Nonlinearity

PAPER

Quasi-periodic solutions for the forced Kirchhoff equation on $\mathbb{T}^d$

To cite this article: Livia Corsi and Riccardo Montalto 2018 *Nonlinearity* **31** 5075

IOP Publishing | London Mathematical Society — Nonlinearity

Nonlinearity **31** (2018) 5075–5109 https://doi.org/10.1088/1361-6544/aad6fe

Quasi-periodic solutions for the forced Kirchhoff equation on $\mathbb{T}^d$

Livia Corsi[1] **and Riccardo Montalto**[2]

[1] School of Mathematics, Georgia Institute of Technology, 686 Cherry St. NW, Atlanta, GA 30332, United States of America

[2] Institut für Mathematik, Universität Zürich Winterthurerstrasse 190, CH-8057 Zürich, CH, Switzerland

E-mail: lcorsi6@math.gatech.edu and riccardo.montalto@math.uzh.ch

Received 21 February 2018, revised 6 July 2018
Accepted for publication 31 July 2018
Published 10 October 2018

CrossMark

Recommended by Dr C Eugene Wayne

Abstract

In this paper we prove the existence of small-amplitude quasi-periodic solutions with Sobolev regularity, for the d-dimensional forced Kirchhoff equation with periodic boundary conditions. This is the first result of this type for a quasi-linear equation in high dimension. The proof is based on a Nash–Moser scheme in Sobolev class and a regularization procedure combined with a multiscale analysis in order to solve the linearized problem at any approximate solution.

Keywords: Kirchhoff equation, quasi-linear PDEs, quasi-periodic solutions, infinite-dimensional dynamical systems, Nash–Moser theory
Mathematics Subject Classification numbers: 37K55, 35L72

1. Introduction and main result

In this paper we consider the forced Kirchhoff equation on the d-dimensional torus $\mathbb{T}^d$

$$\partial_{tt} v - \Big(1 + \int_{\mathbb{T}^d} |\nabla v|^2 \, \mathrm{d}x\Big) \Delta v = \delta f(\omega t, x) \tag{1.1}$$

where $\delta > 0$ is a small parameter, $\omega := \lambda \bar{\omega} \in \mathbb{R}^\nu$, $\lambda \in \mathcal{I} := [1/2, 3/2]$, $\bar{\omega}$ a fixed diophantine vector, i.e.

$$|\bar{\omega} \cdot \ell| \geqslant \frac{\gamma_0}{|\ell|^\nu}, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}, \tag{1.2}$$

and $f : \mathbb{T}^\nu \times \mathbb{T}^d \to \mathbb{R}$ is a sufficiently smooth function with zero average, i.e.

$$\int_{\mathbb{T}^{\nu+d}} f(\varphi,x)\,\mathrm{d}\varphi\,\mathrm{d}x = 0. \tag{1.3}$$

Following [11, 14, 20] we assume also

$$\Big|\sum_{1\leqslant i,j\leqslant\nu} \overline{\omega}_i\overline{\omega}_j p_{ij}\Big| \geqslant \frac{\gamma_0}{|p|^{\nu(\nu+1)}}, \qquad \forall p\in\mathbb{Z}^{\nu(\nu+1)/2}\setminus\{0\}. \tag{1.4}$$

Rescaling $v\mapsto\delta^{\frac{1}{3}}v$, we see that (1.1) takes the form

$$\partial_{tt}v - \Big(1+\varepsilon\int_{\mathbb{T}^d}|\nabla v|^2\,\mathrm{d}x\Big)\Delta v = \varepsilon f(\omega t,x), \quad \varepsilon := \delta^{\frac{2}{3}}. \tag{1.5}$$

Our aim is to prove the existence of quasi-periodic solutions of (1.5) for ε small enough and λ in a large subset of parameters in $\mathcal{I}$. Since ω is nonresonant, finding a quasi-periodic solution with frequency ω is equivalent to find a torus embedding $\varphi\mapsto u(\varphi,\cdot)$ satisfying the equation $F(v)=0$ where

$$F(v)\equiv F(\lambda,v) := (\lambda\bar{\omega}\cdot\partial_\varphi)^2 v - \Big(1+\varepsilon\int_{\mathbb{T}^d}|\nabla v|^2\,\mathrm{d}x\Big)\Delta v - \varepsilon f(\varphi,x) \tag{1.6}$$

acting on the scale of real Sobolev spaces

$$H^s = H^s(\mathbb{T}^{\nu+d}) := \Big\{v(\varphi,x) = \sum_{\substack{\ell\in\mathbb{Z}^\nu\\ j\in\mathbb{Z}^d}} v_{\ell,j}\mathrm{e}^{\mathrm{i}\ell\cdot\varphi}\mathrm{e}^{\mathrm{i}j\cdot x}\in L^2(\mathbb{T}^{\nu+d}) : \|v\|_s^2 := \sum_{\substack{\ell\in\mathbb{Z}^\nu\\ j\in\mathbb{Z}^d}}\langle\ell,j\rangle^{2s}|v_{\ell,j}|^2 < +\infty\Big\} \tag{1.7}$$

where $\langle\ell,j\rangle := \max\{1,|\ell|,|j|\}$. Our main result is the following.

Theorem 1.1. *There exists* $\overline{q} := \overline{q}(\nu,d) > 0$ *such that for all* $q\geqslant\overline{q}$ *and any* $f\in\mathcal{C}^q(\mathbb{T}^\nu\times\mathbb{T}^d)$ *satisfying (1.3) there exist* $s_1 = s_1(\nu,d,q) > 0$*, increasing in q,* $\varepsilon_0=\varepsilon_0(f,\nu,d)>0$ *and for any* $\varepsilon\in(0,\varepsilon_0)$ *a Borel set* $\mathcal{C}_\varepsilon\subseteq\mathcal{I}$ *with asymptotically full Lebesgue measure i.e.*

$$\lim_{\varepsilon\to 0}\mathrm{meas}(\mathcal{C}_\varepsilon) = 1$$

and there exists a function $u_\varepsilon\in\mathcal{C}^1\Big(\mathcal{I},H^{s_1}(\mathbb{T}^\nu\times\mathbb{T}^d)\Big)$ *such that for any* $\lambda\in\mathcal{C}_\varepsilon$*,* $u_\varepsilon(\lambda)$ *is a zero for the functional F appearing in (1.6). Finally, possibly for larger q, there exists* ε_1 *possibly smaller than* ε*, and for all* $\varepsilon\in(0,\varepsilon_1)$ *a Borel set* $\mathcal{O}_\varepsilon\subseteq\mathcal{C}_\varepsilon$ *with asymptotically full Lebesgue measure such that for any* $\lambda\in\mathcal{O}_\varepsilon$ *the found solution is linearly stable.*

The Kirchhoff equation has been introduced for the first time in 1876 by Kirchhoff in dimension 1, without forcing term and with Dirichlet boundary conditions, to describe the transversal free vibrations of a clamped string in which the dependence of the tension on the deformation cannot be neglected. It is a quasi-linear PDE, namely the nonlinear part of the equation contains as many derivatives as the linear differential operator.

Concernig the existence of periodic solutions, Kirchhoff himself observed the existence of a sequence of *normal modes*, namely solutions of the form $v(t,x)=v_j(t)\sin(jx)$ where $v_j(t)$ is 2π-periodic. Under the presence of the forcing term $f(t,x)$ the *normal modes* do not persist[3], since, expanding $v(t,x)=\sum_j v_j(t)\sin(jx)$, $f(t,x)=\sum_j f_j(t)\sin(jx)$, all the components $v_j(t)$ are coupled.

[3] This is true except in the case where f is uni-modal, i.e. $f(t,x)=f_k(t)\sin(kx)$ for some $k\geqslant 1$.

The existence of periodic solutions for the forced Kirchhoff equation in any dimension has been proved by Baldi in [2], while the existence of quasi-periodic solutions in one space dimension under periodic boundary conditions has been proved in [41].

Note that equation (1.5) is a quasi-linear PDE and it is well known that the existence of global solutions (even not periodic or quasi-periodic) for quasi-linear PDEs is not guaranteed, see for instance the non-existence results in [35, 37] for the equation $v_{tt} - a(v_x)v_{xx} = 0$, $a > 0$, $a(v) = v^p$, $p \geqslant 1$, near zero.

The existence of periodic solutions for wave-type equations with unbounded nonlinearities has been proved for instance in [19, 20, 44]. For the water waves equations, which are fully nonlinear PDEs, we mention [1, 31–33]; see also [3] for fully nonlinear Benjamin–Ono equations.

The methods developed in the above mentioned papers do not work for proving the existence of quasi-periodic solutions.

The existence of quasi-periodic solutions for PDEs with unbounded nonlinearities has been developed by Kuksin [36] for KdV and then Kappeler–Pöschel [34]. This approach has been improved by Liu–Yuan [38, 39] to deal with DNLS (derivative nonlinear Schrödinger) and Benjamin–Ono equations. These methods apply to dispersive PDEs like KdV, DNLS but not to derivative wave equation (DNLW) which contains first order derivatives in the nonlinearity. KAM theory for DNLW equation has been recently developed by Berti–Biasco–Procesi in [9, 10]. Such results are obtained via a KAM-like scheme which is based on the so-called *second Melnikov conditions* and provides also the *linear stability* of the solutions.

The existence of quasi-periodic solutions can be also proved by imposing only *first order Melnikov conditions* and the so-called *multiscale approach*. This method has been developed, for PDEs in higher space dimension, by Bourgain in [17, 18, 20] for analytic NLS and NLW, extending the result of Craig–Wayne [21] for 1D wave equation with bounded nonlinearity. Later, this approach has been improved by Berti–Bolle [11, 12] for NLW, NLS with differentiable nonlinearity and by Berti–Corsi–Procesi [14] on compact Lie-groups.

This method is especially convenient in higher space dimension since the second order Melnikov conditions are violated, due to the high multiplicity of the eigenvalues. The drawback is that the linear stability is not guaranteed. Indeed there are very few results concerning the existence and linear stability of quasi-periodic solutions in the case of multiple eigenvalues. We mention [15, 22] for the case of double eigenvalues and [24, 25] in higher space dimension.

All the aforementioned results concern *semi-linear* PDEs, namely PDEs in which the order of the nonlinearity is strictly smaller than the order of the linear part. For quasi-linear (either fully nonlinear) PDEs, the first KAM results have been proved by the *Italian team* in [4–7, 16, 26, 27, 30, 41].

To the best of our knowledge all the results for quasi-linear and fully nonlinear PDEs are only in one space dimension. The result proved in this paper is the first one concerning the existence of quasi-periodic solutions for a quasi-linear PDE in higher space dimension.

The reason why we achieve our result, whereas for other PDEs this is not possible (at least at the present time), is not merely technical and can be roughly explained as follows.

Almost all the literature about the existence of quasi-periodic solutions for dynamical systems in both finite and infinite dimension is ultimately related to a functional Newton scheme. It is well known that in the Newton scheme one has to solve the linearized problem, which in turn means that one has to invert the linearized functional. Such linearized functional is a liner operator acting on a scale of Hilbert spaces, hence one also needs appropriate bounds on the inverse in order to make the scheme convergent. Now, suppose that such linearized operator has the form $\mathcal{L} = \Delta + \varepsilon a(\varphi, x)\Delta$. In order to obtain bounds one wants to reduce this operator to constant coefficients up to a remainder (at least of order zero). Passing to the Fourier side in space, the corresponding symbol is given by $H(x,\xi) = |\xi|^2 + \varepsilon a(\varphi,x)|\xi|^2$ and hence reducing $\mathcal{L}$ to constant coefficients at leading order is equivalent to find a change of variables $(x,\xi) \mapsto (x',\xi')$ such that in the new variables the Hamiltonian $H(x,\xi)$ depends only on ξ'. In the 1D case this is always possible, whereas in dimension higher than one this is possible only in very special cases, due to the Poincaré 'triviality' theorem stating that generically a quasi-integrable Hamiltonian is not integrable; see for instance [29]. Of course there are some cases in which the Hamiltonian $H(x,\xi)$ is integrable (up to lower order terms); see for instance [8, 28, 42]. Indeed in these cases the *complete reduction to constant coefficients* is achieved. However the three papers [8, 28, 42] deal only with linear equations, whereas in the nonlinear case one has to fit the reducibility of the linearized operator with the Newton scheme. For instance, if in our case one tries to follow the above scheme and reduce completely the linearized operator (this is done in [42]), one obtains a bound on the inverse of the linearized operator $\mathcal{L}(u)$ of the form $\|\mathcal{L}(u)^{-1}h\|_s \lesssim_s \|h\|_{s+\sigma} + \|u\|_{2s+\sigma}\|h\|_{s_0+\sigma}$ for $s \geqslant s_0$, where σ is a constant depending only on ν and d. It is well known that a bound of this type is not enough for making the Newton scheme convergent; see [40].

In the present paper we overcome this difficulty as follows. First of all the highest order of our Hamiltonian symbol $H(x,\xi)$ does not depend on x so it is integrable; therefore we perform a reparametrization of time and we also apply a multiplication operator by a function depending only on time, and obtain a transformed operator of the form

$$(\omega \cdot \partial_\theta) - \mu\Delta + \mathcal{R}_2,$$

where μ is a constant ε-close to 1 and $\mathcal{R}_2$ is a bounded operator satisfying decay bounds; see (4.12) and (4.5). Then we do not attempt a reduction scheme for the lower order term $\mathcal{R}_2$ but rather use the multiscale approach. *A priori* this implies that we may not have informations about the linear stability of the solution we find; however the linear stability is obtained *a posteriori*, namely here we prove the existence, then by linearizing on the found solution one can apply theorem 1.2 of [42] and obtain the linear stability of the solution; see theorem 9.1 for details. An *a posteriori* approach of this type has been used for instance in [23] for the NLS on $SU(2)$, $SO(3)$.

Out of curiosity we finally note that our remainder $\mathcal{R}_2$ has a loss of regularity σ which is due to change of variables needed for the reduction up to order zero; see (4.5). We find it interesting that a similar loss of reguarity appears for semi-linear PDEs when the space variable lives on a compact Lie group instead of a torus; see (2.24c) in [14] where such loss is denoted by ν_0.

The paper is organized as follows. After reducing the problem to the zero mean value functions, we introduce the scale of Hilbert spaces and recall some of their properties. In section 4 we discuss some properties of the linearized operator $\mathcal{L}(u)$, and we reduce it to constant coefficients up to a remainder of order zero. We then discuss a Nash–Moser scheme converging on a set A_∞ defined in terms of the reduced operator, and which in principle might be empty. Afterwards in section 6 we introduce a subset $\mathcal{C}_\infty \subseteq A_\infty$ where the multiscale approach can be used. Finally we provide measure esitmates on another subest $\mathcal{C}_\varepsilon \subseteq \mathcal{C}_\infty$, defined in terms of the final solution only. The linear stability is obtained in section 9.

2. Reduction on the zero mean value functions

Following [41], we define the projectors $\Pi_0, \Pi_0^\perp$ as the orthogonal projections

$$\Pi_0 v := v_0(\varphi) = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} v(\varphi,x)\,\mathrm{d}x, \quad \Pi_0^\perp := \mathrm{Id} - \Pi_0,$$

so that writing $v = v_0 + u$, $u := \Pi_0^\perp v$, $f = f_0 + g$, $g := \Pi_0^\perp f$, the equation $F(v) = 0$ (see (1.6)) is equivalent to

$$\begin{cases} (\lambda\bar\omega\cdot\partial_\varphi)^2 u - \left(1+\varepsilon\int_{\mathbb{T}^d}|\nabla u|^2\,\mathrm{d}x\right)\Delta u - \varepsilon g = 0, \\ (\lambda\bar\omega\cdot\partial_\varphi)^2 v_0 - \varepsilon f_0 = 0. \end{cases} \tag{2.1}$$

By (1.2) and (1.3), using that

$$\frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} f_0(\varphi)\,\mathrm{d}\varphi = \frac{1}{(2\pi)^{\nu+d}}\int_{\mathbb{T}^{\nu+d}} f(\varphi,x)\,\mathrm{d}\varphi\,\mathrm{d}x = 0$$

the second equation in (2.1) is easily solved and we get

$$v_0(\varphi) := \varepsilon(\lambda\bar\omega\cdot\partial_\varphi)^{-2} f_0.$$

Then we are reduced to look for zeroes of the nonlinear operator

$$\mathcal{F}(u) \equiv \mathcal{F}(\lambda,u) := (\lambda\bar\omega\cdot\partial_\varphi)^2 u - \left(1+\varepsilon\int_{\mathbb{T}^d}|\nabla u|^2\,\mathrm{d}x\right)\Delta u - \varepsilon g \tag{2.2}$$

acting on Sobolev spaces of functions with zero average in $x \in \mathbb{T}^d$, i.e.

$$H_0^s := \left\{u \in H^s : \int_{\mathbb{T}^d} u(\varphi,x)\,\mathrm{d}x = 0\right\}. \tag{2.3}$$

3. Function spaces, norms, linear operators

Given a family of Sobolev functions $u(\varphi,x;\lambda)$, $\lambda \in \Lambda \subset \mathbb{R}$, we define the Sobolev norm $\|\cdot\|_s$ as

$$\begin{aligned} \|u\|_s &:= \|u\|_s^{\sup} + \|\partial_\lambda u\|_{s-1}^{\sup}, \\ \|u\|_s^{\sup} &:= \sup_{\lambda\in\Lambda}\|u(\cdot;\lambda)\|_s. \end{aligned} \tag{3.1}$$

If $\mu : \Lambda \to \mathbb{R}$, we define

$$\|\mu\| := |\mu|^{\sup} + |\partial_\lambda\mu|^{\sup}, \quad |\mu|^{\sup} := \sup_{\lambda\in\Lambda}|\mu(\lambda)|. \tag{3.2}$$

Note that the classical interpolation result for $\|\cdot\|_s$ holds, i.e. given $u(\cdot;\lambda), v(\cdot;\lambda)$, $\lambda \in \Lambda$, one has

$$\|uv\|_s \leqslant C(s)\|u\|_s\|v\|_{s_0} + C(s_0)\|u\|_{s_0}\|v\|_s, \quad s \geqslant s_0 \tag{3.3}$$

where we fix once and for all

$$s_0 := \left[\frac{\nu+d}{2}\right] + 1 \tag{3.4}$$

and $[x]$ denotes the integer part of $x \in \mathbb{R}$.

For any $N > 0$ let us define the spaces of trigonometric polynomials

$$E_N := \mathrm{span}\left\{\mathrm{e}^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)} : 0 < |(\ell,j)| \leqslant N\right\} \tag{3.5}$$

and the orthogonal projector

$$\Pi_N : L^2(\mathbb{T}^{\nu+d}) \to E_N, \quad \Pi_N^\perp := \mathrm{Id} - \Pi_N; \tag{3.6}$$

of course the following standard smoothing estimates hold:

$$|||\Pi_N u|||_{s+\alpha} \leqslant N^\alpha |||u|||_s, \quad |||\Pi_N^\perp u|||_s \leqslant N^{-\alpha} |||u|||_{s+\alpha}. \tag{3.7}$$

Let us introduce the notations $\lesssim$ and $\lesssim_s$; we write $a \lesssim b$ if there exists a constant $c = c(\nu, d, \gamma_0)$ such that $a < cb$, and $a \lesssim_s b$ if the constant depends also on s.

We now recall some results concerning operators induced by diffeomorphism of the torus.

Lemma 3.1. *Let $\beta(\varphi; \lambda)$ satisfy $|||\beta|||_{s_0+1} \leqslant \delta$ for some δ small enough and $\omega = \lambda\bar{\omega}$ with $\lambda \in \mathcal{I}$. Then the composition operator*

$$\mathcal{B} : u \mapsto \mathcal{B}u, \quad (\mathcal{B}u)(\varphi, x) := u(\varphi + \omega\beta(\varphi), x)\,,$$

satisfies

$$\|\mathcal{B}u\|_s \lesssim_s \|u\|_s + \|\beta\|_{s+s_0}\|u\|_1\,, \qquad \text{for all } s \geqslant 1, \tag{3.8}$$

$$\|(\partial_\lambda \mathcal{B})u\|_s \lesssim_s \|u\|_{s+1} + |||\beta|||_{s+s_0}\|u\|_2, \quad \forall s \geqslant 2. \tag{3.9}$$

Moreover the map $\varphi \mapsto \varphi + \omega\beta(\varphi)$ is invertible with inverse given by $\vartheta \mapsto \vartheta + \omega\breve{\beta}(\vartheta)$. The function $\breve{\beta}$ satisfies the estimate

$$|||\breve{\beta}|||_s \lesssim_s |||\beta|||_{s+s_0}\,. \tag{3.10}$$

Proof. The lemma can be proved arguing as in the proof of lemma B.4 in [3] (using also that by Sobolev embedding $\|\cdot\|_{\mathcal{C}^s} \lesssim \|\cdot\|_{s+s_0}$). The estimate on $\partial_\lambda \mathcal{B}$, follows by differentiating w.r.t. to λ, using the estimate (3.8) and by applying the interpolation estimate (3.3). ■

The following lemma follows directly by applying the classical Moser estimate for composition operators, see [43].

Lemma 3.2 (Composition operator). *Let $f \in \mathcal{C}^q(\mathbb{T}^{\nu+d} \times B_K, \mathbb{R})$, where $B_K := [-K, K]$ for some $K > 0$ large enough. If $u(\cdot; \lambda) \in H^s(\mathbb{T}^{\nu+d})$, $\lambda \in \Lambda$ is a family of Sobolev functions satisfying $\|u\|_{s_0} \leqslant 1$. Then for any $s \geqslant s_0$*

$$|||f(\cdot, u)|||_s \leqslant C(s, f)(1 + |||u|||_s)\,. \tag{3.11}$$

3.1. Linear operators on H_0^s and matrices

Set $\mathbb{Z}_*^d := \mathbb{Z}^d \setminus \{0\}$ and let $B, C \subseteq \mathbb{Z}^\nu \times \mathbb{Z}_*^d$. A bounded linear operator $L : H_B^s \to H_C^s$ is represented, as usual, by a matrix in

$$\mathcal{M}_C^B := \Big\{ \big(M_k^{k'}\big)_{k\in C, k'\in B},\, M_k^{k'} \in \mathbb{C} \Big\}. \tag{3.12}$$

Definition 3.3 (s-decay norm). For any $M \in \mathcal{M}_C^B$ we define its s-decay norm as

$$|M|_s^2 := \sum_{k\in\mathbb{Z}^{\nu+d}} [M(k)]^2 \langle k \rangle^{2s} \tag{3.13}$$

where, for $k = (\ell, j)$ $\langle k \rangle := \max(1, |k|) = \max(1, |\ell|, |j|)$,

$$[M(k)] := \begin{cases} \sup_{h-h'=k, h\in C, h'\in B} |M_h^{h'}|, & k \in C - B, \\ 0, & k \notin C - B. \end{cases} \tag{3.14}$$

If the matrix M depends on a parameter $\lambda \in \Lambda \subseteq \mathbb{R}$, we define

$$\|M\|_s := |M|_s^{\sup} + |\partial_\lambda M|_s^{\sup} \quad \text{where} \quad |M|_s^{\sup} := \sup_{\lambda\in\Lambda} |M(\lambda)|_s.$$

Remark 3.4. Note that if M represent a multiplication operator by a function $a(\varphi, x)$ then

$$|M|_s = \|a\|_s \qquad \text{and} \qquad \|M\|_s = \|\|a\|\|_s.$$

We have the following standard results; see for instance [12] and references therein.

Lemma 3.5 (Interpolation). *For all $s \geqslant s_0$ there is $C(s) > 1$ with $C(s_0) = 1$ such that, for any subset $B, C, D \subseteq \mathbb{Z}^\nu \times \mathbb{Z}^d_*$ and for all $M_1 \in \mathcal{M}^C_D$, $M_2 \in \mathcal{M}^B_C$, one has*

$$|M_1 M_2|_s \leqslant \frac{1}{2}|M_1|_{s_0}|M_2|_s + \frac{C(s)}{2}|M_1|_s|M_2|_{s_0}. \tag{3.15}$$

In particular, one has the algebra property $|M_1 M_2|_s \leqslant C(s)|M_1|_s|M_2|_s$. Similar estimates hold by replacing $|\cdot|_s$ with $\|\cdot\|_s$ if M_1 and M_2 depend on the parameter λ.

Iterating the estimate of the above lemma one easily gets

$$|M^n|_s \leqslant C(s)^n |M|_s^{n-1}|M|_{s_0}, \quad \forall n \in \mathbb{N}, \quad s \geqslant s_0. \tag{3.16}$$

If M depends on the parameter λ, a similar estimate holds by replacing $|\cdot|_s$ with $\|\cdot\|_s$.

Lemma 3.6. *For any $B, C \subseteq \mathbb{Z}^\nu \times \mathbb{Z}^d_*$, let $M \in \mathcal{M}^B_C$. Then*

$$\|Mh\|_s \leqslant C(s)|M|_{s_0}\|h\|_s + C(s)|M|_s\|h\|_{s_0}, \quad \forall h \in H^s_B. \tag{3.17}$$

Of course all the results stated above hold replacing $|\cdot|_s$ by $\|\cdot\|_s$.

4. The linearized operator

In this section we study the linearized operator $\mathcal{L}(u) := D_u\mathcal{F}(u)$ for any $u(\varphi, x; \lambda)$ which is $\mathcal{C}^\infty$ w.r.t. $(\varphi, x) \in \mathbb{T}^{\nu+d}$ and $\mathcal{C}^1$ w.r.t. the parameter $\lambda \in \mathcal{I}$. The linearized operator $\mathcal{L} : H_0^{s+2} \to H_0^s$, $s \geqslant 0$ has the form

$$\mathcal{L} = (\omega\cdot\partial_\varphi)^2 - \big(1 + a(\varphi)\big)\Delta + \mathcal{R}$$
$$a(\varphi) := \varepsilon \int_{\mathbb{T}^d} |\nabla u(\varphi, x)|^2\, \mathrm{d}x, \quad \mathcal{R}[h] := -2\Delta u \int_{\mathbb{T}^d} \Delta u\, h\, \mathrm{d}x, \quad h \in L^2_0(\mathbb{T}^{\nu+d}). \tag{4.1}$$

4.1. Reduction to constant coefficients up to the order zero

In this section we prove the following proposition.

Proposition 4.1. *There exists $\sigma = \sigma(\nu, d) > 0$ such that if*

$$\|\|u\|\|_{s_0+\sigma} \leqslant 1, \tag{4.2}$$

there exists $\delta \in (0,1)$ such that if $\varepsilon\gamma_0^{-1} \leqslant \delta$ then there exist two invertible changes of variables Φ_1, Φ_2 such that

$$\Phi_1 \mathcal{L} \Phi_2 = \mathcal{L}_2 = (\omega \cdot \partial_\vartheta)^2 - \mu\Delta + \mathcal{R}_2$$

where μ is a constant and $\mathcal{R}_2$ is an operator of order 0 satisfying the following properties. The constant $\mu \equiv \mu(\lambda, u(\lambda))$ is $\mathcal{C}^1$ w.r.t. the parameter λ and

$$\|\|\mu - 1\|\| \lesssim \varepsilon, \quad |\partial_u \mu[h]| \lesssim \varepsilon \|h\|_\sigma. \tag{4.3}$$

The changes of variables Φ_1, Φ_2 are $\mathcal{C}^1$ w.r.t. the parameter λ and they satisfy the tame estimates

$$\begin{aligned} &\|\Phi_1^{\pm 1} h\|_s, \|\Phi_2^{\pm 1} h\|_s \lesssim_s \|h\|_s + \|u\|_{s+\sigma} \|h\|_{s_0}, \quad \forall s \geqslant s_0, \\ &\|(\partial_\lambda \Phi_1^{\pm 1}) h\|_{s-1}, \|(\partial_\lambda \Phi_2^{\pm 1}) h\|_{s-1} \lesssim_s \|h\|_s + \|\|u\|\|_{s+\sigma} \|h\|_{s_0}, \quad \forall s \geqslant s_0. \end{aligned} \tag{4.4}$$

The remainder $\mathcal{R}_2$ is self-adjoint in L^2 and satisfies

$$\begin{aligned} &\|\mathcal{R}_2\|_s \lesssim_s \varepsilon(1 + \|\|u\|\|_{s+\sigma}), \quad \forall s \geqslant s_0, \\ &\|\partial_u \mathcal{R}_2[h]\|_s \lesssim_s \varepsilon\Big(\|\|h\|\|_{s+\sigma} + \|\|u\|\|_{s+\sigma} \|\|h\|\|_{s_0+\sigma}\Big), \quad \forall s \geqslant s_0. \end{aligned} \tag{4.5}$$

4.1.1. Step 1: reduction of the highest order. In this section we reduce to constant coefficients the highest order term $a(\varphi)\Delta$ in (4.1). Given a diffeomorphism of the torus $\mathbb{T}^\nu \to \mathbb{T}^\nu$, $\varphi \mapsto \varphi + \omega\alpha(\varphi)$ we consider the induced operator

$$\mathcal{A}h(\varphi, x) := h(\varphi + \omega\alpha(\varphi)) \tag{4.6}$$

where $\alpha : \mathbb{T}^\nu \to \mathbb{R}$ is a small function to be determined. The inverse operator $\mathcal{A}^{-1}$ has the form

$$\mathcal{A}^{-1}h(\vartheta, x) := h(\vartheta + \omega\breve{\alpha}(\vartheta), x) \tag{4.7}$$

where $\vartheta \mapsto \vartheta + \omega\breve{\alpha}(\vartheta)$ is the inverse diffeomorphism of $\varphi \mapsto \varphi + \omega\alpha(\varphi)$. One has the following conjugation rules:

$$\begin{aligned} &\mathcal{A}^{-1} a \mathcal{A} = \mathcal{A}^{-1}[a], \quad \mathcal{A}^{-1} \circ \Delta \circ \mathcal{A} = \Delta, \\ &\mathcal{A}^{-1}(\omega \cdot \partial_\varphi)\mathcal{A} = \mathcal{A}^{-1}\left[1 + \omega \cdot \partial_\varphi \alpha\right] \omega \cdot \partial_\vartheta, \\ &\mathcal{A}^{-1}(\omega \cdot \partial_\varphi)^2 \mathcal{A} = \mathcal{A}^{-1}\left[(1 + \omega \cdot \partial_\varphi \alpha)^2\right](\omega \cdot \partial_\vartheta)^2 + \mathcal{A}^{-1}[(\omega \cdot \partial_\varphi)^2 \alpha]\omega \cdot \partial_\vartheta. \end{aligned} \tag{4.8}$$

By (4.1) and (4.8), one has

$$\mathcal{A}^{-1}\mathcal{L}\mathcal{A} = \mathcal{A}^{-1}\left[(1 + \omega \cdot \partial_\varphi \alpha)^2\right](\omega \cdot \partial_\vartheta)^2 - \mathcal{A}^{-1}[1 + a]\Delta + \mathcal{A}^{-1}[(\omega \cdot \partial_\varphi)^2 \alpha]\omega \cdot \partial_\vartheta + \mathcal{A}^{-1}\mathcal{R}\mathcal{A}. \tag{4.9}$$

We choose the function α so that the coefficient of $(\omega \cdot \partial_\vartheta)^2$ is proportional to the one of the Laplacian Δ, namely we want to solve

$$(1 + \omega \cdot \partial_\varphi \alpha)^2 = \frac{1}{\mu}(1 + a) \tag{4.10}$$

for some constant $\mu \in \mathbb{R}$ to be fixed. Note that by (4.1) and (4.2), one has that $a(\varphi) = O(\varepsilon)$, then for ε small enough $\sqrt{1 + a}$ is well defined and of class $\mathcal{C}^\infty$. Then the equation (4.10) can be written in the form

$$\omega \cdot \partial_\varphi \alpha = \frac{1}{\sqrt{\mu}} \sqrt{1+a} - 1 \tag{4.11}$$

and hence we choose μ so that the rhs of (4.11) has zero average, namely

$$\mu := \Big(\fint_{\mathbb{T}^\nu} \sqrt{1 + a(\varphi)} \, \mathrm{d}\varphi \Big)^2. \tag{4.12}$$

Now, using that $\omega = \lambda \bar{\omega}$ and $\bar{\omega}$ is diophantine, we choose

$$\alpha := (\omega \cdot \partial_\varphi)^{-1} \big[\frac{1}{\sqrt{\mu}} \sqrt{1+a} - 1 \big], \tag{4.13}$$

and in this way, we obtain

$$\begin{aligned} &\mathcal{A}^{-1} \mathcal{L} \mathcal{A} = \rho \mathcal{L}_1, \quad \rho := \mathcal{A}^{-1} [(1 + \omega \cdot \partial_\varphi \alpha)^2], \\ &\mathcal{L}_1 := (\omega \cdot \partial_\vartheta)^2 - \mu \Delta + a_1 \omega \cdot \partial_\vartheta + \mathcal{R}_1, \\ &a_1 := \rho^{-1} \mathcal{A}^{-1} [(\omega \cdot \partial_\varphi)^2 \alpha], \qquad \mathcal{R}_1 := \rho^{-1} \mathcal{A}^{-1} \mathcal{R} \mathcal{A}. \end{aligned} \tag{4.14}$$

Lemma 4.2. *One has* $\int_{\mathbb{T}^\nu} a_1(\vartheta) \, \mathrm{d}\vartheta = 0$.

Proof. By (4.14)

$$a_1(\vartheta) = \mathcal{A}^{-1} \Big[\frac{(\omega \cdot \partial_\varphi)^2 \alpha}{(1 + \omega \cdot \partial_\varphi \alpha)^2} \Big] (\vartheta) = \frac{(\omega \cdot \partial_\varphi)^2 \alpha(\vartheta + \omega \breve{\alpha}(\vartheta))}{(1 + \omega \cdot \partial_\varphi \alpha(\vartheta + \omega \breve{\alpha}(\vartheta)))^2}.$$

Considering the change of variables $\varphi = \vartheta + \omega \breve{\alpha}(\vartheta)$, one gets

$$\begin{aligned} \int_{\mathbb{T}^\nu} a_1(\vartheta) \, \mathrm{d}\vartheta &= \int_{\mathbb{T}^\nu} \frac{(\omega \cdot \partial_\varphi)^2 \alpha(\varphi)}{(1 + \omega \cdot \partial_\varphi \alpha(\varphi))^2} (1 + \omega \cdot \partial_\varphi \alpha(\varphi)) \, \mathrm{d}\varphi \\ &= \int_{\mathbb{T}^\nu} \frac{(\omega \cdot \partial_\varphi)^2 \alpha(\varphi)}{1 + \omega \cdot \partial_\varphi \alpha(\varphi)} \, \mathrm{d}\varphi = \int_{\mathbb{T}^\nu} \omega \cdot \partial_\varphi \log \big(1 + \omega \cdot \partial_\varphi \alpha(\varphi) \big) \, \mathrm{d}\varphi = 0. \end{aligned} \tag{4.15}$$

■

4.1.2. Step 2: reduction of the first order term. The aim of this section is to eliminate the term $a_1(\vartheta) \omega \cdot \partial_\vartheta$ in the operator $\mathcal{L}_1$ defined in (4.14). We conjugate $\mathcal{L}_1$ by means of a multiplication operator

$$\mathcal{B} : h \mapsto b(\vartheta) h$$

where $b : \mathbb{T}^\nu \to \mathbb{R}$ is a function close to 1 to be determined, so that its inverse is given by

$$\mathcal{B}^{-1} : h \mapsto b(\vartheta)^{-1} h.$$

One has the following conjugation rules:

$$\begin{aligned} &\mathcal{B}^{-1} \Delta \mathcal{B} = \Delta, \\ &\mathcal{B}^{-1} \omega \cdot \partial_\vartheta \mathcal{B} = \omega \cdot \partial_\vartheta + b(\vartheta)^{-1} (\omega \cdot \partial_\vartheta b), \\ &\mathcal{B}^{-1} (\omega \cdot \partial_\vartheta)^2 \mathcal{B} = (\omega \cdot \partial_\vartheta)^2 + 2 b(\vartheta)^{-1} (\omega \cdot \partial_\vartheta b) \omega \cdot \partial_\vartheta + b(\vartheta)^{-1} (\omega \cdot \partial_\vartheta)^2 b. \end{aligned} \tag{4.16}$$

By (4.14) and (4.16) one gets

$$\mathcal{L}_2 := \mathcal{B}^{-1}\mathcal{L}_1\mathcal{B} = (\omega\cdot\partial_\vartheta)^2 - \mu\Delta + \Big(b(\vartheta)^{-1}\omega\cdot\partial_\vartheta b + a_1(\vartheta)\Big)\omega\cdot\partial_\vartheta + \mathcal{R}_2 \tag{4.17}$$

where the remainder $\mathcal{R}_2$ is defined as

$$\mathcal{R}_2 := \mathcal{B}^{-1}\mathcal{R}_1\mathcal{B} + b(\vartheta)^{-1}(\omega\cdot\partial_\vartheta)^2 b + a_1(\vartheta)b(\vartheta)^{-1}(\omega\cdot\partial_\vartheta b). \tag{4.18}$$

In order to eliminate the term of order $\omega\cdot\partial_\vartheta$ one has to solve the equation

$$b(\vartheta)^{-1}\omega\cdot\partial_\vartheta b + a_1(\vartheta) = 0. \tag{4.19}$$

Since $b(\vartheta)^{-1}\omega\cdot\partial_\vartheta b = \omega\cdot\partial_\vartheta \log(b(\vartheta))$, the function a_1 has zero average, and recalling that $\omega = \lambda\bar\omega$ with $\bar\omega$ diophantine, the equation (4.19) can be solved by setting

$$b(\vartheta) := \exp\Big(- (\omega\cdot\partial_\vartheta)^{-1}a_1(\vartheta)\Big). \tag{4.20}$$

Then $\mathcal{L}_2$ in (4.17) has the final form

$$\mathcal{L}_2 = \mathcal{D} + \mathcal{R}_2, \quad \mathcal{D} = \mathcal{D}(\lambda, u(\lambda)) := (\omega\cdot\partial_\vartheta)^2 - \mu\Delta, \tag{4.21}$$

and the estimates (4.3)–(4.5) follow similarly to [41]. Indeed they can be proved in an elementary way by using the explicit expressions for $\mathcal{R}_2, \Phi_1, \Phi_2, \mu$ found above and the estimate (3.3), lemmata 3.1, 3.2 and remark 3.4.

Remark 4.3. Note that for $u\equiv 0$ one has $a=0$, $\mu=1$, $\alpha=1$, $\mathcal{A}=\mathbb{1}$, $\rho=1$, $a_1=1$, $b=1$, $\mathcal{B}=\mathbb{1}$ and hence

$$\mathcal{L}_2(0) = \mathcal{L}(0) = (\omega\cdot\partial_\vartheta)^2 - \Delta.$$

In particular $\mathcal{R}_2(0) = 0$.

5. The Nash–Moser scheme

Here we prove the Nash–Moser scheme for parameters λ in a set A_∞ (see below) which in principle might be empty; later we shall prove that A_∞ contains the set $\mathcal{C}_\varepsilon$ mentioned in theorem 1.1 and that $\mathcal{C}_\varepsilon$ has asymptotically full measure.

For any $N>0$ we decompose the operator $\mathcal{L}\equiv\mathcal{L}(u)$ as

$$\mathcal{L}(u) = \mathcal{L}_N(u) + \mathcal{R}_N^\perp(u) \tag{5.1}$$

where

$$\begin{aligned}
\mathcal{L}_N(u) &:= \Phi_1(u)^{-1}(L_N(u) + \Pi_N^\perp)\Phi_2(u)^{-1},\\
L_N(u) &:= D_N(\lambda, u(\lambda)) + R_N\\
D_N(\lambda, u(\lambda)) &:= \Pi_N\mathcal{D}(\lambda, u(\lambda))\Pi_N,\\
R_N(u) &:= \Pi_N\mathcal{R}_2(u)\Pi_N\\
\mathcal{R}_N^\perp(u) &:= \Phi_1(u)^{-1}\Pi_N^\perp\mathcal{L}_2(u)\Pi_N\Phi_2(u)^{-1} + \Phi_1(u)^{-1}\Pi_N\mathcal{L}_2(u)\Pi_N^\perp\Phi_2(u)^{-1}\\
&\quad + \Phi_1(u)^{-1}\Pi_N^\perp\mathcal{L}_2(u)\Pi_N^\perp\Phi_2(u)^{-1} - \Phi_1(u)^{-1}\Pi_N^\perp\Phi_2(u)^{-1}.
\end{aligned} \tag{5.2}$$

Note that, by applying the estimates (4.4) and recalling (4.1), the operator $\mathcal{R}_N^\perp$ satisfies

$$\|\mathcal{R}_N^\perp h\|_{s_0} \lesssim N^{-\mathtt{b}}\big(\|h\|_{s_0+\mathtt{b}+\sigma} + \|u\|_{s_0+\mathtt{b}+\sigma}\|h\|_{s_0+\sigma}\big), \quad \forall \mathtt{b} > 0,$$
$$\|\mathcal{R}_N^\perp h\|_s \lesssim_s \|h\|_{s+\sigma} + \|u\|_{s+\sigma}\|h\|_{s_0+\sigma}, \quad \forall s \geqslant s_0. \tag{5.3}$$

Let $S > s_1 > s_0 + \sigma$ and consider $u \in \mathcal{C}^1(\mathcal{I}, H_0^{s_1})$ such that

$$\|u\|_{s_1} \leqslant 1; \tag{5.4}$$

for any $\tau > 0$, $\delta \in (0, 1/3)$ we define the set

$$\mathfrak{G}_N(u) = \mathfrak{G}_{N,\delta,\tau}(u) := \Big\{\lambda \in \mathcal{I} \ : \ \forall\, s \in [s_1, S], \text{one has } |L_N(\lambda, u(\lambda))^{-1}|_s \lesssim_s N^{\mathfrak{a}+\delta(s-s_1)}(1 + \|u\|_{s+\sigma})\Big\}, \tag{5.5}$$

where $\mathfrak{a} := \tau + \delta s_1$.

For any set $A \subset \mathcal{I}$ and $\eta > 0$ we define

$$\mathcal{N}(A, \eta) := \big\{\lambda \in \mathcal{I} : \text{dist}(\lambda, A) \leqslant \eta\big\}$$

and let

$$N_0 > 0, \quad N_n := N_0^{(3/2)^n}. \tag{5.6}$$

Let us introduce parameters κ_1, κ_2, κ_3, satisfying

$$\kappa_1 > \sigma, \quad \kappa_2 > \max\{3\mathfrak{a} + \frac{3}{2}(s_1 - s_0) + 3 + \frac{9}{4}\kappa_1, 12\mathfrak{a} + 24\},$$
$$\kappa_3 > 6\mathfrak{a} + 6 + 3\delta(S - s_1) + 3\sigma + \frac{3}{2}\kappa_1,$$
$$(1-\delta)(S - s_1) > 2\sigma + 2 + 2\mathfrak{a} + \frac{2}{3}\kappa_3 + \kappa_2, \quad S \leqslant q. \tag{5.7}$$

Note one needs to impose the condition $0 < \delta < \frac{1}{3}$ because the second and the third conditions are compatible only if $(1 - 3\delta)(S - s_1) > 6\mathfrak{a} + 6 + \sigma + \kappa_1$. Recall that q in the third line of (5.7) is the regularity of the forcing term $f(\omega t, x)$ in (1.5).

Theorem 5.1 (Nash–Moser). *For τ, δ, κ_1, κ_2, κ_3 ,s_0, $q \geqslant S > s_1 > s_0 + \sigma$, satisfying (5.7), there are c, $\overline{N}_0$, such that, for all $N_0 \geqslant \overline{N}_0$ and ε_0 small enough such that*

$$\varepsilon_0 N_0^S \leqslant c\,, \tag{5.8}$$

and, for all $\varepsilon \in [0, \varepsilon_0)$ a sequence $\{u_n = u_n(\varepsilon, \cdot)\}_{n \geqslant 0} \subset C^1(\mathcal{I}, H_0^{s_1})$ such that

$(S1)_n$ $u_n(\varepsilon, \lambda) \in E_{N_n}$, $u_n(0, \lambda) = 0$, $\|u_n\|_{s_1} \leqslant 1$.
$(S2)_n$ For all $1 \leqslant i \leqslant n$ one has $\|u_i - u_{i-1}\|_{s_1} \leqslant N_i^{-\kappa_1}$.
$(S3)_n$ Set $u_{-1}{:} = 0$ and define

$$A_n := \bigcap_{i=0}^{n} \mathfrak{G}_{N_i}(u_{i-1})\,. \tag{5.9}$$

For $\lambda \in \mathcal{N}(A_n, N_n^{-\kappa_1/2})$ the function $u_n(\varepsilon, \lambda)$ satisfies $\|\mathcal{F}(u_n)\|_{s_0} \leqslant C N_n^{-\kappa_2}$.
$(S4)_n$ For any $i = 1, \ldots, n$, $\|u_i\|_S \leqslant N_i^{\kappa_3}$.

As a consequence, for all $\varepsilon \in [0, \varepsilon_0)$, the sequence $\{u_n(\varepsilon, \cdot)\}_{n \geqslant 0}$ converges uniformly in $C^1(\mathcal{I}, H_0^{s_1})$ to u_ε with $u_0(\lambda) \equiv 0$, at a superexponential rate

$$\|u_\varepsilon(\lambda) - u_n(\lambda)\|_{s_1} \leqslant N_{n+1}^{-\kappa_1}\,, \quad \forall \lambda \in \mathcal{I}\,, \tag{5.10}$$

and for all $\lambda \in A_\infty := \bigcap_{n \geqslant 0} A_n$ one has $\mathcal{F}(\varepsilon, \lambda, u_\varepsilon(\lambda)) = 0$.

5.1. Proof of theorem 5.1

First of all we note that by differentiating the nonlinear operator $\mathcal{F}$ defined in (2.2) by using (3.3), the following tame properties hold: for any $s \in [s_0, S]$, with $S \leqslant q$, there is $C = C(s)$ such that for any $u, h \in \mathcal{C}^1(\mathcal{I}, H_0^s)$ with $|||u|||_{s_0+2} \leqslant 1$ one has

(F1) $|||\mathcal{F}(\varepsilon, \lambda, u)|||_s \leqslant C(s)(1 + |||u|||_{s+2})$,
(F2) $|||D_u\mathcal{F}(\varepsilon, \lambda, u)[h]|||_s \leqslant C(s)\left(|||h|||_{s+2} + |||u|||_{s+2}|||h|||_{s_0+2}\right)$,
(F3) $|||\mathcal{F}(\varepsilon, \lambda, u + h) - F(\varepsilon, \lambda, u) - D_uF(\varepsilon, \lambda, u)[h]|||_s \leqslant C(s)$
$\left(|||h|||_{s+2}|||h|||_{s_0+2} + |||u|||_{s+2}|||h|||^2_{s_0+2}\right)$.

Lemma 5.2. *Let $\kappa > \mathfrak{a} + 2$ and $|||u|||_{s_1} \leqslant 1$. For any $\lambda \in \mathcal{N}\big(\mathfrak{G}_N(u), 2N^{-\kappa}\big)$, for $s \geqslant s_1$ there exists $\varepsilon_0 = \varepsilon_0(s) \in (0, 1)$ small enough such that if $\varepsilon \leqslant \varepsilon_0$, the operator $L_N(\lambda, u(\lambda))$ is invertible and*

$$\|L_N(u)^{-1}\|_s \lesssim_s N^{2\mathfrak{a}+2+\delta(s-s_1)}(1 + |||u|||_{s+\sigma}). \tag{5.11a}$$

Proof. Let $\lambda \in \mathfrak{G}_N(u)$ and $\lambda' \in \mathcal{I}$ so that $|\lambda - \lambda'| \leqslant 2N^{-\kappa}$. We show by means of a Neumann series argument that $L_N(\lambda', u(\lambda'))$ is invertible, hence we want to bound $L_N(\varepsilon, \lambda', u(\lambda')) - L_N(\varepsilon, \lambda, u(\lambda))$. By (4.3) and (4.5) we have

$$\begin{aligned}|L_N(\varepsilon, \lambda', u(\lambda')) - L_N(\varepsilon, \lambda, u(\lambda))|_s &\lesssim |\Pi_N(\mathcal{D}(\lambda, u(\lambda)) - \mathcal{D}(\lambda', u(\lambda')))\Pi_N|_s \\ &\quad + |\Pi_N(\mathcal{R}_2(u(\lambda)) - \mathcal{R}_2(u(\lambda')))\Pi_N|_s \\ &\lesssim (N^2 + \varepsilon(1 + |||u|||_{s+\sigma}))|\lambda - \lambda'| \lesssim (N^2 + \varepsilon(1 + |||u|||_{s+\sigma}))N^{-\kappa},\end{aligned} \tag{5.12}$$

so that for $s = s_0$, using that $s_0 + \sigma < s_1$ and $|||u|||_{s_1} \leqslant 1$ this reads

$$|L_N(\varepsilon, \lambda', u(\lambda')) - L_N(\varepsilon, \lambda, u(\lambda))|_{s_0} \lesssim N^{-\kappa+2}. \tag{5.13}$$

Setting $A := L_N(\varepsilon, \lambda, u(\lambda))^{-1}(L_N(\varepsilon, \lambda', u(\lambda')) - L_N(\varepsilon, \lambda, u(\lambda)))$, by Neumann series one can write formally

$$L_N(\lambda', u(\lambda'))^{-1} = \sum_{n\geqslant 0}(-1)^n A^n L_N(\lambda, u(\lambda))^{-1},$$

and hence, using (5.12) and (5.13), $\lambda \in \mathfrak{G}_N(u)$ and the interpolation estimate (3.15), we obtain

$$|A|_{s_0} \lesssim N^{2+\mathfrak{a}-\kappa}, \quad |A|_s \lesssim_s N^{\mathfrak{a}+\delta(s-s_1)+2-\kappa}\big(1 + |||u|||_{s+\sigma}\big), \tag{5.14}$$

so that by the estimate (3.16), one obtains

$$\begin{aligned}|L_N(\varepsilon, \lambda', u(\lambda'))^{-1}|_s &\leqslant \Big(\sum_{p\geqslant 0} C(s)^p|A|_s|A|_{s_0}^{p-1}\Big)|L_N(\varepsilon, \lambda, u(\lambda))^{-1}|_{s_0} + \Big(\sum_{p\geqslant 0} C(s_1)^p|A|^p_{s_0}\Big)|L_N(\varepsilon, \lambda, u(\lambda))^{-1}|_s \\ &\lesssim_s N^{\mathfrak{a}+\delta(s-s_1)}(1 + |||u|||_{s+\sigma}).\end{aligned} \tag{5.15}$$

Now for any $\lambda \in \mathcal{N}\big(\mathfrak{G}_N(u), N^{-\kappa}\big)$ by applying (5.2), (4.3) and (4.5) one has

$$|\partial_\lambda L_N(\lambda, u(\lambda))|_s \lesssim_s N^2 + |||u|||_{s+\sigma}. \tag{5.16}$$

Finally, since $\partial_\lambda L_N(\lambda, u(\lambda))^{-1} = -L_N(\lambda, u(\lambda))^{-1}\partial_\lambda L_N(\lambda, u(\lambda))L_N(\lambda, u(\lambda))^{-1}$, applying the estimates (5.15), (3.15) and (5.16) one obtains that

$$|\partial_\lambda L_N(\lambda, u(\lambda))^{-1}|_s \lesssim_s N^{2\mathfrak{a}+2+\delta(s-s_1)}(1 + \|\|u\|\|_{s+\sigma}),$$

so that the assertion follows. ■

The first step of the Nash–Moser algorithm is standard and uses the smallness condition (5.8).

Suppose inductively that u_n is defined in such a way that the properties $(S1)_n - (S4)_n$ hold. We now define u_{n+1}. We write

$$\mathcal{F}(u_n + h) = \mathcal{F}(u_n) + D_u\mathcal{F}(u_n)[h] + \mathcal{Q}(u_n, h) \tag{5.17}$$

where

$$\mathcal{Q}(u_n, h) := \mathcal{F}(u_n + h) - \mathcal{F}(u_n) - D_u\mathcal{F}(u_n)[h], \tag{5.18}$$

so that, using (5.1) with $N = N_n$ and writing $\mathcal{F}(u_n) = \Pi_{N_{n+1}}\mathcal{F}(u_n) + \Pi^\perp_{N_{n+1}}\mathcal{F}(u_n)$ one gets

$$\mathcal{F}(u_n + h) = \mathcal{F}(u_n) + \mathcal{L}_{N_{n+1}}(u_n)[h] + \mathcal{R}^\perp_{N_{n+1}}(u_n)[h] + \mathcal{Q}(u_n, h). \tag{5.19}$$

Note that by applying lemma 5.2, if $\lambda \in \mathcal{N}\big(A_{n+1}, 2N_{n+1}^{-\kappa_1/2}\big)$ (recall (5.9)) the operator $L_{N_{n+1}}(\lambda, u_n(\lambda)) : E_{N_{n+1}} \to E_{N_{n+1}}$ (recall (5.2) and (5.5)) is invertible, implying that $L_{N_{n+1}}(\lambda, u_n(\lambda)) + \Pi^\perp_{N_{n+1}} : H^s_0 \to H^s_0$ is invertible with $\|\big(L_{N_{n+1}}(\lambda, u_n(\lambda)) + \Pi^\perp_{N_{n+1}}\big)^{-1}\|_s \leqslant \|L_{N_{n+1}}(\lambda, u_n(\lambda))^{-1}\|_s \lesssim_s N_{n+1}^{2\mathfrak{a}+2+\delta(s-s_1)}(1 + \|\|u_n\|\|_{s+\sigma})$. Since $\Phi_1(\lambda, u_n(\lambda))$ and $\Phi_2(\lambda, u_n(\lambda))$ are invertible for any $\lambda \in \mathcal{I}$ and satisfy the estimates (4.4) then $\mathcal{L}_{N_{n+1}}(\lambda, u_n(\lambda))$ is also invertible. By the estimates (4.4), the definition of the set $\mathfrak{G}_{N_{n+1}}(u_n)$, the estimate (3.17) and recalling that, by the inductive hypothesis $(S1)_n$ one has $\|\|u_n\|\|_{s_0+\sigma} \leqslant \|\|u_n\|\|_{s_1} \leqslant 1$, we obtain

$$\|\|\mathcal{L}_{N_{n+1}}(u_n)^{-1}[h]\|\|_s \lesssim_s N_{n+1}^{2\mathfrak{a}+2}\|\|h\|\|_s + N_{n+1}^{2\mathfrak{a}+2+\delta(s-s_1)}(1 + \|\|u_n\|\|_{s+\sigma})\|\|h\|\|_{s_0}. \tag{5.20}$$

Let us now define, for $\lambda \in \mathcal{N}\big(A_{n+1}, 2N_{n+1}^{-\kappa_1/2}\big)$,

$$\widetilde{h}_{n+1}(\lambda) := -\Pi_{N_{n+1}}\mathcal{L}_{N_{n+1}}(\lambda, u_n(\lambda))^{-1}\mathcal{F}(\lambda, u_n(\lambda)), \quad \widetilde{u}_{n+1} := u_n + \widetilde{h}_{n+1}. \tag{5.21}$$

Plugging (5.21) into (5.19) one obtains

$$\mathcal{F}(\widetilde{u}_{n+1}) = \Pi^\perp_{N_{n+1}}\mathcal{F}(u_n) + \mathcal{R}^\perp_{N_{n+1}}(u_n)[\widetilde{h}_{n+1}] + \mathcal{Q}(u_n, \widetilde{h}_{n+1}). \tag{5.22}$$

Estimate of $\widetilde{h}_{n+1}$. By applying (5.20), using that $s_1 > s_0 + \sigma > s_0$, the property (3.7) and $\|\|u_n\|\|_{s_1} \leqslant 1$, one gets

$$\begin{aligned}
\|\|\widetilde{h}_{n+1}\|\|_{s_1} &\leqslant N_{n+1}^{s_1-s_0}\|\|\mathcal{L}_{N_{n+1}}(u_n)^{-1}\mathcal{F}(u_n)\|\|_{s_0} \\
&\lesssim N_{n+1}^{s_1-s_0+2\mathfrak{a}+2}\|\|\mathcal{F}(u_n)\|\|_{s_0} \overset{(S3)_n}{\lesssim} N_{n+1}^{s_1-s_0+2\mathfrak{a}+2}N_n^{-\kappa_2}, \\
\|\|\widetilde{h}_{n+1}\|\|_S &\lesssim_S N_{n+1}^{2\mathfrak{a}+2}\|\|\mathcal{F}(u_n)\|\|_S + N_{n+1}^{2\mathfrak{a}+2+\delta(S-s_1)}(1 + \|\|u_n\|\|_{S+\sigma})\|\|\mathcal{F}(u_n)\|\|_{s_1} \\
&\overset{(F1),\,(3.7)}{\lesssim_S} N_{n+1}^{2\mathfrak{a}+2+\delta(S-s_1)+\sigma}(1 + \|\|u_n\|\|_S).
\end{aligned} \tag{5.23}$$

Let us consider a C^∞ cut-off function ψ_{n+1} satisfying

$$\mathrm{supp}(\psi_{n+1}) \subseteq \mathcal{N}\big(A_{n+1}, 2N_{n+1}^{-\frac{\kappa_1}{2}}\big), \quad 0 \leqslant \psi_{n+1} \leqslant 1,$$
$$\psi_{n+1}(\lambda) = 1, \quad \forall \lambda \in \mathcal{N}\Big(A_{n+1}, N_{n+1}^{-\frac{\kappa_1}{2}}\Big)$$

and define an extension of $\widetilde{h}_{n+1}$ to the whole parameter space $\mathcal{I}$ as

$$h_{n+1} := \psi_{n+1}\widetilde{h}_{n+1}, \quad u_{n+1} := u_n + h_{n+1}.$$

Using that $|||\psi_{n+1}||| \lesssim N_{n+1}^{\frac{\kappa_1}{2}}$ and by the estimates (5.23) one has

$$|||h_{n+1}|||_{s_1} \lesssim N_{n+1}^{s_1-s_0+2\mathfrak{a}+2+\frac{\kappa_1}{2}} N_n^{-\kappa_2} \overset{(5.7)}{\lesssim} N_{n+1}^{-\kappa_1}, \tag{5.24a}$$

$$|||h_{n+1}|||_S \lesssim_S N_{n+1}^{2\mathfrak{a}+2+\delta(S-s_1)+\sigma+\frac{\kappa_1}{2}}(1+|||u_n|||_S); \tag{5.24b}$$

in particular $(S2)_{n+1}$ is satisfied. Now

$$|||u_{n+1}|||_S \lesssim_S |||u_n|||_S + N_{n+1}^{2\mathfrak{a}+2+\delta(S-s_1)+\sigma+\frac{\kappa_1}{2}}(1+|||u_n|||_S) \overset{(S4)_n}{\leqslant} C(S)N_{n+1}^{2\mathfrak{a}+2+\delta(S-s_1)+\sigma+\frac{\kappa_1}{2}}N_n^{\kappa_3} \leqslant N_{n+1}^{\kappa_3} \tag{5.25}$$

by (5.7) and by taking $N_0 = N_0(S) > 0$ large enough. Then also $(S4)_{n+1}$ is proved.

Now we estimate $\mathcal{F}(u_{n+1})$ on the set $\mathcal{N}(A_{n+1}, N_{n+1}^{-\frac{\kappa_1}{2}})$. Using again that $|||u_n|||_{s_0+\sigma} \leqslant \|u\|_{s_1} < 1$, one has

$$\begin{aligned}
|||\mathcal{F}(u_{n+1})|||_{s_0} &\overset{(3.7),(5.3),(F3)}{\lesssim} N_{n+1}^{-(S-s_0)}\Big(|||\mathcal{F}(u_n)|||_S + |||h_{n+1}|||_{S+\sigma} + |||u_n|||_{S+\sigma}|||h_{n+1}|||_{s_1}\Big) + |||h_{n+1}|||_{s_0}^2 \\
&\overset{(F1),(3.7),s_1>s_0}{\lesssim} N_{n+1}^{\sigma-(S-s_1)}\Big(1+|||u_n|||_S + |||h_{n+1}|||_S\Big) + N_{n+1}^4|||h_{n+1}|||_{s_0}^2 \\
&\overset{(5.23)}{\lesssim} N_{n+1}^{2\sigma+2+2\mathfrak{a}+(\delta-1)(S-s_1)}(1+|||u_n|||_S) + N_{n+1}^{4\mathfrak{a}+8}|||\mathcal{F}(u_n)|||_{s_0}^2 \\
&\overset{(S3)_n,(S4)_n}{\lesssim} N_{n+1}^{2\sigma+2+2\mathfrak{a}+(\delta-1)(S-s_1)}N_n^{k_3} + N_{n+1}^{4\mathfrak{a}+8}N_n^{-2\kappa_2} \leqslant N_{n+1}^{-\kappa_2}
\end{aligned} \tag{5.26}$$

by (5.7) and taking $N_0 = N_0(S) > 0$ large enough, hence proving $(S3)_{n+1}$. Finally, by using a telescoping argument $u_{n+1} = \sum_{i=0}^{n+1} h_i$, one has

$$|||u_{n+1}|||_{s_1} \overset{(S2)_n}{\leqslant} \sum_{i=0}^{n+1} N_i^{-\kappa_1} \leqslant 1$$

since by taking $N_0 > 0$ is large enough, thus providing $(S1)_{n+1}$.

Clearly the sequence $(u_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{C}^1(\mathcal{I}, H_0^{s_1})$ and therefore the claimed statement follows. ∎

The proof of theorem 5.1 is rather standard and follows the lines of the one in [13, 14]; however here we cannot apply directly the aforementioned results because the subspaces E_N in (3.5) are not invariant under the change of variables $\mathcal{A}$ appearing in (4.6). We also mention that our truncation at the nth step is not $N_0^{2^n}$ but rather $N_0^{\chi^n}$ with $\chi = 3/2$; the reason for this choice is that, since the subspaces E_N are not invariant, we cannot apply the contraction lemma at each step, but really the Newton scheme which converges only for $1 < \chi < 2$.

6. Multiscale analysis

Our aim is to prove that the set A_∞ has asymptotically full measure; in order to do so, following [14] we first prove that A_∞ contains another set $\mathcal{C}_\infty$ and then we show that the set $\mathcal{C}_\infty$ contains another set $\mathcal{C}_\varepsilon$ that has asymptotically full measure.

In order to do so, in addition to the parameters $\tau > 0$, $\delta \in (0, 1/3)$, σ, s_1, s_0, S, $\kappa_1, \kappa_2, \kappa_3$ satisfying (5.7) needed in theorem 5.1, we now introduce other parameters τ_1, χ_0, τ_0, C_1 and add the following constraints

$$\tau > \tau_0 , \quad \tau_1 > 2\chi_0 d , \quad \tau > 2\tau_1 + d + \nu + 1, \quad C_1 \geqslant 2 , \tag{6.1}$$

then, setting $\kappa := \tau + d + \nu + s_0$,

$$\chi_0(\tau - 2\tau_1 - d - \nu) > 3(\kappa + (s_0 + d + \nu)C_1), \qquad \chi_0\delta > C_1, \tag{6.2a}$$

$$s_1 > 3\kappa + \sigma + 2\chi_0(\tau_1 + d + \nu) + C_1 s_0. \tag{6.2b}$$

Note that no restrictions from above on S' are required, i.e. it could be $S' = +\infty$.

Given $\Omega, \Omega' \subset \mathbb{Z}^\nu \times \mathbb{Z}^d_*$, we define

$$\mathrm{diam}(\Omega) := \sup_{k,k'\in\Omega} \mathrm{dist}(k,k'), \qquad \mathrm{dist}(\Omega,\Omega') := \inf_{k\in\Omega,k'\in\Omega'} |k - k'| .$$

Definition 6.1 (Regular/singular sites). We say that the index $k = (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d_*$ is *regular* for a diagonal matrix D, if $|D_{\ell,j}| \geqslant 1$, otherwise we say that k is *singular*.

Definition 6.2 (N-good/N-bad matrices). Let $F \subset \mathbb{Z}^\nu \times \mathbb{Z}^d_*$ be such that $\mathrm{diam}(F) \leqslant 4N$ for some $N \in \mathbb{N}$. We say that a matrix $A \in \mathcal{M}^F_F$ is *N-good* if A is invertible and for all $s \in [s_0, s_2]$ one has

$$|A^{-1}|_s \leqslant N^{\tau+\delta s}.$$

Otherwise we say that A is *N-bad*.

Definition 6.3 ((A, N)-regular, good, bad sites). For any finite $E \subset \mathbb{Z}^\nu \times \mathbb{Z}^d_*$, let $A = D + \varepsilon T \in \mathcal{M}^E_E$ with $D := \mathrm{diag}(D_k)$, $D_k \in \mathbb{C}$. An index $k \in E$ is

- (A, N)-*regular* if there exists $F \subseteq E$ such that $\mathrm{diam}(F) \leqslant 4N$, $\mathrm{dist}(\{k\}, E \setminus F) \geqslant N$ and the matrix A^F_F is N-good.
- (A, N)-*good* if either it is regular for D (definition 6.1) or it is (A, N)-regular. Otherwise k is (A, N)-*bad*.

The above definition could be extended to infinite E.

Let L be as in (5.2). Note that $\mathcal{D}$ in (4.21) is represented by a diagonal matrix

$$D(\lambda) := \mathrm{diag}_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{Z}^d_*} D_{\ell,j}(\lambda), \quad D_{\ell,j}(\lambda) := -(\lambda\bar\omega\cdot\ell)^2 + \mu(\lambda)|j|^2. \tag{6.3}$$

Now for $\theta \in \mathbb{R}$ let us introduce the matrix

$$D(\lambda,\theta) := \mathrm{diag}_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{Z}^d_*} D_{\ell,j}(\lambda,\theta), \quad D_{\ell,j}(\lambda,\theta) := -(\lambda\bar\omega\cdot\ell + \theta)^2 + \mu(\lambda)|j|^2, \tag{6.4}$$

and denote

$$L(\varepsilon, \lambda, \theta, u) := D(\lambda, \theta) + \mathcal{R}_2(u) . \tag{6.5}$$

Lemma 6.4. *For all* $\tau > 1, N > 1, \lambda \in [1/2, 3/2], \ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}^d_*$ *one has*

$$\{\theta \in \mathbb{R} \, : \, |D_{\ell,j}(\lambda,\theta)| \leqslant N^{-\tau}\} \subseteq I_1 \cup I_2 \qquad \text{intervals with } \operatorname{meas}(I_q) \leqslant N^{-\tau}. \tag{6.6}$$

Proof. A direct computation shows

$$\{\theta \in \mathbb{R} \, : \, |D_{\ell,j}| \leqslant N_0^{-\tau}\} = (\theta_{1,-}, \theta_{1,+}) \cup (\theta_{2,-}, \theta_{2,+})$$

with

$$\theta_{1,\pm} = \lambda \overline{\omega} \cdot l + \sqrt{\mu |j|^2 \pm N^{-\tau}}, \quad \theta_{2,\pm} = \lambda \overline{\omega} \cdot l - \sqrt{\mu |j|^2 \pm N^{-\tau}},$$

and hence

$$\operatorname{meas}((\theta_{q,-}, \theta_{q,+})) = \frac{N^{-\tau}}{\sqrt{\mu |j|^2}} + O(N^{-2\tau}), q = 1, 2.$$

Note that by the estimate (4.3), $\mu \approx 1$ and $j \neq 0$ since we are working on the Sobolev space (2.3), so that the assertion follows. ■

For $\tau_0 > 0$, $N_0 \geqslant 1$ we define the set

$$\overline{\mathcal{I}} := \overline{\mathcal{I}}(N_0, \tau_0) := \Big\{ \lambda \in \mathcal{I} \, : \, |(\lambda \bar{\omega} \cdot \ell)^2 - |j|^2| \geqslant N_0^{-\tau_0} \ \text{ for all } k = (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d_* : \, |k| \leqslant N_0 \Big\}. \tag{6.7}$$

In order to perform the multiscale analysis we need finite dimensional truncations of such matrices. Given a parameter family of matrices $L(\theta)$ with $\theta \in \mathbb{R}$ and $N > 1$ for any $k = (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d$ we denote by $L_{N,k}(\theta)$ (or equivalently $L_{N,\ell,j}(\theta)$) the sub-matrix of $L(\theta)$ centered at k, i.e.

$$L_{N,k}(\theta) := L(\theta)_F^F, \quad F := \{k' \in \mathbb{Z}^\nu \times \mathbb{Z}^d_* \, : \, \operatorname{dist}(k, k') \leqslant N\}. \tag{6.8}$$

If $\ell = 0$, instead of the notation (6.8) we shall use the notation

$$L_{N,j}(\theta) := L_{N,0,j}(\theta) \, ,$$

if also $j = 0$ we write

$$L_N(\theta) := L_{N,0}(\theta),$$

and for $\theta = 0$ we denote $L_{N,j} := L_{N,j}(0)$.

Definition 6.5 (*N*-good/*N*-bad parameters). Let $\mathfrak{c}$ be large enough (to be computed). We denote

$$B_N(j_0, \varepsilon, \lambda) := \Big\{ \theta \in \mathbb{R} \, : \, L_{N,j_0}(\varepsilon, \lambda, \theta, u) \text{ is } N\text{-bad} \Big\}. \tag{6.9}$$

A parameter $\lambda \in \mathcal{I}$ is N-good for L if for any $j_0 \in \mathbb{Z}^d$ one has

$$B_N(j_0, \varepsilon, \lambda) \subseteq \bigcup_{q=1}^{N^{\mathfrak{c}}} I_q \, , \quad I_q \text{ intervals with } \operatorname{meas}(I_q) \leqslant N^{-\tau_1}. \tag{6.10}$$

Otherwise we say that λ is N-bad. We denote the set of N-good parameters as

$$\mathcal{G}_N = \mathcal{G}_N(u) := \Big\{ \lambda \in \mathcal{I} \, : \, \lambda \text{ is } N\text{-good for } L \Big\}. \tag{6.11}$$

The following assumption is needed for the multiscale proposition 6.9; we shall verify it later in section 7.

Ansatz 1 (Separation of bad sites). *There exist* $C_1 > 2$, $\hat{N} = \hat{N}(\tau_0) \in \mathbb{N}$ *and* $\hat{\mathcal{I}} \subseteq \overline{\mathcal{I}}$ *(see (6.7)) such that, for all* $N \geqslant \hat{N}$, *and* $\|u\|_{s_1} < 1$ *(with* s_1 *satisfying (6.2b)), if*

$$\lambda \in \mathcal{G}_N(u) \cap \hat{\mathcal{I}},$$

then for any $\theta \in \mathbb{R}$, for all $\chi \in [\chi_0, 2\chi_0]$ and all $j_0 \in \mathbb{Z}^d$ the (L, N)-bad sites $k = (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d_*$ of $L = L_{N^\chi, j_0}(\varepsilon, \lambda, \theta, u)$ admit a partition $\cup_\beta \Omega_\beta$ in disjoint clusters satisfying

$$\text{diam}(\Omega_\beta) \leqslant N^{C_1}, \qquad \text{dist}(\Omega_{\beta_1}, \Omega_{\beta_2}) \geqslant N^2, \text{ for all } \beta_1 \neq \beta_2. \tag{6.12}$$

For $N > 0$, we denote

$$\mathcal{G}^0_N(u) := \Big\{ \lambda \in \mathcal{I} \ : \ \forall\, j_0 \in \mathbb{Z}^d \text{ there is a covering}$$
$$B^0_N(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^\epsilon} I_q, \quad I_q = I_q(j_0) \text{ intervals with } \text{meas}(I_q) \leqslant N^{-\tau_1} \Big\} \tag{6.13}$$

where

$$B^0_N(j_0, \varepsilon, \lambda) := B^0_N(j_0, \varepsilon, \lambda, u) := \Big\{ \theta \in \mathbb{R} \ : \ \|L^{-1}_{N, j_0}(\varepsilon, \lambda, \theta, u)\|_0 > N^{\tau_1} \Big\}. \tag{6.14}$$

We also set

$$J_N(u) := \Big\{ \lambda \in \mathcal{I} \ : \ \|L_N^{-1}(\varepsilon, \lambda, u)\|_0 \leqslant N^{\tau_1} \Big\}. \tag{6.15}$$

Under the smallness condition (5.8), theorem 5.1 applies, thus defining the sequence un and the sets A_n. We now introduce the sets

$$\mathcal{C}_0 := \hat{\mathcal{I}}, \qquad \mathcal{C}_n := \bigcap_{i=1}^n \mathcal{G}^0_{N_i}(u_{i-1}) \bigcap_{i=1}^n J_{N_i}(u_{i-1}) \cap \hat{\mathcal{I}} \tag{6.16}$$

where $\hat{\mathcal{I}}$ is the one appearing in proposition 7.3, $J_N(u)$ in (6.15), and $\mathcal{G}^0_N(u)$ in (6.13).

Theorem 6.6. *Consider parameters satisfying (5.7), (6.1) and (6.2). Then there exists* $\overline{N}_0 \in \mathbb{N}$, *such that, for all* $N_0 \geqslant \overline{N}_0$ *and* $\varepsilon \in [0, \varepsilon_0)$ *with* ε_0 *satisfying (5.8), the following inclusions hold:*

$$(S5)_0 \qquad \|u\|_{s_1} \leqslant 1 \quad \Rightarrow \quad \mathcal{G}_{N_0}(u) = \mathcal{I}$$
$$(S6)_0 \qquad \mathcal{C}_0 \subseteq A_0,$$

and for all $n \geqslant 1$ *(recall the definitions of* A_n *in (5.9))*

$$(S5)_n \qquad \|u - u_{n-1}\|_{s_1} \leqslant N_n^{-\kappa_1} \quad \Rightarrow \quad \bigcap_{i=1}^n \mathcal{G}^0_{N_i}(u_{i-1}) \cap \hat{\mathcal{I}} \subseteq \mathcal{G}_{N_n}(u) \cap \hat{\mathcal{I}},$$

$$(S6)_n \qquad \mathcal{C}_n \subseteq A_n\,.$$

Hence $\mathcal{C}_\infty := \bigcap_{n\geqslant 0} \mathcal{C}_n \subseteq A_\infty := \bigcap_{n\geqslant 0} A_n$.

6.1. Initialization

Property $(S5)_0$ follows from the following lemma.

Lemma 6.7. *For all* $\|u\|_{s_1} \leqslant 1$, $N \leqslant N_0$, *the set* $\mathcal{G}_N(u) = \mathcal{I}$.

Proof. We claim that, for any $\lambda \in [1/2, 3/2]$ and any $j_0 \in \mathbb{Z}^d$, if (recalling the definition (6.4))

$$|D_{\ell,j}(\lambda,\theta)| > N^{-\tau_1}, \quad \forall (\ell,j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d_* \text{ with } |(\ell, j - j_0)| \leqslant N\,, \tag{6.17}$$

then $L_{N,j_0}(\varepsilon,\lambda,\theta)$ is N-good. This implies that

$$B_N(j_0,\varepsilon,\lambda) \subset \bigcup_{|(l,j-j_0)|\leqslant N} \left\{\theta \in \mathbb{R} \,:\, |D_{\ell,j}(\lambda,\theta)| \leqslant N^{-\tau_1}\right\},$$

which in turn, by lemma 6.4, implies the thesis, see (6.10) and (6.11), for some $\mathfrak{c} \geqslant d + \nu + 1$. The above claim follows by a perturbative argument. Indeed, recalling the definition (5.2), for $\|u\|_{s_1} \leqslant 1$, $s_1 = s_2 + \sigma$, we use (4.5) to obtain

$$|(D^{-1}_{N,j_0}(\lambda,\theta))|_{s_2} |R_{N,j_0}(u)|_{s_2} \leqslant \varepsilon C(s_1) |D^{-1}_{N,j_0}(\lambda,\theta)|_{s_2} (1 + \|u\|_{s_2+\sigma}) \overset{(6.17)}{\leqslant} \varepsilon N^{\tau_1} C(s_1) \overset{(5.8)}{\leqslant} \frac{1}{2}.$$

Then we invert L_{N,j_0} by Neumann series and obtain

$$|L^{-1}_{N,j_0}(\varepsilon,\lambda,\theta)|_s \leqslant 2 |D^{-1}_{N,j_0}(\lambda,\theta)|_s \leqslant 2N^{\tau_1} \leqslant N^{\tau+\delta s}, \quad \forall s \in [s_0, s_2]\,,$$

by (6.1), which proves the claim. ■

Lemma 6.8. *Property* $(S6)_0$ *holds.*

Proof. Since $\hat{\mathcal{I}} \subset \overline{\mathcal{I}}$ it is sufficient to prove that $\overline{\mathcal{I}} \subset A_0$. By the definition of A_0 in (5.9) and (5.5), we have to prove that

$$\lambda \in \overline{\mathcal{I}} \quad \Longrightarrow \quad |L^{-1}_{N_0}(\varepsilon,\lambda,0)|_s \lesssim_s N_0^{\mathfrak{a}+\delta(s-s_1)}\,,\ \forall s \in [s_1, S]\,. \tag{6.18}$$

Indeed, if $\lambda \in \overline{\mathcal{I}}$ then $|D_{\ell,j}(\lambda)| \geqslant N_0^{-\tau_0}$, for all $|(\ell,j)| < N_0$, and so $|D_{N_0}(\lambda)^{-1}|_s \leqslant N_0^{\tau_0}$, $\forall s$. Hence the assertion follows immediately by remark 4.3 and (6.1). ■

6.2. Inductive step

By the Nash–Moser theorem 5.1 we know that $(S1)_n$–$(S4)_n$ hold for all $n \geqslant 0$. Assume inductively that $(S5)_i$ and $(S6)_i$ hold for all $i \leqslant n$. In order to prove $(S5)_{n+1}$, we need the following *multiscale proposition* 6.9 which allows to deduce estimates on the $|\cdot|_s$—norm of the inverse of L from informations on the L^2-norm of the inverse L^{-1}, the off-diagonal decay of L, and separation properties of the bad sites.

Proposition 6.9 (Multiscale). *Assume (6.1) and (6.2). For any* $\bar{s} > s_2$, $\Upsilon > 0$ *there exists* $\varepsilon_0 = \varepsilon_0(\Upsilon, s_2) > 0$ *and* $N_0 = N_0(\Upsilon, \bar{s}) \in \mathbb{N}$ *such that, for all* $N \geqslant N_0$, $|\varepsilon| < \varepsilon_0$, $\chi \in [\chi_0, 2\chi_0]$, $E \subset \mathbb{Z}^\nu \times \mathbb{Z}^d_*$ *with* $\mathrm{diam}(E) \leqslant 4N^\chi$, *if the matrix* $A = D + \varepsilon T \in \mathcal{M}^E_E$ *satisfies*

(H1) $|T|_{s_2} \leqslant \Upsilon$,
(H2) $\|A^{-1}\|_0 \leqslant N^{\chi\tau_1}$,
(H3) there is a partition $\{\Omega_\beta\}_\beta$ *of the* (A, N)*-bad sites (definition 6.3) such that*

$$\mathrm{diam}(\Omega_\beta) \leqslant N^{C_1}, \qquad \mathrm{dist}(\Omega_{\beta_1}, \Omega_{\beta_2}) \geqslant N^2, \text{ for } \beta_1 \neq \beta_2,$$

then the matrix A is N^χ*-good and*

$$|A^{-1}|_s \leqslant \frac{1}{4} N^{\chi\tau} \left(N^{\chi\delta s} + \varepsilon |T|_s \right), \quad \forall s \in [s_0, \bar{s}]. \tag{6.19}$$

Note that the bound (6.19) is much more than requiring that the matrix A is N^χ—good, since it holds also for $s > s_2$.

This Proposition is proved by 'resolvent type arguments' and it coincides essentially with [12]-proposition 4.1. The correspondences in the notations of this paper and [12] respectively are the following: $(\tau, \tau_1, d + r, s_2, \bar{s}) \rightsquigarrow (\tau', \tau, b, s_1, S)$, and, since we do not have a potential, we can fix $\Theta = 1$ in definition 4.2 of [12]. Our conditions (6.1) and (6.2) imply conditions (4.4) and (4.5) of [12] for all $\chi \in [\chi_0, 2\chi_0]$ and our (H1) implies the corresponding Hypothesis (H1) of [12] with $\Upsilon \rightsquigarrow 2\Upsilon$. The other hypotheses are the same. Although the s—norm in this paper is different, the proof of [12]-proposition 4.1 relies only on abstract algebra and interpolation properties of the s—norm (which indeed hold also in this case—see section 3.1). Hence it can be repeated verbatim, full details can be found in [14].

Now, we distinguish two cases:

case 1: $(3/2)^{n+1} \leqslant \chi_0$. Then there exists $\chi \in [\chi_0, 2\chi_0]$ (independent of n) such that

$$N_{n+1} = \overline{N}^\chi, \qquad \overline{N} := [N_{n+1}^{1/\chi_0}] \in (N_0^{1/\chi}, N_0). \tag{6.20}$$

This case may occur only in the first steps.

case 2: $(3/2)^{n+1} > \chi_0$. Then there exists a unique $p \in [0, n]$ such that

$$N_{n+1} = N_p^\chi, \qquad \chi = 2^{n+1-p} \in [\chi_0, 2\chi_0). \tag{6.21}$$

Let us start from **case 1** for $n + 1 = 1$; the other (finitely many) steps are identical.

Lemma 6.10. *Property (S5)*$_1$ *holds.*

Proof. We have to prove that $\mathcal{G}^0_{N_1}(u_0) \cap \hat{\mathcal{I}} \subseteq \mathcal{G}_{N_1}(u) \cap \hat{\mathcal{I}}$ where $\|u - u_0\|_{s_1} \leqslant N_1^{-\kappa_1}$. By definition 6.5 and (6.13) it is sufficient to prove that, for all $j_0 \in \mathbb{Z}^d$,

$$B_{N_1}(j_0, \varepsilon, \lambda, u) \subseteq B^0_{N_1}(j_0, \varepsilon, \lambda, u_0),$$

where we stress the dependence on u,u_0 in (6.9) and (6.14). By the definitions (6.14) and (6.9) this amounts to prove that

$$\|L^{-1}_{N_1, j_0}(\varepsilon, \lambda, \theta, u_0)\|_0 \leqslant N_1^{\tau_1} \quad \Longrightarrow \quad L_{N_1 j_0}(\varepsilon, \lambda, \theta, u) \text{ is } N_1 - \text{good}. \tag{6.22}$$

We first claim that $\|L^{-1}_{N_1, j_0}(\varepsilon, \lambda, \theta, u_0)\|_0 \leqslant N_1^{\tau_1}$ implies

$$|L^{-1}_{N_1, j_0}(\varepsilon, \lambda, \theta, u_0)|_s \leqslant \frac{1}{4} N_1^\tau \left(N_1^{\delta s} + |\mathcal{R}_2(u_0)|_s \right) \overset{(4.5)}{\leqslant} \frac{1}{4} N_1^\tau \left(N_1^{\delta s} + \varepsilon (1 + \|\|u_0\|\|_{s+\sigma}) \right), \quad \forall s \in [s_0, S]. \tag{6.23}$$

Indeed we may apply proposition 6.9 to the matrix $A = L_{N_1,j_0}(\varepsilon,\lambda,\theta,u_0)$ with $\bar{s} = S$, $N = \overline{N}$, $N_1 = \overline{N}^\chi$ and $E = \{|l| \leqslant N_1, |j - j_0| \leqslant N_1\}$. Hypothesis (H1) follows by (4.5) and $\|u_0\|_{s_1} \leqslant 1$. Moreover (H2) is $\|L^{-1}_{N_1,j_0}(\varepsilon,\lambda,\theta,u_0)\|_0 \leqslant N_1^{\tau_1}$. Finally (H3) is implied by Ansatz 1 provided we take $N_0^{1/\chi_0} > \hat{N}(\tau_0)$ (recall (6.20)) and noting that $\lambda \in \mathcal{G}_{\overline{N}}(u_0) \cap \hat{\mathcal{I}}$ by lemma 6.7 (since $\overline{N} \leqslant N_0$ then $\mathcal{G}_{\overline{N}}(u_0) = \mathcal{I}$). Hence (6.19) implies (6.23).

We now prove (6.22); we need to distinguish two cases.

case 1. ($|j_0| > N_1^3$). We first show that $B^0_{N_1}(j_0,\varepsilon,\lambda) \subset \mathbb{R} \setminus [-2N_1, 2N_1]$. Recall that if A, A' are self-adjoint matrices, then their eigenvalues $\mu_p(A)$, $\mu_p(A')$ (ranked in nondecreasing order) satisfy

$$|\mu_p(A) - \mu_p(A')| \leqslant \|A - A'\|_0 . \tag{6.24}$$

Threfore all the eigenvalues $\mu_{\ell,j}(\theta)$ of $L_{N_1,j_0}(\varepsilon,\lambda,\theta,u_0)$ are of the form

$$\mu_{\ell,j}(\theta) = \delta_{\ell,j}(\theta) + O(\varepsilon\|\mathcal{R}_2\|_0), \quad \delta_{\ell,j}(\theta) := -(\omega\cdot\ell+\theta)^2 + \mu(u_0)|j|^2 . \tag{6.25}$$

Since $|\omega|_1 = \lambda|\overline{\omega}|_1 \leqslant 3/2$, $|j - j_0| \leqslant N_1$, $|\ell| \leqslant N_1$, one has

$$\delta_{\ell,j}(\theta) \geqslant -\Big(\frac{3}{2}N_1 + |\theta|\Big)^2 + N_1^2 > \frac{1}{2}N_1^2 , \quad \forall |\theta| < 2N_1$$

and this implies $B^0_{N_1}(j_0,\varepsilon,\lambda) \cap [-2N_1, 2N_1] = \emptyset$. Hence the assumption $\|L^{-1}_{N_1,j_0}(\varepsilon,\lambda,\theta,u_0)\|_0 \leqslant N_1^{\tau_1}$ implies $|\theta| < 2N_1$. But then also the eigenvalues of $L_{N_1,j_0}(\varepsilon,\lambda,\theta,u)$ are big since they are also of the form

$$-(\omega\cdot\ell+\theta)^2 + \mu(u)|j|^2 + O(\varepsilon\|\mathcal{R}_2\|_0). \tag{6.26}$$

But then this implies

$$L_{N_1 j_0}(\varepsilon,\lambda,\theta,u) \text{ is } N_1\text{-good.}$$

case 2. ($|j_0| < N_1^3$). Since $\|u - u_0\|_{s_1} \leqslant N_1^{-\kappa_1}$ (recall that $\|u_0\|_{s_1} \leqslant 1$ so $\|u\|_{s_1} \leqslant 2$) then

$$\begin{aligned}|L_{N_1,j_0}(\varepsilon,\lambda,\theta,u_0) - L_{N_1,j_0}(\varepsilon,\lambda,\theta,u)|_{s_2} &\leqslant |L_{N_1,j_0}(\varepsilon,\lambda,\theta,u_0) - L_{N_1,j_0}(\varepsilon,\lambda,\theta,u)|_{s_1-\sigma}\\ &\leqslant |(\mu(u_0) - \mu(u))\mathrm{diag}_{|j-j_0|,|\ell|<N_1}|j|^2 + R_N(u_0) - R_N(u)|_{s_1-\sigma}\\ &\lesssim N_1^6\|u-u_0\|_{s_1} \leqslant \frac{1}{2}.\end{aligned} \tag{6.27}$$

By Neumann series and (6.23) one has $|L^{-1}_{N_1,j_0}(\varepsilon,\lambda,\theta,u)|_s \leqslant N_1^{\tau+\delta s}$ for all $s \in [s_0, s_2]$, namely $L_{N_1,j_0}(\varepsilon,\lambda,\theta,u)$ is N_1-good. ■

Lemma 6.11. *Property (S6)$_1$ holds.*

Proof. Let $\lambda \in \mathcal{C}_1 := \mathcal{G}^0_{N_1}(u_0) \cap J_{N_1}(u_0) \cap \hat{\mathcal{I}}$, see (6.16). By the definitions (5.9) and (5.5), and (S6)$_0$, in order to prove that $\lambda \in A_1$, it is sufficient to prove that $\lambda \in \mathfrak{G}_{N_1}(u_0)$. Since $\lambda \in J_{N_1}(u_0)$ the matrix $\|L^{-1}_{N_1}(\varepsilon,\lambda,u_0)\|_0 \leqslant N_1^{\tau_1}$ (see (6.15)) and so (6.23) holds with $j_0 = 0$, $\theta = 0$. Hence $\lambda \in \mathfrak{G}_{N_1}(u_0)$ ■

Now we consider **case 2**.

Lemma 6.12. $\bigcap_{i=1}^{n+1} \mathcal{G}^0_{N_i}(u_{i-1}) \cap \hat{\mathcal{I}} \subseteq \mathcal{G}_{N_p}(u_n) \cap \hat{\mathcal{I}}$.

Proof. By $(S2)_n$ of theorem 5.1 we get $\|u_n - u_{p-1}\|_{s_1} \leqslant \sum_{i=p}^n \|u_i - u_{i-1}\|_{s_1} \leqslant \sum_{i=p}^n N_i^{-\kappa_1 - 1} \leqslant N_p^{-\kappa_1} \sum_{i=p}^n N_i^{-1} \leqslant N_p^{-\kappa_1}$. Hence $(S5)_p (p \leqslant n)$ implies

$$\bigcap_{i=1}^{n+1} \mathcal{G}^0_{N_i}(u_{i-1}) \cap \hat{\mathcal{I}} \subseteq \bigcap_{i=1}^{p} \mathcal{G}^0_{N_i}(u_{i-1}) \cap \hat{\mathcal{I}} \overset{(S5)_p}{\subseteq} \mathcal{G}_{N_p}(u_n) \cap \hat{\mathcal{I}}$$

proving the lemma. ∎

Lemma 6.13. *Property* $(S5)_{n+1}$ *holds.*

Proof. Fix $\lambda \in \bigcap_{i=1}^{n+1} \mathcal{G}^0_{N_i}(u_{i-1}) \cap \hat{\mathcal{I}}$. Reasoning as in the proof of lemma 6.10, it is sufficient to prove that, for all $j_0 \in \mathbb{Z}^d$, $\|u - u_n\|_{s_1} \leqslant N_{n+1}^{-\kappa_1}$, one has

$$\|L^{-1}_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u_n)\|_0 \leqslant N_{n+1}^{\tau_1} \quad \Longrightarrow \quad L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u) \text{ is } N_{n+1}\text{-good}\,. \tag{6.28}$$

We apply the multiscale proposition 6.9 to the matrix $A = L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u_n)$ with $N^\chi = N_{n+1}$ and $N = N_p$, see (6.21). Assumption (H1) holds and (H2) is $\|L^{-1}_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u_n)\|_0 \leqslant N_{n+1}^{\tau_1}$. Lemma 6.12 implies that $\lambda \in \mathcal{G}_{N_p}(u_n) \cap \hat{\mathcal{I}}$ and therefore also (H3) is satisfied by Ansatz 1. But then proposition 6.9 implies

$$|L^{-1}_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u_n)|_s \leqslant \frac{1}{4} N_{n+1}^{\tau} \left(N_{n+1}^{\delta s} + |\mathcal{R}_2(u_n)|_s \right), \quad \forall s \in [s_0, S]\,. \tag{6.29}$$

Then we can follow word by word the proof of lemma 6.10 (with N_{n+1} instead of N_1, and u_n instead of u_0), i.e. we separate the cases $|j_0| > N_{n+1}^3$ and $|j_0| \leqslant N_{n+1}^3$ and the assertion follows. ∎

Lemma 6.14. *Property* $(S6)_{n+1}$ *holds.*

Proof. Again the proof follows word by word the proof of lemma 6.11 with N_{n+1} instead of N_1, and u_n instead of u_0. ∎

Let us finally define the set

$$\mathcal{C}_\varepsilon := \bigcap_{n \geqslant 0} \bar{\mathcal{G}}^0_{N_0^{2^n}} \cap \bar{J}_{N_0^{2^n}} \cap \tilde{\mathcal{I}} \cap \overline{\mathcal{I}} \tag{6.30}$$

where $\tilde{\mathcal{I}} = \tilde{\mathcal{I}}(N_0)$ is defined in Hypothesis 1, $\overline{\mathcal{I}}$ in (6.7) and, for all $N \in \mathbb{N}$,

$$\bar{J}_N := \left\{ \lambda \in \mathcal{I} \;:\; \|L_N^{-1}(\varepsilon, \lambda, u_\varepsilon(\lambda))\|_0 \leqslant N^{\tau_1}/2 \right\}, \tag{6.31}$$

$$\bar{\mathcal{G}}^0_N := \Big\{ \lambda \in \mathcal{I} \;:\; \forall\, j_0 \in \mathbb{Z}^d \text{ there is a covering}$$
$$\bar{B}^0_N(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^\mathfrak{c}} I_q, \text{ with } I_q = I_q(j_0) \text{ intervals with meas}(I_q) \leqslant N^{-\tau_1} \Big\} \tag{6.32}$$

with

$$\bar{B}^0_N(j_0,\varepsilon,\lambda) := \left\{\theta \in \mathbb{R} \ :\ \|L^{-1}_{N,j_0}(\varepsilon,\lambda,\theta,u_\varepsilon(\lambda))\|_0 > N^{\tau_1}/2\right\}. \qquad (6.33)$$

We have the following result.

Lemma 6.15. $\mathcal{C}_\varepsilon \subseteq \mathcal{C}_\infty$.

Proof. We claim that, for all $n \geqslant 0$, the sets $\bar{\mathcal{G}}^0_{N_n} \subseteq \mathcal{G}^0_{N_n}(u_{n-1})$ and $\bar{J}_{N_n} \subseteq J_{N_n}(u_{n-1})$. These inclusions are a consequence of the super-exponential convergence (5.10) of u_n to u_ε. In view of the definitions (6.32) and (6.13), it is sufficient to prove that, $\forall j_0$, if $\theta \notin \bar{B}^0_{N_n}(j_0,\varepsilon,\lambda)$ then $\|L^{-1}_{N_n,j_0}(\theta,u_{n-1})\|_0 \leqslant N_n^{\tau_1}$, namely $\theta \notin B^0_{N_n}(j_0,\varepsilon,\lambda,u_{n-1})$ (recall (6.14)). Once again we have to distinguish two cases

case 1. ($|j_0| > N_n^3$). In this case, arguing again as in the proof of lemma 6.10 one has $|\theta| < 2N_n$, so the eigenvalues of $L_{N_n,j_0}(\theta,u_{n-1})$ are big and hence $\|L^{-1}_{N_n,j_0}(\theta,u_{n-1})\|_0 \leqslant N_n^{\tau_1}$.

case 2. ($|j_0| \leqslant N_n^3$). One has $\|L^{-1}_{N_n,j_0}(\varepsilon,\lambda,\theta,u_\varepsilon)\|_0 \leqslant N_n^{\tau_1}/2$ by (6.33), and so

$$\begin{aligned}\|L^{-1}_{N_n,j_0}(\theta,u_{n-1})\|_0 &\leqslant \|L^{-1}_{N_n,j_0}(\theta,u_\varepsilon)\|_0 \left\|\left(\mathbb{1} + L^{-1}_{N_n,j_0}(\theta,u_\varepsilon)(L_{N_n,j_0}(\theta,u_{n-1}) - L_{N_n,j_0}(\theta,u_\varepsilon))\right)^{-1}\right\|_0 \\ &\leqslant (N_n^{\tau_1}/2)\,2 = N_n^{\tau_1}\end{aligned}$$

by Neumann series expansions. The inclusion $\bar{J}_{N_n} \subseteq J_{N_n}(u_{n-1})$ follow similarly. ■

Theorem 6.6 and lemma 6.15 are essentially theorem 5.5 and lemma 5.21 of [14] respectively, where (4.5) implies Hypothesis 1 of [14] with $\nu_0 \rightsquigarrow \sigma$, lemma 6.4 implies that Hypothesis 2 of [14] is satisfied and Ansatz 1 here is the separation property of Hypothesys 4 in [14]. However we cannot directly apply the result of [14] for the following reason. The constant μ appearing in (6.3) depends on the function at wich the linearized operator is computed; hence one has

$$L_N(\varepsilon,\lambda,\theta,u) - L_N(\varepsilon,\lambda\theta,v) = (\mu(u)-\mu(v))\Delta + \mathcal{R}_2(u) - \mathcal{R}_2(v).$$

The presence of the term $(\mu(u)-\mu(v))\Delta$ forces us to distinguish the cases $|j_0|$ large, where no small divisor appear, and $|j_0|$ small where one argues by Neumann series as in [14].

In what follows we are going to prove that Ansatz 1 is satisfied and later we shall provide measure estimates for $\mathcal{C}_\varepsilon$, thus concluding the proof of our main theorem 1.1.

7. Proof of Ansatz 1

Given $\Sigma \subseteq \mathbb{Z}^\nu \times \mathbb{Z}^d_*$ we define for $\mathfrak{j} \in \mathbb{Z}^d_*$ the section

$$\Sigma^{(\mathfrak{j})} := \{k = (\ell,\mathfrak{j}) \in \Sigma\}.$$

Definition 7.1. Let θ,λ be fixed and $K > 1$. We denote by Σ_K any subset of singular sites of $D(\lambda,\theta)$ in $\mathbb{Z}^\nu \times \mathbb{Z}^d_*$ such that, for all $\mathfrak{j} \in \mathbb{Z}^d_*$, the cardinality of the section $\Sigma_K^{(\mathfrak{j})}$ satisfies $\#\Sigma_K^{(\mathfrak{j})} \leqslant K$.

Definition 7.2 (Γ-chain). Let $\Gamma \geqslant 2$. A sequence $k_0,\dots,k_m \in \mathbb{Z}^\nu \times \mathbb{Z}^d_*$ with $k_p \neq k_q$ for $0 \leqslant p \neq q \leqslant m$ such that

$$\mathrm{dist}(k_{q+1}, k_q) \leqslant \Gamma, \qquad \text{for all } q = 0, \ldots, m-1, \tag{7.1}$$

is called a Γ-chain of length m.

Proposition 7.3 (Separation of Γ-chains). *There exists $C = C(\nu, d)$ and, for any $N_0 \geqslant 2$ a set $\tilde{\mathcal{I}} = \tilde{\mathcal{I}}(N_0)$ defined as*

$$\tilde{\mathcal{I}} := \tilde{\mathcal{I}}(N_0) := \Big\{ \lambda \in [1/2, 3/2] \ : \ |P(\lambda\overline{\omega})| \geqslant \frac{N_0^{-1}}{1 + |p|^{\nu(\nu+1)}}, \forall \text{ non zero polynomial } P(X) \in \mathbb{Z}[X_1, \ldots, X_\nu] \text{ of the form } P(X) = p_0 + \sum_{1 \leqslant i_1 \leqslant i_2 \leqslant \nu} p_{i_1, i_2} X_{i_1} X_{i_2} \Big\} \tag{7.2}$$

such that, for all $\lambda \in \tilde{\mathcal{I}}$, $\theta \in \mathbb{R}$, and for all K, Γ with $K\Gamma \geqslant N_0$, any Γ-chain of singular sites in Σ_K as in definition 7.1, has length $m \leqslant (\Gamma K)^{C(\nu,d)}$.

Proof. The proof is a slight modification of lemma 4.2 of [12] and lemma 3.5 in [14]. First of all, it is sufficient to bound the length of a Γ-chain of singular sites for $D(\lambda, 0)$. Then we consider the quadratic form

$$Q : \mathbb{R} \times \mathbb{R}^r \to \mathbb{R}, \quad Q(x, j) := -x^2 + \mu|j|^2, \tag{7.3}$$

and the associated bilinear form $\Phi = -\Phi_1 + \Phi_2$ where

$$\Phi_1((x, j), (x', j')) := xx', \qquad \Phi_2((x, j), (x', j')) := \mu j \cdot j'. \tag{7.4}$$

For a Γ-chain of sites $\{k_q = (\ell_q, j_q)\}_{q=0,\ldots,\ell}$ which are singular for $D(\lambda, 0)$ (definition 6.1) we have, recalling (6.3) and setting $x_q := \omega \cdot \ell_q$,

$$|Q(x_q, j_q)| < 2, \qquad \forall q = 0, \ldots, \ell.$$

Moreover, by (7.3) and (7.1), we derive $|Q(x_q - x_{q_0}, j_q - j_{q_0})| \leqslant C|q - q_0|^2\Gamma^2$, $\forall 0 \leqslant q, q_0 \leqslant m$, and so

$$|\Phi((x_{q_0}, j_{q_0}), (x_q - x_{q_0}, j_q - j_{q_0}))| \leqslant C'|q - q_0|^2\Gamma^2. \tag{7.5}$$

Now we introduce the subspace of $\mathbb{R}^{1+d}$ given by

$$\mathcal{S} := \mathrm{Span}_{\mathbb{R}}\{(x_q - x_{q_0}, j_q - j_{q_0}) \ : \ q = 0, \ldots, m\}$$

and denote by $\mathfrak{s} \leqslant d + 1$ the dimension of $\mathcal{S}$. Let $\rho > 0$ be a small parameter specified later on. We distinguish two cases.

Case 1. *For all $q_0 = 0, \ldots, m$ one has*

$$\mathrm{Span}_{\mathbb{R}}\{(x_q - x_{q_0}, j_q - j_{q_0}) \ : \ |q - q_0| \leqslant \ell^\rho, \ q = 0, \ldots, m\} = \mathcal{S}. \tag{7.6}$$

In such a case, we select a basis $f_b := (x_{q_b} - x_{q_0}, j_{q_b} - j_{q_0}) = (\omega \cdot \Delta\ell_{q_b}, \Delta j_{q_b})$, $b = 1, \ldots, \mathfrak{s}$ of $\mathcal{S}$, where $\Delta k_{q_b} = (\Delta\ell_{q_b}, \Delta j_{q_b})$ satisfies $|\Delta k_{q_b}| \leqslant C\Gamma|q_b - q_0| \leqslant C\Gamma m^\rho$. Hence we have the bound

$$|f_{q_b}| \leqslant C\Gamma m^\rho, \qquad b = 1, \ldots, \mathfrak{s}. \tag{7.7}$$

Introduce also the matrix $\Omega = (\Omega_b^{b'})_{b,b'=1}^{\mathfrak{s}}$ with $\Omega_b^{b'} := \Phi(f_{b'}, f_b)$, that, according to (7.4), we write

$$\Omega = \Big(-\Phi_1(f_{b'}, f_b) + \Phi_2(f_{b'}, f_b)\Big)_{b,b'=1}^{\mathfrak{s}} = -X + Y, \tag{7.8}$$

where $X_b^{b'} := (\omega \cdot \Delta\ell_{q_{b'}})(\omega \cdot \Delta\ell_{q_b})$ and $Y_b^{b'} := \mu(\Delta j_{q'_b}) \cdot (\Delta j_{q_b})$. The matrix Y has entries in $\mu\mathbb{Z}$ and the matrix X has rank 1 since each column is

$$X^b = (\omega \cdot \Delta\ell_{q_b}) \begin{pmatrix} \omega & \cdot & \Delta\ell_{q_1} \\ & \vdots & \\ \omega & \cdot & \Delta\ell_{q_{\mathfrak{s}}} \end{pmatrix}, \quad b = 1, \dots, \mathfrak{s}.$$

Then, since the determinant of a matrix with two collinear columns $X^b, X^{b'}$, $b \neq b'$, is zero, we get

$$\begin{aligned} P(\omega) : &= \mu^{d+1}\det(\Omega) = \mu^{d+1}\det(-X+Y) \\ &= \mu^{d+1}(\det(Y) - \det(X^1, Y^2, \dots, Y^{\mathfrak{s}}) - \dots - \det(Y^1, \dots, Y^{\mathfrak{s}-1}, X^{\mathfrak{s}})) \end{aligned}$$

which is a quadratic polinomial as in (7.2) with coefficients $\leqslant C(\Gamma m^\rho)^{2(d+1)}$. Note that $P \not\equiv 0$. Indeed, if $P \equiv 0$ then

$$0 = P(\mathrm{i}\omega) = \mu^{d+1}\det(X+Y) = \mu^{d+1}\det(f_b \cdot f_{b'})_{b,b'=1,\dots,\mathfrak{s}} \neq 0$$

because $\{f_b\}_{b=1}^{\mathfrak{s}}$ is a basis of $\mathcal{S}$. This contradiction proves that $P \not\equiv 0$. But then, by (7.2),

$$\mu^{d+1}|\det(\Omega)| = |P(\omega)| \geqslant \frac{N_0^{-1}}{1+|p|^{\nu(\nu+1)}} \geqslant \frac{N_0^{-1}}{(\Gamma m^\rho)^{C(d,\nu)}},$$

the matrix Ω is invertible and

$$|(\Omega^{-1})_b^{b'}| \leqslant C N_0 (\Gamma m^\rho)^{C'(d,\nu)}. \tag{7.9}$$

Now let $\mathcal{S}^\perp := \mathcal{S}^{\perp\Phi} := \{v \in \mathbb{R}^{s+1} \ : \ \Phi(v,f) = 0, \ \forall f \in \mathcal{S}\}$. Since Ω is invertible, the quadratic form $\Phi_{\mathcal{S}}$ is non-degenerate and so $\mathbb{R}^{d+1} = \mathcal{S} \oplus \mathcal{S}^\perp$. We denote $\Pi_{\mathcal{S}} : \mathbb{R}^{d+1} \to \mathcal{S}$ the projector onto $\mathcal{S}$. Writing

$$\Pi_{\mathcal{S}}(x_{q_0}, j_{q_0}) = \sum_{b'=1}^{d+1} a_{b'} f_{b'}, \tag{7.10}$$

and since $f_b \in \mathcal{S}$, $\forall b = 1, \dots, \mathfrak{s}$, we get

$$w_b := \Phi\big((x_{q_0}, j_{q_0}), f_b\big) = \sum_{b'=1}^{\mathfrak{s}} a_{b'}\Phi(f_{b'}, f_b) = \sum_{b'=1}^{\mathfrak{s}} \Omega_b^{b'} a_{b'}$$

where Ω is defined in (7.8). The definition of f_b, the bound (7.5) and (7.6) imply $|w| \leqslant C(\Gamma m^\rho)^2$. Hence, by (7.9), we deduce $|a| = |\Omega^{-1}w| \leqslant C' N_0 (\Gamma m^\rho)^{C(\nu,d)+2}$, whence, by (7.10) and (7.7),

$$|\Pi_{\mathcal{S}}(x_{q_0}, j_{q_0})| \leqslant N_0 (\Gamma m^\rho)^{C'(\nu,d)}.$$

Therefore, for any $q_1, q_2 = 0, \ldots, m$, one has

$$|(x_{q_1}, j_{q_1}) - (x_{q_2}, j_{q_2})| = |\Pi_{\mathcal{S}}(x_{q_1}, j_{q_1}) - \Pi_{\mathcal{S}}(x_{q_2}, j_{q_2})| \leqslant N_0(\Gamma m^\rho)^{C_1(\nu,d)},$$

which in turn implies $|j_{q_1} - j_{q_2}| \leqslant N_0(\Gamma m^\rho)^{C_1(r,d)}$ for all $q_1, q_2 = 0, \ldots, m$. Since all the j_q have d components (being elements of $\mathbb{Z}^d_*$) they are at most $CN_0^d(\Gamma m^\rho)^{C_1(r,d)d}$. We are considering a Γ-chain in Σ_K (see definition 7.1) and so, for each q_0, the number of $q \in \{0, \ldots, m\}$ such that $j_q = j_{q_0}$ is at most K and hence

$$m \leqslant N_0^d(\Gamma m^\rho)^{C_2(\nu,d)}K \leqslant (\Gamma K)^d(\Gamma m^\rho)^{C_2(\nu,d)}K \leqslant m^{\rho C_2(\nu,d)}(\Gamma K)^{d+C_2(\nu,d)}$$

because of the condition $\Gamma K \geqslant N_0$, Choosing $\rho < 1/(2C_2(\nu,d))$ we get $m \leqslant (\Gamma K)^{2(m+C_2(\nu,d))}$.

Case 2. There is $q_0 = 0, \ldots, m$ such that

$$\dim(\mathrm{Span}_{\mathbb{R}}\{(x_q - x_{q_0}, j_q - j_{q_0}) \ : \ |q - q_0| \leqslant m^\rho, \ q = 0, \ldots, m\}) \leqslant \mathfrak{s} - 1.$$

Then we repeat the argument of case 1 for the sub-chain $\{(\ell_q, j_q) \ : \ |q - q_0| \leqslant m^\rho\}$ and obtain a bound for m^ρ. Since this procedure is applied at most $d + 1$ times, at the end we get a bound like $m \leqslant (\Gamma K)^{C_3(\nu,d)}$. ∎

Corollary 7.4. *Ansatz* 1 *is satisfied.*

The proof of corollary 7.4 follows almost word by word section 5.3 in [14]. However there is a minor issue to be discussed, namely that in section 5.3 in [14] it seems that one needs the index j to be in a lattice, whereas of course this is not the case in the present paper since we reduced to the zero mean valued functions. However the lattice structure is needed only in lemma 5.16 of [14] (see remark 5.17 of [14]). In particular if we replace definition 5.14 of [14] with definition 7.5 below, the argument of [14] can be repeated verbatim.

Definition 7.5. A site $k = (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d$ is

- (L, N)-strongly-regular if $L_{N,k}$ is N-good,
- (L, N)-weakly-singular if, otherwise, $L_{N,k}$ is N-bad,
- (L, N)-strongly-good if either it is regular for $D = D(\lambda, \theta)$ (recall definition 6.1) or all the sites $k' = (\ell', j')$ with $\mathrm{dist}(k, k') \leqslant N$ are (L, N)-strongly-regular. Otherwise k is (L, N)-weakly-bad.

8. Measure estimates

We conclude the proof of theorem 1.1 by showing that the set $\mathcal{C}_\varepsilon$ has asymptotically full measure.

One proceeds differently for $|j_0| \geqslant 6N$ and $|j_0| < 6N$. We assume $N \geqslant N_0 > 0$ large enough and $\varepsilon\|\mathcal{R}_2\|_0 \leqslant 1$.

Lemma 8.1. *For all* $j_0 \in \mathbb{Z}^d_*$, $|j_0| \geqslant 6N$, *and for all* $\lambda \in [1/2, 3/2]$ *one has*

$$\bar{B}^0_N(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^{d+\nu+2}} I_q\,, \quad \text{with } I_q = I_q(j_0) \text{ intervals with } \mathrm{meas}(I_q) \leqslant N^{-\tau_1}\,.$$

Proof. First of all, as in the proof **case 1** in lemma 6.10 we see that $\bar{B}^0_N(j_0, \varepsilon, \lambda) \subset \mathbb{R}\setminus [-2N, 2N]$. Now set $B^{0,+}_N := \bar{B}^0_N(j_0, \varepsilon, \lambda) \cap (2N, +\infty)$, $B^{0,-}_N := \bar{B}^0_N(j_0, \varepsilon, \lambda) \cap (-\infty, -2N)$.

Since

$$\partial_\theta L_{N,j_0}(\varepsilon,\lambda,\theta) = \mathrm{diag}_{|\ell|\leqslant N, |j-j_0|\leqslant N} - 2(\omega\cdot\ell+\theta) \geqslant N\mathbb{1},$$

we apply lemma 5.1 of [11] with $\alpha = N^{-\tau_1}$, $\beta = N$ and $|E| \leqslant CN^{\nu+d}$ and obtain

$$B_N^{0,-} \subset \bigcup_{q=1}^{N^{d+\nu+1}} I_q^-, \quad I_q^- = I_q^-(j_0) \text{ intervals with } \mathrm{meas}(I_q) \leqslant N^{-\tau_1}.$$

We can reason in the same way for $B_N^{0,+}$ and the lemma follows. ■

Consider now $|j_0| < 6N$. We obtain a complexity estimate for $\bar{B}_N^0(j_0,\varepsilon,\lambda)$ by knowing the measure of the set

$$\bar{B}_{2,N}^0(j_0,\varepsilon,\lambda) := \left\{\theta\in\mathbb{R} \ : \ \|L_{N,j_0}^{-1}(\lambda,\varepsilon,\theta)\|_0 > N^{\tau_1}/2\right\}.$$

Lemma 8.2. *For all $|j_0| < 6N$ and all $\lambda\in[1/2,3/2]$ one has*

$$\bar{B}_{2,N}^0(j_0,\varepsilon,\lambda) \subset I_N := [-10\sqrt{d}N, 10\sqrt{d}N].$$

Proof. If $|\theta| > 10\sqrt{d}N$ one has $|\omega\cdot\ell+\theta| \geqslant |\theta| - |\omega\cdot\ell| > (10\sqrt{\nu}-(3/2))N > 8\sqrt{d}N$ and then all the eigenvalues satisfy

$$\mu_{\ell,j}(\theta) = -(\omega\cdot\ell+\theta)^2 + \mu|j|^2 + O(\varepsilon\|\mathcal{R}_2\|_0) \leqslant -62dN^2, \quad \forall|\theta| > 10\sqrt{d}N,$$

proving the lemma. ■

Lemma 8.3. *For all $|j_0| \leqslant 6N$ and all $\lambda\in[1/2,3/2]$ one has*

$$\bar{B}_N^0(j_0,\varepsilon,\lambda) \subset \bigcup_{q=1}^{\hat{C}\mathfrak{M}N^{\tau_1+1}} I_q,\ I_q = I_q(j_0) \text{ intervals with } \mathrm{meas}(I_q) \leqslant N^{-\tau_1}$$

where $\mathfrak{M} := \mathrm{meas}(\bar{B}_{2,N}^0(j_0,\varepsilon,\lambda))$ and $\hat{C} = \hat{C}(d)$.

Proof. This is lemma 5.5 of [11], where our exponent τ_1 is denoted by τ. ■

Lemmas 8.2 and 8.3 imply that for all $\lambda\in[1/2,3/2]$ the set $\bar{B}_N^0(j_0,\varepsilon,\lambda)$ can be covered by $\sim N^{\tau_1+2}$ intervals of length $\leqslant N^{-\tau_1}$. This estimate is not enough. Now we prove that for 'most' λ the number of such intervals does not depend on τ_1, by showing that $\mathfrak{M} = O(N^{\mathfrak{c}-\tau_1})$ where $\mathfrak{c}$ depends only on the dimensions (to be computed). To this purpose first we provide an estimate for the set

$$\boldsymbol{B}_{2,N}^0(j_0,\varepsilon) := \left\{(\lambda,\theta)\in[1/2,3/2]\times\mathbb{R} \ : \ \|L_{N,j_0}^{-1}(\varepsilon,\lambda,\theta)\|_0 > N^{\tau_1}/2\right\}.$$

Then in lemma 8.5 we use Fubini theorem to obtain the desired bound for $\mathrm{meas}(\bar{B}_{2,N}^0(j_0,\varepsilon,\lambda))$.

Lemma 8.4. *For all $|j_0| < 6N$ one has* $\mathrm{meas}(\boldsymbol{B}_{2,N}^0(j_0,\varepsilon)) \lesssim N^{-\tau_1+\nu+d+1}$.

Proof. Let us introduce the variables

$$\zeta = \frac{1}{\lambda^2},\ \eta = \frac{\theta}{\lambda}, \qquad (\zeta,\eta)\in[4/9,4]\times[-20\sqrt{d}N, 20\sqrt{d}N] =: [4/9,4]\times J_N, \tag{8.1}$$

and set

$$L(\zeta,\eta) := \lambda^{-2} L_{N,j_0}(\varepsilon,\lambda,\theta) = \mathrm{diag}_{|\ell|\leqslant N, |j-j_0|\leqslant N}\Big(\big(-\overline{\omega}\cdot\ell+\eta\big)^2 + \zeta\mu(\zeta^{-1/2})|j|^2\Big) + \zeta\mathcal{R}_2(\varepsilon, 1/\sqrt{\zeta}).$$

Note that, since $\|\mu - 1\| \lesssim \varepsilon$, one has

$$\min_{j\in\mathbb{Z}^d_*} \mu|j|^2 \geqslant \frac{1}{2}. \tag{8.2}$$

Then, except for (ζ,η) in a set of measure $O(N^{-\tau_1+\nu+d+1})$ one has

$$\|L(\zeta,\eta)^{-1}\|_0 \leqslant N^{\tau_1}/8. \tag{8.3}$$

Indeed

$$\partial_\zeta L(\zeta,\eta) = \mathrm{diag}_{|\ell|\leqslant N, |j-j_0|\leqslant N}\left(\mu(\zeta^{-1/2})|j|^2 - \frac{1}{2}\zeta^{-1/2}\partial_\lambda\mu(\zeta^{-1/2})\right) + \mathcal{R}_2(\varepsilon,1/\sqrt{\zeta}) - \frac{1}{2}\zeta^{-1/2}\partial_\lambda\mathcal{R}_2 \overset{(8.2)}{\geqslant} \frac{1}{4},$$

for ε small (we used that $\zeta \in [4/9, 4]$ and $|\partial_\lambda\mu| < 1/2$). Therefore lemma 5.1 of [11] implies that for each η, the set of ζ such that at least one eigenvalue of $L(\zeta,\eta)$ has modulus $\leqslant 8N^{-\tau_1}$, is contained in the union of $O(N^{d+\nu})$ intervals with length $O(N^{-\tau_1})$ and hence has measure $\leqslant O(N^{-\tau_1+d+\nu})$. Integrating in $\eta \in J_N$ we obtain (8.3) except in a set with measure $O(N^{-\tau_1+d+\nu+1})$. The same measure estimates hold in the original variables (λ,θ) in (8.1). Finally (8.3) implies

$$\|L^{-1}_{N,j_0}(\varepsilon,\lambda,\theta)\|_0 \leqslant \lambda^{-2}N^{\tau_1}/8 \leqslant N^{\tau_1}/2,$$

for all $(\lambda,\theta) \in [1/2, 2/3] \times \mathbb{R}$ except in a set with measure $\leqslant O(N^{-\tau_1+d+\nu+1})$. ∎

Note that the same argument can be used to show that

$$\mathrm{meas}([1/2,3/2] \setminus \bar{\mathfrak{G}}_N) \leqslant N^{-\tau_1+d+\nu+1} \tag{8.4}$$

where $\bar{\mathfrak{G}}_N$ is defined in (6.31).

Define the set

$$\mathcal{F}_N(j_0) := \left\{\lambda \in [1/2,3/2] \ : \ \mathrm{meas}(\bar{B}^0_{2,N}(j_0,\varepsilon,\lambda)) \geqslant \hat{C}N^{-\tau_1+d+\mathfrak{d}+r+2}\right\} \tag{8.5}$$

where $\hat{C}$ is the constant appearing in lemma 8.3.

Lemma 8.5. *For all* $|j_0| \leqslant 6N$ *one has* $\mathrm{meas}(\mathcal{F}_N(j_0)) = O(N^{-d-1})$.

Proof. By Fubini theorem we have

$$\mathrm{meas}(\boldsymbol{B}^0_{2,N}(j_0,\varepsilon)) = \int_{1/2}^{3/2} \mathrm{d}\lambda\, \mathrm{meas}(\bar{B}^0_{2,N}(j_0,\varepsilon,\lambda)).$$

Now, for any $\beta > 0$, using lemma 8.4 we have

$$\begin{aligned} CN^{-\tau_1+d+\nu+1} &\geqslant \int_{1/2}^{3/2} \mathrm{d}\lambda\, \mathrm{meas}(\bar{B}^0_{2,N}(j_0,\varepsilon,\lambda)) \\ &\geqslant \beta\mathrm{meas}(\{\lambda \in [1/2,3/2] \ : \ \mathrm{meas}(\bar{B}^0_{2,N}(j_0,\varepsilon,\lambda)) \geqslant \beta\}) \end{aligned}$$

and for $\beta = \hat{C}N^{-\tau_1+2d+\nu+2}$ we prove the lemma (recall (8.5)). ∎

Lemma 8.6. *If* $\tau_0 > d + 3\nu + 1$ *then* $\text{meas}([1/2, 3/2] \setminus \overline{\mathcal{I}}) = O(N_0^{-1})$ *where* $\overline{\mathcal{I}}$ *is defined in (6.7).*

Proof. Let us write

$$[1/2, 3/2] \setminus \overline{\mathcal{I}} = \bigcup_{|\ell|,|j|\leqslant N_0} \mathcal{R}_{l,j}, \qquad \mathcal{R}_{\ell,j} := \Big\{ \lambda \in \mathcal{I} \ :\ |(\lambda\overline{\omega}\cdot\ell)^2 - |j|^2| \leqslant N_0^{-\tau_0} \Big\}.$$

Since $j \in \mathbb{Z}^d_*$, then $\mathcal{R}_{0,j} = \emptyset$ if $N_0 > 1$. For $\ell \neq 0$, using the Diophantine condition (1.2), we get $\text{meas}(\mathcal{R}_{\ell,j}) \leqslant CN_0^{-\tau_0+2\nu}$, so that

$$\text{meas}([1/2, 3/2] \setminus \overline{\mathcal{I}}) \leqslant \sum_{|\ell|,|j|\leqslant N_0} \text{meas}(\mathcal{R}_{\ell,j}) \leqslant CN_0^{-\tau_0+d+3\nu} = O(N_0^{-1})$$

because $\tau_0 - d - 3\nu > 1$. ■

The measure of the set $\tilde{\mathcal{I}}$ in (7.2) is estimated in [11]-lemma 6.3 (where $\tilde{\mathcal{I}}$ is denoted by $\tilde{\mathcal{G}}$).

Lemma 8.7. *If* $\gamma < \min(1/4, \gamma_0/4)$ *(where* γ_0 *is that in (1.4)) then* $\text{meas}([1/2, 3/2] \setminus \tilde{\mathcal{I}}) = O(\gamma)$.

To conclude the measure esitimate we note that by the definition in (8.5) for all $\lambda \notin \mathcal{F}_N(j_0)$ one has $\text{meas}(\bar{B}^0_{2,N}(j_0, \varepsilon, \lambda)) < O(N^{-\tau_1+2d+\nu+2})$. Thus for any $\lambda \notin \mathcal{F}_N(j_0)$, applying lemma 8.3 we have

$$\bar{B}^0_N(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^{2d+\nu+4}} I_q\,, \quad I_q \text{ intervals with } \text{meas}(I_q) \leqslant N^{-\tau_1}\,.$$

But then, using also lemma 8.1, we have that (recall (6.32) with $\mathfrak{c} = 2d + \nu + 4$)

$$[1/2, 3/2] \setminus \bar{\mathcal{G}}^0_N \subset \bigcup_{|j_0|\leqslant(\mathfrak{c}+5)\mathfrak{c}^{-1}N} \mathcal{F}_N(j_0)\,.$$

Hence, using lemma 8.5,

$$\text{meas}(\mathcal{I} \setminus \bar{\mathcal{G}}^0_N) \leqslant \sum_{|j_0|\leqslant 6N} \text{meas}(\mathcal{F}_N(j_0)) \leqslant O(N^{-1}).$$

Moreover by (8.4) with $\tau_1 > d + \nu + 2$ we get

$$\text{meas}(\mathcal{I} \setminus \bar{\mathfrak{G}}_N) = O(N^{-1}), \tag{8.6}$$

and finally, lemmas 8.6 and 8.7 with $\gamma = N_0^{-1}$ imply

$$\text{meas}(\mathcal{I} \setminus (\overline{\mathcal{I}} \cap \tilde{\mathcal{I}})) = O(N_0^{-1})\,.$$

Putting these estimates together and recalling the definition (6.30) of $\mathcal{C}_\varepsilon$, we have that

$$\begin{aligned} \text{meas}(\mathcal{I} \setminus \mathcal{C}_\varepsilon) &= \text{meas}\Big(\bigcup_{n\geqslant 0} (\bar{\mathcal{G}}^0_{N_n})^c \bigcup_{n\geqslant 0} (\bar{\mathfrak{G}}_{N_n})^c \cup \tilde{\mathcal{I}}^c \cup \overline{\mathcal{I}}^c \Big) \\ &\leqslant \sum_{n\geqslant 0} \text{meas}(\mathcal{I} \setminus \bar{\mathcal{G}}^0_{N_n}) + \sum_{n\geqslant 0} \text{meas}(\mathcal{I} \setminus \bar{\mathfrak{G}}_{N_n}) + \text{meas}(\mathcal{I} \setminus (\overline{\mathcal{I}} \cap \tilde{\mathcal{I}})) \\ &\overset{(8.6)}{\lesssim} \sum_{n\geqslant 0} N_n^{-1} + N_0^{-1} \lesssim N_0^{-1} \lesssim \varepsilon^{1/(S+1)} \end{aligned} \tag{8.7}$$

i.e. $\mathcal{C}_\varepsilon$ has asymptotically full measure. ■

9. Linear stability

In this section we discuss the linear stability of the quasi-periodic solutions found in theorem 1.1. In order to precisely state the result we need to introduce some more notations. For any $s \geqslant 0$, we define the Sobolev spaces $H^s(\mathbb{T}^d) = H^s(\mathbb{T}^d, \mathbb{C})$, $H^s(\mathbb{T}^d, \mathbb{R})$ respectively of complex and real valued functions

$$H^s(\mathbb{T}^d) := \Big\{u(x) = \sum_{j\in\mathbb{Z}^d} u_j \mathrm{e}^{\mathrm{i}jx} : \|u\|_{H^s_x}^2 := \sum_{j\in\mathbb{Z}^d} \langle j\rangle^{2s}|u_j|^2 < +\infty\Big\}, \quad H^s(\mathbb{T}^d, \mathbb{R}) := \big\{u \in H^s(\mathbb{T}^d) : u = \overline{u}\big\} \tag{9.1}$$

where

$$\langle j\rangle := \max\{1, |j|\}, \quad |j| := \sqrt{j_1^2 + \ldots + j_d^2}, \quad \forall j = (j_1, \ldots, j_d) \in \mathbb{Z}^d.$$

Moreover we define

$$H^s_0(\mathbb{T}^d) := \Big\{u \in H^s(\mathbb{T}^d) : \int_{\mathbb{T}^d} u(x)\,\mathrm{d}x = 0\Big\}, \quad H^s_0(\mathbb{T}^d, \mathbb{R}) := \Big\{u \in H^s(\mathbb{T}^d, \mathbb{R}) : \int_{\mathbb{T}^d} u(x)\,\mathrm{d}x = 0\Big\} \tag{9.2}$$

and introduce the real subspace $\mathbf{H}^s_0(\mathbb{T}^d)$ of $H^s_0(\mathbb{T}^d) \times H^s_0(\mathbb{T}^d)$

$$\mathbf{H}^s_0(\mathbb{T}^d) := \big\{\mathbf{u} := (u, \overline{u}) : u \in H^s_0(\mathbb{T}^d)\big\}, \quad \text{equipped with the norm} \quad \|\mathbf{u}\|_{\mathbf{H}^s_x} := \|u\|_{H^s_x}.$$

Given a linear operator $\mathcal{R} : L^2_0(\mathbb{T}^d) \to L^2_0(\mathbb{T}^d)$ (where $L^2_0(\mathbb{T}^d) := H^0_0(\mathbb{T}^d)$), we define its Fourier coefficients with respect to the exponential basis $\{\mathrm{e}^{\mathrm{i}j\cdot x} : j \in \mathbb{Z}^d \setminus \{0\}\}$ of $L^2_0(\mathbb{T}^d)$ as

$$\mathcal{R}_j^{j'} := \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} \mathcal{R}[\mathrm{e}^{\mathrm{i}j'\cdot x}]\mathrm{e}^{-\mathrm{i}j\cdot x}\,\mathrm{d}x, \qquad \forall j, j' \in \mathbb{Z}^d \setminus \{0\}, \tag{9.3}$$

and we denote by $\overline{\mathcal{R}}$ the linear operator such that $\overline{\mathcal{R}}[u] = \overline{\mathcal{R}[\overline{u}]}$, for any $u \in L^2_0(\mathbb{T}^d)$.

We say that an operator $\mathcal{R}$ is *block diagonal* if $\mathcal{R}_j^{j'} = 0$ for any $j, j' \in \mathbb{Z}^d \setminus \{0\}$ with $|j| \neq |j'|$.

Linearizing the equation (1.5) along a quasi-periodic solution $u(\omega t, x)$, $\omega = \lambda\bar{\omega}$ for $\lambda \in \mathcal{C}_\varepsilon$ one gets a linear wave equation of the form

$$\partial_{tt}h - \Delta h + \varepsilon\mathcal{P}(\omega t)[h] = 0 \tag{9.4}$$

where the linear operator $\mathcal{P}(\omega t)$ is given by

$$\mathcal{P}(\varphi) := -a(\varphi)\Delta - \mathcal{R}(\varphi), \quad \varphi \in \mathbb{T}^\nu$$

$$a(\varphi) = \int_{\mathbb{T}^d} |\nabla u(\varphi, x)|^2\,\mathrm{d}x, \quad \mathcal{R}(\varphi)[h] = -2\Delta u(\varphi, x)\int_{\mathbb{T}^d} \Delta u(\varphi, y)h(y)\,dy, \quad \varphi \in \mathbb{T}^\nu. \tag{9.5}$$

Writing the equation (9.4) as a first order system one obtains

$$\begin{cases} \partial_t h = \psi \\ \partial_t \psi = \Delta h - \varepsilon\mathcal{P}(\omega t)[h]. \end{cases} \tag{9.6}$$

The following theorem holds.

Theorem 9.1. *There exists a strictly positive integer $q_0 = q_0(\nu, d)$ possibly larger than $\bar{q}(\nu, d)$ appearing in theorem 1.1 such that for any $q \geqslant q_0$ there exists $\varepsilon_1 = \varepsilon_1(q, \nu, d) > 0$, possibly smaller than ε_0 of theorem 1.1 and $\mathfrak{S}_q := \mathfrak{S}(q, \nu, d)$, with $1/2 < \mathfrak{S}_q < q$ such that for any $f \in C^q(\mathbb{T}^\nu \times \mathbb{T}^d, \mathbb{R})$ satisfying the hyphothesis (1.3) then for any $\varepsilon \in (0, \varepsilon_1)$ there exists a Borel set $\mathcal{O}_\varepsilon \subset \mathcal{C}_\varepsilon$ of asymptotically full Lebesgue measure, i.e.*

$$|\mathcal{O}_\varepsilon| \to 1 \qquad \text{as} \qquad \varepsilon \to 0, \tag{9.7}$$

such that the following holds: for all $\lambda \in \mathcal{O}_\varepsilon$ *and* $\varphi \in \mathbb{T}^\nu$, *there exists a bounded linear invertible operator* $\mathcal{W}_\infty(\varphi) = \mathcal{W}_\infty(\varphi; \lambda)$ *such that for any* $\frac{1}{2} \leqslant s \leqslant \mathfrak{S}_q$

$$\mathcal{W}_\infty(\varphi) : \mathbf{H}_0^s(\mathbb{T}^d) \to H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$$

satisfying the following property: $(h(t,\cdot), \psi(t,\cdot))$ *is a solution of (9.6) in* $H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$ *if and only if*

$$\mathbf{w}(t,\cdot) = (w(t,\cdot), \overline{w}(t,\cdot)) = \mathcal{W}_\infty(\lambda\bar{\omega}t)^{-1}[(h(t,\cdot), \psi(t,\cdot))]$$

is a solution in $\mathbf{H}_0^s(\mathbb{T}^d)$ *of the PDE with constant coefficients*

$$\partial_t \mathbf{w} = \mathcal{D}_\infty \mathbf{w}, \qquad \mathcal{D}_\infty := \mathrm{i}\begin{pmatrix} -\mathcal{D}_\infty^{(1)} & 0 \\ 0 & \overline{\mathcal{D}}_\infty^{(1)} \end{pmatrix} \tag{9.8}$$

where for any $s \geqslant 1$, $\mathcal{D}_\infty^{(1)} : H_0^s(\mathbb{T}^d) \to H_0^{s-1}(\mathbb{T}^d)$ *is a linear, time-independent,* L^2*-self-adjoint, block-diagonal operator.*

From theorem 9.1 we deduce the linear stability of (1.5), i.e. the following result.

Corollary 9.2. *For any* $\lambda \in \mathcal{O}_\varepsilon$ *and any initial data* $(h^{(0)}, \psi^{(0)}) \in H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$ *with* $1/2 \leqslant s \leqslant \mathfrak{S}_q$, *the solution* $t \in \mathbb{R} \mapsto (h(t,\cdot), \psi(t,\cdot)) \in H_0^{s+\frac{1}{2}}(\mathbb{T}^d, \mathbb{R}) \times H_0^{s-\frac{1}{2}}(\mathbb{T}^d, \mathbb{R})$ *of the Cauchy problem*

$$\begin{cases} \partial_t h = \psi \\ \partial_t \psi = \Delta h - \varepsilon \mathcal{P}(\omega t)[h]. \\ h(0,\cdot) = h^{(0)} \\ \psi(0,\cdot) = \psi^{(0)} \end{cases} \tag{9.9}$$

is stable, namely

$$\sup_{t\in\mathbb{R}} \Big(\|h(t,\cdot)\|_{H_x^{s+\frac{1}{2}}} + \|\psi(t,\cdot)\|_{H_x^{s-\frac{1}{2}}} \Big) \leqslant C_q \big(\|h^{(0)}\|_{H_x^{s+\frac{1}{2}}} + \|\psi^{(0)}\|_{H_x^{s-\frac{1}{2}}} \big)$$

for some constant $C_q = C(q, \nu, d) > 0$.

The proof of this theorem is essentially the same as the one of theorem 1.1 in [42]. The only difference is that in the present case the frequency ω is constrained to the diophantine direction $\bar{\omega}$ in (1.2), namely $\omega = \lambda\bar{\omega}$, $\lambda \in \mathcal{C}_\varepsilon$. Hence we only need to prove the measure estimates for the set $\mathcal{O}_\varepsilon$.

Note that in theorem 1.1 of [42] it is required that the coefficients of the perturbations are sufficiently smooth functions. In the present case, this hypothesis is satisfied by taking the forcing term $f \in C^q$ with q large enough. Indeed, by theorem 1.1, the quasi-periodic solution $u(\omega t, x)$ is in H^S and $S = S(\nu, d, q)$ is increasing w.r.t. to q. By recalling the definition (9.5), we then have that the coefficients of the operator $\mathcal{P}(\varphi)$ are smooth enough if f is smooth enough.

The proof of theorem 1.1 of [42] is based on a reduction procedure which conjugates the vector field

$$\mathcal{L}(\omega t) := \begin{pmatrix} 0 & 1 \\ \Delta - \varepsilon \mathcal{P}(\omega t) & 0 \end{pmatrix}$$

of the equation (9.6) to the constant coefficients block diagonal vector field $\mathcal{D}_\infty$ in (9.8). Such a procedure is developed in sections 3 and 4 of [42]. Actually the linearized Kirchhoff equation (9.4) is a particular case of the wave equations considered in [42]. In section 3 of [42], the vector field $\mathcal{L}(\omega t)$ is conjugated to another one which is an arbitrarily regularizing perturbation of a constant coefficients diagonal vector field. In order to perform this reduction, one needs the frequency ω to be diophantine. This is satified in our setup, since $\omega = \lambda\bar{\omega}$ and $\bar{\omega}$ is a diophantine frequency vector (see (1.2)). Then, in section 4 of [42], it is developed a KAM reducibility scheme which completes the block-diagonalization procedure. Along such an iterative scheme, all the remainders arising at any step are smoothing operators. Therefore one can impose very weak second order Melnikov non resonance conditions with *loss of space derivatives*.

In the rest of this section, we prove the measure estimates (9.7).

9.1. Measure estimates of the set $\mathcal{O}_\varepsilon$

Let us denote by $\sigma_0(\sqrt{-\Delta})$ the spectrum of the operator $\sqrt{-\Delta}$ restricted to the functions with zero average, i.e.

$$\sigma_0(\sqrt{-\Delta}) := \Big\{ |j| = \sqrt{j_1^2 + \ldots + j_d^2} : j = (j_1, \ldots, j_d) \in \mathbb{Z}^d \setminus \{0\} \Big\}.$$

Moreover, given $\mathcal{I}_o \subseteq \mathcal{I} = [1/2, 3/2]$, $\gamma > 0$ and a Lipschitz function $f : \mathcal{I}_o \to \mathbb{R}$ we define

$$|f|^{\mathrm{Lip}(\gamma)} := |f|^{\sup} + \gamma |f|^{\mathrm{lip}},$$
$$|f|^{\sup} := \sup_{\lambda \in \mathcal{I}_o} |f(\lambda)|, \qquad |f|^{\mathrm{lip}} := \sup_{\substack{\lambda_1, \lambda_2 \in \mathcal{I}_o \\ \lambda_1 \neq \lambda_2}} \frac{|f(\lambda_1) - f(\lambda_2)|}{|\lambda_1 - \lambda_2|}$$

the rest of this section is devoted to the proof of the following

Theorem 9.3. *One has that* $\lim_{\varepsilon \to 0} |\mathcal{C}_\varepsilon \setminus \mathcal{O}_\varepsilon| = 0$.

First of all we need some notation. For any $\alpha \in \sigma_0(\sqrt{-\Delta})$ we set

$$n_\alpha := \mathrm{card}\{ j \in \mathbb{Z}^d \setminus \{0\} : |j| = \alpha \} \simeq \alpha^{d-1}, \tag{9.10}$$

and the eigenvalues of $\mathcal{D}_\infty^{(1)}$ in (9.8) are Lipschitz functions $\mu_k^{(\alpha)} : \mathcal{C}_\varepsilon \to \mathbb{R}$ for $k = 1, \ldots, n_\alpha$ of the form

$$\mu_k^{(\alpha)}(\lambda) = m(\lambda)\, \alpha + \mathtt{r}_k^{(\alpha)}(\lambda), \qquad |m - 1|^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon, \qquad |\mathtt{r}_k^{(\alpha)}|^{\mathrm{Lip}(\gamma)} \lesssim \varepsilon\, \alpha^{-1}, \tag{9.11}$$

and $\mathtt{r}_k^{(\alpha)}$ are produced by the algorithm; see [42], section 5.

Then, for any $(\ell, \alpha, \beta) \in \mathbb{Z}^\nu \times \sigma_0(\sqrt{-\Delta}) \times \sigma_0(\sqrt{-\Delta})$, $(\ell, \alpha, \beta) \neq (0, \alpha, \alpha)$, $k = 1, \ldots, n_\alpha$, $j = 1, \ldots, n_\beta$

$$R_{kj}(\ell, \alpha, \beta) := \Big\{ \lambda \in \mathcal{C}_\varepsilon : |\lambda \bar{\omega} \cdot \ell + \mu_k^{(\alpha)}(\lambda) - \mu_j^{(\beta)}(\lambda)| < \frac{2\gamma}{\langle \ell \rangle^\tau \alpha^{\mathtt{d}} \beta^{\mathtt{d}}} \Big\} \tag{9.12}$$

and for any $(\ell, \alpha, \beta) \in \mathbb{Z}^\nu \times \sigma_0(\sqrt{-\Delta}) \times \sigma_0(\sqrt{-\Delta})$, $k = 1, \ldots, n_\alpha$, $j = 1, \ldots, n_\beta$

$$Q_{kj}(\ell,\alpha,\beta) := \Big\{\lambda \in \mathcal{C}_\varepsilon : |\lambda\bar\omega\cdot\ell + \mu_k^{(\alpha)}(\lambda) + \mu_j^{(\beta)}(\lambda)| < \frac{2\gamma(\alpha+\beta)}{\langle\ell\rangle^\tau}\Big\}, \tag{9.13}$$

and we fix the constants $\mathtt{d}$ and τ as

$$\mathtt{d} := 2d+2, \qquad \tau := \nu + 4d. \tag{9.14}$$

Then, setting

$$R(\ell,\alpha,\beta) := \bigcup_{k=1}^{n_\alpha}\bigcup_{j=1}^{n_\beta} R_{kj}(\ell,\alpha,\beta), \quad \forall(\ell,\alpha,\beta)\in\mathbb{Z}^\nu\times\sigma_0(\sqrt{-\Delta})\times\sigma_0(\sqrt{-\Delta}), \quad (\ell,\alpha,\beta)\neq(0,\alpha,\alpha), \tag{9.15}$$

$$Q(\ell,\alpha,\beta) := \bigcup_{k=1}^{n_\alpha}\bigcup_{j=1}^{n_\beta} Q_{kj}(\ell,\alpha,\beta), \quad \forall(\ell,\alpha,\beta)\in\mathbb{Z}^\nu\times\sigma_0(\sqrt{-\Delta})\times\sigma_0(\sqrt{-\Delta}), \tag{9.16}$$

and arguing as in section 5 of [42] one can show that

$$\mathcal{C}_\varepsilon\setminus\mathcal{O}_\varepsilon \subseteq \bigcup_{\substack{(\ell,\alpha,\beta)\in\mathbb{Z}^\nu\times\sigma_0(\sqrt{-\Delta})\times\sigma_0(\sqrt{-\Delta})\\(\ell,\alpha,\beta)\neq(0,\alpha,\alpha)}} R(\ell,\alpha,\beta) \quad \bigcup_{(\ell,\alpha,\beta)\in\mathbb{Z}^\nu\times\sigma_0(\sqrt{-\Delta})\times\sigma_0(\sqrt{-\Delta})} Q(\ell,\alpha,\beta). \tag{9.17}$$

The constant γ appearing in (9.12) and (9.13) can be linked to ε by setting

$$\gamma := \varepsilon^{\mathfrak{a}}, \quad 0<\mathfrak{a}<1. \tag{9.18}$$

Note that this is the same setting as in [42] where the sets $R_{kj}(\ell,\alpha,\beta)$, $Q_{kj}(\ell,\alpha,\beta)$, $R(\ell,\alpha,\beta)$ and $Q(\ell,\alpha,\beta)$ were denoted as $\tilde R_{kj}(\ell,\alpha,\beta)$, $\tilde Q_{kj}(\ell,\alpha,\beta)$, $\tilde R(\ell,\alpha,\beta)$ and $\tilde Q(\ell,\alpha,\beta)$ respectively.

Lemma 9.4. *For ε small enough, the following holds.*

(i) If $R(\ell,\alpha,\beta)\neq\emptyset$, then $\alpha^{-1}+\beta^{-1}\lesssim|\bar\omega\cdot\ell|$. Moreover for any $\alpha,\beta\in\sigma_0(\sqrt{-\Delta})$, $\alpha\neq\beta$, then $R(0,\alpha,\beta)=\emptyset$.

(ii) If $Q(\ell,\alpha,\beta)\neq\emptyset$, then $\alpha+\beta\lesssim|\bar\omega\cdot\ell|$. As a consequence $\alpha,\beta\lesssim\langle\ell\rangle$. Moreover for any $\alpha,\beta\in\sigma_0(\sqrt{-\Delta})$ then $Q(0,\alpha,\beta)=\emptyset$.

Proof. We prove item (i). The proof of item (ii) is similar. Assume that $R(\ell,\alpha,\beta)\neq\emptyset$. Then there exist $k\in\{1,\ldots,n_\alpha\}$, $j\in\{1,\ldots,n_\beta\}$ such that $R_{kj}(\ell,\alpha,\beta)\neq\emptyset$. For any $\lambda\in R_{kj}(\ell,\alpha,\beta)$, one has

$$\begin{aligned}|\mu_k^{(\alpha)}(\lambda)-\mu_j^{(\beta)}(\lambda)| &< \frac{2\gamma}{\langle\ell\rangle^\tau\alpha^{\mathtt{d}}\beta^{\mathtt{d}}}+\frac{3}{2}|\bar\omega\cdot\ell| \\ &\overset{\tau,\mathtt{d}>1}{<} \gamma(\alpha^{-1}+\beta^{-1})+\frac{3}{2}|\bar\omega\cdot\ell|.\end{aligned} \tag{9.19}$$

Furthermore, by (9.11), for ε small enough,

$$\begin{aligned}|\mu_k^{(\alpha)}-\mu_j^{(\beta)}| &\geqslant |m(\lambda)||\alpha-\beta| - |\mathtt{r}_k^{(\alpha)}| - |\mathtt{r}_j^{(\beta)}| \\ &\geqslant \frac{1}{2}|\alpha-\beta| - C_0\varepsilon\big(\alpha^{-1}+\beta^{-1}\big),\end{aligned} \tag{9.20}$$

for some constant $C_0>0$. By lemma A.1-(ii) in [42], one has that $|\alpha-\beta|\geqslant C(\alpha^{-1}+\beta^{-1})$ for some constant $C>0$, hence (9.20) implies that for ε small enough one gets

$$|\mu_k^{(\alpha)}-\mu_j^{(\beta)}| \geqslant C_1(\alpha^{-1}+\beta^{-1}) \tag{9.21}$$

for some constant $C_1 > 0$. Then (9.19) and (9.21) imply that for γ small enough (or ε small enough, see (9.18)) one has $\alpha^{-1} + \beta^{-1} \lesssim |\bar{\omega} \cdot \ell|$. The inequality (9.21) implies also that if $\alpha, \beta \in \sigma_0(\sqrt{-\Delta})$ with $\alpha \neq \beta$, then $R_{kj}(0, \alpha, \beta) = \emptyset$ for any $k \in \{1, \ldots, n_\alpha\}$, $j \in \{1, \ldots, n_\beta\}$. ■

Lemma 9.5. *For ε small enough, the following holds:*

(i) For any $(\ell, \alpha, \beta) \in \mathbb{Z}^\nu \times \sigma_0(\sqrt{-\Delta}) \times \sigma_0(\sqrt{-\Delta})$, $(\ell, \alpha, \beta) \neq (0, \alpha, \alpha)$, if $R(\ell, \alpha, \beta) \neq \emptyset$ then $|R(\ell, \alpha, \beta)| \lesssim \gamma \alpha^{d+1-\mathtt{d}} \beta^{d+1-\mathtt{d}} \langle \ell \rangle^{-\tau}$.

(ii) For any $(\ell, \alpha, \beta) \in \mathbb{Z}^\nu \times \sigma_0(\sqrt{-\Delta}) \times \sigma_0(\sqrt{-\Delta})$, if $Q(\ell, \alpha, \beta) \neq \emptyset$ then $|Q(\ell, \alpha, \beta)| \lesssim \gamma \alpha^{d-1} \beta^{d-1} \langle \ell \rangle^{-\tau}$.

Proof. Let us prove item (i). The proof of item (ii) follows by using similar arguments. Let $(\ell, \alpha, \beta) \in \mathbb{Z}^\nu \times \sigma_0(\sqrt{-\Delta}) \times \sigma_0(\sqrt{-\Delta})$ with $(\ell, \alpha, \beta) \neq (0, \alpha, \alpha)$. By (9.15), it is enough to estimate the measure of the set $R_{kj}(\ell, \alpha, \beta)$ for any $k = 1, \ldots, n_\alpha$, $j = 1, \ldots, n_\beta$. We define

$$\phi(\lambda) := \lambda \bar{\omega} \cdot \ell + \lambda_k^{(\alpha)}(\lambda) - \lambda_j^{(\beta)}(\lambda). \tag{9.22}$$

Hence

$$R_{kj}(\ell, \alpha, \beta) = \Big\{ \lambda \in \mathcal{C}_\varepsilon : |\phi(\lambda)| < \frac{2\gamma}{\langle \ell \rangle^\tau \alpha^{\mathtt{d}} \beta^{\mathtt{d}}} \Big\}.$$

Using that $|\cdot|^{\mathrm{lip}} \leqslant \gamma^{-1} |\cdot|^{\mathrm{Lip}(\gamma)}$, one gets

$$\begin{aligned} |\phi(\lambda_1) - \phi(\lambda_2)| &\geqslant \Big(|\bar{\omega} \cdot \ell| - \gamma^{-1} |m-1|^{\mathrm{Lip}(\gamma)} |\alpha - \beta| - \gamma^{-1} |\mathtt{r}_j^{(\alpha)}|^{\mathrm{Lip}(\gamma)} - \gamma^{-1} |\mathtt{r}_k^{(\beta)}|^{\mathrm{Lip}(\gamma)} \Big) |\lambda_1 - \lambda_2| \\ &\overset{(9.11)}{\geqslant} \Big(|\overline{\omega} \cdot \ell| - C\varepsilon \gamma^{-1} |\alpha - \beta| - C\varepsilon\gamma^{-1} \alpha^{-1} - C\varepsilon\gamma^{-1}\beta^{-1} \Big) |\lambda_1 - \lambda_2|. \end{aligned} \tag{9.23}$$

Since by lemma A.1-(ii) in [42], $|\alpha - \beta| \geqslant C(\alpha^{-1} + \beta^{-1})$, for some $C > 0$, by (9.18), $\varepsilon\gamma^{-1} = \varepsilon^{1-\mathfrak{a}}$, by aplying lemma 9.4-(i), one gets that for ε small enough

$$|\phi(\lambda_1) - \phi(\lambda_2)| \geqslant C(\alpha^{-1} + \beta^{-1}) |\lambda_1 - \lambda_2| \overset{\alpha, \beta \geqslant 1}{\geqslant} C\alpha^{-1}\beta^{-1} |\lambda_1 - \lambda_2|. \tag{9.24}$$

The above estimate implies that

$$\Big| R_{kj}(\ell, \alpha, \beta) \Big| \lesssim \frac{\gamma}{\alpha^{\mathtt{d}-1} \beta^{\mathtt{d}-1} \langle \ell \rangle^\tau}.$$

Finally recalling (9.15) and (9.10), we get the claimed estimate for the measure of $R(\ell, \alpha, \beta)$. ■

Proof of theorem 9.3 concluded. By (9.17), by applying lemmata 9.4 and 9.5, and recalling the definitions of the constants γ, τ and $\mathtt{d}$ made in (9.18) and (9.14), one gets the estimate

$$|\mathcal{C}_\varepsilon \setminus \mathcal{O}_\varepsilon| \lesssim \sum_{\ell \in \mathbb{Z}^\nu, j, j' \in \mathbb{Z}^d} \frac{\varepsilon^{\mathfrak{a}}}{\langle j \rangle^{\mathtt{d}-1-d} \langle j' \rangle^{\mathtt{d}-1-d} \langle \ell \rangle^\tau} + \sum_{\substack{\ell \in \mathbb{Z}^\nu, j, j' \in \mathbb{Z}^d \\ |j|, |j'| \lesssim \langle \ell \rangle}} \frac{\varepsilon^{\mathfrak{a}} \langle j \rangle^{d-1} \langle j' \rangle^{d-1}}{\langle \ell \rangle^\tau} \lesssim \varepsilon^{\mathfrak{a}} \tag{9.25}$$

since the above two series are convergen because of the choices of τ and $\mathtt{d}$ made in (9.14). The proof of theorem 9.3 is then concluded. ■

Note that our proof is very similar to the analogous one in [42]; the only difference is that in lemma 9.4 we have to bound $\alpha^{-1} + \beta^{-1} \lesssim |\bar{\omega} \cdot \ell|$ rather than having the estimate $|\alpha - \beta| \lesssim \langle \ell \rangle$ of lemma 5.1 in [42]. However this is still enough to get the final measure estimate due to the Diophantine condition on $\bar{\omega}$.

Acknowledgments

We warmly thank M Procesi for carefully reading the manuscript, and for her comments and suggestions. We also thank an anonymous referee for his/her comments. LC was supported by NSF grant DMS-1500943. RM was supported by Swiss National Science Foundation, grant *Hamiltonian systems of infinite dimension*, project number: 200020–165537.

References

[1] Alazard T and Baldi P 2015 Gravity capillary standing water waves *Arch. Ration. Mech. Anal.* **217** 741–830
[2] Baldi P 2009 Periodic solutions of forced Kirchhoff equations *Ann. Scuola Normale Super. Pisa Cl. Sci.* **8** 117–41
[3] Baldi P 2013 Periodic solutions of fully nonlinear autonomous equations of Benjamin–Ono type *Ann. Inst. Henri Poincare* C **30** 33–77
[4] Baldi P, Berti M and Montalto R 2014 KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation *Math. Ann.* **359** 471–536
[5] Baldi P, Berti M and Montalto R 2016 KAM for autonomous quasi-linear perturbations of KdV *Ann. Inst. Henri Poincare* C **33** 1589–638
[6] Baldi P, Berti M and Montalto R 2016 KAM for autonomous quasi-linear perturbations of mKdV *Boll. Unione Mat. Ital.* **9** 143–88
[7] Baldi P, Berti M, Haus E and Montalto R 2017 Time quasi-periodic gravity water waves in finite depth (arXiv:1602.02411)
[8] Bambusi D, Grebert B, Maspero A and Robert D 2017 Reducibility of the quantum Harmonic oscillator in d-dimensions with polynomial time dependent perturbation (arXiv:1702.05274)
[9] Berti M, Biasco L and Procesi M 2013 KAM theory for the Hamiltonian DNLW *Ann. Sci. Éc. Norm. Supér.* **46** 301–73
[10] Berti M, Biasco L and Procesi M 2014 KAM theory for the reversible derivative wave equation *Arch. Ration. Mech. Anal.* **212** 905–55
[11] Berti M and Bolle P 2012 Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential *Nonlinearity* **25** 2579–613
[12] Berti M and Bolle P 2013 Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential *J. Eur. Math. Soc.* **15** 229–86
[13] Berti M, Bolle P and Procesi M 2010 An abstract Nash–Moser theorem with parameters and applications to PDEs *Ann. Inst. Henri Poincaré* **1** 377–99
[14] Berti M, Corsi L and Procesi M 2015 An abstract Nash–Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous spaces *Commun. Math. Phys.* **334** 1413–54
[15] Berti M, Kappeler T and Montalto R 2016 Large KAM tori for perturbations of the dNLS equation *Asterisque*. in preparation (arXiv:1603.09252v1)
[16] Berti M and Montalto R 2017 Quasi-periodic standing wave solutions of gravity-capillary water waves *Memoirs of the American Mathematical Society* submitted
[17] Bourgain J 1994 Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE *Int. Math. Res. Not.* (in preparation)
[18] Bourgain J 1998 Quasi-periodic solutions of Hamiltonian perturbations of 2*D* linear Schrödinger equations *Ann. Math.* **148** 363–439
[19] Bourgain J 1999 Periodic solutions of nonlinear wave equations *Harmonic analysis and partial differential equations* (*Chicago Lectures in Mathematics*) (Chicago, IL: University of Chicago Press) pp 69–97

[20] Bourgain J 2005 *Green's Function Estimates for Lattice Schrödinger Operators and Applications* (*Annals of Mathematics Studies* vol 158) (Princeton, NJ: Princeton University Press)
[21] Craig W and Wayne E C 1993 Newton's method and periodic solutions of nonlinear wave equation *Commun. Pure Appl. Math.* **46** 1409–98
[22] Chierchia L and You J 2000 KAM tori for 1D nonlinear wave equations with periodic boundary conditions *Commun. Math. Phys.* **211** 497–525
[23] Corsi L, Haus E and Procesi M 2015 A KAM result on compact Lie groups *Acta App. Math.* **137** 41–59
[24] Eliasson L H and Kuksin S 2010 KAM for nonlinear Schrödinger equation *Ann. Math.* **172** 371–435
[25] Eliasson L H, Grebert B and Kuksin S 2016 KAM for the nonlinear beam equation *Geom. Funct. Anal.* **26** 1588–715
[26] Feola R 2016 KAM for quasi-linear forced hamiltonian NLS (arXiv:1602.01341)
[27] Feola R and Procesi M 2015 Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations *J. Differ. Equ.* **259** 3389–447
[28] Feola R, Giuliani F, Montalto R and Procesi M 2018 Reducibility of first order linear operators on tori via Moser's theorem (arXiv:1801.04224)
[29] Gallavotti G 1986 Quasi integrable mechanical systems *Phénomenes Critiques, Systèmes Aleatoires, Théorie de Jange*, ed K Ostervalder *et al* session XLIII part II (Amsterdam: Elsevier) pp 539–623
[30] Giuliani F 2017 Quasi-periodic solutions for quasi-linear generalized KdV equations *J. Differ. Equ.* **262** 5052–132
[31] Iooss G, Plotnikov P I and Toland J F 2005 Standing waves on an infinitely deep perfect fluid under gravity *Arch. Ration. Mech. Anal.* **177** 367–478
[32] Iooss G and Plotnikov P I 2009 *Small Divisor Problem in the Theory of Three-dimensional Water Gravity Waves* vol 940 (Providence, RI: American Mathematical Society) p 200
[33] Iooss G and Plotnikov P I 2011 Asymmetrical three-dimensional travelling gravity waves *Arch. Ration. Mech. Anal.* **200** 789–880
[34] Kappeler T and Pöschel J 2003 *KAM and KdV* (Berlin: Springer)
[35] Klainermann S and Majda A 1980 Formation of singularities for wave equations including the nonlinear vibrating string *Commun. Pure Appl. Math.* **33** 241–63
[36] Kuksin S 1998 A KAM theorem for equations of the Korteweg-de Vries type *Rev. Math. Math. Phys.* **10** 1–64
[37] Lax P 1964 Development of singularities of solutions of nonlinear hyperbolic partial differential equations *J. Math. Phys.* **5** 611–3
[38] Liu J and Yuan X 2010 Spectrum for quantum Duffing oscillator and small-divisor equation with large-variable coefficient *Commun. Pure Appl. Math.* **63** 1145–72
[39] Liu J and Yuan X 2011 A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations *Commun. Math. Phys.* **307** 629–73
[40] Lojasiewicz S and Zehnder E 1979 An inverse function theorem in Frechet spaces *J. Funct. Anal.* **33** 165–74
[41] Montalto R 2017 Quasi-periodic solutions of forced Kirchhoff equation *Nonlinear Differ. Equ. Appl. NoDEA* **24** 9
[42] Montalto R 2017 A reducibility result for a class of linear wave equations on $\mathbb{T}^d$ *Int. Math. Res. Not.* (https://doi.org/10.1093/imrn/rnx167)
[43] Moser J 1966 Rapidly convergent iteration method and nonlinear partial doifferential equations I *Ann. Scuola Normale Super. Pisa* **20** 265–315
[44] Rabinowitz P H 1967 Periodic solutions of nonlinear hyperbolic partial differential equations part I and II *Commun. Pure Appl. Math.* **20** 145–205
Rabinowitz P H 1969 Periodic solutions of nonlinear hyperbolic partial differential equations part I and II *Commun. Pure Appl. Math.* **22** 15–39