Available online at www.sciencedirect.com
ScienceDirect
J. Differential Equations 245 (2008) 3253–3326
Journal of Differential Equations
www.elsevier.com/locate/jde

Periodic solutions for the Schrödinger equation with nonlocal smoothing nonlinearities in higher dimension

Guido Gentile [a,*], Michela Procesi [a,b]

[a] *Dipartimento di Matematica, Università di Roma Tre, Roma I-00146, Italy*
[b] *Dipartimento di Matematica, Università di Napoli "Federico II," Napoli I-80126, Italy*

Received 7 November 2007

Available online 9 April 2008

Abstract

We consider the nonlinear Schrödinger equation in higher dimension with Dirichlet boundary conditions and with a nonlocal smoothing nonlinearity. We prove the existence of small amplitude periodic solutions. In the fully resonant case we find solutions which at leading order are wave packets, in the sense that they continue linear solutions with an arbitrarily large number of resonant modes. The main difficulty in the proof consists in a "small divisor problem" which we solve by using a renormalisation group approach.

1. Introduction and results

In this paper we prove the existence of small amplitude periodic solutions for a class of nonlinear Schrödinger equations in D dimensions

$$i v_t - \Delta v + \mu v = f\big(x, \Phi(v), \Phi(\bar{v})\big) := \big|\Phi(v)\big|^2 \Phi(v) + F\big(x, \Phi(v), \Phi(\bar{v})\big), \tag{1.1}$$

with Dirichlet boundary conditions on the square $[0, \pi]^D$. Here $D \geqslant 2$ is an integer, μ is a real parameter, Φ is a smoothing operator, which in Fourier space acts as

$$\big(\Phi(u)\big)_k = |k|^{-2s} u_k, \tag{1.2}$$

* Corresponding author.
E-mail addresses: gentile@mat.uniroma3.it (G. Gentile), procesi@mat.uniroma3.it (M. Procesi).

0022-0396/$ – see front matter
doi:10.1016/j.jde.2008.02.037

for some positive s, and F is an analytic odd function, real for real u, such that $F(x,u,\bar u)$ is of order higher than three in $(u,\bar u)$, i.e.

$$F(x,u,\bar u)=\sum_{p=4}^{\infty}\sum_{p_1+p_2=p} a_{p_1,p_2}(x)u^{p_1}\bar u^{p_2}, \qquad F(-x,-u,-\bar u)=-F(x,u,\bar u). \tag{1.3}$$

In particular this implies that the functions a_{p_1,p_2} must be even for odd p and odd for even p. To simplify the analysis we assume that $i^{p+1}a_{p_1,p_2}(x)$ is real. Such an assumption could be removed, see [22].

For $D=2$ we do not impose any further condition on f, whereas for $D\geqslant 3$ we shall consider a more restrictive class of nonlinearities, by requiring

$$f(x,u,\bar u)=\frac{\partial}{\partial \bar u}H(x,u,\bar u)+g(x,\bar u), \qquad \overline{H(x,u,\bar u)}=H(x,u,\bar u), \tag{1.4}$$

i.e. with H a real function and g depending explicitly only on $\bar u$ (besides x) and not on u. Studying Hamiltonian perturbations is quite natural. In fact, we can extend the analysis to more general perturbations by including functions depending only on $\bar u$, even if we cannot provide a physical motivation for them. This limitation is due to technical difficulties which do not occur for $D=2$, where any analytic perturbation is allowed.

A particular case of (1.4) occurs when $H(x,u,\bar u)=F(x,|u|^2)$: this is usually referred to as the *gauge-invariant* case.

In general when looking for small periodic solutions for PDEs one expects to find a "small divisor problem" due to the fact that the eigenvalues of the linear term accumulate to zero in the space of T-periodic solutions, for any T in a positive measure set.

The case of one space dimension was widely studied in the '90 for nonresonant equations by using KAM theory by Kuksin and Pöshel [25,26] and Wayne [32], and by using Lyapunov–Schmidt decomposition by Craig and Wayne [13] and Bourgain [5,9]. The two techniques are somehow complementary. The Lyapunov–Schmidt decomposition is more flexible: it can be successfully adapted to non-Hamiltonian equations and to "resonant" equations, i.e. where the linear frequencies are not rationally independent (see for instance [2–4,8,20,27] and [31] in the case of the nonlinear wave equation). On the other hand, KAM theory provides more information, for instance on the stability of the solutions.

Generally speaking the main feature which is used to solve the small divisor problem (in all the above mentioned techniques) is the "separation of the resonant sites." Such a feature can be described as follows. For instance for $D=1$ consider an equation $\mathbb{D}[u]=f(u)$, where $\mathbb{D}$ is a linear differential operator and $f(u)$ a smooth super-linear function; let λ_k with $k\in\mathbb{Z}^2$ be the linear eigenvalues in the space of T-periodic solutions, so that after rescaling the amplitude and in Fourier space the equation has the form

$$\lambda_k u_k=\varepsilon f_k(u), \tag{1.5}$$

with $\inf_k|\lambda_k|=0$. The separation property for Dirichlet boundary conditions requires:

1. If $|\lambda_k|<\alpha$ then $|k|>C\alpha^{-\delta_0}$ (this is generally obtained by restricting T to a Cantor set).
2. If both $|\lambda_k|<\alpha$ and $|\lambda_h|<\alpha$ then either $h=k$ or $|h-k|\geqslant C(\min\{|h|,|k|\})^{\delta}$.

Here δ_0 and δ are model-dependent parameters, and C is some positive constant. In the case of periodic boundary conditions, 2 should be suitably modified.

It is immediately clear that 2 cannot be satisfied by our equation (1.1) as the linear eigenvalues are

$$\lambda_{n,m} = -\omega n + |m|^2 + \mu, \quad \omega = \frac{2\pi}{T}, \ (n,m) \in \mathbb{Z} \times \mathbb{Z}^D, \tag{1.6}$$

so that all the eigenvalues λ_{n_1,m_1} with $n_1 = n$ and $|m_1| = |m|$ are equal to $\lambda_{n,m}$.

The existence of periodic solutions for $D > 1$ space dimensions was first proved by Bourgain in [6] and [9], by using a Lyapunov–Schmidt decomposition and a technique by Spencer and Frölich to solve the small divisor problem. Again the separation properties are crucial: 1 is assumed and 2 is weakened in the following way:

2′. The sets of $k \in \mathbb{Z}^{D+1}$ such that $|\lambda_k| < 1$ and $R < |k| < 2R$ are separated in clusters, say C_j with $j \in \mathbb{N}$, such that each cluster contains at most R^{δ_1} elements and $\mathrm{dist}(C_i, C_j) \geqslant R^{\delta_2}$, with $0 < \delta_2 \leqslant \delta_1 \ll 1$.

Now, in order to apply Spencer and Frölich's method, one has to control the eigenvalues of appropriate matrices of dimension comparable to $|C_j|$. Such a dimension goes to infinity with R and at the same time the linear eigenvalues go to zero, so that achieving such estimates is a rather delicate question.

Recently Bourgain also proved the existence of quasi-periodic solutions for the nonlinear Schrödinger equation, with local nonlinearities (which corresponds to $s = 0$ in (1.2)), in any dimensions [10]. Still more recently in [15], Eliasson and Kuksin proved the same result by using KAM techniques. We can also mention a very recent paper by Yuan [34], where a variant of the KAM approach was provided to show the existence of quasi-periodic solutions: in this version, stability of the solutions is not obtained, but, conversely, the proof rather simplifies with respect to that given in [15].

In [18], Geng and You have proved, via KAM theory, the existence of quasi-periodic solutions for the NLS with a nonlocal smoothing nonlinearity which does not explicitly depend on the space variables *and* with periodic boundary conditions; under these assumptions they show the existence of a symmetry, which greatly simplifies the analysis. In the case of Dirichlet boundary condition *or* with nonlinearities depending explicitly on x, such as (1.3), this symmetry is broken, so that the results of [18] do not apply to Eq. (1.1) with F depending explicitly on x and/or with Dirichlet boundary conditions.

In this paper we give a different proof of existence of periodic solutions for the nonlinear Schrödinger equation (1.1) with Dirichlet boundary conditions (cf. Theorem 1 below). We use the Lyapunov–Schmidt decomposition and then the so-called (slightly improperly) "Lindstedt series method" [19] to solve the small divisor problem. The main purpose is to establish appropriate techniques and notation in the simplest (nontrivial) possible case which still carries the main difficulties of the D space dimensions. This motivates our choice of Eq. (1.1), with the nonlocal smoothing nonlinearities.

Nonlocal nonlinear Schrödinger equations appear in several contexts in physics, but not of the form (1.1) we are considering (see, for instance, [16,28,29,35,36]). Although such equations do not arise from any physical situation that we are aware of, the regularisation through a smoothing function was already considered in the literature (see [1,30] and [18]), as it provides some nice simplifications when looking for periodic or quasi-periodic or almost-periodic solutions in PDE problems. In our case we can take any regularisation (that is any $s > 0$ is acceptable). We can remove the smoothing (so allowing $s = 0$), but this requires some extra work, which will be discussed elsewhere [22].

Moreover, we are able to find periodic solutions also in some non-Hamiltonian systems and in resonant cases (cf. Theorems 2 to 4 below) where the result was not known in the literature. Such a result represents the main original contribution of our paper: we have preferred to start from the simpler periodic solutions which are usually discussed in the literature (cf. Theorem 1 below) because in that case the formalism is less involved, and hence the proofs simplify in a substantial way.

In particular in the completely resonant case ($\mu = 0$ in (1.1)) we find solutions which in the absence of the perturbation reduce to wave packets, i.e. linear combinations of harmonics centered around suitable frequencies – on the contrary the periodic solutions usually discussed continue a single unperturbed harmonic. To the best of our knowledge the only results in this direction are due to Bourgain: cf. [7] and [8], which have been the main source of inspiration for our work. Solutions of this kind, which arise from the superposition of several harmonics, were already discussed [21] in the one-dimensional case. Also for these solutions the higher dimensions introduce a lot of extra difficulties. With respect to the nonresonant case, the main additional difficulties are related to the nondegeneracy property of the unperturbed solution to be continued. Such a property is very hard to check in general. Bourgain's paper deals with periodic boundary conditions, and explicitly studies the case of quasi-periodic solutions with two frequencies in $D = 2$, where the nondegeneracy property is not the main point. On the other hand, the nondegeneracy property is significantly more difficult in the case of Dirichlet boundary conditions, and we have been able to solve completely the problem only in $D = 2$; we refer to Section 8 for a more detailed discussion.

Let us now describe the general lines of the Lindstedt series approach, which were originally developed by Eliasson [14], Gallavotti [17], and Chierchia and Falcolini [11], in the context of KAM theory for finite-dimensional systems.

The main idea is to consider a "renormalisation" of Eq. (1.5) which can be proved to have solutions. More precisely we consider a new, vector-valued, equation with unknowns $U_j := \{u_k\colon\ k \in C_j\}$

$$\big(\mathbb{D}_j(\omega) + M_j\big)U_j = \varepsilon F_j(U) + L_j U_j, \tag{1.7}$$

where $\{C_j\}_{j\in\mathbb{N}}$ are appropriately chosen clusters (a precise definition will be given below), $\mathbb{D}_j(\omega)$ is the diagonal matrix of the eigenvalues λ_k with $k \in C_j$, $F_j(U)$ is the vector $\{f_k(u)\colon k \in C_j\}$ defined in (1.5), and M_j, L_j are matrices of free parameters. Eq. (1.7) coincides with (1.5) provided $M_j = L_j$ for all $j \in \mathbb{N}$.

The aim then is to proceed as in the one-dimensional renormalisation scheme proposed in [19] and [20]; namely we restrict $(\omega, \{M_j\})$ to a Cantor set and construct both the solution $U_j(\varepsilon, \omega, \{M_h\})$ and $L_j(\varepsilon, \omega, \{M_h\})$ as convergent power series in ε. Then one solves the compatibility equation $M_j = L_j(\varepsilon, \omega, \{M_h\})$; essentially this is done by the implicit function theorem but with the additional complication that L_j is defined for $(\omega, \{M_h\})$ in a Cantor set.

We look for periodic solutions of frequency $\omega = D + \mu - \varepsilon$, with $\varepsilon > 0$, which continue the unperturbed one ($\varepsilon = 0$) with frequency $\omega_0 = D + \mu$. Note that the choice of this particular unperturbed frequency is made only for the sake of definiteness: any other linear frequency would yield the same type of results.

For $\varepsilon \neq 0$ we perform the change of variables

$$\sqrt{\varepsilon} u(x,t) = \Phi\big(v(x,\omega t)\big), \tag{1.8}$$

so that (1.1) becomes

$$\Phi^{-1}(i\omega u_t - \Delta u + \mu u) = \varepsilon |u|^2 u + \frac{1}{\sqrt{\varepsilon}} F(x, \sqrt{\varepsilon} u, \sqrt{\varepsilon}\bar{u}) \equiv \varepsilon f(x,u,\bar{u},\varepsilon), \tag{1.9}$$

with a slight abuse of notation in the definition of f.

We start by considering explicitly the case $F = 0$, for simplicity, so that $f(x,u,\bar{u},\varepsilon) = f(u,\bar{u}) = |u|^2 u$. In that case the problem of the existence of periodic solutions becomes trivial, but the advantage of proceeding this way is that the construction that we are going to envisage extends easily to more general f, with some minor technical adaptations.

We pass to the equation for the Fourier coefficients, by writing

$$u(x,t) = \sum_{n\in\mathbb{Z},\, m\in\mathbb{Z}^D} u_{n,m} \mathrm{e}^{i(nt + m\cdot x)}, \tag{1.10}$$

where $\cdot$ denotes the scalar product, so that (1.9) gives

$$|m|^{2s}\big(-\omega n + |m|^2 + \mu\big) u_{n,m} = \varepsilon \sum_{\substack{n_1+n_2-n_3=n \\ m_1+m_2-m_3=m}} u_{n_1,m_1} u_{n_2,m_2} \bar{u}_{n_3,m_3} \equiv \varepsilon f_{n,m}(u,\bar{u}), \tag{1.11}$$

and the Dirichlet boundary conditions spell

$$u_{n,m} = -u_{n,S_i(m)}, \qquad S_i(e_j) = \big(1 - 2\delta(i,j)\big) e_j \quad \forall i = 1,\dots,D, \tag{1.12}$$

where $\delta(i,j)$ is Kronecker's delta and $S_i(m)$ is the linear operator that changes the sign of the ith component of m.

We proceed as follows. We perform a Lyapunov–Schmidt decomposition separating the P–Q supplementary subspaces. By definition Q is the space of Fourier labels (n,m) such that $u_{n,m}$ solves (1.11) at $\varepsilon = 0$. If $\mu \neq 0$ we impose an irrationality condition on μ, i.e. $\omega_0 n - p \neq 0$, so that Q is defined as

$$Q := \big\{(n,m) \in \mathbb{Z} \times \mathbb{Z}^D \colon\, n = 1,\ m_i = \pm 1\ \forall i\big\}. \tag{1.13}$$

By the Dirichlet boundary conditions, calling $V = \{1,1,\dots,1\}$, for all $(1,m) \in Q$ we have that $u_{1,m} = \pm u_{1,V}$; see (1.12). Then (1.11) naturally splits into two sets of equations: the Q equations, for (n,m) such that $n = 1$ and $|m| = \sqrt{D}$, and the P equations, for all the other values of (n,m). We first solve the P equations keeping $q := u_{1,V}$ as a parameter. Then we consider the Q equations and solve them via the implicit function theorem.

We look for solutions of (1.11) such that $u_{n,m} \in \mathbb{R}$ for all (n,m); this is possible as one can find real solutions for the bifurcation equations in Q, and then the recursive P–Q equations are closed on the subspace of real $u_{n,m}$. The same condition can be imposed also in the more general case (1.3), provided the functions $i^{p_1+p_2+1} a_{p_1,p_2}(x)$ are real, as we are assuming.

For $\mu \neq 0$ we shall construct periodic solutions which are analytic both in time and space, and not only sub-analytic, as for instance in [6]. This is due to the presence of the smoothing nonlinearity.

Theorem 1. *Consider Eq.* (1.9), *with Φ defined by* (1.2) *for arbitrary $s > 0$ and F given by* (1.3) *if $D = 2$ and by* (1.3) *and* (1.4) *if $D \geqslant 3$. There exist a Cantor set $\mathfrak{M} \subset (0, \mu_0)$ and a constant ε_0 such that the following holds. For all $\mu \in \mathfrak{M}$ there exists a Cantor set $\mathfrak{E}(\mu) \subset (0, \varepsilon_0)$, such that for all $\varepsilon \in \mathfrak{E}(\mu)$ the equation admits a solution $u(x,t)$, which is 2π-periodic in time, analytic in time and in space, such that*

$$\left| u(x,t) - (2i)^D q_0 \mathrm{e}^{it} \prod_{i=1}^{D} \sin x_i \right| \leqslant C\varepsilon, \qquad q_0 = \sqrt{D^s 3^{-D}}, \tag{1.14}$$

uniformly in (x,t). The set $\mathfrak{M}$ has full measure and for all $\mu \in \mathfrak{M}$ the set $\mathfrak{E} = \mathfrak{E}(\mu)$ has positive Lebesgue measure and

$$\lim_{\varepsilon \to 0^+} \frac{\operatorname{meas}(\mathfrak{E} \cap [0,\varepsilon])}{\varepsilon} = 1, \tag{1.15}$$

where meas *denotes the Lebesgue measure.*

In [22], by refining the analysis we show that the result extends to the case $s = 0$.
For $\mu = 0$ the following result extends Theorem 1 of [21] to the higher-dimensional case.

Theorem 2. *Consider Eq.* (1.9) *with $\mu = 0$, $D \geqslant 2$, Φ defined by* (1.2) *and F given by* (1.3) *and* (1.4). *There exist a constant ε_0 and a Cantor set $\mathfrak{E} \subset (0, \varepsilon_0)$, such that for all $\varepsilon \in \mathfrak{E}$ the equation admits a solution $u(x,t)$, which is 2π-periodic in time, sub-analytic in time and in space, satisfying* (1.14) *and* (1.15).

Remark. For $\mu \neq 0$ we could consider unperturbed periodic solutions with other frequencies and we would obtain the same kind of results as in Theorem 1, with only some trivial changes of notation in the proofs. For $\mu = 0$ and if the functions $a_{p_1,p_2}(x)$ in (1.3) are constant, we could easily extend Theorem 2 to unperturbed solutions with different frequencies (as the proof of Lemma 8.3 shows). Considering nonconstant a_{p_1,p_2}'s would require some extra work.

For $D = 2$ the following result extends Theorem 2 of [21].

Theorem 3. *Consider Eq.* (1.9) *with $\mu = 0$, $D = 2$, Φ defined by* (1.2) *and F given by* (1.3) *and* (1.4). *Let $\mathfrak{K}$ any interval in $\mathbb{R}_+$. For $N > 4$ there exist sets $\mathcal{M}_+$ of N vectors in $\mathbb{Z}_+^2$ and sets of real amplitudes a_m with $m \in \mathcal{M}_+$ such that the following holds. Define*

$$q_0(x,t) = -4 \sum_{m \in \mathcal{M}_+} a_m \mathrm{e}^{i \frac{|m|^2}{2} t} \sin(m_1 x_1) \sin(m_2 x_2). \tag{1.16}$$

There are a finite set $\mathfrak{K}_0$ of points in $\mathfrak{K}$, a positive constant ε_0 and a set $\mathfrak{E} \subset (0, \varepsilon_0)$ (*all depending on $\mathcal{M}_+$*), *such that for all $s \in \mathfrak{K} \setminus \mathfrak{K}_0$ and $\varepsilon \in \mathfrak{E}$, Eq.* (1.9) *admits a solution $u(x,t)$, which is 2π-periodic in time, sub-analytic in time and space, such that*

$$\left|u(x,t) - q_0(x,t)\right| \leqslant C\varepsilon, \tag{1.17}$$

uniformly in (x,t). *Finally*

$$\lim_{\varepsilon \to 0^+} \frac{\operatorname{meas}(\mathfrak{E} \cap [0,\varepsilon])}{\varepsilon} = 1, \tag{1.18}$$

where meas *denotes the Lebesgue measure.*

In the case $D > 2$ we can still find a solution of the leading order of the Q equations of the form (1.16); however in order to prove the existence of a solution $u(x,t)$ of the full equation we need a "nondegeneracy condition," namely that some finite-dimensional matrix (denoted by $J_{1,1}$ and defined in Section 8) is invertible.

Theorem 4. *Consider Eq.* (1.9) *with* $\mu = 0$, $D \geqslant 2$, Φ *defined by* (1.2) *and* F *given by* (1.3) *and* (1.4). *There exist sets* $\mathcal{M}_+$ *of* N *vectors in* $\mathbb{Z}_+^D$ *and sets of real amplitudes* a_m *with* $m \in \mathcal{M}_+$ *such that the* Q *equations at* $\varepsilon = 0$ *have the solution*

$$q_0(x,t) = (2i)^D \sum_{m \in \mathcal{M}_+} a_m e^{i \frac{|m|^2}{D} t} \prod_{i=1}^{D} \sin(m_i x_i). \tag{1.19}$$

The set $\mathcal{M}_+$ *identifies a finite order matrix* $J_{1,1}$ (*depending analytically on the parameter* s). *For* $N > 1$ *if* $\det J_{1,1} = 0$ *is not an identity in* s *then the following holds. There are a finite set* $\mathfrak{K}_0$ *of points in* $\mathfrak{K}$, *a positive constant* ε_0 *and a set* $\mathfrak{E} \subset (0,\varepsilon_0)$ (*all depending on* $\mathcal{M}_+$), *such that for all* $s \in \mathfrak{K} \setminus \mathfrak{K}_0$ *and* $\varepsilon \in \mathfrak{E}$, *Eq.* (1.9) *admits a solution* $u(x,t)$, *which is* 2π*-periodic in time, sub-analytic in time and space, such that*

$$\left|u(x,t) - q_0(x,t)\right| \leqslant C\varepsilon, \tag{1.20}$$

uniformly in (x,t), *and* $\mathfrak{E}$ *satisfies the property* (1.18).

The existence of periodic solutions in the completely resonant case $\mu = 0$ holds "for most values of the parameter s." Essentially in the proof we use that $\det J_{1,1}$ is not identically zero in s (which can be explicitly proved for $D = 2$). Therefore it is not obvious how to extend Theorems 2 to 4 to the case $s = 0$ even by following the strategy in [22]. Indeed one has in principle the further difficulty of proving that $\det J_{1,1} \neq 0$ for a given value of s, in particular for $s = 0$. In [22] we have proved this result in the case of a_{p_1,p_2} constant. In the general case, for given sets $\mathcal{M}_+$, one can check such a property numerically, but of course in full generality the problem remains open.

Note that the sets $\mathcal{M}_+$ for which Theorems 3 and 4 hold are different – as the construction given in the proof shows – from those considered by Bourgain, where all vectors in $\mathcal{M}_+$ have the same modulus. In particular, when considering periodic solutions, in our case also in the gauge-preserving case we find small divisor problems, contrary to what happens for the solutions explicitly described by Bourgain.

2. Technical set-up and propositions

2.1. Separation of the small divisors

Let us require that μ is strongly nonresonant (and in a full measure set), i.e. that there exist $1 \gg \gamma_0 > 0$ and $\tau_0 > 1$ such that

$$\big|(D+\mu)n - p - a\mu\big| \geqslant \frac{\gamma_0}{|n|^{\tau_0}} \quad \forall a = 0, 1,\ (n, p) \in \mathbb{Z}^2,\ (n, p) \neq (1, D),\ n \neq 0. \tag{2.1}$$

We shall denote by $\mathfrak{M}$ the set of values $\mu \in (0, \mu_0)$ which satisfy (2.1). For $\mu \in \mathfrak{M}$ and ε_0 small enough we shall restrict ε to a large relative measure set $\mathfrak{E}_0(\gamma) \subset (0, \varepsilon_0)$ by imposing the Diophantine conditions (recall that $\omega = D + \mu - \varepsilon$)

$$\mathfrak{E}_0(\gamma) := \left\{\varepsilon \in (0, \varepsilon_0)\colon\ |\omega n - p| \geqslant \frac{\gamma}{n^{\tau_1}}\ \forall (n, p) \in \mathbb{N}^2\right\}, \tag{2.2}$$

for some $\tau_1 > \tau_0 + 1$ and $\gamma \leqslant \gamma_0/2$; see Appendix A.1. These conditions guarantee the "separation of the resonant sites," due to the regularising nonlinearity, for all pairs (n, m) and (n', m') such that $n \neq n'$; indeed we have the following result.

Lemma 2.1. *Fix $s_0 \in \mathbb{R}$. For all $\varepsilon \in \mathfrak{E}_0(\gamma)$ if for some $p \geqslant p_1$, $n, n_1 \in \mathbb{N}$ one has*

$$p^{s_0}|\omega n - p - \mu| \leqslant \gamma/2, \qquad p_1^{s_0}|\omega n_1 - p_1 - \mu| \leqslant \gamma/2, \tag{2.3}$$

then either $n = n_1$ and $p = p_1$ or $|n - n_1| \geqslant p_1^{s_0/\tau_1}$ and $n + n_1 \geqslant B_0 p_1$ for some constant B_0.

Proof. If $n - n_1 \neq 0$ one has $\gamma/|n - n_1|^{\tau_1} \leqslant |\omega(n - n_1) - (p - p_1)| \leqslant \gamma/p_1^{s_0}$, so that one obtains $p_1^{s_0} \leqslant |n - n_1|^{\tau_1}$. If $n = n_1$ then $|p - p_1| \leqslant \gamma/p_1^{s_0}$, hence $p = p_1$. Finally the inequality $n + n_1 \geqslant B_0 p_1$ follows immediately from (2.3), with the constant B_0 depending on ω and μ. □

Remark. Note that if s_0 is small enough one can always bound $B_0 p_1 \geqslant p_1^{s_0/\tau_1}$.

We shall now use the following lemma to reorder our space index set $\mathbb{Z}^D$. The proof is deferred to Appendix A.2 (see also [12,9]). Through all the paper, for any given finite set A we denote with $|A|$ the number of elements of A.

Lemma 2.2. *For all $\alpha > 0$ small enough one can write $\mathbb{Z}^D = \bigcup_{j \in \mathbb{N}} \Lambda_j$ such that*

(i) *all $m \in \Lambda_j$ are on the same sphere, i.e. for all $j \in \mathbb{N}$ there exists $p_j \in \mathbb{N}$ such that $|m|^2 \equiv p_j$ $\forall m \in \Lambda_j$;*
(ii) *Λ_j has d_j elements such that $|\Lambda_j| \equiv d_j \leqslant C_1 p_j^{\alpha}$, for some j-independent constant C_1;*
(iii) *for all $i \neq j$ such that Λ_j and Λ_i are on the same sphere (i.e. such that $p_j = p_i$) one has*

$$\operatorname{dist}(\Lambda_i, \Lambda_j) \geqslant C_2 p_j^{\beta}, \quad \beta = \frac{2\alpha}{2D + (D+2)!D^2}, \tag{2.4}$$

for some j-independent constant C_2;

(iv) *if* $d_j > 1$ *then for any* $m \in \Lambda_j$ *there exists* $m' \in \Lambda_j$ *such that* $|m - m'| < C_2 p_j^{\beta}$, *so that one has* $\operatorname{diam}(\Lambda_j) \leqslant C_1 C_2 p_j^{\alpha+\beta}$.

If $D = 2$ *one can take* $d_j = 2$ *for all* j *and* $\beta = 1/3$.

Remarks. (1) Essentially Lemma 2.2 assures that the points located on the intersection of the lattice $\mathbb{Z}^D$ with a sphere of any given radius r can be divided into a finite number of clusters, containing each just a few elements (that is of order r^{α}, $\alpha \ll 1$) and not too close to each other (that is at a distance not less than of order r^{β}, $\beta > 0$; in fact one has $\beta < \alpha$).

(2) In fact the proof given in Appendix A.2 shows that $\operatorname{diam}(\Lambda_j) <$ const. $p_j^{\alpha/D}$.

By definition we call Λ_1 the list of vectors m such that $m_i = \pm 1$ (that is $p_j = D$). In the following we shall take $\alpha \ll \min\{s, 1\}$, with s given in (1.2).

2.2. *Renormalised P–Q equations*

For $(n, j) \neq (1, 1)$, let us define

$$U_{n,j} = \{u_{n,m}\}_{m \in \Lambda_j}, \tag{2.5}$$

which is a vector in $\mathbb{R}^{d_j}$. Recall that $p_j = |m|^2$ if $m \in \Lambda_j$; the equations for $U_{n,j}$ are then by definition

$$p_j^s \delta_{n,j} U_{n,j} = \varepsilon F_{n,j}, \tag{2.6}$$

where

$$\delta_{n,j} = -\omega n + p_j + \mu, \qquad F_{n,j} = \{f_{n,m}\}_{m \in \Lambda_j}. \tag{2.7}$$

We introduce the ε-dependent

$$y_{n,j} := p_j^{s_2} \delta_{n,j}, \tag{2.8}$$

where the exponent $s_2 < s$ will be fixed in the forthcoming Definition 2.5(iv), and we define the *renormalised P equations* (for $(n, j) \neq (1, 1)$) as

$$p_j^s \big(\delta_{n,j} I + p_j^{-s} \bar{\chi}_1(y_{n,j}) M_{n,j}\big) U_{n,j} = \eta F_{n,j} + L_{n,j} U_{n,j}, \tag{2.9}$$

where I (the identity), $M_{n,j}$ and $L_{n,j}$ are $d_j \times d_j$ matrices and $\bar{\chi}_1$ is a C^{∞} nonincreasing function such that (see Fig. 2 below)

$$\begin{cases} \bar{\chi}_1(x) = 1, & \text{if } |x| < \gamma/8, \\ \bar{\chi}_1(x) = 0, & \text{if } |x| > \gamma/4, \end{cases} \tag{2.10}$$

and $\bar{\chi}_1'(x) < C\gamma^{-1}$ for some positive constant C (the prime denotes derivative with respect to the argument).

Clearly (2.9) coincides with (2.6), hence with (1.11), provided

$$\eta = \varepsilon, \qquad \bar{\chi}_1(y_{n,j}) M_{n,j} = L_{n,j}, \tag{2.11}$$

for all $(n, j) \neq (1, 1)$. The matrices $L_{n,j}$ will be called the *counterterms*.

We complete the renormalised P equations with the *renormalised Q equations*

$$D^s q = \sum_{\substack{n_1+n_2-n_3=1 \\ n_i=1}} \sum_{\substack{m_1+m_2-m_3=V \\ m_i \in \Lambda_1}} u_{n_1,m_1} u_{n_2,m_2} u_{n_3,m_3} + \sum^*_{\substack{n_1+n_2-n_3=1 \\ m_1+m_2-m_3=V}} u_{n_1,m_1} u_{n_2,m_2} u_{n_3,m_3}, \tag{2.12}$$

where the symbol $\sum^*$ implies the restriction to the triples of (n_i, m_i) such that at least one has not $n_i = |m_i|^2 = 1$. It should be noticed that the second sum vanishes at $\eta = 0$.

2.3. Matrix spaces

Here we introduce some notations and properties that we shall need in the following.

Definition 2.3. Let A be a $d \times d$ real-symmetric matrix, and denote with $A(i, j)$ and $\lambda^{(i)}(A)$ its entries and its eigenvalues, respectively. Given a list $\underline{m} := \{m_1, \dots, m_d\}$ with $m_i \in \mathbb{Z}^D$ and a positive number σ, we define the norms

$$|A|_\infty := \max_{i,j \leqslant d} \left|A(i, j)\right|, \qquad |A|_{\sigma,\underline{m}} := \max_{i,j \leqslant d} \left|A(i, j)\right| \mathrm{e}^{\sigma |m_i - m_j|^\rho},$$
$$\|A\| := \frac{1}{\sqrt{d}} \sqrt{\operatorname{tr}\left(A^T A\right)} = \sqrt{\frac{1}{d} \sum_{i,j=1}^{d} A(i, j)^2}, \tag{2.13}$$

with ρ depending on D. For fixed $\underline{m} = \{m_1, \dots, m_d\} \in \mathbb{Z}^{dD}$ we call $\mathcal{A}(\underline{m})$ the space of $d \times d$ real-symmetric matrices A with norm $|A|_{\sigma,\underline{m}}$.

Lemma 2.4. *Given a matrix $A \in \mathcal{A}(\underline{m})$, the following properties hold.*

(i) *The norm $\|A\|$ is a smooth function in the coefficients $A(i, j)$.*
(ii) *One has $\frac{1}{\sqrt{d}} \|A\| \leqslant |A|_\infty \leqslant \sqrt{d} \|A\|$.*
(iii) *One has $\frac{1}{\sqrt{d}} \max_i \sqrt{\lambda^{(i)}(A^T A)} \leqslant \|A\| \leqslant \max_i \sqrt{\lambda^{(i)}(A^T A)}$.*
(iv) *For invertible A one has*

$$\partial_{A(i,j)} A^{-1}(h, l) = -A^{-1}(h, i) A^{-1}(j, l), \qquad \partial_{A(i,j)} \|A\| = \frac{A(i, j)}{d \|A\|}. \tag{2.14}$$

Proof. Item (i) follows by the invariance of the characteristic polynomial under change of coordinates.

Items (ii) and (iii) are trivial.

The first relation in item (iv) follows by the definition of differential as

$$D_A f(A)[B] \equiv \partial_\varepsilon f(A+\varepsilon B)|_{\varepsilon=0}. \tag{2.15}$$

Now by Taylor expansion we get $D_A(A)^{-1}[B] = -A^{-1}BA^{-1}$. The second relation is trivial. □

Remark. Note that for A symmetric one has $\sqrt{\lambda^{(i)}(A^T A)} = |\lambda^{(i)}(A)|$.

Definition 2.5. Let $\{\Lambda_j\}_{j=1}^{\infty}$ be the partition of $\mathbb{Z}^D$ introduced in Lemma 2.2. Fix α small enough with respect to $\min\{s,1\}$, with s given in (1.2). Call $\Omega \subset \mathbb{Z}\times\mathbb{N}$ the set of indices $(n,j) \neq (1,1)$ such that

$$-\frac{1}{2} + (D+\mu-\varepsilon_0)n < p_j < (D+\mu)n + \frac{1}{2}. \tag{2.16}$$

For ε_0 small enough (2.16) in particular implies $n>0$, hence $\Omega \subset \mathbb{N}^2$. With each $(n,j)\neq(1,1)$ we associate the list $\Lambda_j = \{m_j^{(1)},\dots,m_j^{(d_j)}\}$, with $d_j \leqslant C_1 p_j^{\alpha}$, and a $d_j\times d_j$ real-symmetric matrix $M_{n,j}\in\mathcal{A}(\Lambda_j)$ (see Definition 2.3), such that $M_{n,j}=0$ if $(n,j)\notin\Omega$.

- (i) We call $\mathcal{M}$ the space of all matrices which belong to a space $\mathcal{A}(\Lambda_j)$ for some $j\in\mathbb{N}$, and for $A\in\mathcal{A}(\Lambda_j)$ we set $|A|_\sigma = |A|_{\sigma,\Lambda_j}$.
- (ii) We denote the eigenvalues of $\bar{\chi}_1(y_{n,j})M_{n,j}$ with $p_j^{\alpha}\nu_{n,j}^{(i)}$, so that $\nu_{n,j}^{(i)} \leqslant C|M_{n,j}|_\infty \leqslant C|M_{n,j}|_\sigma$, for some constant C.
- (iii) For invertible $\delta_{n,j}I + p_j^{-s}\bar{\chi}_1(y_{n,j})M_{n,j}$ we define $x_{n,j}$ and $\nu_{n,j}$ by setting

$$x_{n,j} = \left|\delta_{n,j} + p_j^{-s+2\alpha}\nu_{n,j}\right| = \left\|\left(\delta_{n,j}I + p_j^{-s}\bar{\chi}_1(y_{n,j})M_{n,j}\right)^{-1}\right\|^{-1}, \tag{2.17}$$

  where the norm $\|A\|$ is introduced in Definition 2.3 – notice that $\nu_{n,j}$, hence $x_{n,j}$, depends both on ε and M;
- (iv) We call $s_1 = s-2\alpha$ and set $s_2 = s_1/4$ in (2.8).

Remark. Note that the eigenvalues $\nu_{n,j}^{(i)}$ are proportional to $\bar{\chi}_1(y_{n,j})$, hence vanish for $|y_{n,j}| > \gamma/4$.

Lemma 2.6. *There exists a positive constant C such that one has $|\nu_{n,j}| \leqslant C|M_{n,j}|_\infty \leqslant C|M_{n,j}|_\sigma$.*

Proof. For notational simplicity set $M_{n,j}=M$, $\delta_{n,j}=\delta$, $p_j=p$, $d_j=d$, $x_{n,j}=x$, $\nu_{n,j}=\nu$, $\nu_{n,j}^{(i)}=\nu_i$, and define $\lambda_i = \delta + p^{-s+\alpha}\nu_i$, with $|\nu_i|\leqslant C|M|_\infty$ (see Definition 2.5(ii)). Then one has

$$x = \left|\delta + p^{-s+2\alpha}\nu\right| = \left(\frac{1}{d}\sum_{i=1}^{d}\frac{1}{\lambda_i^2}\right)^{-1/2} \leqslant C_1^{1/2}p^{\alpha/2}\min_i|\lambda_i| \leqslant C_1^{1/2}p^{\alpha/2}\Big(|\delta| + p^{-s+\alpha}\min_i|\nu_i|\Big).$$

We distinguish between two cases.

1. If there exists $i = i_0$ such that $|\delta| < 2p^{-s+\alpha}|\nu_{i_0}|$ then one obtains

$$x \leqslant 2C_1^{1/2} p^{-s+3\alpha/2}|\nu_{i_0}| + p^{-s+3\alpha/2} \min_i |\nu_i| \leqslant 4C_1^{1/2} p^{-s+2\alpha}|\nu_{i_0}|.$$

Therefore, if $|\delta| < p^{-s+2\alpha}|\nu|/2$ one has

$$p^{-s+2\alpha}|\nu|/2 < x < 4C_1^{1/2} p^{-s+2\alpha}|\nu_{i_0}| \leqslant 4CC_1^{1/2} p^{-s+2\alpha}|M|_\infty,$$

hence $|\nu| \leqslant \text{const.}\,|M|_\infty$. If $|\delta| \geqslant p^{-s+2\alpha}|\nu|/2$ one has, by the assumption on δ, $p^{-s+2\alpha}|\nu|/2 \leqslant |\delta| < 2p^{-s+\alpha}|\nu_{i_0}| \leqslant 4p^{-s+2\alpha}|\nu_{i_0}|$, and the same bound follows.

2. If $|\delta| \geqslant 2p^{-s+\alpha}|\nu_i|$ for all $i = 1, \dots, d$, then one has

$$x = |\delta| \left(\frac{1}{d} \sum_{i=1}^{d} \frac{1}{(1 + \delta^{-1} p^{-s+\alpha} \nu_i)^2} \right)^{-1/2} = |\delta| + O\Big(p^{-s+\alpha} \max_i \nu_i\Big),$$

so that $|\nu| \leqslant \text{const.}\, p^{-\alpha} C|M|_\infty$. $\square$

Remark. The space of lists $M = \{M_{n,j}\}_{(n,j)\in\mathbb{N}^2}$ such that $M_{n,j} \in \mathcal{M}$ (cf. Definition 2.5(i)) and $|M|_\sigma = \sup_{n,j} |M_{n,j}|_\sigma < \infty$ is a Banach space, that we denote with $\mathcal{B}$.

Definition 2.7. We define $\mathfrak{D}_0 = \{(\varepsilon, M)\colon\ 0 < \varepsilon \leqslant \varepsilon_0,\ |M|_\sigma \leqslant C_0\varepsilon_0\}$, for a suitable positive constant C_0, and $\mathfrak{D}(\gamma) \subset \mathfrak{D}_0$ as the set of all $(\varepsilon, M) \in \mathfrak{D}_0$ such that $\varepsilon \in \mathfrak{E}_0(\gamma)$ and

$$\left| \omega n - \left(p_j + \mu + \frac{\nu_{n,j}}{p_j^{s_1}} \right) \right| \geqslant \frac{\gamma}{|n|^\tau} \quad \forall (n,j) \in \Omega,\ (n,j) \neq (1,1),\ n \neq 0, \tag{2.18}$$

for some $\tau > \tau_0 + 1 + D$.

Remark. We shall call *Melnikov conditions* the Diophantine conditions in (2.2) and (2.18). We shall call (2.2) the *second Melnikov conditions*, as they will be used to bound the difference of the momenta of comparable lines of the forthcoming tree formalism.

2.4. Main propositions

We state the propositions which represent our main technical results. Theorem 1 is an immediate consequence of Propositions 1 and 2 below.

Proposition 1. *Assume that $(\varepsilon, M) \in \mathfrak{D}(\gamma)$. There exist positive constants c_0, K_0, K_1, σ, η_0, Q_0 such that the following holds true. It is possible to find a sequence of matrices $L \in \mathcal{B}$,*

$$L := \big\{ L_{n,j}(\eta, \varepsilon, M; q) \big\}_{(n,j)\in\mathbb{N}^2\setminus\{(1,1)\}}, \tag{2.19}$$

such that the following holds.

(i) *There exists a unique solution $U_{n,j}(\eta, M, \varepsilon; q)$, with $(n,j) \in \mathbb{Z} \times \mathbb{N} \setminus \{(1,1)\}$, of Eq.* (2.9) *which is analytic in η, q for $|\eta| \leqslant \eta_0$, $|q| \leqslant Q_0$, $\eta_0 Q_0^2 \leqslant c_0$ and such that*

$$\big| U_{n,j}(\eta, M, \varepsilon; q)(a) \big| \leqslant |\eta| q^3 K_0 \mathrm{e}^{-\sigma(|n| + |p_j|^{1/2})}. \tag{2.20}$$

(ii) *The sequence* $L_{n,j}(\eta,\varepsilon,M;q)$ *is analytic in* η *and uniformly bounded for* $(\varepsilon,M)\in\mathfrak{D}(\gamma)$ *as*

$$\left|L(\eta,\varepsilon,M;q)\right|_\sigma \leqslant K_0|\eta|q^2. \tag{2.21}$$

(iii) *The functions* $U_{n,j}(\eta,\varepsilon,M;q)$ *and* $L_{n,j}(\eta,\varepsilon,M;q)$ *can be extended on the set* $\mathfrak{D}_0$ *to* C^1 *functions, denoted by* $U^E_{n,j}(\eta,\varepsilon,M;q)$ *and* $L^E_{n,j}(\eta,\varepsilon,M;q)$*, such that*

$$L^E_{n,j}(\eta,\varepsilon,M;q)=L_{n,j}(\eta,\varepsilon,M;q),\qquad U^E_{n,j}(\eta,\varepsilon,M;q)=U_{n,j}(\eta,\varepsilon,M;q), \tag{2.22}$$

for all $(\varepsilon,M)\in\mathfrak{D}(2\gamma)$.

(iv) *The extended Q-equation, obtained from* (2.12) *by substituting* $U_{n,j}(\eta,\varepsilon,M;q)$ *with* $U^E_{n,j}(\eta,\varepsilon,M;q)$*, has a solution* $q^E(\eta,\varepsilon,M)$*, which is a true solution of* (2.12) *for* $(\varepsilon,M)\in\mathfrak{D}(2\gamma)$; *with an abuse of notation we shall call*

$$U^E_{n,j}(\eta,\varepsilon,M)=U^E_{n,j}\big(\eta,\varepsilon,M;q^E(\eta,\varepsilon,M)\big),$$
$$L^E_{n,j}(\eta,\varepsilon,M)=L^E_{n,j}\big(\eta,\varepsilon,M;q^E(\eta,\varepsilon,M)\big).$$

(v) *The functions* $L^E_{n,j}(\eta,\varepsilon,M)$ *satisfy the bounds*

$$\left|L^E(\eta,\varepsilon,M)\right|_\sigma \leqslant |\eta|K_1,\qquad \left|\partial_\varepsilon L^E_{n,j}(\eta,\varepsilon,M)\right|_\sigma \leqslant |\eta|K_1|n|^{1+s_2},$$
$$\sum_{(n,j)\in\Omega}\sum_{a,b=1}^{d_j}\left|\partial_{M_{n,j}(a,b)}L^E(\eta,\varepsilon,M)\right|_\sigma \mathrm{e}^{-\sigma|m_a-m_b|^\rho}\leqslant |\eta|K_1, \tag{2.23}$$

with ρ *depending on* D*, and one has*

$$\left|U^E_{n,j}(\eta,\varepsilon,M)\right|\leqslant |\eta|K_1\mathrm{e}^{-\sigma(|n|+|p_j|^{1/2})}, \tag{2.24}$$

uniformly for $(\varepsilon,M)\in\mathfrak{D}_0$.

Once we have proved Proposition 1, we solve the compatibility equation for the extended counterterm function $L^E_{n,j}$ $(\eta=\varepsilon,\varepsilon,M)$, which is well defined provided we choose ε_0 so that $\varepsilon_0<\eta_0$.

Proposition 2. *For all* $(n,j)\in\Omega$*, there exist* C^1 *functions* $M_{n,j}(\varepsilon):(0,\varepsilon_0)\to\mathfrak{D}_0$ *(with an appropriate choice of* C_0*) such that*

(i) $M_{n,j}(\varepsilon)$ *verifies*

$$\bar{\chi}_1(y_{n,j})M_{n,j}(\varepsilon)=L^E_{n,j}\big(\varepsilon,\varepsilon,M(\varepsilon)\big), \tag{2.25}$$

and is such that

$$\left|M_{n,j}(\varepsilon)\right|_\sigma\leqslant K_2\varepsilon,\qquad \left|\partial_\varepsilon M_{n,j}(\varepsilon)\right|_\sigma\leqslant K_2\big(1+|\varepsilon n|\big)|n|^{s_2}, \tag{2.26}$$

for a suitable constant K_2;

(ii) *the set* $\mathfrak{A} \equiv \mathfrak{A}(2\gamma)$, *defined as*

$$\mathfrak{A} = \{\varepsilon \in \mathfrak{E}_0(\gamma): (\varepsilon, M(\varepsilon)) \in \mathfrak{D}(2\gamma)\}, \tag{2.27}$$

has large relative Lebesgue measure, namely $\lim_{\varepsilon\to 0^+} \varepsilon^{-1} \operatorname{meas}(\mathfrak{A} \cap (0, \varepsilon)) = 1$.

Proof of Theorem 1. By Proposition 1(i) for all $(\varepsilon, M) \in \mathfrak{D}(\gamma)$ we can find a sequence $L_{n,j}(\eta, \varepsilon, M)$ so that there exists a unique solution $U_{n,j}(\eta, \varepsilon, M)$ of (2.6) for all $|\eta| \leqslant \eta_0$, where η_0 depends only on γ for ε_0 small enough. By Proposition 1(iii) the sequence $L_{n,j}(\eta, \varepsilon, M)$ and the solution $U_{n,j}(\eta, \varepsilon, M)$ can be extended to C^1 functions (denoted by $L^E_{n,j}(\eta, \varepsilon, M)$ and $U^E_{n,j}(\eta, \varepsilon, M)$) for all $(\varepsilon, M) \in D$. Moreover $L^E_{n,j}(\eta, \varepsilon, M) = L_{n,j}(\eta, \varepsilon, M)$ and $U^E_{n,j}(\eta, \varepsilon, M) = U_{n,j}(\eta, \varepsilon, M)$ for all $(\varepsilon, M) \in \mathfrak{D}(2\gamma)$.

Eq. (2.8) coincides with our original (2.6) provided the compatibility equations (2.10) are satisfied. Now we fix $\varepsilon_0 < \eta_0$ so that $L^E_{n,j}(\eta = \varepsilon, \varepsilon, M)$ and $U^E_{n,j}(\eta = \varepsilon, \varepsilon, M)$ are well defined. By Proposition 2(i) there exists a sequence of matrices $M_{n,j}(\varepsilon)$ which satisfies the extended compatibility equations (2.24). Finally by Proposition 2(ii) the Cantor set $\mathfrak{A}(2\gamma)$ is well defined and of large relative measure.

For all $\varepsilon \in \mathfrak{A}(2\gamma)$ the pair $(\varepsilon, M(\varepsilon))$ is by definition in $\mathfrak{D}(2\gamma)$ so that by Proposition 1(iii) one has

$$L_{n,j}(\varepsilon, \varepsilon, M(\varepsilon)) = L^E_{n,j}(\varepsilon, \varepsilon, M(\varepsilon)), \qquad u(\varepsilon, \varepsilon, M(\varepsilon); x, t) = u^E(\varepsilon, \varepsilon, M(\varepsilon); x, t), \tag{2.28}$$

so that $U_{n,j}(\varepsilon, \varepsilon, M(\varepsilon))$ solves (2.8) for $\eta = \varepsilon$. So by Proposition 2(i) $M(\varepsilon)$ solves the true compatibility equations (2.10), $\bar{\chi}_1(y_{n,j})M_{n,j}(\varepsilon) = L_{n,j}(\varepsilon, \varepsilon, M(\varepsilon))$, for all $\varepsilon \in \mathfrak{A}(2\gamma)$. Then $u(\varepsilon, \varepsilon, M(\varepsilon); x, t)$ is a true nontrivial solution of our (1.9) in $\mathfrak{A}(2\gamma)$. Then by setting $\mathfrak{E}(\mu) = \mathfrak{A}(2\gamma)$ the result follows. □

3. Recursive equations and tree expansion

In this section we find a formal solution $U_{n,j}$ of (2.9) as a power series on η; the solution $U_{n,j}$ is parameterised by the matrices $L_{n',j'}$ and it will be written in the form of a tree expansion.

We assume for $L_{n,j}(\eta, \varepsilon, M)$ and $U_{n,j}(\eta, \varepsilon, M)$, with $(n, j) \neq (1, 1)$, a formal series expansion in η, i.e.

$$L_{n,j}(\eta, \varepsilon, M) = \sum_{k=1}^{\infty} \eta^k L^{(k)}_{n,j}, \qquad U_{n,j}(\eta, \varepsilon, M) = \sum_{k=1}^{\infty} \eta^k U^{(k)}_{n,j}, \tag{3.1}$$

for all $(n, j) \neq (1, 1)$. Note that (3.1) naturally defines the vector components $u^{(k)}_{n,m}$, $m \in \Lambda_j$.

By definition we set

$$U^{(0)}_{1,1} = \{u_{1,m}: m \in \Lambda_1\}, \qquad u_{1,V} = q, \qquad U^{(k)}_{1,1} = 0, \quad k \neq 0, \tag{3.2}$$

where $V = (1, 1, \dots, 1)$. Inserting the series expansion in (2.9) we obtain for all $(n, j) \neq (1, 1)$ the recursive equations

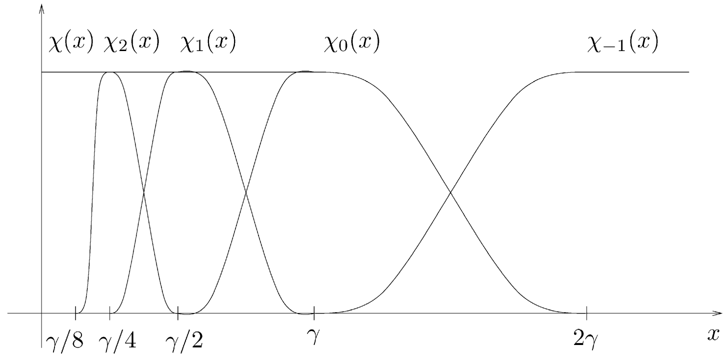

Fig. 1. Graphs of some of the C^∞ compact support functions $\chi_h(x)$ partitioning the unity. The function $\chi(x)$ is given by the envelope of all functions but $\chi_{-1}(x)$.

$$p_j^s\big(\delta_{n,j} I + p_j^{-s}\bar{\chi}_1(y_{n,j})M_{n,j}\big)U_{n,j}^{(k)} = F_{n,j}^{(k)} + \sum_{r=1}^{k-1} L_{n,j}^{(r)} U_{n,j}^{(k-r)}, \tag{3.3}$$

while for $(n, j) = (1, 1)$ we have

$$q = f_{1,V}. \tag{3.4}$$

In (3.3), for $m_a \in \Lambda_j$, where $a = 1, \dots, d_j$, $F_{n,j}^{(k)}(a)$ is defined as

$$F_{n,j}^{(k)}(a) = \sum_{k_1+k_2+k_3=k-1} \sum_{\substack{n_1+n_2-n_3=n \\ m_1+m_2-m_3=m_a}} u_{n_1,m_1}^{(k_1)} u_{n_2,m_2}^{(k_2)} u_{n_3,m_3}^{(k_3)}, \tag{3.5}$$

where each $u_{n_i,m_i}^{(k_i)}$ is a component of some $U_{n_i,j_i}^{(k_i)}$. Recall that we are assuming for the time being $f(u,\bar{u}) = |u|^2 u$ and we are looking for solutions with real Fourier coefficients $u_{n,m}$.

3.1. Multiscale analysis

It is convenient to rewrite (3.3) introducing the following scale functions.

Definition 3.1. Let $\chi(x)$ be a C^∞ nonincreasing function such that $\chi(x) = 0$ if $|x| \geqslant 2\gamma$ and $\chi(x) = 1$ if $|x| \leqslant \gamma$; moreover, if the prime denotes derivative with respect to the argument, one has $|\chi'(x)| \leqslant C\gamma^{-1}$ for some positive constant C. Let $\chi_h(x) = \chi(2^h x) - \chi(2^{h+1}x)$ for $h \geqslant 0$, and $\chi_{-1}(x) = 1 - \chi(x)$; see Fig. 1. Then

$$1 = \chi_{-1}(x) + \sum_{h=0}^{\infty} \chi_h(x) = \sum_{h=-1}^{\infty} \chi_h(x). \tag{3.6}$$

We can also write

$$1 = \bar{\chi}_1(x) + \bar{\chi}_0(x) + \bar{\chi}_{-1}(x), \tag{3.7}$$

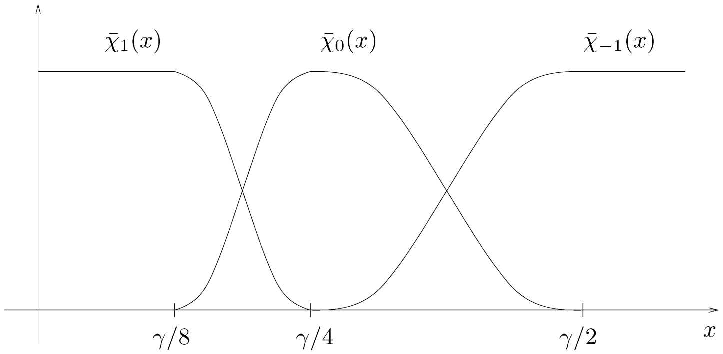

Fig. 2. Graphs of the C^∞ functions partitioning the unity $\bar\chi_{-1}(x)$, $\bar\chi_0(x)$ and $\bar\chi_1(x)$.

with $\bar\chi_1(x) = \chi(8x)$ (cf. (2.8) and Fig. 2), $\bar\chi_{-1}(x) = 1 - \chi(4x)$, and $\bar\chi_0(x) = \chi_2(x) = \chi(4x) - \chi(8x)$.

Remark. Note that $\chi_h(x) \neq 0$ implies $2^{-h-1}\gamma < |x| < 2^{-h+1}\gamma$ if $h \geqslant 0$ and $\gamma < |x|$ if $h = -1$. In particular if $\chi_h(x) \neq 0$ and $\chi_{h'}(x) \neq 0$ for $h \neq h'$ then $|h - h'| = 1$.

Definition 3.2. We denote (recall (2.17) and that $s_1 = s - 2\alpha$)

$$x_{n,j} \equiv x_{n,j}(\varepsilon, M) = \left|\delta_{n,j} + \frac{\nu_{n,j}}{p_j^{s_1}}\right|. \tag{3.8}$$

For $h = -1, 0, 1, 2, \dots, \infty$ and $i = -1, 0, 1$ we define $G_{n,j,h,i}(\varepsilon, M)$ as follows:

(i) for $i = -1, 0$, we set $G_{n,j,h,i} = 0$ for $h \neq -1$ and $G_{n,j,-1,i} = 0$ for all (ε, M) such that $\bar\chi_i(y_{n,j}) = 0$;
(ii) similarly we set $G_{n,j,h,1} = 0$ for all (ε, M) such that $\chi_h(x_{n,j}) = 0$;
(iii) otherwise we set

$$\begin{cases} G_{n,j,-1,i} = \bar\chi_i(y_{n,j}) p_j^{-s}\left(\delta_{n,j} I + \dfrac{\bar\chi_1(y_{n,j})\, M_{n,j}}{p_j^s}\right)^{-1}, & i = -1, 0, \\ G_{n,j,h,1} = \bar\chi_1(y_{n,j})\chi_h(x_{n,j}) p_j^{-s}\left(\delta_{n,j} I + \dfrac{\bar\chi_1(y_{n,j}) M_{n,j}}{p_j^s}\right)^{-1}, & h \geqslant -1. \end{cases} \tag{3.9}$$

Then $G_{n,j,h,i}$ will be called the *propagator* on scale h.

Remarks. (1) If $p_j^\alpha \nu_{n,j}^{(i)}$ are the eigenvalues of $\bar\chi_1(y_{n,j}) M_{n,j}$ (cf. Definition 2.5) one has by Lemma 2.4

$$\min_i \left|\delta_{n,j} + p_j^{-s+\alpha}\nu_{n,j}^{(i)}\right| \leqslant x_{n,j} \leqslant \min_i \sqrt{d_j}\left|\delta_{n,j} + p_j^{-s+\alpha}\nu_{n,j}^{(i)}\right|, \tag{3.10}$$

so that $\delta_{n,j} I + p_j^{-s} \bar{\chi}_1(y_{n,j}) M_{n,j}$ is invertible where $G_{n,j,h,i}(\varepsilon, M)$ is not identically zero; this implies that $G_{n,j,h,i}(\varepsilon, M)$ is well defined (and C^∞) on all $\mathfrak{D}_0$ (as given in Definition 2.7).

(2) If $i = -1, 0$, then for $(\varepsilon, M) \in \mathfrak{D}_0$ the denominators are large. Indeed $i \neq 1$ implies $|y_{n,j}| \geqslant \gamma/8$, hence $|\delta_{n,j}| \geqslant p_j^{-s_2}\gamma/8$, whereas $|p_j^{-s_1}\nu_{n,j}| \leqslant p_j^{-s_1} C C_0 |\varepsilon_0| \leqslant \text{const.}\, p_j^{-s_2}\varepsilon_0$ in $\mathfrak{D}_0$ (with C as in Lemma 2.6 and C_0 as in Definition 2.7), so that $x_{n,j} = |\delta_{n,j} + p_j^{-s_1}\nu_{n,j}| \geqslant |\delta_{n,j}|/2$. Then

$$
\begin{aligned}
|G_{n,j,-1,i}|_\infty &= p_j^{-s} \left| \left(\delta_{n,j} I + \frac{\bar{\chi}_1(y_{n,j})\, M_{n,j}}{p_j^s} \right)^{-1} \right|_\infty \\
&\leqslant C_1^{1/2} p_j^{-s+\alpha/2} \left| \delta_{n,j} + \frac{\nu_{n,j}}{p_j^{s_1}} \right|^{-1} \leqslant 2 C_1^{1/2} p_j^{-s+\alpha/2+s_2} |y_{n,j}|^{-1} \\
&\leqslant \frac{16}{\gamma} C_1^{1/2} p_j^{-3s/4},
\end{aligned}
\tag{3.11}
$$

where we have also used Lemma 2.4(ii).

(3) Notice that $G_{n,j,-1,-1}$ is a diagonal matrix (cf. (3.9) and notice that $\bar{\chi}_{-1}(y_{n,j})\bar{\chi}_1(y_{n,j}) = 0$ identically).

Inserting the multiscale decomposition (3.6) and (3.7) into (3.3) we obtain

$$
U_{n,j}^{(k)} = \sum_{i=-1,0,1} \sum_{h=-1}^{\infty} U_{n,j,h,i}^{(k)}, \tag{3.12}
$$

with

$$
U_{n,j,h,i}^{(k)} = G_{n,j,h,i} F_{n,j}^{(k)} + \delta(i,1)\, G_{n,j,h,1} \left(\sum_{h_1=-1}^{\infty} \sum_{i_1=0,1} \sum_{r=1}^{k-1} L_{n,j,h}^{(r)} U_{n,j,h_1,i_1}^{(k-r)} \right), \tag{3.13}
$$

where $\delta(i,j)$ is Kronecker's delta, and we have used that $h = -1$ for $i \neq 1$ and written

$$
L_{n,j}^{(r)} = \sum_{h=-1}^{\infty} \bar{\chi}_1(y_{n,j}) \chi_h(x_{n,j})\, L_{n,j,h}^{(r)}, \tag{3.14}
$$

with the functions $L_{n,j,h}^{(r)}$ to be determined.

3.2. Tree expansion

Eq. (3.13) can be applied recursively until we obtain the Fourier components $u_{n,m}^{(k)}$ as (formal) polynomials in the variables $G_{n,j,h,i}$, q and $L_{n,j,h}^{(r)}$ with $r < k$. It turns out that $u_{n,m}^{(k)}$ can be written as sums over *trees* (see Lemma 3.6 below), defined in the following way.

A (connected) graph $\mathcal{G}$ is a collection of points (vertices) and lines connecting all of them. The points of a graph are most commonly known as graph vertices, but may also be called *nodes* or points. Similarly, the lines connecting the nodes of a graph are most commonly known as graph

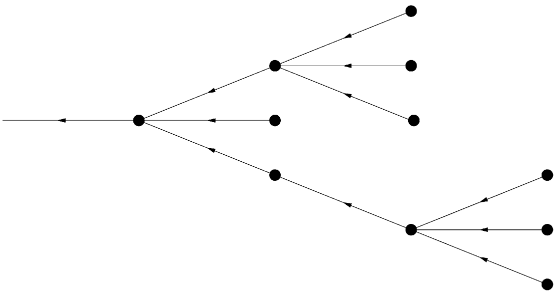

Fig. 3. Example of an unlabelled tree (only internal nodes with 1 and 3 entering lines are taken into account, according to the diagrammatic rules in Section 3.3).

edges, but may also be called branches or simply *lines*, as we shall do. We denote with $V(\mathcal{G})$ and $L(\mathcal{G})$ the set of nodes and the set of lines, respectively. A path between two nodes is the minimal subset of $L(\mathcal{G})$ connecting the two nodes. A graph is planar if it can be drawn in a plane without graph lines crossing.

Definition 3.3. A *tree* is a planar graph $\mathcal{G}$ containing no closed loops. One can consider a tree $\mathcal{G}$ with a single special node v_0: this introduces a natural partial ordering on the set of lines and nodes, and one can imagine that each line carries an arrow pointing toward the node v_0. We can add an extra (oriented) line ℓ_0 exiting the special node v_0; the added line will be called the *root line* and the point it enters (which is not a node) will be called the *root* of the tree. In this way we obtain a *rooted tree* θ defined by $V(\theta) = V(\mathcal{G})$ and $L(\theta) = L(\mathcal{G}) \cup \ell_0$. A *labelled tree* is a rooted tree θ together with a label function defined on the sets $L(\theta)$ and $V(\theta)$.

We shall call *equivalent* two rooted trees which can be transformed into each other by continuously deforming the lines in the plane in such a way that the latter do not cross each other (i.e. without destroying the graph structure). We can extend the notion of equivalence also to labelled trees, simply by considering equivalent two labelled trees if they can be transformed into each other in such a way that also the labels match. An example of tree is illustrated in Fig. 3.

Given two nodes $v, w \in V(\theta)$, we say that $w \prec v$ if v is on the path connecting w to the root line. We can identify a line with the nodes it connects; given a line $\ell = (v, w)$ we say that ℓ enters v and exits (or comes out of) w. Given two comparable lines ℓ and ℓ_1, with $\ell_1 \prec \ell$, we denote with $\mathcal{P}(\ell_1, \ell)$ the path of lines connecting ℓ_1 to ℓ; by definition the two lines ℓ and ℓ_1 do not belong to $\mathcal{P}(\ell_1, \ell)$. We say that a node v is along the path $\mathcal{P}(\ell_1, \ell)$ if at least one line entering or exiting v belongs to the path. If $\mathcal{P}(\ell_1, \ell) = \emptyset$ there is only one node v along the path (such that ℓ_1 enters v and ℓ exits v).

In the following we shall deal mostly with labelled trees: for simplicity, where no confusion can arise, we shall call them just trees.

We call *internal nodes* the nodes such that there is at least one line entering them; we call *internal lines* the lines exiting the internal nodes. We call *end-points* the nodes which have no entering line. We denote with $L(\theta)$, $V_0(\theta)$ and $E(\theta)$ the set of lines, internal nodes and end-points, respectively. Of course $V(\theta) = V_0(\theta) \cup E(\theta)$.

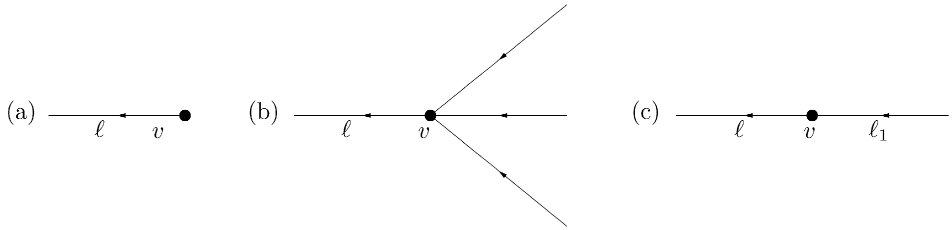

Fig. 4. Labels associated to the nodes and lines of the trees. (a) The line ℓ exits the end-point v: one associate with ℓ the labels i_ℓ, h_ℓ, n_ℓ and m_ℓ, and with v the labels n_v, m_v and k_v, with the constraints $i_\ell = -1$, $h_\ell = -1$, $n_\ell = n_v = 1$, $m_\ell = m_v \in \Lambda_1$, $k_v = 0$. (b) The line ℓ exits the node v with $s_v = 3$: one associate with ℓ the labels i_ℓ, h_ℓ, n_ℓ, j_ℓ, m_ℓ, m'_ℓ, a_ℓ, b_ℓ, and with v the label k_v, with the constraints $(n_\ell, j_\ell) \neq (1,1)$, $m_\ell = \Lambda_{j_\ell}(a_\ell)$, $m'_\ell = \Lambda_{j_\ell}(b_\ell)$, $k_v = 1$. (c) The line ℓ exits the node v with $s_v = 1$: one associate with ℓ the labels i_ℓ, h_ℓ, n_ℓ, j_ℓ, m_ℓ, m'_ℓ, a_ℓ, b_ℓ, and with v the labels k_v, a_v, b_v, j_v and n_v, with the constraints $(n_\ell, j_\ell) \neq (1,1)$, $m_\ell = \Lambda_{j_\ell}(a_\ell)$, $m'_\ell = \Lambda_{j_\ell}(b_\ell)$, $k_v \geqslant 1$, $a_v = b_\ell$, $b_v = a_{\ell_1}$, $n_\ell = n_{\ell_1}$, $j_\ell = j_{\ell_1}$. Other constraints are listed in the text.

3.3. Diagrammatic rules

We associate with the nodes (internal nodes and end-points) and lines of any tree θ some labels, according to the following rules; see Fig. 4 for reference.

(1) For each node v there are s_v entering lines, with $s_v \in \{0,1,3\}$; if $s_v = 0$ then $v \in E(\theta)$.
(2) With each end-point $v \in E(\theta)$ one associates the *mode* labels (n_v, m_v), with $m_v \in \Lambda_1$ and $n_v = 1$. One also associates with each end-point an *order* label $k_v = 0$, and a *node factor* $\eta_v = \pm q$, with the sign depending on the sign of the permutation from m_v to V: one can write $\eta_v = (-1)^{|m_v - V|_1/2} q$, where $|x|_1$ is the l_1-norm of x.
(3) With each line $\ell \in L(\theta)$ not exiting an end-point, one associates the *index* label $j_\ell \in \mathbb{N}$ and the *momenta* $(n_\ell, m_\ell, m'_\ell) \in \mathbb{Z} \times \mathbb{Z}^D \times \mathbb{Z}^D$ such that $(n_\ell, j_\ell) \neq (1,1)$ and $m_\ell, m'_\ell \in \Lambda_{j_\ell}$. One has $p_{j_\ell} = |m_\ell|^2 = |m'_\ell|^2$ (see Lemma 2.2(ii) for notations). The momenta define $a_\ell, b_\ell \in \{1, \ldots, d_j\}$, with $d_{j_\ell} = |\Lambda_{j_\ell}| \leqslant C_1 p^{\alpha}_{j_\ell}$, such that $m_\ell = \Lambda_{j_\ell}(a_\ell)$, $m'_\ell = \Lambda_{j_\ell}(b_\ell)$.
(4) With each line $\ell \in L(\theta)$ not exiting an end-point one associates a *type* label $i_\ell = -1, 0, 1$. If $i_\ell = -1$ then $m_\ell = m'_\ell$.
(5) With each line $\ell \in L(\theta)$ not exiting an end-point one associates the scale label $h_\ell \in \mathbb{N} \cup \{-1, 0\}$. If $i_\ell = 0, -1$ then $h_\ell = -1$; if two lines ℓ, ℓ' have $(n_\ell, j_\ell) = (n_{\ell'}, j_{\ell'})$, then $|i_\ell - i_{\ell'}| \leqslant 1$ and if moreover $i_\ell = i_{\ell'} = 1$ then also $|h_\ell - h_{\ell'}| \leqslant 1$.
(6) If $\ell \in L(\theta)$ exits an end-point v then $h_\ell = -1$, $i_\ell = -1$, $j_\ell = 1$, $n_\ell = 1$ and $m_\ell = m_v$.
(7) With each line $\ell \in L(\theta)$ except the root line one associates a sign $\sigma(\ell) = \pm 1$ such that for all $v \in V_0(\theta)$ one has

$$1 = \sum_{\ell \in L(v)} \sigma(\ell), \tag{3.15}$$

where $L(v)$ is the set of the s_v lines entering v. One does not associate any label σ to the root line ℓ_0.
(8) If $s_v = 1$ the labels n_{ℓ_1}, j_{ℓ_1} of the line entering v are the same as the labels n_ℓ, j_ℓ of the line ℓ exiting v, and one defines $j_v = j_\ell$, $a_v = b_\ell$, $b_v = a_{\ell_1}$. One associates with such v an order label $k_v \in \mathbb{N}$ and with ℓ a type label $i_\ell = 1$.
(9) If $s_v = 3$ then $k_v = 1$. If ℓ is the line exiting v and ℓ_1, ℓ_2, ℓ_3 are the lines entering v one has

$$n_\ell = \sigma(\ell_1)n_{\ell_1} + \sigma(\ell_2)n_{\ell_2} + \sigma(\ell_3)n_{\ell_3} = \sum_{\ell' \in L(v)} \sigma(\ell')n_{\ell'} \tag{3.16}$$

and

$$m'_\ell = \sigma(\ell_1)m_{\ell_1} + \sigma(\ell_2)m_{\ell_2} + \sigma(\ell_3)m_{\ell_3} = \sum_{\ell' \in L(v)} \sigma(\ell')m_{\ell'}, \tag{3.17}$$

with $L(v)$ defined after (3.15).

(10) With each line $\ell \in L(\theta)$ one associates the *propagator*

$$g_\ell := G_{n_\ell, j_\ell, h_\ell, i_\ell}(a_\ell, b_\ell), \tag{3.18}$$

if ℓ does not exit an end-point and $g_\ell = 1$ otherwise.

(11) With each internal node $v \in V_0(\theta)$ one associates a *node factor* η_v such that $\eta_v = 1/3$ for $s_v = 3$ and $\eta_v = L^{(k_v)}_{n_\ell, j_\ell, h_\ell}(a_v, b_v)$ for $s_v = 1$.

(12) Finally one defines the *order* of a tree as

$$k(\theta) = \sum_{v \in V(\theta)} k_v. \tag{3.19}$$

Definition 3.4. We call $\Theta^{(k)}$ the set of all the nonequivalent trees of order k defined according to the diagrammatic rules. We call $\Theta^{(k)}_{n,m}$ the set of all the nonequivalent trees of order k and with labels (n, m) associated to the root line.

Lemma 3.5. *For all $\theta \in \Theta^{(k)}$ and for all lines $\ell \in L(\theta)$ one has $|n_\ell|, |m_\ell|, |m'_\ell| \leqslant Bk$, for some constant B.*

Proof. By definition of order one has $|V_0(\theta)| \leqslant k$ and by induction one proves $|E(\theta)| \leqslant 2|V_0(\theta)| + 1$ (by using that $s_v \leqslant 3$ for all $v \in V_0(\theta)$). Hence $|E(\theta)| \leqslant 2k + 1$. Each end-point v contributes $n_v = \pm 1$ to the momentum n_ℓ of any line ℓ following v, so that $|n_\ell| \leqslant 2k + 1$ for all lines $\ell \in L(\theta)$.

Let θ_ℓ be the tree with root line ℓ and let $k(\theta_\ell)$ be its order. Then the bounds $|m_\ell|, |m'_\ell| \leqslant 2k(\theta_\ell) + 1$ can be proved by induction on $k(\theta_\ell)$ as follows. If v is the internal node which ℓ exits and $s_v = 3$, call ℓ_1, ℓ_2, ℓ_3 the lines entering v (the case $s_v = 1$ can be discussed in the same way, and it is even simpler) and for $i = 1, \ldots, 3$ denote by θ_i the tree with root line ℓ_i and by k_i the corresponding order. Then $k_1 + k_2 + k_3 = k(\theta_\ell) - 1$, so that by the inductive hypothesis one has

$$m'_\ell = m_{\ell_1} + m_{\ell_2} + m_{\ell_3} \quad \Longrightarrow \quad \left|m'_\ell\right| \leqslant \sum_{i=1}^{3} (2k_i + 1) \leqslant 2k(\theta_\ell) + 1,$$

and hence also $|m_\ell| = |m'_\ell| \leqslant 2k(\theta_\ell) + 1$. □

The coefficients $u^{(k)}_{n,m}$ can be represented as sums over the trees defined above; this is in fact the content of the following lemma.

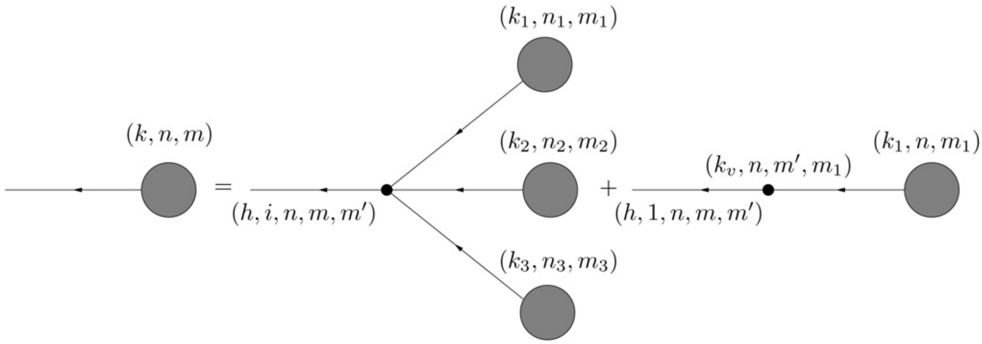

Fig. 5. Graphical representation of (3.20); the sums are understood; note that $\sum_j \sigma(\ell_j)m_j = m'$ in the first summand and $k_v + k_1 = k$ in the second summand.

Lemma 3.6. *The coefficients $u^{(k)}_{n,m}$ can be written as*

$$u^{(k)}_{n,m} = \sum_{\theta \in \Theta^{(k)}_{n,m}} \mathrm{Val}(\theta), \tag{3.20}$$

where

$$\mathrm{Val}(\theta) = \Bigg(\prod_{\ell \in L(\theta)} g_\ell\Bigg)\Bigg(\prod_{v \in V(\theta)} \eta_v\Bigg). \tag{3.21}$$

Proof. The proof is done by induction on $k \geqslant 1$. For $k = 1$ it reduces just to a trivial check.

Now, let us assume that (3.20) holds for $k' < k$, and use that $u^{(0)}_{n,m} = q\delta(n,1) \cdot \prod_{i=1}^{D}(\pm\delta(m_i, \pm 1))$. If we set $m = \Lambda_j(a)$, we have (see Fig. 5)

$$\begin{aligned} u^{(k)}_{n,m} = {} & \sum_{h=-1}^{\infty} \sum_{i=-1,0,1} \sum_{b=1}^{d_j} G_{n,j,h,i}(a,b) \sum_{k_1+k_2+k_3=k} \sum_{\substack{n_1+n_2-n_3=n \\ m_1+m_2-m_3=\Lambda_j(b)}} u^{(k_1)}_{n_1,m_1} u^{(k_2)}_{n_2,m_2} u^{(k_3)}_{n_3,m_3} \\ & + \sum_{h=-1}^{\infty} \sum_{b,b'=1}^{d_j} G_{n,j,h,1}(a,b) \sum_{r=1}^{k-1} L_{n,j,h}(b,b')\, u^{(k-r)}_{n,\Lambda_j(b')}. \end{aligned} \tag{3.22}$$

Consider a tree $\theta \in \Theta^{(k)}_{n,m}$ such that $m = \Lambda_j(a)$, $s_{v_0} = 3$ and $h_{\ell_0} = h$, if ℓ_0 is the root line of θ and v_0 is defined in 3.3. Let $\theta_1, \theta_2, \theta_3$ be the sub-trees whose root lines ℓ_1, ℓ_2, ℓ_3 enter v_0. By (3.15) one has $\sum_{j=1}^{3} \sigma(\ell_j) m_{\ell_j} = m'_{\ell_0}$, with $m'_{\ell_0} = \Lambda_j(b)$ for $b = b_{\ell_0}$. Then we have

$$\mathrm{Val}(\theta) = G_{n,j,h,i}(a,b)\, \mathrm{Val}(\theta_1)\, \mathrm{Val}(\theta_2)\, \mathrm{Val}(\theta_3), \tag{3.23}$$

and we reorder the lines so that $\sigma(\ell_3) = -1$, which produces a factor 3.

In the same way consider a tree $\theta \in \Theta^{(k)}_{n,m}$ such that $m = \Lambda_j(a)$, $s_{v_0} = 1$ and $h_{\ell_0} = h$, with the same notations as before. Let θ_1 be the sub-tree whose root line ℓ_1 enters v_0. Set $k_{v_0} = r$, $m_{v_0} = \Lambda_j(b)$, $m'_{v_0} = \Lambda_j(b')$, where $b = b_{\ell_0}$ and $b' = a_{\ell_1}$. Then

$$\mathrm{Val}(\theta) = G_{n,j,h,1}(a,b) L^{(r)}_{n,j,h}(b,b')\,\mathrm{Val}(\theta_1), \tag{3.24}$$

so that the proof is complete. □

3.4. Clusters and resonances

In the preceding section we have found a power series expansion for $U_{n,j}$ solving (2.9) and parameterised by $L_{n,j}$. However for general values of $L_{n,j}$ such an expansion is not convergent, as one can easily identify contributions at order k which are $O(k!^{\xi})$, for a suitable constant ξ. In this section we show that it is possible to choose the parameters $L_{n,j}$ in a proper way to cancel such "dangerous" contributions; in order to do this we have to identify the dangerous contributions and this will be done through the notion of *clusters* and *resonances*.

Definition 3.7. Given a tree $\theta \in \Theta^{(k)}_{n,m}$ a *cluster* T on scale h is a connected maximal set of nodes and lines such that all the lines ℓ have a scale label $\leqslant h$ and at least one of them has scale h; we shall call $h_T = h$ the scale of the cluster. We shall denote by $V(T)$, $V_0(T)$ and $E(T)$ the set of nodes, internal nodes and the set of end-points, respectively, which are contained inside the cluster T, and with $L(T)$ the set of lines connecting them. Finally $k_T = \sum_{V(T)} k_v$ will be called the order of T.

Therefore an inclusion relation is established between clusters, in such a way that the innermost clusters are the clusters with lowest scale, and so on. A cluster T can have an arbitrary number of lines entering it (*entering lines*), but only one or zero line coming out from it (*exiting line* or *root line* of the cluster); we shall denote the latter (when it exists) with ℓ^1_T. Notice that by definition all the external lines have $i_\ell = 1$.

Definition 3.8. We call 1-*resonance* on scale h a cluster T of scale $h_T = h$ with only one entering line ℓ_T and one exiting line ℓ^1_T of scale $h^{(e)}_T > h+1$, with $|V(T)| > 1$ and such that

(i) one has

$$n_{\ell^1_T} = n_{\ell_T} \geqslant 2^{(h-2)/\tau}, \qquad m'_{\ell^1_T} \in \Lambda_{j_{\ell_T}}, \tag{3.25}$$

(ii) if for some $\ell \in L(T)$ not on the path $\mathcal{P}(\ell_T, \ell^1_T)$ one has $n_\ell = n_{\ell_T}$, then $j_\ell \neq j_{\ell_T}$.

We call 2-*resonance* a set of lines and nodes which can be obtained from a 1-resonance by setting $i_{\ell_T} = 0$.

Finally we call *resonances* the 1- and 2-resonances. The line ℓ^1_T of a resonance will be called the root line of the resonance. The root lines of the resonances will be also called resonant lines.

Remarks. (1) A 2-resonance is not a cluster, but it is well defined due to condition (ii) of the 1-resonances. Indeed, such a condition implies that there is a one-to-one correspondence between 1-resonances and 2-resonances.

(2) The reason why we do not include in the definition of 1-resonances the clusters which satisfy only condition (i), i.e. such that there is a line $\ell \in L(T) \setminus \mathcal{P}(\ell_T, \ell^1_T)$ with $n_\ell = n_{\ell_T}$ and $j_\ell = j_{\ell_T}$, is that these clusters do not give any problems and can be easily controlled, as will become clear in the proof of Lemma 4.1; cf. also the subsequent Remark (1).

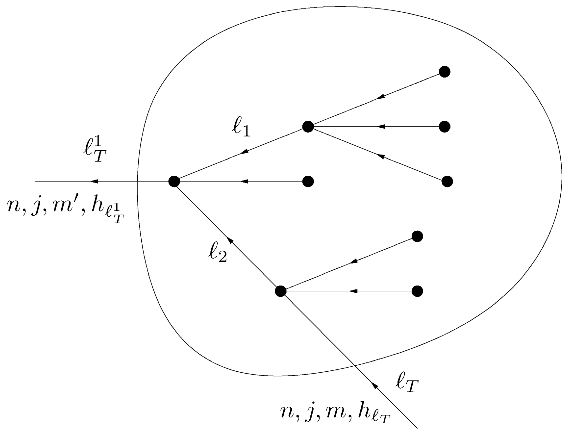

Fig. 6. Example of resonance T. We have set $j_{\ell_T^1} = j$, $n_{\ell_T^1} = n$, $m'_{\ell_T^1} = m'$, $m_{\ell_T} = m$, so that $n_{\ell_T} = n$ and $j_{\ell_T} = j$, by (3.25). Moreover, if $h_T = h$ is the scale of T, one has $h_{\ell_T} \geqslant h+1$ by definition of cluster and $h_{\ell_T^1} = h_T^{(e)} > h+1$ by definition of resonance. For any line $\ell \in L(T)$ one has $h_\ell \leqslant h$ and there is at least one line on scale h. The path $\mathcal{P}(\ell_T, \ell_T^1)$ consists of the line ℓ_1. If $n_{\ell_2} = n$ then $j_{\ell_2} \neq j$ by the condition (ii).

(3) The 2-resonances are included among the resonances for the following reason. The 1-resonances are the dangerous contributions, and we shall cancel them by a suitable choice of the counterterms. Such a choice automatically cancels out the 2-resonances.

An example of resonance is illustrated in Fig. 6. We associate a numerical value with the resonances as done for the trees. To do this we need some further notations.

Definition 3.9. The trees $\theta \in \mathcal{R}^{(k)}_{h,n,j}$ with $n \geqslant 2^{(h-2)/\tau}$ and $(n, j) \in \Omega$ are defined as the trees $\theta \in \Theta^{(k)}_{h,n,m}$ with the following modifications:

(a) there is a single end-point, called e, carrying the labels n_e, m_e such that $n_e = n$, $m_e \in \Lambda_j$; if ℓ_e is the line exiting from e then we associate with it a propagator $g_{\ell_e} = 1$, a label $m_{\ell_e} = m_e$ and a label $\sigma_{\ell_e} \in \{\pm 1\}$;
(b) the root line ℓ_0 has $i_{\ell_0} = 1$, $n_{\ell_0} = n$ and $m'_{\ell_0} \in \Lambda_j$ and the corresponding propagator is $g_{\ell_0} = 1$;
(c) one has $\max_{\ell \in L(\theta) \setminus \{\ell_0, \ell_e\}} h_\ell = h$.

A cluster T (and consequently a resonance) on scale $h_T \leqslant h$ for $\theta \in \mathcal{R}^{(k)}_{h,n,j}$ is defined as a connected maximal set of nodes $v \in V(\theta)$ and lines $\ell \in L(\theta) \setminus \{\ell_0, \ell_e\}$ such that all the lines ℓ have a scale label $\leqslant h_T$ and at least one of them has scale h_T.

We define the set $\mathcal{R}^{(k)}$ as the set of trees belonging to $\mathcal{R}^{(k)}_{h,n,j}$ for some triple (h, n, j).

Remark. The entering line ℓ_e has no label m'_{ℓ_e}, while the root line has no label m_{ℓ_0}. Both carry no scale label. Recall that by the diagrammatic rule (7) the root line ℓ_0 has no σ label.

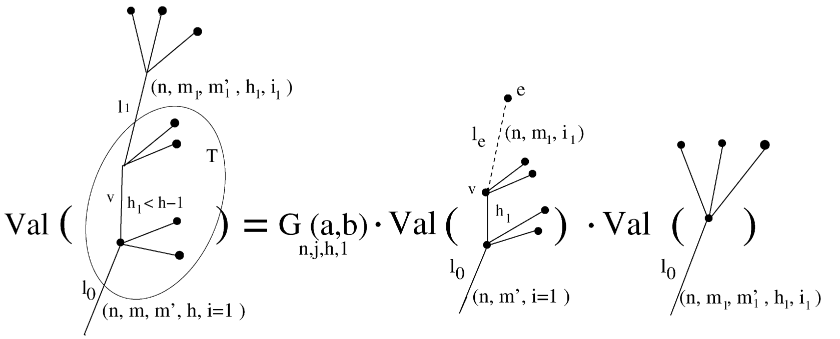

Fig. 7. We associate with the resonance T (enclosed in an ellipse and such that $m = \Lambda_j(a)$, $m' = \Lambda_j(b)$, $m_1, m'_1 \in \Lambda_j$) the tree $\theta_T \in \mathcal{R}_{h_1,n,j}$, and vice versa.

Lemma 3.10. *Let B be the same constant as in Lemma* 3.5. *For all $\theta \in \mathcal{R}^{(k)}_{h,n,j}$ and for all ℓ not in the path $\mathcal{P}(\ell_e, \ell_0)$ one has $|n_\ell| \leqslant Bk$ and $|m_\ell|, |m'_\ell| \leqslant Bk$. For ℓ on such a path one has* $\min\{|n_\ell - n_e|, |n_\ell + n_e|\} \leqslant Bk$.

Proof. For the lines not along the path $\mathcal{P} = \mathcal{P}(\ell_e, \ell_0)$ the proof is as for Lemma 3.5. If a line ℓ is along the path $\mathcal{P}$ then one can write $n_\ell = n^0_\ell \pm n_e$, where n^0_ℓ is the sum of the labels $\pm n_v$ of all the end-points preceding ℓ but e. The signs depend on the labels $\sigma(\ell')$ of the lines ℓ' preceding ℓ; in particular the sign in front of n_e depends on the labels $\sigma(\ell')$ of the lines $\ell' \in \mathcal{P}(\ell_e, \ell)$, in agreement with to (3.16). Then the last assertion follows by reasoning once more as in the proof of Lemma 3.5. □

The definition of value of the trees in $\mathcal{R}^{(k)}$ is identical to that given in (3.21) for the trees in $\Theta^{(k)}$.

Let us now consider a tree θ with a resonance T whose exiting line is the root line ℓ_0 of θ, let θ_1 be the tree atop the resonance. Given a resonance T, there exists a unique $\theta_T \in \mathcal{R}^{(k)}_{h,n,j}$, with $n = n_{\ell_0}$, $j = j_{\ell_0}$ and $h = h_T$, such that (see Fig. 7)

$$\mathrm{Val}(\theta) = g_{\ell_0}\, \mathrm{Val}(\theta_T)\, \mathrm{Val}(\theta_1), \tag{3.26}$$

so that we can call, with a slight abuse of language, $\mathrm{Val}(\theta_T)$ the value of the resonance T.

3.5. *Choice of the parameters $L_{n,j}$*

With a suitable choice of the parameters $L_{n,j,h}$ the functions $u^{(k)}_{n,m}$ can be rewritten as sum over "renormalised" trees defined below.

Definition 3.11. We define the set of *renormalised trees* $\Theta^{(k)}_{R,n,m}$ defined as the trees in $\Theta^{(k)}_{n,m}$ with no resonances nor nodes with $s_v = 1$. In the same way we define $\mathcal{R}^{(k)}_{R,h,n,j}$. We call $\mathcal{R}^{(k)}_{R,h,n,j}(a,b)$ the set of trees $\theta \in \mathcal{R}^{(k)}_{R,h,n,j}$ such that the entering line has $m_e = \Lambda_j(b)$ while the root line has $m'_{\ell_0} = \Lambda_j(a)$. Finally we define the sets $\Theta^{(k)}_R$ and $\mathcal{R}^{(k)}_R$ as the sets of trees belonging to $\Theta^{(k)}_{R,n,m}$ for some n, m and, respectively, to $\mathcal{R}^{(k)}_{R,h,n,j}$ for some h, n, j.

We extend the notion of resonant line by including also the lines coming out from a node v with $s_v = 1$. This leads to the following definition.

Definition 3.12. A resonant line is either the root line of a resonance (see Definition 3.8) or the line exiting a node v with $s_v = 1$.

The following result holds.

Lemma 3.13. *For all k, n, m one has*

$$u_{n,m}^{(k)} = \sum_{\theta \in \Theta_{R,n,m}^{(k)}} \mathrm{Val}(\theta), \tag{3.27}$$

provided we choose in (3.14)

$$\begin{cases} L_{n,j,h}^{(k)}(a,b) = -\displaystyle\sum_{h_1 < h-1} \sum_{\theta \in \mathcal{R}_{R,h_1,n,j}^{(k)}(a,b)} \mathrm{Val}(\theta), & (n,j) \in \Omega, \\ L_{n,j,h}^{(k)}(a,b) = 0, & (n,j) \notin \Omega, \end{cases} \tag{3.28}$$

where $\mathcal{R}_{R,h_1,n,j}^{(k)}(a,b)$ is as in Definition 3.11.

Proof. First note that by definition $L_{n,j,h} = 0$ if $(n,j) \notin \Omega$. We proceed by induction on k. For $k = 1$ (3.28) holds as $\Theta_{R,n,m}^{(1)} \equiv \Theta_{n,m}^{(1)}$. Then we assume that (3.28) holds for all $r < k$. By (3.13) one has $U_{n,j,h,i}^{(k)} = G_{n,j,h,i} F_{n,j}^{(k)}$for $i = -1, 0$, and

$$U_{n,j,h,1}^{(k)} = G_{n,j,h,1} F_{n,j}^{(k)} + G_{n,j,h,1} \left(\sum_{h_2=-1}^{\infty} \sum_{i_2=1,0} \sum_{r=1}^{k-1} L_{n,j,h}^{(r)} U_{n,j,h_2,i_2}^{(k-r)} \right), \tag{3.29}$$

where $F_{n,j}^{(k)}$ is a function of the coefficients $u_{n',m'}^{(r')}$ with $r' < k$. By the inductive hypothesis each $u_{n',m'}^{(r')}$ can be expressed as a sum over trees in $\Theta_{R,n',m'}^{(r')}$. Therefore $(G_{n,j,h,i} F_{n,j}^{(k)})(a)$ is given by the sum over the trees $\theta \in \Theta_{n,m}^{(k)}$, with $m = \Lambda_j(a)$ and $s_{v_0} = 3$ (v_0 is introduced in Definition 3.3), such that only the root line ℓ_0 of θ can be resonant. Note that ℓ_0 can be resonant only if $i = i_{\ell_0} = 1$. If ℓ_0 is nonresonant then $\theta \in \Theta_{R,n,m}^{(k)}$, so that the assertion holds trivially for $i \neq 1$.

For $i = 1$ we split the coefficients of $G_{n,j,h,1} F_{n,j}^{(k)}$ as sum of two terms: the first one, denoted $G_{n,j,h,1} J_{n,j}^{(k)}$, is the sum over all trees belonging to $\Theta_{R,n,m}$ for $m \in \Lambda_j$ with $s_{v_0} = 3$ and the second one is sum of trees with value

$$\mathrm{Val}(\theta) = g_{\ell_0} \mathrm{Val}(\theta_T) \mathrm{Val}(\theta_1), \tag{3.30}$$

with $\theta_T \in \mathcal{R}_{R,h_1,n,j}^{(r)}$ and $\theta_1 \in \Theta_{R,n,m'}^{(k-r)}$ with $m' = \Lambda_j(b)$ for some r and some b; by definition of resonance we have $h_1 < h - 1$.

We get terms of this type for all θ_T and θ_1 so that

$$F_{n,j}^{(k)}(a) = J_{n,j}^{(k)}(a) + \sum_{b=1}^{d_j} \sum_{h_2=-1}^{\infty} \sum_{i_2=1,0} \sum_{r=1}^{k-1} \sum_{h_1<h-1} \Bigg(\sum_{\theta \in \mathcal{R}^{(r)}_{R,h_1,n,j}(a,b)} \mathrm{Val}(\theta) \Bigg) U^{(k-r)}_{n,j,h_2,i_2}(b), \quad (3.31)$$

where the sum over $h_1 < h-1$ of the terms between parentheses gives $-L^{(r)}_{n,j,h}(a,b)$ by the first line in (3.28) Therefore all the terms but $J^{(k)}_{n,j}(a)$ in (3.31) cancel out the term between parentheses in (3.29), and only the term $G_{n,j,h,i} J^{(k)}_{n,j}(a)$ is left in (3.29). On the other hand $G_{n,j,h,i} J^{(k)}_{n,j}(a)$ is by definition the sum over all trees in $\Theta^{(k)}_{R,n,m}$, so that the assertion follows also for $i = 1$. □

Remarks. (1) The proof of Lemma 3.13 justifies why we included into the definition of resonances (cf. Definition 3.8) also the 2-resonances, even if the latter are not clusters. Indeed in (3.29) we have to sum also over $i_2 = 0$.

(2) Note that $\mathrm{Val}(\theta)$ is a monomial of degree $2k+1$ in q for $\theta \in \Theta^{(k)}_{R,n.m}$, and it is a monomial of degree $2k$ in q for $\theta \in \Theta^{(k)}_{R,n,m}$.

In the next section we shall prove that the matrices $L^{(k)}_{n,j,h}$ are symmetric (we still have to show that the matrices are well defined). For this we shall need the following result.

Lemma 3.14. *For all trees $\theta \in \mathcal{R}_{R,h,n,j}(a,b)$ there exists a tree $\theta_1 \in \mathcal{R}_{R,h,n,j}(b,a)$ such that* $\mathrm{Val}(\theta) = \mathrm{Val}(\theta_1)$.

Proof. Given a tree $\theta \in \mathcal{R}_{R,h,n,j}(a,b)$ consider the path $\mathcal{P} = \mathcal{P}(\ell_e, \ell_0)$, and set $\mathcal{P} = \{\ell_1, \dots, \ell_N\}$, with $\ell_0 \succ \ell_1 \succ \dots \succ \ell_N \succ \ell_{N+1} = \ell_e$. We construct a tree $\theta_1 \in \mathcal{R}_{R,h,n,j}(b,a)$ in the following way.

1. We shift the σ_ℓ labels down the path $\mathcal{P}$, so that $\sigma_{\ell_k} \to \sigma_{\ell_{k+1}}$ for $k = 1, \dots, N$, ℓ_0 acquires the label σ_{ℓ_1}, while ℓ_e loses its label σ_{ℓ_e} (which becomes associated with the line ℓ_N).

4. For all the lines $\ell \in \mathcal{P}$ we exchange the labels m_ℓ, m'_ℓ, so that $m_{\ell_k} \to m'_{\ell_k}$, $m'_{\ell_k} \to m_{\ell_k}$ for $k = 1, \dots, N$, while one has simply $m'_{\ell_0} \to m_{\ell_e}$ and $m_{\ell_e} \to m'_{\ell_0}$ for the root and entering lines.

3. For any pair $\ell_1(v), \ell_2(v)$ of lines not on the path $\mathcal{P}$ and entering the node v along the path, we exchange the corresponding labels σ_ℓ, i.e. $\sigma_{\ell_1(v)} \to \sigma_{\ell_2(v)}$ and $\sigma_{\ell_2(v)} \to \sigma_{\ell_1(v)}$.

2. The line ℓ_e becomes the root line, and the line ℓ_0 becomes the entering line.

As a consequence of item 2 the ordering of nodes and lines along the path $\mathcal{P}$ is reversed (in particular the arrows of all the lines $\ell \in \mathcal{P} \cup \{\ell_T, \ell_T^1\}$ are reversed). On the contrary the ordering of all the lines and nodes outside $\mathcal{P}$ is not changed by the operations above. This means that all propagators and node factors of lines and nodes, respectively, which do not belong to $\mathcal{P}$ remain the same.

Then the symmetry of M, hence of the propagators, implies the result. □

4. Bryuno lemma and bounds

In the previous section we have shown that, with a suitable choice of the parameters $L_{n,j}$, we can express the coefficients $u^{(k)}_{n,m}$ as sums over trees belonging to $\Theta^{(k)}_{R,n,m}$. We show in this section that such an expansion is indeed convergent if η is small enough and $(\varepsilon, M) \in \mathfrak{D}(\gamma)$ (see Definition 2.7).

4.1. Bounds on the trees in $\Theta_R^{(k)}$

Given a tree $\theta \in \Theta_{R,n,m}^{(k)}$, we call $\mathfrak{S}(\theta, \gamma)$ the set of $(\varepsilon, M) \in \mathfrak{D}_0$ such that for all $\ell \in L(\theta)$ one has

$$\begin{cases} 2^{-h_\ell - 1}\gamma \leqslant |x_{n_\ell, j_\ell}| \leqslant 2^{-h_\ell + 1}\gamma, & h_\ell \neq -1, \\ |x_{n_\ell, j_\ell}| \geqslant \gamma, & h_\ell = -1, \end{cases} \tag{4.1}$$

with $x_{n,j}$ defined in (2.17), and

$$\begin{cases} |y_{n_\ell, j_\ell}| \leqslant 2^{-2}\gamma, & i_\ell = 1, \\ 2^{-3}\gamma \leqslant |y_{n_\ell, j_\ell}| \leqslant 2^{-1}\gamma, & i_\ell = 0, \\ 2^{-2}\gamma \leqslant |y_{n_\ell, j_\ell}|, & i_\ell = -1, \end{cases} \tag{4.2}$$

with $y_{n,j}$ defined in (2.8). In other words we can have $\mathrm{Val}(\theta) \neq 0$ only if $(\varepsilon, M) \in \mathfrak{S}(\theta, \gamma)$.

We call $\mathfrak{D}(\theta, \gamma) \subset \mathfrak{D}_0$ the set of (ε, M) such that

$$|x_{n_\ell, j_\ell}| \geqslant \frac{\gamma}{|n_\ell|^\tau}, \tag{4.3}$$

for all lines $\ell \in L(\theta)$ such that $i_\ell = 1$, and

$$|\delta_{n_\ell, j_\ell} - \delta_{n_{\ell_1}, j_{\ell_1}}| \geqslant \frac{\gamma}{|n_\ell - n_{\ell_1}|^\tau}, \tag{4.4}$$

for all pairs of lines $\ell_1 \prec \ell \in L(\theta)$ such that $n_\ell \neq n_{\ell_1}$, $i_\ell, i_{\ell_1} = 0, 1$ and $\prod_{\ell' \in \mathcal{P}(\ell_1, \ell)} \sigma(\ell')\sigma(\ell_1) = 1$ (the last condition implies that $|n_\ell - n_{\ell_1}|$ is bounded by the sum of $|n_v|$ of the nodes v preceding ℓ but not ℓ_1). This means that $\mathfrak{D}(\theta, \gamma)$ is the set of (ε, M) verifying the Melnikov conditions (2.2) and (2.18) in θ.

In order to bound $\mathrm{Val}(\theta)$ we will use the following result (Bryuno lemma).

Lemma 4.1. *Given a tree $\theta \in \Theta_{R,n,m}^{(k)}$ such that $\mathfrak{D}(\theta, \gamma) \cap \mathfrak{S}(\theta, \gamma) \neq \emptyset$, then the scales h_ℓ of θ obey*

$$N_h(\theta) \leqslant \max\{0, ck(\theta)2^{(2-h)\beta/\tau} - 1\}, \tag{4.5}$$

where $N_h(\theta)$ is the number of lines ℓ with $i_\ell = 1$ and scale h_ℓ greater or equal than h, and c is a suitable constant.

Proof. For $(\varepsilon, M) \in \mathfrak{D}(\theta, \gamma) \cap \mathfrak{S}(\theta, \gamma)$ both (4.1) and (4.3) hold. Moreover by Lemma 3.5 one has $|n| \leqslant Bk(\theta)$. This implies that one can have $N_h(\theta) \geqslant 1$ only if $k(\theta)$ is such that $k(\theta) > k_0 := B^{-1}2^{(h-1)/\tau}$. Therefore for values $k(\theta) \leqslant k_0$ the bound (4.5) is satisfied.

If $k(\theta) > k_0$, we proceed by induction by assuming that the bound holds for all trees θ' with $k(\theta') < k(\theta)$. Define $E_h := c^{-1}2^{(-2+h)\beta/\tau}$: so we have to prove that $N_h(\theta) \leqslant$

$\max\{0, k(\theta)E_h^{-1} - 1\}$. In the following we shall assume that c is so large that all the assertions we shall make hold true.

Call ℓ_0 the root line of θ and $\ell_1, \dots, \ell_m$ the $m \geqslant 0$ lines on scale $\geqslant h-1$ which are the closest to ℓ_0 and such that $i_{\ell_s} = 1$ for $s = 1, \dots, m$.

If the root line ℓ_0 of θ is on scale $< h$ then

$$N_h(\theta) = \sum_{i=1}^{m} N_h(\theta_i), \tag{4.6}$$

where θ_i is the sub-tree with ℓ_i as root line.

By construction $N_{h-1}(\theta_i) \geqslant 1$, so that $k(\theta_i) > B^{-1}2^{(h-2)/\tau}$ and therefore for c large enough (recall that $\beta < \alpha \ll 1$) one has $\max\{0, k(\theta_i)E_h^{-1} - 1\} = k(\theta_i)E_h^{-1} - 1$, and the bound follows by the inductive hypothesis.

If the root line ℓ_0 has $i_{\ell_0} = 1$ and scale $\geqslant h$ then $\ell_1, \dots, \ell_m$ are the entering line of a cluster T.

By denoting again with θ_i the sub-tree having ℓ_i as root line, one has

$$N_h(\theta) = 1 + \sum_{i=1}^{m} N_h(\theta_i), \tag{4.7}$$

so that, by the inductive assumption, the bound becomes trivial if either $m = 0$ or $m \geqslant 2$.

If $m = 1$ then one has a cluster T with two external lines $\ell_T^1 = \ell_0$ and $\ell_T = \ell_1$, such that $h_{\ell_1} \geqslant h - 1$ and $h_{\ell_0} \geqslant h$. Then, for the assertion to hold in such a case, we have to prove that $(k(\theta) - k(\theta_1))E_h^{-1} \geqslant 1$. For $(\varepsilon, M) \in \mathfrak{S}(\theta, \gamma) \cap \mathfrak{D}(\theta, \gamma)$ one has

$$\min\bigl\{|n_{\ell_0}|, |n_{\ell_1}|\bigr\} \geqslant 2^{(h-2)/\tau}, \tag{4.8}$$

and, by definition, one has $i_{\ell_0} = i_{\ell_1} = 1$, hence $|y_{n_{\ell_0}, j_{\ell_0}}|, |y_{n_{\ell_1}, j_{\ell_1}}| \leqslant \gamma/4$ (see (4.2)), so that we can apply Lemma 2.1.

We distinguish between two cases.

1. If $n_{\ell_0} \neq n_{\ell_1}$, by Lemma 2.1 with $s_0 = s_2$ (and the subsequent Remark) one has

$$\begin{aligned} |n_{\ell_0} \pm n_{\ell_1}| &\geqslant \text{const.} \min\bigl\{|n_{\ell_0}|, |n_{\ell_1}|\bigr\}^{s_2/\tau_1} \geqslant \text{const.} \min\bigl\{|n_{\ell_0}|, |n_{\ell_1}|\bigr\}^{s_2/\tau} \\ &\geqslant \text{const.}\, 2^{(h-2)s_2/\tau^2} \geqslant E_h, \end{aligned}$$

   where we have used that $s_2/\tau^2 \geqslant \beta/\tau$ for α small enough. Therefore $B(k(\theta) - k(\theta_1)) \geqslant \min_{a=\pm 1} |n_{\ell_0} + a n_{\ell_1}| \geqslant E_h$.
2. If $n_{\ell_0} = n_{\ell_1}$, consider the path $\mathcal{P} = \mathcal{P}(\ell_1, \ell_0)$. Now consider the nodes along the path, and call ℓ_i the lines entering these nodes and θ_i the sub-trees which have such lines as root lines. If m_i denotes the momentum label m_{ℓ_i} one has, by Lemma 3.5, $|m_i| \leqslant Bk(\theta_i)$.

Call $\bar{\ell}$ the line on the path $\mathcal{P} \cup \{\ell_1\}$ closest to ℓ_0 such that $i_{\bar{\ell}} \neq -1$ (that is all lines ℓ along the path $\mathcal{P}(\bar{\ell}, \ell_0)$ have $i_\ell = -1$).

2.1. If $|n_{\bar\ell}| \leqslant |n_{\ell_0}|/2$ then, by the conservation law (3.16) one has $k(\theta) - k(\theta_1) > B^{-1}|n_{\ell_0}|/2 \geqslant E_h$.

2.2. If $|n_{\bar\ell}| \geqslant |n_{\ell_0}|/2$ we distinguish between the two following cases.

2.2.1. If $n_{\bar\ell} \neq n_{\ell_0}$ $(= n_{\ell_1})$ then by Lemma 2.1 and (4.2) one finds

$$|n_{\bar\ell} \pm n_{\ell_0}| \geqslant \text{const.} \min\bigl\{|n_{\bar\ell}|, |n_{\ell_0}|\bigr\}^{s_2/\tau} \geqslant \text{const.}\, 2^{-s_2/\tau} 2^{(h-2)s_2/\tau^2} > E_h.$$

2.2.2. If $n_{\bar\ell} = n_{\ell_0}$ then we have two further sub-cases.

2.2.2.1. If $j_{\ell_0} \neq j_{\bar\ell}$, then $|m_{\bar\ell} - m'_{\ell_0}| \geqslant C_2 p_{j_{\ell_0}}^{\beta} \geqslant C|n_{\ell_0}|^{\beta}$, for some constant C. For all the lines ℓ along the path $\mathcal{P}(\bar\ell, \ell_0)$ one has $i_\ell = -1$, hence $m_\ell = m'_\ell$ (cf. Remark (3) after Definition 3.2), so that $|m_{\bar\ell} - m'_{\ell_0}| \leqslant \sum_i |m_i| \leqslant B(k(\theta) - k(\theta_1))$, and the assertion follows once more by using (4.8).

2.2.2.2. If $j_{\ell_0} = j_{\bar\ell}$ then $i_{\bar\ell} = 0$ because $h_{\bar\ell} \leqslant h - 2$ and one would have $|h_{\bar\ell} - h| = |h_{\bar\ell} - h_{\ell_0}| \leqslant 1$ if $i_{\bar\ell} = 1$. As 2-resonances (as well as 1-resonances) are not possible there exists a line ℓ' (again with $i_{\ell'} = 0$ because $h_{\ell'} \leqslant h - 2$), not on the path $\mathcal{P}(\bar\ell, \ell_0)$, such that $j_{\ell'} = j_{\ell_0}$ and $|n_{\ell'}| = |n_{\ell_0}| > 2^{(h-2)/\tau}$; cf. condition (ii) in Definition 3.8. In this case one has $k(\theta) - k(\theta_1) > B^{-1}|n_{\ell'}| \geqslant E_h$.

This completes the proof of the lemma. □

Remarks. (1) It is just the notion of 2-resonance and property (ii) in Definition 3.8, which makes nontrivial case 2.2.2.2 in the proof of Lemma 4.1.

(2) Note that in the discussion of case 2.2.2.1 we have proved that $k(\theta) - k(\theta_1) \geqslant \text{const.}|n_{\ell_0}|^{\beta}$ (using once more that $s_2/\tau \geqslant \beta$ for α, hence β, small enough with respect to s).

The Bryuno lemma implies the following result.

Lemma 4.2. *There is a positive constant D_0 such that for all trees $\theta \in \Theta^k_{R,n,m}$ and for all $(\varepsilon, M) \in \mathfrak{D}(\theta, \gamma) \cap \mathfrak{S}(\theta, \gamma)$ one has*

$$\text{(i)} \quad \bigl|\mathrm{Val}(\theta)\bigr| \leqslant D_0^k q^{2k+1} \left(\prod_{h=1}^{\infty} 2^{hN_h(\theta)}\right) \prod_{\ell \in L(\theta)} p_{j_\ell}^{-3s/4},$$

$$\text{(ii)} \quad \bigl|\partial_\varepsilon \mathrm{Val}(\theta)\bigr| \leqslant D_0^k q^{2k+1} \left(\prod_{h=1}^{\infty} 2^{2hN_h(\theta)}\right) \prod_{\ell \in L(\theta)} p_{j_\ell}^{-s_2-\alpha},$$

$$\text{(iii)} \quad \sum_{(n',j') \in \Omega} \sum_{a',b'=1}^{d_{j'}} \bigl|\partial_{M_{n',j'}(a',b')} \mathrm{Val}(\theta)\bigr| \leqslant D_0^k q^{2k+1} \left(\prod_{h=1}^{\infty} 2^{2hN_h(\theta)}\right) \prod_{\ell \in L(\theta)} p_{j_\ell}^{-s_2-\alpha} \tag{4.9}$$

if $|n| < Bk$ and $|m| < Bk$, with B given as in Lemma 3.5, *and* $\mathrm{Val}(\theta) = 0$ *otherwise.*

Proof. By Lemma 3.5 we know that $\Theta^k_{R,n,m}$ is empty if $|n| > Bk$ or $|m| > Bk$. We first extract the factor q^{2k+1} by noticing that a renormalised tree of order k has $2k+1$ end-points (cf. the proof of Lemma 3.5).

For $(\varepsilon, M) \in \mathfrak{D}(\theta,\gamma) \cap \mathfrak{S}(\theta,\gamma)$ the bounds (4.5) hold. First of all we bound all propagators g_ℓ such that $i_\ell = -1, 0$ with $16C_1^{1/2}\gamma^{-1}|p_{j_\ell}|^{-3s/4}$ according to (3.11). For the remaining g_ℓ we use the inequalities (4.1) due to the scale functions: by Lemma 2.4(ii) one has $|G_{n,j,h,1}(a,b)| \leqslant \sqrt{d_j}\, p_j^{-s}|\delta_{n,j} + p_j^{-s_1}\nu_{n,j}|$, so that we can bound the propagators g_ℓ proportionally to $2^{h_\ell}|p_{j_\ell}|^{-3s/4}$. This proves the bound (i) in (4.9); notice that the product over the scale labels is convergent.

When deriving $\mathrm{Val}(\theta)$ with respect to ε we get a sum of trees with a distinguished line, say ℓ, whose propagator g_ℓ is substituted with $\partial_\varepsilon g_\ell$ in the tree value. For simplicity, in the following set $j = j_\ell$, $h_\ell = h$ and $n = n_\ell$.

Let us first consider the case $i_\ell = -1, 0$ (so that g_ℓ is given by the first line of (3.9)), and recall Lemma 2.4(ii) and (iii). Bounding the derivative $\partial_\varepsilon g_\ell$ we obtain, instead of the bound on g_ℓ, a factor

$$C\gamma^{-1}\frac{p_j^{s_2}|n|C_1^{1/2}p_j^{\alpha/2}}{p_j^{s}|\delta_{n,j} + p_j^{-s_1}\nu_{n,j}|} \leqslant CC_1^{1/2}\frac{16}{\gamma^2}|n|p_j^{-(s-2s_2-\alpha/2)}, \tag{4.10}$$

arising when the derivative acts on $\bar\chi_i(y_{n,j})$ (here and in the following factors $C\gamma^{-1}$ is a bound on the derivative of χ with respect to its argument), and a factor

$$\frac{2|n|C_1p_j^{\alpha}}{p_j^{s}(\delta_{n,j} + p_j^{-s_1}\nu_{n,j})^2} \leqslant 2C_1\frac{16^2}{\gamma^2}|n|p_j^{-(s-2s_2-\alpha)}, \tag{4.11}$$

arising when the derivative acts on the matrix $(\delta_{n,j}I + p_j^{-s}\bar\chi(y_{n,j})M_{n,j})^{-1}$.

If $i_\ell = 1$ then the propagator is given by the second line in (3.9), so that both summands arising from the derivation of the function $\bar\chi_1(y_{n,j})$ and of the matrix $(\delta_{n,j}I + p_j^{-s}\bar\chi(y_{n,j})M_{n,j})^{-1}$ are there, and they are both bounded proportionally to $p_j^{-s_2-\alpha}|n|2^{2h}$ (recall that $s_2 = (s-2\alpha)/4$). Moreover (see Lemma 2.4(iv)) there is also an extra summand containing a factor

$$2C\gamma^{-1}C_1^{7/2}\frac{|n|p_j^{7\alpha/2}2^{h+1}}{p_j^{s}|\delta_{n,j} + p_j^{-s_1}\nu_{n,j}|} \leqslant \text{const.}\, p_j^{-s+4\alpha}|n|2^{2h}, \tag{4.12}$$

arising when the derivative acts on $\chi_h(x_{n,j})$. Indeed, by setting $A = (\delta_{n,j}I + p_j^{-s}\bar\chi_1(y_{n,j}) \cdot M_{n,j})^{-1}$, so that $x_{n,j} = \|A\|^{-1}$, one has

$$\partial_\varepsilon x_{n,j} = \frac{1}{d_j^{1/2}\|A\|^3}\sum_{i,k,h,l=1}^{d_j} A(i,k)A(i,h)A(l,k)\partial_\varepsilon A^{-1}(h,l), \tag{4.13}$$

which implies (4.12). For $\alpha \ll s$ we can bound $s - 4\alpha$ with $s_2 + \alpha$.

Finally we can bound each $n = n_\ell$ with Bk (see Lemma 3.5). All the undistinguished lines in the tree (i.e. all lines $\ell' \neq \ell$ in $L(\theta)$) can be bounded as in item (i). This proves the bound (ii) in (4.9).

The derivative with respect to $M_{n',j'}(a',b')$ gives a sum of trees with a distinguished line ℓ (as in the previous case (ii)), with the propagator $\partial_{M_{n',j'}(a',b')} g_\ell$ replacing g_ℓ. Notice that ℓ must carry the labels n', j'. We have two contributions, one arising from the derivative of the matrix and the other one (provided $i_\ell = 1$) arising from the derivative of the scale function χ_{h_ℓ} (there is no contribution analogous to (4.10) because $y_{n,j}$ does not depend on M). By reasoning as in the case (ii) we obtain a factor proportional to $2^{2h} p_{j_\ell}^{-s_2-\alpha}$.

The sums over the labels $(n', j') \in \Omega$ and $a', b' = 1, \dots, d_{j'}$ can be bounded as follows. By Lemma 3.5 one has $|n'| < Bk$. Then j' must be such that $p_{j'} = O(n')$, which implies that the number of values which j' can assume is at most proportional to $|n'|^{D-1}$, and a', b' vary in $\{1, \dots, d_{j'}\}$, with $d_{j'} \leqslant C_1 p_{j'}^{\alpha} \leqslant \text{const.}|n'|^{\alpha}$. Therefore we obtain an overall factor proportional to $k^{1+(D-1)+2\alpha} \leqslant k^{1+D} \leqslant C^k$ for some constant C. Hence also the bound (iii) of (4.9) is proved. □

4.2. Bounds on the trees in $\mathcal{R}_R^{(k)}$

Given a tree $\theta \in \mathcal{R}_{R,h,n,j}$, we call $\tilde{\mathfrak{S}}(\theta, \gamma)$ set of $(\varepsilon, M) \in \mathfrak{D}_0$ such that (4.1) holds for all $\ell \in L(\theta) \setminus \{\ell_e, \ell_0\}$, and (4.2) holds for all $\ell \in L(\theta)$. Let $\tilde{\mathfrak{D}}(\theta, \gamma) \subset \mathfrak{D}_0$ be the set of (ε, M) such that (4.3) holds for all $\ell \in L(\theta) \setminus \{\ell_e, \ell_0\}$, and (4.4) holds for all pairs $\ell_1 \prec \ell \in L(\theta)$ such that

(i) $n_{\ell_1} \neq n_\ell$, $i_\ell, i_{\ell_1} = 0, 1$ and $\prod_{\ell' \in \mathcal{P}(\ell_1, \ell)} \sigma(\ell')\sigma(\ell_1) = 1$;
(ii) either both ℓ, ℓ_1 are on the path $\mathcal{P}(\ell_e, \ell_0)$ or none of them is on such a path.

The following lemma will be useful.

Lemma 4.3. *Given a tree $\theta \in \mathcal{R}^{(k)}_{R,\bar{h},n,j}(a,b)$ such that $\tilde{\mathfrak{D}}(\theta, \gamma) \cap \tilde{\mathfrak{S}}(\theta, \gamma) \neq \emptyset$ then there are two positive constants B_2 and B_3 such that*

(i) *a line ℓ on the path $\mathcal{P}(\ell_e, \ell_0)$ can have $i_\ell \neq -1$ only if $k \geqslant B_2|n|^\beta$;*
(ii) *one has $k \geqslant B_3|m_a - m_b|^\rho$ with $1/\rho = 1 + \alpha/\beta = 1 + D(1 + D(D+2)!/2)$.*

Proof. (i) One can proceed very closely to case 2 in the proof of Lemma 4.1, with ℓ_e and ℓ playing the role of ℓ_1 and $\bar{\ell}$, respectively – see Remark (2) after the proof of Lemma 4.1. We omit the details.

(ii) By Lemma 2.2, for all $m_a, m_b \in \Lambda_j$ one has $|m_a - m_b| \leqslant C_2 p_j^{\alpha+\beta}$. For $(n, j) \in \Omega$ this implies that $|m_a - m_b| \leqslant \text{const.}|n|^{\alpha+\beta}$. If $k \geqslant B_2|n|^\beta$ then one has $k \geqslant \text{const.}|m_a - m_b|^{\beta/(\alpha+\beta)}$, the statement holds true (recall that α/β is given by (2.4)). If $k < B_2|n|^\beta$ then by item (i) all the lines ℓ on the path $\mathcal{P}(\ell_e, \ell_0)$ have $i_\ell = -1$, hence $m_\ell = m'_\ell$. Then by calling, as in the proof of Lemma 4.1, θ_i the sub-trees whose root lines enter the nodes of $\mathcal{P}(\ell_e, \ell_0)$ and m_i the momentum label m_{ℓ_i}, we obtain $|m_{\ell_e} - m_{\ell_0}| \leqslant \sum_i |m_i| \leqslant Bk$, and the assertion follows once more. □

The following generalisation of Lemma 4.1 holds.

Lemma 4.4. *Given tree $\theta \in \mathcal{R}^{(k)}_{R,\bar{h},n,j}$ such that $\tilde{\mathfrak{D}}(\theta, \gamma) \cap \tilde{\mathfrak{S}}(\theta, \gamma) \neq \emptyset$ then the scales h_ℓ of θ obey, for all $h \leqslant \bar{h}$,*

$$N_h(\theta) \leqslant \max\{0, ck(\theta)2^{(2-h)\beta/\tau}\}, \tag{4.14}$$

where $N_h(\theta)$ and c are defined as in Lemma 4.1.

Proof. To prove the lemma we consider a slightly different class of trees with respect to $\mathcal{R}^{(k)}_{R,\bar{h},n,j}$, which we denote by $\mathcal{R}^{(k)}_{R,\bar{h}}$. The differences are as follows:

(i) the root line has scale labels $h_{\ell_0} \leqslant \bar{h}$ and $i_{\ell_0} \in \{-1, 0, 1\}$,

(ii) we remove the condition $n_e = n_{\ell_0}$, $j_e = j_{\ell_0}$, and require only that $|n_e| > 2^{(\bar{h}-2)/\tau}$.

Notice that, for all $\theta \in \mathcal{R}^{(k)}_{R,\bar{h},n,j}$, among the three sub-trees entering the root, two are in $\Theta^{(k_1)}_R$ and $\Theta^{(k_2)}_R$, respectively, and one is in $\mathcal{R}^{(k_3)}_{R,\bar{h}_1}$, with $\bar{h}_1 \leqslant \bar{h}$ (recall that by definition $h_\ell \leqslant \bar{h}$ for all $\ell \in L(\theta)$), and $k_1 + k_2 + k_3 = k - 1$. Hence we shall prove (4.14) for the trees $\theta \in \mathcal{R}^{(k)}_{R,\bar{h}}$, for which we can proceed by induction on k.

For $(\varepsilon, M) \in \tilde{\mathfrak{D}}(\theta, \gamma) \cap \tilde{\mathfrak{S}}(\theta, \gamma)$ we have both (4.1) and (4.3) for all $\ell \in L(\theta) \setminus \{\ell_0, \ell_e\}$. Moreover by Lemma 3.10 we have $Bk(\theta) \geqslant |n_\ell + an_e|$, where $a = 0$ if ℓ is not on the path $\mathcal{P} = \mathcal{P}(\ell_e, \ell_0)$ and $a \in \{\pm 1\}$ otherwise.

For ℓ not on the path $\mathcal{P}$ one can have $h_\ell \geqslant h$ only if $k(\theta)$ is such that $k(\theta) > k_0 = B^{-1}2^{(h-1)/\tau}$ (cf. the proof of Lemma 4.1). If all lines not along the path $\mathcal{P}$ have scales $< h$, consider the line ℓ on the path $\mathcal{P}$ with scale $h_\ell \geqslant h$ which is the closest to ℓ_e (the case in which such a line does not exist is trivial, because it yields $N_h(\theta) = 0$)). Then ℓ is the exiting line of a cluster T with ℓ_e as entering line. Note that we have both $|n_\ell| \geqslant 2^{(h-1)/\tau}$ and $|n_e| > 2^{(\bar{h}-2)/\tau}$, with $\bar{h} \geqslant h$. As T cannot be a resonance, if $n_\ell = n_e$ then either $j_\ell \neq j_e$, so that

$$k_T > \min\{B_2|n_e|^\beta, B^{-1}C_2|p_{j_e}|^\beta\} > \text{const.}\, 2^{(h-1)\beta/\tau}$$

(cf. Lemma 4.3(i) and case 2.2.2.1 in the proof of Lemma 4.1), or $j_\ell = j_e$, so that

$$k_T > B^{-1}2^{(\bar{h}-2)/\tau} \geqslant B^{-1}2^{(h-2)/\tau}$$

(cf. case 2.2.2.2 in the proof of Lemma 4.1). If on the contrary $n_\ell \neq n_e$, by Lemma 2.1 one has $Bk(\theta) \geqslant \text{const.} \min\{|n_\ell \pm n_e|\} \geqslant \text{const.}\, 2^{(h-2)s_2/\tau^2}$. Therefore there exists a constant $\tilde{B}$ such that for values $k(\theta) \leqslant \tilde{k}_0 := \tilde{B}^{-1}2^{(h-1)s_2/\tau^2}$ the bound (4.14) is satisfied.

If $k(\theta) > \tilde{k}_0$, we assume that the bound holds for all trees θ' with $k(\theta') < k(\theta)$. Define $E_h = c^{-1}2^{(-2+h)\beta/\tau}$: we want to prove that $N_h(\theta) \leqslant \max\{0, k(\theta)E_h^{-1}\}$.

We proceed exactly as in the proof of Lemma 4.1. The only difference is that, when discussing case 2.2.1, one can deduce $|n_{\bar{\ell}} \pm n_{\ell_0}| \geqslant \text{const.} \min\{|n_{\ell_0}|, |n_{\bar{\ell}}|\}^{s_2/\tau} \geqslant \text{const.}\, 2^{(h-2)s_2/\tau^2} > E_h$ by using that the quantity n_e cancels out as the line $\bar{\ell}$ is along the path $\mathcal{P}$. □

The following result is an immediate consequence of the previous lemma.

Lemma 4.5. *For fixed k the matrices $L^{(k)}_{n,j}$ are symmetric*; *moreover the following identity holds*:

$$L^{(k)}_{n,j} = -\bar{\chi}_1(y_{n,j}) \sum_{h=-1}^{\infty} C_h(x_{n,j}) \sum_{\theta \in \mathcal{R}^{(k)}_{R,h,n,j}} \text{Val}(\theta), \tag{4.15}$$

where, by definition,

$$C_h(x) = \sum_{h_1=h+2}^{\infty} \chi_h(x). \tag{4.16}$$

Proof. The previous analysis has shown that the matrices $L_{n,j}^{(k)}$ are well defined. Then the matrices are symmetric by Lemma 3.14, where we have established a one-to-one correspondence between the trees contributing to $L_{n,j}^{(k)}(a,b)$ and those contributing to $L_{n,j}^{(k)}(b,a)$ such that the corresponding trees have the same value. Identity (4.15) follows from the definitions (3.14) and (3.28). □

Remark. Notice that $C_h(x) = 1$ when $|x| \leqslant 2^{-h-2}\gamma$ and $C_h(x) = 0$ when $|x| \geqslant 2^{-h-1}\gamma$.

Lemma 4.6. *Given a tree $\theta \in \mathcal{R}_{R,h,n,j}^{(k)}(a,b)$, for $(\varepsilon, M) \in \tilde{\mathfrak{D}}(\theta,\gamma) \cap \tilde{\mathfrak{S}}(\theta,\gamma)$ and $\sigma > 0$ one has*

$$\text{(i)} \quad \bigl|\mathrm{Val}(\theta)\bigr| \leqslant \bigl(Dq^2\bigr)^k 2^{-h} \left(\prod_{h'=-1}^{h} 2^{h' N_{h'}(\theta)}\right) \mathrm{e}^{-\sigma|m_a - m_b|^{\rho}} \prod_{\ell \in L(\theta)} p_{j_\ell}^{-3s/4},$$

$$\text{(ii)} \quad \bigl|\partial_\varepsilon \mathrm{Val}(\theta)\bigr| \leqslant \bigl(Dq^2\bigr)^k 2^{-h} |n| \left(\prod_{h'=-1}^{h} 2^{2h' N_{h'}(\theta)}\right) \mathrm{e}^{-\sigma|m_a - m_b|^{\rho}} \prod_{\ell \in L(\theta)} p_{j_\ell}^{-s_2-\alpha},$$

$$\text{(iii)} \quad \sum_{(n',j') \in \Omega} \sum_{a',b'=1}^{d_{j'}} \bigl|\partial_{M_{n',j'}(a',b')} \mathrm{Val}(\theta)\bigr| \leqslant \bigl(Dq^2\bigr)^k 2^{-h} \left(\prod_{h'=-1}^{h} 2^{2h' N_{h'}(\theta)}\right) \mathrm{e}^{-\sigma|m_a - m_b|^{\rho}} \prod_{\ell \in L(\theta)} p_{j_\ell}^{-s_2-\alpha}, \tag{4.17}$$

for some constant D depending on σ and γ.

Proof. The proof follows the same lines as that of Lemma 4.2. We first extract the factor q^{2k} by noticing that a renormalised tree in $\mathcal{R}_R^{(k)}$ has $2k$ end-points. To extract the factor 2^{-h} we recall that there is at least a line $\ell \neq \ell_0$ on scale $h_\ell = h$: then $N_h(\theta) \geqslant 1$ and by (4.14) we obtain $k > \text{const.}\, 2^{h\beta/\tau}$, so that $C^k 2^{-h} \geqslant 1$ for a suitable constant C. To extract the factor $\mathrm{e}^{-\sigma|m_a - m_b|^{\rho}}$ we use Lemma 4.3(ii) to deduce $\tilde{C}^k \mathrm{e}^{-\sigma|m_a-m_b|^{\rho}} \geqslant 1$. Hence the bound (i) in (4.17) follows.

When applying the derivative with respect to ε to $\mathrm{Val}(\theta)$ we reason exactly as in Lemma 4.2; the only difference is that we bound $|n_\ell| < |n| + Bk$, which provides in the bound (4.17) an extra factor $|n|$ with respect to the bound (ii) in (4.9).

The derivative with respect to $M_{n',j'}(a',b')$ gives a sum of trees with a distinguished line ℓ carrying the propagator $\partial_{M_{n',j'}(a',b')} g_\ell$ instead of g_ℓ. As in case (iii) of Lemma 4.2 we have two contributions, one when the derivative acts on the matrix and the other (if $i_\ell = 1$) when the derivative acts on χ_{h_ℓ}; by the same arguments as in Lemma 4.2(ii) we obtain a factor of order $2^{2h} p_{j_\ell}^{-s_2-\alpha}$.

By Lemma 3.10 one has $\min\{|n'-n|, |n'+n|\} < Bk$, so that the sum over n' is finite and proportional to k. The sum over j', a', b' produces a factor proportional to $|n'|^{(D-1)+2\alpha}$ – reason as in the proof of (4.9)(iii) in Lemma 4.2. This provides an overall factor of order $|n'|^D$. If $k \geqslant B_2|n|^\beta$ (with B_2 defined in Lemma 4.3) this factor can be bounded by C^k for some constant C. If $k < B_2|n|^\beta$ then, by Lemma 4.3(i), one must have $i_\ell = -1$, hence $m_\ell = m'_\ell$, for all lines ℓ on the path $\mathcal{P}(\ell_e, \ell_0)$: then if $a' \neq b'$ necessarily the line ℓ, which the derivative is applied to, is not on such a path, and the possible values of j', a', b' are bounded proportionally to k^D. If $a' = b'$ either $\ell \notin \mathcal{P}(\ell_e, \ell_0)$ – and we can reason as before – or $\ell \in \mathcal{P}(\ell_e, \ell_0)$: in the last case we use the conservation law (3.17) of the momenta (m_ℓ, m'_ℓ), and we obtain again at most k^D terms. □

Remark. For any fixed $\sigma > 0$ the constant D in (4.17) is proportional to $\tilde{C}$, hence grows exponentially in σ. As we shall need for $\tilde{C}$ to be at worst proportional to $1/\varepsilon_0$ (in order to have convergence of the series (3.27)), this means that σ can be taken as large as $O(|\log \varepsilon_0|)$.

We are now ready to prove the first part of Proposition 1.

Proposition 1(i)–(ii). *There exist constants c_0, K_0, Q_0 and σ such that the following bounds hold for all $(\varepsilon, M) \in \mathfrak{D}(\gamma)$, $q < Q_0$ and $\eta \leqslant \eta_1 = c_0 Q_0^{-2}$:*

$$\begin{aligned} |u_{n,m}| &< K_0|\eta|q^3 \mathrm{e}^{-\sigma(|n|+|m|)}, \qquad & |L_{n,j}|_\sigma &< K_0|\eta|q^2, \\ |\partial_\varepsilon L_{n,j}|_\sigma &< K_0|n|^{1+s_2}|\eta|q^2, \qquad & |\partial_\eta L_{n,j}|_\sigma &< K_0 q^2, \end{aligned} \tag{4.18}$$

for all $(n, j) \neq (1, 1)$. Moreover the operator norm of the derivative with respect to $M_{n,j}$ is bounded as

$$\begin{aligned} \|\partial_M L\|_\sigma &:= \sup_{A\in\mathcal{B}} \frac{|\partial_M L[A]|_\sigma}{|A|_\sigma} \\ &\leqslant \sup_{n,j\in\Omega} \sup_{a,b=1,\dots,d_j} \sum_{n',j'} \sum_{a',b'=1}^{d_{j'}} \left|\partial_{M_{n',j'}(a',b')} L_{n,j}(a,b)\right| \mathrm{e}^{\sigma(|m_a-m_b|^\rho - |m_{a'}-m_{b'}|^\rho)} \\ &< K_0|\eta|q^2, \end{aligned} \tag{4.19}$$

where the space $\mathcal{B}$ is defined in the Remark after the proof of Lemma 2.6.

Proof. By definition $\mathfrak{D}(\gamma)$ is contained in all $\mathfrak{D}(\theta, \gamma)$ and in all $\tilde{\mathfrak{D}}(\theta, \gamma)$, so that we can use Lemmas 4.2 and 4.6 to bound the values of trees. First we fix an unlabelled tree θ and sum over the values of the labels: we can modify independently all the end-point labels, the scales, the type labels and the momenta m_ℓ if $i_\ell \neq -1$ (one has $m_\ell = m'_\ell$ for $i_\ell = -1$). Fixed (ε, M) and (n_ℓ, j_ℓ) there are only $d_{j_\ell} = O(p_{j_\ell}^\alpha)$ possible values for m_ℓ. This reduces the factors $p_{j_\ell}^{-s_2-\alpha}$ to $p_{j_\ell}^{-s_2}$ in the bounds (4.9) and (4.17). By summing over the type and scale labels $\{i_\ell, h_\ell\}_{\ell\in L(\theta)}$ (recall that after fixing the mode labels and ε there are only two possible values for each h_ℓ such that $\mathrm{Val}(\theta) \neq 0$), we obtain a factor 4^k, and summing over the possible end-point labels provides another factor $2^{(D+1)(2k+1)}$. Finally we bound the number of unlabelled trees of order k by $\bar{C}^k$ for a suitable constant $\bar{C}$ [23]. In (4.9) we can bound

$$\prod_{h=-1}^{\infty} 2^{hN_h(\theta)} = \exp\left(\log 2 \sum_{h=-1}^{\infty} hN_h(\theta)\right) \leqslant \exp\left(\text{const.}\, k \sum_{h=-\infty}^{\infty} h2^{-h\beta/2\tau}\right) \leqslant C^k, \tag{4.20}$$

for a suitable constant C, and an analogous bound holds for the products over the scales in (4.17).

Since (see (3.1) and (3.20))

$$u_{n,m}=\sum_{k=1}^{\infty}\eta^k\sum_{\theta\in\Theta^{(k)}_{R,n,m}}\operatorname{Val}(\theta),\tag{4.21}$$

and, by Lemma 3.5, $\Theta^{(k)}_{R,n,m}$ is empty if $k<B^{-1}|n|$ or $k<B^{-1}|m|$, we obtain the first bound in (4.18).

Using (4.17)(i), we bound the sum on $\theta\in\mathcal{R}^{(k)}_{R,h,n,j}$ exactly in the same way. The main difference is that $\mathcal{R}^{(k)}_{R,h,n,j}(a,b)$ is empty if $|m_a-m_b|>B_3^{-1}k^{1/\rho}$, by Lemma 4.3(ii). Then by Lemma 4.5, we obtain the second bound in (4.18).

As for the third bound in (4.18), we have

$$\begin{aligned}\partial_\varepsilon L_{n,j}=&-\bar\chi_1(y_{n,j})\sum_{h=-1}^{\infty}C_h(x_{n,j})\sum_{\theta\in\mathcal{R}^{(k)}_{R,n,j,h}}\partial_\varepsilon\operatorname{Val}(\theta)\\&-\bar\chi_1(y_{n,j})\sum_{h=-1}^{\infty}\big(\partial_\varepsilon C_h(x_{n,j})\big)\sum_{\theta\in\mathcal{R}^{(k)}_{R,n,j,h}}\operatorname{Val}(\theta)\\&-\big(\partial_\varepsilon\bar\chi_1(y_{n,j})\big)\sum_{h=-1}^{\infty}C_h(x_{n,j})\sum_{\theta\in\mathcal{R}^{(k)}_{R,n,j,h}}\operatorname{Val}(\theta),\end{aligned}\tag{4.22}$$

where the first summand is treated, just like in the previous cases, by using (4.17)(ii) instead of (4.17)(i). In the other summands $\operatorname{Val}(\theta)$ is bounded exactly as in the previous cases, but the derivative with respect to ε gives in the second summand an extra factor proportional to $|n|2^h p_j^{3\alpha}$ – appearing only for those values of h such that $\chi_h(x_{n,j})$ is nonzero (and for each value of ε there are only two such values so that the sum over h is finite) – and in the third summand a factor proportional to $|n|\,p_j^{s_2}$. We omit the details, which can be easily worked out by reasoning as for (4.10) and (4.12) in the proof of Lemma 4.2. Finally we bound 2^h by C^k as in the proof of Lemma 4.6.

The fourth bound in (4.18) follows trivially by noting that to any order k the derivative with respect to η of η^k produces $k\eta^{k-1}$.

Finally, one can reason in the same way about the derivative with respect to $M_{n,j}$, by using (4.17)(iii), so that (4.19) follows. □

5. Whitney extension and implicit function theorems

5.1. Extension of U and L

In this section we extend the function $L_{n,j}$, defined in $\mathfrak{D}(\gamma)$, to the larger set $\mathfrak{D}_0$. The extended function $L^E_{n,j}$ is a Whitney extension of $L_{n,j}$, see [33].

Lemma 5.1. *The following statements hold true.*

(i) *Given* $\theta \in \mathcal{R}^{(k)}_{R,h,n,j}$*, we can extend* $\mathrm{Val}(\theta)$ *to a function, called* $\mathrm{Val}^E(\theta)$*, defined and* C^1 *in* $\mathfrak{D}_0$*, and* $L_{n,j}(\eta,\varepsilon,M;q)$ *to a function* $L^E_{n,j} \equiv L^E_{n,j}(\eta,\varepsilon,M;q)$ *such that*

$$L^E_{n,j} = -\bar{\chi}_1(y_{n,j}) \sum_{h=-1}^{\infty} C_h(x_{n,j}) \sum_{k=1}^{\infty} \eta^k \sum_{\theta\in\mathcal{R}^{(k)}_{R,n,j,h}} \mathrm{Val}^E(\theta) \tag{5.1}$$

satisfies for any $(\varepsilon,M)\in\mathfrak{D}_0$ *the same bounds in* (4.18) *and* (4.19) *as* $L^E_{n,j}(\eta,\varepsilon,M;q)$ *in* $\mathfrak{D}(\gamma)$*. Furthermore* $\mathrm{Val}(\theta) = \mathrm{Val}^E(\theta)$ *for any* $(\varepsilon,M)\in\tilde{\mathfrak{D}}(\theta,2\gamma)\subset\tilde{\mathfrak{D}}(\theta,\gamma)$ *and* $\mathrm{Val}^E(\theta)=0$ *for* $(\varepsilon,M)\in\mathfrak{D}_0\setminus\tilde{\mathfrak{D}}(\theta,\gamma)$.

(ii) *In the same way, given* $\theta\in\Theta^{(k)}_{R,n,m}$*, we can extend* $\mathrm{Val}(\theta)$ *to a function* $\mathrm{Val}^E(\theta)$ *defined and* C^1 *in* $\mathfrak{D}_0$*, and* $U_{n,j}(\eta,\varepsilon,M;q)$ *to a function* $U^E_{n,j}(\eta,\varepsilon,M;q)$ *such that* $u^E_{n,m}\equiv u^E_{n,m}(\eta,\varepsilon,M;q)$*, given by*

$$u^E_{n,m} = \sum_{k=1}^{\infty}\eta^k \sum_{\theta\in\Theta^{(k)}_{R,n,m}} \mathrm{Val}^E(\theta), \tag{5.2}$$

satisfies for any $(\varepsilon,M)\in\mathfrak{D}_0$ *the same bounds in* (4.18) *as* $u_{n,m}$ *in* $\mathfrak{D}(\gamma)$.

Furthermore $\mathrm{Val}(\theta)=\mathrm{Val}^E(\theta)$ *for any* $(\varepsilon,M)\in\mathfrak{D}(\theta,2\gamma)\subset\mathfrak{D}(\theta,\gamma)$ *and* $\mathrm{Val}^E(\theta)=0$ *for* $(\varepsilon,M)\in\mathfrak{D}_0\setminus\mathfrak{D}(\theta,\gamma)$.

Proof. We prove first the statement for the case $\theta\in\mathcal{R}^{(k)}_{R,h,n,j}$. We use the C^∞ compact support function $\chi_{-1}(t):\mathbb{R}\to\mathbb{R}^+$, introduced in Definition 3.1. Recall that $\chi_{-1}(t)$ equals 0 if $|t|<\gamma$ and 1 if $|t|\geqslant 2\gamma$, and $|\partial_t\chi_{-1}(t)|\leqslant C\gamma^{-1}$, for some constant C.

Given a tree $\theta\in\mathcal{R}^{(k)}_{R,h,n,j}$, we define

$$\begin{aligned}\mathrm{Val}^E(\theta) = {} & \prod_{\substack{\ell\in L(\theta)\setminus\{\ell_e,\ell_0\}\\ i_\ell=1}} \chi_{-1}\big(|x_{n_\ell,j_\ell}||n_\ell|^\tau\big)\\ & \times \prod^{**}_{\ell_1,\ell_2\in L(\theta)} \chi_{-1}\big(|\delta_{n_{\ell_1},j_{\ell_1}}-\delta_{n_{\ell_2},j_{\ell_2}}||n_{\ell_1}-n_{\ell_2}|^{\tau_1}\big)\,\mathrm{Val}(\theta),\end{aligned} \tag{5.3}$$

where $\prod^{**}_{\ell_1,\ell_2\in L(\theta)}$ is the product on the pairs $\ell_1\prec\ell_2\in L(\theta)$ such that $\prod_{\ell\in\mathcal{P}(\ell_1,\ell_2)}\sigma(\ell)\sigma(\ell_1)=1$, $i_{\ell_j}=1,0$, $n_{\ell_1}\neq n_{\ell_2}$, and either both ℓ_1,ℓ_2 are on the path connecting e to v_0 or both of them are not on such a path. The sign $\prod_{\ell\in\mathcal{P}(\ell_1,\ell_2)}\sigma(\ell)\sigma(\ell_1)$ is such that $|n_{\ell_1}-n_{\ell_2}|\leqslant n$.

By definition $\mathrm{Val}^E(\theta)=\mathrm{Val}(\theta)$ for $(\varepsilon,M)\in\tilde{\mathfrak{D}}(\theta,2\gamma)$ as in this set the scale functions χ_{-1} in the above formula are identically equal to 1.

By definition $\mathrm{supp}(\mathrm{Val}^E(\theta))\subset\tilde{\mathfrak{D}}(\theta,\gamma)$, as the scale functions χ_{-1} in the above formula are identically equal to 0 in the complement of $\tilde{\mathfrak{D}}(\theta,\gamma)$ with respect to $\mathfrak{D}_0$.

To bound the derivatives the only fact that prevents us from simply applying (4.17)(ii)–(iii) is the presence of the extra terms due to the derivatives of the χ_{-1} functions. Each factor of the first product in (5.3), when derived, produces an extra factor proportional to $2^{h_\ell} p_{j_\ell}^{3\alpha} |n_\ell|^{\tau+1}$. Note that a summand of this kind appears only if $i_\ell = 1$ and (ε, M) is such that

$$2^{-h_\ell - 1}\gamma \leqslant x_{n_\ell, j_\ell} \leqslant \frac{2\gamma}{|n_\ell|^\tau}. \tag{5.4}$$

This implies $|n_\ell| < 2^{(h_\ell+1)/\tau}$, so that the presence of the extra factor simply produces, in (4.17)(ii), a larger constant D and a larger exponent – say 4 – instead of 2 in the factor $2^{2h' N_{h'}(\theta)}$. Each factor of the second product produces an extra factor $|n_{\ell_1} - n_{\ell_2}|^{\tau_1+1}$, which can be bounded by C^k.

Therefore the derivatives of $L^E_{n,j}$ respect the same bounds (4.18) as $L_{n,j}$ modulo a redefinition of the constants c_0, K_0. These bounds are uniform (independent of (n, j)) and can be performed also for higher order derivatives, hence $L^E_{n,j}$ is a C^1 function of (ε, M).

We proceed in the same way for $\theta \in \Theta_{R,n,m}$:

$$\mathrm{Val}^E(\theta) = \prod_{\ell \in L(\theta):\, i_\ell = 1} \chi_{-1}\big(|x_{n_\ell, m_\ell}||n_\ell|^\tau\big) \prod^{***}_{\ell_1, \ell_2 \in L(\theta)} \chi_{-1}\big(|\delta_{n_{\ell_1}, j_{\ell_1}} \delta_{n_{\ell_2}, j_{\ell_2}}||n_{\ell_1} - n_{\ell_2}|^{\tau_1}\big)\,\mathrm{Val}(\theta), \tag{5.5}$$

where now the product $\prod^{***}$ runs on the pairs of lines $\ell_1 \prec \ell_2$ such that $\prod_{\ell \in \mathcal{P}(\ell_1, \ell_2)} \sigma(\ell)\sigma(\ell_1) = 1$, $i_{\ell_j} = 1, 0$, and $n_{\ell_1} \neq n_{\ell_2}$. □

Proposition 1(iii). *L^E is differentiable in $(\varepsilon, M) \in \mathfrak{D}_0$ and satisfies the bounds*

$$\big|\partial_\varepsilon L^E_{n,j}(a,b)\big| < C_1 |n|^{1+s_2} \mathrm{e}^{-\sigma |m_a - m_b|^\rho} |\eta| q^2,$$
$$\sum_{(n',j') \in \Omega} \sum_{a',b'=1}^{d_{j'}} \big|\partial_{M_{n',j'}(a',b')} L^E_{n,j}(a,b)\big| \mathrm{e}^{|m_a - m_b|^\rho} < C_1 |\eta|, \tag{5.6}$$

where C_1 is a suitable constant.

Proof. Simply combine the proof of Lemma 5.1 with that of Proposition 1(ii). □

5.2. The extended Q equation

Going back to (2.12), we can extend it to all $\mathfrak{D}_0$ by using $U^E_{n,j}$ instead of $U_{n,j}$ for all $(n, j) \neq (1, 1)$; we obtain the equation

$$D^s q = f_{1,V}\big(u^E\big) = \sum_{\substack{n_1+n_2-n_3=1 \\ m_1+m_2-m_3=V}} u^E_{n_1,m_1} u^E_{n_2,m_2} u^E_{n_3,m_3}. \tag{5.7}$$

The leading order is obtained for $n_i = 1$ and $m_i \in \Lambda_1$ for all $i = 1, 2, 3$, namely at $\eta = 0$ we have a nonlinear algebraic equation for q,

$$D^s q = 3^D q^3, \tag{5.8}$$

with solution $q_0 = \sqrt{D^s 3^{-D}}$. We can now prove the following result.

Proposition 1(iv). *There exists η_0 such that for all $|\eta| \leqslant \eta_0$ and $(\varepsilon, M) \in \mathfrak{D}_0$, Eq.* (5.7) *has a solution $q^E(\varepsilon, M; \eta)$, which is analytic in η and C^1 in (ε, M); moreover*

$$\left|\partial_\varepsilon q^E(\varepsilon, M; \eta)\right| \leqslant K|\eta|, \qquad \sum_{(n,j)\in\Omega} \sum_{a,b=1}^{d_j} \left|\partial_{M_{n,j}(a,b)} q^E(\varepsilon, M; \eta)\right|_\sigma \leqslant K|\eta|, \tag{5.9}$$

for a suitable constant K, and $q^E = q$ for $(\varepsilon, M) \in \mathfrak{D}(2\gamma)$.

Proof. Set $Q_0 := 2q_0$. Then there exists η_1 such that u^E is analytic in η, q for $|\eta| \leqslant \eta_1$ and $|q| \leqslant Q_0$ and C^1 in (ε, M). By the implicit function theorem, there exists $\eta_0 \leqslant \eta_1$ such that for all $|\eta| \leqslant \eta_0$ there is a solution $q^E \equiv q^E(\eta, \varepsilon, M)$ of the Q equations (5.7) such that $|q^E| < 3q_0/2$. By definition of the extension u^E, Eq. (5.7) coincides (2.12) on $\mathfrak{D}(2\gamma)$. The bounds on the derivatives follow from Lemmas 4.2 and 5.1. □

We now define

$$\begin{aligned} U^E_{n,j}(\eta, \varepsilon, M) &= U^E_{n,j}\big(\eta, \varepsilon, M; q^E(\eta, \varepsilon, M)\big), \\ L^E_{n,j}(\eta, \varepsilon, M) &= L^E_{n,j}\big(\eta, \varepsilon, M; q^E(\eta, \varepsilon, M)\big). \end{aligned} \tag{5.10}$$

Proposition 1(v). *There exists a positive constant K_1 such that the matrices $L^E_{n,j}(\eta, \varepsilon, M)$ satisfy the bounds*

$$\begin{gathered} \left|L^E(\eta, \varepsilon, M)\right|_\sigma \leqslant |\eta| K_1, \qquad \left|\partial_\varepsilon L^E_{n,j}(\eta, \varepsilon, M)\right|_\sigma \leqslant |\eta| K_1 |n|^{1+s_2}, \\ \sum_{(n,j)\in\Omega} \sum_{a,b=1}^{d_j} \left|\partial_{M_{n,j}(a,b)} L^E(\eta, \varepsilon, M)\right|_\sigma \mathrm{e}^{-\sigma|m_a - m_b|^\rho} \leqslant |\eta| K_1, \end{gathered} \tag{5.11}$$

and the coefficients $U^E_{n,j}(\eta, \varepsilon, M)$ satisfy the bounds

$$\left|U^E_{n,j}(\eta, \varepsilon, M)\right| \leqslant |\eta| K_1 \mathrm{e}^{-\sigma(|n| + |p_j|^{1/2})}, \tag{5.12}$$

uniformly for $(\varepsilon, M) \in \mathfrak{D}_0$.

Proof. It follows trivially from the bounds (5.9) and from the bounds of Lemma 5.1. □

6. Proof of Proposition 2

6.1. Proof of Proposition 2(i)

Let us consider the compatibility equation (2.11) where $L_{n,j} = L^E_{n,j}(\eta, \varepsilon, M)$. One can rewrite (2.11) as

$$\bar{\chi}_1(y_{n,j}) M_{n,j} = L^E_{n,j}(\eta, \varepsilon, M) \equiv \eta \bar{\chi}_1(y_{n,j}) \tilde{L}^E(\eta, \varepsilon, M), \tag{6.1}$$

with $\tilde{L}^E(\eta, \varepsilon, M) = O(1)$, so that (6.1) has for $\eta = 0$ the trivial solution $M_{n,j} = 0$.

The bounds of Proposition 1(v) imply that the Jacobian of the application $\tilde{L}^E(\eta, \varepsilon, M) : \mathcal{B} \to \mathcal{B}$ is bounded in the operator norm ($\mathcal{B}$ is defined in the Remark before Definition 2.7). Thus there exists η_0, K_2 such that, for $|\eta| \leqslant \eta_0$ and for all $(\varepsilon, M) \in \mathfrak{D}(2\gamma)$, we can apply the implicit function theorem to (6.1) and obtain a solution $M_{n,j}(\eta, \varepsilon)$, which satisfies the bounds

$$|M_{n,j}|_\sigma \leqslant K_2|\eta|, \qquad \left|\partial_\varepsilon M_{n,j}(\eta, \varepsilon)\right|_\sigma \leqslant K_2|n|^{1+s_2}|\eta|, \qquad \left|\partial_\eta M_{n,j}(\eta, \varepsilon)\right|_\sigma \leqslant K_2, \tag{6.2}$$

for a suitable constant K_2.

Finally we fix $\varepsilon_0 \leqslant \eta_0$, $\eta = \varepsilon$ and set (with an abuse of notation) $M_{n,j}(\varepsilon) = M_{n,j}(\eta = \varepsilon, \varepsilon)$, so that, by noting that

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} M_{n,j}(\varepsilon) = \partial_\eta M_{n,j}(\eta, \varepsilon) + \partial_\varepsilon M_{n,j}(\eta, \varepsilon), \tag{6.3}$$

we deduce from (6.2) the bound (2.26).

6.2. Proof of Proposition 2(ii) – measure estimates

We now study the measure of the set (2.27). By definition this is given by the set of $\varepsilon \in \mathfrak{E}_0(\gamma)$ such that the further Diophantine conditions

$$x_{n,j}(\varepsilon) := \left\| \left(\delta_{n,j} I + p_j^{-s} \bar{\chi}_1(y_{n,j}) M_{n,j}(\varepsilon)\right)^{-1} \right\|^{-1} \geqslant \frac{2\gamma}{|n|^\tau} \tag{6.4}$$

are satisfied for all $(n, j) \in \Omega$ such that $(n, j) \neq (1, 1)$. Recall that $(n, j) \in \Omega$ implies $n > 0$. By Lemma 2.4(iii) one has

$$x_{n,j}(\varepsilon) \geqslant \min_i \left|\lambda^{(i)}\left(\delta_{n,j} I + p_j^{-s} \bar{\chi}_1(y_{n,j}) M_{n,j}\right)\right| = \min_i \left|\delta_{n,j} + p_j^{-s+\alpha} \nu_{n,j}^{(i)}(\varepsilon)\right|, \tag{6.5}$$

since the matrices are symmetric. Recall that $p_j^\alpha \nu_{n,j}^{(i)}(\varepsilon)$ are the eigenvalues of $\bar{\chi}_1(y_{n,j}) M_{n,j}^{(i)}(\varepsilon)$ and that $d_j \leqslant C_1 p_j^\alpha$ (cf. Definition 2.5).

Then we impose the conditions

$$\left|\delta_{n,j} + p_j^{-s+\alpha} \nu_{n,j}^{(i)}(\varepsilon)\right| \geqslant \frac{2\gamma}{n^\tau} \quad \forall (n, j) \in \Omega \setminus \left\{(1, 1)\right\},\ i = 1, \ldots, d_j, \tag{6.6}$$

and recall that $M_{n,j} = 0$ (i.e. $\nu_{n,j}^{(i)} = 0$) if $(n, j) \notin \Omega$, so that for $(n, j) \notin \Omega$ the Diophantine conditions (6.6) are surely verified, by (2.1).

Call $\mathfrak{A}$ the set of values of $\varepsilon \in \mathfrak{E}_0(\gamma)$ which verify (6.6). We estimate the measure of the subset of $\mathfrak{E}_0(\gamma)$ complementary to $\mathfrak{A}$, i.e. the set defined as union of the sets

$$\mathfrak{I}_{n,j,i} := \left\{\varepsilon \in \mathfrak{E}_0(\gamma)\colon \left|\delta_{n,j} + p_j^{-s+\alpha} \nu_{n,j}^{(i)}(\varepsilon)\right| \leqslant \frac{2\gamma}{n^\tau}\right\}, \tag{6.7}$$

for $(n, j) \in \Omega$ and $i = 1, \ldots, d_j$.

Given n, the condition $(n,j)\in\Omega$ implies that p_j can assume at most $\varepsilon_0 n+1$ different values – cf. (2.16). On a $(D-1)$-dimensional sphere of radius R there are at most $O(R^{D-1})$ integer points, hence the number of values which j can assume is bounded proportionally to $n^{D-1}(1+\varepsilon_0 n)$. Finally i assumes $d_j\leqslant C_1 p_j^{\alpha}$ values.

Since $\mu\in\mathfrak{M}$ we have, for $n\leqslant(\gamma_0/4\varepsilon_0)^{1/(\tau_0+1)}$,

$$\left|\delta_{n,j}+p_j^{-s+\alpha}\nu_{n,j}^{(i)}(\varepsilon)\right|\geqslant\left|(D+\mu)n-p_j-\mu\right|-2\varepsilon_0 n\geqslant\frac{\gamma_0}{2n^{\tau_0}}, \tag{6.8}$$

so that we have to discard the sets $\mathfrak{I}_{n,j,i}$ only for $n\geqslant(\gamma_0/4\varepsilon_0)^{1/(\tau_0+1)}$.

Let us now recall that for a symmetric matrix $M(\varepsilon)$ depending smoothly on a parameter ε, the eigenvalues are C^1 in ε [24]. Then the measure of each $\mathfrak{I}_{n,j,i}$ can be bounded from above by

$$\frac{4\gamma}{n^{\tau}}\sup_{\varepsilon\in\mathfrak{E}_0(\gamma)}\left|\left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(\delta_{n,j}+p_j^{-s+\alpha}\nu_{n,j}^{(i)}(\varepsilon)\right)\right)^{-1}\right|, \tag{6.9}$$

where one has

$$\left|\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(\delta_{n,j}+p_j^{-s+\alpha}\nu_{n,j}^{(i)}(\varepsilon)\right)\right|\geqslant\frac{n}{2}. \tag{6.10}$$

This can be obtained as follows. Proving (6.10) requires to find lower bounds for

$$\left|\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(-\omega n+p_j+\mu+p_j^{-s}\bar{\chi}(y_{n,j})\lambda_{n,j}^{(i)}(\varepsilon)\right)\right|,$$

where $\lambda_{n,j}^{(i)}(\varepsilon)$ are the eigenvalues of $M_{n,j}(\varepsilon)$ (i.e. $\bar{\chi}_1(y_{n,j})\lambda_{n,j}^{(i)}(\varepsilon)=p_j^{\alpha}\nu_{n,j}^{(i)}(\varepsilon)$). The eigenvalues $\lambda_{n,j}(\varepsilon)^{(i)}$ are C^1 in ε, so that, by Lidskii's lemma [24], one has

$$\left|\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\lambda_{n,j}^{(i)}(\varepsilon)\right|\leqslant d_j\left|\frac{\mathrm{d}}{\mathrm{d}\varepsilon}M_{n,j}\right|_{\infty}\leqslant C_1K_2\left(1+\varepsilon_0 n^{1+s_2}\right)n^{\alpha}, \tag{6.11}$$

where we have used (6.3) and (6.2). Since $s_2+\alpha\leqslant s-\alpha$, we obtain

$$\left|\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\left(-\omega n+p_j+\mu+p_j^{-s}\bar{\chi}(y_{n,j})\lambda_{n,j}^{(i)}(\varepsilon)\right)\right|\geqslant\frac{n}{2},$$

which implies (6.10).

Recall that p_j is bounded proportionally to n. Then for fixed n we have to sum over const.$(1+\varepsilon_0 n)n^{D-1}$ values of j and over $d_j\leqslant C_1p_j^{\alpha}\leqslant$ const. $n^{\alpha}\leqslant$ const. n.

Therefore we have

$$\begin{aligned}\sum_{(n,j)\in\Omega}\sum_{i=1}^{d_j}\operatorname{meas}(\mathfrak{I}_{n,j,i})&\leqslant\text{const.}\sum_{n\geqslant(\gamma/4\varepsilon_0)^{-1/(\tau_0+1)}}\gamma|n|^{D}\left(\frac{1}{|n|^{\tau+1}}+\frac{\varepsilon_0}{|n|^{\tau}}\right)\\&\leqslant\text{const.}\left(\varepsilon_0^{(\tau-D)/(\tau_0+1)}+\varepsilon_0^{1+(\tau-D-1)/(\tau_0+1)}\right),\end{aligned} \tag{6.12}$$

provided $\tau > D+1$. Therefore the measure is small compared to that of $\mathfrak{E}_0(\gamma)$ – which is of order ε_0– if $\tau > \max\{\tau_0 + D + 1, D+1\} = \tau_0 + D + 1$.

7. Generalisations and proof of Theorem 1

7.1. Eq. (1.4): Proof of Theorem 1 in $D > 2$

In order to consider Eq. (1.4) we only need to make few generalisations. By our assumptions

$$f(x,u,\bar{u}) = g(x,\bar{u}) + \partial_{\bar{u}} H(x,u,\bar{u}),$$

with H real-valued. For simplicity we discuss explicitly only the case with odd p in (1.3)—hence the functions $a_{p_1,p_2}(x)$ are even and real. Considering also even p should require considering an expansion in $\sqrt{\varepsilon}$: this would not introduce any technical difficulties, but on the other hand would require a deeper change in notations.

We modify the tree expansion, analogously to what done in [21]. The change of variables (1.8) transforms each monomial in (1.3) into a monomial $\varepsilon^{(p_1+p_2-1)/2} a_{p_1,p_2}(x)\, u^{p_1} \bar{u}^{p_2}$; we can take into account the contributions arising from $g(x,\bar{u})$, by considering the corresponding Taylor expansion and putting $p_1 = 0$ and $p_2 \geqslant 3$. All the other contributions are such $p_1 a_{p_1,p_2} = (p_2 + 1) a_{p_2+1,p_1-1}$ (by the reality of H and of the functions a_{p_1,p_2}).

Each new monomial produces internal nodes of order $k_v = (p_{v,1} + p_{v,2} - 1)/2 \in \mathbb{N}$, such that $k_v \geqslant 2$, with $p_{v,1} + p_{v,2}$ entering lines among which $p_{v,1}$ have sign $\sigma = 1$ and $p_{v,2}$ have sign $\sigma = -1$; note that the case previously discussed corresponds to $(p_{v,1}, p_{v,2}) = (2,1)$. Hence, with the notations of Section 3.3, we can write $s_v = p_{v,1} + p_{v,2}$, with s_v odd.

Each internal node v has labels $k_v, p_{v,1}, p_{v,2}, m_v$, with the mode label $m_v \in \mathbb{Z}^D$. The node factor associated with v is $a_{p_{v,1},p_{v,2},m_v}$, namely the Fourier coefficient with index m_v in the Fourier expansion of the function $a_{p_{v,1},p_{v,2}}$; by the analyticity assumption on the nonlinearity the Fourier coefficients decay exponentially in m, that is

$$|a_{p_{v,1},p_{v,2},m_v}| \leqslant A_1 \mathrm{e}^{-A_2|m|}, \tag{7.1}$$

for suitable constants A_1 and A_2.

The conservation laws (3.16) and (3.17) have to be suitably changed. We can still write that n_ℓ is given by the right-hand side, the only difference being that $L(v)$ contain s_v lines (and each line $\ell \in L(v)$ has its own sign $\sigma(\ell)$). On the contrary (3.17) for m'_ℓ has to be changed in a more relevant way: indeed one has

$$m'_\ell = m_v + \sum_{\ell' \in L(v)} \sigma(\ell') m_{\ell'}, \tag{7.2}$$

with $L(v)$ defined as before.

The order of any tree θ is still defined as in (3.19), and, more generally, all the other labels are defined exactly as in Section 3.3.

The first differences appear when one tries to bound the momenta of the lines in terms of the order of the tree. In fact one has

$$\left|E(\theta)\right| \leqslant 1 + \sum_{v \in V_0(\theta)} (s_v - 1), \tag{7.3}$$

which reduces to the formula given in the proof of Lemma 3.5 only for $s_v \leqslant 3$. One has $s_v = 2k_v + 1$, so that

$$\sum_{v \in V_0(\theta)} (s_v - 1) = 2 \sum_{v \in V_0(\theta)} k_v = 2k, \tag{7.4}$$

and one can still bound $|n_\ell| \leqslant Bk$ for any tree $\theta \in \Theta^{(k)}$ and any line $\ell \in L(\theta)$.

The conservation law (7.1) gives, for any line $\ell \in L(\theta)$,

$$\max\left\{|m_\ell|, \left|m'_\ell\right|\right\} \leqslant Bk + \sum_{v \in V_0(\theta)} |m_v|, \tag{7.5}$$

for some constant B. The bound in (7.5) is obtained by reasoning as in the proof of Lemma 3.5; the last sum is due to the mode labels of the internal nodes. Thus the bound on n_ℓ in Lemma 3.5 still holds, while the bounds on m_ℓ, m'_ℓ have to be replaced with (7.5). The same observation applies to Lemma 3.10.

Also Lemma 3.14 still holds. The proof proceeds as follows. The tree θ_1 which one associates with each $\theta \in \mathcal{R}_{R,n,j,h}(a,b)$ is the tree in $\mathcal{R}_{R,n,j,h}(b,a)$ defined as follows.

1. As in the proof of Lemma 3.14.
2. As in the proof of Lemma 3.14.
3. Let $\bar{v}$ be a node along the path $\mathcal{P} = \mathcal{P}(\ell_e, \ell_0)$ and let $\ell_1, \ldots, \ell_s$, with $s = s_{\bar{v}}$ be the lines entering $\bar{v}$; suppose that ℓ_1 is the line belonging to the path $\mathcal{P} \cup \{\ell_e\}$. If $\sigma(\ell_1) = 1$ we change $m_{\bar{v}} \to -m_{\bar{v}}$ and we change all the signs of the other lines, i.e. $\sigma(\ell_i) \to -\sigma(\ell_i)$ for $i = 2, \ldots, s$, whereas if $\sigma(\ell) = -1$ we do not change anything.
4. As in the proof of Lemma 3.14.

Then one can easily check that the form (1.4) of the nonlinearity implies that the tree θ_1 is well defined (as an element of $\mathcal{R}_{R,n,j,h}(b,a)$) and has the same value as θ.

Remarks. (1) Note that item 3 above reduces to item 3 of Lemma 3.14 if $s_v = 3$ for each internal node v.

(2) If the node $\bar{v}$ has $p_{\bar{v},1} = 0$ (i.e. the monomial associated to it arises from the function $g(x, \bar{u})$) then the operation in item 3 is empty.

Therefore we can conclude that the counterterms are still symmetric.

The analysis of Section 4 can be performed almost unchanged. Here we confine ourselves to show the few changes that one has to take into account.

The first relevant difference appears in Lemma 4.1. Because of the presence of the mode labels of the internal nodes the bound (4.5) on $N_h(\theta)$ does not hold anymore, and it has to be replaced with

$$N_h(\theta) \leqslant \max\left\{0, c\left(k(\theta) + \sum_{v \in V_0(\theta)} |m_v|\right) 2^{(2-h)\beta/\tau} - 1\right\}, \tag{7.6}$$

for a suitable constant c. The proof of (7.6) proceeds as the proof of Lemma 4.1 in Section 4. We use that in (4.7) for $m=1$ one has

$$k(\theta)-k(\theta_1)+\sum_{v\in V_0(\theta)}|m_v|-\sum_{v\in V_0(\theta_1)}|m_v|=k_T+\sum_{v\in V_0(T)}|m_v|, \tag{7.7}$$

and, except for item 2.2.2.1, we simply bound the right-hand side of (7.7) with k_T. The only item which requires a different argument is item 2.2.2.1, where instead of the bound $|m_{\bar\ell}-m_{\ell_0}|\leqslant \sum_i |m_i|$ we have, by (7.5),

$$|m_{\bar\ell}-m_{\ell_0}|\leqslant \sum_{v\in\mathcal{P}(\bar\ell,\ell_0)}|m_v|+\sum_i |m_i|\leqslant Bk_T+\sum_{v\in V_0(T)}|m_v|\leqslant \max\{B,1\}\bigg(k_T+\sum_{v\in V_0(T)}|m_v|\bigg),$$

where $v\in\mathcal{P}(\bar\ell,\ell_0)$ means that the node v is along the path $\mathcal{P}(\bar\ell,\ell_0)$ (i.e. $\ell_v\in\mathcal{P}(\bar\ell,\ell_0)\cup\ell_0$) and the sum over i is over all sub-trees which have the root lines entering one of such nodes.

Remark. Note that if the coefficients $a_{p_1,p_2}(x)$ in (1.3) are just constants (i.e. do not depend on x), then $m_v\equiv 0$ and (7.6) reduces to (4.5).

Moreover in (4.9) we have a further product

$$\prod_{v\in V_0(\theta)} A_1\mathrm{e}^{-A_2|m_v|}, \tag{7.8}$$

while the product of the factors $2^{hN_h(\theta)}$ can be written as

$$\Bigg(\prod_{h=-1}^{h_0} 2^{hN_h(\theta)}\Bigg)\Bigg(\prod_{h=h_0+1}^{\infty} 2^{hN_h(\theta)}\Bigg)\leqslant 2^{h_0k}\prod_{h=h_0+1}^{\infty} 2^{hN_h(\theta)}, \tag{7.9}$$

with h_0 to be fixed, where the last product, besides a contribution which can be bounded as in (4.20), gives a further contribution

$$\prod_{h=h_0}^{\infty}\prod_{v\in V_0(\theta)} 2^{ch|m_v|2^{(2-h)\beta/\tau}}\leqslant \prod_{v\in V_0(\theta)}\exp\Bigg(\text{const.}|m_v|\sum_{h=h_0}^{\infty} h2^{-h\beta/\tau}\Bigg), \tag{7.10}$$

so that we can use part of the exponential factors in (7.8) to compensate the exponential factors in (7.10), provided h_0 is large enough (depending on τ).

Another consequence of (7.2) is in Lemma 4.3: item (ii) has to be replaced with

$$|m_a-m_b|\leqslant \text{const.}\,k^{1/\rho}+\sum_{v\in V_0(\theta)}|m_v|, \tag{7.11}$$

because each internal node v contributes a momentum m_v to the momenta of the lines following v. Up to this observation, the proof of (7.11) proceeds as in the proof of Lemma 4.3.

Therefore also the bound (4.14) of Lemma 4.4 has to changed into (7.6), for all $h \leqslant \bar{h}$. The proof proceeds as that of Lemma 4.4 in Section 4, with the changes outlined above when dealing with case 2.2.2.1.

The property (7.11) reveals itself also in the proof of Lemma 4.6. More precisely, in order to extract a factor $\mathrm{e}^{-\sigma|m_a-m_b|^\rho}$, we use that (7.11) implies (recall that $\rho < 1$)

$$|m_a - m_b|^\rho \leqslant C\left(k + \sum_{v\in V_0(\theta)} |m_v|\right), \tag{7.12}$$

for a suitable ρ-dependent constant C, so that we can write, for some other constant $\tilde{C}$,

$$\mathrm{e}^{\sigma|m_a-m_b|^\rho} \leqslant \tilde{C}^k \prod_{v\in V_0(\theta)} \mathrm{e}^{\sigma|m_v|}, \tag{7.13}$$

where σ has to be chosen so small (e.g. $|\sigma| < A_2/4$, with A_2 given in (7.1)) that the last product in (7.13) can be controlled by part of the exponentially decaying factors $\mathrm{e}^{-A_2|m_v|}$ associated with the internal nodes. This means, in particular, that σ cannot be arbitrarily large when ε_0 becomes small (cf. the Remark after the proof of Lemma 4.6).

As in (4.9) also in (4.17) there are the further factors (7.8), which can be dealt with exactly as in the previous case.

Besides the issues discussed above, there is no other substantial change with respect to the analysis of Sections 4 to 6.

7.2. Eq. (1.1) in dimension 2: Proof of Theorem 1 in $D = 2$

We can consider more general nonlinearities in the case $D = 2$, that is of the form (1.3) without the simplifying assumption (1.4). Indeed in such a case the counterterms are 2×2 matrices (cf. Lemma 2.2), so that we can bound $x_{n,j}$ by the absolute value of the determinant of $\delta_{n,j} I + \bar{\chi}_1(y_{n,j}) M_{n,j} p_j^{-s}$, which is a C^2 function of ε (we have proved only C^1 but it should be obvious that we can bound as many derivatives of $L^E_{n,j}$ as we need to, possibly by decreasing the domain of convergence of the functions involved).

Set for notational simplicity $\bar{M}_{n,j} = \bar{\chi}_1(y_{n,j}) M_{n,j}$. Let us evaluate the measure of the Cantor set

$$\mathfrak{E}_1 = \left\{\varepsilon \in \mathfrak{E}_0(\gamma)\colon \left|\delta^2_{n,j} + p_j^{-s} \operatorname{tr} \bar{M}_{n,j} \delta_{n,j} + p_j^{-2s} \det \bar{M}_{n,j}\right| \geqslant 2\gamma |n|^{-\tau}\right\}, \tag{7.14}$$

following the scheme of Section 6. Here we are using explicitly that for $D = 2$ one has

$$\det\left(\delta_{n,j} I + p_j^{-s} \bar{M}_{n,j}\right) = \delta^2_{n,j} + p_j^{-s} \operatorname{tr} \bar{M}_{n,j} \delta_{n,j} + p_j^{-2s} \det \bar{M}_{n,j}, \tag{7.15}$$

because $M_{n,j}$ is a 2×2 matrix.

We estimate the measure of the complement of $\mathfrak{E}_1$ with respect to $\mathfrak{E}_0(\gamma)$, which is the union of the sets

$$\mathfrak{I}_{n,j} := \left\{\varepsilon \in \mathfrak{E}_0(\gamma)\colon \left|\delta^2_{n,j} + p_j^{-s} a_{n,j} \delta + p_j^{-2s} b_{n,j}\right| \leqslant \frac{2\gamma}{|n|^\tau}\right\}, \tag{7.16}$$

where $a_{n,j} = \operatorname{tr} \bar{M}_{n,j}$ and $b_{n,j} = \det \bar{M}_{n,j}$.

Given n the condition $(n, j) \in \Omega$ implies that p_j can assume at most $\varepsilon_0 n + 1$ different values. On a one-dimensional sphere of radius R there are less than R integer points, so the number of values j can assume is bounded proportionally to $n(\varepsilon_0 n + 1)$.

Since $\mu \in \mathfrak{M}$ we have for $|n| \leqslant n_0(\gamma_0^2/\varepsilon_0)^{1/2\tau_0}$, with some constant n_0,

$$\left|\delta_{n,j}\left(\delta_{n,j} + p_j^{-s} a_{n,j}\right) + p_j^{-2s} b_{n,j}\right| \geqslant \left(\left|(D+\mu)n - p_j - \mu\right| - 2\varepsilon_0|n|\right)^2 - \text{const.}\, \varepsilon_0 \geqslant \frac{\gamma_0^2}{2|n|^{2\tau_0}}, \tag{7.17}$$

so that

$$\left|\delta_{n,j}\left(\delta_{n,j} + p_j^{-s} a_{n,j}\right) + p_j^{-2s} b_{n,j}\right| \geqslant \frac{\gamma}{|n|^{\tau}}, \tag{7.18}$$

provided $\gamma < \gamma_0^2/2$ and $\tau > 2\tau_0$.

The measure of each $\mathfrak{I}_{n,j}$ can be bounded from above by

$$\frac{2\gamma}{|n|^{\tau}} \sup_{\varepsilon \in \mathfrak{E}_0(\gamma)} \left| \left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon} \left(\delta_{n,j}^2 + p_j^{-s} a_{n,j}\delta_{n,j} + p_j^{-2s} b_{n,j}\right)\right)^{-1} \right|. \tag{7.19}$$

In order to control the derivative we restrict ε to the Cantor set

$$\mathfrak{E}_2 = \left\{ \varepsilon \in \mathfrak{E}_0(\gamma) \colon \left|\delta_{n,j}\left(2n + p_j^{-s} a'_{n,j}(\varepsilon)\right) + n p_j^{-s} a_{n,j} + p_j^{-2s} b'_{n,j}(\varepsilon)\right| \geqslant \frac{\gamma}{|n|^{\tau_2}} \right.$$
$$\left. \text{for all } (n, j) \in \Omega \right\}, \tag{7.20}$$

with $a'_{n,j}(\varepsilon) = \mathrm{d}a_{n,j}(\varepsilon)/\mathrm{d}\varepsilon$ and $b'_{n,j}(\varepsilon) = \mathrm{d}b_{n,j}(\varepsilon)/\mathrm{d}\varepsilon$. On this set we have (recall that p_j is bounded proportionally to $|n|$)

$$\sum_{(n,j)\in\Omega} \operatorname{meas}(\mathfrak{I}_{n,j}) \leqslant \text{const.} \sum_{n \geqslant n_0(\gamma/\varepsilon_0)^{1/2\tau_0}} \frac{n + \varepsilon_0 n^2}{|n|^{\tau - \tau_2}} \leqslant \text{const.}\, \varepsilon_0^{(\tau - \tau_2 - 2)/2\tau_0}, \tag{7.21}$$

provided $\tau > \tau_2 + 3$. Hence $\operatorname{meas}(\mathfrak{I}_{n,j})$ is small with respect to ε_0 provided $\tau > 2\tau_0 + \tau_2 + 2$.

Finally let us study the measure of $\mathfrak{E}_2$. The bounds (6.2) – and their proofs to deal with the second derivatives – imply

$$|a_{n,j}|, |b_{n,j}| \leqslant C\varepsilon_0, \qquad |a'_{n,j}|, |b'_{n,j}| \leqslant C\left(1 + \varepsilon_0|n|^{1+s_2}\right),$$
$$\left|a''_{n,j}\right|, \left|b''_{n,j}\right| \leqslant C\left(1 + \varepsilon_0|n|^{2+2s_2}\right), \tag{7.22}$$

for some constant C.

Let us call $\mathfrak{I}^1_{n,j}$ the complement of $\mathfrak{E}_2$ with respect to $\mathfrak{E}_0(\gamma)$ at fixed $(n, j) \in \Omega$. As in estimating the set $\mathfrak{I}_{n,j}$ in (7.16) we can restrict the analysis to the values of n such that

$n > n_1(\gamma_0/\varepsilon_0)^{1/2\tau_0}$, possibly with a constant n_1 different from n_0. Then we need a lower bound on the derivative, which gives

$$\left| |n| \left(2n + a'_{n,j} p_j^{-s}\right) + \delta_{n,j} a''_{n,j} p_j^{-s} + b''_{n,j} p_j^{-2s} \right| \geqslant \frac{n^2}{2} \tag{7.23}$$

(recall that $\delta_{n,j} < 1/2$). Hence we get

$$\sum_{(n,j)\in\Omega} \operatorname{meas}\left(\mathfrak{I}^1_{n,j}\right) \leqslant \text{const.} \sum_{n \geqslant n_1(\gamma/\varepsilon_0)^{1/2\tau_0}} \frac{(n+\varepsilon_0 n^2)}{|n|^{\tau_2 - 2}} \leqslant \text{const.}\, \varepsilon_0^{(\tau_2-4)/2\tau_0}, \tag{7.24}$$

provided $\tau_2 > 5$. Again the measure is small with respect to ε_0 provided $\tau_2 > 2\tau_0 + 4$. For $\tau_0 > 1$ this gives $\tau_2 > 6$ and therefore $\tau > 2\tau_0 + \tau_2 + 2 > 10$.

Remark. The argument given above applies only when $D = 2$, because only in such a case the matrices $M_{n,j}$ are of finite n-independent size (cf. Lemma 2.2). A generalisation to the case $D > 2$ should require some further work.

8. Proofs of Theorems 2–4

Let us now consider (1.9) with $\mu = 0$, under the conditions (1.3) if $D = 2$ and both (1.3) and (1.4) if $D \geqslant 3$. Note that for $\mu = 0$ one has $\omega = D - \varepsilon$.

The Q subspace is infinite-dimensional, namely (1.13) is replaced by

$$Q := \left\{(n,m) \in \mathbb{N} \times \mathbb{Z}^D \colon\ Dn = |m|^2\right\}, \tag{8.1}$$

so that Q contains as many elements as the set of $m \in \mathbb{Z}^D$ such that $|m|^2/D \in \mathbb{N}$.

As in [21] our strategy will be as follows: first, we shall find a finite-dimensional solution of the bifurcation equation, hence we shall prove that it is nondegenerate in Q and eventually we shall solve both the P and Q equations iteratively.

A further difficulty comes from the separation of the resonant sites. Indeed conditions (2.1) and (2.2) are fulfilled now only for those (n, p) such that $Dn \neq p$. This implies that Lemma 2.1 does not hold: given $p_i^{s_0}|\omega n_i - p_i| \leqslant \gamma/2$ for $i = 1, 2$ it is possible that $D(n_1 - n_2) = p_1 - p_2$ and in such a case we have at most $|p_1 - p_2| \leqslant \gamma/\varepsilon p_2^{s_0}$, which in general provides no separation at all. Hence we cannot use anymore the second Melnikov conditions.

We then replace Lemma 2.2 a by more general result (cf. Lemma 8.4 below), due to Bourgain; consequently we deal with a more complicated renormalised P equations.

8.1. *The Q equations*

In [21] we considered the one-dimensional case and used the integrable cubic term in order to prove the existence of finite-dimensional subsets of Q such that there exists a solution of the bifurcation equation with support on those sets.

In order to extend this result to $D \geqslant 2$ we start by considering (1.9) projected on the Q subspace. We set $u_{n,m} = q_m$ if $(n,m) \in Q$, so that the Q equations become

$$|m|^{2(1+s)} D^{-1} q_m = \sum_{\substack{m_1+m_2-m_3=m \\ n_1+n_2-n_3=|m|^2/D}} u_{n_1,m_1} u_{n_2,m_2} u_{n_3,m_3}.$$

Setting $q_m = a_m + Q_m$, with $Q_m = O(\eta)$, the leading order provides a relatively simple equation, as shown by Bourgain in [9]:

$$|m|^{2(1+s)} D^{-1} a_m = \sum_{\substack{m_1,m_2,m_3 \\ m_1+m_2-m_3=m \\ \langle m_1-m_3, m_2-m_3\rangle=0}} a_{m_1} a_{m_2} a_{m_3}, \tag{8.2}$$

which will be called the *bifurcation equation*. One can easily find finite-dimensional sets $\mathcal{M}$ such that

(i) if $m \in \mathcal{M}$ then $S_i(m) \in \mathcal{M}$ $\forall i = 1, \dots, D$ (S_i is defined in (1.12)),
(ii) if $m_1, m_2, m_3 \in \mathcal{M}$ and $\langle m_1 - m_3, m_2 - m_3\rangle = 0$, then $m_1 + m_2 - m_3 \in \mathcal{M}$.

Remarks. (1) Condition (i) implies that $\mathcal{M}$ is completed described by its intersection $\mathcal{M}_+$ with $\mathbb{Z}_+^D$.

(2) Clearly (8.2) admits a solution with support on sets respecting (i) and (ii) above. An example is as follows. For all r the set $\mathcal{M}_+(r) := \{m \in \mathbb{Z}_+^D\colon |m| = r\}$ is a finite-dimensional set on which (8.2) is closed.

(3) We look for a solution of (8.2) which satisfies the Dirichlet boundary conditions. Hence we study (8.2) as an equation for a_m with $m \in \mathcal{M}_+$.

(4) Note that the bifurcation equation (8.2) is only apparently equal to that considered by Bourgain, because of the Dirichlet boundary conditions. The latter introduce a lot of symmetries – and hence of degeneracies – which make the analysis much more complicated: in particular checking the invertibility of the forthcoming matrix J requires a lot of work.

Finding nontrivial solutions of (1.9) by starting from solutions of the bifurcation equation like those of the example may however be complicated, so we shall prove the existence of solutions under the following, more restrictive, conditions.

Lemma 8.1. *There exist finite sets $\mathcal{M}_+ \subset \mathbb{Z}_+^D$ such that $|m|^2$ is divided by D for all $m \in \mathcal{M}_+$ and* (8.2) *is equivalent to*

$$\begin{cases} |m|^{2(1+s)} D^{-1} - 2^{D+1} A^2 - \big(3^D - 2^{D+1}\big) a_m^2 = 0, & a_m \in \mathcal{M}_+, \\ a_m = 0, & a_m \in \mathbb{Z}_+^D \setminus \mathcal{M}_+, \end{cases} \tag{8.3}$$

with $A^2 := \sum_{m \in \mathcal{M}_+} a_m^2$.

Proof. The idea is to choose the $m \in \mathcal{M}_+$ so that $|m|^2 \in D\mathbb{N}$, (8.3) is equivalent to (8.2) and has a nontrivial solution. We choose $\mathcal{M}_+$ so that the following conditions are fulfilled:

(a) setting $N := |\mathcal{M}_+|$ and $\min_{m\in\mathcal{M}_+} |m| = |m_1|$ (this only implies a reordering of the elements of $\mathcal{M}_+$), we impose

$$2^{D+1} \sum_{m\in\mathcal{M}_+\setminus\{m_1\}} |m|^{2+2s} \leqslant \big(3^D + 2^{D+1}(N-2)\big)|m_1|^{2+2s}; \tag{8.4}$$

(b) the identity $\langle m_1 - m_3, m_2 - m_3\rangle = 0$ can be verified only if either $m_1 - m_3 = 0$ or $m_2 - m_3 = 0$ or $|(m_1)_i| = |(m_2)_i| = |(m_3)_i|$ for all $i = 1,\dots,D$ ($(m_j)_i$ is the ith component of the vector m_j).

An easy calculation shows that under conditions (b) Eq. (8.2) assumes the form

$$a_m\big(|m|^{2(1+s)}D^{-1} - 2^{D+1}A^2 - \big(3^D - 2^{D+1}\big)a_m^2\big) = 0, \tag{8.5}$$

and hence is equivalent to (8.3). Now, in order to find a nontrivial solution to (8.5) we must impose

$$|m_1|^{2+2s} = \min_{m\in\mathcal{M}_+} |m|^{2+2s} \geqslant 2^{D+1}A^2 D, \tag{8.6}$$

with A determined by

$$D\big(2^{D+1}(N-1) + 3^D\big)A^2 = M, \quad N = |\mathcal{M}_+|,\ M := \sum_{m\in\mathcal{M}_+} |m|^{2+2s}. \tag{8.7}$$

As in the one-dimensional case [21], if we fix N then (8.6) is equivalent to condition (a), i.e. (8.4), which is an upper bound on the moduli of the remaining $m_i \in \mathcal{M}_+ \setminus \{m_1\}$. Then there exist sets of the type described above at least for $N = 1$.

To complete the proof (for all $N \in \mathbb{N}$) we have still to show that sets $\mathcal{M}_+$ verifying the conditions (a) and (b) exist. The existence of sets with $N = 1$ is trivial, an iterative method of construction for any N is then provided in Appendix A.3. □

Remark. The compatibility condition (8.4) requires for the harmonics of the periodic solution to be large enough, and not too spaced from each other. Therefore, once we have proved that the solutions of the bifurcation equation can be continued for $\varepsilon \neq 0$, we can interpret the corresponding periodic solutions as perturbed wave packets. The same result was found in $D = 1$ in [21].

We have proved that the bifurcation equation admits a nontrivial solution

$$q^{(0)}(x,t) = \sum_{m\in\mathcal{M}_+} q_m^{(0)} \mathrm{e}^{i\frac{|m|^2}{D}t}(2i)^D \prod_{i=1}^{D} \sin(m_i x_i), \tag{8.8}$$

with $q_m^{(0)} = a_m$ for $m \in \mathbb{Z}_+^D$ and extended to all $\mathbb{Z}^D$ by imposing the Dirichlet boundary conditions.

We can set $q_m = q_m^{(0)} + Q_m$ for all $m \in \mathbb{Z}^D$ and split the Q equations in a bifurcation equation (8.3) and a recursive linear equation for Q_m:

$$|m|^{2+2s} D^{-1} Q_m - 2 \sum_{\substack{m_1,m_2,m_3 \\ m_1+m_2-m_3=m \\ \langle m_1-m_3, m_2-m_3\rangle=0}} Q_{m_1} q_{m_2}^{(0)} q_{m_3}^{(0)} - \sum_{\substack{m_1,m_2,m_3 \\ m_1+m_2-m_3=m \\ \langle m_1-m_3, m_2-m_3\rangle=0}} q_{m_1}^{(0)} q_{m_2}^{(0)} Q_{m_3}$$

$$= \sum^*_{\substack{m_1+m_2-m_3=m \\ n_1+n_2-n_3=|m|^2/D}} u_{n_1,m_1} u_{n_2,m_2} u_{n_3,m_3}, \tag{8.9}$$

where for all $(n,m) \in Q$ one has $u_{n,m} \equiv q_m$ and $*$ in the last sum means that the sum is restricted to the triples (n_i, m_i) such that if at least two of u_{n_i,m_i} are $q_{m_i}^{(0)}$ then the label (n,m) of the third one must not belong to Q.

By using once more the Dirichlet boundary conditions, we can see (8.9) as an equation for the coefficients Q_m with $m \in \mathbb{Z}_+^D$. In particular the left-hand side yields an infinite-dimensional matrix J acting on $\mathbb{Z}_+^D$. We need to invert this matrix.

Lemma 8.2. *For all D and for all choices of $\mathcal{M}_+$ as in Lemma* 8.1, *one has that J is a block-diagonal matrix, with finite-dimensional blocks, whose sizes are bounded from above by some constant M_1 depending only on D and $\mathcal{M}_+$.*

The result above is trivial for $D = 2$ and requires some work for $D > 2$, see Appendix A.4. In any case it is not enough to ensure that the matrix J is invertible.

Lemma 8.3. *For $N = 1$ and any $D \geqslant 2$ and for $N > 4$ and $D = 2$ there exist sets $\mathcal{M}_+$ such that the matrix J is invertible outside a discrete set of values of s.*

Proof. We can write $J = \text{diag}\{|m|^{2+2s}/D - 2^{D+1}A^2\} + Y$, where A is defined in (8.7) and with $|Y|_\infty$ bounded by a constant depending only on D and $\mathcal{M}_+$. Therefore for M_0 large enough we can write J as

$$J = \begin{pmatrix} J_{1,1} & 0 \\ 0 & J_{2,2} \end{pmatrix},$$

where $J_{1,1}$ is an $M_0 \times M_0$ matrix, and $J_{2,2}$ is – by the definition of M_0 – invertible.

To ensure the invertibility of $J_{1,1}$ we notice that $\det J_{1,1} = 0$ is an analytic equation for the parameter s, and therefore is either identically satisfied or has only a denumerable set of solutions with no accumulation points. For all s outside such denumerable set J is invertible.

For $N = 1$ and $\mathcal{M}_+ = \{V \equiv (1, \dots, 1)\}$ the Dirichlet boundary conditions imply that we only need to consider those $m \in \mathbb{Z}_+^D$ with strictly positive components. For all such m either $m = V$ or $|m|^2 > D$. This implies that $J_{1,1}$ has two diagonal blocks: a 1×1 block involving $\mathcal{M}_+$ and a block involving m such that $|m|^2 > D$. The first block is trivially found to be nonzero. In the second block the off-diagonal entries all depend linearly on D^{2s}, and for all m the diagonal entry with index m is $|m|^{2(1+s)}/D$ plus a term depending linearly on D^{2s}: therefore in the limit $s \to \infty$ this block is invertible. Hence $\det J_{1,1} = 0$ is not an identity in s.

If $N > 1$ we restrict our attention to the case $D = 2$, where we can describe the matrix $J_{1,1}$ with sufficient precision. We have the following sub-lemma.

Sub-lemma. *For $D=2$ and $N>4$ consider $\mathcal{M}_+$ as a point in $\mathbb{C}^{2N}$.*

(i) *The set of points $\mathcal{M}_+$ which either do not respect Lemma* 8.1 *or are such that* $\det J_{1,1}=0$ *identically in s is contained in a proper algebraic variety $\mathcal{W}$.*
(ii) *Provided that $|m_1|$ is large enough one can always find integer points which do not belong to $\mathcal{W}$ and respect* (8.4) *for all s in some open interval.*

The proof is in Appendix A.5. Then the assertion for $D=2$ and $N>4$ follows immediately. □

Remark. On the basis of Lemma 8.2 one expects that invertibility of J holds in more general cases. However, proving that for a given $\mathcal{M}_+$ the function $\det J_{1,1}$ is not identically zero can quite lengthly.

The two cases envisaged in Lemma 8.3, where invertibility of J can be explicitly proved, lead to Theorems 2 and 3.

Theorem 4 covers applies to the general case in which $J_{1,1}$ is known *a priori* to be invertible. Of course, given a set $\mathcal{M}_+$ verifying the conditions of Lemma 8.1 one can check, through a finite number of operations, whether $J_{1,1}$ is invertible, and, if it is, then the analysis below ensures the existence of periodic solutions. Indeed, the forthcoming analysis of the P equations applies without any further assumption in all cases in which $J_{1,1}$ is invertible because of Lemma 8.2.

8.2. Renormalised P equations

The following lemma (Bourgain lemma) will play a fundamental role in the forthcoming discussion. A proof is provided in Appendix A.6.

Lemma 8.4. *For all sufficiently small α we can partition $\mathbb{Z}^D=\bigcup_{j\in\mathbb{N}}\Delta_j$ so that, setting*

$$p_j=\min_{m\in\Delta_j}|m|^2,\qquad \Phi(m)=\big(m,|m|^2\big), \tag{8.10}$$

there exist j-independent constants C_1 and C_2 such that

$$|\Delta_j|\leqslant C_1p_j^{\alpha},\qquad \operatorname{dist}\big(\Phi(\Delta_i),\Phi(\Delta_j)\big)\geqslant C_2\min\big\{p_i^{\beta},p_j^{\beta}\big\},\qquad \operatorname{diam}(\Delta_j)<C_1C_2p_j^{\alpha+\beta}, \tag{8.11}$$

with $\beta=\alpha/(1+2^{D-1}D!(D+1)!)D$.

Remarks. (1) For fixed ε, $\omega n-|m|^2$ can be small only if n is the integer nearest to $|m|^2/\omega$.

(2) For any (m_1,n_1) and (n_2,m_2) such that $m_1\in\Delta_j$, $m_2\in\Delta_{j'}$ for $j'\neq j$, and n_i is the integer nearest to $|m_i|^2/\omega$, $i=1,2$, one has

$$|m_1-m_2|+|n_1-n_2|\geqslant C_3\min\big\{p_j^{\beta},p_{j'}^{\beta}\big\}, \tag{8.12}$$

for some constant C_3 independent of ω.

(3) As in Lemma 2.2 also here one could prove that in fact $\mathrm{diam}(\Delta_j) < \mathrm{const.}\, p_j^{\alpha/D}$; see Appendix A.6 for details.

Definition 8.5. We call $\mathcal{C}_j$ the sets of $(n,m) \in \mathbb{Z} \times \mathbb{Z}^D$ such that $m \in \Delta_j$, $Dn \neq |m|^2$ and $-1/2 + (D-\varepsilon_0)n \leqslant |m|^2 \leqslant Dn + 1/2$. We set $\delta_{n,m} = -\omega n + |m|^2$ and $d_j = |\mathcal{C}_j|$, and define the d_j-dimensional vectors and the $d_j \times d_j$ matrices

$$U_j = \{u_{n,m}\}_{(n,m)\in\mathcal{C}_j}, \qquad \mathbb{D}_j = \mathrm{diag}\left\{\left(\frac{|m|^2}{p_j}\right)^s \delta_{n,m}\right\}_{(n,m)\in\mathcal{C}_j},$$
$$\hat{\chi}_{1,j} = \mathrm{diag}\left\{\sqrt{\bar{\chi}_1(\delta_{n,m})}\right\}_{(n,m)\in\mathcal{C}_j} \tag{8.13}$$

parameterised by $j \in \mathbb{N}$.

Remark. Notice that for each pair $(n,m),(n',m') \in \mathcal{C}_j$ we have $|(n,m)-(n',m')| \leqslant C(\varepsilon_0 p_j/D + p_j^{2\alpha})$ for a suitable constant C.

We define the renormalised P equations

$$\begin{cases} u_{n,m} = \eta \dfrac{f_{n,m}}{|m|^{2s}\delta_{n,m}}, & (n,m) \notin \mathcal{C} := \bigcup_{j\in\mathbb{N}} \mathcal{C}_j, \quad Dn \neq |m|^2, \\ p_j^s\big(\mathbb{D}_j + p_j^{-s}\hat{M}_j\big)U_j = \eta F_j + L_j U_j, & j \in \mathbb{N}, \end{cases} \tag{8.14}$$

where $\hat{M}_j = \hat{\chi}_{1,j} M_j \hat{\chi}_{1,j}$, and the parameter η and the counterterms L_j will have to satisfy eventually the identities

$$\eta = \varepsilon, \qquad \hat{M}_j = L_j, \tag{8.15}$$

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

Remark. We note that d_j can be as large as $O(\varepsilon_0 p_j^{1+\alpha})$, hence can be large with respect to p_j. However for given ε the matrix $A_j = \mathbb{D}_j + p_j^{-s}\hat{M}_j$ is diagonal apart from a $C_1 p_j^{\alpha} \times C_1 p_j^{\alpha}$ (ε-depending) block. This implies that the matrix A_j has at most p_j^{α} eigenvalues which are different from $|m|^{2s}\delta_{n,m}$. This can be proved as follows. Consider the entry $A_j(a,b)$, with $a,b \in \mathcal{C}_j$, with $a=(n_1,m_1)$ and $b=(n_2,m_2)$. The nondiagonal part can be nonzero only if $\sqrt{\bar{\chi}_1(\delta_{n_1,m_1})\bar{\chi}_1(\delta_{n_2,m_2})}M(a,b) \neq 0$, which requires $|\delta_{n_i,m_i}| \leqslant \gamma/4$ for $i=1,2$. Therefore for fixed ε, m_1 and m_2 one has only one possible value for each n_i, i.e. the integer closest to $\omega^{-1}|m_i|^2$. This proves the assertion because $|\Delta_j| \leqslant C_1 p_j^{\alpha}$ and for all $(n,m) \in \mathcal{C}_j$ one has $m \in \Delta_j$.

Definition 2.3 and Lemma 2.4 still hold, with $\mathbb{Z}^{dD}$ replaced with $\mathbb{Z}^{d(D+1)}$ in the definitions of $\mathcal{A}(\underline{m})$. Definitions 2.5(i)–(ii) can be maintained with (n,j) replaced by j, while (iii) becomes

$$x_j = \left\| \hat{\chi}_{1,j}\big(\mathbb{D}_j + p_j^{-s}\hat{M}_j\big)^{-1}\hat{\chi}_{1,j} \right\|^{-1}, \tag{8.16}$$

where the norm $\|\cdot\|$ is defined according to Definition 2.3, with d replaced with $C_1 p_j^\alpha$, which is a bound on the size of the nonsingular block of the matrix $\hat\chi_{1,j}(\mathbb{D}_j + p_j^{-s}\hat M_j)^{-1}\hat\chi_{1,j}$.

Finally there is no parameter s_2. Equivalently we can set $s_2 = 0$, which leads to identify $y_{n,m}$ with $\delta_{n,m}$ (cf. (2.8)): this explains why there is no need to introduce the further parameters $y_{n,m}$.

The main Propositions 1 and 2 in Section 2.4 still hold with the following changes.

1. $(n, j) \in \mathbb{Z} \times \mathbb{N}$ (or Ω) has to be always replaced either with $j \in \mathbb{N}$ or with $(n,m) \notin \mathcal{C}$, for $Dn \neq |m|^2$ – the set $\mathcal{C}$ is defined in (8.14).
2. In Proposition 1, q (i.e. the solution of the Q equations) is not a parameter any more: it is substituted with the solution, say $q^{(0)}$, of the bifurcation equation (8.2), whose Fourier coefficients can be incorporated in the list of positive constants given at the beginning of the statement.
3. In Proposition 1(i) the bound (2.20) becomes

$$\left|u_{n,m}(\eta, M, \varepsilon)\right| \leqslant K_0|\eta| e^{-\sigma(|n|^{1/4}+|m|^{1/4})}, \tag{8.17}$$

for some constant K_0, namely we have only sub-analyticity in space and time.
4. In Proposition 1(v) one must replace s_2 with s in the first line of (2.23) and in (2.26), and $e^{-\sigma|m_a-m_b|^\rho}$ with $e^{-\sigma|(n_a,m_a)-(n_b,m_b)|^\rho}$ in the second line of (2.23), for a suitable constant ρ.

8.3. Multiscale analysis

The multiscale analysis follows in essence the same ideas as in the previous sections, but there are a few changes, that we discuss here. It turns out to be more convenient to replace the functions $\chi_h(x)$ with new functions $\tilde\chi_h(x) = \chi_h(32x)$, in order to have $\tilde\chi_{-1}(x_j) = 1$ when $\bar\chi_1(\delta_{n,m}) \neq 1$ for all $(n,m) \in \mathcal{C}_j$. This only provides an extra factor 32 in the estimates. For notational simplicity in the following we shall drop the tilde.

Let us call $A_j = \mathbb{D}_j + p_j^{-s}\hat M_j$. Note that

$$1 = \bar\chi_1(\delta_{n,m}) + \bar\chi_0(\delta_{n,m}) + \bar\chi_{-1}(\delta_{n,m}) \quad \forall (n,m) \in \mathcal{C}_j. \tag{8.18}$$

Introduce a *block multi-index* $\vec{\mathbf{b}}$, defined as a d_j-dimensional vector with components $\mathbf{b}(a) \in \{1, 0, -1\}$, and set

$$\bar\chi_{j,\vec{\mathbf{b}}} = \prod_{a=1}^{d_j} \bar\chi_{\mathbf{b}(a)}(\delta_{n(a),m(a)}). \tag{8.19}$$

For any $\vec{\mathbf{b}}$ we can consider the permutation $\pi_{\vec{\mathbf{b}}}$ which reorders $(\mathbf{b}(1), \ldots, \mathbf{b}(d_j))$ into $(\mathbf{b}_{\pi_{\vec{\mathbf{b}}}}(1), \ldots, \mathbf{b}_{\pi_{\vec{\mathbf{b}}}}(d_j))$ in such a way that the first N_1 elements are 1, the following N_2 elements are 0, and the last $N_3 = d_j - N$, with $N = N_1 + N_2$, elements are -1. The permutation $\pi_{\vec{\mathbf{b}}}$ induces a permutation matrix $P_{\vec{\mathbf{b}}}$ such that $P_{\vec{\mathbf{b}}} A_j P_{\vec{\mathbf{b}}}^{-1}$ can be written in the block form

$$P_{\vec{\mathbf{b}}} A_j P_{\vec{\mathbf{b}}}^{-1} = \begin{pmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{1,2}^T & A_{2,2} & A_{2,3} \\ A_{1,3}^T & A_{2,3}^T & A_{3,3} \end{pmatrix}, \tag{8.20}$$

where the block $A_{1,1}$, $A_{2,2}$ and $A_{3,3}$ contain all the entries $A_j(a,b)$ with $\mathbf{b}(a)=\mathbf{b}(b)=1$, with $\mathbf{b}(a)=\mathbf{b}(b)=0$ and $\mathbf{b}(a)=\mathbf{b}(b)=-1$, respectively, while the nondiagonal blocks are defined consequently.

Then for all $\vec{\mathbf{b}}$ such that $\bar{\chi}_{j,\vec{\mathbf{b}}}\neq 0$ we can write

$$A_j = P_{\vec{\mathbf{b}}}\begin{pmatrix} A_{1,1} & A_{1,2} & 0 \\ A_{1,2}^T & A_{2,2} & 0 \\ 0 & 0 & A_{3,3}\end{pmatrix} P_{\vec{\mathbf{b}}}^{-1}, \tag{8.21}$$

where we have used that if $\bar{\chi}_{j,\vec{\mathbf{b}}}\neq 0$ then the blocks $A_{1,3}$ and $A_{2,3}$ are zero. Furthermore, for the same reason, the block $A_{3,3}$ is a diagonal matrix. Note that $N \leqslant C_1 p_j^{\alpha}$ by the Remark after (8.15).

The first $N\times N$ block of A_j in general is not block-diagonal, but it can be transformed into a block-diagonal matrix. Indeed, we have

$$A_j = S_{j,\vec{\mathbf{b}}}\tilde{A}_{j,\vec{\mathbf{b}}}S_{j,\vec{\mathbf{b}}}^T, \qquad \tilde{A}_{j,\vec{\mathbf{b}}} = \begin{pmatrix} \tilde{A}_{1,1} & 0 & 0 \\ 0 & A_{2,2} & 0 \\ 0 & 0 & A_{3,3}\end{pmatrix}, \qquad S_{j,\vec{\mathbf{b}}} = P_{\vec{\mathbf{b}}}\begin{pmatrix} I & B & 0 \\ 0 & I & 0 \\ 0 & 0 & I\end{pmatrix}, \tag{8.22}$$

where

$$\tilde{A}_{1,1} = A_{1,1} - A_{1,2}A_{2,2}^{-1}A_{1,2}^T, \qquad B = A_{1,2}A_{2,2}^{-1}, \tag{8.23}$$

while I and 0 are the identity and the null matrix (in the correct spaces). Of course also the matrices $A_{i,j}$ depend on $\vec{\mathbf{b}}$ even if we are not making explicit such a dependence.

The invertibility of $A_{2,2}$ is ensured by the condition $\mathbf{b}(a)=0$ for the indices $a=N_1+1,\dots,N$. The inverse $A_{2,2}^{-1}$ can by bounded proportionally to $1/\gamma$ in the operator norm. Then also A_j can be inverted provided $\tilde{A}_{1,1}$ is invertible, i.e. provided $\det \tilde{A}_{1,1}\neq 0$. Hence in the following we shall assume that this is the case (and we shall check that this holds true whenever it appears; see in particular (8.29) below).

Hence for all $\vec{\mathbf{b}}$ such that $\bar{\chi}_{j,\vec{\mathbf{b}}}\neq 0$ we can write

$$A_j^{-1} = S_{j,\vec{\mathbf{b}}}^{-T}\tilde{A}_{j,\vec{\mathbf{b}}}^{-1}S_{j,\vec{\mathbf{b}}}^{-1}, \tag{8.24}$$

and set

$$\mathbb{G}_{j,\vec{\mathbf{b}},1} = p_j^{-s}S_{j,\vec{\mathbf{b}}}^{-T}\begin{pmatrix} \tilde{A}_{1,1}^{-1} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0\end{pmatrix} S_{j,\vec{\mathbf{b}}}^{-1},$$

$$\mathbb{G}_{j,\vec{\mathbf{b}},0} = p_j^{-s}S_{j,\vec{\mathbf{b}}}^{-T}\begin{pmatrix} 0 & 0 & 0 \\ 0 & A_{2,2}^{-1} & 0 \\ 0 & 0 & 0\end{pmatrix} S_{j,\vec{\mathbf{b}}}^{-1}, \qquad \mathbb{G}_{j,\vec{\mathbf{b}},-1} = p_j^{-s}S_{j,\vec{\mathbf{b}}}^{-T}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & A_{3,3}^{-1}\end{pmatrix} S_{j,\vec{\mathbf{b}}}^{-1}, \tag{8.25}$$

so that (8.24) gives

$$p_j^{-s}A_j^{-1} = \mathbb{G}_{j,\vec{\mathbf{b}},-1} + \mathbb{G}_{j,\vec{\mathbf{b}},0} + \mathbb{G}_{j,\vec{\mathbf{b}},1}, \tag{8.26}$$

for all $\vec{\mathbf{b}}$ such that $\bar{\chi}_{j,\vec{\mathbf{b}}} \neq 0$. We can define $\mathbb{G}_{j,\vec{\mathbf{b}},i}$ also for $\vec{\mathbf{b}}$ such that $\bar{\chi}_{j,\vec{\mathbf{b}}} = 0$, simply by setting $\mathbb{G}_{j,\vec{\mathbf{b}},i} = 0$ for such $\vec{\mathbf{b}}$. Then we define the propagators

$$G_{j,\vec{\mathbf{b}},i,h} = \begin{cases} \bar{\chi}_{j,\vec{\mathbf{b}}} \chi_h(x_j) \mathbb{G}_{j,\vec{\mathbf{b}},1}, & \text{if } i=1 \text{ and } \chi_h(x_j) \neq 0, \\ \bar{\chi}_{j,\vec{\mathbf{b}}} \mathbb{G}_{j,\vec{\mathbf{b}},i}, & \text{if } i=0,-1 \text{ and } h=-1, \\ 0, & \text{otherwise}, \end{cases} \tag{8.27}$$

so that we obtain

$$\begin{aligned} p_j^{-s} A_j^{-1} &= p_j^{-s} \sum_{\vec{\mathbf{b}}} \bar{\chi}_{j,\vec{\mathbf{b}}} A_j^{-1} = \sum_{\vec{\mathbf{b}}} \bar{\chi}_{j,\vec{\mathbf{b}}} \left[(\mathbb{G}_{j,\vec{\mathbf{b}},-1} + \mathbb{G}_{j,\vec{\mathbf{b}},0}) + \sum_{h=-1}^{\infty} \chi_h(x_j) \mathbb{G}_{j,\vec{\mathbf{b}},1} \right] \\ &= \sum_{\vec{\mathbf{b}}} \sum_{i=-1,0,1} \sum_{h=-1}^{\infty} G_{j,\vec{\mathbf{b}},i,h}, \end{aligned} \tag{8.28}$$

which provides the multiscale decomposition.

Remark. Only the propagator $G_{j,\vec{\mathbf{b}},1,h}$ can produce small divisors, because the diagonal propagator $G_{j,\vec{\mathbf{b}},-1,-1}$ and the nondiagonal propagator $G_{j,\vec{\mathbf{b}},0,-1}$ have denominators which are not really small. We can bound $|G_{j,\vec{\mathbf{b}},i,-1}|_\sigma$ for $i=-1,0$ by using a Neumann expansion, since by definition in the corresponding blocks one has $|\delta_{n,m}| \geqslant \gamma/8$ and $|M_j|_\sigma \leqslant C\varepsilon_0$.

Hence we can bound the propagators as

$$|G_{j,\vec{\mathbf{b}},i,-1}|_\sigma \leqslant C\gamma^{-1} p_j^{-s}, \quad i=0,-1, \qquad |G_{j,\vec{\mathbf{b}},1,h}|_\infty \leqslant 2^h C\gamma^{-1} p_j^{-s+\alpha}, \tag{8.29}$$

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

Recall that we are assuming $|J^{-1}|_\sigma \leqslant C$ for some s-dependent constant C.

We write the counterterms as

$$L_j = \sum_{h=-1}^{\infty} \chi_h(x_j) \sum_{\vec{\mathbf{b}}} \bar{\chi}_{j,\vec{\mathbf{b}}} L_{j,\vec{\mathbf{b}},h}, \tag{8.30}$$

where by definition $L_{j,\vec{\mathbf{b}},h}(a,b) = 0$ if either $\mathbf{b}(a) = -1$ or $\mathbf{b}(b) = -1$.

With this modifications to (3.9) the multiscale expansion follows as in Section 3.1, with $j = (n,m)$:

$$U_j^{(k)} = \sum_{i=-1,0,1} \sum_{\vec{\mathbf{b}}} \sum_{h=-1}^{\infty} U_{j,\vec{\mathbf{b}},i,h}^{(k)}, \tag{8.31}$$

with

$$\begin{cases} u^{(k)}_{n,m} = \dfrac{f^{(k)}_{n,m}}{|m|^{2s}\delta_{n,m}}, \quad (n,m) \notin \bigcup_{j\in\mathbb{N}} \mathcal{C}_j, \quad Dn \neq |m|^2, \\ U^{(k)}_{j,\vec{\mathbf{b}},i,h} = G_{j,\vec{\mathbf{b}},i,h} F^{(k)}_j + \delta(i,1) G_{j,\vec{\mathbf{b}},1,h} \left(\displaystyle\sum_{h_1=-1}^{\infty} \sum_{\vec{\mathbf{b}}_1 \neq \vec{0}} \sum_{i_1=0,1,-1} \sum_{r=1}^{k-1} L^{(r)}_{j,\vec{\mathbf{b}},h} U^{(k-r)}_{j,\vec{\mathbf{b}}_1,i_1,h_1} \right), \quad j \in \mathbb{N}, \\ u^{(k)}_{n,m} = q^{(k)}_m = J^{-1} \displaystyle\sum_{k_1+k_2+k_3=k} \sideset{}{^*}\sum_{\substack{m_1+m_2-m_3=m \\ n_1+n_2-n_3=|m|^2/D}} u^{(k_1)}_{n_1,m_1} u^{(k_2)}_{n_2,m_2} u^{(k_3)}_{n_3,m_3}, \quad Dn = |m|^2, \end{cases} \tag{8.32}$$

where $*$ has the same meaning as in (8.9) and $\delta(i, j)$ as usual is Kronecker's delta.

8.4. Tree expansion

We only give the differences with respect to Section 3.2.

(1) As in Section 3.2.
(2) One has $(n_v, m_v) \in Q$ and the node factor is $\eta_v = q^{(0)}_{m_v}$.
(3) We add a further label r, p, q to the lines to evidence which term of (8.32) we are considering. We also associate with each line ℓ a label $j_\ell \in \mathbb{Z}_+$, with the constraints $j_\ell \in \mathbb{N}$ if ℓ is a p-line and $j_\ell = 0$ otherwise.
(4) The momenta are: $(n_\ell, m_\ell), (n'_\ell, m'_\ell) \in \mathcal{C}_{j_\ell}$ for a p-line, $(n_\ell, m_\ell), (n'_\ell, m'_\ell) \in Q$, with $|m_\ell - m'_\ell| \leqslant M_1$, for a q-line, and finally $(n_\ell, m_\ell) = (n'_\ell, m'_\ell) \notin \bigcup_{j\in\mathbb{N}} \mathcal{C}_j \cup Q$ for an r-line. For a p-line the momenta define the labels $a_\ell, b_\ell \in \{1, \dots, d_j\}$, with $d_{j_\ell} = |\mathcal{C}_{j_\ell}|$, such that $(n_\ell, m_\ell) = \mathcal{C}_{j_\ell}(a_\ell)$ and $(n'_\ell, m'_\ell) = \mathcal{C}_{j_\ell}(b_\ell)$. For a q-line the momenta define a_ℓ, b_ℓ such that $(n_\ell, m_\ell) = Q(a_\ell)$ and $(n'_\ell, m'_\ell) = Q(b_\ell)$.
(5) Each p-line carries also a *block label* $\vec{\mathbf{b}}_\ell$ with components $\mathbf{b}_\ell(a) = -1, 0, 1$, where $a = 1, \dots, d_{j_\ell}$.
(6) Both r-lines and q-lines ℓ have $i_\ell = -1$ and $h_\ell = -1$.
(7) One must replace (n_ℓ, j_ℓ) with j_ℓ. Moreover if two lines ℓ and ℓ' have $j_\ell = j_{\ell'}$ then $|\mathbf{b}_\ell(a) - \mathbf{b}_{\ell'}(a)| \leqslant 1$ and if $h_\ell \neq -1$ then $\vec{\mathbf{b}}_\ell \neq \vec{0}$ (by the definition of functions χ_h).
(8) One has $n_\ell = n_v$ instead of $n_\ell = 1$ for lines ℓ coming out from end-points.
(9) One must replace (n_ℓ, j_ℓ) with j_ℓ.
(10) Eq. (3.16) becomes

$$n'_\ell = \sigma(\ell_1) n_{\ell_1} + \sigma(\ell_2) n_{\ell_2} + \sigma(\ell_3) n_{\ell_3} = \sum_{\ell' \in L(v)} \sigma(\ell') n_{\ell'} \tag{8.33}$$

(that is n_ℓ is replaced with n'_ℓ), while (3.17) does not change.
(11) The propagator G_ℓ of any line ℓ is given by $g_\ell = G_{j_\ell,\vec{\mathbf{b}}_\ell,i_\ell,h_\ell}(a_\ell, b_\ell)$, as defined in (8.27), if ℓ is a p-line, while it is given by $g_\ell = J^{-1}(a_\ell, b_\ell)$ if ℓ is a q-line and by $g_\ell = 1/\delta_{n_\ell,m_\ell}|m_\ell|^{2s}$ if ℓ is an r-line.
(12) The node factor for $s_v = 1$ is $\eta_v = L^{(k_v)}_{j_\ell,\vec{\mathbf{b}}_\ell,h_\ell}(a_v, b_v)$, where ℓ is the line exiting v.

The set $\Theta_j^{(k)}$ is defined as in Definition 3.4, with j instead of n, m, by taking into account also the new rules listed above. This will lead to a tree representation (3.20) for (8.32), which can be proved as for Lemma 3.6.

In Lemma 3.5 the estimate $|n_\ell| \leqslant Bk$ does not hold any more because there is no longer conservation of the momenta n_ℓ (i.e. (3.18) has been replaced with (8.33)), and all the bounds on the momenta should be modified into $|n_\ell|, |n'_\ell|, |m_\ell|, |m'_\ell| \leqslant Bk^{1+4\alpha}$ for some constant B. This can be proved by induction on the order of the tree. The bound is trivially true to first order. It is also trivially true if either the root line has $i = -1$ or it is q-line or an r-line (one just needs to choose B appropriately). Suppose now that the root line is a p-line with $i \neq -1$: call v_0 the node which the root line exits. If $s_{v_0} = 3$, call $\theta_1, \theta_2, \theta_3$ the sub-trees with root lines ℓ_1, ℓ_2, ℓ_3, respectively, entering the node v_0. We have $|(n_{\ell_i}, m_{\ell_i})| \leqslant k_i^{1+4\alpha}$ by the inductive hypothesis, and by definition $|(n'_\ell, m'_\ell)| \leqslant \sum_{i=1}^3 Bk_i^{1+4\alpha} \leqslant B(k-1)^{1+4\alpha}$. Then $|(n_\ell, m_\ell)| \leqslant B(k-1)^{1+4\alpha} + C_2(k-1)^{2\alpha(1+4\alpha)} \leqslant Bk^{1+4\alpha}$. If $s_v = 1$ the proof is easier.

8.5. *Clusters and resonances*

Definition 3.7 of cluster is unchanged, while Definition 3.8 of resonance becomes as follows.

Definition 8.6. We call 1-*resonance* on scale $h \geqslant 0$ a cluster T of scale $h(T) = h$ with only one entering line ℓ_T and one exiting line ℓ_T^1 of scale $h_T^{(e)} > h + 1$ with $|V(T)| > 1$ and such that

(i) one has

$$\text{(a)} \quad j_{\ell_T^1} = j_{\ell_T}, \qquad \text{(b)} \quad p_{j_{\ell_T}} \geqslant 2^{(h-2)/\tau}, \tag{8.34}$$

(ii) for all $\ell \in L(T)$ not on the path $\mathcal{P}(\ell_T, \ell_T^1)$ one has $j_\ell \neq j_{\ell_T}$.

We call 2-*resonance* a set of lines and nodes which can be obtained from a 1-resonance by setting $i_{\ell_T} = 0, -1$. Resonances are defined as the sets which are either 1-resonances or 2-resonances. Differently from 3.8 we do not include among the resonant lines the lines exiting a 2-resonance.

Definition 3.9 is unchanged provided that we replace (n, j) with j, we require $p_j \geqslant 2^{(h-2)/\tau}$, we associate with the node e the labels $(n_e, m_e) \in \mathcal{C}_j$ and with ℓ_0 the labels $(n_{\ell_0}, m_{\ell_0}) \in \mathcal{C}_j$.

Since we do not have the conservation of the momentum n, Lemma 3.10 does not hold in the same form: the bounds have to be weakened into $|n_\ell|, |m_\ell|, |n'_\ell|, |m'_\ell| \leqslant Bk^{1+4\alpha}$ for the lines ℓ not along the $\mathcal{P}(\ell_e, \ell_0)$, and $|n_\ell|, |m_\ell|, |n'_\ell|, |m'_\ell| \leqslant B(|n| + k)^{1+4\alpha}$ for the lines along the path.

8.6. *Choice of the counterterms*

The choice of the counterterm (8.30) is not unique and therefore is rather delicate.

Resonances produce contributions that make the power series to diverge. We want to eliminate such divergences with a careful choice of the counterterms.

The sets $\Theta_{R,j}^{(k)}$ and $\mathcal{R}_{R,h,j}^{(k)}$ are defined slightly differently with respect to Definition 3.11.

Definition 8.7. We denote by $\Theta_{R,j}^{(k)}$ the set of renormalised trees defined as the trees in $\Theta_j^{(k)}$ with the following differences:

(i) The trees do not contain any 1-resonance T with $\vec{\mathbf{b}}_{\ell_T^1} = \vec{\mathbf{b}}_{\ell_T}$.

(ii) If a node v has $s_v = 1$ then $\vec{\mathbf{b}}_\ell \neq \vec{\mathbf{b}}_{\ell'}$, where ℓ and ℓ' are the lines exiting and entering, respectively, the node v. The factor $\eta_v = L^{(k_v)}_{j_\ell,\vec{\mathbf{b}}_\ell,h_\ell}$ associated with v will be defined in (8.39).

(iii) The propagators of any line ℓ entering any 1-resonance T (recall that by (i) one has $\vec{\mathbf{b}}_{\ell_T^1} \neq \vec{\mathbf{b}}_{\ell_T}$, where $\ell_T = \ell$), is

$$g_\ell = \chi_{h_\ell}(x_{j_\ell})\bar{\chi}_{j_\ell,\vec{\mathbf{b}}_\ell}\big(\mathbb{G}_{j_\ell,\vec{\mathbf{b}}_{\ell_T^1},0}(a_\ell,b_\ell) + \mathbb{G}_{j_\ell,\vec{\mathbf{b}}_{\ell_T^1},-1}(a_\ell,b_\ell) - \mathbb{G}_{j_\ell,\vec{\mathbf{b}}_\ell,0}(a_\ell,b_\ell) - \mathbb{G}_{j_\ell,\vec{\mathbf{b}}_\ell,-1}(a_\ell,b_\ell)\big), \tag{8.35}$$

and the same holds for the propagator of any line ℓ with $i_\ell = 1$ entering a node v with $s_v = 1$.

In the same way we define $\mathcal{R}^{(k)}_{R,h,j}$. We call $\mathcal{R}^{(k)}_{R,h,j}(a,b)$ the set of trees $\theta \in \mathcal{R}^{(k)}_{R,h,j}$ such that the entering line has $m_e = \mathcal{C}_j(a)$ while the root line has $m'_{\ell_0} = \mathcal{C}_j(b)$. Finally we define the sets $\Theta^{(k)}_R$ and $\mathcal{R}^{(k)}_R$ as the sets of trees belonging to $\Theta^{(k)}_{R,j}$ for some j and, respectively, to $\mathcal{R}^{(k)}_{R,h,j}$ for some h, j.

By proceeding as in Section 3.5 we introduce the following matrices:

$$\mathcal{T}^{(k)}_{j,h}(a,b) = \sum_{h_1<h-1} \sum_{\theta\in\mathcal{R}^{(k)}_{R,j,h_1}(a,b)} \operatorname{Val}(\theta). \tag{8.36}$$

We use a different symbol for such matrices, as we shall see that the counterterms will not be identified with the matrices in (8.35), even if they will be related to them. We shall see that, by the analog of Lemma 3.14, the matrices $\mathcal{T}^{(k)}_{j,h}$ are symmetric.

To define the counterterms L_j we note that, in order to cancel at least the 1-resonances, we need the following condition:

$$G_{j,\vec{\mathbf{b}},1,h}\big(L^{(k)}_{j,\vec{\mathbf{b}},h} + \mathcal{T}^{(k)}_{j,h}\big)\mathbb{G}_{j,\vec{\mathbf{b}},1} = 0. \tag{8.37}$$

Moreover in order to solve the compatibility condition we need a solution $L_{j,\vec{\mathbf{b}},h}(a,b)$ which is proportional to $\bar{\chi}_1(\delta_{n(a),m(a)})\bar{\chi}_1(\delta_{n(b),m(b)})$, and clearly the solution $L^{(k)}_{j,\vec{\mathbf{b}},h} + \mathcal{T}^{(k)}_{j,h} = 0$ does not comply with this requirement. However, since $G_{j,\vec{\mathbf{b}},1,h}$ is not invertible, (8.37) does not imply $L^{(k)}_{j,\vec{\mathbf{b}},h} = -\mathcal{T}^{(k)}_{j,h}$; indeed there exists a solution such that $L_{j,\vec{\mathbf{b}},h}(a,b) \neq 0$ only if $\mathbf{b}(a) = \mathbf{b}(b) = 1$. This solution does not cancel the resonances T with $\vec{\mathbf{b}}_{\ell_T^1} \neq \vec{\mathbf{b}}_{\ell_T}$, and does not even touch the 2-resonances. Nevertheless, if (8.37) holds, we shall see that we are left only with 2-resonances and partially cancelled 1-resonances, which admit better bounds (see (8.17)).

By definition $L^{(k)}_{j,\vec{\mathbf{b}},h}(a,b) = 0$ if either $\mathbf{b}(a)$ or $\mathbf{b}(b)$ is equal to -1. Then (8.37) reduces to the following equation for the matrix $X = L^{(k)}_{j,\vec{\mathbf{b}},h} + \mathcal{T}^{(k)}_{j,h}$:

$$\begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix}^T S_{j,\vec{\mathbf{b}}}^{-1} X S_{j,\vec{\mathbf{b}}}^{-T} \begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix} = 0 \quad \Longrightarrow \quad X_{1,1} - \left(B X_{1,2}^T + X_{1,2} B^T\right) + B X_{2,2} B^T = 0, \tag{8.38}$$

where we define

$$P_{\vec{\mathbf{b}}}^{-1} X P_{\vec{\mathbf{b}}} = \begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ X_{1,2}^T & X_{2,2} & X_{2,3} \\ X_{1,3}^T & X_{2,3}^T & X_{3,3} \end{pmatrix}.$$

In (8.38) there are two matrices which act as free parameters. A (nonunique) solution is

$$P_{\vec{\mathbf{b}}}^{-1} L_{j,\vec{\mathbf{b}},h}^{(k)} P_{\vec{\mathbf{b}}} = \begin{pmatrix} L_{1,1}^{(k)} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} I & -B & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1} \mathcal{T}_{j,h} P \begin{pmatrix} I & 0 & 0 \\ -B & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{8.39}$$

In this definition, $L_{j,\vec{\mathbf{b}},h}^{(k)}(a,b) \neq 0$ only if $\mathbf{b}(a) = \mathbf{b}(b) = 1$, so that $L_{j,\vec{\mathbf{b}},h}^{(k)}$ has the correct factors $\bar{\chi}_1$, that is $L_{j,\vec{\mathbf{b}},h}^{(k)} = \hat{\chi}_{1,j} \hat{L}_{j,\vec{\mathbf{b}},h}^{(k)} \hat{\chi}_{1,j}$ for a suitable $\hat{L}_{j,\vec{\mathbf{b}},h}^{(k)}(a,b)$. Moreover the 1-resonances with $\vec{\mathbf{b}}_{\ell_T^1} = \vec{\mathbf{b}}_{\ell_T}$ are cancelled, while the 2-resonances with $\vec{\mathbf{b}}_{\ell_T^1} = \vec{\mathbf{b}}_{\ell_T}$ are untouched since $L_{j,\vec{\mathbf{b}},h}(G_{j,\vec{\mathbf{b}},0,-1} + G_{j,\vec{\mathbf{b}},-1,-1}) = 0$.

Let us now consider a 1-resonance T with $\vec{\mathbf{b}}_{\ell_T^1} \neq \vec{\mathbf{b}}_{\ell_T}$. We can write (by setting $\vec{\mathbf{b}} = \vec{\mathbf{b}}_{\ell_T^1}$)

$$\begin{aligned} &G_{j,\vec{\mathbf{b}},1,h}\big(L_{j,\vec{\mathbf{b}},h}^{(k)} + \mathcal{T}_{j,h}^{(k)}\big) G_{j,\vec{\mathbf{b}}_1,1,h_1} \\ &\quad = G_{j,\vec{\mathbf{b}},1,h}\big(L_{j,\vec{\mathbf{b}},h}^{(k)} + \mathcal{T}_{j,h}^{(k)}\big) \chi_{h_1}(x_j) \bar{\chi}_{j,\vec{\mathbf{b}}_1}\big(p_j^{-s} A_j^{-1} - \mathbb{G}_{j,\vec{\mathbf{b}}_1,0} - \mathbb{G}_{j,\vec{\mathbf{b}}_1,-1}\big) \\ &\quad = G_{j,\vec{\mathbf{b}},1,h}\big(L_{j,\vec{\mathbf{b}},h}^{(k)} + \mathcal{T}_{j,h}^{(k)}\big) \chi_{h_1}(x_j) \bar{\chi}_{j,\vec{\mathbf{b}}_1}(\mathbb{G}_{j,\vec{\mathbf{b}},0} + \mathbb{G}_{j,\vec{\mathbf{b}},-1} - \mathbb{G}_{j,\vec{\mathbf{b}}_1,0} - \mathbb{G}_{j,\vec{\mathbf{b}}_1,-1}), \end{aligned} \tag{8.40}$$

which does not vanish since $\vec{\mathbf{b}}_{\ell_T^1} \neq \vec{\mathbf{b}}_{\ell_T}$. In that case we say that the 1-resonance is *regularised*.

Then Lemma 3.13 holds true, with $L_{n,j,h}^{(k)}$ substituted with $T_{j,h}^{(k)}$, provided that in the definition of renormalised trees (cf. Definition 3.11) we add the condition that all 1-resonances T with $\vec{\mathbf{b}}_{\ell_T^1} \neq \vec{\mathbf{b}}_{\ell_T}$ and all the nodes with $s_v = 1$ and $i_v = 1$ are regularised.

Also Lemma 3.14 is still true, as the property for the matrix to be symmetric depends only the nonlinearity.

8.7. Bryuno lemma in $\Theta_R^{(k)}$

The set $\mathfrak{S}(\theta, \gamma)$ is defined by (4.1), provided we substitute (n, j) with j and γ with $\gamma/32$. (4.2) is replaced by

$$\begin{cases} |\delta_{n(a),m(a)}| \leqslant 2^{-2}\gamma, & \mathbf{b}_\ell(a) = 1, \\ 2^{-3}\gamma \leqslant |\delta_{n(a),m(a)}| \leqslant 2^{-1}\gamma, & \mathbf{b}_\ell(a) = 0, \\ 2^{-2}\gamma \leqslant |\delta_{n(a),m(a)}|, & \mathbf{b}_\ell(a) = -1, \end{cases} \tag{8.41}$$

for all $j_\ell \geqslant 1$ and $a = 1, \dots, d_{j_\ell}$.

For the definition of the set $\mathfrak{D}(\theta,\gamma)$ we require only the condition (4.3), which becomes

$$|x_{j_\ell}| \geqslant \frac{\gamma}{p_{j_\ell}^{\tau}}. \tag{8.42}$$

We define $N_h(\theta)$ as the set of lines ℓ with $i_\ell = 1$ and scale $h_\ell \geqslant h$, which do not enter any resonance. Then, with this new definition of $N_h(\theta)$, one has $N_h(\theta) \leqslant \max\{0, ck(\theta)2^{(2-h)\beta/2\tau} - 1\}$ (note in the exponent the extra factor $1/2$ with respect to the bound (4.5)). The proof follows the same lines as the proof of Lemma 4.1 in Section 4.1, with the following minor changes.

In order to have a line on scale h we need that $Bk^{1+4\alpha} \geqslant Cp_{j_\ell} \geqslant C2^{(h-1)/\tau}$ for some constant C. We proceed as in the proof of Lemma 4.1, up to (4.8), where again n_{ℓ_i} should be substituted with $p_{j_{\ell_i}}$ with $i = 0, 1$. Define $E_h = c^{-1}2^{(-2+h)\beta/2\tau}$.

1. If $j_{\ell_1} = j_{\ell_0}$ then, since ℓ_1 by hypothesis does not enter a (regularised) resonance, there exists a line ℓ' with $i_{\ell'} \in \{0,-1\}$, not along the path $\mathcal{P}(\bar\ell, \ell_0)$, such that $j_{\ell'} = j_{\ell_0}$. By the Remark after Definition 8.5, we know that $|n_{\ell'}| \geqslant |n_{\ell_0}/2| > 2^{(h-2)/\tau}$. In this case one has $(k(\theta) - k(\theta_1))^{1+4\alpha} > B^{-1}|n_{\ell'}| \geqslant E_h$.
2. If $j_{\ell_1} \neq j_{\ell_0}$ then we call $\bar\ell \in \mathcal{P}(\ell_0, \ell_1)$ the line with $i \neq -1$ which is the closest to ℓ_0.

2.1. If $p_{j_{\bar\ell}} \leqslant p_{j_{\ell_0}}/2$ then $(k(\theta) - k(\theta_1))^{1+4\alpha} \geqslant Cp_{j_{\ell_0}}$.
2.2. If $p_{j_{\bar\ell}} > p_{j_{\ell_0}}/2$ then one reasons as in case 2.2 of Lemma 4.1, with the following differences.

2.2.1. If $j_{\ell_0} \neq j_{\bar\ell}$, then $|(n_{\bar\ell}, m_{\bar\ell}) - (n_{\ell_0}, m_{\ell_0})| \geqslant \text{const.}\, p_{j_{\ell_0}}^{\beta}$. For all the lines ℓ along the path $\mathcal{P}(\bar\ell, \ell_0)$ one has $i_\ell = -1$, hence either $n_\ell = n'_\ell$ and $m_\ell = m'_\ell$ (if ℓ is a p-line) or $|m_\ell - m'_\ell| \leqslant M_1$ (if ℓ is a q-line), so that $|(n_{\bar\ell}, m_{\bar\ell}) - (n_{\ell_0}, m_{\ell_0})| \leqslant 2B(k(\theta) - k(\theta_1))^{1+4\alpha}$, with the same meaning for the symbols as in Section 4.1, and the assertion follows once more by using (4.8).
2.2.2. If $j_{\ell_0} = j_{\bar\ell}$ then there are two further sub-cases.

2.2.2.1. If $\bar\ell$ does not enter any resonance, we proceed as in item 1.
2.2.2.2. If $\bar\ell$ enters a resonance, then we continue up to the next line $\tilde\ell$ on the same path with $i \neq -1$. If $j_{\tilde\ell} \neq j_{\ell_0}$ the proof is concluded as in 2.2.1 since $2Bk^{1+4\alpha} \geqslant |(n_{\bar\ell}, m_{\bar\ell}) - (n_{\tilde\ell}, m_{\tilde\ell})| \geqslant C_1 p_j^{\beta}$. Likewise – using item 2.2.2.1 – the proof is concluded if the line $\tilde\ell$ does not enter a resonance. If $\tilde\ell$ enters a resonance with $j_{\tilde\ell} = j_{\ell_0}$, we proceed until we reach a line with $i \neq -1$ which either has $j \neq j_{\ell_0}$ or does not enter a resonance: this is surely possible, because by definition ℓ_1 does not enter a resonance and $j_{\ell_1} \neq j_{\ell_0}$. This completes the proof of the lemma.

Lemma 4.2 holds with $|n|, |m| \leqslant Bk^{1+4\alpha}$ and $q = 1$, and with $p_j^{-3s/4}$ in all the lines of (4.9). The proof is the same (recall that we can set $s_2 = 0$); we only need to substitute p_j^{α} (which bounded the dimension of the nondiagonal block) with d_j. In (iii) the labels (n', j') should be substituted by j'.

8.8. Bryuno lemma in $\mathcal{R}_R^{(k)}$

The definitions of $\tilde{\mathfrak{S}}(\theta,\gamma)$ and $\tilde{\mathfrak{D}}(\theta,\gamma)$ are changed exactly as $\mathfrak{S}(\theta,\gamma)$ and $\mathfrak{D}(\theta,\gamma)$, respectively, in the previous Section 8.7.

Definition 8.8. We divide $\mathcal{R}_{R,h,j}$ into two sets $\mathcal{R}^1_{R,h,j}$ and $\mathcal{R}^2_{R,h,j}$: $\mathcal{R}^1_{R,h,j}$ contains all the trees such that either $\mathcal{P}(\ell_0,\ell_e)=\emptyset$ or at least one line $\ell\in\mathcal{P}(\ell_0,\ell_e)$ has $j_\ell\neq j$, and $\mathcal{R}^2_{R,h,j}=\mathcal{R}_{R,h,j}\setminus\mathcal{R}^1_{R,h,j}$. This naturally yields a decomposition $\mathcal{R}^{(k)}_{R,h,j}=\mathcal{R}^{(k,1)}_{R,h,j}\cup\mathcal{R}^{(k,2)}_{R,h,j}$ for all $k\in\mathbb{N}$.

The two properties (i) and (ii) of Lemma 4.3 should be restated as follows.

(i) *There exists a positive constant B_2 such that if $k\leqslant B_2 p_j^{\beta/(1+4\alpha)}$ then $\mathcal{R}^{(k,1)}_{R,j,h}$ contains only trees with $\mathcal{P}(\ell_0,\ell_e)=\emptyset$;*
(ii) *for all $\theta\in\mathcal{R}^{(k,1)}_{R,h,j}(a,b)$ we have $B_3|(n(a),m(a))-(n(b),m(b))|^\rho\leqslant k$, with ρ a constant depending on D, for a suitable positive constant B_3.*

The proof of (i) can be obtained by reasoning as in the cases 2.1 and 2.2.1 of Section 8.7, while that of (ii) proceeds as in the proof of Lemma 4.3(ii).

For the trees in $\mathcal{R}^{(k,2)}_{R,h,j}$ all the lines ℓ along the path $\mathcal{P}(\ell_0,\ell_e)$ have $j_\ell=j$, and we can bound the product of the corresponding propagators as

$$\left(\prod_{\ell\in\mathcal{P}(\ell_0,\ell_e)}4C\gamma^{-1}p_{j_\ell}^{-s}\right)\exp\left(-\sigma\sum_{\ell\in\mathcal{P}(\ell_0,\ell_e)}\left|(n_\ell,m_\ell)-\left(n'_\ell,m'_\ell\right)\right|^\rho\right)$$
$$\leqslant C^k e^{-\sigma|(n_{\ell_0},m_{\ell_0})-(n_{\ell_e},m_{\ell_e})|^\rho}, \tag{8.43}$$

where the factor 4 is due to regularisation of the propagators with $i_\ell=1$ (see (8.35)), and we have used (8.29) to bound $|G_{j,\vec{\mathbf{b}},i,-1}|_\sigma$ for $i=0,-1$. Hence also $|\mathrm{Val}(\theta)|_\sigma$ is bounded by C^k.

Lemma 4.4 and properties (i) and (ii) of Lemma 4.6 are modified exactly as the corresponding 4.1 and 4.2. In (4.17)(ii) $|n|$ should be substituted by $|p_j|$. Finally (4.17)(iii) should be replaced with

$$\sum_{j'\in\mathbb{N}}\sum_{a',b'=1}^{d_{j'}}\left|\partial_{M_{j'}(a',b')}\mathrm{Val}(\theta)\right|\leqslant D^k 2^{-h}\left(\prod_{h'=-1}^{h}2^{2h'N_{h'}(\theta)}\right)\prod_\ell p_{j_\ell}^{3s/4}, \tag{8.44}$$

which can be proved as follows.

1. Let us first consider $\mathcal{R}^1_{R,j,h}$. We have no difficulty in bounding the sums and derivatives applied on lines $\ell\notin\mathcal{P}(\ell_0,\ell_e)$. By the analog of Lemma 4.3 discussed above, if $B_2 k\leqslant p_j^{\beta/(1+4\alpha)}$ then $\mathcal{P}(\ell_0,\ell_e)=\emptyset$ and we have no problem. Otherwise we have at most $(2p_j+k)^{1+4\alpha}$ possible values of (n,m) and (n',m') which can be associated with a line ℓ along the path $\mathcal{P}(\ell_0,\ell_e)$ and by our assumption one has $(2p_j+k)^{1+4\alpha}\leqslant C^k$ for some constant C.

2. If all the lines $\ell \in \mathcal{P}(\ell_0, \ell_e)$ have $j_\ell = j$ then the sums with $a' \neq b'$ contain at most $C_1^2 p_j^{2\alpha}$ terms, whereas the sums with $a' = b'$ contain at most k terms, since there are at most k lines on $\mathcal{P}(\ell_0, \ell_e)$.

The rest of Section 4 is unchanged. In Section 5.1 we remove the second Melnikov condition (the $**$ and $***$ products) in (5.3) and (5.5).

8.9. Measure estimates

By definition we have to evaluate the measure of the set

$$\left\{\varepsilon\colon\ \left\|\hat{\chi}_{1,j}\left(\mathbb{D}_j + p_j^{-s}\hat{M}_j\right)^{-1}\hat{\chi}_{1,j}\right\|^{-1} \geqslant \frac{2\gamma}{p_j^\tau}\ \forall j \in \mathbb{N}\right\}. \tag{8.45}$$

By Lemma 2.4(iii) one has

$$x_j \geqslant \min_{i=1,\dots,d_j}\left|\lambda^{(i)}\left(\mathbb{D}_j + p_j^{-s}\hat{M}_j\right)\right|, \tag{8.46}$$

since the matrices are symmetric and the minimum is attained for some i such that $\bar{\chi}_1(\delta_{n(i),m(i)}) \neq 0$.

The set (8.45) contains the set

$$\mathfrak{E} = \left\{\varepsilon \in (0, \varepsilon_0)\colon\ \left|\lambda^{(i)}\left(\mathbb{D}_j + p_j^{-s}\hat{M}_j\right)\right| \geqslant \frac{2\gamma}{p_j^\tau}\ \forall i = 1, \dots, d_j,\ \forall j \in \mathbb{N}\right\}. \tag{8.47}$$

We estimate the measure of the subset of $(0, \varepsilon_0)$ complementary to $\mathfrak{E}$, i.e. the set defined as union of the sets

$$\mathfrak{I}_{j,i} := \left\{\varepsilon \in (0, \varepsilon_0)\colon\ \left|\lambda^{(i)}\left(\mathbb{D}_j + p_j^{-s}\hat{M}_j\right)\right| \leqslant \frac{2\gamma}{p_j^\tau}\right\}, \tag{8.48}$$

for $j \in \mathbb{N}$ and $i = 1, \dots, d_j$.

First we notice that if $|p_j| \leqslant C/\varepsilon_0^\xi$, for appropriate constants ξ and C, then, by Lidskii's Lemma [24],

$$\left|\lambda^{(i)}\left(\mathbb{D}_j + p_j^{-s}\hat{M}_j\right)\right| \geqslant \left(\frac{|m(i)|^2}{p_j}\right)^s\left(-Dn(i) + \left|m(i)\right|^2\right) - C\left(\varepsilon_0 p_j + p_j^\alpha\right)p^{2\alpha}(\varepsilon_0 + \varepsilon_0 p_j), \tag{8.49}$$

which implies that

$$\lambda^{(i)}\left(\mathbb{D}_j + p_j^{-s}\hat{M}_j\right) \geqslant \frac{1}{2},$$

as soon as $\xi = 1/2$ and C is suitably small. Therefore we have to discard the sets $\mathfrak{I}_{j,i}$ only for $p_j \geqslant C/\varepsilon_0^\xi$.

Let us now recall that for a symmetric matrix $M(x)$ depending analytically on a parameter x, the derivatives of the eigenvalues are: $\partial_x \lambda^{(i)}(x) = \langle v_i, \partial_x M(x) v_i \rangle$, where v_i are the corresponding eigenvectors [24].

Since $\mathbb{D}_j$ depends linearly – and therefore analytically – on ε we consider $\lambda_i(x,\varepsilon) := \lambda^{(i)}(\mathbb{D}_j(x) + p_j^{-s}\hat{M}_j)$ with x, ε independent parameters.

Clearly $|\partial_x \lambda_i(x,\varepsilon)| \geqslant p_j$, and, by Lidskii's Lemma again,

$$\left|\partial_\varepsilon \lambda_i(x,\varepsilon)\right| \leqslant p_j^{-s} \sum_{i=1}^{d_j} \left|\lambda^{(i)}(\partial_\varepsilon \hat{M}_j)\right|.$$

Now $\hat{M}_j$ is a $d_j \times d_j$ matrix which for each fixed $\bar{\varepsilon}$ has only a nonzero block of size p_j^α; the properties of the functions $\bar{\chi}_{j,1}$ imply that also $\partial_\varepsilon \hat{M}_j$ has only a nonzero block of size p_j^α. So one has

$$\left|\partial_\varepsilon \lambda_i(x,\varepsilon)\right| \leqslant C\big(1 + \varepsilon_0 p_j^{1-s+5\alpha}\big),$$

for some constant C.

Then the measure of each $\mathfrak{I}_{j,i}$ can be bounded from above by

$$\frac{4\gamma}{p_j^\tau} \sup_{\varepsilon\in(0,\varepsilon_0)} \left| \left(\frac{\mathrm{d}}{\mathrm{d}\varepsilon} \lambda^{(i)}\big(\mathcal{D}_j(\varepsilon) + p_j^{-s}\hat{M}_j(\varepsilon)\big) \right)^{-1} \right| \leqslant \frac{8\gamma}{p_j^{\tau+1}}. \tag{8.50}$$

Therefore we have

$$\sum_{j\in\mathbb{N}} \sum_{i=1}^{d_j} \operatorname{meas}(\mathfrak{I}_{j,i}) \leqslant \text{const.} \sum_{p\geqslant C/\varepsilon_0^\xi} \gamma p^{D+\alpha} \left(\frac{1}{p^{\tau+1}} \right) \leqslant \text{const.}\big(\varepsilon_0^{\xi(\tau-D-\alpha)}\big), \tag{8.51}$$

provided $\tau > D + \alpha + 1/\xi$, so that the measure of the complementary of $\mathfrak{E}$ is small in $(0,\varepsilon_0)$ if $\tau > D + \alpha + 1/\xi$.

Acknowledgments

We wish to thank Prof. C. Procesi for useful discussions on the results in Appendix A.4. The paper was written when one of us (M.P.) was partially supported by INDAM.

Appendix A

A.1. Preliminary measure estimate

We estimate the measure of the complement of $\mathfrak{E}_0(\gamma)$, defined in (2.2), with respect to the set $(0,\varepsilon_0)$, under the condition $\mu \in \mathfrak{M}$. For all $n, p \in \mathbb{N}$ we consider the set

$$\mathfrak{I}_{n,p} = \left\{ \varepsilon \in (0,\varepsilon_0)\colon\ |\omega n - p| \leqslant \frac{\gamma}{n^{\tau_1}} \right\}. \tag{A.1.1}$$

The measure of such a set is bounded proportionally to $|n|^{-(\tau_1+1)}$. Moreover one has

$$\sum_{n,p=1}^{\infty} \operatorname{meas}(\mathfrak{I}_{n,p}) \leqslant \text{const.} \sum_{n=1}^{\infty} |n|^{-(\tau_1+1)} + \text{const.}\, \varepsilon_0 \sum_{n=1}^{\infty} |n|^{-\tau_1}, \tag{A.1.2}$$

because the number of values that p can assume is at most $1+\varepsilon_0 n$ (simply note that $|\omega n - p| \geqslant 1/2$ if p is not the integer closest to ωn and $|\omega - D - \mu| \leqslant \varepsilon_0$).

Finally we note that, by (2.1), for $n < (\gamma_0/2\varepsilon_0)^{1/(\tau_0+1)}$ one has

$$|\omega n - p| \geqslant \big|(D+\mu)n - p\big| - \varepsilon_0|n| \geqslant \gamma |n|^{-\tau_0}, \tag{A.1.3}$$

provided $\gamma \leqslant \gamma_0/2$. Hence the sum in (A.1.2) can be restricted to $n \geqslant (\gamma_0/2\varepsilon_0)^{1/(\tau_0+1)}$, so that

$$\sum_{n,p} \operatorname{meas}(\mathfrak{I}_{n,p}) \leqslant \text{const.}\, \varepsilon_0^{\tau_1/(\tau_0+1)} + \text{const.}\, \varepsilon_0^{1+(\tau_1-1)/(\tau_0+1)}, \tag{A.1.4}$$

which is infinitesimal in ε_0 provided $\tau_1 > \tau_0 + 1$.

A.2. Proof of the separation Lemma 2.2

Let $D \in \mathbb{N}$ be fixed, $D \geqslant 2$. For all $D > d \geqslant 1$ and for all $r > 1$ let $S^d(r)$ denote a d-sphere of radius r, $S_0^d(r)$ a sphere $S^d(r)$ centred at the origin and $B(S^d(r))$ the ball with boundary $S^d(r)$. A ball $B(S^d(r))$ determines uniquely a $(d+1)$-dimensional subspace P^{d+1} in which it lies. Consider now a sphere $S^{d-1}(\Gamma)$ contained in $S^d(r)$: this determines a d-dimensional subspace – say H – of P^{d+1}. H divides P^{d+1} in two parts, and hence divides $B(S^d(r))$ in two as well. We call these two parts spherical caps, we call $B(S^{d-1}(\Gamma))$ the base of the cap and Γ the radius of the cap. Finally the maximum of the distance between a point in the cap and the base will be called the height of the cap – and will be denoted by h.

Lemma A.2.1. *Given any $D > d \geqslant 1$ there exist two constants $C(D,d)$, $C'(D,d)$ such that the following holds. For all $0 < \varepsilon \ll 1$ and for all d-spheres $S_0^d(r)$ there exist $N = N(\varepsilon, r, D, d)$ sets of integer points Λ_α (depending on the sphere and on ε, d and D), such that*

$$S_0^d(r) \cap \mathbb{Z}^D = \bigcup_{\alpha=1}^{N} \Lambda_\alpha, \qquad |\Lambda_\alpha| \leqslant C(D,d) \max\{r^\varepsilon, d+2\},$$
$$\operatorname{dist}(\Lambda_\alpha, \Lambda_\beta) \geqslant C'(D,d)\, r^{\delta(\varepsilon,d)}, \tag{A.2.1}$$

with $\delta(\varepsilon,d) := 2\varepsilon/d(d+2)!$.

The proof of this lemma follows easily from the following result.

Lemma A.2.2. *There exist constants C and C' such that the following holds. Let $n_1, \ldots, n_k \in S^d(r) \cap \mathbb{Z}^D$. If for all $i = 1, \ldots, k-1$ one has $|n_i - n_{i+1}| < C r^{\delta(\varepsilon,d)}$ then $k < C' \max\{r^\varepsilon, d+2\}$.*

Facts. We group here some simple facts on points on a sphere; see Fig. 8.

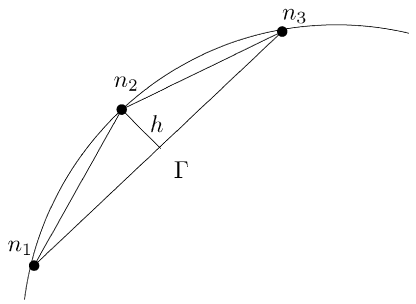

Fig. 8. Simplex generated by three points n_1, n_2, n_3 on the sphere $S^1(r)$. Γ is the base of the spherical cap in which the three points are contained and h the height. If h is small then the volume (= area) of the cap is of order Γ^3/r, with $\Gamma = O(|n_1 - n_2| + |n_1 - n_3|)$.

A. For any sphere $S^d(r)$ one has $|S^d(r) \cap \mathbb{Z}^D| \leqslant \bar{C}(D,d)\, r^d$, for some constant $\bar{C}(D,d)$.
B. $j+1$ affinely independent points $n_1, \ldots, n_{j+1}$ in $\mathbb{Z}^D$ generate a j-dimensional simplex, which is their convex envelope, and determine uniquely a $(j-1)$-dimensional sphere on which the points lie. Setting $v_i = n_1 - n_{i+1}$ for $i = 1, \ldots, j$, the volume of the simplex generated by $n_1, \ldots, n_{j+1}$ is

$$\frac{1}{j!}\left|\det N N^T\right|^{\frac{1}{2}}, \qquad N = \begin{pmatrix} v_{11} & \ldots & v_{1D} \\ \ldots & \ldots & \ldots \\ v_{j1} & \ldots & v_{jD} \end{pmatrix}, \tag{A.2.2}$$

and, since N has integer coefficients, the volume of the simplex is bounded from below by $1/j!$.
C. The volume of a spherical cap of height h and radius Γ is $O(\Gamma^d h)$; some trivial trigonometry tells us that if h is $o(r)$ then $\Gamma^2/h = O(r)$ and hence we can write $\Gamma^d h = O(\Gamma^{d+2}/r)$.
D. Given two balls $S^d(R)$ and $S^d(r)$, $S^d(R)$ determines two spherical caps on $B(S^d(r))$ with basis $B(S^d(r) \cap S^d(R))$: clearly the radius of the caps is smaller than R and r.
E. Consider $j+2$ affinely independent points $n_1, \ldots, n_{j+2}$ in $\mathbb{Z}^D$ which determine the sphere $S^j(r_j)$ for some r_j. If for all $i = 1, \ldots, j+1$ one has $|n_1 - n_{i+1}| \leqslant K = o(r_j)$ then all the points $n_1, \ldots, n_{j+2}$ are contained in a spherical cap which does not contain the center of $S^j(r_j)$ and of radius $\Gamma \leqslant K$. This implies that also $h = o(r_j)$ and the volume of the cap is bounded by $c(j)\Gamma^{j+2}/r_j \geqslant 1/j!$ for some constant $c(j)$. Therefore one has the bound $r_j \leqslant C(j) K^{j+2}$ with $C(j) = j!/c(j)$.

Proof of Lemma A.2.1. For $k \leqslant d+1$ the assertion is trivially satisfied, hence we can assume from now on $k \geqslant d+2$. For $i = 1, \ldots, d$, let k_i be the largest index such that $n_1, \ldots, n_{k_i}$ are contained in an $(i+1)$-dimensional affine subspace and therefore define an i-sphere – say $S^i(r_i)$ in $S^d(r)$ for some $r_i \leqslant r$. The proof is performed by induction.

Step 1. Call $S^1(r_1)$ – with $r_1 \leqslant r$ – the 1-sphere determined by n_1, n_2 and n_3 (which are surely affinely independent); by hypothesis $|n_1 - n_i| \leqslant 2Cr^{\delta(\varepsilon,d)}$ for $i = 2, 3$. Provided $r_1 \geqslant r^{2\delta(\varepsilon,d)}$ then n_1, n_2, n_3 respect the assumptions of Fact E with $K = 2Cr^{\delta(\varepsilon,d)}$, and we have $r_1 < A(D,1) r^{\alpha_1 \delta}$, with $\delta = \delta(\varepsilon,d)$, $\alpha_1 = 3$ and a suitably large constant $A(D,1)$. If $r_1 \leqslant r^{2\delta(\varepsilon,d)}$ then surely $r_1 \leqslant A(D,1) r^{\alpha_1 \delta}$. By Fact A we have at most $\bar{C}(D,1)\, r_1 \leqslant \bar{C}(D,1) A(D,1) r^{3\delta}$ integer vectors on $S^1(r_1) \cap S^d(r)$ and therefore $k_1 \leqslant C(D,1) r^{3\delta}$.

Inductive hypothesis. For $i = 1, \dots, j$ let $A(D, i)$ be suitably large constants and set $\alpha_i = (i + 2)!/2$. Assume that for all $i \leqslant j - 1$, we have proven that $n_1, \dots, n_{k_i}$ lie on an i-sphere of radius $r_i \leqslant A(D, i)r^{\alpha_i\delta}$ – so that $k_i \leqslant C(D, i)r^{i\alpha_i\delta}$.

Step $j > 1$. By definition if $k_{j-1} < k$ the point $n_{k_{j-1}+1}$ is such that $n_1, n_2, n_3, n_{k_1+1}, \dots, n_{k_{j-1}+1}$ are $j + 2$ affinely independent points which determine a j-sphere $S^j(r_j)$, for some $r_j \leqslant r$. By definition of r_{j-1} one has $|n_i - n_1| \leqslant 2r_{j-1} \leqslant 2A(D, j - 1)r^{\alpha_{j-1}\delta}$ for all $i = 2, \dots, k_{j-1}$ and clearly $|n_{k_{j-1}+1} - n_1| \leqslant Cr^{\delta} + 2r_j \leqslant B_j r^{\alpha_{j-1}\delta}$, for a suitable constant B_j. If $r_j \leqslant r^{2a_{j-1}\delta}$ the inductive hypothesis is proven, otherwise if $r_j > r^{2\alpha_{j-1}\delta}$ then we can apply Fact E with $K = B_j r^{\alpha_{j-1}\delta}$. We have $r_j \leqslant A(D, j)r^{(j+2)\alpha_{j-1}\delta} \equiv A(D, j)r^{\alpha_j\delta}$ by setting $\alpha(j) = (j + 2)\alpha_{j-1} = (j + 2)!/2$. □

Remarks. (1) A careful look at the proof of Lemma A.2.2 shows that in Lemma A.2.1 one can choose $C(D, d)$ and $C'(D, d)$ as functions of the only D.

(2) In the proof of Lemma A.2.2 the construction in step 1 shows that if one takes three vectors n_1, n_2 and n_3 on a 1-sphere $S^1(r_1)$ then (with the notations used in the proof of the lemma) one has $\max\{n_1 - n_2, n_1 - n_3\} > C_1 r_1^{1/3}$. Therefore for $d = 1$ these sets Λ_j can be chosen in such a way that each set contains at most two elements, and the distance between two distinct sets on the same sphere $S^1(r)$ is larger than a universal constant times $r^{1/3}$.

Lemma A.2.1 implies that it is possible to decompose the set $\mathbb{Z}^D \cup S_0^D(r)$ as the union of sets Δ such that $\mathrm{diam}(\Delta) < \text{const.}\, r^{\delta+\varepsilon}$ (cf. [6, p. 399]), and $|\Delta| < \text{const.}\, r^{D(\delta+\varepsilon)}$. Hence, if we take α small enough and we set $\beta = \delta$ and $\alpha = D(\delta + \varepsilon)$, by using that $\varepsilon/\delta = (d+2)!d/2$, Lemma 2.1 follows.

A.3. Constructive scheme for Lemma 8.1

Here we prove that the sets $\mathcal{M}_+$ verifying the conditions (a) and (b) in the proof of Lemma 8.1 are nonempty. The proof consists in providing explicitly a construction.

1. Fix a list of parameters $\alpha_2, \dots, \alpha_N \in \mathbb{R}$ such that $\alpha_i < \alpha_{i-1}$ for $i = 2, \dots, N$, with $\alpha_1 = 1$, and

$$2^D \sum_{i=2}^{N} \alpha_i^{2+2s} \leqslant 3^D + 2^D(N - 2). \tag{A.3.1}$$

2. Given $r \in \mathbb{R}^+$ and for $i = 2, \dots, N$ consider the regions $\mathcal{R}_i(r) := \{x \in \mathbb{Z}_+^D\colon \alpha_{i-1}r \leqslant |x| \leqslant \alpha_i r\}$ with r so big that it is not possible to cover any of the $\mathcal{R}_i(r)$ with $3N^2 2^{2D}$ planes and spheres.
3. Choose an integer vector $m_1 \in \mathbb{Z}_+^D$ such that $|m_1|^2 = r^2$ is divided by D, and construct the "orbit" $\mathcal{O}(m_1) := \{m \in \mathbb{Z}^D\colon |m_i| = |(m_1)_i|\}$.
4. For each pair $m, m' \in \mathcal{O}(m_1)$ consider the two planes orthogonal to $m - m'$ and passing respectively through m and m', and the sphere which has $m - m'$ as diameter (there are at most $3 \cdot 2^{D-1}(2^D - 1)$ planes and spheres).
5. Choose the second integer vector $m_2 \in \mathcal{R}_2(r)$ such that $|m_2|^2$ divides D and the orbit $\mathcal{O}(m_2)$ does not lie on any of the planes and spheres defined at step 4.

6. For each pair $m, m' \in \mathcal{O}(m_1) \cup \mathcal{O}(m_2)$ proceed as in step 4. We have at most further $3 \cdot 2^D(2^{D+1} - 1)$ planes and spheres.
7. Then we proceed iteratively. When we arrive to m_N we have to remove at most $3N2^{D-1}(N2^D - 1)$ planes and spheres.

A.4. Blocks of the matrix J

Write $\mathcal{M} = \{m_1, \dots, m_M\}$, with $M = 2^D N$, and set $\mathcal{V} = \{v = (m, m')\colon m, m' \in \mathcal{M},\ m \neq m'\}$: clearly $L := |\mathcal{V}| = M(M-1)$. We call *alphabet* the set $\mathcal{V}$ and *letters* the elements (vectors) of $\mathcal{V}$. We call *word* of length $\ell \geqslant 1$ any string $v_1 v_2 \dots v_\ell$, with $v_k \in \mathcal{V}$ for $k = 1, \dots, \ell$. Let us denote with $\mathcal{A}$ the set of all words with letters in the alphabet $\mathcal{V}$ plus the empty set (which can be seen as a word of length 0).

For $v \in \mathcal{V}$ with $v = (m_i, m_j)$ we write $v(1) = m_i$ and $v(2) = m_j$. Given two words $a = v_1 \dots v_n$ and $b = v'_1 \dots v'_{n'}$ we can construct a new word $ab = v_1 \dots v_n v'_1 \dots v'_{n'}$ of length $n + n'$. Finally we can introduce a map $a \to w(a)$, which associates with any letter $v \in \mathcal{V}$ the vector $v(1) - v(2)$, to any word $a = v_1 \dots v_n$ the vector $w(a) = w(v_1) + \dots + w(v_n)$ and finally $w(\emptyset) = 0$. We say that a is a *loop* if $w(a) = 0$.

Remarks. (1) Given a set $\mathcal{M}$ let $\mathcal{V}$ be the corresponding alphabet. If $|\mathcal{M}| = M$ then $|\mathcal{V}| = L(M) = M(M-1)$. If we add an element m_{N+1} to $\mathcal{M}$ so to obtain a new set $\mathcal{M}' = \mathcal{M} \cup \{m_{N+1}\}$, then the corresponding alphabet $\mathcal{V}'$ contains all the letters of $\mathcal{V}$ plus other $2M$ letters. We can imagine that this alphabet is obtained through $2M$ steps, by adding one by one the $2M$ new letters. In this way, we can imagine that the length L of the alphabet can be increased just by 1.

(2) By construction $w(v_1 v_2) = w(v_2 v_1)$. In particular $w(a)$ depends only on the letters of a (each with its own multiplicity), but not on the order they appear within a.

Define a matrix J, such that

(i) $J_{jk} = J(q_j, q_k)$, with $q_j, q_k \in \mathbb{Z}^D$,
(ii) $J(q, q') \neq 0$ if there exist $m_1, m_2 \in \mathcal{M}$ such that $q - m_1 = q' - m_2$ and $\langle m' - m_2, m_1 - m_2 \rangle = 0$, and $J(q, q') = 0$ otherwise.

A sequence $C = \{q_0, q_1, \dots, q_n\}$ will be called a *chain* if $J(q_{k-1}, q_k) \neq 0$ for $k = 1, \dots, n$. We call $n = |C|$ the length of the chain C. A chain can be seen as a pair of a vector and a word, that is $C = (q_0; a)$, where $q_0 \in \mathbb{Z}^D$ and $a = v_1 \dots v_n$, with $w(v_k) = q_k - q_{k-1}$. Note that, by definition of the matrix J, given a chain C as above, one has

$$q_k = q_{k-1} + w(v_k), \qquad \bigl\langle q_k - v_k(2), w(v_k) \bigr\rangle = 0, \tag{A.4.1}$$

for all $k = 1, \dots, n$.

Lemma A.4.1. *Given a chain $C = (q_0; a)$, if the word a contains a string $v_0 a_0 v_0$, with $v_0 \in \mathcal{V}$ and $a_0 \in \mathcal{A}$, then $\langle w(v_0 a_0), w(v_0) \rangle = 0$.*

Proof. As the word a of C contains the string $v_0 a_0 v_0$, by (A.4.1) there exists $j \geqslant 1$ such that

$$\bigl\langle q_j - v_0(2), w(v_0) \bigr\rangle = 0, \qquad \bigl\langle q_j + v_0 + w(a_0) - v_0(2), w(v_0) \bigr\rangle = 0,$$

so that $\langle w(v_0) + w(a_0), w(v_0) \rangle = 0$. $\square$

Lemma A.4.2. *Given a chain $C = (q_0; a)$, if the word a contains a string $a_0 b_0 a_0$, with $a_0, b_0 \in \mathcal{A}$ and a_0 containing all the letters of the alphabet $\mathcal{V}$, then $a_0 b_0$ is a loop.*

Proof. For any $v \in \mathcal{V}$ we can write $a_0 = a_1 v a_2$, with $a_1, a_2 \in \mathcal{A}$ depending on v. Then $a_0 b_0 a_0 = a_1 v a_2 b_0 a_1 v a_2$. Consider the string $v a_2 b_0 a_1 v$: by Lemma A.4.1 one has $\langle w(v a_2 b_0 a_1), w(v)\rangle = 0$. On the other hand (cf. Remark (2) after the definition of loop) one has $w(v a_2 b_0 a_1) = w(a_1 v a_2 b_0) = w(a_0 b_0)$, so that $\langle w(a_0 b_0), w(v)\rangle = 0$. As v is arbitrary we conclude that

$$\big\langle w(a_0 b_0), w(v)\big\rangle = 0 \quad \forall v \in \mathcal{V} \quad \Longrightarrow \quad w(a_0 b_0) = 0,$$

i.e. $a_0 b_0$ is a loop. □

Lemma A.4.3. *There exists K such that if a word has length $k \geqslant K$ then the word contains a loop. The value of K depends only on the number of letters of the alphabet.*

Proof. The proof is by induction on the length L of the alphabet $\mathcal{V}$ (cf. Remark (1) after the definition of loop).

For $L = 1$ the assertion is trivially satisfied. Assume that for given L there exists an integer $K(L)$ such that any word of length $K(L)$ containing at most L distinct letters has a loop: we want to show that then if the alphabet has $L + 1$ letters there exists $K(L + 1)$ such that any word of the alphabet with length $K(L + 1)$ has also a loop.

Let $N(L)$ be the number of words of length $K(L)$ written with the letters of an alphabet $\mathcal{V}$ with $|\mathcal{V}| = L + 1$. Consider a word $a = a_1 \ldots a_{N(L)+1}$, where each a_k has length $K(L)$. We want to show by contradiction that a contains a loop. If this is not the case, by the inductive assumption for each k either a_k contains a loop or it must contain all the $L + 1$ letters. As all words a_k have length $K(L)$ and there are $N(L) + 1$ of them, at least two words, say a_i and a_j with $i < j$, must be equal to each other. Therefore we can write $a = a_1 \ldots a_{i-1} a_i b a_i a_{j+1} \ldots a_{N(L)+1}$, where $b = a_{i+1} \ldots a_{j-1}$ if $j > i + 1$ and $b = \emptyset$ if $j = i + 1$. Hence a contains the string $a_i b a_i$, with a_i containing all the letters. Hence by Lemma A.4.2 one has $w(a_i b) = 0$, i.e. $a_i b$ is a loop. □

Remark. Note that the proof of Lemma A.4.3 implies

$$K(L + 1) \leqslant K(L)\big(N(L) + 1\big) \leqslant \prod_{\ell=1}^{L} \big(N(\ell) + 1\big), \tag{A.4.2}$$

which provides a bound on the maximal length of the chains in terms of the length of the alphabet $\mathcal{V}$.

Lemma 8.2 follows immediately from the results above, by noting that all the spheres with diameter a vector $v(1) - v(2)$ with $v \in \mathcal{V}$ are inside a compact ball of $\mathbb{Z}^D$.

A.5. Invertibility of J for $D = 2$

In the following we assume $D = 2$ and $N > 4$. We prove the Sub-lemma of Lemma 8.3.

We first prove that (i) implies (ii). As seen in Appendix A.3 condition (8.4) is implied by

$$\alpha_i \leqslant \left(\frac{|m_i|}{|m_1|}\right)^{2+2s} \leqslant \alpha_{i+1} \quad \forall i = 2, \ldots, N-1, \tag{A.5.1}$$

where the $\alpha_i > 1$ are fixed in Appendix A.3.

For $|m_1|$ large enough, (A.5.1) contains a $2N$-dimensional ball of arbitrarily large radius. By definition an algebraic variety is the set of solutions of some polynomial equations and therefore cannot contain all the positive integer points of a ball provided the radius is large enough (depending on the degree of the polynomial).

To prove (i) let us start with some notations. We consider $\mathbb{Z}^{2N}$ as a lattice in $\mathbb{C}^{2N}$, we denote $x = \{x_1, \ldots, x_N\} \equiv \mathcal{M}_+ \in \mathbb{C}^{2N}$, where each x_i is a point in $\mathbb{C}^2$; we denote the points in $\mathcal{M}$ still as $m_i \in \mathbb{C}^2$, and for each point $x_i \in \mathcal{M}_+$ we have the orbit $\mathcal{O}(x_i) \in \mathcal{M}$ i.e. the four points in $\mathcal{M}$ obtained by changing the signs of the components of x_i.

Definition A.5.1.

(i) Given two points m_i, m_j in $\mathcal{M}$ we consider: the circle with diameter $m_i - m_j$ (curve of type 1) and the two lines orthogonal to $m_i - m_j$ and passing respectively through m_i (curve of type 2) and through m_j (curve of type 3). Note that the curve is identified by the couple (m_i, m_j) and by the type label. We call $\mathcal{C}$ the finite set of distinct curves obtained in this way for all couples $m_i \neq m_j$ in $\mathcal{M}$.

(ii) Let C be a curve in $\mathcal{C}$ identified by the couple (m_i, m_j). We say that a point m' is g-linked by (m_i, m_j) to $m \in C$ if one has either (1) $m' = -m + m_i + m_j$, if C is a curve of type 1, or (2) $m' = m + (m_j - m_i)$, if C is a curve of type 2, or (3) $m' = m - (m_j - m_i)$, if C is a curve of type 3. Notice that in case (1) also m' is on the circle, while in cases (2) and (3) m' is on a curve of type 3 and 2, respectively. We say that two points $m, m' \in \mathbb{Z}^2_+$ are linked by (m_i, m_j) if there are two points $\bar{m} \in \mathcal{O}(m)$ and $\bar{m}' \in \mathcal{O}(m')$ such that $\bar{m}, \bar{m}'$ are g-linked by (m_i, m_j).

(iii) Given $\mathcal{M}_+$ we consider the set H of points $y_j \notin \mathcal{M}$ which lie on the intersection of two curves in C, counted with their multiplicity. Set $r := |H|$: we denote the list of intersection points as $y = \{y_1, \ldots, y_{r(N)}\}$. Note that r depend only on N.

We first prove that the points $x \in \mathbb{C}^{2N}$ which do not satisfy Lemma 8.1 lie on an algebraic variety. As seen in Appendix A.3, Lemma 8.1 is verified by requiring that if either a curve of type 1 contains three points in $\mathcal{M}$ or a curve of type 2 or 3 contains two points in $\mathcal{M}$, then such points are on the same orbit. It is clear (see Appendix A.3) that this condition can be achieved by requiring that x does not belong to some proper algebraic variety, say $\mathcal{W}_a$, in $\mathbb{C}^{2N}$.

Let us now consider the set of points $x \in \mathbb{Z}^{2N}_+$ where $\det J_{1,1}$ is identically equal to zero (as a function of s); since $J_{1,1}$ is a block diagonal matrix we factorise the single blocks and treat them separately. The matrix $J_{1,1}$ has some simple blocks which we can describe explicitly. Recall that

$$16A^2 = c_1 \sum_{i=1}^{N} |x_i|^2, \qquad 2a^2_{x_i} = (1 - c_1)|x_i|^2 - c_1 \sum_{\substack{j=1,\ldots,N \\ j \neq i}} |x_j|^2, \tag{A.5.2}$$

where $c_1 = 8/(8N + 1)$.

1. For all $m \in \mathbb{Z}_+^2$ such that m does not belong to any curve $C \in \mathcal{C}$ one has $Y_{m,m'} = 0$ for all m'; by considering the limit $s \to \infty$ one can easily check that $J_{m,m} = |m|^{2+2s}/2 - 8A^2 = 0$ is never an identity in s (independently of the choice of $\mathcal{M}_+$).
2. For all linked couples $m, m' \in \mathbb{Z}_+^2$ such that each point belongs to one and only one curve one has either a diagonal block $|m|^{2+2s}/2 - 8A^2 - 4a_{x_i}^2$ for some $x_i \in \mathcal{M}_+$ if $m = m'$, or a 2×2 matrix

$$\begin{pmatrix} |m|^{2+2s}/2 - 8A^2 & -2a_{m_i}a_{m_j} \\ -2a_{m_i}a_{m_j} & |m'|^{2+2s}/2 - 8A^2 \end{pmatrix}$$

   if $m \neq m'$ and (m_i, m_j) is the couple linking m' to m. In both cases a trivial check of the limit $s \to \pm\infty$ will ensure that the determinant is not identically null.
3. There is a block matrix containing all and only the elements of $\mathcal{M}_+$. Such a matrix is easily obtained by differentiating the left-hand side of (8.5):

$$-2\begin{pmatrix} a_{m_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{m_N} \end{pmatrix}\begin{pmatrix} 9 & 8 & \dots & 8 \\ 8 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 8 \\ 8 & \dots & 8 & 9 \end{pmatrix}\begin{pmatrix} a_{m_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{m_N} \end{pmatrix}.$$

   Since all the a_{m_i} are nonzero we only need to prove that the matrix in the middle is invertible, which is trivially true since the determinant is an odd integer.

We now have considered all those blocks in $J_{1,1}$ whose invertibility can be easily checked directly. We are left with the intersection points in $H' := H \cap \mathbb{Z}_+^2 \setminus \mathcal{M}_+$ and all those points m' which are linked to some $y_j \in H'$. We call $\tilde{J}$ the restriction of $J_{1,1}$ to such points; the crucial property of $\tilde{J}$ is that it is a $K \times K$ matrix with K bounded by above by some constant depending only on N.

We will impose that $\tilde{J}$ is invertible at $s = 0$ by requiring that $\mathcal{M}_+$ does not lie on an appropriate algebraic variety in $\mathbb{C}^{2N}$.

By definition the points in H (and the points linked to them) are algebraic functions of $x \in \mathbb{C}^{2N}$. By construction $\tilde{J}_{m,m} - Y_{m,m} = |m|^{2+2s}/2 - 8A^2$ and moreover $Y_{m,m'}$ contains a contribution $-2a_{m_i}a_{m_j}$ for each couple (m_i, m_j) linking m' to m. We want to prove that for $s = 0$ the equation $\det \tilde{J} = 0$ (which is an equation for $x \in \mathbb{C}^{2N}$) defines a proper algebraic variety, say $\mathcal{W}_f$, in $\mathbb{C}^{2N}$.

We consider the space $\mathbb{C}^T := \mathbb{C}^{2N} \times \mathbb{C}^N \times \mathbb{C}^{2r}$ and, with an abuse of notation, we denote the generic point in $\mathbb{C}^T$ by $(x, a, y) = (x_1, \dots, x_N, a_{x_1}, \dots, a_{x_N}, y_1, \dots, y_r)$ (therefore we consider (x, a, y) as independent variables). Note that $\det \tilde{J} = 0$ is a polynomial equation in $\mathbb{C}^T$. We call $\mathcal{W}_b$ the algebraic variety defined by requiring both that the a_{x_i} satisfy (A.5.2) and that each y_j lies on at least two curves of C ($\mathcal{W}_b$ is equivalent to a finite number of copies of $\mathbb{C}^{2N}$).

We now recall a standard theorem in algebraic geometry which states: *Let W be an algebraic variety in $\mathbb{C}^{n+m}$ and let Π be the projection $\mathbb{C}^{n+m} \to \mathbb{C}^n$ then $\overline{\Pi(W)}$ is an algebraic variety* (clearly it may be the whole $\mathbb{C}^n$!) We set $n = 2N$ (the first $2N$ variables), $m = 2r + N$ and apply the stated theorem to $\overline{\Pi(\mathcal{W}_b \cap \mathcal{W}_f)}$; we now only need to prove that the algebraic variety we have obtained is proper; to do so it is convenient to treat separately the invertibility conditions of each single block of $\tilde{J}$.

The first step is to simplify as far as possible the structure of the intersections and therefore of the matrix $\tilde{J}$. The simplest possible block involving an intersection point y_j is such that

(i) only two curves in $\mathcal{C}$ pass through y_j;
(ii) the two points linked to y_j (by the couples of points in $\mathcal{M}$ identifying the curves) are not intersection points.

Such a configuration gives either a 3×3 matrix or a 2×2 matrix – if one of the curves is either an horizontal or vertical line or a circle centred at the origin.

Definition A.5.2. We say that a curve $C \in \mathcal{C}$ depends on the two – possibly equal – variables $x_i, x_j \in \mathbb{C}^2$ if C is identified by the couple (m_i, m_j), such that $m_i \in \mathcal{O}(x_i)$ and $m_j \in \mathcal{O}(x_j)$.

The negation of (i) is that y_j is on (at least) three curves of $\mathcal{C}$: such a condition defines a proper algebraic variety in $\mathbb{C}^T$, say $\mathcal{W}_j$. We now consider the projection of $\mathcal{W}_b \cap \mathcal{W}_j$ on $\mathbb{C}^{2N}$: its closure is an algebraic variety and either it is proper or the triple intersection occurs for any choice of x (which unfortunately can indeed happen due to the symmetries introduced by the Dirichlet boundary conditions).

Three curves in $\mathcal{C}$ depend on at most six variables in $\mathbb{C}^2$. If four or more of such variables are different then at least one variable, say x_k, appears only once. By moving x_k in $\mathbb{C}^2$ we can move arbitrarily one of the curves, while the other two (which do not depend on x_k) remain fixed. This implies that the triple intersection cannot hold true for all values of x_k and thus $\overline{\Pi(\mathcal{W}_b \cap \mathcal{W}_j)}$ is a proper variety in $\mathbb{C}^{2N}$.

In the same way the negation of (ii) is that one point linked to y_j lies on (at least) two curves of $\mathcal{C}$ (one curve is fixed by the fact that the point is linked to y_j); again the intersection is determined by six points in $\mathcal{M}$ and the same reasoning holds.

We call $\mathcal{W}_c$ the variety in $\mathbb{C}^T$ defined by the union of all those $\mathcal{W}_j$ such that $\overline{\Pi(\mathcal{W}_b \cap \mathcal{W}_j)}$ is proper.

In $\mathcal{W}_b \setminus \mathcal{W}_c$ we can now classify the possible blocks appearing in $\tilde{J}$ (notice that only intersection points which are integer-valued have to be taken into account when constructing the blocks in $\tilde{J}$).

1. We have a list of at most 3×3 blocks corresponding to the intersection points of type (i)–(ii). Such intersection points are identified by two curves which can depend on at most four different variables x_{i_k} with $k = 1, \dots, 4$.
2. There are more complicated blocks corresponding to multiple intersections (or intersection points linked to each other), which occur for all $x \in \mathbb{C}^{2N}$ due to symmetry. As we have proved above the curves defining such intersections depend on at most three different variables x_{i_k}.

In any given block, call it B_h, the contribution from Y involves only terms of the form $-2a_{m_i}a_{m_j}$ such that $m_i, m_j \in \bigcup_{k=1}^4 \mathcal{O}(x_{i_k})$. Each a_{m_j} depends on all the components of x; in particular, $a_{m_j}^2$ can be written as a term depending only on the x_{i_k} plus the term $-\frac{1}{2}c_1 \sum_{j \neq i_1, \dots, i_4} x_j^2$. Since by hypothesis $N > 4$ and $k \leqslant 4$ the second sum is surely nonempty.

Finally one has the diagonal contributions (from $J - Y$): $|y_j|^2 - \frac{1}{2}c_1 \sum_{j \neq i_1, \dots, i_4} x_j^2 + z$, where z is a polynomial function in the x_{i_k}'s.

In the limit $\sum_{j\neq i_1,\dots,i_4} x_j^2 \to \infty$ the terms depending on the x_{i_k}'s become irrelevant and we are left with a matrix (of unknown size) whose entries, apart from the common factor $-\frac{1}{2}c_1 \sum_{j\neq i_1,\dots,i_4} x_j^2$, are integer numbers. It is easily seen that these numbers are odd on the diagonal, while all the off-diagonal terms are even; indeed Y contributes only even entries while $J-Y$ is diagonal and odd due to the term $8A^2$. Thus the determinant (apart form the common factors) is odd and hence the equation $\det B_h = 0$ is not an identity on $\mathcal{W}_b$. If we call $\mathcal{W}_h$ the variety in $\mathbb{C}^T$ defined by $\det B_h = 0$ then $\overline{\Pi(\mathcal{W}_b \cap \mathcal{W}_h)}$ is surely proper. Finally we call $\mathcal{W}_d$ the union of all the $\mathcal{W}_h$ and set $\mathcal{W}_f = \mathcal{W}_d \cup \mathcal{W}_c \cup \mathcal{W}_c$.

A.6. Proof of the separation Lemma 8.4

The following proof is adapted from [9]. Given $\varepsilon > 0$ define $\delta = \delta(\varepsilon, D) = \varepsilon/2^{D-1} \cdot D!(D+1)!$. Then Lemma 8.4 follows from the results below.

Lemma A.6.1. *Let $x \in \mathbb{R}^d$. Assume that there exist d vectors $\Delta_1, \dots, \Delta_d$, which are linearly independent in $\mathbb{Z}^d$, and such that $|\Delta_k| \leqslant A_1$ and $|x \cdot \Delta_k| \leqslant A_2$ for all $k = 1, \dots, d$. Then $|x| \leqslant C(d) A_1^{d-1} A_2$ for some constant $C(d)$ depending only on d.*

Proof. Call $\beta_k \in [0, \pi/2]$ the angle between Δ_k and the direction of the vector x. Without any loss of generality we can assume $\beta_k \geqslant \beta_d$ for all $k = 1, \dots, d-1$. Set $\beta'_d = \pi/2 - \beta_d$. One has $\beta'_d > 0$ because $\Delta_1, \dots, \Delta_d$ are linearly independent.

Consider the simplex generated by the vectors $\Delta_1, \dots, \Delta_d$. By the fact 2 in the proof of Lemma 8.4 one has, for some d-dependent constant $C(d)$,

$$1 \leqslant C(d) |\Delta_1| |\Delta_2| \cdots |\Delta_d| |\sin \alpha_{1,2}| |\sin \alpha_{12,3}| \cdots |\sin \alpha_{1\dots(d-1),d}|, \tag{A.6.1}$$

where $\alpha_{1\dots(j-1),j}$, $j \geqslant 2$, is the angle between the vector Δ_j and the plane generated by the vectors $\Delta_1, \dots, \Delta_{j-1}$. Hence

$$1 \leqslant C(d) A_1^{d-1} |\Delta_d| |\sin \alpha_{1\dots(d-1),d}|. \tag{A.6.2}$$

Moreover one has

$$|x \cdot \Delta_d| = |x| |\Delta_d| |\cos \beta_d| = |x| |\Delta_d| |\sin \beta'_d| \geqslant |x| |\Delta_d| |\sin \alpha_{1\dots(d-1),d}|, \tag{A.6.3}$$

so that, from (A.6.2) and (A.6.3), we obtain $|x| A_1^{-(d-1)} \leqslant C(d) A_2$, so that the assertion follows. □

Lemma A.6.2. *There exist constants C and C' such that the following holds. Let $n_1, \dots, n_k \in \mathbb{Z}^D$ be a sequence of distinct elements such that $|\Phi(n_j) - \Phi(n_{j+1})| \leqslant C r^\delta$. Then $k \leqslant C' \max\{r^\varepsilon, D+2\}$.*

Proof. Since the vectors n_j are on the lattice $\mathbb{Z}^D$ there exist a constant $C_1(D)$ and $j_0 \leqslant k/2$ such that $|n_{j_0}| > C_1(D) k^{1/D}$. Set $\Delta_j = n_j - n_{j_0}$. By assumption one has $|\Phi(n_j) - \Phi(n_{j+1})| \leqslant C r^\delta$, hence $|\Phi(n_j) - \Phi(n_{j_0})| \leqslant C(j - j_0) r^\delta$ for all $j_0 + 1 \leqslant j \leqslant k$. Then $|\Phi(n_j) - \Phi(n_{j_0})| \leqslant A_1 := C J_1 r^\delta$ for all $j_0 + 1 \leqslant j \leqslant j_0 + J_1$. Fix $J_1 = k^{1/\alpha(D)}$, with $\alpha(n) = 2n(n+1)$. By using that

$$\Phi(n_j) - \Phi(n_{j_0}) = \left(\Delta_j, 2\Delta_j \cdot n_{j_0} + |\Delta_j|^2\right), \tag{A.6.4}$$

we find $|\Delta_j| \leqslant A_1$ and $|n_{j_0} \cdot \Delta_j| \leqslant A_2 := A_1^2$ for all $j_0 + 1 \leqslant j \leqslant j_0 + J_1$.

If $\mathrm{Span}\{\Delta_{j_0+1}, \dots, \Delta_{j_0+J_1}\} = D$ then by Lemma A.6.1 one has $|n_{j_0}| \leqslant C(D)A_1^{D+1}$. Then, for this relation to be not in contradiction with $|n_{j_0}| > C_1(D)k^{1/D}$, we must have $C_1(D)k^{1/D} < C(D)A_1^{D+1}$, hence $k \leqslant C_2(D)r^{\alpha(D)\delta}$ for some constant $C_2(D)$.

If $\mathrm{Span}\{\Delta_{j_0+1}, \dots, \Delta_{j_0+J_1}\} \leqslant D - 1$ then there exists a subspace H_1 with $\dim(H_1) = D - 1$ such that $n_j \in n_{j_0} + H_1$ for $j_0 + 1 \leqslant j \leqslant j_0 + J_1$. Choose $j_1 \leqslant J_1/2$ such that $P_{H_1} n_{j_1} := n_{j_1} - n_{j_0} \in H_1$ satisfies $|P_{H_1} n_{j_1}| > C(D-1)J_1^{1/(D-1)}$, and fix $J_2 = J_1^{1/\alpha(D-1)}$. Redefine $\Delta_j = n_j - n_{j_1}$ for $j \geqslant j_1 + 1$, $A_1 = CJ_2 r^\delta$ and $A_2 = A_1^2$: by reasoning as in the previous case we find again $|\Delta_j| \leqslant A_1$ and $|n_{j_1} \cdot \Delta_j| \leqslant A_2$ for all $j_0 + 1 \leqslant j \leqslant j_0 + J_1$.

If $\mathrm{Span}\{\Delta_{j_1+1}, \dots, \Delta_{j_1+J_2}\} = D - 1$ then by Lemma A.6.1 one has $|n_{j_1}| \leqslant C(D-1)A_1^D$, which implies $C_1(D-1)J_1^{1/D-1} < C(D)A_1^D$. By using the new definition of A_1, we obtain $J_1 \leqslant C_2(D-1)r^{\alpha(D-1)\delta}$, hence $k \leqslant C_3(D)r^{\alpha(D-1)\alpha(D)\delta}$ for some other constant $C_3(D)$.

If $\mathrm{Span}\{\Delta_{j_1+1}, \dots, \Delta_{j_0+J_2}\} \leqslant D - 2$ then there exists a subspace H_2 with $\dim(H_1) = D - 1$ such that $n_j \in n_{j_1} + H_2$ for $j_1 + 1 \leqslant j \leqslant j_1 + J_2$. Then we iterate the construction until either we find $k \leqslant C_{n+2}(D)r^{\alpha(D-1)\dots\alpha(D-n)\delta}$ for some $n \leqslant D - 1$ and some constant $C_{n+2}(D)$ or we arrive at a subspace H_{D-1} with $\dim(H_{D-1}) = 1$.

In the last case the vectors $\Delta_{j_{D-2}+1}, \dots, \Delta_{j_{D-2}+J_{D-1}}$, with $J_{D-1} = J_{D-2}^{1/\alpha(2)}$, are linearly dependent by construction, so that they lie all on the same line. Therefore, we can find at least $J_{D-1}/2$ of them, say the first $J_{D-1}/2$, with decreasing distance from the origin. If we set $n_{j_{D-2}+1} = a$, $n_{j_{D-2}+J_{D-1}/2} = b$, and $n_{j_{D-2}+1} - n_{j_{D-2}+J_{D-1}}/2 = c$, and sum over $j_{D-2} + 1 \leqslant j \leqslant j_{D-2} + J_{D-1}/2$ the inequalities

$$|n_j - n_{j-1}| + |n_j|^2 - |n_{j-1}|^2 \leqslant \text{const.}\left|\Phi(n_j) - \Phi(n_{j-1})\right| \leqslant \text{const.}\, Cr^\delta, \tag{A.6.5}$$

we obtain

$$|c| + |c|^2 \leqslant |c| + |a|^2 - |b|^2 \leqslant \text{const.}\, Cr^\delta \frac{J_{D-1}}{2}, \tag{A.6.6}$$

where $|c| \geqslant J_{D-1}/2$. Hence $J_{D-1} \leqslant (Cr^\delta)^2$.

By collecting together all the bound above we find $k \leqslant C_D(D)r^{2\alpha(D)\dots\alpha(2)\delta}$, so that, by defining $C' = C_D(D)$ and using that $\varepsilon/\delta = \alpha(D)\dots\alpha(2) = 2^{D-1}D!(D+1)!$, the assertion follows. □

Lemma A.6.3. *There exist constants ε', δ', C and C' such that the following holds. Given $n_0 \in \mathbb{Z}^D$ there exists a set $\Delta \subset \mathbb{Z}^D$, with $n_0 \in \Delta$, such that* $\mathrm{diam}(\Delta) < C'r^{\varepsilon'}$ *and $|\Phi(x) - \Phi(y)| > C'r^{\delta'}$ for all $x \in \Delta$ and $y \notin \Delta$.*

Proof. Cf. [9, p. 399], which proves the assertion with $\varepsilon' = \delta + \varepsilon$ and $\delta' = \delta = \delta(\varepsilon, D)$. □

Lemma A.6.4. *Let Δ be as in Lemma* A.6.3. *There exists a constant C'' such that one has $|\Delta| \leqslant C''r^{D(\varepsilon+\delta)}$.*

Proof. The bound follows from Lemma A.6.3 and from the fact that $\mathrm{diam}(\Delta) < C'r^\varepsilon$, by using that the points in Δ are distinct lattice points in $\mathbb{R}^D$. □

References

[1] D. Bambusi, M. Berti, A Birkhoff–Lewis-type theorem for some Hamiltonian PDEs, SIAM J. Math. Anal. 37 (1) (2005) 83–102.

[2] D. Bambusi, S. Paleari, Families of periodic solutions of resonant PDEs, J. Nonlinear Sci. 11 (1) (2001) 69–87.

[3] M. Berti, Ph. Bolle, Periodic solutions of nonlinear wave equations with general nonlinearities, Comm. Math. Phys. 243 (2) (2003) 315–328.

[4] M. Berti, Ph. Bolle, Cantor families of periodic solutions for completely resonant nonlinear wave equations, Duke Math. J. 134 (2) (2006) 359–419.

[5] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, Int. Math. Res. Not. 11 (1994), 475ff., approx. 21 pp. (electronic).

[6] J. Bourgain, Construction of periodic solutions of nonlinear wave equations in higher dimension, Geom. Funct. Anal. 5 (1995) 629–639.

[7] J. Bourgain, Nonlinear Schrödinger equations, in: Hyperbolic Equations and Frequency Interactions, Park City, UT, 1995, in: IAS/Park City Math. Ser., vol. 5, Amer. Math. Soc., Providence, RI, 1999, pp. 3–157.

[8] J. Bourgain, Periodic solutions of nonlinear wave equations, in: Harmonic Analysis and Partial Differential Equations, Chicago, IL, 1996, in: Chicago Lectures in Math., University of Chicago Press, Chicago, IL, 1999, pp. 69–97.

[9] J. Bourgain, Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. of Math. (2) 148 (2) (1998) 363–439.

[10] J. Bourgain, Green's Function Estimates for Lattice Schrödinger Operators and Applications, Ann. of Math. Stud., vol. 158, Princeton Univ. Press, Princeton, NJ, 2005.

[11] L. Chierchia, C. Falcolini, A direct proof of a theorem by Kolmogorov in Hamiltonian systems, Ann. Sc. Norm. Super. Pisa Cl. Sci. (4) 21 (4) (1994) 541–593.

[12] B. Connes, Sur les coefficients des séries trigonométriques convergentes sphériquement, C. R. Acad. Sci. Paris Sér. A–B 283 (4) (1976) A159–A161.

[13] W. Craig, C.E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, Comm. Pure Appl. Math. 46 (1993) 1409–1498.

[14] L.H. Eliasson, Absolutely convergent series expansions for quasi periodic motions, Math. Phys. Electron. J. 2 (1996), Paper 4, 33 pp. (electronic).

[15] L.H. Eliasson, S. Kuksin, KAM for nonlinear Schrödinger equation, preprint, 2006, Ann. of Math., in press.

[16] A.E. Filippov, Nonlinear nonlocal Schrödinger equation in the context of quantum mechanics, Phys. Lett. A 215 (1–2) (1996) 32–39.

[17] G. Gallavotti, Twistless KAM tori, Comm. Math. Phys. 164 (1) (1994) 145–156.

[18] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher-dimensional spaces, Comm. Math. Phys. 262 (2) (2006) 343–372.

[19] G. Gentile, V. Mastropietro, Construction of periodic solutions of the nonlinear wave equation with Dirichlet boundary conditions by the Lindstedt series method, J. Math. Pures Appl. (9) 83 (8) (2004) 1019–1065.

[20] G. Gentile, V. Mastropietro, M. Procesi, Periodic solutions for completely resonant nonlinear wave equations with Dirichlet boundary conditions, Comm. Math. Phys. 256 (2) (2005) 437–490.

[21] G. Gentile, M. Procesi, Conservation of resonant periodic solutions for the one-dimensional nonlinear Schrödinger equation, Comm. Math. Phys. 262 (3) (2006) 533–553.

[22] G. Gentile, M. Procesi, Periodic solutions for a class of nonlinear partial differential equations in higher dimension, preprint, 2008.

[23] F. Harary, E.M. Palmer, Graphical Enumeration, Academic Press, New York, London, 1973.

[24] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, Berlin, New York, 1976.

[25] S.B. Kuksin, Nearly Integrable Infinite-Dimensional Hamiltonian Systems, Lecture Notes in Math., vol. 1556, Springer-Verlag, Berlin, 1994.

[26] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. (2) 143 (1) (1996) 149–179.

[27] B.V. Lidskiĭ, E.I. Shul'man, Periodic solutions of the equation $u_{tt} - u_{xx} + u^3 = 0$, Funct. Anal. Appl. 22 (4) (1988) 332–333.

[28] Y. Matsuno, Asymptotic solutions of the nonlocal nonlinear Schrödinger equation in the limit of small dispersion, Phys. Lett. A 309 (1–2) (2003) 83–89.

[29] D.E. Pelinovsky, R.H.J. Grimshaw, Nonlocal models for envelope waves in a stratified fluid, Stud. Appl. Math. 97 (4) (1996) 369–391.

[30] J. Pöschel, On the construction of almost periodic solutions for a nonlinear Schrödinger equation, Ergodic Theory Dynam. Systems 22 (5) (2002) 1537–1549.
[31] M. Procesi, Quasi-periodic solutions for completely resonant nonlinear wave equations in 1D and 2D, Discrete Contin. Dyn. Syst. 13 (3) (2005) 541–552.
[32] C.E. Wayne, Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory, Comm. Math. Phys. 127 (3) (1990) 479–528.
[33] H. Whitney, Analytic extensions of differential functions defined in closed sets, Trans. Amer. Math. Soc. 36 (1) (1934) 63–89.
[34] X. Yuan, A KAM theorem with applications to partial differential equations of higher dimension, Comm. Math. Phys. 275 (1) (2007) 97–137.
[35] V. Zharnitsky, E. Grenier, Ch.K.R.T. Jones, S.K. Turitsyn, Stabilizing effects of dispersion management, Phys. D 152/153 (2001) 794–817.
[36] O.V. Zhdaneev, G.N. Serezhnikov, A.Yu. Trifonov, A.V. Shapovalov, Semiclassical trajectory-coherent states of the nonlinear Schrödinger equation with unitary nonlinearity, Russian Phys. J. 42 (7) (1999) 598–606.