Furthermore call $V_p(\theta)$, $p = 1, 2$, the set of nodes $v \in V(\theta)$ with branching number $p_v = p$, and $L_0(\theta)$ the set of lines $\ell \in L(\theta)$ with $\boldsymbol{\nu}_\ell = \mathbf{0}$ which do not come out from endpoints. Then one has $|L_0(\theta)| = |V_2(\theta)|$.

We say that two trees are equivalent if they can be transformed into each other by continuously deforming the lines in such a way that the latter do not cross each other. Define $\Theta_{k,j,\boldsymbol{\nu}}$ the set of inequivalent trees with labels $j, \boldsymbol{\nu}$ associated with the root line, and with $k_\theta = \sum_{v \in P(\theta)} k_v = |P(\theta)| - |V_2(\theta)| = k$. The number of inequivalent trees in $\Theta_{k,j,\boldsymbol{\nu}}$ with fixed assignments of the mode labels $\{\boldsymbol{\nu}_v\}_{v \in P(\theta)}$ can be bounded by a constant to the power k: indeed a tree of order k has $P(\theta) \le 2k$, so that the number of unlabelled trees of order k can be bounded by the number of random walks with $4k$ steps, i.e. by 2^{4k}, and all labels except the mode labels assume a finite number of values.

We can summarise the considerations above into the following formal statement.

Lemma 8 *Let $\boldsymbol{\omega}$ be a Bryuno vector. One has*

$$
\begin{aligned}
u^{(k)}_{j,\boldsymbol{\nu}} &= \sum_{\theta \in \Theta_{k,j,\boldsymbol{\nu}}} \mathrm{Val}(\theta), \qquad && k \ge 1,\ j = 1, 2, \\
\mu^{(k)} &= \sum_{\theta \in \Theta_{k,3,\mathbf{0}}} \mathrm{Val}(\theta), \qquad && k \ge 1,
\end{aligned} \tag{3.27}
$$

with the tree value $\mathrm{Val}(\theta)$ *given by*

$$
\mathrm{Val}(\theta) = \Big(\prod_{\ell \in L(\theta)} g_\ell \Big) \Big(\prod_{v \in P(\theta)} F_v \Big). \tag{3.28}
$$

with the propagators g_ℓ given by (3.26), and the factors F_v given by (3.23), (3.24) and (3.25). One has $\mu^{(k)} \in i\mathbb{R}$ for all $k \ge 1$.

Even if (3.28) looks the same as (3.20), the meaning of the symbols is different.

The formal series (3.27) is well defined, as it is easy to check, but to order k, in general, we obtain for $\mathrm{Val}(\theta)$ bounds growing like $k!$ to some positive powers, so that summability is prevented if we try to estimate the series (3.27) by taking the absolute values of the tree values. To give a meaning to the formal series, we have to exploit some remarkable cancellations between the tree values. This can be showed by introducing a suitable resummation criterion of the series, which lead to a new series in which to any order k each tree value can be bounded proportionally to a constant to the power k. This will be done next.

4 Renormalised series

Consider a tree θ, and suppose that each line ℓ carries a further label $n_\ell \in \mathbb{Z}_+ \cup \{-1\}$, the *scale* label. We say that a connected set of lines and nodes $T \subset L(\theta)$ is a *cluster* on scale n_T if (i) all lines in T have scales no smaller than n_T, (ii) at least one line in T is on scale n_T, and (iii) it is maximal (which means that the lines connected to T but not belonging to it are on scales less than n_T). If T contains only one node (and no lines) we set $n_T = -1$, as in the case in which all lines in T are on scale -1.

If θ is a linear tree then all clusters have only one entering line, while in nonlinear trees clusters can have any number of entering lines. On the contrary a cluster, in both linear and nonlinear trees, can have only either zero or one exiting line. We call *external lines* of a cluster T the lines which are either entering or exiting lines for T.

We say that the cluster T is a *self-energy cluster* if (i) T has only one entering line and only one exiting line, (ii) the entering line carries the same momentum and component label as the exiting line, and (iii) no line along the path of lines connecting the external lines has vanishing momentum.

A self-energy cluster by construction can contain other self-energy clusters. We say that a self-energy cluster is a *renormalised self-energy cluster* if it does not contain any other self-energy clusters. We say that a tree θ is a *renormalised tree* if it does not contain any self-energy clusters. Given a self-energy cluster T, denote by $V(T)$, $E(T)$ and $L(T)$ the set of nodes, the set of endpoints and the set of lines, respectively, contained in T, and set $P(T) = V(T) \cup E(T)$. Call $V_p(T)$ the set of nodes $v \in V(T)$ with $p_v = p$, and $L_0(T)$ the set of lines $\ell \in L(T)$ with $\boldsymbol{\nu}_\ell = \mathbf{0}$ which do not come out from endpoints. Set $k_T = |P(T)| - |V_2(T)|$. An example of self-energy cluster is given in Figure 10.

Define the *self-energy value* $\mathcal{V}_T(\boldsymbol{\omega} \cdot \boldsymbol{\nu})$ as

$$\mathcal{V}_T(\boldsymbol{\omega} \cdot \boldsymbol{\nu}) = \varepsilon^{k_T} \Big(\prod_{\ell \in L(T)} g_\ell^{\mathcal{R}} \Big) \Big(\prod_{v \in P(T)} F_v \Big), \tag{4.1}$$

with the factors F_v defined as in Section 3 and the *renormalised propagators* $g_\ell^{\mathcal{R}}$ still to be defined.

The renormalised self-energy clusters can be of two kinds: those in which both external lines are attached to the same node, and those in which there is a nontrivial path of lines connecting the external lines. Those of the first type can be seen as obtained from the expansion of the white square representing a node of in a linear tree.

Consider a renormalised self-energy cluster T of the second kind. Call v_{in} and v_{out} the nodes which the entering line ℓ_{in} and the exiting line ℓ_{out} of T, respectively, are attached to. Then add a further node v_0 and a further line ℓ_0 and consider the set $\tilde{T}$, with $V(\tilde{T}) = V(T) \cup \{v_0\}$ and $L(\tilde{T}) = L(T) \cup \{\ell_0\}$, constructed as follows. Detach the line ℓ_{out} from v_{out} add attach it to the node v_0, and connect the node v_0 to the node v_{out} through the line ℓ_0 (oriented from v_{out} to v_0). Finally detach the line ℓ_{in} from v_{in} and reattach it to the node v_0 (so that $p_{v_0} = 2$). The last operation can be performed in two ways (ℓ_{in} can be above or below ℓ_0), hence it generates two renormalised self-energy clusters T' and T''. We call, shortly, *shift operation* the mechanism described above; cf. Figure 11.

Lemma 9 *For each renormalised self-energy cluster T of the second kind there is one and only one pair of renormalised self-energy clusters T' and T'' of the first kind which can be obtained from T through the shift operation.*

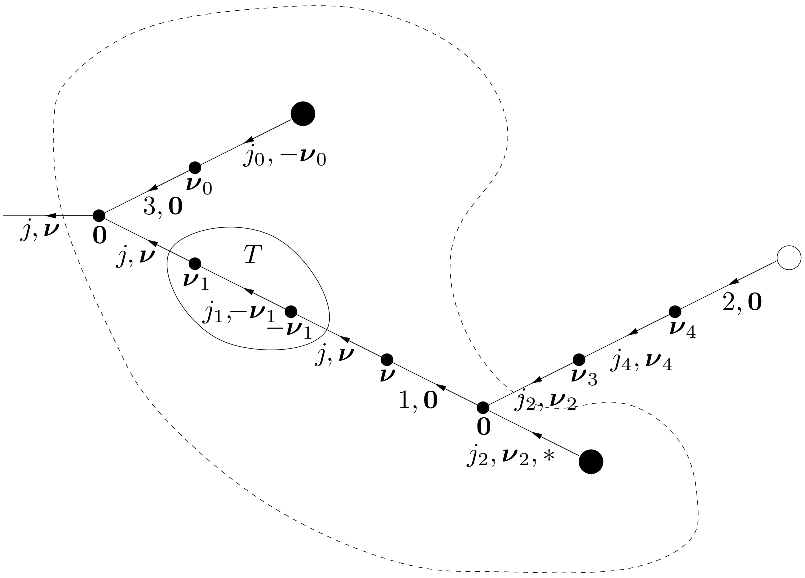

Figure 10: Example of self-energy cluster. Let T be the set of nodes and lines inside the solid line, i.e. the set consisting of the line ℓ with momentum $-\boldsymbol{\nu}_1$ and of the two nodes v_1 and v_2, with mode labels $\boldsymbol{\nu}_1$ and $-\boldsymbol{\nu}_1$, respectively, connected by such a line. Then T is a self-energy cluster if the scale of the line ℓ is strictly less than the scales of both the line ℓ_2 entering v_2 and the line ℓ_1 exiting v_1, i.e. if $n_\ell < \min\{n_{\ell_1}, n_{\ell_2}\}$. In such a case ℓ_1 and ℓ_2 become the external lines of T. The set of nodes and lines inside the dotted line cannot be a self-energy cluster, even if it is a cluster and $\boldsymbol{\nu} = \boldsymbol{\nu}_2$, because the path of lines between the external lines contains a line with vanishing momentum.

Proof. The proof is a simple application of the diagrammatic rules described in Section 3. ∎

This allows us to introduce a notion of equivalence between renormalised self-energy clusters. Then we can consider the renormalised self-energy clusters as triples of equivalent renormalised self-energy clusters $\{T, T', T''\}$.

Assume that $\boldsymbol{\omega}$ be a Bryuno vector. Define

$$C_0 = \sum_{n=0}^{\infty} 2^{n(d-1)} \alpha_n, \qquad \alpha_n = \alpha_n(\omega) := \inf_{|\boldsymbol{\nu}| \le 2^n} |\boldsymbol{\omega} \cdot \boldsymbol{\nu}|, \tag{4.2}$$

and set $\alpha_n = C_0 \gamma_n$. If the sum in (4.2) diverges, redefine C_0 by writing $2^{n(d-2)}$ instead of $2^{n(d-1)}$ (so that convergence is assured because $\alpha_n \le |\boldsymbol{\omega}| 2^{-n(d-1)}$, by Dirichlet's theorem [29]), and replace γ_n with $\gamma_n 2^{-n}$ in the following multiscale decomposition – see the definition of the compact support functions χ_n after (4.4), – and in the Diophantine conditions (4.14).

Note that $n' > n$ implies $\gamma_{n'} \le \gamma_n$, while $\gamma_{n'} < \gamma_n$ implies $n' > n$.

Set $\mathbb{Z}^d_* = \mathbb{Z}^d \setminus \{\mathbf{0}\}$, and define

$$n(\boldsymbol{\nu}) = \{n \in \mathbb{Z}_+ : 2^{n-1} < |\boldsymbol{\nu}| \le 2^n\} \tag{4.3}$$

for all $\boldsymbol{\nu} \in \mathbb{Z}^d_*$.

Let $\psi(x)$ a non-decreasing $C^\infty(\mathbb{R})$ function defined in $\mathbb{R}$, such that

$$\psi(x) = \begin{cases} 1\,, & \text{for } |x| \ge C_1\,, \\ 0\,, & \text{for } |x| \le C_1/2\,, \end{cases} \tag{4.4}$$

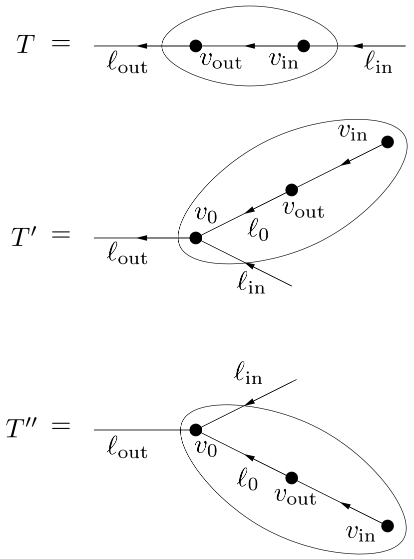

Figure 11: Examples of renormalised self-energy clusters belonging to the same equivalence class: T is a renormalised self-energy cluster of the second type, while T' and T'' are renormalised self-energy clusters of the first kind. The self-energy values of T' and T'' are equal to each other: in fact the trees containing such renormalised self-energy clusters can be obtained from each other by permuting the entering lines of v_0. The external lines ℓ_{in} and ℓ_{out} do not belong to the self-energy clusters, and have been drawn only to help visualising the structure of the self-energy clusters.

with the constant $C_1 \le C_0$ to be defined later. Set also $\chi(x) := 1 - \psi(x)$, and define, for all $n \in \mathbb{Z}_+$, $\chi_n(x) := \chi(\beta^{-1}\gamma_n^{-1}x)$ and $\psi_n(x) := \psi(\beta^{-1}\gamma_n^{-1}x)$, with $\beta = 1/4$.

Define

$$\Delta_0(x) = \left(\frac{1}{2}\left(\frac{1}{x^2} + \frac{1}{(x+2\lambda_0)^2}\right)\right)^{-1/2}, \tag{4.5}$$

and, setting $\mathcal{M}_1^{[0]}(x) := 0$ and $\mathcal{M}_2^{[0]}(x) := \lambda_0$, define for $n \ge 1$ and $j = 1, 2$

$$\begin{aligned} \mathcal{M}_j^{[\le n]}(x) &= \sum_{p=0}^{n} \mathcal{M}_j^{[p]}(x), \\ \mathcal{M}_j^{[n]}(x) &= \chi_0(\Delta_0(x)) \dots \chi_{n-1}(\Delta_0(x)) M_j^{[n]}(x), \\ M_j^{[n]}(x) &= \frac{i}{2} \sum_{k=1}^{\infty} \sum_{T \in \mathcal{S}_{k,j,n-1}} \mathcal{V}_T(x), \end{aligned} \tag{4.6}$$

where $\mathcal{S}_{k,j,n}$ is the set of all renormalised self-energy clusters T on scale n with $|P(T)| - |V_2(T)| = k$ and with component label j associated with both external lines. For $n = 0$ we interpret $\mathcal{M}_j^{[\le 0]}(x) = \mathcal{M}_j^{[0]}(x)$. One has

$$\min\{|x|, |x+2\lambda_0|\} \le \Delta_0(x) \le \sqrt{2}\min\{|x|, |x+2\lambda_0|\}. \tag{4.7}$$

Then the *renormalised propagator* is defined as $g_\ell^{\mathcal{R}} = g_\ell$ if $\boldsymbol{\nu}_\ell = \mathbf{0}$ and $g_\ell^{\mathcal{R}} = g_{j_\ell}^{[n_\ell]}(\boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell)$ if $\boldsymbol{\nu}_\ell \ne \mathbf{0}$, with

$$g_j^{[n]}(x) = -i \frac{\chi_0(\Delta_0(x)) \dots \chi_{n-1}(\Delta_0(x)) \psi_n(\Delta_0(x))}{x + 2\mathcal{M}_j^{[\le n]}(x)}, \tag{4.8}$$

so that we see that $g^{[n]}(x) \neq 0$ implies

$$\frac{1}{2}\beta\gamma_n C_1 \leq \Delta_0(x) \leq \beta\gamma_{n-1} C_1. \tag{4.9}$$

We associate a scale label n_ℓ also with lines with vanishing momentum, by setting $n_\ell = -1$.

Note that $\mathcal{M}_j^{[\leq n]}(x)$ is defined in terms of propagators on scales $n' < n$, hence in terms of $\mathcal{M}_{j'}^{[n']}(x')$, with $n' < n$: this means that (4.6) provides a recursive definition of $\mathcal{M}_j^{[\leq n]}(x)$, hence it makes sense. Note also that self-energy clusters on scale -1 (in particular those consisting of a single node) are not taken into account in (4.6); this will be motivated by Lemma 10 below.

Define the *tree value* $\mathrm{Val}(\theta)$ as

$$\mathrm{Val}(\theta) = \Big(\prod_{\ell \in L(\theta)} g_\ell^{\mathcal{R}} \Big) \Big(\prod_{v \in P(\theta)} F_v \Big). \tag{4.10}$$

Then, if $\Theta^{\mathcal{R}}_{j,k,\boldsymbol{\nu}}$ is the set of inequivalent renormalised trees with labels $j, \boldsymbol{\nu}$ associated with the root line and with $|P(\theta)| - |V_2(\theta)| = k$, set

$$u^{[k]}_{j,\boldsymbol{\nu}} = \sum_{\theta \in \Theta^{\mathcal{R}}_{j,k,\boldsymbol{\nu}}} \mathrm{Val}(\theta), \tag{4.11}$$

with $u^{[k]}_{3,\mathbf{0}} := \mu^{[k]}$, and define the function $\overline{u}(t) = (u_1(t), u_2(t))$ as

$$\overline{u}_j(t) = \sum_{k=1}^{\infty} \varepsilon^k u_j^{[k]}(t), \qquad u_j^{[k]}(t) = \sum_{\boldsymbol{\nu} \in \mathbb{Z}^d} \mathrm{e}^{i\boldsymbol{\nu}\cdot\boldsymbol{\omega} t} u^{[k]}_{j,\boldsymbol{\nu}}, \tag{4.12}$$

and the *counterterm* $\overline{\mu}$ as

$$\overline{\mu} = \sum_{k=1}^{\infty} \varepsilon^k \mu^{[k]}, \tag{4.13}$$

that we call the *renormalised series* for $u(t)$ and μ, respectively.

Lemma 10 *The self-energy clusters on scale -1 have values which cancel out exactly when summed together, hence there is no contributions arising from them to $\mathcal{M}_j^{[\leq n]}(x)$.*

Proof. The self-energy clusters on scale -1 are those represented in Figure 12. Hence they would contribute to $\mathcal{M}_j^{[\leq n]}(x)$ a value $f_{11,\mathbf{0}} + i\mu^{[1]}$ for $j = 1$ and $f_{22,\mathbf{0}} - i\mu^{[1]}$ for $j = 2$. By the very definition of $\mu^{[1]}$ one has $i\mu^{[1]} = -f_{11,\mathbf{0}}$, so that $f_{11,\mathbf{0}} + i\mu^{[1]} = 0$ for $j = 1$. For $j = 2$ one has $f_{22,\mathbf{0}} - i\mu^{[1]} = f_{22,\mathbf{0}} + f_{11,\mathbf{0}} = 0$, where we used that $f \in \mathfrak{m}$, so that $\mathrm{tr}\, f = 0$. ■

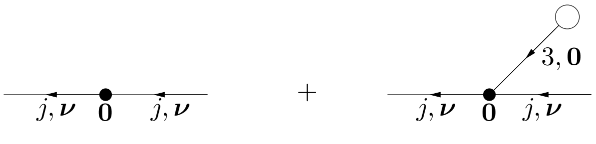

Figure 12: Self-energy clusters on scale -1 contributing to $\mathcal{M}_j^{[\leq n]}(\boldsymbol{\omega} \cdot \boldsymbol{\nu})$. The external lines do not belong to the self-energy clusters, and have been drawn only to help visualising the structure of the self-energy clusters.

For higher values of n, $\mathcal{M}_1^{[n]}(x)$ and $\mathcal{M}_2^{[n]}(x)$ are no longer equal to each other. However, we shall see that there is a deep symmetry yielding $\mathcal{M}_1^{[n]}(0) = -\mathcal{M}_2^{[n]}(-2\lambda_0)$ (cf. Lemma 15). Moreover, the cancellation mechanism which leads to Lemma 10 still works for any n, and implies partial cancellations, as Lemma 16 will show.

Assume the Diophantine conditions

$$\begin{aligned} |\boldsymbol{\omega}\cdot\boldsymbol{\nu}| &> C_1\gamma_{n(\boldsymbol{\nu})}, \\ |\boldsymbol{\omega}\cdot\boldsymbol{\nu}\pm 2\lambda_0| &> C_1\gamma_{n(\boldsymbol{\nu})}, \end{aligned} \tag{4.14}$$

for all $\boldsymbol{\nu}\in\mathbb{Z}^d_*$ and all $n\geq 0$. For $C_1\leq C_0$ the conditions in the first line are automatically satisfied by definition. The condition in the second line, called the *(first) Melnikov condition*, instead, have to be explicitly required with the constant C_1 – the same as in (4.4) – still to be fixed.

Let Λ_0 be the set in which λ_0 varies, and call Λ_0^* the subset of values $\lambda_0\in\Lambda_0$ for which the conditions (4.14) are satisfied. Of course Λ_0 has to be such that for $\lambda_0\in\Lambda_0$ one has $\lambda=\lambda_0+\mu\in[a,b]$, but for the time being we ignore such a constraint.

5 Convergence of the renormalised series

In this Section we assume that $\lambda_0\in\Lambda_0^*$. Hence the Diophantine conditions (4.14) are satisfied. We want to study the renormalised series for u and μ, with the aim of showing first that they converge, so that the functions $\overline{u}$ and $\overline{\mu}$ are well defined, second that $\overline{u}$ solves the equations (2.19) provided one fixes $\mu=\overline{\mu}$ and both $\overline{u}$ and $\overline{\mu}$ are analytic in ε, third that the relative measure of the set Λ_0^* is large. Finally we have to check that the last property implies that the set of values of λ in $[a,b]$ for which (2.10) is reducible also is of large measure; this will be done in Section 6.

We note since now that for any renormalised self-energy cluster T one has $|L(T)|=|P(T)|-1$, so that $|L(T)|-|L_0(T)|=k_T-1$. Moreover if $T\in\mathcal{S}_{k,j,n}$, with $n\geq 0$, then $k_T\geq 2$, because there must be at least one line on scale n.

In the following by saying that some property holds "for ε small enough" we mean that there exists a constant ε_0 (not necessarily the same in all the statements) such that (i) $\varepsilon_0 C_1^{-1}\ll 1$, and (ii) for $|\varepsilon|<\varepsilon_0$ that property is satisfied. Define also

$$|x+2\rho_0(x)| := \min\{|x|,|x+2\lambda_0|\}, \tag{5.1}$$

so that $\rho_0(x)$ is either 0 (if $x+\lambda_0\geq 0$) or λ_0 (if $x+\lambda_0<0$).

An important remark is that in the forthcoming Lemmata 11 and 12 the results hold unchanged if, in (4.9), we replace β with 2β in the upper bound and β with $\beta/2$ in the lower bound. Why this is important will be explained in the proof of Lemma 16.

Lemma 11 *Let θ be a renormalised tree. Call $N_n(\theta)$ the number of lines in $L(\theta)$ on scale n. One has*

$$N_n(\theta)\leq K2^{-n}M(\theta), \qquad M(\theta):=\sum_{v\in P(\theta)}|\boldsymbol{\nu}_v| \tag{5.2}$$

for a suitable constant K.

Proof. First note that if $N_n(\theta)\neq 0$ then, by (4.9), there exists a line $\ell\in L(\theta)$ such that $C_1\gamma_{n-1}>\beta C_1\gamma_{n-1}\geq|x+2\rho_0(x)|>C_1\gamma_{n(\boldsymbol{\nu}_\ell)}$, hence $n(\boldsymbol{\nu}_\ell)\geq n$, thence $M(\theta)\geq|\boldsymbol{\nu}_\ell|>2^{n-1}$.

Then we prove by induction that

$$N_n(\theta)\neq 0 \quad\Longrightarrow\quad N_n(\theta)\leq 22^{-n}M(\theta)-1. \tag{5.3}$$

If the root line of θ is not on scale n the bound (5.3) follows by induction. If the root line ℓ of θ is on scale n consider the lines $\ell_1,\ldots,\ell_p$ on scales $\geq n$ such that no line along the paths connecting any of them to the root line is on scale $\geq n$. If $p\geq 2$ again the bound follows inductively. If $p=1$ call θ_1 the subtree with root line ℓ_1, and call T the set of points and lines between ℓ_1 and ℓ (that is which precede ℓ but not ℓ_1). Denote by $P(T)$ the set of points in T, and define $M(T) := \sum_{v \in P(T)} |\boldsymbol{\nu}_v|$. Call $\boldsymbol{\nu}$ and $\boldsymbol{\nu}'$ the momenta associated with ℓ and ℓ_1, respectively, and set $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}$ and $x' = \boldsymbol{\omega} \cdot \boldsymbol{\nu}'$. One has $N_n(\theta) = 1 + N_n(\theta_1)$, and both $|x + 2\rho_0(x)|$ and $|x' + 2\rho_0(x')|$ are less than $\beta C_1 \gamma_{n-1}$, so that $|(x - x') + 2(\rho_0(x) - \rho_0(x'))| \leq 2\beta C_1 \gamma_{n-1} < C_1 \gamma_{n-1}$.

If there is (at least) a line ℓ' with $\boldsymbol{\nu}_{\ell'} = \mathbf{0}$ along the path of lines between the external lines ℓ and ℓ_1, then there exist two disjoint sets T_1 and T_2, with $P(T) = P(T_1) \cup P(T_2)$ and $L(T) = L(T_1) \cup L(T_2) \cup \{\ell'\}$, such that both $M(T_1)$ and $M(T_2)$ are greater than $|\boldsymbol{\nu}|$. Since ℓ is on scale n one has $|\boldsymbol{\nu}| > 2^{n-1}$, so that $M(T) \geq \max\{M(T_1), M(T_2)\} \geq 2^{n-1}$. If there is no line with zero momentum between the external lines, then $\boldsymbol{\nu} \neq \boldsymbol{\nu}'$, otherwise T would be a renormalised self-energy cluster. Therefore by the second Diophantine conditions (4.14), one obtains $n(\boldsymbol{\nu} - \boldsymbol{\nu}') \geq n$, so that $M(T) \geq 2^{n-1}$ also in such a case.

Hence, by the inductive hypothesis

$$N_n(\theta) \leq 1 + \big(22^{-n} M(\theta_1) - 1\big) \leq 1 - 22^{-n} M(T) + \big(22^{-n} M(\theta) - 1\big) \leq 22^{-n} M(\theta) - 1, \tag{5.4}$$

and the bound (5.3) follows. ■

Lemma 12 *Let T be a renormalised self-energy cluster. Call $N_n(T)$ the number of lines in $L(T)$ on scale n, with $n \leq n_T$. One has*

$$N_n(T) \leq K 2^{-n} M(T), \qquad M(T) := \sum_{v \in P(T)} |\boldsymbol{\nu}_v| > 2^{n_T - 1}, \tag{5.5}$$

with the same constant K as in (5.2).

Proof. We first prove the bound on $M(T)$. By construction T must contain at least a line ℓ on scale n_T, so that $|x_\ell + 2\rho_0(x_\ell)| \leq \beta C_1 \gamma_{n_T - 1}$, with $x_\ell = \boldsymbol{\omega} \cdot \boldsymbol{\nu}_\ell$. Write $\boldsymbol{\nu}_\ell = \boldsymbol{\nu}^0_{\ell_0} + \sigma_\ell \boldsymbol{\nu}$, where $\boldsymbol{\nu}$ is the momentum associated with the entering line of T and $\sigma_\ell = 0, 1$, and set $x = \boldsymbol{\omega} \cdot \boldsymbol{\nu}$ and $x^0_\ell = \boldsymbol{\omega} \cdot \boldsymbol{\nu}^0_\ell$. The entering line of T has scale strictly larger than n_T, so that $|x + 2\rho_0(x)| \leq \beta C_1 \gamma_{n_T - 1}$. If $M(T) \leq 2^{n_T - 1}$ then $|\boldsymbol{\nu}^0_\ell| \leq M(T) \leq 2^{n_T - 1}$, hence $n(\boldsymbol{\nu}^0_\ell) \leq n_T - 1$, so that $|x^0_\ell + 2\rho_0(x^0_\ell)| > C_1 \gamma_{n(\boldsymbol{\nu}^0_\ell)} \geq C_1 \gamma_{n_T - 1}$, by the Diophantine conditions (4.14). Then one has

$$\begin{aligned} C_1 \gamma_{n_T - 1} &> |x_\ell + 2\rho_0(x_\ell)| + \sigma_\ell |x + 2\rho_0(x)| \\ &\geq |x^0_\ell + 2(\rho_0(x_\ell) - \sigma_\ell \rho_0(x)| > C_1 \gamma_{n_T - 1}, \end{aligned} \tag{5.6}$$

which leads to a contradiction.

Next we pass to the bound on $N_n(T)$. Consider a subset G_0 of the lines of a tree θ between two lines ℓ_{out} and ℓ_{in} Set $G = G_0 \cup \{\ell_{\text{in}}\} \cup \{\ell_{\text{out}}\}$. Let $[n_{\text{in}}], [n_{\text{out}}]$ be the scales of the lines ℓ_{out} and ℓ_{in}, respectively, and suppose that $n_{\text{in}}, n_{\text{out}} \geq n$, while all lines in G_0 (if any) have scales $n' \leq n_T - 1$. Note that in general G_0 is not even a cluster unless $n_{\text{in}}, n_{\text{out}} \geq n_T$. Then we can prove that if $N_n(G_0) \neq 0$ then $N_n(G_0) \leq 22^{-n} \sum_{v \in P(G_0)} |\boldsymbol{\nu}_v| - 1$, where $P(G_0)$ is the set of points preceding ℓ_{out} and following ℓ_{in}.

If G_0 has no lines then the mode $\boldsymbol{\nu}_0$ of the (only) node between ℓ_{out} and ℓ_{in} is such that $|\boldsymbol{\nu}_0| \geq 2^{n-1}$, by the second Diophantine conditions (4.14), and the statement is true. Hence we proceed inductively on the number of lines in G_0. If no line of G_0 on the path $\mathcal{P}(G)$ connecting the external lines of G has scale n then the lines in G_0 on scale n (if any) belong to trees with root on $\mathcal{P}(G)$, and the statement follows from the bound (5.3) for trees given in the proof of Lemma 11. If there is a line $\ell \in \mathcal{P}(G)$ on scale n, then call G_1 and G_2 the disjoint subsets of G such that $G_1 \cup G_2 \cup \{\ell\} = G$. Then $G_1 \cup \{\ell\}$ and $G_2 \cup \{\ell\}$ have the same structure of G itself, but each has less lines. Hence, again the inductive assumption yields the result.

Therefore, as a particular case, by choosing $G_0 = T$, with $T \in \mathcal{S}_{k,j,n_T - 1}$, the bound for $N_n(G_0)$ implies the bound on $N_n(T)$ we are looking for. ■

Lemma 13 *Assume that the propagators $g_j^{[p]}(x)$ can be uniformly bounded for all $0 \le p \le n-1$ as*

$$\left| g_j^{[p]}(x) \right| \le K_1 C_1^{-1} \gamma_p^{-1}, \tag{5.7}$$

for some positive constant K_1. Then one has

$$|\mathcal{V}_T(\boldsymbol{\omega}\cdot\boldsymbol{\nu})| \le |\varepsilon|^{k_T} D_1^{k_T} C_1^{-(k_T-1)} \gamma_{m_0}^{-k_T} \mathrm{e}^{-\kappa_0 M(T)/2}, \tag{5.8}$$

for a suitable constant D_1. If also the derivatives of the propagators are bounded uniformly as

$$\left| \partial_x g_j^{[p]}(x) \right| \le K_2 C_1^{-2} \gamma_p^{-3}, \tag{5.9}$$

for some positive constant K_2, one has also

$$\left| \frac{\mathrm{d}}{\mathrm{d}x}\, \mathcal{V}_T(x)|_{x=\boldsymbol{\omega}\cdot\boldsymbol{\nu}} \right| \le |\varepsilon|^{k_T} D_2^{k_T} C_1^{-k_T} \gamma_{m_0}^{-k_T-2} \mathrm{e}^{-\kappa_0 M(T)/2}, \tag{5.10}$$

for a suitable constant D_2.

Proof. For any renormalised self-energy cluster T consider the corresponding self-energy value (4.1). The product of factors F_v can be bounded as

$$\prod_{v\in P(T)} F_v \le F_0^{k_T} \prod_{v\in V_1(T)\cup E(T)} \mathrm{e}^{-\kappa_0|\boldsymbol{\nu}_v|}, \tag{5.11}$$

while the product of propagators can be bounded, for any $m_0 \in \mathbb{N}$, as

$$\prod_{v\in L(T)} g_\ell^{\mathcal{R}} \le C_1^{-(k_T-1)} \gamma_{m_0}^{-k_T} \exp\left(K \sum_{n=m_0+1}^{\infty} \frac{1}{2^n} \log\frac{1}{\gamma_n} M(T) \right), \tag{5.12}$$

where the first bound (5.5) of Lemma 12 has been used. If we choose m_0 such that

$$K \sum_{n=m_0+1}^{\infty} \frac{1}{2^n} \log\frac{1}{\gamma_n} \le \frac{\kappa_0}{12}, \tag{5.13}$$

then we obtain (5.8). Such m_0 exists because $\boldsymbol{\omega}$ is a Bryuno vector; cf. (3.2).

Call $\mathcal{P}(T)$ the path of lines $\ell \in L(T)$ which are between the external lines of T. Then the derivative of $\mathcal{V}_T(x)$ can be written as

$$\partial_x \mathcal{V}_T(x) = \varepsilon^{k_T} \sum_{\ell\in\mathcal{P}(T)} \partial_x g_\ell^{\mathcal{R}} \Big(\prod_{\ell'\in L(T)\setminus\ell} g_{\ell'}^{\mathcal{R}} \Big) \Big(\prod_{v\in P(T)} F_v \Big), \tag{5.14}$$

so that, by reasoning as in the previous case, using the bounds (5.9) and choosing again m_0 as in (5.13), we obtain (5.10). ■

Lemma 14 *$\mathcal{M}_j^{[\le n]}|\mathbb{R}$ is real for all $n \ge 0$ and $j = 1, 2$.*

Proof. The proof is by induction on n. For $n = 0$ the assertion is trivially satisfied. Then assume that it holds for all $n' < n$.

Let T be a renormalised self-energy cluster contributing to $\mathcal{M}_j^{[n]}(x)$ through (4.6). Denote by $v_{\rm in}$ and $v_{\rm out}$ the nodes in $V(T)$ which the entering line $\ell_{\rm in}$ and the exiting line $\ell_{\rm out}$ of T are attached to, respectively. Call $\mathcal{P}(T)$ the set of lines and nodes between the external lines of T.

Together with T consider also the renormalised self-energy cluster T' obtained as follows. Detach the line $\ell_{\rm in}$ from $v_{\rm in}$ and attach it to the node $v_{\rm out}$, and detach the line $\ell_{\rm out}$ from $v_{\rm out}$ and attach it to the node $v_{\rm in}$. Consistently, orient all lines along the path $\mathcal{P}(T)$ between the external lines of T in the opposite direction, i.e. from $v_{\rm out}$ to $v_{\rm in}$. Finally change the mode labels of all nodes along $\mathcal{P}(T)$, i.e. of all nodes $v \in V(\mathcal{P}(T))$, if $V(\mathcal{P}(T))$ denotes the set of nodes along $\mathcal{P}(T)$. The latter operation is possible because of the following reason. Each line entering a node $v \in V(\mathcal{P}(T))$ has zero momentum: indeed for each node v with branching number $p_v = 2$ one of the three lines connected with v must have zero momentum (cf. Figure 8), and by definition of self-energy cluster such a line cannot lay on $\mathcal{P}(T)$. Hence $\sum_{v\in V(\mathcal{P}(T))} \boldsymbol{\nu}_v = \mathbf{0}$. Note also that each line entering a node $v \in V(\mathcal{P}(T))$ is the root line of a tree contributing to $\mu^{[k_v]}$, for some k_v. Along the path $\mathcal{P}(T)$ the propagators have not changed because of the operation above (cf. the analogous discussion in the proof of Lemma 7), by the inductive hypothesis. The node factors are changed as described in the proof of Lemma 7. As a consequence, when we sum over all possible renormalised self-energy clusters, we find $\mathcal{M}_j^{[\le n]}(x) = \mathcal{M}_j^{[\le n]*}(x)$, which proves the assertion. ■

Lemma 15 *Assume that the propagators $g_j^{[p]}(x)$ and their derivatives can be uniformly bounded for all $0 \le p \le n-1$ as in (5.7) and (5.9), for some constants K_1 and K_2. Then one has $\mathcal{M}_1^{[n]}(0) = -\mathcal{M}_2^{[n]}(-2\lambda_0)$ for all $n \ge 1$.*

Proof. Write $M_1^{[n]}(x)$ according to (4.6). For any T contributing to $M_1^{[n]}(x)$ we construct a renormalised self-energy cluster T' contributing to $M_2^{[n]}(x)$ as follows. Call $\mathcal{P}(T)$ the path of lines and nodes between the external lines of T, and denote with $V(\mathcal{P}(T))$ and $L(\mathcal{P}(T))$ the set of nodes and the set of lines, respectively, along $\mathcal{P}(T)$. If $L(\mathcal{P}(T)) = \emptyset$ the assertion trivially follows from (3.23). Hence in the following assume $L(\mathcal{P}(T)) \neq \emptyset$.

By definition of self-energy cluster all $\ell \in L(\mathcal{P}(T))$ have momentum different from zero, while all lines connected to a node $v \in V(\mathcal{P}(T))$ have zero momentum (cf. Figure 8). Hence $\sum_{v\in V(\mathcal{P}(T))} \boldsymbol{\nu}_v = \mathbf{0}$. The nodes $v \in V(\mathcal{P}(T))$ are totally ordered, so that we can number them $v_0, v_1, \ldots, v_N$, if $N = |L(\mathcal{P}(T))|$. The self-energy cluster T' is obtained through three steps: (i) first, we associate to each node v_i, $i = 0, \ldots, N$, the mode label and the node factor of the node v_{N-i} in T, – in other words we revert the order of the nodes, – (ii) next, we write all node factors $f_{11,\boldsymbol{\nu}_v}$ and $f_{22,\boldsymbol{\nu}_v}$ as $f_{11,\boldsymbol{\nu}_v} = -f_{22,\boldsymbol{\nu}_v}$ and $f_{22,\boldsymbol{\nu}_v} = -f_{11,\boldsymbol{\nu}_v}$, – by using that $\operatorname{tr} f = 0$, – (iii) finally we change consistently the component labels j_ℓ of the lines $\ell \in L(\mathcal{P}(T))$, – which means that each label $j = 1$ is changed into $j = 2$ and vice versa.

If $\ell \in L(T)$ is the line connecting, say, v_k to v_{k-1} for some $k = 1, \ldots, N$, we still call ℓ the line in $L(T')$ which connects v_{N-k+1} to v_{N-k}. For each line $\ell \in L(T)$ we can write its momentum as $\boldsymbol{\nu}_\ell = \boldsymbol{\nu}_\ell^0 + \boldsymbol{\nu}$, where $\boldsymbol{\nu}_\ell^0$ is the sum of the mode labels of the nodes $v \in V(\mathcal{P}(T))$ preceding v in T and $\boldsymbol{\nu}$ is the momentum of the line entering T. Then the corresponding line ℓ in $L(T')$ will have momentum $-\boldsymbol{\nu}_\ell^0 + \boldsymbol{\nu}$. Therefore each propagator $g_j^{[n]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell + \boldsymbol{\omega}\cdot\boldsymbol{\nu})$ in T is changed into $g_{3-j}^{[n]}(-\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell + \boldsymbol{\omega}\cdot\boldsymbol{\nu})$ in T'.

From the very definition of the propagators one sees immediately that, by setting $x_\ell^0 = \boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell^0$ and $x = \boldsymbol{\omega}\cdot\boldsymbol{\nu}$, one has

$$\begin{aligned} g_1^{[n_\ell]}(x_\ell^0) &= g_2^{[n_\ell]}(x_\ell^0 - 2\lambda_0) = -g_2^{[n_\ell]}(-x_\ell^0 - 2\lambda_0), \\ g_2^{[n_\ell]}(x_\ell^0) &= g_1^{[n_\ell]}(x_\ell^0 + 2\lambda_0) = -g_1^{[n_\ell]}(-x_\ell^0 - 2\lambda_0). \end{aligned} \tag{5.15}$$

Now compute $\operatorname{Val}(T)$ for $x = 0$ and $\operatorname{Val}(T')$ for $x = -2\lambda_0$. Of course the node factors do not depend on the momenta, so that

$$\prod_{v\in V(\mathcal{P}(T))} F_v = (-1)^{|J(\mathcal{P}(T))|} \prod_{v\in V(\mathcal{P}(T'))} F_v, \tag{5.16}$$

where $J(\mathcal{P}(T))$ is the set of nodes $v \in V(\mathcal{P}(T))$ with $j_{\ell_v} = j_{\ell'_v}$. It is immediate to realise that $|J(\mathcal{P}(T))|$ has the same parity of $|V(\mathcal{P}(T))|$, – see the proof of Lemma 7 for a similar argument.

By using (5.15) we obtain also

$$\prod_{\ell\in L(\mathcal{P}(T))} g_\ell^{\mathcal{R}}\Big|_{x=0} = (-1)^{|L(\mathcal{P}(T))|} \prod_{\ell\in L(\mathcal{P}(T'))} g_\ell^{\mathcal{R}}\Big|_{x=-2\lambda_0}. \tag{5.17}$$

Finally we have

$$\Big(\prod_{v\in P(T)\setminus V(\mathcal{P}(T))} F_v\Big)\Big(\prod_{\ell\in L(T)\setminus L(\mathcal{P}(T))} g_\ell^{\mathcal{R}}\Big) = \Big(\prod_{v\in P(T')\setminus V(\mathcal{P}(T'))} F_v\Big)\Big(\prod_{\ell\in L(T')\setminus L(\mathcal{P}(T'))} g_\ell^{\mathcal{R}}\Big) \tag{5.18}$$

for all $x \in \mathbb{R}$, so that, by using that $(-1)^{|J(\mathcal{P}(T))|}(-1)^{|L(\mathcal{P}(T))|} = (-1)^{|V(\mathcal{P}(T))|+|L(\mathcal{P}(T))|} = -1$, we find $\mathcal{V}_T(0) = -\mathcal{V}_{T'}(-2\lambda_0)$. Then the assertion follows. ∎

Lemma 16 *Assume that the propagators $g_j^{[p]}(x)$ and their derivatives can be uniformly bounded for all $0 \le p \le n-1$ as in (5.7) and (5.9), for some constants K_1 and K_2. Then for ε small enough and $n \ge 1$ one has $\mathcal{M}_1^{[n]}(0) = \mathcal{M}_2^{[n]}(-2\lambda_0) = 0$, and*

$$\begin{aligned} \left|\mathcal{M}_1^{[n]}(x)\right| &\le B_1 \mathrm{e}^{-\kappa_1 2^n} |\varepsilon|^2 \min\{C_1^{-1}, |x| C_1^{-2}\}, \\ \left|\mathcal{M}_2^{[n]}(x)\right| &\le B_1 \mathrm{e}^{-\kappa_1 2^n} |\varepsilon|^2 \min\{C_1^{-1}, |x+2\lambda_0| C_1^{-2}\}, \end{aligned} \tag{5.19}$$

for suitable n-independent constants B_1 and κ_1.

Proof. By using the definitions in (4.6) and noting that all sums are controlled, we see that the bound (5.8) implies the bound $|\mathcal{M}_j^{[n]}(x)| \le B_1 \mathrm{e}^{-\kappa_1 2^n} |\varepsilon|^2 C_1^{-1}$ for both $j=1$ and $j=2$.

The proof of the other bounds is more subtle. Let us start with the case $j=1$.

Let T be a renormalised self-energy cluster. First consider the case $C_1 \gamma_{n(M(T))} \le 4|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$, where $\boldsymbol{\nu}$ is the momentum associated with the entering line of T. In that case one can extract from the last product in (5.11) a factor $\mathrm{e}^{-\kappa_0 M(T)/4} \le \mathrm{e}^{-\kappa_0 2^{n(M(T))}/8}$. Since $\boldsymbol{\omega}$ is a Bryuno number then $a_n := 2^{-n}\log 1/\alpha_n$ tends to zero as $n\to\infty$, hence for $\boldsymbol{\omega}\cdot\boldsymbol{\nu}$ small enough one has $\mathrm{e}^{-\kappa_0 2^{n(M(T))}/8} \le (C_0\gamma_{n(M(T))})^{\kappa_0/8a_{n(M(T))}} \le C_0\gamma_{n(M(T))} \le 4C_0 C_1^{-1}|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$, which implies the bound (5.19).

Then we consider the case $C_1 n(M(T)) > 4|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$. In that case for any line $\ell \in L(T)$ and for any $n < n_\ell$, by the Diophantine conditions (4.14), one has $|x_\ell^0 + 2\rho_0(x_\ell^0)| > C_1\gamma_{n(\boldsymbol{\nu}_\ell^0)}$, where $x_\ell = \boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell$ and $x_\ell^0 = \boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell^0$, with $\boldsymbol{\nu}_\ell = \boldsymbol{\nu}_\ell^0 + \sigma_\ell^0$, $\sigma_\ell = 0,1$. Since $|\boldsymbol{\nu}_\ell^0| \le M(T)$, then $C_1\gamma_{n(\boldsymbol{\nu}_\ell^0)} \ge C_1\gamma_{n(M(T))} > 4|\boldsymbol{\omega}\cdot\boldsymbol{\nu}|$, which yields

$$2\left|x_\ell^0 + 2\rho_0(x_\ell^0)\right| \ge |x_\ell + 2\rho_0(x_\ell)| \ge \frac{1}{2}\left|x_\ell^0 + 2\rho_0(x_\ell^0)\right|. \tag{5.20}$$

Such a property is important for the following reason. It can happen, by the properties of the compact support functions, that a line ℓ is such that $g_{j_\ell}^{[n_\ell]}(x_\ell^0) \ne 0$, whereas $g_{j_\ell}^{[n_\ell]}(x_\ell^0 + x) = 0$. On the other hand in order to exploit the cancellations describe below we have to consider also renormalised self-energy clusters containing lines of this kind. Then (5.20) says that in such cases, even if the bounds (4.7) are not satisfied, one still has bounds of the same form with the only difference that β is replaced with 2β in the upper bound and with $\beta/2$ in the lower bound. But this is enough to apply both Lemma 11 and Lemma 12.

For any renormalised self-energy cluster we consider the renormalised self-energy clusters which belong to the same equivalence class. Assume that T is that of the second kind and that T' and T'' are those of the first type. The corresponding self-energy values differ because of two facts: (i) the value of T' and T'' has an extra overall factor $-1/2$, deriving from the product of the propagator i times the node factor $i/2$, and (ii) for all lines along the path between the external lines of T the propagators depend also on $\boldsymbol{\omega}\cdot\boldsymbol{\nu}$. The latter statement means that if ℓ is one of such lines then $g_\ell^{\mathcal{R}} = g_{j_\ell}^{[n_\ell]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell^0 + \boldsymbol{\omega}\cdot\boldsymbol{\nu})$ for $\ell \in L(T)$, while $g_\ell^{\mathcal{R}} = g_{j_\ell}^{[n_\ell]}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}_\ell^0)$ for $\ell \in L(T')$ and $\ell \in L(T'')$. Finally, the two renormalised self-energy clusters T' and T'' have the same values.

Therefore $\mathcal{V}_{T'}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) = \mathcal{V}_{T'}(0)$ and $\mathcal{V}_{T''}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) = \mathcal{V}_{T''}(0) = \mathcal{V}_{T'}(0)$, hence

$$\begin{aligned} \mathcal{V}_{T'}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) + \mathcal{V}_{T''}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) + \mathcal{V}_{T}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) &= \mathcal{V}_{T'}(0) + \mathcal{V}_{T''}(0) + \mathcal{V}_{T}(0) \\ &+ \left(\mathcal{V}_{T}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) - \mathcal{V}_{T}(0)\right) = \mathcal{V}_{T}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) - \mathcal{V}_{T}(0), \end{aligned} \tag{5.21}$$

as $\mathcal{V}_{T'}(0) = \mathcal{V}_{T''}(0) = -\mathcal{V}_{T}(0)/2$. By writing

$$\mathcal{V}_{T}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) - \mathcal{V}_{T}(0) = \boldsymbol{\omega}\cdot\boldsymbol{\nu} \int_0^1 \mathrm{d}s \, \frac{\mathrm{d}}{\mathrm{d}x} \, \mathcal{V}_{T}(x)|_{x=s\boldsymbol{\omega}\cdot\boldsymbol{\nu}} \tag{5.22}$$

and using (5.10) the bound (5.19) follows once more.

The case $j = 2$ follows easily from Lemma 15. Indeed for any renormalised self-energy cluster T we can write

$$\mathcal{V}_{T}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) = \mathcal{V}_{T}(-2\lambda_0) + \left(\mathcal{V}_{T}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) - \mathcal{V}_{T}(-2\lambda_0)\right), \tag{5.23}$$

where

$$\mathcal{V}_{T}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}) - \mathcal{V}_{T}(-2\lambda_0) = (\boldsymbol{\omega}\cdot\boldsymbol{\nu} + 2\lambda_0) \int_0^1 \mathrm{d}s \, \frac{\mathrm{d}}{\mathrm{d}x} \, \mathcal{V}_{T}(x)|_{x=-2\lambda_0+s(\boldsymbol{\omega}\cdot\boldsymbol{\nu}+2\lambda_0)} \tag{5.24}$$

can be bounded by using (5.10), while

$$\frac{i}{2}\sum_{k=1}^{\infty} \sum_{T\in\mathcal{S}_{k,2,n-1}} \mathcal{V}_{T}(-2\lambda_0) = \mathcal{M}_2^{[n]}(-2\lambda_0) = -\mathcal{M}_1^{[n]}(0) = 0, \tag{5.25}$$

so that the assertion is proved also in such a case. ∎

Lemma 17 *Assume that the propagators* $g_j^{[p]}(x)$ *are differentiable, and that, together with their derivatives, they can be uniformly bounded for all* $0 \le p \le n-1$ *as in (5.7) and (5.9), for suitable constants* K_1 *and* K_2*. Then for* ε *small enough* $\mathcal{M}_j^{[\le n]}(x)$ *is differentiable in* x*, and one has*

$$\begin{aligned} &\left|\mathcal{M}_j^{[\le n]}(x') - \mathcal{M}_j^{[\le n]}(x) - \partial_x \mathcal{M}_j^{[\le n]}(x)\,(x'-x)\right| = o(\varepsilon^2 C_1^{-2}|x'-x|), \\ &\left|\partial_x \mathcal{M}_j^{[\le n]}(x)\right| \le B_2|\varepsilon|^2 C_1^{-2}, \end{aligned} \tag{5.26}$$

for a suitable constant B_2*.*

Proof. By writing $\mathcal{M}_j^{[\le n]}(x)$ according to (4.6), one finds immediately that the function is differentiable if the propagators are differentiable, and that the derivative satisfies the bound in (5.19). The factor ε^2 is due to the fact that a self-energy cluster T depending explicitly on x has at least $k_T = 2$. ∎

Lemma 18 *Assume that the propagators* $g_j^{[p]}(x)$ *and their derivatives can be uniformly bounded for all* $0 \le p \le n-1$ *as in (5.7) and (5.9), for some constants* K_1 *and* K_2*. Then for* ε *small enough one has*

$$\left|x + 2\mathcal{M}_j^{[\le n]}(x)\right| \ge \frac{1}{2}\Delta_0(x) \tag{5.27}$$

as far as $g_j^{[n]}(x) \ne 0$*.*

Proof. By Lemma 16 one has $\mathcal{M}_1^{[\leq n]}(0) = 0$ and $\mathcal{M}_2^{[\leq n]}(-2\lambda_0) = \lambda_0$. Set $j(x) = 1$ when $\rho_0(x) = 0$ and $j(x) = 2$ when $\rho_0(x) = \lambda_0$, so that one can write

$$\begin{aligned} x + 2\mathcal{M}_{j(x)}^{[\leq n]}(x) &= x + 2\mathcal{M}_{j(x)}^{[\leq n]}(-2\rho_0(x)) + \left(2\mathcal{M}_{j(x)}^{[\leq n]}(x) - 2\mathcal{M}_{j(x)}^{[\leq n]}(-2\rho_0(x))\right) \\ &= x + 2\rho_0(x) + 2\left(\mathcal{M}_{j(x)}^{[\leq n]}(x) - \mathcal{M}_{j(x)}^{[\leq n]}(-2\rho_0)\right), \end{aligned} \tag{5.28}$$

where $|\mathcal{M}_{j(x)}^{[\leq n]}(x) - \mathcal{M}_{j(x)}^{[\leq n]}(-2\rho_0)| \leq \text{const.}|\varepsilon|^2 C_1^{-2}|x + 2\rho_0(x)|$, by Lemma 17. Then by (4.7) one has

$$\left|x + 2\mathcal{M}_{j(x)}^{[\leq n]}(x)\right| \geq \left(1 - \text{const.}|\varepsilon|^2 C_1^{-2}\right)|x + 2\rho_0(x)| \geq \frac{1 - \text{const.}|\varepsilon|^2 C_1^{-2}}{\sqrt{2}}\Delta_0(x). \tag{5.29}$$

Since $|x + 2\mathcal{M}_{3-j(x)}^{[\leq n]}(x)| \geq (1 - \text{const.}|\varepsilon|^2 C_1^{-2})|x + 2\mathcal{M}_{j(x)}^{[\leq n]}(x)|$, the bound follows. ■

Lemma 19 *The propagators $g_j^{[n]}(x)$ satisfy the bounds (5.7) and (5.9) for all $n \geq 0$.*

Proof. The proof can be performed by induction. For $n = 1$ the bounds (5.7) and (5.9) are trivially satisfied, as $\mathcal{M}_1^{[0]}(x) = 0$ and $\mathcal{M}_2^{[0]}(x) = \lambda_0$, because of the Diophantine conditions (4.14).

The difference for $n > 1$ is that now the propagators depend also on the functions $\mathcal{M}_j^{[p]}(x)$, $p < n$, appearing in the denominators and the compact support functions. Then assume (5.7) and (5.9) for all $p < n$. Then one has $|g_j^{[n]}(x)| \leq \text{const.}\psi_n(\Delta_0(x))/\Delta_0(x) \leq \text{const.}C_1^{-1}\gamma_n^{-1}$, by Lemma 18. Moreover

$$\begin{aligned} \partial_x g_j^{[n]}(x) &= -i\sum_{p=0}^{n-1}\chi_0(\Delta_0(x))\ldots\partial\chi_p(\Delta_0(x))\ldots\chi_{n-1}(\Delta_0(x))\psi_n(\Delta_0(x))\frac{\partial_x\Delta_0(x)}{x + 2\mathcal{M}_j^{[\leq n]}(x)} \\ &- i\chi_0(\Delta_0(x))\ldots\chi_{n-1}(\Delta_0(x))\partial\psi_n(\Delta_0(x))\frac{\partial_x\Delta_0(x)}{x + 2\mathcal{M}_j^{[\leq n]}(x)} \\ &+ i\chi_0(\Delta_0(x))\ldots\chi_{n-1}(\Delta_0(x))\psi_n(\Delta^{[n]}(x))\frac{1 + 2\partial_x\mathcal{M}_j^{[\leq n]}(x)}{(x + 2\mathcal{M}_j^{[\leq n]}(x))^2}, \end{aligned} \tag{5.30}$$

where ∂ denotes derivative with respect to the argument.

One checks immediately that for all $p \geq 0$

$$\partial\chi_p(x) \leq \text{const.}C_1^{-1}\gamma_p^{-1}, \qquad \partial\psi_p(x) \leq \text{const.}C_1^{-1}\gamma_p^{-1}, \qquad \partial_x\Delta_0(x) \leq \text{const.}, \tag{5.31}$$

so that the derivative $\partial_x\mathcal{M}_j^{[\leq n]}(x)$ can be bounded through (5.26), because of the inductive hypothesis.

Hence, by using once more (4.9) and Lemma 18 to bound the denominators, we obtain from (5.30)

$$\left|\partial_x g_j^{[n']}(x)\right| \leq \text{const.}C_1^{-2}\left(\sum_{p=0}^{n-1}\gamma_p^{-1}\gamma_n^{-1} + \gamma_n^{-1}\gamma_n^{-1} + \gamma_n^{-2}\right) \leq \text{const.}C_1^{-2}\gamma_n^{-3}, \tag{5.32}$$

which proves the assertion. ■

Lemma 20 *Let $\lambda \in \Lambda_0^*$. There exists $\varepsilon_0 > 0$ such that for $|\varepsilon| < \varepsilon_0$ the coefficients $u_{j,\boldsymbol{\nu}}^{[k]}$, $j = 1, 2$, and $\mu^{[k]}$ are bounded by*

$$\left|u_{j,\boldsymbol{\nu}}^{[k]}\right| \leq B\,\mathrm{e}^{-\kappa|\boldsymbol{\nu}|}|\varepsilon|^k, \qquad \left|\mu^{[k]}\right| \leq B\,|\varepsilon|^k, \tag{5.33}$$

for suitable k-independent constants B and κ. One can take $\varepsilon_0 = O(C_1\gamma_{m_0})$, with m_0 depending on κ_0.

Proof. For any tree $\theta \in \Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu}}$ the value $\mathrm{Val}(\theta)$ can be bounded by using the bounds (5.11) for the factors F_v and the bounds (5.7), proved in Lemma 19, for the propagators. Summation over the Fourier labels can be performed by using an exponential decay factor $\mathrm{e}^{-\kappa_0 M(T)/4}$ which can be extracted from (5.11). Summation over the other labels and over the number of unlabelled trees can be easily bounded as a constant to the power k. ■

Lemma 21 *The function $\overline{u}(t)$ solves (3.3) for all $\boldsymbol{\nu} \neq \mathbf{0}$, provided $\mu = \overline{\mu}$.*

Proof. We write

$$\begin{aligned} \overline{u}_j(t) &= \overline{u}_{j,\mathbf{0}} + \sum_{\boldsymbol{\nu}\in\mathbb{Z}^d} e^{i\boldsymbol{\omega}\cdot\boldsymbol{\nu}} \overline{u}_{j,\boldsymbol{\nu}}, \qquad \overline{u}_{j,\boldsymbol{\nu}} = \sum_{n=0}^{\infty} \overline{u}_{j,\boldsymbol{\nu},n}, \\ \overline{u}_{j,\boldsymbol{\nu},n} &= \sum_{k=1}^{\infty} \varepsilon^k \sum_{\theta\in\Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},n}} \mathrm{Val}(\theta), \end{aligned} \tag{5.34}$$

where $\Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},n}$ is the set of trees in $\Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu}}$ with root line on scale n.

An important property of the compact support functions is that

$$1 = \sum_{n=0}^{\infty} \Psi_n(x), \qquad \Psi_n(x) := \chi_0(\Delta_0(x)) \dots \chi_{n-1}(\Delta_0(x))\psi_n(\Delta_0(x)), \tag{5.35}$$

where the summand for $n = 0$ is meant as $\psi_0(\Delta_0(x))$. More generally one has for all $s \geq 1$

$$1 = \sum_{n=p}^{\infty} \Psi_{p,n}(x), \qquad \Psi_{p,n}(x) := \chi_p(\Delta_0(x)) \dots \chi_{n-1}(\Delta_0(x))\psi_n(\Delta_0(x)), \tag{5.36}$$

where again the summand for $n = p$ is meant as $\psi_p(\Delta_0(x))$.

We can rewrite the equation (3.3) as

$$u_{j,\boldsymbol{\nu}} = g_j(x)\Phi_{j,\boldsymbol{\nu}}(u), \qquad \Phi_j = \varepsilon f_{j1} + i\mu + \varepsilon f_{j1}u_1 + \varepsilon f_{j2}u_2 + (-1)^{j+1} i\mu u_j, \tag{5.37}$$

where $x = \boldsymbol{\omega}\cdot\boldsymbol{\nu}$, $g_j(x) = -i(x + 2\mathcal{M}^{[0]}_j(x))^{-1}$, with $\mathcal{M}^{[0]}_1(x) = 0$ and $\mathcal{M}^{[0]}_2(x) = \lambda_0$.

By using (5.35) we can write

$$\begin{aligned} g_j(x)\Phi_{j,\boldsymbol{\nu}}(\overline{u}) &= g_j(x) \sum_{n=0}^{\infty} \Psi_n(x)\Phi_{j,\boldsymbol{\nu}}(\overline{u}) \\ &= g_j(x) \sum_{n=0}^{\infty} \Psi_n(x) \left(g^{[n]}_j(x)\right)^{-1} \left(g^{[n]}_j(x)\Phi_{j,\boldsymbol{\nu}}(\overline{u})\right), \end{aligned} \tag{5.38}$$

where $\Psi_n(x)(g^{[n]}_j(x))^{-1} = i(x + 2\mathcal{M}^{[\leq n]}_j(x))$, and

$$g^{[n]}_j(x)\Phi_{j,\boldsymbol{\nu}}(\overline{u}) = \sum_{k=1}^{\infty} \varepsilon^k \sum_{\theta\in\overline{\Theta}^{\mathcal{R}}_{k,j,\boldsymbol{\nu},n}} \mathrm{Val}(\theta), \tag{5.39}$$

where $\overline{\Theta}^{\mathcal{R}}_{k,j,\boldsymbol{\nu},n}$ differs from $\Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},n}$ as it contains also trees which can have one renormalised self-energy cluster T with exiting line given by the root line of θ. In such a case if p is the line of the entering line of T, then $p \geq 0$ and the scale n_T of T is such that $n_T + 1 \leq \min\{n, p\}$, by definition of cluster.

Then we have

$$\begin{aligned}
&\sum_{n=0}^{\infty} \Psi_n(x) \left(g_j^{[n]}(x)\right)^{-1} \left(g_j^{[n]}(x)\Phi_{j,\boldsymbol{\nu}}(u)\right) \\
&\qquad = i \sum_{n=0}^{\infty} \left(x + 2\mathcal{M}_j^{[\le n]}(x)\right) \sum_{k=1}^{\infty} \varepsilon^k \sum_{\theta \in \Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},n}} \mathrm{Val}(\theta) \\
&\qquad -2i \sum_{n=1}^{\infty} \Psi_n(x) \sum_{p=n}^{\infty} \sum_{s=1}^{n} M_j^{[s]}(x) \sum_{k=1}^{\infty} \varepsilon^k \sum_{\theta \in \Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},p}} \mathrm{Val}(\theta) \\
&\qquad -2i \sum_{n=2}^{\infty} \Psi_n(x) \sum_{p=1}^{n-1} \sum_{s=1}^{p} M_j^{[s]}(x) \sum_{k=1}^{\infty} \varepsilon^k \sum_{\theta \in \Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},p}} \mathrm{Val}(\theta),
\end{aligned} \tag{5.40}$$

and we can use the definitions (5.34) to write

$$\sum_{k=1}^{\infty} \varepsilon^k \sum_{\theta \in \Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},n}} \mathrm{Val}(\theta) = \overline{u}_{j,\boldsymbol{\nu},n}, \qquad \sum_{k=1}^{\infty} \varepsilon^k \sum_{\theta \in \Theta^{\mathcal{R}}_{k,j,\boldsymbol{\nu},p}} \mathrm{Val}(\theta) = \overline{u}_{j,\boldsymbol{\nu},p}, \tag{5.41}$$

in the second line and, respectively, in the third and fourth lines.

Then the sum of the third and fourth lines in (5.40) gives

$$\begin{aligned}
&-2i \left(\sum_{n=1}^{\infty} \Psi_n(x) \sum_{p=n}^{\infty} \sum_{s=1}^{n} M_j^{[s]}(x)\, \overline{u}_{j,\boldsymbol{\nu},p} + \sum_{n=2}^{\infty} \Psi_n(x) \sum_{p=1}^{n-1} \sum_{s=1}^{p} M_j^{[s]}(x)\, \overline{u}_{j,\boldsymbol{\nu},p} \right) \\
&\qquad = -2i \sum_{p=1}^{\infty} \overline{u}_{j,\boldsymbol{\nu},p} \left(\sum_{s=1}^{p} \sum_{n=s}^{p} M_j^{[s]}(x)\Psi_n(x) + \sum_{s=1}^{p} \sum_{n=p+1}^{\infty} M_j^{[s]}(x)\Psi_n(x) \right) \\
&\qquad = -2i \sum_{p=1}^{\infty} \overline{u}_{j,\boldsymbol{\nu},p} \sum_{s=1}^{p} M_j^{[s]}(x) \sum_{n=s}^{\infty} \Psi_n(x).
\end{aligned} \tag{5.42}$$

If we define

$$\Xi_n(x) := \chi_0(\Delta_0(x)) \ldots \chi_{n-1}(\Delta_0(x))\chi_n(\Delta_0(x)), \tag{5.43}$$

then in (5.42) we can write

$$\sum_{n=s}^{\infty} \Psi_n(x) = \Xi_{s-1}(x) \sum_{q=s}^{\infty} \Psi_{s,q} = \Xi_{s-1}(x), \tag{5.44}$$

where the property (5.36) has been used. Hence in (5.42) we have

$$\sum_{s=1}^{p} M_j^{[s]}(x) \sum_{n=s}^{\infty} \Psi_n(x) = \sum_{s=1}^{p} M_j^{[s]}(x)\Xi_{s-1}(x) = \mathcal{M}_j^{[\le p]}(x) - \mathcal{M}_j^{[0]}(x), \tag{5.45}$$

where the factor $\mathcal{M}_j^{[0]}(x)$ has been subtracted as the sum over s starts from $s = 1$ and not from $s = 0$.

If we insert (5.40) into (5.38), by taking into account (5.42) and (5.45), we obtain

$$\begin{aligned}
g_j(x)\Phi_{j,\boldsymbol{\nu}}(\overline{u}) &= g_j(x) \left(i \sum_{n=0}^{\infty} \left(x + 2\mathcal{M}_j^{[\le n]}(x)\right) - 2i \sum_{n=1}^{\infty} \left(\mathcal{M}_j^{[\le n]}(x) - \mathcal{M}_j^{[0]}(x)\right) \right) \overline{u}_{j,\boldsymbol{\nu},n} \\
&= g_j(x) \sum_{n=0}^{\infty} i \left(x + 2\mathcal{M}_j^{[0]}(x)\right) \overline{u}_{j,\boldsymbol{\nu},n} = \sum_{n=0}^{\infty} \overline{u}_{j,\boldsymbol{\nu},n} = \overline{u}_{j,\boldsymbol{\nu}},
\end{aligned} \tag{5.46}$$

so that (5.37) is satisfied for $u = \overline{u}$. ■

Lemma 22 *The function $\overline{u}(t)$ solves the system of differential equations (2.19) for all $t \in \mathbb{R}$, provided $\mu = \overline{\mu}$. Moreover the function H defined in (2.20) satisfies $H(\overline{u}(t)) = 0$ for all $t \in \mathbb{R}$. Both $\overline{u}(t)$ and $\overline{\mu}$ are analytic in ε.*

Proof. Because of Lemma 21, to show that $\overline{u}(t)$ is a solution it is enough to prove that $\overline{u}$ solves (3.4), that is $0 = \Phi_{j,\mathbf{0}}(\overline{u})$, with $\Phi_j(u)$ defined in (5.33). But this is obvious by construction.

The claim on $H(\overline{u}(t))$ follows if the solution is in $\mathfrak{M}$, so that (2.16) is satisfied. But again this follows from the construction of the solution.

Finally the statement about analyticity easily follows from the construction of the renormalised series. The series defining $\overline{u}(t)$ in (4.12) and $\overline{\mu}$ in (4.11) can be viewed as power series in ε with coefficients depending on ε. The coefficients depend on ε through the propagators, and in fact are analytic in ε (for ε small enough). Hence the series themselves define functions which are analytic in ε. ■

A result analogous to Lemma 22, in particular analyticity of the conjugation and of the counterterm, was proved by using renormalisation group techniques in [25, 19]. In the case of the Schrödinger equation it was also obtained in [3], with techniques similar to those used in this paper; cf. also [10, 6, 18, 2] for related issues. See also [11], Chapter 9, for resummed series defining analytic functions, in the context of maximal KAM tori.

6 Reducibility on a large measure set

So far we have proved that, as far as $\lambda_0 \in \Lambda_0^*$, the function $\overline{u}(t)$ solves (2.19). We still have to prove that the relative measure of Λ_0^* with respect to Λ_0 is large, and we have to see what this means for the parameter $\lambda \in [a, b]$. We shall find that the subset $\Lambda^* \subset [a, b]$ of values λ for which the construction envisaged in the previous sections works is of large measure; this will complete the proof of Theorem 1.

As a consequence of Lemma 19, we have that (5.19), (5.26), and (5.27) hold for all $n \geq 0$.

For each $\boldsymbol{\nu} \in \mathbb{Z}_*^d$ we have to exclude all values $\lambda_0 \in \Lambda_0^*$ such that $|\boldsymbol{\omega} \cdot \boldsymbol{\nu} + 2\lambda_0| \leq C_1 \gamma_{n(\boldsymbol{\nu})}$. If we consider λ_0 as a function of an auxiliary parameter $t \in [-1, 1]$, we can write

$$\boldsymbol{\omega} \cdot \boldsymbol{\nu} + 2\lambda_0(t) = t C_1 \gamma_{n(\boldsymbol{\nu})}, \qquad t \in [-1, 1], \tag{6.1}$$

so that

$$\frac{\mathrm{d}\lambda_0}{\mathrm{d}t} = \frac{C_1}{2} \gamma_{n(\boldsymbol{\nu})}. \tag{6.2}$$

Then for each $\boldsymbol{\nu} \in \mathbb{Z}_*^d$ we have to exclude all values of t in $[-1, 1]$.

Lemma 23 *There exists $\varepsilon_0 > 0$ and $\sigma > 0$ such that for all $|\varepsilon| < \varepsilon_0$ the Lebesgue measure of the set $\Lambda_0 \setminus \Lambda_0^*$ is bounded proportionally to $|\varepsilon|^\sigma$.*

Proof. The set Λ_0^* is obtained by imposing the Diophantine conditions (4.14). Then we can bound

$$\operatorname{meas}(\Lambda_0^*) = \int_{\Lambda_0^*} \mathrm{d}\lambda_0 = \sum_{\boldsymbol{\nu} \in \mathbb{Z}_*^d} \int_{-1}^{1} \mathrm{d}t \, \frac{\mathrm{d}\lambda_0}{\mathrm{d}t} = \sum_{\boldsymbol{\nu} \in \mathbb{Z}_*^d} C_1 \gamma_{n(\boldsymbol{\nu})}, \tag{6.3}$$

where we can write

$$\sum_{\boldsymbol{\nu} \in \mathbb{Z}_*^d} \gamma_{n(\boldsymbol{\nu})} = \sum_{n=0}^{\infty} \sum_{2^{n-1} < |\boldsymbol{\nu}| \leq 2^n} \gamma_n \leq \text{const.} \sum_{n=0}^{\infty} 2^{n(d-1)} \gamma_n \leq \text{const.} \tag{6.4}$$

By inserting (6.4) into (6.3) we obtain $\operatorname{meas}(\Lambda_0^*) \leq \text{const.} C_1$, hence Λ_0^* is a set of measure proportional to C_1.

By Lemma 20 we can take $C_1 = |\varepsilon|^\sigma$, with $0 < \sigma < 1$. Hence the measure of the discarded set can be bounded proportionally to $|\varepsilon|^\sigma$. ■

In the following write $\overline{u}(t, \lambda_0)$, $\mu(\lambda_0)$ and $\mathcal{M}_j^{[\leq n]}(x, \lambda_0)$ to make explicit the dependence of $\overline{u}(t)$, μ and $\mathcal{M}_j^{[\leq n]}(x)$ on λ_0.

Lemma 24 *Assume $\lambda_0, \lambda_0' \in \Lambda_0^*$. One has*

$$\begin{aligned} &\left|\mathcal{M}_j^{[\leq n]}(x, \lambda_0') - \mathcal{M}_j^{[\leq n]}(x, \lambda_0) - \partial_{\lambda_0}\mathcal{M}_j^{[\leq n]}(x, \lambda_0)\right| = o(\varepsilon^2 C_1^{-2} |\lambda_0' - \lambda_0|), \\ &\left|\partial_{\lambda_0}\left(\mathcal{M}_j^{[\leq n]}(x, \lambda_0) - \mathcal{M}_j^{[0]}(x, \lambda_0)\right)\right| \leq A_1 |\varepsilon|^2 C_1^{-2}, \end{aligned} \tag{6.5}$$

for a suitable constant A_1. In particular $\mathcal{M}_j^{[\leq n]}(x, \lambda_0)$ can be extended in all $\overline{\Lambda}_0$ to a differentiable function (Whitney extension), whose derivative satisfies the bound in (6.5).

Proof. The proof is by induction. $\mathcal{M}_j^{[\leq n]}(x, \lambda_0)$ can be written according to (4.6). Hence it depends recursively on $\mathcal{M}_{j'}^{[\leq n']}(x)$, $n' < n$, through the propagators, and one can express $\partial_{\lambda_0}\mathcal{M}_j^{[\leq n]}(x, \lambda_0)$ as sum of derivatives of self-energy values,

$$\partial_{\lambda_0}\mathcal{V}_T(x) = \varepsilon^{k_T} \sum_{\ell \in \mathcal{P}(T)} \partial_{\lambda_0} g_\ell^{\mathcal{R}} \Big(\prod_{\ell' \in L(T) \setminus \ell} g_{\ell'}^{\mathcal{R}} \Big) \Big(\prod_{v \in P(T)} F_v \Big). \tag{6.6}$$

For $n = 0$ the assertion is trivially satisfied, as $g_j^{[0]}(x) = -i(x + 2\mathcal{M}_j^{[0]}(x, \lambda_0))^{-1}$, with $\mathcal{M}_1^{[0]}(x, \lambda_0)) = 0$ and $\mathcal{M}_2^{[0]}(x, \lambda_0)) = \lambda_0$. Then for all $\lambda_0 \in \Lambda_0^*$ one has $\partial_{\lambda_0} g_1^{[0]}(x, \lambda_0) = 0$ and $\partial_{\lambda_0} g_2^{[0]}(x, \lambda_0) = -2i(x + 2\lambda_0)^{-2}$, and one can consider the Whitney extension of $\mathcal{M}_j^{[0]}(x, \lambda_0))$ in all $\overline{\Lambda}_0$.

For $n \geq 1$ assume that all $\mathcal{M}_{j'}^{[\leq n']}(x, \lambda_0)$, $n' < n$, can be extended to differentiable functions in $\overline{\Lambda}_0$ and satisfy the bounds in (6.5). Then the derivative $\partial_{\lambda_0} g_\ell^{\mathcal{R}}$, in (6.6), can be bounded because of the inductive hypothesis. Simply, one reasons as in the proof of Lemma 19, and (6.5) follows. ■

The Whitney differentiability of $\mathcal{M}_j^{[\leq n]}(x, \lambda_0)$ implies also that of $\overline{u}(t, \lambda_0)$ and $\mu(\lambda_0)$. Hence the following result follows immediately from Lemma 24.

Lemma 25 *The renormalised series for $\overline{u}(t)$ and μ converge to functions which are differentiable in the sense of Whitney in Λ_0^*.*

Now, we can conclude the proof of theorem 1.

Lemma 26 *Call Λ^* the subset of $[a, b]$ for which the system (2.19) is reducible. There exists $\varepsilon_0 > 0$ such that for all $|\varepsilon| < \varepsilon_0$ the Lebesgue measure of the set $[a, b] \setminus \Lambda^*$ is bounded proportionally to $|\varepsilon|^\sigma$, with σ as in Lemma 23.*

Proof. Write $\lambda_0 + \mu = \lambda$. We want to fix the set Λ_0 so that for $\lambda_0 \in \Lambda_0^* \subset \Lambda_0$ one has $\lambda \in \Lambda := [a, b]$. Write $\Lambda_0 = [a_0, b_0]$, with $a_0 = a - \varepsilon f_{11,\mathbf{0}} + A\varepsilon^2 C_1^{-1}$ and $b_0 = b - \varepsilon f_{11,\mathbf{0}} - A\varepsilon^2 C_1^{-1}$, where A is a constant such that for all $\lambda_0 \in \Lambda_0^*$ and all $|\varepsilon| < \varepsilon_0$ one has $|\mu - \varepsilon f_{11,\mathbf{0}}| < A\varepsilon^2 C_1^{-1}$ (this is possible by Lemma 20). Then $\mathrm{meas}(\Lambda_0) = \mathrm{meas}(\Lambda) - 2A\varepsilon^2 C_1^{-1}$, whereas $\mathrm{meas}(\Lambda_0^*) = \mathrm{meas}(\Lambda_0) - O(|\varepsilon|^\sigma)$ by Lemma 23. Call Λ^* the subset of values $\lambda \in \Lambda$ such that $\lambda = \lambda_0 + \mu$, for $\lambda \in \Lambda_0^*$ and $\mu = \mu(\lambda_0)$. By construction $\Lambda^* \subset \Lambda$.

By Lemma 25 the function $\lambda_0 \to \mu(\lambda_0)$ is differentiable in the sense of Whitney, so that

$$\frac{\mathrm{d}\lambda}{\mathrm{d}\lambda_0} = \frac{\mathrm{d}}{\mathrm{d}\lambda_0}(\lambda_0 + \mu) = 1 + \frac{\mathrm{d}\mu}{\mathrm{d}\lambda_0}, \qquad \left|\frac{\mathrm{d}\mu}{\mathrm{d}\lambda_0}\right| \leq \mathrm{const.}|\varepsilon|^2 C_1^{-1}, \tag{6.7}$$

where we explicitly used that the first contribution to μ depending on λ_0 has size $O(\varepsilon^2 C_1^{-1})$.

Therefore

$$\mathrm{meas}(\Lambda \setminus \Lambda^*) = \int_{\Lambda\setminus\Lambda^*} \mathrm{d}\lambda \le -2A\varepsilon^2 + \int_{\Lambda_0\setminus\Lambda_0^*} \mathrm{d}\lambda_0 \left| \frac{\mathrm{d}\lambda}{\mathrm{d}\lambda_0} \right| \le \mathrm{const.}|\varepsilon|^{\sigma}, \tag{6.8}$$

because $\mathrm{meas}(\Lambda_0 \setminus \Lambda_0^*) = O(|\varepsilon|^{\sigma})$ by Lemma 23, and the assertion is proved. ■

So far we assumed $0 \notin [a,b]$. If $0 \in [a,b]$ we can discard a subset $\Lambda_1 \subset [a,b]$ around the origin, of measure $O(|\varepsilon|^{\sigma})$, such that for all $\lambda \in [a,b] \setminus \Lambda_1$ one has $|\lambda| > \mathrm{const.}|\varepsilon|^{\sigma}$. Then $|\lambda_0|$ is bounded below proportionally to $|\varepsilon|^{\sigma}$, because $|\lambda - \lambda_0| = |\mu| = O(|\varepsilon|)$ and $\sigma < 1$. Though, this does not modify the bounds of the previous sections. Indeed the only difference is that also the propagators with vanishing momentum (that is on scale -1) are bounded proportionally to $|\varepsilon|^{-\sigma}$ – like those with non-zero momentum $\boldsymbol{\nu}_\ell$, which are bounded proportionally to $|\varepsilon|^{-\sigma}\gamma_{n(\boldsymbol{\nu}_\ell)}^{-1}$ – and the bounds were obtained by using that one has at worst a factor $|\varepsilon|^{-\sigma}$ per line. Then one can restrict the analysis to $[a,b] \setminus \Lambda_1$, and the same conclusions hold.

Acknowledgments. I'm indebted to Giovanni Gallavotti for useful discussions.

References

[1] A. Avila, R. Krikorian, *Reducibility or non-uniform hyperbolicity for quasi-periodic Schrödinger cocycles*, Preprint, to appear on Ann. Math.

[2] M. Bartuccelli, G. Gentile, *Lindstedt series for perturbations of isochronous systems*, Rev. Math. Phys. **14** (2002), no. 2, 121–171.

[3] F. Bonetto, G. Gallavotti, G. Gentile, V. Mastropietro, *Quasi-linear flows on tori: regularity of their linearization*, Comm. Math. Phys. **192** (1998), no. 3, 707–736.

[4] Ch.-Q. Cheng, *Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems*, Comm. Math. Phys **177** (1996), no. 3, 529–559.

[5] E.I. Dinaburg, Ja.G. Sinaĭ, *The one-dimensional Schrödinger equation with quasiperiodic potential*, Funkcional. Anal. i Priložen. **9** (1975), no. 4, 8–21; English translation: Functional Anal. Appl. **9** (1975), no. 4, 279–289 (1976).

[6] L.H. Eliasson, *Hamiltonian systems with linear normal form near an invariant torus*, Nonlinear dynamics (Bologna, 1988), 11–29, World Sci. Publishing, Teaneck, NJ, 1989.

[7] L.H. Eliasson, *Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation*, Comm. Math. Phys. **146** (1992), no. 3, 447–482.

[8] L.H. Eliasson, *Reducibility and point spectrum for linear quasi-periodic skew-products*, Proceedings of the International Congress of Mathematicians, Vol. II (Berlin, 1998), Doc. Math. 1998, Extra Vol. II, 779–787.

[9] L.H. Eliasson, *On the discrete one-dimensional quasi-periodic Schrödinger equation and other smooth quasi-periodic skew products*, Hamiltonian systems with three or more degrees of freedom (S'AgarÃ^3, 1995), 55–61, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., 533, Kluwer Acad. Publ., Dordrecht, 1999.

[10] G. Gallavotti, *A criterion of integrability for perturbed nonresonant harmonic oscillators. "Wick ordering" of the perturbations in classical mechanics and invariance of the frequency spectrum*, Comm. Math. Phys. **87** (1982/83), no. 3, 365–383.

[11] G. Gallavotti, F. Bonetto, G. Gentile, *Aspects of ergodic, qualitative and statistical theory of motion*, Texts and Monographs in Physics, Springer, Berlin, 2004.

[12] G. Gallavotti, G. Gentile, A. Giuliani, *Fractional Lindstedt series*, Preprint, to appear on J. Math. Phys.

[13] G. Gallavotti, G. Gentile, *Hyperbolic low-dimensional invariant tori and summations of divergent series*, Comm. Math. Phys. **227** (2002), no. 3, 421–460.

[14] G. Gentile, *Quasi-periodic solutions for two-level systems*, Comm. Math. Phys. **242** (2003), no. 1, 221–250.

[15] G. Gentile, *Degenerate lower-dimensional tori under the Bryuno condition*, Preprint.

[16] G. Gentile, D.A. Cortez, J.C.A. Barata, *Stability for quasi-periodically perturbed Hill's equations*, Comm. Math. Phys. **260** (2005), no. 2, 403-443.

[17] G. Gentile, G. Gallavotti, *Degenerate elliptic resonances*, Comm. Math. Phys. **257** (2005), no. 2, 319–362.

[18] G. Gentile, V. Mastropietro, *Methods for the analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications*, Rev. Math. Phys. **8** (1996), no. 3, 393–444.

[19] H. Koch, J. Lopes Dias, *Renormalization of Diophantine skew flows, with applications to the reducibility problem*, Preprint.

[20] R. Krikorian, *Réductibilité des systèmes produits-croisés à valeurs dans des groupes compacts*, Astérisque **259** (1999), vi+216 pp.

[21] R. Krikorian, *Global density of reducible quasi-periodic cocycles on* $\mathbf{T}^1 \times \mathrm{SU}(2)$, Ann. of Math. **154** (2001), no. 2, 269–326.

[22] A. Iserles, *Expansions that grow on trees*, Notices Amer. Math. Soc. **49** (2002), no. 4, 430–440.

[23] A. Iserles, S.P. Nørsett, *On the solution of linear differential equations in Lie groups*, R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **357** (1999), no. 1754, 983–1019.

[24] À. Jorba, C. Simó, *On the reducibility of linear differential equations with quasiperiodic coefficients*, J. Differential Equations **98** (1992), no. 1, 111–124.

[25] J. Lopes Dias, *A normal form theorem for Brjuno skew-systems through renormalization*, Preprint.

[26] J. Moser, J. Pöschel, *An extension of a result by Dinaburg and Sinai on quasiperiodic potentials*, Comment. Math. Helv. **59** (1984), no. 1, 39–85.

[27] L. Pastur, A. Figotin, *Spectra of random and almost-periodic operators*, Grundlehren der Mathematischen Wissenschaften 297, Springer, Berlin, 1992.

[28] H. Rüssmann, *On the one-dimensional Schrödinger equation with a quasiperiodic potential*, Nonlinear dynamics (International Conference, New York, 1979), pp. 90–107, Ann. New York Acad. Sci., 357, New York Acad. Sci., New York, 1980.

[29] W.M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics, 785, Springer, Berlin, 1980.

[30] Ju. Xu, Q. Zheng, *On the reducibility of linear differential equations with quasiperiodic coefficients which are degenerate*, Proc. Amer. Math. Soc. **126**(1998), no. 5, 1445–1451.