

# An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds

Massimiliano Berti<sup>1</sup>, Livia Corsi<sup>1</sup>, Michela Procesi<sup>2</sup>

<sup>1</sup> Dipartimento di Matematica, Università di Napoli “Federico II”, Napoli, I-80126, Italy

<sup>2</sup> Dipartimento di Matematica, Università di Roma “La Sapienza”, Roma, I-00185, Italy

E-mail: m.berti@unina.it, livia.corsi@unina.it, mprocesi@mat.uniroma1.it

## Abstract

We prove an abstract Implicit Function Theorem with parameters for smooth operators defined on sequence scales, modeled for the search of quasi-periodic solutions of PDEs. The tame estimates required for the inverse linearised operators at each step of the iterative scheme are deduced via a multiscale inductive argument. The Cantor like set of parameters where the solution exists is defined in a non inductive way. This formulation completely decouples the iterative scheme from the measure theoretical analysis of the parameters where the small divisors non-resonance conditions are verified. As an application, we deduce the existence of quasi-periodic solutions for forced NLW and NLS equations on *any* compact Lie group or manifold which is homogeneous with respect to a compact Lie group, extending previous results valid only for tori. A basic tool of harmonic analysis is the highest weight theory for the irreducible representations of compact Lie groups.

**Keywords:** Quasi-periodic solutions for PDEs; Nash-Moser theory; small divisor problems; Non-linear Schrödinger and wave equations; analysis on compact Lie groups

**MSC classification:** 37K55; 58C15; 35Q55; 35L05

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>An implicit function theorem with parameters on sequence spaces</b>	<b>7</b>
2.1	Linear operators on $H^s$ and matrices . . . . .	8
2.2	Main abstract results . . . . .	10
<b>3</b>	<b>Applications to PDEs</b>	<b>14</b>
3.1	Analysis on Lie groups . . . . .	14
3.2	Proof of Theorem 1.1 for NLW . . . . .	16
3.3	Proof of Theorem 1.1 for NLS . . . . .	22
<b>4</b>	<b>An abstract Nash-Moser theorem</b>	<b>24</b>

4.1	Proof of Theorem 4.2 . . . . .	26
4.2	Initialisation of the Nash-Moser scheme . . . . .	26
4.3	Iterative step . . . . .	27
<b>5</b>	<b>Proof of Theorem 2.16</b>	<b>31</b>
5.1	Initialisation . . . . .	33
5.2	Inductive step . . . . .	34
5.3	Separation properties . . . . .	37
5.4	Regularity . . . . .	39
<b>A</b>	<b>Proof of the multiscale Proposition 5.8</b>	<b>40</b>

# 1 Introduction

In the last years several works have been devoted to the search of quasi-periodic solutions of Hamiltonian PDEs in higher space dimensions, like analytic nonlinear Schrödinger (NLS) and nonlinear wave (NLW) equations on  $\mathbb{T}^n$ . A major difficulty concerns the verification of the so-called Melnikov non-resonance conditions. The first successful approach, due to Bourgain [9], [10], used a Newton iterative scheme which requires only the minimal (first-order) Melnikov conditions, which are verified inductively at each step of the iteration.

In this paper we prove an abstract, differentiable Nash-Moser Implicit Function Theorem with parameters for smooth operators defined on Hilbert sequence scales. As applications, we prove the existence of quasi-periodic solutions with Sobolev regularity of the forced nonlinear wave equation

$$u_{tt} - \Delta u + mu = \varepsilon f(\omega t, x, u), \quad x \in \mathbb{M}, \tag{1.1}$$

and the nonlinear Schrödinger equation

$$iu_t - \Delta u + mu = \varepsilon f(\omega t, x, u), \quad x \in \mathbb{M}, \tag{1.2}$$

where  $\mathbb{M}$  is *any* compact Lie group or manifold which is homogeneous with respect to a compact Lie group, namely there exists a compact Lie group which acts on  $\mathbb{M}$  transitively and differentiably. In (1.1)-(1.2) we denote by  $\Delta$  the Laplace-Beltrami operator, the “mass”  $m > 0$ , the parameter  $\varepsilon > 0$  is small, and the frequency vector  $\omega \in \mathbb{R}^d$  is non-resonant, see (1.4)-(1.6) below.

Examples of compact connected Lie groups are the standard torus  $\mathbb{T}^n$ , the special orthogonal group  $SO(n)$ , the special unitary group  $SU(n)$ , and so on. Examples of (compact) manifolds homogeneous with respect to a compact Lie group are the spheres  $S^n$ , the real and complex Grassmannians, and the moving frames, namely, the manifold of the  $k$ -ples of orthonormal vectors in  $\mathbb{R}^n$  with the natural action of the orthogonal group  $O(n)$  and many others; see for instance [11].

The study of (1.1)-(1.2) on a manifold  $\mathbb{M}$  which is homogeneous with respect to a compact Lie group  $G$  is reduced to that of (1.1)-(1.2) on the Lie group  $G$  itself. Indeed  $\mathbb{M}$  is diffeomorphic to  $\mathbb{M} = G/N$  where  $N$  is a closed subgroup of  $G$  and the Laplace-Beltrami operator on  $\mathbb{M}$  can be identified with the Laplace-Beltrami operator on  $G$ , acting on functions invariant under  $N$  (see [7]-Theorem 2.7 and [17, 18, 23]).

Concerning regularity we assume that the nonlinearity  $f \in C^q(\mathbb{T}^d \times \mathbb{M} \times \mathbb{R}; \mathbb{R})$ , resp.  $\mathbf{f}(\varphi, x, u) \in C^q(\mathbb{T}^d \times \mathbb{M} \times \mathbb{C}; \mathbb{C})$  in the real sense (namely as a function of  $\operatorname{Re}(u), \operatorname{Im}(u)$ ), for some  $q$  large enough. We also require that

$$\mathbf{f}(\omega t, x, u) = \partial_{\bar{u}} H(\omega t, x, u), \quad H(\varphi, x, u) \in \mathbb{R}, \quad \forall u \in \mathbb{C}, \quad (1.3)$$

so that the NLS equation (1.2) is Hamiltonian.

We assume that the frequency  $\omega$  has a fixed direction, namely

$$\omega = \lambda \bar{\omega}, \quad \lambda \in \mathcal{I} := [1/2, 3/2], \quad |\bar{\omega}|_1 := \sum_{p=1}^d |\omega_p| \leq 1, \quad (1.4)$$

for some fixed diophantine vector  $\bar{\omega}$ , i.e.  $\bar{\omega}$  satisfies

$$|\bar{\omega} \cdot l| \geq 2\gamma_0 |l|^{-d}, \quad \forall l \in \mathbb{Z}^d \setminus \{0\}. \quad (1.5)$$

For the NLW equation (1.1) we assume also the quadratic diophantine condition

$$\left| \sum_{1 \leq i, j \leq d} \bar{\omega}_i \bar{\omega}_j p_{ij} \right| \geq \frac{\gamma_0}{|p|^{d(d+1)}}, \quad \forall p \in \mathbb{Z}^{d(d+1)/2} \setminus \{0\} \quad (1.6)$$

which is satisfied for all  $|\bar{\omega}|_1 \leq 1$  except a set of measure  $O(\gamma_0^{1/2})$ , see Lemma 6.1 in [5].

The search of quasi-periodic solutions of (1.1)-(1.2) reduces to finding solutions  $u(\varphi, x)$  of

$$(\omega \cdot \partial_\varphi)^2 u - \Delta u + mu = \varepsilon f(\varphi, x, u), \quad i\omega \cdot \partial_\varphi u - \Delta u + mu = \varepsilon \mathbf{f}(\varphi, x, u, \bar{u}), \quad (1.7)$$

in some Sobolev space  $H^s$  of both the variables  $(\varphi, x)$ , see Section 2.

**Theorem 1.1.** *Let  $\mathbb{M}$  be any compact Lie group or manifold which is homogeneous with respect to a compact Lie group. Consider the NLW equation (1.1), and assume (1.4)-(1.6); for the NLS equation (1.2)-(1.3) assume only (1.4)-(1.5). Then there are  $s, q \in \mathbb{R}$  such that, for any  $f, \mathbf{f} \in C^q$  and for all  $\varepsilon \in [0, \varepsilon_0]$  with  $\varepsilon_0 > 0$  small enough, there is a map*

$$u_\varepsilon \in C^1([1/2, 3/2], H^s), \quad \sup_{\lambda \in [1/2, 3/2]} \|u_\varepsilon(\lambda)\|_s \rightarrow 0, \quad \text{as } \varepsilon \rightarrow 0,$$

and a Cantor-like set  $\mathcal{C}_\varepsilon \subset [1/2, 3/2]$ , satisfying  $\operatorname{meas}(\mathcal{C}_\varepsilon) \rightarrow 1$  as  $\varepsilon \rightarrow 0$ , such that, for any  $\lambda \in \mathcal{C}_\varepsilon$ ,  $u_\varepsilon(\lambda)$  is a solution of (1.7), with  $\omega = \lambda \bar{\omega}$ . Moreover if  $f, \mathbf{f} \in C^\infty$  then the solution  $u_\varepsilon(\lambda)$  is of class  $C^\infty$  both in time and space.

Actually Theorem 1.1 is deduced by the abstract *Implicit Function Theorems 2.16, 2.18* (and *Corollary 2.17*) on scales of Hilbert sequence spaces. We postpone their precise formulations to Section 2.2, since some preparation is required.

Theorem 1.1 is a first step in the direction of tackling the very hard problem of finding quasi-periodic solutions for NLW and NLS on any compact Riemannian manifold, if ever true. This is an open problem also for periodic solutions. In the particular case that the manifold  $\mathbb{M} = \mathbb{T}^n$  is a  $n$ -dimensional torus, Theorem 1.1 is proved in [4] for NLS, and, in [5], for NLW.

So far, the literature about quasi-periodic solutions is restricted to NLS and NLW on tori (which are compact commutative Lie groups). The first results were proved for the interval  $[0, \pi]$  by Kuksin

[19, 20], Wayne [27], Pöschel [22, 21], or for the 1-dimensional circle  $\mathbb{T}$  by Craig-Wayne [13], Bourgain [8], and Chierchia You [12]. For higher dimensional tori  $\mathbb{T}^n$ ,  $n \geq 2$ , the first existence results have been obtained by Bourgain [9, 10] for NLS and NLW with Fourier multipliers via a multiscale analysis, recently applied by Wang [26] for completely resonant NLS. Using KAM techniques, Eliasson-Kuksin [14, 15] proved existence and stability of quasi-periodic solutions for NLS with Fourier multipliers, see also Procesi-Xu [25]. Then Geng-Xu-You [16] proved KAM results for the cubic NLS in dimension 2 and Procesi-Procesi [24] in any dimension  $n$  and polynomial nonlinearity.

The reason why these results are confined to tori is that these proofs require specific properties of the eigenvalues and the eigenfunctions must be the exponentials or, at least, strongly “localized close to exponentials”. Recently, Berti-Bolle [4, 5] have extended the multiscale analysis to deal with NLS and NLW on  $\mathbb{T}^d$  with a multiplicative potential. In such a case the eigenfunctions may not be localized close to the exponentials.

In the previous paper [7], Berti-Procesi proved existence of periodic solutions for NLW and NLS on any compact Lie group or manifold homogenous with respect to a compact Lie group. Main difficulties concern the eigenvalues of the Laplace-Beltrami operator, with their unbounded multiplicity, and the rule of multiplications of the eigenfunctions. A key property which is exploited is that, for a Lie group, the product of two eigenfunctions is a finite linear combinations of them (as for the exponentials or the spherical harmonics). From a dynamical point of view, it implies, roughly speaking, that only finitely many normal modes are strongly coupled.

Theorem 1.1 extends the result in [7] to the harder quasi-periodic setting. As already said, it is deduced by the abstract *Implicit Function Theorems 2.16-2.18*. These results rely on the Nash-Moser iterative Theorem 4.2 and a multiscale inductive scheme for deducing tame estimates for the inverse linearised operators at each step of the iteration, see Section 5. A main advantage of Theorem 2.16 is that the Cantor-like set of parameters  $\mathcal{C}_\varepsilon$  in (2.33) for which a solution exists is defined in terms of the “solution”  $u_\varepsilon$ , and it is *not* inductively defined as in previous approaches. This formulation completely decouples the Nash-Moser iteration from the discussion about the measure of the parameters where all the required “non-resonance” conditions are verified. The possibility to impose the non-resonance conditions through the “final solution” was yet observed in [3] (in a Lyapunov-Schmidt context) and in [2] for a KAM theorem. In the present case the Cantor set  $\mathcal{C}_\varepsilon$  is rather involved. Nevertheless we are able to provide efficient measure estimates in the applications. This simplifies considerably the presentation because the measure estimates are not required at each step. In conclusion, in order to apply Theorems 2.16-2.18, one does not need to know the multiscale techniques nor the Nash-Moser approach: they can be used as a *black box*.

We believe that Theorems 2.16-2.18 can be applied to several other cases. The abstract hypotheses can be verified by informations of the harmonic analysis on the manifold. In the case of compact Lie groups and homogeneous manifolds, Theorem 1.1 follows by using only the harmonic analysis in [7] (see Section 3), which stems from the informations on the eigenvalues and eigenspaces of the Laplace-Beltrami operator provided by the highest weight theory, see [23].

We find it convenient to use a Nash-Moser scheme because the eigenvalues of the Laplacian are highly degenerate and the second order Melnikov non resonance conditions required for the KAM reducibility scheme might not be satisfied.

*Informal presentation of the ideas and techniques.* Many nonlinear PDEs (such as (1.7)) can be seen as implicit function equations of the form

$$F(\varepsilon, \lambda, u) = 0,$$

having for  $\varepsilon = 0$  the trivial solution  $u(t, x) = 0$ , for each parameter  $\lambda \in \mathcal{I}$ . Clearly, due to the small divisors, the standard Implicit Function Theorem fails, and one must rely on some Nash-Moser or KAM quadratic scheme. They are rapidly convergent iterative algorithms based on the Newton method and hence need some informations about the invertibility of the linearisation  $L(\varepsilon, \lambda, u)$  of  $F$  at any function  $u$  close to zero.

Due to the Hamiltonian structure, the operator  $L(\varepsilon, \lambda, u)$  is self-adjoint and it is easy to obtain informations on its eigenvalues, implying the invertibility of  $L(\varepsilon, \lambda, u)$  with bounds of the  $L^2$ -norm of  $L^{-1}(\varepsilon, \lambda, u)$  for “most” parameters  $\lambda$ . However these informations are not enough to prove the convergence of the algorithm: one needs estimates on the high Sobolev norm of the inverse which do not follow only from bounds on the eigenvalues. Usually this property is implied by a sufficiently fast polynomial off-diagonal decay of the matrices which represent the inverse operators.

In the case of the interval  $[0, \pi]$ , the eigenvalues of  $L(\varepsilon, \lambda, u)$  are often distinct, a property which enables to diagonalise  $L(\varepsilon, \lambda, u)$  via a smooth change of variables (reducibility) implying very strong estimates of the inverse operator in high Sobolev norm. This method automatically implies also the stability of the solution. Unfortunately, the eigenvalues of  $\Delta$  are not simple already on  $\mathbb{T}$  (a fortiori neither on  $\mathbb{T}^n$ ,  $n \geq 2$ ), so that generalising these *reducibility* methods is complicated and strongly depends on the equation. For NLS it is obtained in [14].

However, the convergence of the Nash-Moser scheme only requires “tame” estimates of the inverse in high Sobolev norm (for instance like the one in (4.4)) which may be obtained under weaker spectral hypotheses. In the case of NLS and NLW on  $\mathbb{T}^n$  these estimates have been obtained in [4, 5] via a *multiscale analysis* on Sobolev spaces (see [10] in an analytic setting). Informally, the multiscale method is a way to prove an off-diagonal decay for the inverse of a finite-dimensional invertible matrix with off-diagonal decay, by using informations on the invertibility (in high norm) of a great number of principal minors of order  $N$  much smaller than the dimension of the matrix. The polynomial off-diagonal decay of a matrix implies that it defines a “tame” operator between Sobolev spaces.

In this paper we extend these techniques also to the case of compact Lie groups and manifolds which are homogeneous with respect to a compact Lie group (in the latter case we “lift-up” the equation to the Lie group). Two key points concern

1. the matrix representation of a multiplication operator  $u \mapsto bu$ ,
2. the properties of the eigenvalues of the Laplace-Beltrami operator.

The multiplication rules for the eigenfunctions, together with the numeration of the eigenspaces provided the highest weight theory, implies that the multiplication operator by a Sobolev function  $b \in H^s(\mathbb{M})$  is represented in the eigenfunction basis as a block matrix with off-diagonal decay, as stated precisely in Lemmas 3.1, 3.2 (proved in [7]). The block structure of this matrix takes into account the (large) multiplicity of the degenerate eigenvalues of  $\Delta$  on  $\mathbb{M}$  (several blocks could correspond to the same eigenvalue). This in principle could be a problem because one can not hope to achieve any off-diagonal decay property for the matrices restricted to such blocks. However, as in [7], we do not need such a decay, being sufficient to control only the  $L^2$ -operator norm on these blocks. Interestingly, properties of this type have been used by Bambusi, Delort, Grébert, Szeftel [1] for Birkhoff normal form results of Klein-Gordon equations on Zoll manifolds (a main difficulty in [1] is to verify the Birkhoff normal form non-resonance conditions).

Concerning item 2, the eigenvalues of the Laplace-Beltrami operator on a Lie group are very similar to those on a torus, as stated in (3.6). This enables to prove “separation properties” of clusters of

singular/bad sites (i.e. Fourier indices with a corresponding small divisor) à la Bourgain [9], [10]. Thanks to the off-diagonal decay property proved in item 1 such “resonant” clusters interact only weakly. As in the case of  $\mathbb{T}^n$  (where the eigenvalues of  $-\Delta$  are  $|j|^2$ ,  $j \in \mathbb{Z}^n$ ) one does not gather into the same cluster all the indexes corresponding to the same eigenvalue. The reason is that such clusters would not satisfy the needed separation properties.

We now give some more detail about the proof. In the usual PDE applications the function spaces decompose as a direct sum of eigenspaces of  $L(0, \lambda, 0)$ , each of them being a direct product of the exponentials (for the time-direction) and the eigenspaces of  $-\Delta + m$  (space direction). Hence we decompose  $u = \sum_k u_k$  with  $k = (l, j) \in \mathbb{Z}^d \times \Lambda_+$  ( $l \in \mathbb{Z}^d$  is the time-Fourier component and  $j \in \Lambda_+$  the space-Fourier component). In particular  $L(0, \lambda, 0)$  is a diagonal operator which is proportional to the identity on each eigenspace. Moreover the dependence on  $l$  appears only through a scalar function  $\mathfrak{D}_j(\omega \cdot l)$ , see (2.24a).

In Theorems 2.16-2.18 we revisit in a more abstract way the strategy of [4, 5], obtaining a unified and more general result for smooth operators  $F(\varepsilon, \lambda, u)$  acting on a Hilbert scale of sequences spaces. These results are based on three hypotheses which allow the possibility of passing from  $L^2$ -norm estimates to high Sobolev norm bounds for the inverse linearized operators. We try to explain the meaning of this assumptions:

- i. the linearised operator can be written as the sum of a diagonal part  $D = D(\lambda)$  (which is the linearised operator at  $\varepsilon = 0, u = 0$ ) plus a perturbation which has off-diagonal decay and is Töplitz in the time indices (see Hypothesis 1),
- ii. a uniform lower bound for the derivative of  $\mathfrak{D}_j(y)$  on the set where  $|\mathfrak{D}_j(y)|$  is small (see Hypothesis 2),
- iii. an assumption on the length of chains of “singular sites” (see Hypothesis 3).

Under these hypotheses Theorem 2.16 implies the existence, for  $\varepsilon$  sufficiently small, of a function  $u = u_\varepsilon(\lambda)$  which is a solution of the equation  $F(\varepsilon, \lambda, u) = 0$  for all  $\lambda$  in the Cantor-like set  $\mathcal{C}_\varepsilon$  in (2.33), which is defined only in terms of the eigenvalues of submatrices of  $L(\varepsilon, \lambda, u_\varepsilon(\lambda))$ . Roughly speaking the set  $\mathcal{C}_\varepsilon$  is defined as the intersection of two families of sets:

- 1. the sets  $\overline{\mathfrak{G}}_N$  of parameters  $\lambda$  for which the  $N$ -truncation of  $L(\varepsilon, \lambda, u_\varepsilon)$  is invertible in  $L^2$  with good bounds of the  $L^2$ -norm of the inverse (see (2.34)),
- 2. the sets  $\overline{\mathfrak{G}}_N^0$  of parameters  $\lambda$  for which the principal minors of order  $N$  having a small eigenvalue are *separated* (see (2.35)).

Technically the sets of type 2 are defined by exploiting the time-covariance property (2.29) and analysing the complexity of the real parameter  $\theta$  (see (2.25)) for which the  $N$ -truncation of the time-translated matrix  $L(\varepsilon, \lambda, \theta, u_\varepsilon)$  have a small eigenvalue: since  $\omega = \lambda\bar{\omega}$  is diophantine these two definitions are equivalent.

Finally we underline two main differences with respect to the abstract Nash-Moser theorem in [6]. The first is that the tame estimates (4.4) required for the inverse linearized operators are much weaker than in [6]. Note, in particular, that the tame exponent in (4.4) grows like  $\sim \delta s$  (this corresponds to an unbounded loss of derivatives as  $s$  increases). This improvement is necessary to deal with

quasi-periodic solutions. The second difference is that in [6] the measure issue was not yet completely decoupled from the Nash-Moser iteration, as, on the contrary, it is achieved in this paper thanks to the introduction of the set  $\mathcal{C}_\varepsilon$  in (2.33).

The paper is essentially self-contained. The Appendix A contains the proof of the multiscale proposition 5.8 which follows verbatim as in [4]. We have added it for the convenience of the reader.

*Acknowledgements.* We thank L. Biasco, P. Bolle, C. Procesi for many useful comments.

## 2 An implicit function theorem with parameters on sequence spaces

We work on a scale of Hilbert sequence spaces defined as follows. We start from an index set

$$\mathfrak{K} := \mathfrak{J} \times \mathfrak{A} = \mathbb{Z}^d \times \Lambda_+ \times \mathfrak{A} \quad (2.1)$$

where  $\Lambda_+ \subset \Lambda$  is contained in a  $r$ -dimensional lattice (in general not orthogonal)

$$\Lambda := \left\{ j \in \mathbb{R}^r : j = \sum_{p=1}^r j_p \mathbf{w}_p, j_p \in \mathbb{Z} \right\} \quad (2.2)$$

generated by independent vectors  $\mathbf{w}_1, \dots, \mathbf{w}_r \in \mathbb{R}^r$ . The set  $\mathfrak{A}$  is finite, and in the applications will be either  $\mathfrak{A} = \{1\}$  (for NLW) or  $\mathfrak{A} = \{1, -1\}$  (for NLS). Given  $k \in \mathfrak{K}$  we denote

$$k = (i, \mathbf{a}) = (l, j, \mathbf{a}) \in \mathbb{Z}^d \times \Lambda_+ \times \mathfrak{A}, \quad |k| = |i| := \max(|l|, |j|), \quad |j| := |j|_\infty = \max_p |j_p| \in \mathbb{N}. \quad (2.3)$$

If  $\mathfrak{A} = \{1\}$  we simply write  $k = (l, j)$ .

We require that  $\Lambda_+$  has a *product* structure, namely that

$$j = \sum_{p=1}^r j_p \mathbf{w}_p, j' = \sum_{p=1}^r j'_p \mathbf{w}_p \in \Lambda_+ \quad \Rightarrow \quad j'' = \sum_{p=1}^r j''_p \mathbf{w}_p \in \Lambda_+ \text{ if } \forall p \min(j_p, j'_p) \leq j''_p \leq \max(j_p, j'_p). \quad (2.4)$$

Condition (2.4) will be used only in order to prove Lemma 5.16. It could be probably weakened. In the applications it is satisfied.

To each  $j \in \Lambda_+$  we associate a ‘‘multiplicity’’  $d_j \in \mathbb{N}$ . Then, for  $s \geq 0$ , we define the (Sobolev) scale of Hilbert sequence spaces

$$H^s := H^s(\mathfrak{K}) := \left\{ u = \{u_k\}_{k \in \mathfrak{K}}, u_k \in \mathbb{C}^{d_j} : \|u\|_s^2 := \sum_{k \in \mathfrak{K}} \langle w_k \rangle^{2s} \|u_k\|_0^2 < \infty \right\} \quad (2.5)$$

where  $\|\cdot\|_0$  denotes the  $L^2$ -norm in  $\mathbb{C}^{d_j}$  and the weights  $\langle w_k \rangle := \max(c, 1, w_k)$  satisfy

$$c|k| \leq w_k \leq C|k|, \quad \forall k \in \mathfrak{K},$$

for suitable constants  $0 < c \leq C$ . In the applications the weights  $w_k$  are related to the eigenvalues of the Laplacian, see Examples 1, 2 below.

**Remark 2.1.** *The abstract Theorem 2.16 does not require any bound on the multiplicity  $d_j$ . In the applications we use the polynomial bound (3.4) for Lemmas 3.1, 3.2 and for the measure estimates.*

For any  $B \subseteq \mathfrak{K}$  we define the subspaces

$$H_B^s := \{u \in H^s : u_k = 0 \text{ for } k \notin B\}. \quad (2.6)$$

If  $B$  is a finite set the space  $H_B^s = H_B$  does not depend on  $s$  and it is included in  $\cap_{s \geq 0} H^s$ .

Finally, for  $k = (i, \mathfrak{a}), k' = (i', \mathfrak{a}') \in \mathfrak{K}$  we denote

$$\text{dist}(k, k') := \begin{cases} 1, & i = i', \mathfrak{a} \neq \mathfrak{a}', \\ |i - i'|, & \text{otherwise,} \end{cases} \quad (2.7)$$

where  $|i|$  is defined in (2.3).

**Remark 2.2.** *In principle  $i - i'$  may not be in  $\mathbb{Z}^d \times \Lambda_+$  because  $\Lambda_+$  is not a lattice. However, since  $\Lambda_+ \subset \Lambda$  we can always compute  $|i - i'|$  by considering  $i - i' \in \mathbb{Z}^d \times \Lambda$ . In order to avoid this problem we will extend our vectors by setting them to zero on  $(\mathbb{Z}^d \times \Lambda \times \mathfrak{A}) \setminus \mathfrak{K}$ .*

All the constants that will appear in the sequel may depend on the index set  $\mathfrak{K}$ , the weights  $w_k$  and on  $s$ . We will evidence only the dependence on  $s$ .

## 2.1 Linear operators on $H^s$ and matrices

Let  $B, C \subseteq \mathfrak{K}$ . A bounded linear operator  $L : H_B^s \rightarrow H_C^s$  is represented, as usual, by a matrix in

$$\mathcal{M}_C^B := \left\{ (M_k^{k'})_{k \in C, k' \in B}, M_k^{k'} \in \text{Mat}(d_j \times d_{j'}, \mathbb{C}) \right\}. \quad (2.8)$$

It is useful to evidence a bigger block structure. We decompose

$$B = \bar{B} \times \mathfrak{B}, \quad \bar{B} := \text{Proj}_{\mathbb{Z}^d \times \Lambda_+} B, \quad \mathfrak{B} := \text{Proj}_{\mathfrak{A}} B$$

and  $C = \bar{C} \times \mathfrak{C}$ , defined in the same way. Now, for  $i = (l, j) \in \bar{C}$ ,  $i' = (l', j') \in \bar{B}$ , we consider the matrix

$$M_{\{i\}}^{\{i'\}} := \{M_{i, \mathfrak{a}}^{i', \mathfrak{a}'}\}_{\mathfrak{a} \in \mathfrak{C}, \mathfrak{a}' \in \mathfrak{B}}, \quad M_{\{i\}}^{\{i'\}} \in \text{Mat}(|\mathfrak{C}|d_j \times |\mathfrak{B}|d_{j'}, \mathbb{C}),$$

where  $|\mathfrak{B}|, |\mathfrak{C}|$  denote the cardinality of  $\mathfrak{B}, \mathfrak{C} \subseteq \mathfrak{A}$  respectively. In the same way, given a vector  $v := \{v_k\}_{k \in \bar{C} \times \mathfrak{C}}$ , for  $i = (l, j) \in \bar{C}$ , we set  $v_{\{i\}} := \{v_{i, \mathfrak{a}}\}_{\mathfrak{a} \in \mathfrak{C}}$ .

**Remark 2.3.** *The difference with respect to [4], [5] is that the dimension of the matrix blocks  $M_{\{i\}}^{\{i'\}}$  may not be uniformly bounded. They are scalars for the NLW equation in [5] and, in [4], for NLS, at most  $1 \times 2, 2 \times 1$  or  $2 \times 2$  matrices, because  $d_j = d_{j'} = 1$  and  $1 \leq |\mathfrak{B}|, |\mathfrak{C}| \leq 2$ .*

We endow  $\text{Mat}(|\mathfrak{C}|d_j \times |\mathfrak{B}|d_{j'}, \mathbb{C})$  with the  $L^2$ -operator norm, which we denote  $\|\cdot\|_0$ . Note that whenever a multiplication is possible one has the algebra property.

**Definition 2.4.** (*s*-decay norm) *For any  $M \in \mathcal{M}_C^B$  we define its  $s$ -norm*

$$|M|_s^2 := K_1 \sum_{i \in \mathbb{Z}^d \times \Lambda} [M(i)]^2 \langle i \rangle^{2s} \quad (2.9)$$



where  $\langle i \rangle := \max(1, |i|)$ ,

$$[M(i)] := \begin{cases} \sup_{h-h'=i, h \in \overline{C}, h' \in \overline{B}} \|M_{\{h\}}^{\{h'\}}\|_0, & i \in \overline{C} - \overline{B}, \\ 0, & i \notin \overline{C} - \overline{B}, \end{cases} \quad (2.10)$$

and  $K_1 > 4 \sum_{i \in \mathbb{Z}^d \times \Lambda} \langle i \rangle^{-2s_0}$ .

We denote by  $(\mathcal{M}^s)_C^B \subset \mathcal{M}_C^B$  the set of matrices with finite  $s$ -norm  $|\cdot|_s$ . If  $B, C$  are finite sets then  $(\mathcal{M}^s)_C^B = \mathcal{M}_C^B$  does not depend on  $s$ , and, for simplicity, we drop the apex  $s$ .

Note that the norm  $|\cdot|_s \leq |\cdot|_{s'}$  for  $s \leq s'$ .

The norm defined in (2.9) is a variation of that introduced in Definition 3.2 of [4]. The only difference concerns the dimensions of the blocks  $M_{\{i\}}^{\{i'\}}$  as noted in Remark 2.3. However, since the matrices  $M_{\{i\}}^{\{i'\}}$  are measured with the operator norm  $\|\cdot\|_0$  the algebra and interpolation properties of the norm  $|\cdot|_s$  follow similarly to [4], as well as all the properties in section 3-[4]. Indeed, given  $M \in \mathcal{M}_C^B$  we introduce the Töpliz matrix

$$\mathcal{M} := (\mathcal{M}_{\{i\}}^{\{i'\}}) \in \mathcal{M}_C^B, \quad \mathcal{M}_{\{i\}}^{\{i'\}} := [M(i - i')] \mathbb{1}_{|\mathfrak{C}|_{d_j} \times |\mathfrak{B}|_{d_{j'}}} \quad (2.11)$$

which has the same decay norm

$$|M|_s = |\mathcal{M}|_s. \quad (2.12)$$

**Lemma 2.5.** *Let  $M_1 \in \mathcal{M}_D^C$  and  $M_2 \in \mathcal{M}_C^B$ . Then  $M_1 M_2 \in \mathcal{M}_D^B$  satisfies  $|M_1 M_2|_s \leq |\mathcal{M}_1 \mathcal{M}_2|_s$ .*

*Proof.* For  $i' \in \overline{B}$ ,  $i \in \overline{D}$ , we have

$$\begin{aligned} \left\| (M_1 M_2)_{\{i\}}^{\{i'\}} \right\|_0 &\leq \sum_{q \in \overline{C}} \left\| (M_1)_{\{i\}}^{\{q\}} \right\|_0 \left\| (M_2)_{\{q\}}^{\{i'\}} \right\|_0 \stackrel{(2.10)}{\leq} \sum_{q \in \overline{C}} [(M_1)(i - q)] [(M_2)(q - i')] \\ &= \sum_{q \in \overline{C}} [(M_1)(i - q)] [(M_2)(q - i')] \left\| \mathbb{1}_{|\mathfrak{D}|_{d_j} \times |\mathfrak{B}|_{d_{j'}}} \right\|_0 \\ &= \left\| \sum_{q \in \overline{C}} [M_1(i - q)] \mathbb{1}_{|\mathfrak{D}|_{d_j} \times |\mathfrak{C}|_{d_q}} [M_2(q - i')] \mathbb{1}_{|\mathfrak{C}|_{d_q} \times |\mathfrak{B}|_{d_{j'}}} \right\|_0 \stackrel{(2.11)}{=} \left\| (\mathcal{M}_1 \mathcal{M}_2)_{\{i\}}^{\{i'\}} \right\|_0. \end{aligned}$$

Therefore  $[(M_1 M_2)(i - i')] \leq \|(\mathcal{M}_1 \mathcal{M}_2)_{\{i\}}^{\{i'\}}\|_0$  and the lemma follows.  $\blacksquare$

In what follows we fix  $s \geq s_0 > (d + r)/2$ .

**Lemma 2.6. (Interpolation)** *For all  $s \geq s_0$  there is  $C(s) > 1$  with  $C(s_0) = 1$  such that, for any subset  $B, C, D \subseteq \mathfrak{K}$  and for all  $M_1 \in \mathcal{M}_D^C$ ,  $M_2 \in \mathcal{M}_C^B$ , one has*

$$|M_1 M_2|_s \leq \frac{1}{2} |M_1|_{s_0} |M_2|_s + \frac{C(s)}{2} |M_1|_s |M_2|_{s_0}. \quad (2.13)$$

*In particular, one has the algebra property  $|M_1 M_2|_s \leq C(s) |M_1|_s |M_2|_s$ .*

*Proof.* For the Töpliz matrices  $\mathcal{M}_1, \mathcal{M}_2$  the interpolation inequality (2.13) follows as usual (with  $C(s_0) \leq 1$  possibly taking  $K_1$  larger). Hence Lemma 2.5 and (2.12) imply (2.13).  $\blacksquare$

The  $s$ -norm of a matrix also controls the  $\|\cdot\|_s$  norm (see [4]-Lemma 3.5).

**Lemma 2.7.** For any  $B, C \subseteq \mathfrak{K}$ , let  $M \in \mathcal{M}_C^B$ . Then

$$\|Mh\|_s \leq C(s)|M|_{s_0}\|h\|_s + C(s)|M|_s\|h\|_{s_0}, \quad \forall h \in H_B^s. \quad (2.14)$$

*Proof.* Regarding a vector  $h = \{h_k\}_{k \in \overline{B} \times \mathfrak{B}}$  as a column matrix, its  $s$ -decay norm is  $|h|_s^2 = K_1 \sum_{i \in \overline{B}} \langle i \rangle^{2s} \|h_{\{i\}}\|_0^2$ . Hence (2.14) follows by Lemma 2.6 because  $c(s)\|h\|_s \leq |h|_s \leq c'(s)\|h\|_s$ . ■

We conclude this section stating further properties of the  $s$ -norm: such lemmata are proved word by word as Lemmas 3.6, 3.7 3.8 and 3.9 of [4] respectively.

**Lemma 2.8. (Smoothing)** Let  $M \in \mathcal{M}_C^B$  and  $N \geq 2$ . For all  $s' \geq s \geq 0$  the following hold.

(i) If  $M_k^{k'} = 0$  for all  $\text{dist}(k', k) < N$  (recall the definition (2.7)), then

$$|M|_s \leq N^{-(s'-s)}|M|_{s'}. \quad (2.15)$$

(ii) If  $M_k^{k'} = 0$  for all  $\text{dist}(k', k) > N$ , then

$$|M|_{s'} \leq N^{s'-s}|M|_s, \quad |M|_s \leq N^{s+d+r}\|M\|_0. \quad (2.16)$$

**Lemma 2.9. (Decay along lines)** Let  $M \in \mathcal{M}_C^B$  and denote by  $M_k$ ,  $k \in C$ , its  $k$ -th line. Then

$$|M|_s \leq |\mathfrak{C}|K_2 \max_{k \in C} |M_k|_{s+d+r}, \quad \forall s \geq 0. \quad (2.17)$$

**Lemma 2.10.** Let  $M \in \mathcal{M}_C^B$ . Then  $\|M\|_0 \leq |M|_{s_0}$ .

**Definition 2.11.** We say that a matrix  $M \in \mathcal{M}_C^B$  is left invertible if there exists  $N \in \mathcal{M}_B^C$  such that  $NM = \mathbf{1}_B$ . In such a case  $N$  is called a left inverse of  $M$ .

A matrix  $M$  is left-invertible if and only if it is injective. The left inverse is, in general, not unique. In what follows we shall denote by  $^{[-1]}M$  any left inverse of  $M$  when this does not causes ambiguity.

**Lemma 2.12. (Perturbation of left-invertible matrices)** Let  $M \in \mathcal{M}_C^B$  be a left invertible matrix. Then for any  $P \in \mathcal{M}_C^B$  such that  $^{[-1]}M|_{s_0}|P|_{s_0} \leq 1/2$  there exists a left inverse of  $M + P$  such that

$$^{[-1]}(M + P)|_{s_0} \leq 2|^{[-1]}M|_{s_0}, \quad |^{[-1]}(M + P)|_s \leq C(s)(|^{[-1]}M|_s + |^{[-1]}M|_{s_0}^2|P|_s), \quad (2.18)$$

for any  $s \geq s_0$ . Moreover, if  $\|^{[-1]}M\|_0\|P\|_0 \leq 1/2$ , then there is a left inverse of  $M + P$  which satisfies

$$\|^{[-1]}(M + P)\|_0 \leq 2\|^{[-1]}M\|_0. \quad (2.19)$$

## 2.2 Main abstract results

We consider a non-linear operator

$$F(\varepsilon, \lambda, u) = D(\lambda)u + \varepsilon f(u) \quad (2.20)$$

where  $\varepsilon > 0$  is small, the parameter  $\lambda \in \mathcal{I} \subset [1/2, +\infty)$ , and  $D(\lambda)$  is a diagonal linear operator  $D(\lambda) : H^{s+\nu} \rightarrow H^s$  such that

$$\|D(\lambda)h\|_s, \|\partial_\lambda D(\lambda)h\|_s \leq C(s)\|h\|_{s+\nu} \quad (2.21)$$

(in the applications  $D(\lambda) = i\lambda\bar{\omega} \cdot \partial_\varphi - \Delta + m$  or  $(\lambda\bar{\omega} \cdot \partial_\varphi)^2 - \Delta + m$ ) whose action on the subspace associated to a fixed index  $k$  is scalar, namely

$$D(\lambda) = \text{diag}(D_k(\lambda)\mathbb{1}_{d_j})_{k \in \mathfrak{K}}. \quad (2.22)$$

We assume that, for some  $s_0 > (d+r)/2$ , the nonlinearity  $f \in C^2(B_1^{s_0}, H^{s_0})$  (where  $B_1^{s_0}$  denotes the unit ball in  $H^{s_0}$ ) and the following ‘‘tame’’ properties hold: given  $S' > s_0$ , for all  $s \in [s_0, S']$  there exists a constant  $C(s)$  such that for any  $\|u\|_{s_0} \leq 2$ ,

$$(f1) \quad \|df(u)[h]\|_s \leq C(s)(\|u\|_s \|h\|_{s_0} + \|h\|_s),$$

$$(f2) \quad \|d^2 f(u)[h, v]\|_s \leq C(s) \left( \|u\|_s \|h\|_{s_0} \|v\|_{s_0} + \|h\|_s \|v\|_{s_0} + \|h\|_{s_0} \|v\|_s \right)$$

hold. Our goal is to find  $u = u_\varepsilon(\lambda) \in H^s$  for suitable  $s$  which solves the equation  $F(\varepsilon, \lambda, u_\varepsilon(\lambda)) = 0$  at least for ‘‘some’’ values of  $\lambda \in \mathcal{I}$ .

Then we assume further properties on the linearised operator

$$L = L(\varepsilon, \lambda, u) = D(\lambda) + \varepsilon T(u), \quad D(\lambda) = \text{diag}(D_k(\lambda)\mathbb{1}_{d_j})_{k \in \mathfrak{K}}. \quad (2.23)$$

where  $T(u)$  is the matrix which represents the bounded linear operator  $df(u)$ , see (2.8).

**Hypothesis 1.** *Let  $\bar{\omega} \in \mathbb{R}^d$  satisfy (1.5). There exists a function  $\mathfrak{D} : \Lambda_+ \times \mathfrak{A} \times \mathbb{R} \rightarrow \mathbb{C}$  and  $\nu_0 > 0$  such that*

$$\text{(Covariance)} \quad D_{(l,j,\mathfrak{a})}(\lambda) = \mathfrak{D}_{j,\mathfrak{a}}(\lambda\bar{\omega} \cdot l), \forall \lambda \in \mathcal{I} \quad (2.24a)$$

$$\text{(Töplitz in time)} \quad T \in \mathcal{M}_{\mathfrak{K}}^{\mathfrak{K}} : T_{(l,j,\mathfrak{a})}^{(l',j',\mathfrak{a}')} = T_{(j,\mathfrak{a})}^{(j',\mathfrak{a}')} (l - l') \quad (2.24b)$$

$$\text{(Off-diagonal decay)} \quad |T(u)|_{s-\nu_0} \leq C(s)(1 + \|u\|_s), \quad (2.24c)$$

$$\text{(Lipschitz)} \quad |T(u) - T(u')|_{s-\nu_0} \leq C(s)(\|u - u'\|_s + (\|u\|_s + \|u'\|_s)\|u - u'\|_{s_0}), \quad (2.24d)$$

for all  $\|u\|_{s_0}, \|u'\|_{s_0} \leq 2$  and  $s_0 + \nu_0 < s < S'$ .

For any  $\theta \in \mathbb{R}$  we set

$$D(\lambda, \theta) = \text{Diag}(D_k(\lambda, \theta)\mathbb{1}_{d_j}), \quad D_k(\lambda, \theta) := \mathfrak{D}_{j,\mathfrak{a}}(\lambda\bar{\omega} \cdot l + \theta) \quad (2.25a)$$

$$L(\varepsilon, \lambda, \theta, u) := D(\lambda, \theta) + \varepsilon T(u). \quad (2.25b)$$

We need the following information about the unperturbed small divisors.

**Hypothesis 2. (Initialisation)** *There are  $n$  such that for all  $\tau_1 > 1$ ,  $N > 1$ ,  $\lambda \in \mathcal{I}$ ,  $l \in \mathbb{Z}^d$ ,  $j \in \Lambda_+$ ,  $\mathfrak{a} \in \mathfrak{A}$ , the set*

$$\{\theta \in \mathbb{R} : |D_{(l,j,\mathfrak{a})}(\lambda, \theta)| \leq N^{-\tau_1}\} \subseteq \bigcup_{q=1}^n I_q \quad \text{intervals with } \text{meas}(I_q) \leq N^{-\tau_1}. \quad (2.26)$$

We now distinguish which unperturbed small divisors are actually small or not.

**Definition 2.13. (Regular/singular sites)** *We say that the index  $k \in \mathfrak{K}$  is regular for a matrix  $D := \text{diag}(D_k\mathbb{1}_{d_j})$ ,  $D_k \in \mathbb{C}$ , if  $|D_k| \geq 1$ , otherwise we say that  $k$  is singular.*

We need an assumption which provides separation properties of clusters of singular sites.

For any  $\Sigma \subset \mathfrak{K}$  and  $\tilde{j} \in \Lambda_+$  we denote the section of  $\Sigma$  at fixed  $\tilde{j}$  by

$$\Sigma^{(\tilde{j})} := \{k = (l, \tilde{j}, \mathbf{a}) \in \Sigma\}.$$

**Definition 2.14.** *Let  $\theta, \lambda$  be fixed and  $K > 1$ . We denote by  $\Sigma_K$  any subset of singular sites of  $D(\lambda, \theta)$  in  $\mathfrak{K}$  such that, for all  $\tilde{j} \in \Lambda_+$ , the cardinality of the section  $\Sigma_K^{(\tilde{j})}$  satisfies  $\#\Sigma_K^{(\tilde{j})} \leq K$ .*

**Definition 2.15. ( $\Gamma$ -Chain)** *Let  $\Gamma \geq 2$ . A sequence  $k_0, \dots, k_\ell \in \mathfrak{K}$  with  $k_p \neq k_q$  for  $0 \leq p \neq q \leq \ell$  such that*

$$\text{dist}(k_{q+1}, k_q) \leq \Gamma, \quad \text{for all } q = 0, \dots, \ell - 1, \quad (2.27)$$

*is called a  $\Gamma$ -chain of length  $\ell$ .*

**Hypothesis 3. (Separation of singular sites)** *There exists a constant  $\mathfrak{s}$  and, for any  $N_0 \geq 2$ , a set  $\tilde{\mathcal{I}} = \tilde{\mathcal{I}}(N_0)$  such that, for all  $\lambda \in \tilde{\mathcal{I}}$ ,  $\theta \in \mathbb{R}$ , and for all  $K, \Gamma$  with  $K\Gamma \geq N_0$ , any  $\Gamma$ -chain of singular sites in  $\Sigma_K$  as in Definition 2.14, has length  $\ell \leq (\Gamma K)^\mathfrak{s}$ .*

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

$$L_{N,i}(\theta) := L(\theta)_F^F, \quad F := \{k' \in \mathfrak{K} : \text{dist}(k, k') \leq N\}. \quad (2.28)$$

If  $l = 0$ , instead of the notation (2.28) 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  $L_{N,j} := L_{N,j}(0)$ .

By hypothesis 1, the matrix  $L = L(\varepsilon, \lambda, \theta, u)$  has the following covariance property in time

$$L_{N,l,j}(\varepsilon, \lambda, \theta, u) = L_{N,j}(\varepsilon, \lambda, \theta + \lambda \bar{\omega} \cdot l, u). \quad (2.29)$$

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

$$\bar{\mathcal{I}} := \bar{\mathcal{I}}(N_0, \tau_0) := \left\{ \lambda \in \mathcal{I} : |D_k(\lambda)| \geq N_0^{-\tau_0} \text{ for all } k = (i, \mathbf{a}) \in \mathfrak{K} : |i| \leq N_0 \right\}. \quad (2.30)$$

**Theorem 2.16.** *Let  $\mathfrak{e} > d + r + 1$ . Assume that  $F$  in (2.20) satisfies (2.21)-(2.22), (f1)-(f2) and Hypotheses 1, 2, 3 with  $S'$  large enough, depending on  $\mathfrak{e}$ . Then, there are  $\tau_1 > 1$ ,  $\bar{N}_0 \in \mathbb{N}$ ,  $s_1$ ,  $S \in (s_0 + \nu_0, S' - \nu_0)$  with  $s_1 < S$  (all depending on  $\mathfrak{e}$ ) and  $c(S) > 0$  such that for all  $N_0 \geq \bar{N}_0$ , if the smallness condition*

$$\varepsilon N_0^S < c(S) \quad (2.31)$$

*holds, then there exists a function  $u_\varepsilon \in C^1(\mathcal{I}, H^{s_1 + \nu})$  with  $u_0(\lambda) = 0$ , which solves*

$$F(\varepsilon, \lambda, u_\varepsilon(\lambda)) = 0 \quad (2.32)$$

for all  $\lambda \in \mathcal{C}_\varepsilon \subset \mathcal{I}$  defined in (2.33) below. The set  $\mathcal{C}_\varepsilon$  is defined in terms of the “solution”  $u_\varepsilon(\lambda)$ , as

$$\mathcal{C}_\varepsilon := \bigcap_{n \geq 0} \bar{\mathcal{G}}_{N_0^{2^n}}^0 \cap \bar{\mathfrak{G}}_{N_0^{2^n}} \cap \tilde{\mathcal{I}} \cap \bar{\mathcal{I}} \quad (2.33)$$

where  $\tilde{\mathcal{I}} = \tilde{\mathcal{I}}(N_0)$  is defined in Hypothesis 3,  $\bar{\mathcal{I}}$  in (2.30), and, for all  $N \in \mathbb{N}$ ,

$$\bar{\mathfrak{G}}_N := \left\{ \lambda \in \mathcal{I} : \|L_N^{-1}(\varepsilon, \lambda, u_\varepsilon(\lambda))\|_0 \leq N^{\tau_1}/2 \right\}, \quad (2.34)$$

$$\begin{aligned} \bar{\mathcal{G}}_N^0 := & \left\{ \lambda \in \mathcal{I} : \forall j_0 \in \Lambda_+ \text{ there is a covering} \right. \\ & \left. \bar{B}_N^0(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^\varepsilon} I_q, \text{ with } I_q = I_q(j_0) \text{ intervals with } \text{meas}(I_q) \leq N^{-\tau_1} \right\} \end{aligned} \quad (2.35)$$

with

$$\bar{B}_N^0(j_0, \varepsilon, \lambda) := \left\{ \theta \in \mathbb{R} : \|L_{N, j_0}^{-1}(\varepsilon, \lambda, \theta, u_\varepsilon(\lambda))\|_0 > N^{\tau_1}/2 \right\}. \quad (2.36)$$

Finally, if the tame estimates (f1)-(f2), (2.24c), (2.24d) hold up to  $S' = +\infty$  then  $u_\varepsilon(\lambda) \in \cap_{s \geq 0} H^s$ .

In applications, it is often useful to work in appropriate closed subspaces  $\widehat{H}^s(\mathfrak{R}) \subset H^s(\mathfrak{R})$  which are invariant under the action of  $F$ . The following corollary holds:

**Corollary 2.17.** *Assume, in addition to the hypotheses of Theorem 2.16, that  $F(\varepsilon, \lambda, \cdot) : \widehat{H}^{s+\nu}(\mathfrak{R}) \rightarrow \widehat{H}^s(\mathfrak{R})$ ,  $\forall s > s_0$ . Then the function  $u_\varepsilon$  provided by Theorem 2.16 belongs to  $C^1(\mathcal{I}, \widehat{H}^{s_1+\nu}(\mathfrak{R}))$ .*

In Theorem 2.16 the Cantor like  $\mathcal{C}_\varepsilon$  defined in (2.33) may be empty. In order to prove that it has asymptotically full measure we need more informations. We fix  $N_0 = \lceil \varepsilon^{-1/(S+1)} \rceil$  so that the smallness condition (2.31) is satisfied for  $\varepsilon$  small enough.

**Theorem 2.18.** *Let  $N_0 = \lceil \varepsilon^{-1/(S+1)} \rceil$  with  $\varepsilon$  small enough so that (2.31) holds. Assume, in addition to the hypotheses of Theorem 2.16, that for all  $N \geq N_0$ ,*

$$\text{meas}(\mathcal{I} \setminus \bar{\mathcal{G}}_N^0), \text{meas}(\mathcal{I} \setminus \bar{\mathfrak{G}}_N) = O(N^{-1}), \quad \text{meas}(\mathcal{I} \setminus (\bar{\mathcal{I}} \cap \tilde{\mathcal{I}})) = O(N_0^{-1}). \quad (2.37)$$

Then  $\mathcal{C}_\varepsilon$  satisfies, for some  $K > 0$ ,

$$\text{meas}(\mathcal{I} \setminus \mathcal{C}_\varepsilon) \leq K\varepsilon^{1/(S+1)}. \quad (2.38)$$

*Proof.* Let us denote  $N_n = N_0^{2^n}$ . By the explicit expression (2.33) we have

$$\begin{aligned} \text{meas}(\mathcal{I} \setminus \mathcal{C}_\varepsilon) &= \text{meas} \left( \bigcup_{n \geq 0} (\bar{\mathcal{G}}_{N_n}^0)^c \bigcup_{n \geq 0} (\bar{\mathfrak{G}}_{N_n})^c \cup \tilde{\mathcal{I}}^c \cup \bar{\mathcal{I}}^c \right) \\ &\leq \sum_{n \geq 0} \text{meas}(\mathcal{I} \setminus \bar{\mathcal{G}}_{N_n}^0) + \sum_{n \geq 0} \text{meas}(\mathcal{I} \setminus \bar{\mathfrak{G}}_{N_n}) + \text{meas}(\mathcal{I} \setminus (\bar{\mathcal{I}} \cap \tilde{\mathcal{I}})) \\ &\stackrel{(2.37)}{\leq} C_0 \sum_{n \geq 0} N_n^{-1} + C_1 N_0^{-1} \leq C' N_0^{-1} \leq K\varepsilon^{1/(S+1)} \end{aligned} \quad (2.39)$$

which proves (2.38). ■

In the applications to NLW and NLS the conditions (2.37) will be verified taking  $\tau_0, \tau_1$  large, with a suitable  $\varepsilon$ , see Proposition 3.6.

### 3 Applications to PDEs

Now we apply Theorems 2.16-2.18 to the NLW and NLS equations (1.1)-(1.2). To be precise, when  $\mathbf{M}$  is a manifold which is homogeneous with respect to a compact Lie group, we rely on Corollary 2.17.

We briefly recall the relevant properties of harmonic analysis on compact Lie groups that we need, referring to [23] (and [7]) for precise statements and proofs.

A compact manifold  $\mathbf{M}$  which is homogeneous with respect to a compact Lie group is, up to an isomorphism, diffeomorphic to

$$\mathbf{M} = G/N, \quad G := \mathbf{G} \times \mathbb{T}^{r_2}, \quad (3.1)$$

where  $\mathbf{G}$  is a simply connected compact Lie group,  $\mathbb{T}^{r_2}$  is a torus, and  $N$  is a closed subgroup of  $G$ . Then, a function on  $\mathbf{M}$  can be seen as a function defined on  $G$  which is invariant under the action of  $N$ , and the space  $H^s(\mathbf{M}, \mathbb{C})$  (or  $H^s(\mathbf{M}, \mathbb{R})$ ) can be identified with the subspace

$$\widehat{H}^s := \widehat{H}^s(G, \mathbb{C}) := \left\{ u \in H^s(G) : u(x) = u(xg), \quad \forall x \in G = \mathbf{G} \times \mathbb{T}^{r_2}, g \in N \right\}. \quad (3.2)$$

Moreover, the Laplace-Beltrami operator on  $\mathbf{M}$  can be identified with the Laplace-Beltrami operator on the Lie group  $G$ , acting on functions invariant under  $N$  (see Theorem 2.7, [7]). Then we “lift” the equations (1.1)-(1.2) on  $G$  and we use harmonic analysis on Lie groups.

#### 3.1 Analysis on Lie groups

Any simply connected compact Lie group  $\mathbf{G}$  is the product of a finite number of simply connected Lie groups of simple type (which are classified and come in a finite number of families).

Let  $\mathbf{G}$  be of simple type, with dimension  $\mathfrak{d}$  and rank  $r$ . Denote by  $\mathfrak{w}_1, \dots, \mathfrak{w}_r \in \mathbb{R}^r$  the fundamental weights of  $\mathbf{G}$  and consider the cone of dominant weights

$$\Lambda_+(\mathbf{G}) := \left\{ j = \sum_{p=1}^r j_p \mathfrak{w}_p : j_p \in \mathbb{N} \right\} \subset \Lambda := \left\{ j = \sum_{p=1}^r j_p \mathfrak{w}_p : j_p \in \mathbb{Z} \right\}.$$

Note that  $\Lambda_+(\mathbf{G})$  satisfies (2.4) and indexes the finite dimensional irreducible representations of  $\mathbf{G}$ .

The eigenvalues and the eigenfunctions of the Laplace-Beltrami operator  $\Delta$  on  $\mathbf{G}$  are

$$\mu_j := -|j + \rho|_2^2 + |\rho|_2^2, \quad \mathbf{f}_{j,\sigma}(x), \quad x \in \mathbf{G}, \quad j \in \Lambda_+(\mathbf{G}), \quad \sigma = 1, \dots, d_j, \quad (3.3)$$

where  $\rho := \sum_{i=1}^r \mathfrak{w}_i$ ,  $|\cdot|_2$  denotes the euclidean norm on  $\mathbb{R}^r$ , and  $\mathbf{f}_j(x)$  is the (unitary) matrix associated to an irreducible unitary representation  $(R_{V_j}, V_j)$  of  $\mathbf{G}$ , precisely

$$(\mathbf{f}_j(x))_{h,k} = \langle R_{V_j}(x)v_h, v_k \rangle, \quad v_h, v_k \in V_j,$$

where  $(v_h)_{h=1, \dots, \dim V_j}$  is an orthonormal basis of the finite dimensional euclidean space  $V_j$  with scalar product  $\langle \cdot, \cdot \rangle$ . We denote by  $\mathcal{N}_j$  the corresponding eigenspace of  $\Delta$ . The *degeneracy of the eigenvalue*  $\mu_j$  satisfies

$$d_j \leq |j + \rho|_2^{\mathfrak{d}-r}. \quad (3.4)$$

The Peter-Weyl theorem implies the orthogonal decomposition

$$L^2(\mathbf{G}) = \bigoplus_{j \in \Lambda_+(\mathbf{G})} \mathcal{N}_j.$$

Many informations on the eigenvalues  $\mu_j$  are known. There exists an integer  $\mathfrak{Z} \in \mathbb{N}$  such that (see [7]-Lemma 2.6) the fundamental weights satisfy

$$w_i \cdot w_p \in \mathfrak{Z}^{-1}\mathbb{Z}, \quad \forall i, p = 1, \dots, r, \quad (3.5)$$

so that, in particular,

$$\mu_j := -|j + \rho|_2^2 + |\rho|_2^2, \quad |j|_2^2, \quad \rho \cdot j, \quad |\rho|_2^2 \in \mathfrak{Z}^{-1}\mathbb{Z}, \quad \forall j \in \Lambda_+(\mathfrak{G}). \quad (3.6)$$

For a product group,  $L^2(\mathfrak{G}_1 \times \mathfrak{G}_2) = L^2(\mathfrak{G}_1) \otimes L^2(\mathfrak{G}_2)$  and all the irreducible representations are obtained by the tensor product of the irreducible representations of  $\mathfrak{G}_1$  and  $\mathfrak{G}_2$ . Hence we extend all the above properties to any compact Lie group  $\mathfrak{G}$ . For simplicity we still denote the dimension of the group as  $\mathfrak{d}$  and the rank as  $r$ . In particular  $\Lambda_+(G) = \Lambda_+(\mathfrak{G}) \times \mathbb{Z}^{r_2}$  (see (3.1)) is the index set for the irreducible representations of  $G$ , with indices  $j = (j_1, j_2)$ ,  $j_1 \in \Lambda_+(\mathfrak{G})$ ,  $j_2 \in \mathbb{Z}^{r_2}$ , and  $\rho \rightsquigarrow (\rho, 0)$ .

We denote the indices  $i = (l, j) \in \mathbb{Z}^d \times \Lambda_+(G)$ , so that  $L^2(\mathbb{T}^d \times \mathfrak{G} \times \mathbb{T}^{r_2})$  naturally decomposes as product of subspaces  $\mathcal{N}_k$  of the form

$$\mathcal{N}_k := \langle e^{i\varphi \cdot l} \rangle \otimes \mathcal{N}_j = \langle e^{i\varphi \cdot l} \rangle \otimes \mathcal{N}_{j_1} \otimes \langle e^{ix_2 \cdot j_2} \rangle.$$

We also set

$$|i| := \max(|l|, |j|), \quad |l| := |l|_\infty, \quad |j| := |j|_\infty = \max_i |j_i|, \quad \langle i \rangle := \max(1, |i|).$$

The Sobolev spaces  $H^s(\mathbb{T}^d \times G)$  and  $H^s(\mathbb{T}^d \times G) \times H^s(\mathbb{T}^d \times G)$ , for a Lie group  $G$ , can be now identified with sequence spaces introduced in Section 2.

**Example 1.** Let  $\mathfrak{A} := \{1\}$ ,  $\Lambda_+ := \Lambda_+(G)$  be the cone of fundamental weights and, for  $k = (l, j) \in \mathfrak{K} = \mathbb{Z}^d \times \Lambda_+$ , let  $w_k := \sqrt{|l|_2^2 + |j + \rho|_2^2}$ . Then we may identify  $H^s(\mathfrak{K})$  with the Sobolev space  $H^s(\mathbb{T}^d \times G)$ .

**Example 2.** Let  $\mathfrak{A} := \{1, -1\}$ ,  $\Lambda_+ := \Lambda_+(G)$  be the cone of fundamental weights and, for  $k = (l, j, \mathfrak{a}) \in \mathfrak{K} := \mathbb{Z}^d \times \Lambda_+ \times \mathfrak{A}$ , let  $w_k := \sqrt{|l|_2^2 + |j + \rho|_2^2}$ . Then we may identify  $H^s(\mathfrak{K})$  with the Sobolev space  $H^s(\mathbb{T}^d \times G) \times H^s(\mathbb{T}^d \times G)$ .

The final fundamental property that we exploit concerns the off-diagonal decay of the block matrix which represents the multiplication operator, see (2.8). The block structure of this matrix takes into account the (large) multiplicity of the degenerate eigenvalues of  $\Delta$ . We remark that several blocks could correspond to the same eigenvalue (as in the case of the torus). The next lemmas, proved in [7], are ultimately connected to the fact that the product of two eigenfunctions of the Laplace operator on a Lie group is a finite sum of eigenfunctions, see [7]-Theorem 2.10.

The forthcoming Lemmas are a reformulation of Proposition 2.19 and Lemma 7.1 in [7] respectively, and they require the polynomial bound (3.4).

**Lemma 3.1.** ([7]-Proposition 2.19) Let  $\mathfrak{K}$  be as in Example 1 and  $b \in H^s(\mathbb{T}^d \times G)$  be real valued. Then the multiplication operator  $B : u(\varphi, x) \mapsto b(\varphi, x)u(\varphi, x)$  is self-adjoint in  $L^2$  and, for any  $s > (d + \mathfrak{d})/2$ ,

$$\|B_k^{k'}\|_0 \leq \frac{C(s)\|b\|_s}{\langle k - k' \rangle^{s - (d + \mathfrak{d})/2}}, \quad \forall k, k' \in \mathbb{Z}^d \times \Lambda_+,$$

where  $B_k^{k'} \in \text{Mat}(d_j \times d_{j'}, \mathbb{C})$ , see (2.8).

**Lemma 3.2.** ([7]-Lemma 7.1) Let  $\mathfrak{K}$  be as in Example 2. Consider  $a, b, c \in H^s(\mathbb{T}^d \times G)$  with  $a, b$  real valued. Then the multiplication operator with matrix

$$B = \begin{pmatrix} a(\varphi, x) & c(\varphi, x) \\ \bar{c}(\varphi, x) & b(\varphi, x) \end{pmatrix}$$

is self-adjoint in  $L^2$  and, for any  $s > (d + \mathfrak{d})/2$ ,

$$\|B_{\{i\}}^{\{i'\}}\|_0 \leq C(s) \frac{\max(\|a\|_s, \|b\|_s, \|c\|_s)}{\langle i - i' \rangle^{s - (d + \mathfrak{d})/2}}, \quad \forall i, i' \in \mathbb{Z}^d \times \Lambda_+.$$

**Corollary 3.3.** Let  $B$  be a linear operator as in the previous two Lemmas. Then, for all  $s > (d + \mathfrak{d})/2$ ,

$$|B|_s \leq C(s) \max(\|a\|_{s + \nu_0}, \|b\|_{s + \nu_0}, \|c\|_{s + \nu_0}), \quad \nu_0 := (2d + \mathfrak{d} + r + 1)/2.$$

### 3.2 Proof of Theorem 1.1 for NLW

We apply Theorems 2.16-2.18 to the operator

$$\begin{aligned} F(\varepsilon, \lambda, \cdot) : H^{s+2}(\mathbb{T}^d \times \mathbb{M}, \mathbb{R}) &\longrightarrow H^s(\mathbb{T}^d \times \mathbb{M}, \mathbb{R}) \\ u &\longmapsto (\lambda \bar{\omega} \cdot \partial_\varphi)^2 u - \Delta u + m u - \varepsilon f(\varphi, x, u) \end{aligned}$$

which can be extended to  $H^{s+2}(\mathbb{T}^d \times \mathbf{G} \times \mathbb{T}^{r_2}, \mathbb{R}) \rightarrow H^s(\mathbb{T}^d \times \mathbf{G} \times \mathbb{T}^{r_2}, \mathbb{R})$  such that for all  $u \in H^{s+2}(\mathbb{T}^d) \otimes \widehat{H}^{s+2}$  one has  $F(\varepsilon, \lambda, u) \in H^s(\mathbb{T}^d) \otimes \widehat{H}^s$  where  $\widehat{H}^s$  is defined in (3.2).

Setting  $\mathfrak{A} := \{1\}$ ,  $\Lambda_+ := \Lambda_+(G)$ ,  $G := \mathbf{G} \times \mathbb{T}^{r_2}$ , we are in the functional setting of Example 1. The Hypothesis (2.21)-(2.22) holds with  $\nu = 2$  and the interpolation estimates (f1)-(f2) are verified provided that  $f(\varphi, x, u) \in C^q$  for  $q$  large enough and  $s_0 > (d + \mathfrak{d})/2 \geq (d + r)/2$ .

**Remark 3.4.** We require  $s_0 > (d + \mathfrak{d})/2$  in view of the embedding  $H^{s_0}(\mathbb{T}^d \times G) \hookrightarrow L^\infty(\mathbb{T}^d \times G)$  which, in turn, implies the algebra and interpolation properties of the spaces  $H^s(\mathbb{T}^d \times G)$ ,  $s \geq s_0$ . The weaker bound  $s_0 > (d + r)/2$  is sufficient in order to prove the algebra and interpolation properties of the decay norm  $|\cdot|_s$  (see Section 2.1), which hold with no constraint on the multiplicity  $d_j$ .

The linearised operator

$$D(\lambda) - \varepsilon g(\varphi, x), \quad D(\lambda) := (\lambda \bar{\omega} \cdot \partial_\varphi)^2 - \Delta + m, \quad g(\varphi, x) := (\partial_u f)(\varphi, x, u(\varphi, x)),$$

is represented, in the Fourier basis  $e^{il \cdot \varphi} \mathbf{f}_j(x)$ , as in (2.23) with

$$D_k(\lambda) := D_{(l,j)}(\lambda) = -(\lambda \bar{\omega} \cdot l)^2 + m - \mu_j \tag{3.7}$$

and  $T(u)$  is the matrix associated to the multiplication operator by  $-g(\varphi, x)$ . Corollary 3.3 implies that  $T(u) \in \mathcal{M}^{s - \nu_0}$  for all  $u \in H^s(\mathbb{T}^d \times G)$  and the estimates (2.24c), (2.24d) hold by interpolation. Hypothesis 1 holds with  $\mathfrak{D}_j(y) = -y^2 + m - \mu_j$ .

Also Hypothesis 2 holds: a direct computation shows that

$$\{\theta \in \mathbb{R} : |D_{(l,j)}(\lambda, \theta)| \leq N^{-\tau_1}\} \subseteq \bigcup_{q=1,2} I_q, \quad I_q \text{ intervals with } \text{meas}(I_q) \leq \frac{4N^{-\tau_1}}{\sqrt{m - \mu_j}} + O(N^{-2\tau_1}),$$



and Hypothesis 2 holds with  $\mathbf{n} = 16/\sqrt{m}$ .

Hypothesis 3 about the length of chains of singular sites follows as in [5] because the eigenvalues of the Laplace-Beltrami operator are very similar to those on a torus, see (3.6). For  $\gamma > 0$  let

$$\begin{aligned} \tilde{\mathcal{I}} := \tilde{\mathcal{I}}(\gamma) := & \left\{ \lambda \in [1/2, 3/2] : |P(\lambda\bar{\omega})| \geq \frac{\gamma}{1 + |\rho|^{d(d+1)}}, \forall \text{ non zero polynomial} \right. \\ & \left. P(X) \in \mathbb{Z}[X_1, \dots, X_d] \text{ of the form } P(X) = p_0 + \sum_{1 \leq i_1, i_2 \leq d} p_{i_1, i_2} X_{i_1} X_{i_2} \right\}. \end{aligned} \quad (3.8)$$

**Lemma 3.5.** *For all  $N_0 \geq 2$ , Hypothesis 3 is satisfied with  $\tilde{\mathcal{I}}$  defined in (3.8) and  $\gamma = N_0^{-1}$ .*

*Proof.* The proof follows Lemma 4.2 of [5]. First of all, it is sufficient to bound the length of a  $\Gamma$ -chain of singular sites for  $D(\lambda, 0)$ . Then we consider the quadratic form

$$Q : \mathbb{R} \times \mathbb{R}^r \rightarrow \mathbb{R}, \quad Q(x, j) := -x^2 + |j|_2^2, \quad (3.9)$$

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

$$\Phi_1((x, j), (x', j')) := xx', \quad \Phi_2((x, j), (x', j')) := j \cdot j'. \quad (3.10)$$

For a  $\Gamma$ -chain of sites  $\{k_q = (l_q, j_q)\}_{q=0, \dots, \ell}$  which are singular for  $D(\lambda, 0)$  (Definition 2.13) we have, recalling (3.7), (3.6), and setting  $x_q := \omega \cdot l_q$ ,

$$|Q(x_q, j_q + \rho)| < 2 + |m - |\rho|_2^2|, \quad \forall q = 0, \dots, \ell.$$

Moreover, by (3.9), (2.27), we derive  $|Q(x_q - x_{q_0}, j_q - j_{q_0})| \leq C|q - q_0|^2 \Gamma^2$ ,  $\forall 0 \leq q, q_0 \leq \ell$ , and so

$$|\Phi((x_{q_0}, j_{q_0} + \rho), (x_q - x_{q_0}, j_q - j_{q_0}))| \leq C'|q - q_0|^2 \Gamma^2. \quad (3.11)$$

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

$$\mathcal{S} := \text{Span}_{\mathbb{R}}\{(x_q - x_{q_0}, j_q - j_{q_0}) : q = 0, \dots, \ell\}$$

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

**Case 1.** *For all  $q_0 = 0, \dots, \ell$  one has*

$$\text{Span}_{\mathbb{R}}\{(x_q - x_{q_0}, j_q - j_{q_0}) : |q - q_0| \leq \ell^\delta, q = 0, \dots, \ell\} = \mathcal{S}. \quad (3.12)$$

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

$$|f_b| \leq C\Gamma\ell^\delta, \quad b = 1, \dots, \mathfrak{s}. \quad (3.13)$$

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 (3.10), we write

$$\Omega = \left( -\Phi_1(f_{b'}, f_b) + \Phi_2(f_{b'}, f_b) \right)_{b, b'=1}^{\mathfrak{s}} = -X + Y, \quad (3.14)$$

where  $X_b^{b'} := (\omega \cdot \Delta l_{q_{b'}})(\omega \cdot \Delta l_{q_b})$  and  $Y_b^{b'} := (\Delta j_{q_{b'}}) \cdot (\Delta j_{q_b})$ . By (3.5) the matrix  $Y$  has entries in  $\mathfrak{Z}^{-1}\mathbb{Z}$  and the matrix  $X$  has rank 1 since each column is

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

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

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

which is a quadratic polynomial as in (3.8) with coefficients  $\leq C(\Gamma\ell^\delta)^{2(r+1)}$ . Note that  $P \not\equiv 0$ . Indeed, if  $P \equiv 0$  then

$$0 = P(i\omega) = \mathfrak{Z}^{r+1} \det(X + Y) = \mathfrak{Z}^{r+1} \det(f_b \cdot f_{b'})_{b,b'=1,\dots,\mathfrak{s}} > 0$$

because  $\{f_b\}_{b=1}^{\mathfrak{s}}$  is a basis of  $\mathcal{S}$ . This contradiction proves that  $P \neq 0$ . But then, by (3.8),

$$\mathfrak{Z}^{r+1} |\det(\Omega)| = |P(\omega)| \geq \frac{\gamma}{1 + |p|^{d(d+1)}} \geq \frac{\gamma}{(\Gamma\ell^\delta)^{C(d,r)}},$$

the matrix  $\Omega$  is invertible and

$$|(\Omega^{-1})_b^{b'}| \leq C\gamma^{-1}(\Gamma\ell^\delta)^{C'(d,r)}. \quad (3.15)$$

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

$$\Pi_{\mathcal{S}}(x_{q_0}, j_{q_0} + \rho) = \sum_{b'=1}^{r+1} a_{b'} f_{b'}, \quad (3.16)$$

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

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

where  $\Omega$  is defined in (3.14). The definition of  $f_b$ , the bound (3.11) and (3.12) imply  $|w| \leq C(\Gamma\ell^\delta)^2$ . Hence, by (3.15), we deduce  $|a| = |\Omega^{-1}w| \leq C'\gamma^{-1}(\Gamma\ell^\delta)^{C(r,d)+2}$ , whence, by (3.16) and (3.13),

$$|\Pi_{\mathcal{S}}(x_{q_0}, j_{q_0} + \rho)| \leq \gamma^{-1}(\Gamma\ell^\delta)^{C'(r,d)}.$$

Therefore, for any  $q_1, q_2 = 0, \dots, \ell$ , one has

$$|(x_{q_1}, j_{q_1}) - (x_{q_2}, j_{q_2})| = |\Pi_{\mathcal{S}}(x_{q_1}, j_{q_1} + \rho) - \Pi_{\mathcal{S}}(x_{q_2}, j_{q_2} + \rho)| \leq \gamma^{-1}(\Gamma\ell^\delta)^{C_1(r,d)},$$

which in turn implies  $|j_{q_1} - j_{q_2}| \leq \gamma^{-1}(\Gamma\ell^\delta)^{C_1(r,d)}$  for all  $q_1, q_2 = 0, \dots, \ell$ . Since all the  $j_q$  have  $r$  components (being elements of  $\Lambda_+(G)$ ) they are at most  $C\gamma^{-r}(\Gamma\ell^\delta)^{C_1(r,d)r}$ . We are considering a  $\Gamma$ -chain in  $\Sigma_K$  (see Definition 2.14) and so, for each  $q_0$ , the number of  $q \in \{0, \dots, \ell\}$  such that  $j_q = j_{q_0}$  is at most  $K$  and hence

$$\ell \leq \gamma^{-r}(\Gamma\ell^\delta)^{C_2(r,d)} K \leq (\Gamma K)^r (\Gamma\ell^\delta)^{C_2(r,d)} K \leq \ell^{\delta C_2(r,d)} (\Gamma K)^{r+C_2(r,d)}$$

because of the condition  $\Gamma K \geq N_0$  (Hypothesis 3) and  $N_0 = \gamma^{-1}$ . Choosing  $\delta < 1/(2C_2(r, d))$  we get  $\ell \leq (\Gamma K)^{2(r+C_2(r, d))}$ .

**Case 2.** There is  $q_0 = 0, \dots, \ell$  such that

$$\dim(\text{Span}_{\mathbb{R}}\{(x_q - x_{q_0}, j_q - j_{q_0}) : |q - q_0| \leq \ell^\delta, q = 0, \dots, \ell\}) \leq \mathfrak{s} - 1.$$

Then we repeat the argument of Case 1 for the sub-chain  $\{(l_q, j_q) : |q - q_0| \leq \ell^\delta\}$  and obtain a bound for  $\ell^\delta$ . Since this procedure should be applied at most  $r + 1$  times, at the end we get a bound like  $\ell \leq (\Gamma K)^{C_3(r, d)}$ .  $\blacksquare$

We have verified the hypotheses of Theorem 2.16 and Corollary 2.17. The next proposition proves that also the assumptions (2.37) in Theorem 2.18 hold.

**Proposition 3.6.** *Fix  $\tau_0 > r + 3d + 1$ ,  $\tau_1 > d + \mathfrak{d} + 2$  and  $\mathfrak{e} = d + \mathfrak{d} + r + 4$ . There exists  $N_0 \in \mathbb{N}$  (possibly larger than the  $N_0$  found in Theorem 2.16) such that (2.37) holds.*

The proof of Proposition 3.6 -which will continue until the end of this section- follows by basic properties of the eigenvalues of a self-adjoint operator, which are a consequence of their variational characterisation. Proposition 3.6 is indeed a reformulation of Proposition 5.1 of [5]. With respect to [5], the eigenvalues of the Laplacian in (3.6) are different and the index set  $\Lambda$  is not an orthonormal lattice.

**Remark 3.7.** *There are two positive constants  $\mathfrak{c} < \mathfrak{C}$  such that  $\mathfrak{c}|j| \leq |j|_2 \leq \mathfrak{C}|j|$ . Hence if  $|j| > \alpha \mathfrak{c}^{-1}N$ ,  $N > 2|\rho|_2$ , then the eigenvalues  $\mu_j$  in (3.6) satisfy  $-\mu_j > \alpha(\alpha - 1)N^2$ . On the other hand if  $|j| \leq \alpha \mathfrak{c}^{-1}N$  then  $-\mu_j \leq \alpha(\alpha + 1)(\mathfrak{C}/\mathfrak{c})^2N^2$ .*

Recall that if  $A, A'$  are self-adjoint matrices, then their eigenvalues  $\mu_p(A)$ ,  $\mu_p(A')$  (ranked in nondecreasing order) satisfy

$$|\mu_p(A) - \mu_p(A')| \leq \|A - A'\|_0. \quad (3.17)$$

We study finite dimensional restrictions of the the self-adjoint operator  $L(\varepsilon, \lambda) = L(\varepsilon, \lambda, u_\varepsilon(\lambda)) = D(\lambda) + \varepsilon T(\varepsilon, \lambda)$ .

One proceeds differently for  $|j_0| \geq (c + 5)c^{-1}N$  and  $|j_0| < (c + 5)c^{-1}N$ . We assume  $N \geq N_0 > 0$  large enough and  $\varepsilon \|T\|_0 \leq 1$ .

**Lemma 3.8.** *For all  $j_0 \in \Lambda_+(G)$ ,  $|j_0| \geq (c + 5)c^{-1}N$ , and for all  $\lambda \in [1/2, 3/2]$  one has*

$$B_N^0(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^{d+\mathfrak{d}+2}} I_q, \quad \text{with } I_q = I_q(j_0) \text{ intervals with } \text{meas}(I_q) \leq N^{-\tau_1}.$$

*Proof.* We first show that  $B_N^0(j_0, \varepsilon, \lambda) \subset \mathbb{R} \setminus [-2N, 2N]$ . Indeed by (3.17) all the eigenvalues  $\mu_{l,j,\sigma}(\theta)$ ,  $\sigma = 1, \dots, d_j$ , of  $L_{N,j_0}(\varepsilon, \lambda, \theta)$  (recall that  $d_j$  denotes the degeneracy of the eigenvalues  $\mu_j$  in (3.3)), are of the form

$$\mu_{l,j,\sigma}(\theta) = \delta_{l,j}(\theta) + O(\varepsilon \|T\|_0), \quad \delta_{l,j}(\theta) := -(\omega \cdot l + \theta)^2 + m - \mu_j. \quad (3.18)$$

Since  $|\omega|_1 = \lambda |\bar{\omega}|_1 \leq 3/2$ ,  $|j - j_0| \leq N$ ,  $|l| \leq N$ , one has, by Remark 3.7,

$$\delta_{l,j}(\theta) \geq -\left(\frac{3}{2}N + |\theta|\right)^2 + 20N^2 > 7N^2, \quad \forall |\theta| < 2N.$$

By (3.18) we deduce  $\mu_{l,j,\sigma}(\theta) \geq 6N^2$  and this implies  $B_N^0(j_0, \varepsilon, \lambda) \cap [-2N, 2N] = \emptyset$ . Now set  $B_N^{0,+} := B_N^0(j_0, \varepsilon, \lambda) \cap (2N, +\infty)$ ,  $B_N^{0,-} := B_N^0(j_0, \varepsilon, \lambda) \cap (-\infty, -2N)$ . Since

$$\partial_\theta L_{N,j_0}(\varepsilon, \lambda, \theta) = \text{diag}_{|l| \leq N, |j-j_0| \leq N} -2(\omega \cdot l + \theta) \mathbf{1}_{d_j} \geq N \mathbf{1},$$

we apply Lemma 5.1 of [5] with  $\alpha = N^{-\tau_1}$ ,  $\beta = N$  and  $|E| \leq CN^{\mathfrak{d}+d}$  (this is due to the bound  $d_j \leq |j + \rho|_2^{0-r}$ ) and obtain

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

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

Consider now  $|j_0| < (c+5)c^{-1}N$ . We obtain a complexity estimate for  $B_N^0(j_0, \varepsilon, \lambda)$  by knowing the measure of the set

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

**Lemma 3.9.** *For all  $|j_0| < (c+5)c^{-1}N$  and all  $\lambda \in [1/2, 3/2]$  one has*

$$B_{2,N}^0(j_0, \varepsilon, \lambda) \subset I_N := [-gN, gN], \quad g := (2c+8)\mathbb{C}c^{-1}.$$

*Proof.* If  $|\theta| > gN$  one has  $|\omega \cdot l + \theta| \geq |\theta| - |\omega \cdot l| > (g - (3/2))N > (2c+6)\mathbb{C}c^{-1}N$ . Using Remark 3.7, all the eigenvalues

$$\mu_{l,j,\sigma}(\theta) = -(\omega \cdot l + \theta)^2 + m - \mu_j + O(\varepsilon\|T\|_0) \leq -(\mathbb{C}c^{-1}N)^2, \quad \forall |\theta| > gN,$$

proving the lemma.  $\blacksquare$

**Lemma 3.10.** *For all  $|j_0| \leq (c+5)c^{-1}N$  and all  $\lambda \in [1/2, 3/2]$  one has*

$$B_N^0(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{\hat{C}\mathfrak{M}N^{\tau_1+1}} I_q, \quad I_q = I_q(j_0) \text{ intervals with } \text{meas}(I_q) \leq N^{-\tau_1}$$

where  $\mathfrak{M} := \text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda))$  and  $\hat{C} = \hat{C}(r)$ .

*Proof.* This is Lemma 5.5 of [5], where our exponent  $\tau_1$  is denoted by  $\tau$ .  $\blacksquare$

Lemmas 3.9 and 3.10 imply that for all  $\lambda \in [1/2, 3/2]$  the set  $B_N^0(j_0, \varepsilon, \lambda)$  can be covered by  $\sim N^{\tau_1+2}$  intervals of length  $\leq N^{-\tau_1}$ . This estimate is not enough. Now we prove that for “most”  $\lambda$  the number of such intervals does not depend on  $\tau_1$ , showing that  $\text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda)) = O(N^{\varepsilon-\tau_1})$  where  $\varepsilon = d + \mathfrak{d} + r + 4$  has been fixed in Proposition 3.6. To this purpose first we provide an estimate for the set

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

Then in Lemma 3.12 we use Fubini Theorem to obtain the desired bound for  $\text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda))$ .

**Lemma 3.11.** *For all  $|j_0| < (c+5)c^{-1}N$  one has  $\text{meas}(B_{2,N}^0(j_0, \varepsilon)) \leq CN^{-\tau_1+\mathfrak{d}+d+1}$  for some  $C > 0$ .*

*Proof.* Let us introduce the variables

$$\zeta = \frac{1}{\lambda^2}, \quad \eta = \frac{\theta}{\lambda}, \quad (\zeta, \eta) \in [4/9, 4] \times [-2gN, 2gN] =: [4/9, 4] \times J_N, \quad (3.19)$$

and set

$$L(\zeta, \eta) := \lambda^{-2} L_{N, j_0}(\varepsilon, \lambda, \theta) = \text{diag}_{|l| \leq N, |j-j_0| \leq N} \left( (-\bar{\omega} \cdot l + \eta)^2 + \zeta(-\mu_j + m) \right) \mathbf{1}_{d_j} + \varepsilon \zeta T(\varepsilon, 1/\sqrt{\zeta}).$$

Note that

$$\min_{j \in \Lambda_+(G)} -\mu_j + m \geq m. \quad (3.20)$$

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

$$\|L(\zeta, \eta)^{-1}\|_0 \leq N^{\tau_1} / 8. \quad (3.21)$$

Indeed

$$\partial_\zeta L(\zeta, \eta) = \text{diag}_{|l| \leq N, |j-j_0| \leq N} \left( (-\mu_j + m) \mathbf{1}_{d_j} \right) + \varepsilon T(\varepsilon, 1/\sqrt{\zeta}) - \frac{\varepsilon}{2} \zeta^{-1/2} \partial_\lambda T \stackrel{(3.20)}{\geq} \frac{m}{2} \mathbf{1},$$

for  $\varepsilon$  small (we used that  $\zeta \in [4/9, 4]$ ). Therefore Lemma 5.1 of [5] implies that for each  $\eta$ , the set of  $\zeta$  such that at least one eigenvalue of  $L(\zeta, \eta)$  has modulus  $\leq 8N^{-\tau_1}$ , is contained in the union of  $O(N^{d+d})$  intervals with length  $O(N^{-\tau_1})$  and hence has measure  $\leq O(N^{-\tau_1 + d + d})$ . Integrating in  $\eta \in J_N$  we obtain (3.21) except in a set with measure  $O(N^{-\tau_1 + d + d + 1})$ . The same measure estimates hold in the original variables  $(\lambda, \theta)$  in (3.19). Finally (3.21) implies

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

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

The same argument implies that

$$\text{meas}([1/2, 3/2] \setminus \bar{\mathfrak{G}}_N) \leq N^{-\tau_1 + d + d + 1} \quad (3.22)$$

where  $\bar{\mathfrak{G}}_N$  is defined in (2.34).

Define the set

$$\mathcal{F}_N(j_0) := \left\{ \lambda \in [1/2, 3/2] : \text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda)) \geq \hat{C} N^{-\tau_1 + d + d + r + 2} \right\} \quad (3.23)$$

where  $\hat{C}$  is the constant appearing in Lemma 3.10.

**Lemma 3.12.** *For all  $|j_0| \leq (c+5)c^{-1}N$  one has  $\text{meas}(\mathcal{F}_N(j_0)) = O(N^{-r-1})$ .*

*Proof.* By Fubini Theorem we have

$$\text{meas}(B_{2,N}^0(j_0, \varepsilon)) = \int_{1/2}^{3/2} d\lambda \text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda)).$$

Now, for any  $\beta > 0$ , using Lemma 3.11 we have

$$\begin{aligned} CN^{-\tau_1 + d + d + 1} &\geq \int_{1/2}^{3/2} d\lambda \text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda)) \\ &\geq \beta \text{meas}(\{\lambda \in [1/2, 3/2] : \text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda)) \geq \beta\}) \end{aligned}$$

and for  $\beta = \hat{C} N^{-\tau_1 + r + d + d + 2}$  we prove the lemma (recall (3.23)). ■

**Lemma 3.13.** *If  $\tau_0 > r + 3d + 1$  then  $\text{meas}([1/2, 3/2] \setminus \bar{\mathcal{I}}) = O(N_0^{-1})$  where  $\bar{\mathcal{I}}$  is defined in (2.30).*

*Proof.* Let us write

$$[1/2, 3/2] \setminus \bar{\mathcal{I}} = \bigcup_{|l|, |j| \leq N_0} \mathcal{R}_{l,j}, \quad \mathcal{R}_{l,j} := \left\{ \lambda \in \Lambda : |(\lambda \bar{\omega} \cdot l)^2 + \mu_j - m| \leq N_0^{-\tau_0} \right\}.$$

Since  $-\mu_j + m \geq m > 0$ , then  $\mathcal{R}_{0,j} = \emptyset$  if  $N_0 > m^{-1/\tau_0}$ . For  $l \neq 0$ , using the Diophantine condition (1.5), we get  $\text{meas}(\mathcal{R}_{l,j}) \leq CN_0^{-\tau_0+2d}$ , so that

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

because  $\tau_0 - r - 3d > 1$ . ■

The measure of the set  $\tilde{\mathcal{I}}$  in (3.8) is estimated in [5]-Lemma 6.3 (where  $\tilde{\mathcal{I}}$  is denoted by  $\tilde{\mathcal{G}}$ ).

**Lemma 3.14.** *If  $\gamma < \min(1/4, \gamma_0/4)$  (where  $\gamma_0$  is that in (1.6)) then  $\text{meas}([1/2, 3/2] \setminus \tilde{\mathcal{I}}) = O(\gamma)$ .*

*Proof of Proposition 3.6 completed.* By the definition in (3.23) for all  $\lambda \notin \mathcal{F}_N(j_0)$  one has  $\text{meas}(B_{2,N}^0(j_0, \varepsilon, \lambda)) < O(N^{-\tau_1+r+\mathfrak{d}+d+2})$ . Thus for any  $\lambda \notin \mathcal{F}_N(j_0)$ , applying Lemma 3.10 we have

$$B_N^0(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^{r+\mathfrak{d}+d+4}} I_q, \quad I_q \text{ intervals with } \text{meas}(I_q) \leq N^{-\tau_1}.$$

But then, using also Lemma 3.8, we have that (recall (2.35) with  $\varepsilon = r + \mathfrak{d} + d + 4$ )

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

Hence, using Lemma 3.12,

$$\text{meas}(\mathcal{I} \setminus \bar{\mathcal{G}}_N^0) \leq \sum_{|j_0| \leq (c+5)c^{-1}N} \text{meas}(\mathcal{F}_N(j_0)) \leq O(N^{-1})$$

which is the first bound in (2.37). The second bound follows by (3.22) with  $\tau_1 > d + \mathfrak{d} + 2$ . Finally, Lemmas 3.13 and 3.14 with  $\gamma = N_0^{-1}$  implies the third estimate in (2.37). ■

### 3.3 Proof of Theorem 1.1 for NLS

In order to apply Theorems 2.16-2.18 to the Hamiltonian NLS, we start by defining two extensions  $\mathfrak{F}(u, v), \mathfrak{H}(u, v)$  of class  $C^q(\mathbb{T}^d \times \mathbb{M} \times \mathbb{C}^2; \mathbb{C})$  (in the real sense) of  $\mathbf{f}(u)$  in such a way that  $\mathfrak{F}(u, \bar{u}) = \mathfrak{H}(u, \bar{u}) = \mathbf{f}(u)$  and  $\partial_u \mathfrak{F}(u, \bar{u}) = \partial_v \mathfrak{H}(u, \bar{u}) \in \mathbb{R}$ ,  $\partial_{\bar{u}} \mathfrak{F}(u, \bar{u}) = \partial_{\bar{u}} \mathfrak{H}(u, \bar{u}) = \partial_{\bar{v}} \mathfrak{F}(u, \bar{u}) = \partial_{\bar{v}} \mathfrak{H}(u, \bar{u}) = 0$

and  $\partial_v \mathfrak{F}(u, \bar{u}) = \overline{\partial_u \mathfrak{H}(u, \bar{u})}^1$ . Then we “double” the NLS equation, namely we look for a zero of the vector NLS operator

$$F(\varepsilon, \lambda, u^+, u^-) := \begin{cases} i\lambda\bar{\omega} \cdot \partial_\varphi u^+ - \Delta u^+ + mu^+ - \varepsilon \mathfrak{F}(\varphi, x, u^+, u^-) \\ -i\lambda\bar{\omega} \cdot \partial_\varphi u^- - \Delta u^- + mu^- - \varepsilon \mathfrak{H}(\varphi, x, u^+, u^-) \end{cases} \quad (3.24)$$

on the space  $H^s(\mathbb{T}^d \times G) \times H^s(\mathbb{T}^d \times G)$ . Note that (3.24) reduces to (1.2) on the invariant subspace

$$\mathcal{U} := \{u = (u^+, u^-) \in H^s \times H^s : u^- = \overline{u^+}\}.$$

Setting  $\mathfrak{A} = \{1, -1\}$ ,  $\Lambda_+ := \Lambda_+(G)$ , we are in the functional setting of Example 2, namely  $H^s(\mathbb{T}^d \times G) \times H^s(\mathbb{T}^d \times G) \equiv H^s(\mathfrak{R})$  where  $\mathfrak{R} = \mathbb{Z}^d \times \Lambda_+(G) \times \{1, -1\}$ . Then equation (3.24) is of the form (2.20) with

$$D(\lambda) := \begin{pmatrix} i\lambda\bar{\omega} \cdot \partial_\varphi - \Delta + m & 0 \\ 0 & -i\lambda\bar{\omega} \cdot \partial_\varphi - \Delta + m \end{pmatrix}, \quad f(u) := - \begin{pmatrix} \mathfrak{F}(\varphi, x, u^+, u^-) \\ \mathfrak{H}(\varphi, x, u^+, u^-) \end{pmatrix},$$

and (2.21) holds with  $\nu = 2$ . Again the interpolation estimates (f1)–(f2) are verified if  $s_0 \geq (\mathfrak{d} + d)/2$ .

In the Fourier basis  $e^{il \cdot \varphi} \mathbf{f}_j(x)$  the operator  $D(\lambda)$  is represented by an infinite dimensional matrix as in (2.23), with  $D_{(l,j,\mathfrak{a})}(\lambda) = -\mathfrak{a}\lambda\bar{\omega} \cdot l - \mu_j + m$  and

$$T_{(l',j')}^{(l,j)} := \begin{pmatrix} P_{j'}^j(l-l') & Q_{j'}^j(l-l') \\ \overline{Q}_{j'}^j(l-l') & P_{j'}^j(l-l') \end{pmatrix},$$

where  $P_{j'}^j(l), Q_{j'}^j(l)$  are the matrix representation in the Fourier basis of the multiplication operators

$$P(\varphi, x) := -\partial_{u^+} \mathfrak{F}(\varphi, x, u^+(\varphi, x), u^-(\varphi, x)), \quad Q(\varphi, x) := -\partial_{u^-} \mathfrak{H}(\varphi, x, u^+(\varphi, x), u^-(\varphi, x)).$$

By the Hamiltonian assumption (1.3), the constraints on  $\mathfrak{F}, \mathfrak{H}$  and  $(u^-, u^+) \in \mathcal{U}$ , it results  $P(\varphi, x) \in \mathbb{R}$  and  $(T_{(l',j')}^{(l,j)})^\dagger = T_{(l,j)}^{(l',j')}$ . Hypothesis 1 is then verified because (2.24a) holds with  $\mathfrak{D}_{j,\mathfrak{a}}(y) = -\mathfrak{a}y - \mu_j + m$ ,

$$T_{(l,j,\mathfrak{a})}^{(l',j',\mathfrak{a})} = P_{j'}^j(l-l'), \quad \mathfrak{a} = \pm 1, \quad T_{(l,j,1)}^{(l',j',-1)} = Q_{j'}^j(l-l'),$$

Corollary 3.3 implies that  $T \in \mathcal{M}^{s-\nu_0}$  and the estimates (2.24c), (2.24d) hold by interpolation.

We introduce the additional parameter  $\theta$  and following (2.25) we define the matrices

$$L(\varepsilon, \lambda, \theta, u) := D(\lambda, \theta) + \varepsilon T(\lambda, u), \quad D(\lambda, \theta) := D(\lambda) + \theta Y, \quad Y := \text{diag}_{(l,j) \in \mathbb{Z}^d \times \Lambda_+(G)} \begin{pmatrix} -\mathbf{1}_{d_j} & 0 \\ 0 & \mathbf{1}_{d_j} \end{pmatrix}$$

so that Hypothesis 2 holds with  $\mathbf{n} = 2$ .

Hypothesis 3 and the measure estimates needed in Theorem 2.18 are obtained as in the case of NLW, following [4] instead of [5].

<sup>1</sup>where, for  $u = r + is$ ,  $v = a + ib$  we set  $\partial_u := (\partial_r - i\partial_s)/2$ ,  $\partial_{\bar{u}} := (\partial_r + i\partial_s)/2$ ,  $\partial_v := (\partial_a - i\partial_b)/2$ ,  $\partial_{\bar{v}} := (\partial_a + i\partial_b)/2$ . A possible extension is the following: writing  $\mathbf{f}(\varphi, x, r + is) = \mathbf{f}_1(r, s) + i\mathbf{f}_2(r, s)$  we consider

$$\begin{aligned} \mathfrak{F}(\varphi, x, r + is, a + ib) &:= (1+i)\mathbf{f}_1\left(\frac{r+a}{2}, r-a+s\right) - i\mathbf{f}_1\left(\frac{r+a}{2} - \frac{s+b}{2}, r-a+s\right) \\ &\quad + \mathbf{f}_2\left(2a-r-(s+b), \frac{r-a}{2} + \frac{s-b}{2}\right) + (-1+i)\mathbf{f}_2\left(a, \frac{s-b}{2}\right) \\ \mathfrak{H}(\varphi, x, r + is, a + ib) &:= (1+i)\mathbf{f}_1\left(\frac{r+a}{2}, \frac{a-r}{2} + \frac{s-b}{2}\right) - i\mathbf{f}_1\left(\frac{r+a}{2} - \frac{s+b}{2}, \frac{r-a}{2} + \frac{s-b}{2}\right) \\ &\quad + \mathbf{f}_2\left(\frac{r+a}{2} + \frac{s+b}{2}, \frac{a-r}{2} + \frac{s-b}{2}\right) - (1+i)\mathbf{f}_2\left(a, \frac{s-b}{2}\right). \end{aligned}$$

## 4 An abstract Nash-Moser theorem

Let us consider a scale of Banach spaces  $(X_s, \|\cdot\|_s)_{s \geq 0}$ , such that

$$\forall s \leq s', \quad X_{s'} \subseteq X_s \quad \text{and} \quad \|u\|_s \leq \|u\|_{s'}, \quad \forall u \in X_{s'},$$

and define  $X := \bigcap_{s \geq 0} X_s$ .

We assume that there is a non-decreasing family  $(E^{(N)})_{N \geq 0}$  of subspaces of  $X$  such that  $\bigcup_{N \geq 0} E^{(N)}$  is dense in  $X_s$  for any  $s \geq 0$ , and that there are projectors

$$\Pi^{(N)} : X_0 \rightarrow E^{(N)}$$

satisfying: for any  $s \geq 0$  and any  $\nu \geq 0$  there is a positive constant  $C := C(s, \nu)$  such that

$$(P1) \quad \|\Pi^{(N)} u\|_{s+\nu} \leq CN^\nu \|u\|_s \quad \text{for all } u \in X_s$$

$$(P2) \quad \|(\mathbf{1} - \Pi^{(N)})u\|_s \leq CN^{-\nu} \|u\|_{s+\nu} \quad \text{for all } u \in X_{s+\nu}.$$

In every Banach scale with smoothing operators satisfying (P1)-(P2) as above, the following interpolation inequality holds (see Lemma 1.1 in [6]): for all  $s_1 < s_2$ ,  $t \in [0, 1]$ ,

$$\|u\|_{ts_1+(1-t)s_2} \leq K(s_1, s_2) \|u\|_{s_1}^t \|u\|_{s_2}^{1-t}. \quad (4.1)$$

**Remark 4.1.** *The sequence spaces  $H^s(\mathfrak{R})$  defined in (2.5) admit spaces  $E^{(N)} := \{u = \{u_k\}_{k \in \mathfrak{R}} : u_k = 0 \text{ for } |k| > N\}$  whose corresponding projectors  $\Pi^{(N)}$  satisfy (P1) – (P2).*

Let us consider a parameter family of  $C^2$  maps  $F : [0, \varepsilon_0) \times \mathcal{I} \times X_{s_0+\nu} \rightarrow X_{s_0}$  for some  $s_0 \geq 0$ ,  $\nu > 0$ ,  $\varepsilon_0 > 0$  and  $\mathcal{I}$  an interval in  $\mathbb{R}$ . We assume

$$(F0) \quad F(0, \lambda, 0) = 0 \quad \text{for any } \lambda \in \mathcal{I},$$

and the following tame properties: given  $S' > s_0$ ,  $\forall s \in [s_0, S')$ , for all  $\|u\|_{s_0} \leq 1$ ,  $(\varepsilon, \lambda) \in [0, \varepsilon_0) \times \mathcal{I}$ ,

$$(F1) \quad \|\partial_\lambda F(\varepsilon, \lambda, u)\|_s \leq C(s)(1 + \|u\|_{s+\nu}),$$

$$(F2) \quad \|D_u F(\varepsilon, \lambda, 0)[h]\|_s \leq C(s)\|h\|_{s+\nu},$$

$$(F3) \quad \|D_u^2 F(\varepsilon, \lambda, u)[h, v]\|_s \leq C(s)(\|u\|_{s+\nu}\|h\|_{s_0}\|v\|_{s_0} + \|h\|_{s+\nu}\|v\|_{s_0} + \|h\|_{s_0}\|v\|_{s+\nu}),$$

$$(F4) \quad \|\partial_\lambda D_u F(\varepsilon, \lambda, u)[h]\|_s \leq C(s)(\|h\|_{s+\nu} + \|u\|_{s+\nu}\|h\|_{s_0}).$$

In application the following assumption is often verified

$$\exists \widehat{X}_s \subset X_s \text{ closed subspaces of } X_s, \quad s \geq 0, \quad \text{such that } F : \widehat{X}_{s+\nu} \rightarrow \widehat{X}_s. \quad (4.2)$$

In order to prove the existence of a zero for  $F$  we shall follow a Nash-Moser approach whose main assumption concerns the invertibility of the linearised operators

$$L^{(N)}(\varepsilon, \lambda, u) := \Pi^{(N)} D_u F(\varepsilon, \lambda, u)|_{E^{(N)}}$$



in a neighborhood of  $u = 0$ .

We introduce parameters  $\sigma, \tau > 0, s_1 > s_0, \delta \in (0, 1), S \in (s_0, S')$  satisfying

$$\sigma \geq \max\{2(\tau + \delta s_1) + 3\nu + 2, 4(\tau + \delta s_1 + \nu)\}, \quad 2(2(\tau + \delta s_1) + \nu + 3 + \sigma) \leq S - s_1 \leq 4(\sigma + 1). \quad (4.3)$$

Define the sets

$$J_{\tau, \delta}^{(N)} := \left\{ (\lambda, u) \in \mathcal{I} \times E^{(N)} : \|u\|_{s_1} \leq 1, L^{(N)}(\varepsilon, \lambda, u) \text{ is invertible and, } \forall s \in [s_1, S], \right. \\ \left. \varepsilon \in [0, \varepsilon_0), \|L^{(N)}(\varepsilon, \lambda, u)^{-1}[h]\|_s \leq C(s)N^\mu(\|h\|_s + N^{\delta(s-s_1)}\|u\|_s\|h\|_{s_1}), \mu := \tau + \delta s_1 \right\}. \quad (4.4)$$

For  $K > 0$  and  $u \in C^1(\mathcal{I}, E^{(N)})$  satisfying  $\|u\|_{s_1} \leq 1, \|\partial_\lambda u\|_{s_1} \leq K$ , we set

$$\mathbf{G}_{\tau, \delta}^{(N)}(u) := \{\lambda \in \mathcal{I} : (\lambda, u(\lambda)) \in J_{\tau, \delta}^{(N)}\}. \quad (4.5)$$

Given  $N_0 \in \mathbb{N}$  set  $N_n := N_0^{2^n}$  and denote with  $E_n, \Pi_n, J_{\tau, \delta}^n$  the subspace  $E^{(N_n)}$ , the projector  $\Pi^{(N_n)}$  and the set  $J_{\tau, \delta}^{(N_n)}$  respectively. Given any set  $A$  and a positive real number  $\eta$  we denote by  $\mathcal{N}(A, \eta)$  the open neighborhood of  $A$  with width  $\eta$ .

**Theorem 4.2. (Nash-Moser)** *Assume (F0)-(F4). Then, for all  $\tau > 0, \delta \in (0, 1/4), \sigma, s_1 > s_0, S < S'$ , satisfying (4.3), there are  $c, \bar{N}_0, K_0 > 0$ , such that, for all  $N_0 \geq \bar{N}_0$  and  $\varepsilon_0$  small enough such that*

$$\varepsilon_0 N_0^S \leq c, \quad (4.6)$$

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

$$(S1)_n \quad u_n(\varepsilon, \lambda) \in E_n, u_n(0, \lambda) = 0, \|u_n\|_{s_1} \leq 1 \text{ and } \|\partial_\lambda u_n\|_{s_1} \leq K_0 N_0^{\sigma/2}.$$

$$(\widehat{S1})_n \quad \text{If (4.2) holds then } u_n(\varepsilon, \lambda) \in E_n \cap \widehat{X}_s.$$

$$(S2)_n \quad \text{For all } 1 \leq i \leq n \text{ one has } \|u_i - u_{i-1}\|_{s_1} \leq N_i^{-\sigma-1} \text{ and } \|\partial_\lambda(u_i - u_{i-1})\|_{s_1} \leq N_i^{-1-\nu}.$$

(S3)<sub>n</sub> Set  $u_{-1} := 0$  and define

$$A_n := \bigcap_{i=0}^n \mathbf{G}_{\tau, \delta}^{(N_i)}(u_{i-1}). \quad (4.7)$$

For  $\lambda \in \mathcal{N}(A_n, N_n^{-\sigma/2})$  the function  $u_n(\varepsilon, \lambda)$  solves the equation  $\Pi_n F(\varepsilon, \lambda, u) = 0$ .

(S4)<sub>n</sub> Setting  $B_n := 1 + \|u_n\|_S$  and  $B'_n := \|\partial_\lambda u_n\|_S$  one has

$$B_n \leq 2N_{n+1}^p, \quad p := \mu + \frac{\nu}{2} + 1, \quad (4.8a)$$

$$B'_n \leq 2N_{n+1}^q, \quad q := 2\mu + \nu + 2 + (\sigma/2). \quad (4.8b)$$

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

$$\|u_\varepsilon(\lambda) - u_n(\lambda)\|_{s_1} \leq N_{n+1}^{-\sigma-1}, \quad \forall \lambda \in \mathcal{I}, \quad (4.9)$$

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

Finally, if (4.2) holds then  $u_\varepsilon(\lambda) \in \widehat{X}_{s_1+\nu}$ .

## 4.1 Proof of Theorem 4.2

Taylor formula and (F0)–(F4) imply the following tame properties: for any  $s \in [s_0, S]$  there is  $C = C(s)$  such that for any  $u, h \in X_s$  with  $\|u\|_{s_0} \leq 2$  and  $\|h\|_{s_0} \leq 1$  one has

$$(F5) \quad \|F(\varepsilon, \lambda, u)\|_s \leq C(s)(\varepsilon + \|u\|_{s+\nu}),$$

$$(F6) \quad \|D_u F(\varepsilon, \lambda, u)[h]\|_s \leq C(s)(\|u\|_{s+\nu}\|h\|_{s_0} + \|h\|_{s+\nu})$$

$$(F7) \quad \|F(\varepsilon, \lambda, u+h) - F(\varepsilon, \lambda, u) - D_u F(\varepsilon, \lambda, u)[h]\|_s \leq C(s)(\|u\|_{s+\nu}\|h\|_{s_0}^2 + \|h\|_{s+\nu}\|h\|_{s_0}).$$

The following Lemma follows as in [6], Lemma 2.2.

**Lemma 4.3.** *Let  $(\lambda, u) \in J_{\tau, \delta}^{(N)}$  with  $\|u\|_{s_1} \leq 1$ . For  $\varepsilon$  small enough, there exists  $c = c(S) > 0$  such that, if  $|\lambda' - \lambda| + \|h\|_{s_1} \leq cN^{-(\mu+\nu)}$ ,  $h \in E^{(N)}$ , then  $L^{(N)}(\varepsilon, \lambda', u+h)$  is invertible and*

$$\left\| L^{(N)}(\varepsilon, \lambda', u+h)^{-1}[v] \right\|_{s_1} \leq 2N^\mu \|v\|_{s_1}, \quad \forall v \in E^{(N)}, \quad (4.10a)$$

$$\left\| L^{(N)}(\varepsilon, \lambda', u+h)^{-1}[v] \right\|_S \leq 2N^\mu \|v\|_S + KN^\mu (N^{\mu+\nu}(\|u\|_S + \|h\|_S) + N^{\delta(S-s_1)}\|u\|_S) \|v\|_{s_1}. \quad (4.10b)$$

## 4.2 Initialisation of the Nash-Moser scheme

Set  $A_0 := \mathbf{G}_{\tau, \delta}^{(N_0)}(0)$ . By (4.5) we have that  $\lambda \in A_0$  if and only if  $(\lambda, 0) \in J_{\tau, \delta}^{(N_0)}$ . Therefore, if  $N_0$  is large enough Lemma 4.3 ensures that for all  $\lambda \in \mathcal{N}(A_0, 2N_0^{-\sigma/2})$  the operator  $L^{(N_0)}(\varepsilon, \lambda, 0)$  is invertible for  $\varepsilon$  small enough and

$$\|L^{(N_0)}(\varepsilon, \lambda, 0)^{-1}\|_{s_1} \leq 2N_0^\mu, \quad \|L^{(N_0)}(\varepsilon, \lambda, 0)^{-1}\|_S \leq 2N_0^\mu. \quad (4.11)$$

Let us denote

$$\begin{aligned} L_0 &:= L^{(N_0)}(\varepsilon, \lambda, 0), \quad r_{-1} := \Pi_0 F(\varepsilon, \lambda, 0) \\ R_{-1}(u) &:= \Pi_0 (F(\varepsilon, \lambda, u) - F(\varepsilon, \lambda, 0) - D_u F(\varepsilon, \lambda, 0)[u]). \end{aligned}$$

We look for a solution of the equation  $\Pi_0 F(\varepsilon, \lambda, u) = 0$ , as a fixed point of the map

$$\mathcal{H}_0 : E_0 \longrightarrow E_0, \quad u \longmapsto \mathcal{H}_0(u) := -L_0^{-1}(r_{-1} + R_{-1}(u)).$$

Let us show that  $\mathcal{H}_0$  is a contraction in the set  $\mathbf{B}_{\rho_0} = \{u \in E_0 : \|u\|_{s_1} \leq \rho_0 := c_0 N_0^\mu \varepsilon\}$ , for all  $\varepsilon \in [0, \varepsilon_0(N_0)]$  and some  $c_0 = c_0(s_1)$ . We bound

$$\|r_{-1}\|_{s_1} \leq \|F(\varepsilon, \lambda, 0)\|_{s_1} \stackrel{(F5)}{\leq} C(s_1)\varepsilon, \quad \|R_{-1}\|_{s_1} \leq C(s_1)\|\Pi_0 u\|_{s_1+\nu}^2 \stackrel{(P1)}{\leq} C(s_1)N_0^\nu \|u\|_{s_1}^2, \quad (4.12)$$

so that for any  $u \in \mathbf{B}_{\rho_0}$  one has

$$\|\mathcal{H}_0(u)\|_{s_1} \stackrel{(4.11), (4.12)}{\leq} 2N_0^\mu C(s_1)(\varepsilon + N_0^\nu \|u\|_{s_1}^2) \stackrel{def}{\leq} 2C(s_1)N_0^\mu \varepsilon + 2C(s_1)N_0^{\mu+\nu} \rho_0^2 \leq \rho_0$$

where we have set  $c_0 = 4C(s_1)$  and using (4.6). This means that  $\mathcal{H}_0$  maps  $\mathbf{B}_{\rho_0}$  into itself. In the same way (using (F3)) we obtain  $\|D\mathcal{H}_0(u)[h]\|_{s_1} \leq \|h\|_{s_1}/2$  so that  $\mathcal{H}_0$  is a contraction on  $(\mathbf{B}_{\rho_0}, \|\cdot\|_{s_1})$  and hence it admits a unique fixed point  $\tilde{u}_0(\varepsilon, \lambda)$  for all  $\lambda \in \mathcal{N}(A_0, 2N_0^{-\sigma/2})$ .

Now, for  $\lambda \in \mathcal{N}(A_0, 2N_0^{-\sigma/2})$  one has, by (F0), that  $\tilde{u}_0(0, \lambda) = 0$ . The Implicit Function Theorem ensures that  $\tilde{u}_0(\varepsilon, \cdot) \in C^1(\mathcal{N}(A_0, 2N_0^{-\sigma/2}); \mathbf{B}_{\rho_0})$  and  $\partial_\lambda \tilde{u}_0 = L^{(N_0)}(\varepsilon, \lambda, \tilde{u}_0)^{-1}[\Pi_0 \partial_\lambda F(\varepsilon, \lambda, \tilde{u}_0)]$ . Hence

$$\|\partial_\lambda \tilde{u}_0\|_{s_1} \stackrel{(4.10a)}{\leq} 2N_0^\mu \|\Pi_0 \partial_\lambda F(\varepsilon, \lambda, \tilde{u}_0)\|_{s_1} \stackrel{(F1), (P1)}{\leq} 2N_0^\mu C(s_1)(1 + N_0^{\mu+\nu} c_0 \varepsilon) \leq C(s_1) N_0^\mu$$

for  $\varepsilon$  small.

We now define  $u_0 := \psi_0 \tilde{u}_0 : [0, \varepsilon_0] \times \mathcal{I} \rightarrow E_0$  where  $\psi_0$  is a  $C^1(\mathcal{I}, \mathbb{R})$  cut-off function such that  $0 \leq \psi_0 \leq 1$  and

- $\psi_0(\lambda) = 1$  for  $\lambda \in \mathcal{N}(A_0, N_0^{-\sigma/2})$  and  $\psi_0(\lambda) = 0$  for  $\lambda \notin \mathcal{N}(A_0, 2N_0^{-\sigma/2})$ ,
- $|\partial_\lambda \psi_0| \leq CN_0^{\sigma/2}$ .

Of course  $u_0$  satisfies (S3)<sub>0</sub>. Moreover  $u_0(0, \lambda) = 0$  and satisfies also (S1)<sub>0</sub>, since

$$\|u_0\|_{s_1} \leq 1/2, \quad \|\partial_\lambda u_0\|_{s_1} \leq CN_0^{\sigma/2} + C(s_1)N_0^\mu \leq C(s_1)N_0^{\sigma/2} \quad (4.13)$$

because  $\sigma > 2\mu$ . Finally, the bounds (S4)<sub>0</sub> follow in the same way.

### 4.3 Iterative step

Suppose inductively that we have defined  $u_n \in C^1(\mathcal{I}, E_n)$  such that properties (S1)<sub>n</sub> – (S4)<sub>n</sub> hold. We define  $u_{n+1}$  as follows. For  $h \in E_{n+1}$  let us write

$$\Pi_{n+1} F(\varepsilon, \lambda, u_n(\varepsilon, \lambda) + h) = r_n + L_{n+1}[h] + R_n(h),$$

where

$$\begin{aligned} r_n &:= \Pi_{n+1} F(\varepsilon, \lambda, u_n), & L_{n+1} &:= L_{n+1}(\varepsilon, \lambda) := L^{(N_{n+1})}(\varepsilon, \lambda, u_n(\varepsilon, \lambda)), \\ R_n(h) &:= \Pi_{n+1} (F(\varepsilon, \lambda, u_n(\varepsilon, \lambda) + h) - F(\varepsilon, \lambda, u_n) - D_u F(\varepsilon, \lambda, u_n)[h]). \end{aligned}$$

Note that (F7) and (S1)<sub>n</sub> imply

$$\|R_n(h)\|_s \leq C(s)(\|u_n\|_{s+\nu} \|h\|_{s_1}^2 + \|h\|_{s+\nu} \|h\|_{s_1}), \quad (4.14)$$

and for  $\lambda \in \mathcal{N}(A_n, N_n^{-\sigma/2})$ , we use (S3)<sub>n</sub> to obtain

$$r_n = \Pi_{n+1} F(\varepsilon, \lambda, u_n) - \Pi_n F(\varepsilon, \lambda, u_n) = \Pi_{n+1} (\mathbf{1} - \Pi_n) F(\varepsilon, \lambda, u_n). \quad (4.15)$$

If  $A_{n+1} = \emptyset$  we define  $u_i := u_n$  for all  $i > n$ , otherwise we proceed as follows. By definition (4.7), for all  $\lambda \in A_{n+1}$ , the operator  $L_{n+1}(\varepsilon, \lambda)$  is invertible. We also note that for  $N_0$  large enough, one has

$$\mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2}) \subset \mathcal{N}(A_n, N_n^{-\sigma/2}). \quad (4.16)$$

**Lemma 4.4.** For  $\varepsilon$  small enough,  $\forall \lambda \in \mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2})$ , the operator  $L_{n+1}(\varepsilon, \lambda)$  is invertible and

$$\|L_{n+1}^{-1}[v]\|_{s_1} \leq 2N_{n+1}^\mu \|v\|_{s_1}, \quad (4.17a)$$

$$\|L_{n+1}^{-1}[v]\|_S \leq KN_{n+1}^\mu \left( \|v\|_S + \left( N_{n+1}^\xi B_n + N_{n+1}^{(\mu+\nu)-\sigma/2} B'_n \right) \|v\|_{s_1} \right) \quad (4.17b)$$

where  $\xi := \max\{\mu + \nu, \delta(S - s_1)\}$ .

*Proof.* We apply Lemma 4.3. For  $\lambda \in \mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2})$  there is  $\lambda' \in A_{n+1}$  such that, setting  $h(\lambda) := u_n(\varepsilon, \lambda) - u_n(\varepsilon, \lambda')$  one has (use (S1)<sub>n</sub>)

$$|\lambda - \lambda'| + \|h\|_{s_1} \leq 3N_{n+1}^{-\sigma/2} (1 + KN_0^{\sigma/2}) \leq cN_{n+1}^{-(\mu+\nu)}$$

using (4.3) and  $N_0$  large. ■

For all  $\lambda \in \mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2})$ , let consider the map

$$\mathcal{H}_{n+1} : E_{n+1} \longrightarrow E_{n+1}, \quad h \longmapsto \mathcal{H}_{n+1}(h) := -L_{n+1}^{-1}[r_n + R_n]. \quad (4.18)$$

**Lemma 4.5.** For  $N_0$  large enough  $\mathcal{H}_{n+1}$  has a unique fixed point  $\tilde{h}_{n+1} = \mathcal{H}(\tilde{h}_{n+1})$  in  $\mathbf{B}_{\rho_{n+1}} := \{h \in E_{n+1} : \|h\|_{s_1} \leq \rho_{n+1} := N_{n+1}^{-\sigma-1}\}$ . Moreover, the following estimate holds

$$\|\tilde{h}_{n+1}\|_{s_1} \leq 2C(s)N_n^{-(s-s_1)+2\tau+2\delta s_1+\nu} (1 + \|u_n\|_s), \quad \forall s \in (s_1, S'). \quad (4.19)$$

*Proof.* Let us prove that  $\mathcal{H}_{n+1}$  is a contraction in  $\mathbf{B}_{\rho_{n+1}}$ . For  $\lambda \in \mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2})$  we use (4.17a) and the definition of  $\mathcal{H}_{n+1}$  in order to bound  $\|\mathcal{H}_{n+1}(h)\|_{s_1} \leq 2N_{n+1}^\mu (\|r_n\|_{s_1} + \|R_n(h)\|_{s_1})$ . Now we have

$$\|r_n\|_{s_1} \stackrel{(4.15),(P2),(F5)}{\leq} C(S)N_{n+1}^{-(S-s_1)/2} (\varepsilon + \|u_n\|_{S+\nu}) \stackrel{(P1)}{\leq} C(S)N_{n+1}^{-(S-s_1-\nu)/2} B_n \quad (4.20)$$

where  $B_n := 1 + \|u_n\|_S$ , see (S4)<sub>n</sub>. On the other hand

$$\|R_n(h)\|_{s_1} \stackrel{(4.14),(P1)}{\leq} C(s_1)N_{n+1}^\nu \|h\|_{s_1}^2 \leq C(s_1)N_{n+1}^\nu \rho_{n+1}^2. \quad (4.21)$$

Therefore

$$2N_{n+1}^\mu (\|r_n\|_{s_1} + \|R_n(h)\|_{s_1}) \leq \rho_{n+1} \quad (4.22)$$

is implied by  $N_{n+1}^{-(S-s_1-\nu)/2} B_n + N_{n+1}^\nu \rho_{n+1}^2 \leq \rho_{n+1} N_{n+1}^{-\mu-1}$ , which in turn follows by (4.8a), (4.3). So  $\mathcal{H}_{n+1}(\mathbf{B}_{\rho_{n+1}}) \subseteq \mathbf{B}_{\rho_{n+1}}$ . The derivative  $D_h \mathcal{H}_{n+1}(h)[v] = -L_{n+1}^{-1} \Pi_{n+1} (D_u F(\varepsilon, \lambda, u_n + h)[v] - D_u F(\varepsilon, \lambda, u_n)[v])$ , satisfies

$$\begin{aligned} \|D_h \mathcal{H}_{n+1}(h)[v]\|_{s_1} &\stackrel{(4.17a)}{\leq} 2N_{n+1}^\mu \|D_u F(\varepsilon, \lambda, u_n + h)[v] - D_u F(\varepsilon, \lambda, u_n)[v]\|_{s_1} \\ &\stackrel{(F3)}{\leq} 4N_{n+1}^\mu \left( \|u_n\|_{s_1+\nu} \|h\|_{s_1} \|v\|_{s_1} + \|h\|_{s_1+\nu} \|v\|_{s_1} + \|h\|_{s_1} \|v\|_{s_1+\nu} \right) \\ &\stackrel{(P1)}{\leq} 12N_{n+1}^{\mu+\nu} \rho_{n+1} \|v\|_{s_1} \stackrel{\text{def}}{\leq} 12N_{n+1}^{\mu+\nu-\sigma-1} \|v\|_{s_1} \leq \|v\|_{s_1/2} \end{aligned}$$

using (4.3) and  $N_0$  is large. Then the Contraction Lemma implies the existence of a unique fixed point  $\tilde{h}_{n+1} = \mathcal{H}_{n+1}(\tilde{h}_{n+1})$ . Now, by (4.17a), and the first inequality in (4.21), we get

$$\|\tilde{h}_{n+1}\|_{s_1} \leq 2N_{n+1}^\mu (\|r_n\|_{s_1} + C(s_1)N_{n+1}^\nu \|\tilde{h}_{n+1}\|_{s_1}^2) \leq 2N_{n+1}^\mu \|r_n\|_{s_1} + 2N_{n+1}^{\mu+\nu} C(s_1) \rho_{n+1} \|\tilde{h}_{n+1}\|_{s_1}.$$

Using (4.20) with  $S \rightsquigarrow s$ ,  $B_n \rightsquigarrow (1 + \|u_n\|_s)$ , and  $2N_{n+1}^{\mu+\nu}C(s_1)\rho_{n+1} < 1/2$ , the bound (4.19) follows.  $\blacksquare$

For  $\lambda \in \mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2})$ , let  $\tilde{h}_{n+1}(\varepsilon, \lambda)$  be the unique solution of  $h = \mathcal{H}_{n+1}(h)$ . It results  $\tilde{h}_{n+1}(0, \lambda) = 0$ .

**Remark 4.6.** If (4.2) holds then  $L_{n+1}, \mathcal{H}_{n+1} : E_{n+1} \cap \widehat{X}_s \rightarrow E_{n+1} \cap \widehat{X}_s$  and so  $\tilde{h}_{n+1}(\varepsilon, \lambda) \in E_{n+1} \cap \widehat{X}_s$ .

**Lemma 4.7.**  $\|\tilde{h}_{n+1}\|_S \leq N_{n+1}^{2p}$  where  $2p = 2(\tau + \delta s_1) + \nu + 2$ , see (4.8a).

*Proof.* Using (4.17b), (4.22),  $\xi \geq \mu + \nu$  (see Lemma 4.4), we estimate

$$\begin{aligned} \|\tilde{h}_{n+1}\|_S &\leq KN_{n+1}^\mu \left( \|r_n\|_S + \|R_n(\tilde{h}_{n+1})\|_S + N_{n+1}^{\xi-\mu-1} \rho_{n+1} (B_n + N_{n+1}^{-\sigma/2} B'_n) \right) \\ &\leq KN_{n+1}^\mu \left( N_{n+1}^{\nu/2} B_n (1 + N_{n+1}^{\nu/2} \rho_{n+1}^2) + N_{n+1}^\nu \rho_{n+1} \|\tilde{h}_{n+1}\|_S + N_{n+1}^{\xi-\mu-1} \rho_{n+1} (B_n + N_{n+1}^{-\sigma/2} B'_n) \right) \end{aligned}$$

where in the last bound we used that,  $\forall \|h\|_{s_1} \leq \rho_{n+1}$ ,

$$\|r_n\|_S \stackrel{(P1),(F5)}{\leq} CN_{n+1}^{\nu/2} B_n, \quad \|R_n(h)\|_S \stackrel{(P1),(F7)}{\leq} K(N_{n+1}^\nu B_n \rho_{n+1}^2 + N_{n+1}^\nu \rho_{n+1} \|h\|_S). \quad (4.23)$$

Now, since  $N_{n+1}^{\nu+\mu} \rho_{n+1} \leq 1/2$  by (4.3), we shift  $\|\tilde{h}_{n+1}\|_S$  on the l.h.s. and obtain

$$\|\tilde{h}_{n+1}\|_S \leq K((N_{n+1}^{(\nu/2)+\mu} + N_{n+1}^{\xi-\sigma-2})B_n + N_{n+1}^{\xi-(3\sigma/2)-2}B'_n)$$

and the lemma follows by (4.8), (4.3).  $\blacksquare$

**Lemma 4.8.** The map  $\tilde{h}_{n+1}$  is in  $C^1(\mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2}); \mathbf{B}_{\rho_{n+1}})$  and

$$\|\partial_\lambda \tilde{h}_{n+1}\|_{s_1} \leq N_{n+1}^{-\nu-1}/2, \quad \|\partial_\lambda \tilde{h}_{n+1}\|_S \leq N_{n+1}^{2q}. \quad (4.24)$$

*Proof.* For  $\lambda \in \mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2})$  we have  $U_{n+1}(\lambda, \tilde{h}_{n+1}(\lambda)) \equiv 0$  where we have set  $U_{n+1}(\lambda, h) := \Pi_{n+1}F(\varepsilon, \lambda, u_n(\lambda) + h)$ . Hence

$$0 = \frac{d}{d\lambda}(U_{n+1}(\lambda, \tilde{h}_{n+1}(\lambda))) = (\partial_\lambda U_{n+1})(\lambda, \tilde{h}_{n+1}(\lambda)) + (D_h U_{n+1})(\lambda, \tilde{h}_{n+1}(\lambda)) \partial_\lambda \tilde{h}_{n+1}(\lambda). \quad (4.25)$$

On the other hand, since  $\|\tilde{h}_{n+1}\|_{s_1} \leq N_{n+1}^{-\sigma-1} \ll cN_{n+1}^{-(\mu+\nu)}$  for  $N_0$  large enough (recall that  $\mu + \nu < \sigma$ ), so that we apply Lemma 4.3 and obtain

$$\|(D_h U_{n+1}(\lambda, \tilde{h}_{n+1}))^{-1}[v]\|_{s_1} \leq 4N_{n+1}^\mu \|v\|_{s_1}, \quad (4.26a)$$

and, using also Lemma 4.7,

$$\left\| \left( (D_h U_{n+1}(\lambda, \tilde{h}_{n+1}))^{-1}[v] \right) \right\|_S \leq 4N_{n+1}^\mu \|v\|_S + KN_{n+1}^\mu \left( B_n \left( N_{n+1}^{\nu+\mu} + N_{n+1}^{\delta(S-s_1)} \right) + N_{n+1}^{\mu+\nu+2p} \right) \|v\|_{s_1}. \quad (4.26b)$$

Therefore by the Implicit Function Theorem we have  $\tilde{h}_{n+1} \in C^1(\mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2}); \mathbf{B}_{\rho_{n+1}})$  and

$$\partial_\lambda \tilde{h}_{n+1} = -((D_h U_{n+1})(\lambda, \tilde{h}_{n+1}(\lambda)))^{-1} (\partial_\lambda U_{n+1})(\lambda, \tilde{h}_{n+1}(\lambda)).$$

Now, by (4.16) we use (S3)<sub>n</sub> to deduce

$$\begin{aligned}\partial_\lambda U_{n+1}(\lambda, h) &= \Pi_{n+1}((\partial_\lambda F)(\varepsilon, \lambda, u_n + h) - (\partial_\lambda F)(\varepsilon, \lambda, u_n)) \\ &\quad + \Pi_{n+1}((D_u F)(\varepsilon, \lambda, u_n + h) - (D_u F)(\varepsilon, \lambda, u_n)) [\partial_\lambda u_n] \\ &\quad + \Pi_{n+1}(\mathbf{1} - \Pi_n)((\partial_\lambda F)(\varepsilon, \lambda, u_n) + (D_u F)(\varepsilon, \lambda, u_n)) [\partial_\lambda u_n].\end{aligned}$$

Now, by (F1)-(F4), (F6), (P1)-(P2), (S1)<sub>n</sub>, (4.26a), we get

$$\|\partial_\lambda \tilde{h}_{n+1}\|_{s_1} \leq CN_{n+1}^\mu (N_{n+1}^{\nu-\sigma-1} + N_{n+1}^{-\frac{S-s_1}{2}+(\nu/2)} (N_0^{\sigma/2} B_n + B'_n)) \leq N_{n+1}^{-\nu-1}/2$$

by (4.8) and (4.3). Now to get the estimate for the  $S$ -norm we use (4.26b) and obtain

$$\begin{aligned}\left\| \left( D_h U_{n+1}(\varepsilon, \lambda, \tilde{h}) \right)^{-1} [\partial_\lambda U_{n+1}(\varepsilon, \lambda, \tilde{h}_{n+1})] \right\|_S &\leq 4N_{n+1}^\mu \|\partial_\lambda U_{n+1}(\varepsilon, \lambda, \tilde{h}_{n+1})\|_S + KN_{n+1}^\mu \\ &\quad \times \left( B_n \left( N_{n+1}^{\nu+\mu} + N_{n+1}^{\delta(S-s_1)} \right) + N_{n+1}^{\mu+\nu+2p} \right) \|\partial_\lambda U_{n+1}(\varepsilon, \lambda, \tilde{h}_{n+1})\|_{s_1} \\ &\leq 4N_{n+1}^\mu (N_0^{\sigma/2} \|u_n\|_{S+\nu} + \|\tilde{h}_{n+1}\|_{S+\nu} + \|\partial_\lambda u_n\|_{S+\nu}) + KN_{n+1}^\mu \\ &\quad \times \left( B_n \left( N_{n+1}^{\nu+\mu} + N_{n+1}^{\delta(S-s_1)} \right) + N_{n+1}^{\mu+\nu+2p} \right) N_{n+1}^{-\mu-\nu-1} \leq N_{n+1}^{2q}\end{aligned}$$

by (4.8) and (4.3). ■

Let us define

$$h_{n+1}(\varepsilon, \lambda) := \psi_{n+1}(\lambda) \tilde{h}_{n+1}(\varepsilon, \lambda) \tag{4.27}$$

where  $\psi_{n+1}$  is a  $C^1$  cut-off function such that  $0 \leq \psi_{n+1} \leq 1$  and

- $\psi_{n+1}(\lambda) = 1$  for  $\lambda \in \mathcal{N}(A_{n+1}, N_{n+1}^{-\sigma/2})$  and  $\psi_{n+1}(\lambda) = 0$  for  $\lambda \notin \mathcal{N}(A_{n+1}, 2N_{n+1}^{-\sigma/2})$ ,
- $|\partial_\lambda \psi_{n+1}| \leq N_{n+1}^{\sigma/2}$ .

Then, by Lemma 4.5, (4.24), we get

**Lemma 4.9.** *One has  $h_{n+1} \in C^1(\mathcal{I}; \mathbf{B}_{\rho_{n+1}})$  and*

$$h_{n+1}(0, \lambda) = 0, \quad \|h_{n+1}\|_{s_1} \leq N_{n+1}^{-\sigma-1}, \quad \|\partial_\lambda h_{n+1}\|_{s_1} \leq N_{n+1}^{-\nu-1}. \tag{4.28}$$

We now conclude the proof of Theorem 4.2. Let

$$u_{n+1} := u_n + h_{n+1}. \tag{4.29}$$

We want to show that (S1)<sub>n+1</sub>-(S4)<sub>n+1</sub> are satisfied. Property (S1)<sub>n+1</sub> follows by (4.13) and (4.28). Moreover Remark 4.6 implies that  $u_{n+1} \in E_{n+1} \cap \widehat{X}_s$ , i.e.  $(\widehat{S1})_{n+1}$  holds. Property (S2)<sub>n+1</sub> is (4.28). Property (S3)<sub>n+1</sub> follows by the definition (4.27) and since  $\tilde{h}_{n+1}(\varepsilon, \lambda)$  solves  $\Pi_{n+1} F(\varepsilon, \lambda, u_n(\varepsilon, \lambda) + h) = 0$ , for all  $\lambda \in A_{n+1}$ . Finally

$$B_{n+1} \leq B_n + \|h_{n+1}\|_S \stackrel{(S4)_n, \text{Lem.4.7}}{\leq} 2N_{n+1}^p + N_{n+1}^{2p} \leq 2N_{n+1}^p$$

and (S4)<sub>n</sub>, Lemma 4.7, (4.24), imply

$$B'_{n+1} \leq B'_n + \|\partial_\lambda h_{n+1}\|_S \leq 2N_{n+1}^q + N_{n+1}^{(\sigma/2)+2p} + N_{n+1}^{2q} \leq 2N_{n+1}^q$$

because  $q = 2p + (\sigma/2)$ . Hence also  $(S4)_{n+1}$  follows.  $\blacksquare$

Note that so far the set  $A_\infty := \bigcap_{n \geq 0} A_n$  where  $u_\varepsilon$  is a solution of  $F(\varepsilon, \lambda, u_\varepsilon(\lambda)) = 0$  (Theorem 4.2) could have zero measure or even be the empty set. The goal of the next section is to show that, under further assumptions on  $F$  (i.e. those of Theorem 2.16), the set  $\mathcal{C}_\varepsilon$  in (2.33) (which is defined in a non inductive way) is contained in  $A_\infty$ .

## 5 Proof of Theorem 2.16

We now specialise the abstract Nash-Moser Theorem 4.2 to the scale of sequence spaces  $X^s = H^s(\mathfrak{K})$  and to operators  $F$  of the form (2.20) satisfying the assumptions of Theorem 2.16. In particular (2.21)-(2.22), (f1)-(f2), imply the assumptions (F0)-(F4) of Theorem 4.2.

In addition to the parameters  $\tau > 0$ ,  $\delta \in (0, 1/4)$ ,  $\sigma$ ,  $s_1 > s_0$ ,  $S < S'$  satisfying (4.3) needed in Theorem 4.2, we now introduce other parameters  $\tau_1$ ,  $\chi_0$ ,  $\tau_0$ ,  $C_1$  and add the following constraints

$$s_1 > s_0 + \nu_0, \quad S < S' - \nu_0, \quad \tau > \tau_0, \quad \tau_1 > 2\chi_0 d, \quad \tau > 2\tau_1 + d + r + 1, \quad C_1 \geq 2, \quad (5.1)$$

then, setting  $\kappa := \tau + d + r + s_0$ ,  $s_2 := s_1 - \nu_0$

$$\chi_0(\tau - 2\tau_1 - d - r) > 3(\kappa + (s_0 + d + r)C_1), \quad \chi_0\delta > C_1, \quad (5.2a)$$

$$s_2 > 3\kappa + 2\chi_0(\tau_1 + d + r) + C_1 s_0, \quad 2\delta s_1 > \nu_0. \quad (5.2b)$$

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

**Remark 5.1.** *In the applications, the constants  $\tau_0$ ,  $\tau_1$  have to be taken large enough, in order to verify condition (2.37). Nevertheless, all the constraints (4.3), (5.1), (5.2) may be verified.*

Given  $\Omega, \Omega' \subset \mathfrak{K}$ , we define

$$\text{diam}(\Omega) := \sup_{k, k' \in \Omega} \text{dist}(k, k'), \quad \text{dist}(\Omega, \Omega') := \inf_{k \in \Omega, k' \in \Omega'} \text{dist}(k, k'),$$

where  $\text{dist}(\cdot, \cdot)$  is defined in (2.7).

**Definition 5.2.** ( *$N$ -good/ $N$ -bad matrices*). *Let  $F \subset \mathfrak{K}$  be such that  $\text{diam}(F) \leq 4N$  for some  $N \in \mathbb{N}$ . We say that a matrix  $A \in \mathcal{M}_F^E$  is  $N$ -good if  $A$  is invertible and for all  $s \in [s_0, s_2]$  one has*

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

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

**Definition 5.3.** ( *$(A, N)$ -regular, good, bad sites*). *For any finite  $E \subset \mathfrak{K}$ , let  $A = D + \varepsilon T \in \mathcal{M}_E^E$  with  $D := \text{diag}(D_k \mathbf{1}_{d_j})$ ,  $D_k \in \mathbb{C}$ . An index  $k \in E$  is*

- *$(A, N)$ -regular if there exists  $F \subseteq E$  such that  $\text{diam}(F) \leq 4N$ ,  $\text{dist}(\{k\}, E \setminus F) \geq N$  and the matrix  $A_F^E$  is  $N$ -good.*
- *$(A, N)$ -good if either it is regular for  $D$  (Definition 2.13) or it is  $(A, N)$ -regular. Otherwise  $k$  is  $(A, N)$ -bad.*

The above definition could be extended to infinite  $E$ .

**Definition 5.4. ( $N$ -good/ $N$ -bad parameters).** For  $L$  as in (2.25), we denote

$$B_N(j_0, \varepsilon, \lambda) := \left\{ \theta \in \mathbb{R} : L_{N, j_0}(\varepsilon, \lambda, \theta, u) \text{ is } N\text{-bad} \right\}. \quad (5.3)$$

A parameter  $\lambda \in \mathcal{I}$  is  $N$ -good for  $L$  if for any  $j_0 \in \Lambda_+$  one has

$$B_N(j_0, \varepsilon, \lambda) \subseteq \bigcup_{q=1}^{N^\varepsilon} I_q, \quad I_q \text{ intervals with } \text{meas}(I_q) \leq N^{-\tau_1} \quad (5.4)$$

where  $\varepsilon$  is the parameter introduced in Theorem 2.16. Otherwise we say that  $\lambda$  is  $N$ -bad. We denote the set of  $N$ -good parameters as

$$\mathcal{G}_N = \mathcal{G}_N(u) := \left\{ \lambda \in \mathcal{I} : \lambda \text{ is } N\text{-good for } L \right\}. \quad (5.5)$$

Note that the above definition deals only with finite dimensional truncations of  $L$ .

The following assumption is needed for the multiscale Proposition 5.8.

**Hypothesis 4. (Separation of bad sites)** There exist  $C_1 = C_1(\mathfrak{K}) > 2$ ,  $\hat{N} = \hat{N}(\mathfrak{K}, \tau_0) \in \mathbb{N}$  and  $\hat{\mathcal{I}} \subseteq \bar{\mathcal{I}}$  (see (2.30)) such that, for all  $N \geq \hat{N}$ , and  $\|u\|_{s_1} < 1$  (with  $s_1$  satisfying (5.2b)), 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  $j_0 \in \Lambda_+$  the  $(L, N)$ -bad sites  $k = (l, j, \mathbf{a}) \in \mathfrak{K}$  of  $L = L_{N \times, j_0}(\varepsilon, \lambda, \theta, u)$  admit a partition  $\cup_\alpha \Omega_\alpha$  in disjoint clusters satisfying

$$\text{diam}(\Omega_\alpha) \leq N^{C_1}, \quad \text{dist}(\Omega_\alpha, \Omega_\beta) \geq N^2, \quad \text{for all } \alpha \neq \beta. \quad (5.6)$$

For  $N > 0$ , we denote

$$\mathcal{G}_N^0(u) := \left\{ \lambda \in \mathcal{I} : \forall j_0 \in \Lambda_+ \text{ there is a covering} \right. \\ \left. B_N^0(j_0, \varepsilon, \lambda) \subset \bigcup_{q=1}^{N^\varepsilon} I_q, \quad I_q = I_q(j_0) \text{ intervals with } \text{meas}(I_q) \leq N^{-\tau_1} \right\} \quad (5.7)$$

where

$$B_N^0(j_0, \varepsilon, \lambda) := B_N^0(j_0, \varepsilon, \lambda, u) := \left\{ \theta \in \mathbb{R} : \|L_{N, j_0}^{-1}(\varepsilon, \lambda, \theta, u)\|_0 > N^{\tau_1} \right\}. \quad (5.8)$$

We also set

$$\mathfrak{G}_N(u) := \left\{ \lambda \in \mathcal{I} : \|L_N^{-1}(\varepsilon, \lambda, u)\|_0 \leq N^{\tau_1} \right\}. \quad (5.9)$$

Under the smallness condition (4.6), Theorem 4.2 applies, thus defining the sequence  $u_n$  and the sets  $A_n$ . We now introduce the sets

$$\mathcal{C}_0 := \hat{\mathcal{I}}, \quad \mathcal{C}_n := \bigcap_{i=1}^n \mathcal{G}_{N_i}^0(u_{i-1}) \bigcap_{i=1}^n \mathfrak{G}_{N_i}(u_{i-1}) \cap \hat{\mathcal{I}} \quad (5.10)$$

where  $\hat{\mathcal{I}}$  is defined in Hypothesis 4,  $\mathfrak{G}_N(u)$  in (5.9), and  $\mathcal{G}_N^0(u)$  in (5.7).



**Theorem 5.5.** *Assume that  $F$  in (2.20) satisfies (2.21)-(2.22), (f1)-(f2) and Hypotheses 1, 2 and 4. Assume that the parameters satisfy (4.3), (5.1), (5.2). Then there exists  $\bar{N}_0 \in \mathbb{N}$ , such that, for all  $N_0 \geq \bar{N}_0$  and  $\varepsilon \in [0, \varepsilon_0)$  with  $\varepsilon_0$  satisfying (4.6), the following inclusions hold:*

$$(S5)_0 \quad \|u\|_{s_1} \leq 1 \quad \Rightarrow \quad \mathcal{G}_{N_0}(u) = \mathcal{I}$$

$$(S6)_0 \quad \mathcal{C}_0 \subseteq A_0,$$

and for all  $n \geq 1$  (recall the definitions of  $A_n$  in (4.7))

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

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

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

## 5.1 Initialisation

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

**Lemma 5.6.** *For all  $\|u\|_{s_1} \leq 1$ ,  $N \leq N_0$ , the set  $\mathcal{G}_N(u) = \mathcal{I}$ .*

*Proof.* We claim that, for any  $\lambda \in \mathcal{I}$  and any  $j_0 \in \Lambda_+$ , if

$$|D_k(\lambda, \theta)| > N^{-\tau_1}, \quad \forall k = (l, j, \mathbf{a}) \in \mathfrak{K} \text{ with } |(l, j - j_0)| \leq N, \quad (5.11)$$

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

$$B_N(j_0, \varepsilon, \lambda) \subset \bigcup_{|(l, j - j_0)| \leq N} \{\theta \in \mathbb{R} : |D_k(\lambda, \theta)| \leq N^{-\tau_1}\},$$

which in turn, by Hypothesis 2, implies the thesis, see (5.4), (5.5), for some  $\varepsilon \geq d + r + 1$ . The above claim follows by a perturbative argument. Indeed for  $\|u\|_{s_1} \leq 1$ ,  $s_1 = s_2 + \nu_0$ , we use (2.24c) to obtain

$$\varepsilon |(D_{N, j_0}^{-1}(\lambda, \theta))|_{s_2} |T_{N, j_0}(u)|_{s_2} \leq \varepsilon C(s_1) |D_{N, j_0}^{-1}(\lambda, \theta)|_{s_2} (1 + \|u\|_{s_2 + \nu_0}) \stackrel{(5.11)}{\leq} \varepsilon N^{\tau_1} C(s_1) \stackrel{(4.6)}{\leq} \frac{1}{2}.$$

Then we invert  $L_{N, j_0}$  by Neumann series and Lemma 2.12 implies

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

by (5.1), which proves the claim.  $\blacksquare$

**Lemma 5.7.** *Property  $(S6)_0$  holds.*

*Proof.* Since  $\hat{\mathcal{I}} \subset \bar{\mathcal{I}}$  it is sufficient to prove that  $\bar{\mathcal{I}} \subset A_0$ . By the definition of  $A_0$  in (4.7), (4.5), (4.4), we have to prove that

$$\lambda \in \bar{\mathcal{I}} \quad \Longrightarrow \quad \|L_{N_0}^{-1}(\varepsilon, \lambda, 0)[h]\|_s \leq C(s) N_0^{\tau + \delta s_1} \|h\|_s, \quad \forall s \in [s_1, S]. \quad (5.12)$$

Indeed, if  $\lambda \in \bar{\mathcal{I}}$  then  $|D_k(\lambda)| \geq N_0^{-\tau_0}$ , for all  $|k| < N_0$ , and so  $|D(\lambda)^{-1}|_s \leq N_0^{\tau_0}$ ,  $\forall s$ . Since  $\varepsilon |D(\lambda)^{-1} T(0)|_{s_1} \leq \varepsilon N_0^{\tau_0} |T(0)|_{s_1} < 1/2$  for  $\varepsilon$  small enough, Lemma 2.12 implies

$$|L_{N_0}^{-1}(\varepsilon, \lambda, 0)|_{s_1} \leq 2N_0^{\tau_0}, \quad |L_{N_0}^{-1}(\varepsilon, \lambda, 0)|_s \leq C(s) N_0^{\tau_0} (1 + \varepsilon N_0^{\tau_0} |T(0)|_s), \quad \forall s > s_1,$$

and, by (5.1), (2.24c) and Lemma 2.7, the estimate (5.12) follows.  $\blacksquare$

## 5.2 Inductive step

By the Nash-Moser Theorem 4.2 we know that (S1)<sub>n</sub>–(S4)<sub>n</sub> hold for all  $n \geq 0$ . Assume inductively that (S5)<sub>i</sub> and (S6)<sub>i</sub> hold for all  $i \leq n$ . In order to prove (S5)<sub>n+1</sub>, we need the following *multiscale Proposition* 5.8 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 5.8. (Multiscale)** *Assume (5.1), (5.2). 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  $N \geq N_0$ ,  $|\varepsilon| < \varepsilon_0$ ,  $\chi \in [\chi_0, 2\chi_0]$ ,  $E \subset \mathfrak{K}$  with  $\text{diam}(E) \leq 4N^\chi$ , if the matrix  $A = D + \varepsilon T \in \mathcal{M}_E^E$  satisfies*

$$(H1) \quad |T|_{s_2} \leq \Upsilon,$$

$$(H2) \quad \|A^{-1}\|_0 \leq N^{\chi\tau_1},$$

(H3) *there is a partition  $\{\Omega_\alpha\}_\alpha$  of the  $(A, N)$ -bad sites (Definition 5.3) such that*

$$\text{diam}(\Omega_\alpha) \leq N^{C_1}, \quad \text{dist}(\Omega_\alpha, \Omega_\beta) \geq N^2, \quad \text{for } \alpha \neq \beta,$$

*then the matrix  $A$  is  $N^\chi$ -good and*

$$|A^{-1}|_s \leq \frac{1}{4} N^{\chi\tau} \left( N^{\chi\delta s} + \varepsilon |T|_s \right), \quad \forall s \in [s_0, \bar{s}]. \quad (5.13)$$

Note that the bound (5.13) is much more than requiring that the matrix  $A$  is  $N^\chi$ -good, since it holds also for  $s > s_2$ .

This Proposition is proved by “resolvent type arguments” and it coincides essentially with [4]-Proposition 4.1. The correspondences in the notations of this paper and [4] 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 [4]. Our conditions (5.1), (5.2) imply conditions (4.4) and (4.5) of [4] for all  $\chi \in [\chi_0, 2\chi_0]$  and our (H1) implies the corresponding Hypothesis (H1) of [4] with  $\Upsilon \rightsquigarrow 2\Upsilon$ . The other hypotheses are the same. Although the  $s$ -norm in this paper is different, the proof of [4]-Proposition 4.1 relies only on abstract algebra and interpolation properties of the  $s$ -norm (which indeed hold also in this case – see section 2.1). Hence it can be repeated verbatim and we report it in the Appendix for completeness.

Now, we distinguish two cases:

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

$$N_{n+1} = \bar{N}^\chi, \quad \bar{N} := [N_{n+1}^{1/\chi_0}] \in (N_0^{1/\chi}, N_0). \quad (5.14)$$

This case may occur only in the first steps.

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

$$N_{n+1} = N_p^\chi, \quad \chi = 2^{n+1-p} \in [\chi_0, 2\chi_0]. \quad (5.15)$$

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

**Lemma 5.9.** *Property (S5)<sub>1</sub> holds.*

*Proof.* We have to prove that  $\mathcal{G}_{N_1}^0(u_0) \cap \hat{\mathcal{I}} \subseteq \mathcal{G}_{N_1}(u) \cap \hat{\mathcal{I}}$  where  $\|u - u_0\|_{s_1} \leq N_1^{-\sigma}$ . By Definition 5.4 and (5.7) it is sufficient to prove that, for all  $j_0 \in \Lambda_+$ ,

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

where we stress the dependence on  $u, u_0$  in (5.3), (5.8). By the definitions (5.8), (5.3) this amounts to prove that

$$\|L_{N_1, j_0}^{-1}(\varepsilon, \lambda, \theta, u_0)\|_0 \leq N_1^{\tau_1} \implies L_{N_1, j_0}(\varepsilon, \lambda, \theta, u) \text{ is } N_1\text{-good.} \quad (5.16)$$

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

$$|L_{N_1, j_0}^{-1}(\varepsilon, \lambda, \theta, u_0)|_s \leq \frac{1}{4} N_1^{\tau_1} \left( N_1^{\delta s} + \varepsilon |T(u_0)|_s \right), \quad \forall s \in [s_0, S]. \quad (5.17)$$

Indeed we may apply Proposition 5.8 to the matrix  $A = L_{N_1, j_0}(\varepsilon, \lambda, \theta, u_0)$  with  $\bar{s} = S$ ,  $N = \bar{N}$ ,  $N_1 = \bar{N}^X$  and  $E = \{|l| \leq N_1, |j - j_0| \leq N_1\} \times \mathfrak{A}$ . Hypothesis (H1) with  $\Upsilon = 3C(s_1)$  follows by (2.24c) and  $\|u_0\|_{s_1} \leq 1$ . Moreover (H2) is  $\|L_{N_1, j_0}^{-1}(\varepsilon, \lambda, \theta, u_0)\|_0 \leq N_1^{\tau_1}$ . Finally (H3) is implied by Hypothesis 4 provided we take  $N_0^{1/\chi_0} > \hat{N}(\mathfrak{K}, \tau_0)$  (recall (5.14)) and noting that  $\lambda \in \mathcal{G}_{\bar{N}}(u_0) \cap \hat{\mathcal{I}}$  by Lemma 5.6 (since  $\bar{N} \leq N_0$  then  $\mathcal{G}_{\bar{N}}(u_0) = \mathcal{I}$ ). Hence (5.13) implies (5.17).

We now prove (5.16) by a perturbative argument. Since  $\|u - u_0\|_{s_1} \leq N_1^{-\sigma}$  (recall that  $\|u_0\|_{s_1} \leq 1$  so  $\|u\|_{s_1} \leq 2$ ) then, for  $\nu_1 := \max(\nu, \nu_0)$ ,

$$\begin{aligned} |L_{N_1, j_0}(\varepsilon, \lambda, \theta, u_0) - L_{N_1, j_0}(\varepsilon, \lambda, \theta, u)|_{s_2} &\leq |L_{N_1, j_0}(\varepsilon, \lambda, \theta, u_0) - L_{N_1, j_0}(\varepsilon, \lambda, \theta, u)|_{s_1 - \nu_1} \\ &\stackrel{(2.21), (2.24d)}{\leq} C \|u - u_0\|_{s_1} \leq C N_1^{-\sigma} < 1/2. \end{aligned}$$

By Neumann series (see Lemma 2.12) and (5.17) one has  $|L_{N_1, j_0}^{-1}(\varepsilon, \lambda, \theta, u)|_s \leq N_1^{\tau_1 + \delta s}$  for all  $s \in [s_0, s_2]$ , namely  $L_{N_1, j_0}(\varepsilon, \lambda, \theta, u)$  is  $N_1$ -good.  $\blacksquare$

**Lemma 5.10.** *Property (S6)<sub>1</sub> holds.*

*Proof.* Let  $\lambda \in \mathcal{C}_1 := \mathcal{G}_{N_1}^0(u_0) \cap \mathfrak{G}_{N_1}(u_0) \cap \hat{\mathcal{I}}$ , see (5.10). By the definitions (4.7), (4.5), and (S6)<sub>0</sub>, in order to prove that  $\lambda \in A_1$ , it is sufficient to prove that  $(\lambda, u_0(\lambda)) \in J_{\tau, \delta}^{(N_1)}$ . Since  $\lambda \in \mathfrak{G}_{N_1}(u_0)$  the matrix  $\|L_{N_1}^{-1}(\varepsilon, \lambda, u_0)\|_0 \leq N_1^{\tau_1}$  (see (5.9)) and so (5.17) holds with  $j_0 = 0, \theta = 0$ . Hence, using (2.24c) and  $u_0 \in E_0$ , we have

$$|L_{N_1}^{-1}(\varepsilon, \lambda, u_0)|_s \leq \frac{1}{4} N_1^{\tau_1} (N_1^{\delta s} + \varepsilon C(s)(1 + N_1^{\nu_0/2} \|u_0\|_s)), \quad \forall s \in [s_0, S],$$

that, by Lemma 2.7,  $\delta s_1 > \nu_0/2$ , (P1), implies,  $(S1)_n, \forall s \in [s_1, S]$ ,

$$\|L_{N_1}^{-1}(\varepsilon, \lambda, u_0)[h]\|_s \leq C(s) N_1^{\tau_1 + \delta s_1} (\|h\|_s + N_1^{\delta(s-s_1)} \|u_0\|_s \|h\|_{s_1}), \quad \forall h \in E_1,$$

which is the inequality in (4.4) with  $N = N_1, u = u_0$ . Hence  $(\lambda, u_0(\lambda)) \in J_{\tau, \delta}^{(N_1)}$ .  $\blacksquare$

Now we consider **case 2**.

**Lemma 5.11.**  $\bigcap_{i=1}^{n+1} \mathcal{G}_{N_i}^0(u_{i-1}) \cap \hat{\mathcal{I}} \subseteq \mathcal{G}_{N_p}(u_n) \cap \hat{\mathcal{I}}$ .

*Proof.* By  $(S2)_n$  of Theorem 4.2 we get  $\|u_n - u_{p-1}\|_{s_1} \leq \sum_{i=p}^n \|u_i - u_{i-1}\|_{s_1} \leq \sum_{i=p}^n N_i^{-\sigma-1} \leq N_p^{-\sigma} \sum_{i=p}^n N_i^{-1} \leq N_p^{-\sigma}$ . Hence  $(S5)_p$  ( $p \leq n$ ) implies

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

proving the lemma. ■

**Lemma 5.12.** *Property  $(S5)_{n+1}$  holds.*

*Proof.* Fix  $\lambda \in \bigcap_{i=1}^{n+1} \mathcal{G}_{N_i}^0(u_{i-1}) \cap \hat{\mathcal{I}}$ . Reasoning as in the proof of Lemma 5.9, it is sufficient to prove that, for all  $j_0 \in \Lambda_+$ ,  $\|u - u_n\|_{s_1} \leq N_{n+1}^{-\sigma}$ , one has

$$\|L_{N_{n+1}, j_0}^{-1}(\varepsilon, \lambda, \theta, u_n)\|_0 \leq N_{n+1}^{\tau_1} \implies L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u) \text{ is } N_{n+1}\text{-good.} \quad (5.18)$$

We apply the multiscale Proposition 5.8 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 (5.15). Assumption (H1) holds and (H2) is  $\|L_{N_{n+1}, j_0}^{-1}(\varepsilon, \lambda, \theta, u_n)\|_0 \leq N_{n+1}^{\tau_1}$ . Lemma 5.11 implies that  $\lambda \in \mathcal{G}_{N_p}(u_n) \cap \hat{\mathcal{I}}$  and therefore also (H3) is satisfied since we are assuming Hypothesis 4. But then Proposition 5.8 implies

$$|L_{N_{n+1}, j_0}^{-1}(\varepsilon, \lambda, \theta, u_n)|_s \leq \frac{1}{4} N_{n+1}^\tau \left( N_{n+1}^{\delta s} + \varepsilon |T(u_n)|_s \right), \quad \forall s \in [s_0, S]. \quad (5.19)$$

Finally, for  $\|u - u_n\|_{s_1} \leq N_{n+1}^{-\sigma}$  (recall that  $\|u_n\|_{s_1} \leq 1$  so  $\|u\|_{s_1} \leq 2$ ) one has

$$\begin{aligned} |L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u_n) - L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u)|_{s_2} &= |L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u_n) - L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u)|_{s_1 - \nu_1} \\ &\leq C \|u - u_n\|_{s_1} \leq C N_{n+1}^{-\sigma}, \end{aligned}$$

where the second bound follows by (2.21) and (2.24d) with  $\nu_1 = \max(\nu, \nu_0)$ . Hence (5.19) and Lemma 2.12 imply  $|L_{N_{n+1}, j_0}(\varepsilon, \lambda, \theta, u)^{-1}|_s \leq N_{n+1}^{\tau + \delta s}$  for all  $s \in [s_0, s_2]$ , proving (5.18). ■

**Lemma 5.13.** *Property  $(S6)_{n+1}$  holds.*

*Proof.* Follow word by word the proof of Lemma 5.10 with  $N_{n+1}$  instead of  $N_1$ , and  $u_n$  instead of  $u_0$ . Since  $\lambda \in \mathfrak{G}_{N_{n+1}}(u_n)$  (see (5.9)) the bound (5.19) holds with  $j_0 = 0$ ,  $\theta = 0$ , and so

$$|L_{N_{n+1}}^{-1}(\varepsilon, \lambda, u_n)|_s \leq \frac{1}{4} N_{n+1}^\tau \left( N_{n+1}^{\delta s} + \varepsilon |T(u_n)|_s \right), \quad \forall s \in [s_0, S]. \quad (5.20)$$

■

### 5.3 Separation properties

In order to complete the proof of Theorem 2.16 we show that Hypothesis 3 implies Hypothesis 4.

**Proposition 5.14.** *Hypothesis 3 implies Hypothesis 4 with  $C_1 = (\mathbf{e} + d + r + 3)\mathbf{s} + 3$ ,  $\hat{\mathcal{I}} := \tilde{\mathcal{I}} \cap \bar{\mathcal{I}}$  and  $\hat{N}$  large enough.*

We split the proof of Proposition 5.14 in several Lemmas. For  $\|u\|_{s_1} < 1$ , we consider  $L := L(\lambda, \theta, u) = D(\lambda, \theta) + \varepsilon T(u)$  defined in (2.25).

**Definition 5.15.** *A site  $k = (i, \mathbf{a}) \in \mathfrak{K}$  is*

- *$(L, N)$ -strongly-regular if  $L_{N,i}$  is  $N$ -good,*
- *$(L, N)$ -weakly-singular if, otherwise,  $L_{N,i}$  is  $N$ -bad,*
- *$(L, N)$ -strongly-good if either it is regular for  $D = D(\lambda, \theta)$  (recall Definition 2.13) or all the sites  $k' = (i', \mathbf{a}')$  with  $\text{dist}(k, k') \leq N$  are  $(L, N)$ -strongly-regular. Otherwise  $k$  is  $(L, N)$ -weakly-bad.*

The above definition differs from that of  $(L, N)$ -good matrix (Definition 5.3) in the following way. Here we do not introduce a finite subset  $E$  but study the infinite dimensional matrix  $L$ , and require invertibility conditions on the  $N$ -dimensional submatrices centered at a strongly-regular point  $k$  (with respect to [4] we use a different notation, see Definition 5.1-[4]).

**Lemma 5.16.** *For any  $j_0 \in \Lambda_+$ ,  $\chi \in [\chi_0, 2\chi_0]$  consider  $k = (l, j, \mathbf{a})$  such that  $|l|, |j - j_0| \leq N^\chi$ , if  $k$  is  $(L, N)$ -strongly-good then  $k$  is  $(L_{N^\chi, j_0}, N)$ -good.*

*Proof.* Set  $N' = N^\chi$  and (recall the definition (2.2))

$$E = \mathfrak{I} \times \mathfrak{S} \times \mathfrak{A}, \quad \mathfrak{I} := [-N', N']^d \cap \mathbb{Z}^d, \quad \mathfrak{S} := \left( j_0 + \left\{ \sum_{p=1}^r \alpha_p \mathbf{w}_p : \alpha_p \in [-N', N'] \right\} \right) \cap \Lambda_+.$$

If  $k \in E$  is regular then it is  $(L, N)$ -good. If  $k = (l, j, \mathbf{a}) \in E$  is singular but  $(L, N)$ -strongly-regular, we define the neighborhood  $F_N = F_N(k)$  as

$$F_N := \mathfrak{I}_N \times \mathfrak{S}_N \times \mathfrak{A}, \quad \mathfrak{I}_N := \left( \prod_{q=1}^d I_q \right) \cap \mathbb{Z}^d, \quad \mathfrak{S}_N := \left\{ \sum_{i=1}^r \beta_p \mathbf{w}_p : \beta_p \in J_p \right\} \cap \Lambda_+,$$

where the intervals  $J_p \in \mathbb{R}$  are defined as follows (we set  $a_p := (j_0)_p - N'$ ,  $b_p := (j_0)_p + N'$ ):

$$\begin{aligned} j_p - a_p > N, \quad b_p - j_p > N &\quad \Rightarrow \quad J_p := [j_p - N, j_p + N], \\ j_p - a_p \leq N, \quad b_p - j_p > N &\quad \Rightarrow \quad J_p := [a_p, a_p + 2N] \\ j_p - a_p > N, \quad b_p - l_p \leq N &\quad \Rightarrow \quad J_p := [b_p - 2N, b_p], \end{aligned}$$

same for  $I_q$ . By construction  $\text{dist}(k, E \setminus F_N) > N$  and  $\text{diam}(F_N) \leq 2N < 4N$ . Moreover, by (2.4) there exists  $\bar{k} \in E$  with  $\text{dist}(\bar{k}, k) < N$  such that

$$F_N = \left( (\bar{l} + [-N, N]^d) \times \left( \bar{j} + \left\{ \sum_{p=1}^r \alpha_p \mathbf{w}_p : \alpha_p \in [-N, N] \right\} \right) \times \mathfrak{A} \right) \cap \mathfrak{K}.$$

Then, since  $k$  is  $(L, N)$ -strongly-regular, the  $|(L_{F_N}^{F_N})^{-1}|_s = |(L_{N, \bar{k}})^{-1}|_s \leq N^{\tau + \delta s}$  proving the lemma. ■

**Remark 5.17.** The assumption that  $\Lambda_+$  has the “product structure” (2.4) has been used only in Lemma 5.16 above.

The Diophantine condition (1.5) implies (since  $\lambda \geq 1/2$ ) that

$$|\omega \cdot l| = \lambda |\bar{\omega} \cdot l| \geq \gamma_0 |l|^{-d}, \quad \forall l \in \mathbb{Z}^d \setminus \{0\}. \quad (5.21)$$

**Lemma 5.18.** Let  $\lambda$  be  $N$ -good for  $L$ , see Definition 5.4. Then, for any  $j_0 \in \Lambda_+$ ,  $\chi \in [\chi_0, 2\chi_0]$ , the cardinality

$$\#\{k = (l, j_0, \mathbf{a}) \in \mathfrak{K} : k \text{ is } (L, N)\text{-weakly-singular, } |l| \leq 2N^\chi\} \leq |\mathfrak{A}|N^\epsilon. \quad (5.22)$$

*Proof.* By Definition 5.15, if  $(l, j_0, \mathbf{a})$  is  $(L, N)$ -weakly-singular, then  $L_{N, l, j_0}(\varepsilon, \lambda, \theta)$  is  $N$ -bad. By (2.29) this means that  $L_{N, j_0}(\varepsilon, \lambda, \theta + \lambda \bar{\omega} \cdot l)$  is  $N$ -bad, i.e.  $\theta + \lambda \bar{\omega} \cdot l \in B_N(j_0, \varepsilon, \lambda)$ , see (5.3). By assumption  $\lambda$  is  $N$ -good for  $L$  and hence (5.4) holds. We claim that in each interval  $I_q$  there is at most one element  $\theta + \omega \cdot l'$  with  $\omega = \lambda \bar{\omega}$  and  $|l'| \leq 2N^\chi$ . This, of course, imply (5.22). Indeed, if there are  $l' \neq l''$  with  $|l'|, |l''| \leq 2N^\chi$  such that  $\theta + \omega \cdot l', \theta + \omega \cdot l'' \in I_q$ , then

$$|\omega \cdot (l' - l'')| = |(\omega \cdot l' + \theta) - (\omega \cdot l'' + \theta)| \leq |I_q| \leq N^{-\tau_1}. \quad (5.23)$$

On the other hand (5.21) implies

$$|\omega \cdot (l' - l'')| \geq \frac{\gamma_0}{|l' - l''|^d} \geq \frac{\gamma_0}{(4N^\chi)^d} = 4^{-d} \gamma_0 N^{-\chi d}. \quad (5.24)$$

Clearly (5.23) and (5.24) are in contradiction for  $N \geq N_0$  large, because  $\tau_1 > 2\chi_0 d$ , see (5.1).  $\blacksquare$

**Corollary 5.19.** Let  $\lambda$  be  $N$ -good for  $L$ . Then, for all  $j_0 \in \Lambda_+$ , the number of  $(L, N)$ -weakly-bad sites  $(l, j_0, \mathbf{a}) \in \mathfrak{K}$  with  $|l| \leq N^\chi$  is bounded from above by  $N^{\epsilon+r+d+1}$ . Hence the  $(L, N)$ -weakly-bad sites are a set  $\Sigma_K$  as in Definition 2.14 with  $K = N^{\epsilon+d+r+1}$ .

*Proof.* By Lemma 5.18 above, the set of  $(L, N)$ -weakly-singular sites  $(l, j, \mathbf{a})$  with  $|l| \leq N^\chi + N$ ,  $|j - j_0| \leq N$  has cardinality at most  $CN^\epsilon \times N^r$ . Each  $(L, N)$ -weakly-bad site  $(l, j_0, \mathbf{a})$  with  $|l| \leq N^\chi$  is included in some  $N$ -ball centered at an  $(L, N)$ -weakly-singular site and each of these balls contains at most  $CN^d$  sites with  $j = j_0$ . Therefore there are at most  $CN^{\epsilon+r} \times N^d$  of such bad sites.  $\blacksquare$

**Definition 5.20.** Given two sites  $k, k' \in \mathfrak{K}$  we say that  $k \cong k'$  if there exists a  $\Gamma$ -chain  $\{k_q\}_{q=0}^\ell$  (Definition 2.15) of  $(L, N)$ -weakly-bad sites connecting  $k$  to  $k'$ , namely  $k_0 = k$  and  $k_\ell = k'$ .

*Proof of Proposition 5.14 completed.* Let  $\lambda \in \mathcal{G}_N \cap \tilde{\mathcal{I}}$ , see Definition 5.4 and recall that  $\tilde{\mathcal{I}}$  is introduced in Hypothesis 3. Set  $\Gamma = N^2$ . A  $N^2$ -chain of  $(L, N)$ -weakly-bad sites is formed by sites which are singular for  $D(\lambda, \theta)$ , see Definition 5.15. Corollary 5.19 implies that  $\#\Sigma_K^{(\tilde{j})} \leq K = N^{\epsilon+d+r+1}$ ,  $\forall \tilde{j} \in \Lambda_+$ , so that Hypothesis 3 (for  $\lambda \in \tilde{\mathcal{I}}(N_0)$  and since  $\Gamma K = N^{\epsilon+d+r+3} > N_0$ ) implies

$$\ell \leq (N^2 N^{\epsilon+d+r+1})^s = N^{(\epsilon+d+r+3)s}. \quad (5.25)$$

The equivalence relation introduced in Definition 5.20 induces a partition of the  $(L, N)$ -weakly-bad sites in disjoint equivalence classes  $\Omega_\alpha$  with (recall  $\Gamma = N^2$ )

$$\text{dist}(\Omega_\alpha, \Omega_\beta) > N^2, \quad \text{diam}(\Omega_\alpha) \leq N^2 \ell \stackrel{(5.25)}{\leq} N^{2+(\epsilon+d+r+3)s} \leq N^{C_1},$$

which is (5.6). We have verified Hypothesis 4 with  $C_1 = 3 + (\mathbf{e} + d + r + 3)\mathbf{s}$  and  $\hat{\mathcal{I}} := \tilde{\mathcal{I}} \cap \bar{\mathcal{I}}$ .  $\blacksquare$

The assumptions of Theorem 2.16 imply those of Theorems 4.2 and 5.5, by Proposition 5.14. We fix the parameters to satisfy (4.3), (5.1), (5.2). Note that, by Proposition 5.14, the constant  $C_1$  is large with  $\mathbf{e}$  and so, by (5.2), the constant  $S'$  has to be large with  $\mathbf{e}$ . Then, Theorem 4.2 implies the existence of a solution  $u_\varepsilon(\lambda)$  of  $F(\varepsilon, \lambda, u_\varepsilon(\lambda)) = 0$  for all  $\lambda \in A_\infty := \bigcap_{n \geq 0} A_n$  and Theorem 5.5 implies that  $C_\infty \subseteq A_\infty$ .

**Lemma 5.21.**  $C_\varepsilon \subseteq C_\infty$  where the set  $C_\varepsilon$  is defined in (2.33).

*Proof.* We claim that, for all  $n \geq 0$ , the sets  $\bar{\mathcal{G}}_{N_n}^0 \subseteq \mathcal{G}_{N_n}^0(u_{n-1})$  and  $\bar{\mathfrak{G}}_{N_n} \subseteq \mathfrak{G}_{N_n}(u_{n-1})$ . These inclusions are a consequence of the super-exponential convergence (4.9) of  $u_n$  to  $u_\varepsilon$ . In view of the definitions (2.35) and (5.7), it is sufficient to prove that  $B_{N_n}^0(j_0, \varepsilon, \lambda, u_{n-1}) \subseteq \bar{B}_{N_n}^0(j_0, \varepsilon, \lambda)$ ,  $\forall j_0$ . Equivalently, if  $\theta \notin \bar{B}_{N_n}^0(j_0, \varepsilon, \lambda)$  then  $\|L_{N_n, j_0}^{-1}(\theta, u_{n-1})\|_0 \leq N_n^{\tau_1}$ , namely  $\theta \notin B_{N_n}^0(j_0, \varepsilon, \lambda, u_{n-1})$  (recall (5.8)). Indeed,  $\|L_{N_n, j_0}^{-1}(\varepsilon, \lambda, \theta, u_\varepsilon)\|_0 \leq N_n^{\tau_1}/2$  by (2.36), and so

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

by Neumann series expansions and using Lemma 2.10, (2.24d), (4.9), and (4.3). The inclusions  $\bar{\mathfrak{G}}_{N_n} \subseteq \mathfrak{G}_{N_n}(u_{n-1})$  follow similarly.  $\blacksquare$

The last conclusion of Theorem 2.16 is proved in the next section.

## 5.4 Regularity

We now consider the case  $S' = +\infty$ . The key estimate is the following upper bound for the divergence of the high Sobolev norm of the approximate solutions  $u_n$ , which extends  $(S4)_n$ . It requires only a small modification of Lemma 4.7.

**Lemma 5.22.** For all  $s \geq S$  one has  $B_n(s) := 1 + \|u_n\|_s \leq C(s)N_n^{2p}$  where  $2p := 2(\tau + \delta s_1) + \nu + 2$ .

*Proof.* Given  $s$ , we take  $n_0(s) \in \mathbb{N}$  such that  $N_{n_0(s)} > N_0(\Upsilon, s)$  where  $N_0(\Upsilon, s)$  is introduced in the multiscale Proposition 5.8. For all  $n \leq n_0(s)$ , the required bound  $B_n(s) \leq C(s)N_n^{2(\tau + \delta s_1) + \nu + 2}$  holds taking  $C(s)$  large enough. We now prove the same bound for  $n > n_0(s)$ . In this case, Proposition 5.8 implies the estimate (5.20) also for  $s > S$ , see (5.13). Then (2.24c) and (P1) imply

$$|L_{N_{n+1}}^{-1}(\varepsilon, \lambda, u_n)|_s \leq C(s)N_{n+1}^\tau (N_{n+1}^{\delta s} + N_{n+1}^{\nu_0/2} \|u_n\|_s). \quad (5.26)$$

Hence, for all  $h \in E_{n+1}$ , using (2.14), (5.26),  $(S1)_n$ ,  $\nu_0/2 < \delta s_1$ ,

$$\|L_{N_{n+1}}^{-1}(\varepsilon, \lambda, u_n)[h]\|_s \leq C'(s)N_{n+1}^{\tau + \delta s_1} (\|h\|_s + (N_{n+1}^{\delta(s-s_1)} + B_n(s))\|h\|_{s_1}). \quad (5.27)$$

Now, as in (4.20), (4.21), we get

$$\begin{aligned} \|r_n + R_n(\tilde{h}_{n+1})\|_{s_1} &\leq C(s)N_{n+1}^{-(s-s_1-\nu)/2} B_n(s) + C(s_1)N_{n+1}^\nu \|\tilde{h}_{n+1}\|_{s_1}^2 \\ &\leq C'(s)N_{n+1}^{-(s-s_1-\nu)/2} B_n(s) \end{aligned} \quad (5.28)$$

using (4.19) and  $N_{n+1}^{\nu+\tau+\delta s_1} N_{n+1}^{-\sigma-1} \leq 1$  (by (4.3)). Since  $\tilde{h}_{n+1}$  defined in Lemma 4.5 is the fixed point of  $\mathcal{H}_{n+1}$  in (4.18), we have, using (5.27), (4.23) (with  $S \rightsquigarrow s$ ), (4.3),  $\delta \in (0, 1/4)$ , (5.28)

$$\|\tilde{h}_{n+1}\|_s \leq C'(s) N_{n+1}^{\tau+\delta s_1+(\nu/2)} B_n(s) (1 + N_{n+1}^{-(s-s_1)/2} B_n(s)) + C'(s) N_{n+1}^{\tau+\delta s_1+\nu} \rho_{n+1} \|\tilde{h}_{n+1}\|_s. \quad (5.29)$$

For all  $n > n_0(s)$  (possibly larger) we have  $C'(s) N_{n+1}^{\tau+\delta s_1+\nu-\sigma-1} < 1/2$  (see (4.3)). Moreover  $N_{n+1}^{-(s-s_1)/2} B_n(s) \leq 2$  (by (P1) and  $(S1)_n$ ) and (5.29) implies

$$\|\tilde{h}_{n+1}\|_s \leq C(s) N_{n+1}^{\tau+\delta s_1+(\nu/2)} B_n(s). \quad (5.30)$$

Therefore (recall (4.29) and (4.27))  $B_{n+1}(s) \leq B_n(s) + \|\tilde{h}_{n+1}\|_s \leq C'(s) N_{n+1}^{\tau+\delta s_1+(\nu/2)} B_n(s)$ , and so the sequence  $B_n(s) N_n^{-2(\tau+\delta s_1)-\nu-2}$  is bounded.  $\blacksquare$

By (4.27), Lemma 5.22 and the estimate (5.30) imply that (use also Lemma 4.7 for  $s_1 < s \leq S$ )

$$\|h_n\|_s \leq \|\tilde{h}_n\|_s \leq C(s) N_n^{2(\tau+\delta s_1)+\nu+1} \leq C(s) N_n^{2p}, \quad \forall s > s_1. \quad (5.31)$$

Now for all  $s > s_1$  let  $s' = 2s - s_1 > s$ . The interpolation inequality (4.1) implies

$$\|h_n\|_s \leq C(s) \|h_n\|_{s_1}^{1/2} \|h_n\|_{s'}^{1/2} \stackrel{(S2)_n, (5.31)}{\leq} C(s) N_n^{-\frac{\sigma+1}{2}+p} \stackrel{(4.3)}{\leq} C'(s) N_n^{-1}$$

which implies  $\|u_\varepsilon\|_s \leq \sum_{n \geq 0} \|h_n\|_s < \infty$ , i.e.  $u_\varepsilon \in H^s$  for all  $s$ .  $\blacksquare$

*Proof of Corollary 2.17.* Since (4.2) holds with  $\widehat{X}_s = \widehat{H}^s(\mathfrak{R})$  by assumption, the solution  $u_\varepsilon \in \widehat{H}^{s_1+\nu}(\mathfrak{R})$  by the last sentence of the Nash-Moser Theorem 4.2.  $\blacksquare$

## A Proof of the multiscale Proposition 5.8

We first prove a lemma about left invertible block diagonal matrices.

**Lemma A.1.** *Let  $\mathcal{D} \in \mathcal{M}_C^B$  be a left invertible block diagonal matrix, with  $B \subseteq C$  i.e.*

$$\mathcal{D}_{k'}^k = \begin{cases} \mathcal{D}_{k'}^k, & \text{if } (k, k') \in \cup_\alpha (\Omega_\alpha \times \Omega'_\alpha), \\ 0 & \text{if } (k, k') \notin \cup_\alpha (\Omega_\alpha \times \Omega'_\alpha), \end{cases} \quad (\text{A.1})$$

where  $\{\Omega_\alpha\}_\alpha$  is a partition of  $B$ , i.e.  $\cup_\alpha \Omega_\alpha = B$ ,  $\Omega_\alpha \cap \Omega_\beta = \emptyset$ ,  $\forall \alpha \neq \beta$ , and the family  $\{\Omega'_\alpha\}_\alpha$  is such that  $\cup_\alpha \Omega'_\alpha \subseteq C$ ,  $\Omega_\alpha \subseteq \Omega'_\alpha$  and  $\Omega'_\alpha \cap \Omega'_\beta = \emptyset$ ,  $\forall \alpha \neq \beta$ . Then  $\mathcal{D}$  has a block diagonal left inverse and, given any left-inverse  $L \in \mathcal{M}_B^C$  of  $\mathcal{D}$ , its restriction

$$(R)_k^{k'} = \begin{cases} L_k^{k'}, & \text{if } (k, k') \in \cup_\alpha (\Omega_\alpha \times \Omega'_\alpha), \\ 0 & \text{if } (k, k') \notin \cup_\alpha (\Omega_\alpha \times \Omega'_\alpha), \end{cases} \quad (\text{A.2})$$

is a left inverse of  $\mathcal{D}$ .



*Proof.*  $\mathcal{D}$  is left invertible and block diagonal, hence each block is left invertible. This produces a block diagonal left inverse of  $\mathcal{D}$ . In order to prove that  $R$  in (A.2) is a left inverse of  $\mathcal{D}$  it is sufficient to show that  $(L - R)\mathcal{D} = 0$ . Indeed, for any  $k \in B$  there is a (unique) index  $\alpha$  such that  $k \in \Omega_\alpha$ , and for any  $k' \in B$  one has

$$((L - R)\mathcal{D})_k^{k'} = \sum_{q \notin \Omega'_\alpha} (L - R)_k^q \mathcal{D}_q^{k'}, \quad (\text{A.3})$$

since, by definition (A.2),  $(L - R)_k^q = 0$  if  $q \in \Omega'_\alpha$ . Now, if  $k' \in \Omega_\alpha$  then  $\mathcal{D}_q^{k'} = 0$  for all  $q \notin \Omega'_\alpha$  (see (A.1)) and hence (A.3) implies  $((L - R)\mathcal{D})_k^{k'} = 0$ . If, otherwise,  $k' \in \Omega_\beta$  for some  $\beta \neq \alpha$ , then, by (A.3) and  $\mathcal{D}_q^{k'} = 0$  for all  $q \notin \Omega'_\beta$ , we have

$$((L - R)\mathcal{D})_k^{k'} = \sum_{q \in \Omega'_\beta} (L - R)_k^q \mathcal{D}_q^{k'} = \sum_{q \in \Omega'_\beta} L_k^q \mathcal{D}_q^{k'} = (LD)_k^{k'} = (\mathbf{1}_B)_k^{k'} = 0$$

where in the second equality we used that  $R_k^q = 0, \forall q \in \Omega'_\beta$  (since  $\Omega'_\beta \cap \Omega'_\alpha = \emptyset$ ), and in the third that  $\mathcal{D}_q^{k'} = 0, \forall q \notin \Omega'_\beta$ .  $\blacksquare$

Call  $G$  the set of the  $(A, N)$ -good sites and  $B$  the set of  $(A, N)$ -bad sites, see Definition 5.3. Let  $\Pi_B, \Pi_G$  be the projectors on the subspaces  $H_B, H_G$  (see (2.6)) and decompose  $u = u_B + u_G$ ,  $u_B := \Pi_B u$ ,  $u_G := \Pi_G u$ .

**Lemma A.2. (Semi-reduction on the good sites).** *There exists  $N_1 = N_1(\Upsilon)$  such that, for  $N \geq N_1$ , there exist  $\mathcal{B} \in \mathcal{M}_G^B$  and  $\mathcal{G} \in \mathcal{M}_G^E$  satisfying (recall that  $\kappa = \tau + d + r + s_0$ )*

$$|\mathcal{G}|_{s_0} \leq cN^\kappa, \quad |\mathcal{B}|_{s_0} \leq c\varepsilon\Upsilon, \quad (\text{A.4})$$

for some  $c = c(s_2)$  and, for all  $s \geq s_0$ ,

$$|\mathcal{G}|_s \leq C(s)N^{2\kappa}(N^{s-s_0} + \varepsilon N^{-d-r}|T|_{s+d+r}), \quad |\mathcal{B}|_s \leq C(s)\varepsilon N^\kappa(N^{s-s_0} + \varepsilon N^{-d-r}|T|_{s+d+r}), \quad (\text{A.5})$$

such that if  $u$  solves  $Au = h$  then

$$u_G = \mathcal{B}u_B + \mathcal{G}h. \quad (\text{A.6})$$

Conversely, if  $u_G = \mathcal{B}u_B + \mathcal{G}h$  then, for all  $k$  regular, one has  $(Au)_k = h_k$ .

*Proof.* We first prove that there exist matrices  $W, R \in \mathcal{M}_G^E$ , satisfying

$$|W|_{s_0} \leq N^\kappa, \quad |R|_{s_0} \leq C(s_2)\varepsilon\Upsilon, \quad (\text{A.7})$$

$$|W|_s \leq C(s)N^{\kappa+s-s_0}, \quad |R|_s \leq \varepsilon C(s)N^\kappa(N^{s-s_0} + N^{-d-r}|T|_{s+d+r}), \quad \forall s \geq s_0, \quad (\text{A.8})$$

such that if  $u$  solves  $Au = h$  then

$$u_G + Ru = Wh. \quad (\text{A.9})$$

Indeed, fix  $k \in G$ . If  $k$  is regular set  $F = \{k\}$ , while, if  $k$  is singular but  $(A, N)$ -regular, let  $F \subset E$  with  $\text{diam}(F) \leq 4N$  be such that  $\text{dist}(k, E \setminus F) \geq N$  and  $A_F^F$  is  $N$ -good. If  $u$  solves  $Au = h$  then  $A_F^F u_F + A_F^{E \setminus F} u_{E \setminus F} = h_F$ , and hence

$$u_F + Qu_{E \setminus F} = (A_F^F)^{-1}h_F, \quad Q := (A_F^F)^{-1}A_F^{E \setminus F} = \varepsilon(A_F^F)^{-1}T_F^{E \setminus F}. \quad (\text{A.10})$$

Lemma 2.6, the fact that  $A_F^E$  is  $N$ -good (see Definition 5.2) and the Hypothesis (H1) imply

$$|Q|_{s_2} \leq C(s_2)\varepsilon\Upsilon N^{\tau+\delta s_2}. \quad (\text{A.11})$$

Moreover, since  $\text{diam}(F) \leq 4N$ , (2.13) and the first inequality in (2.16), we get

$$|Q|_{s+d+r} \leq \varepsilon C(s)N^{(\delta-1)s_0}(\Upsilon N^{s+d+r+\tau} + N^{\tau+s_0}|T|_{s+d+r}). \quad (\text{A.12})$$

Projecting (A.10) onto  $\{k\}$  we obtain  $u_k + \sum_{k' \in E} R_k^{k'} u_{k'} = \sum_{k' \in E} W_k^{k'} h_{k'}$  where

$$R_k^{k'} := \begin{cases} Q_k^{k'}, & \text{if } k' \in E \setminus F, \\ 0 & \text{if } k' \in F, \end{cases} \quad \text{and} \quad W_k^{k'} := \begin{cases} [(A_F^E)^{-1}]_k^{k'}, & \text{if } k' \in F, \\ 0 & \text{if } k' \in E \setminus F \end{cases} \quad (\text{A.13})$$

which is (A.9).

If  $k$  is regular (see Definition 2.13) then  $F = \{k\}$ , and, for  $\varepsilon$  small, one has  $\|(A_k^k)^{-1}\|_0 \leq 2$ . Then the  $k$ -th line of the matrix  $R$  is bounded by  $|R_k|_{s_0+d+r} \leq \varepsilon |(A_k^k)^{-1} T_k|_{s_0+d+r} \leq 2\varepsilon \Upsilon$  by (5.2b).

If  $k$  is singular but  $(A, N)$ -regular then  $R_k^{k'} = 0$  for  $\text{dist}(k', k) \leq N$  and

$$|R_k|_{s_0+d+r} \stackrel{(2.15)}{\leq} N^{-(s_2-s_0-d-r)} |R_k|_{s_2} \leq N^{-(s_2-s_0-d-r)} |Q|_{s_2} \stackrel{(A.11)}{\leq} C(s_2)\Upsilon\varepsilon,$$

using also that  $\tau + s_0 + d + r - (1 - \delta)s_2 < 0$ , see (5.2b) and  $\delta \in (0, 1/4)$ . But then, Lemma 2.9 implies the second inequality in (A.7). The first inequality in (A.7) follows in the same way. The first estimate in (A.8) is a consequence of the first in (A.7) and (2.16). The second estimate in (A.8) follows by

$$|R|_s \stackrel{(2.17)}{\leq} K \sup_{k \in G} |R_k|_{s+d+r} \leq K |Q|_{s+d+r} \stackrel{(A.12)}{\leq} C(s)\varepsilon N^\kappa (N^{s-s_0} + N^{-(d+r)}) |T|_{s+d+r}.$$

We rewrite (A.9) as  $(\mathbb{1}_G + R^G)u_G = Wh - R^B u_B$  where  $R^G$  denotes the restriction  $R^G : G \rightarrow G$ . By (A.7), if  $|\varepsilon| \leq \varepsilon_0(s_2)$  is small enough, then  $|R^G|_{s_0} < 1/2$  and, by Lemma 2.12,

$$|(\mathbb{1}_G + R^G)^{-1}|_{s_0} < 2 \quad (\text{A.14})$$

$$|(\mathbb{1}_G + R^G)^{-1}|_s \leq C(s)(1 + |R^G|_s) \leq C(s)(1 + \varepsilon N^\kappa (N^{s-s_0} + N^{-d-r}) |T|_{s+d+r}), \quad \forall s \geq s_0. \quad (\text{A.15})$$

Then we obtain (A.6) with  $\mathcal{G} := (\mathbb{1}_G + R^G)^{-1}W$ ,  $\mathcal{B} := -(\mathbb{1}_G + R^G)^{-1}R^B$ . The bounds (A.4), (A.5) follow by Lemma 2.6 and (A.14), (A.15), (A.7), (A.8). Finally, (A.6) is equivalent to (A.9) which, for  $k$  regular, gives  $u_k + (A_k^k)^{-1} \sum_{k' \neq k} A_k^{k'} u_{k'} = (A_k^k)^{-1} h_k$ . The final assertion follows.  $\blacksquare$

**Lemma A.3. (Reduction on the bad sites).** *If  $u$  solves  $Au = h$  then*

$$A' u_B = Zh \quad \text{where} \quad A' := A_E^B + A_E^G \mathcal{B} \in \mathcal{M}_E^B, \quad Z := \mathbb{1}_E - A_E^G \mathcal{G} \in \mathcal{M}_E^E \quad (\text{A.16})$$

satisfy

$$|A'|_{s_0} \leq c_0, \quad |A'|_s \leq C(s)N^\kappa (N^{s-s_0} + \varepsilon N^{-d-r} |T|_{s+d+r}), \quad (\text{A.17a})$$

$$|Z|_{s_0} \leq cN^\kappa, \quad |Z|_s \leq C(s)N^{2\kappa} (N^{s-s_0} + \varepsilon N^{-d-r} |T|_{s+d+r}). \quad (\text{A.17b})$$

Moreover  $(A^{-1})_B^E$  is a left inverse of  $A'$ .

*Proof.* If  $u$  solves  $Au = h$  then  $A^G u_G + A^B u_B = h$  (we denote  $A_E^G = A^G$  same for  $A^B$ ) and, by (A.6) we deduce  $(A^G \mathcal{B} + A^B)u_B = h - A^G \mathcal{G}h$ , which is (A.16). By the last assertion of Lemma A.2, for any  $k$  regular and for all  $h$ , we have  $((A^G \mathcal{B} + A^B)u_B)_k = (h - A^G \mathcal{G}h)_k$  identically, namely the lines

$$A'_k = 0 \quad \text{and} \quad Z_k = 0, \quad \text{for all } k \text{ regular.} \quad (\text{A.18})$$

That is, denoting  $R \subset E$  the set of the regular sites in  $E$ , we have  $\Pi_R A' = 0$ . Then (A.17) follow by applying the interpolation estimates (2.13), (A.4), (A.5),  $\kappa > d + r$ , and  $|D_{E \setminus R}|_s \leq 1$  for all  $s$ .

Finally,  $(A^{-1})_B$  is a left inverse of  $A'$  because  $A^{-1}A' = A^{-1}(A_E^B + A_E^G \mathcal{B}) = \mathbb{1}_E^B + \mathbb{1}_E^G \mathcal{B}$  which, in turn, implies  $(A^{-1})_B A' = \mathbb{1}_B$ .  $\blacksquare$

**Lemma A.4. (Left inverse with decay).** *The matrix  $A'$  in (A.16) has a left inverse  ${}^{[-1]}A'$  such that*

$$|{}^{[-1]}A'|_s \leq C(s)N^{2\chi\tau_1 + \kappa + 2(s_0 + d + r)C_1}(N^{C_1 s} + \varepsilon|T|_{s+d+r}), \quad \forall s \geq s_0. \quad (\text{A.19})$$

*Proof.* Let us define the matrix  $\mathcal{D} \in \mathcal{M}_E^B$  as

$$\mathcal{D}_{k'}^k := \begin{cases} (A')_{k'}^k, & \text{if } (k, k') \in \cup_\alpha (\Omega_\alpha \times \Omega'_\alpha), \\ 0 & \text{if } (k, k') \notin \cup_\alpha (\Omega_\alpha \times \Omega'_\alpha), \end{cases}$$

where the family  $\{\Omega_\alpha\}_{\alpha \in \mathbb{Z}}$  is the one in Hypothesis (H3) of Proposition 5.8 and  $\Omega'_\alpha := \{k \in E : \text{dist}(k, \Omega_\alpha) \leq N^2/4\}$ . First of all we prove that  $\mathcal{D}$  admits a left inverse  $\mathcal{W}$  with  $\|\mathcal{W}\|_0 \leq 2N^{\chi\tau_1}$ . Indeed, setting  $\mathcal{R} := A' - \mathcal{D}$ , we have that  $\mathcal{R}_{k'}^k = 0$  for all  $k, k'$  such that  $\text{dist}(k, k') < N^2/4$  and

$$|\mathcal{R}|_{s_0} \stackrel{(2.15)}{\leq} 4^{s_2} N^{-2(s_2 - s_0 - d - r)} |\mathcal{R}|_{s_2 - d - r} \leq 4^{s_2} N^{-2(s_2 - s_0 - d - r)} |A'|_{s_2 - d - r} \leq C(s_2) N^{2\kappa - s_2} \quad (\text{A.20})$$

using (A.17a), (H1). But then, by Lemma 2.10,

$$\|\mathcal{R}\|_0 \|(A^{-1})_B\|_0 \leq |\mathcal{R}|_{s_0} \|A^{-1}\|_0 \stackrel{(\text{A.20}), (\text{H2})}{\leq} C(s_2) N^{2\kappa - s_2 + \chi\tau_1} \stackrel{(5.2b)}{\leq} 1/2,$$

for  $N$  large enough. Now, since  $(A^{-1})_B$  is a left inverse of  $A'$  (Lemma A.3), Lemma 2.12 implies that  $\mathcal{D} = A' - \mathcal{R}$  has a left inverse  $\mathcal{W}$  such that (see (2.19))

$$\|\mathcal{W}\|_0 \leq 2\|(A^{-1})_B\|_0 \leq 2\|A^{-1}\|_0 \leq 2N^{\chi\tau_1}, \quad (\text{A.21})$$

by hypothesis (H2) of Proposition 5.8. Now Lemma A.1 allows to define a block diagonal left inverse of  $\mathcal{D}$ , denoted by  ${}^{[-1]}\mathcal{D}$ , as the restriction of  $\mathcal{W}$  as in (A.2). Since  $\text{diam}(\Omega_\alpha) \leq N^{C_1}$  (Hypothesis (H3)) then  $\text{diam}(\Omega'_\alpha) \leq 2N^{C_1}$  and so  ${}^{[-1]}\mathcal{D}_{k'}^k = 0$  if  $\text{dist}(k, k') > 2N^{C_1}$ . Therefore, for any  $s \geq 0$  one has

$$|{}^{[-1]}\mathcal{D}|_s \stackrel{(2.16)}{\leq} C(s) N^{(s+d+r)C_1} \|{}^{[-1]}\mathcal{D}\|_0 \leq C(s) N^{(s+d+r)C_1 + \chi\tau_1}, \quad (\text{A.22})$$

by (A.21) and Lemma A.1. Finally,  $A' = \mathcal{D} + \mathcal{R}$  and, (A.22), (A.20) imply

$$|{}^{[-1]}\mathcal{D}|_{s_0} |\mathcal{R}|_{s_0} \leq C(s_2) N^{(s_0 + d + r)C_1 + \chi\tau_1 + 2\kappa - s_2} \leq 1/2 \quad (\text{A.23})$$

by (5.2) for  $N$  large enough. But then, Lemma 2.12 implies

$$|{}^{[-1]}A'|_{s_0} \leq 2|{}^{[-1]}\mathcal{D}|_{s_0} \leq C(s_0) N^{(s_0 + d + r)C_1 + \chi\tau_1}, \quad (\text{A.24})$$

and, using (2.18), (A.17a), (A.22), we obtain (A.19). ■

*Proof of Proposition 5.8 completed.* By Lemmas A.2, A.3 and A.4, if  $u$  solves  $Au = h$  then  $u_G = \mathcal{G}h + \mathcal{B}u_B$ ,  $u_B = ({}^{[-1]}A')Zh$ , which in turn implies

$$(A^{-1})_B = ({}^{[-1]}A')Z, \quad (A^{-1})_G = \mathcal{G} + \mathcal{B}({}^{[-1]}A')Z = \mathcal{G} + \mathcal{B}(A^{-1})_B.$$

Then (2.13), (A.24), (A.19), (A.17b), (2.16), (A.4), (A.5) imply, setting  $\zeta := 2\tau_1 + d + r + 2\chi^{-1}(\kappa + C_1(s_0 + d + r))$ , for all  $s \in [s_0, \bar{s}]$ , that

$$|(A^{-1})_B|_s + |(A^{-1})_G|_s \leq C(s)N^{\chi\zeta}(N^{sC_1} + \varepsilon|T|_s) \leq N^{\chi\tau}(N^{\chi\delta s} + \varepsilon|T|_s)/4$$

by (5.2a) and  $N \geq N_0(\bar{s})$  large enough. Thus (5.13) is proved. ■

## References

- [1] Bambusi D., Delort J.M., Grébert B., Szeftel J., *Almost global existence for Hamiltonian semilinear Klein-Gordon equations with small Cauchy data on Zoll manifolds*, Comm. Pure Appl. Math. 60, 11, 1665-1690, 2007.
- [2] Berti M., Biasco L., *Branching of Cantor manifolds of elliptic tori and applications to PDEs*, Comm. Math. Phys, 305, 3, 741-796, 2011.
- [3] Berti M., Bolle P., *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Mathematical Journal, 134, issue 2, 359-419, 2006.
- [4] Berti M., Bolle P., *Quasi-periodic solutions with Sobolev regularity of NLS on  $\mathbb{T}^d$  with a multiplicative potential*, Journal European Math. Society, 15, 229-286, 2013.
- [5] Berti M., Bolle P., *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity, 25, 2579-2613, 2012.
- [6] Berti M., Bolle P., Procesi M., *An abstract Nash-Moser theorem with parameters and applications to PDEs*, Ann. I. H. Poincaré, 1, 377-399, 2010.
- [7] Berti M., Procesi M., *Nonlinear wave and Schrödinger equations on compact Lie groups and Homogeneous spaces*, Duke Math. J., 159, 479-538, 2011.
- [8] Bourgain J., *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices, no. 11, 1994.
- [9] Bourgain J., *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Annals of Math. 148, 363-439, 1998.
- [10] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, 2005.
- [11] Bröcker T., Tom Dieck T., *Representations of Compact Lie Groups*, Grad. Texts in Math. 98 Springer, New York, 1995.
- [12] Chierchia L., You J., *KAM tori for 1D nonlinear wave equations with periodic boundary conditions*, Comm. Math. Phys. 211, 497-525, 2000.

- [13] Craig W., Wayne C. E., *Newton's method and periodic solutions of nonlinear wave equation*, Comm. Pure Appl. Math. 46, 1409-1498, 1993.
- [14] Eliasson L.H., Kuksin S., *KAM for non-linear Schrödinger equation*, Annals of Math., 172, 371-435, 2010.
- [15] Eliasson L. H., Kuksin S., *On reducibility of Schrödinger equations with quasiperiodic in time potentials*, Comm. Math. Phys, 286, 125-135, 2009.
- [16] Geng J., Xu X., You J., *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math. 226, 5361-5402, 2011.
- [17] Helgason S., *Geometric Analysis on Symmetric Spaces*, Math. Surveys Monogr. 39, Amer. Math. Soc. Providence, 1994.
- [18] Helgason S., *Groups and Geometric Analysis: Integral Geometry, Invariant Differential Operators, and Spherical Functions*, corrected reprint of the 1984 original, Math. Surveys Monogr. 83. Amer. Math. Soc., Providence, 2000.
- [19] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional Anal. i Prilozhen., 21, 22-37, 95, 1987.
- [20] Kuksin S., *Analysis of Hamiltonian PDEs*, Oxford Lecture series in Mathematics and its applications 19, Oxford University Press, 2000.
- [21] Kuksin S., Pöschel J., *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. Math. (2) 143 149-79, 1996.
- [22] Pöschel J., *A KAM-Theorem for some nonlinear PDEs*, Ann. Sc. Norm. Pisa, 23, 119-148, 1996.
- [23] Procesi C., *Lie groups: an approach through invariants and representations*, Universitext, Springer, New York, 2007.
- [24] Procesi M., Procesi C., *A KAM algorithm for the resonant nonlinear Schrödinger equation*, preprint 2013.
- [25] Procesi M., Xu X., *Quasi-Töplitz Functions in KAM Theorem*, SIAM J.Math. Anal. 45, 4, 2148-2181, 2013.
- [26] Wang W. M., *Supercritical nonlinear Schrödinger equations I: quasi-periodic solutions*, 2011.
- [27] Wayne E., *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. 127, 479-528, 1990.

This research was supported by the European Research Council under FP7 and partially by the PRIN2009 grant “Critical point theory and perturbative methods for nonlinear differential equations”.