# Quasi-periodic solutions of nonlinear wave equations on $\mathbb{T}^d$

Massimiliano Berti
SISSA, Via Bonomea 265
34136, Trieste, Italy
berti@sissa.it.
&
Philippe Bolle,
Avignon Université, LMA EA2151
F-84000 Avignon, France,
philippe.bolle@univ-avignon.fr.

# Preface

Many Partial Differential Equations (PDEs) arising in physics can be seen as infinite dimensional Hamiltonian systems

$$\partial_t z = J(\nabla_z H)(z)\,, \quad z \in E\,, \tag{0.0.1}$$

where the Hamiltonian function $H : E \to \mathbb{R}$ is defined on an *infinite* dimensional Hilbert space $E$ of functions $z := z(x)$, and $J$ is a non-degenerate antisymmetric operator.

Main examples are the *nonlinear wave* equation (NLW)

$$u_{tt} - \Delta u + V(x)u + g(x,u) = 0\,, \tag{0.0.2}$$

the nonlinear Schrödinger equation (NLS), the beam equation, and the higher dimensional membrane equation, the water waves equations, i.e. the Euler equations of Hydrodynamics describing the evolution of an incompressible irrotational fluid under the action of gravity and surface tension, as well as its approximate models like the Korteweg de Vries (KdV) equation, the Boussinesq, Benjamin-Ono, Kadomtsev-Petviashvili (KP) equations, $\ldots$, among many others. We refer to [95] for a general introduction to Hamiltonian PDEs.

In this Monograph we shall adopt a "Dynamical Systems" point of view, regarding the nonlinear wave equation (0.0.2), equipped with periodic boundary conditions $x \in \mathbb{T}^d := (\mathbb{R}/2\pi\mathbb{Z})^d$, as an infinite dimensional Hamiltonian system, and we shall prove the existence of Cantor families of finite dimensional invariant tori, filled by quasi-periodic solutions of (0.0.2). The first results in this direction are due to Bourgain [40]. The search of invariant sets for the flow is an essential change of paradigm in the study of hyperbolic equations, with respect to the more traditional pursuit of the initial value problem. This perspective has allowed to find many new results, inspired by finite dimensional Hamiltonian systems, for Hamiltonian PDEs.

When the space variable $x$ belongs to a bounded domain like a compact interval $x \in [0, \pi]$ (with Dirichlet boundary conditions), or $x \in \mathbb{T}^d$ (periodic boundary conditions), or, more generally, $x$ belongs to a *compact* manifold, the dynamics of a Hamiltonian PDE (0.0.1), like (0.0.2), is expected to have a "recurrent" behaviour in time, with many *periodic* and *quasi-periodic* solutions, i.e. solutions (defined for *all* times) of the form

$$u(t) = U(\omega t) \in E \qquad \text{where} \qquad \mathbb{T}^n \ni \varphi \mapsto U(\varphi) \in E \tag{0.0.3}$$

is $2\pi$-periodic in the angular variables $\varphi := (\varphi_1, \ldots, \varphi_n)$ and the frequency vector $\omega \in \mathbb{R}^n$ is nonresonant, namely $\omega \cdot \ell \neq 0$, $\forall \ell \in \mathbb{Z}^n \setminus \{0\}$. When $n = 1$ the solution $u(t)$ is periodic in time, with period $2\pi/\omega$. If $U(\omega t)$ is a quasi-periodic solution then, since the orbit $\{\omega t\}_{t \in \mathbb{R}}$ is *dense* on $\mathbb{T}^n$, the torus-manifold $U(\mathbb{T}^n) \subset E$ is invariant under the flow of (0.0.1).

Notice that all the solutions of the linear wave equation (0.0.2) with $g = 0$,

$$u_{tt} - \Delta u + V(x)u = 0\,, \quad x \in \mathbb{T}^d\,, \tag{0.0.4}$$

are of this form. Indeed the self-adjoint operator $-\Delta + V(x)$ possesses a complete $L^2$-orthonormal basis of eigenfunctions $\Psi_j(x)$, $j \in \mathbb{N}$, with eigenvalues $\lambda_j \to +\infty$,

$$(-\Delta + V(x))\,\Psi_j(x) = \lambda_j\,\Psi_j(x)\,, \quad j \in \mathbb{N}\,. \tag{0.0.5}$$

Supposing for simplicity that $-\Delta + V(x) > 0$, the eigenvalues $\lambda_j = \mu_j^2$, $\mu_j > 0$, are positive, and all the solutions of (0.0.4) are

$$\sum_{j \in \mathbb{N}} \alpha_j \cos(\mu_j t + \theta_j)\Psi_j(x)\,, \quad \alpha_j,\ \theta_j \in \mathbb{R}\,, \tag{0.0.6}$$

which, according to the resonance properties of the linear frequencies $\mu_j = \mu_j(V)$, are periodic, quasi-periodic, or almost-periodic in time (i.e. quasi-periodic with infinitely many frequencies).

What happens to these solutions under the effect of the nonlinearity $g(x, u)$ ?

There exist special nonlinear equations for which all the solutions are still periodic, quasi-periodic or almost-periodic in time, for example the sine-Gordon equation, KdV, $1d$ cubic-NLS, .... These are completely integrable PDEs. However, for generic nonlinearities, one expects, in analogy with the celebrated Poincaré non-existence theorem of prime integrals for nearly integrable Hamiltonian systems, that this is not the case.

On the other hand, for sufficiently small Hamiltonian perturbations of a non degenerate integrable system in $\mathbb{T}^n \times \mathbb{R}^n$, the classical KAM –Kolmogorov-Arnold-Moser– theorem proves the persistence of quasi-periodic solutions with *Diophantine* frequency vector $\omega \in \mathbb{R}^n$, i.e. satisfying for some $\gamma > 0$ and $\tau \geq n - 1$, the non-resonance condition

$$|\omega \cdot \ell| \geq \frac{\gamma}{|\ell|^\tau}\,, \quad \forall \ell \in \mathbb{Z}^n \setminus \{0\}\,. \tag{0.0.7}$$

Such frequencies form a Cantor set of $\mathbb{R}^n$ of positive measure if $\tau > n - 1$. These quasi-periodic solutions (which densely fill invariant Lagrangian tori) were constructed by Kolmogorov [92] and Arnold [2] for analytic systems using an iterative Newton scheme, then modified by Moser [97]-[98] with the introduction of smoothing operators in order to deal with merely differentiable perturbations. This scheme then gave rise to abstract Nash-Moser implicit function theorems like the ones due to Zehnder in [117, 118], see also [100], [80].

What happens for infinite dimensional systems like PDEs?

- The central question of KAM theory for PDEs is: *do "most" of the periodic, quasi-periodic, almost-periodic solutions of an integrable PDE (linear or nonlinear) persist, just slightly deformed, under the effect of a nonlinear perturbation?*

KAM theory for Partial Differential Equations started a bit more than thirty years ago with the pioneering works of Kuksin [93] and Wayne [115], about existence of quasi-periodic solutions for semilinear perturbations of 1-dimensional linear wave and Schrödinger equations in the interval $[0, \pi]$. These results are based on an extension of the KAM perturbative approach developed for the search of lower dimensional tori in finite dimensional systems, see [99], [53], [102], and relies on the verification of the so called second order Melnikov non-resonance conditions.

Nowadays KAM theory for 1-$d$ partial differential equations has reached a satisfactory level of comprehension, including bifurcation of small amplitude solutions [96], [104], [16], perturbations of large finite gap solutions [94], [95], [32], [90], [28], extension to periodic boundary conditions [51], [34], [45], [67], use of weak non-degeneracy conditions [11], non-linearities with derivatives [86], [119], [18] up to quasi-linear ones [6]-[8], [62], including water-waves equations [30], [5], applications to quantum harmonic oscillator [77], [9]-[10]. We describe these developments more in detail in section 1.3.

On the other hand, KAM theory for multidimensional PDEs still contains few results and a satisfactory picture is under construction. If the space dimension $d$ is 2 or more, major difficulties are the following:

1. the eigenvalues $\mu_j^2$ of the Sturm-Liouville operator $-\Delta + V(x)$ in (0.0.5) appear in huge clusters of increasing size. For example, if $V(x) = 0$, and $x \in \mathbb{T}^d$, they are
$$\mu_j = |j|^2 = j_1^2 + \ldots + j_d^2 \,, \quad j = (j_1, \ldots, j_d) \in \mathbb{Z}^d .$$
2. The eigenfunctions $\Psi_j(x)$ may be "not localized" with respect to the exponentials, i.e., roughly speaking, the elements $(\Psi_j, e^{\mathrm{i}k \cdot x})_{L^2}$ of the matrix which expresses the change of basis between $(\Psi_j)$ and $(e^{\mathrm{i}k \cdot x})$, do not decay rapidly to zero as the distance $||k| - j| \to +\infty$.

The first existence result of time periodic solutions for the nonlinear wave equation

$$y_{tt} - \Delta y + my = y^3 + \text{h.o.t.} \,, \quad x \in \mathbb{T}^d \,, \ d \geq 2 \,,$$

has been proved by Bourgain in [35], extending the Craig-Wayne approach [51], originally developed if $x \in \mathbb{T}$. Further existence results of periodic solutions have been proved in Berti-Bolle [21] for merely differentiable nonlinearities, Berti-Bolle-Procesi [25] for Zoll manifolds, Gentile-Procesi [71] using Lindstedt series techniques, and Delort [52] for NLS using paradifferential calculus.

The first breakthrough result about existence of quasi-periodic solutions for space multidimensional PDEs was due to Bourgain [37] for analytic Hamiltonian NLS equations of the form

$$\mathrm{i}u_t = \Delta u + M_\sigma u + \varepsilon \partial_{\bar u} H(u, \bar u) \tag{0.0.8}$$

with $x \in \mathbb{T}^2$, where $M_\sigma = \mathrm{Op}(\sigma_j)$ is a Fourier multiplier supported on finitely many sites $\mathbb{S} \subset \mathbb{Z}^2$, i.e. $\sigma_j = 0$, $\forall j \in \mathbb{Z}^d \setminus \mathbb{S}$. The $\sigma_j$, $j \in \mathbb{S}$, play the role of external parameters used to verify suitable non-resonance conditions. Notice that the eigenfunctions of $\Delta + M_\sigma$ are the exponentials $e^{\mathrm{i}j\cdot x}$ and so the above mentioned problem 2 is non present.

Later on, using tools of sub-harmonic analysis previously developed for quasi-periodic Anderson localization theory, in Bourgain-Goldstein-Schlag [41], [39], Bourgain [40] was able to extend this result in any space dimension $d$, and also for analytic nonlinear wave equations of the form

$$y_{tt} - \Delta y + M_\sigma y + \varepsilon F'(y) = 0\,, \quad x \in \mathbb{T}^d\,. \tag{0.0.9}$$

We also mention the existence results of quasi-periodic solutions of Bourgain-Wang [42]-[43] for NLS and NLW under a random perturbation. The stochastic case is a priori easier than the deterministic one because it is simpler to verify the non-resonance conditions with a random variable.

The main analysis for proving the existence of quasi-periodic solutions of (0.0.9) concerns finite dimensional restrictions of the quasi-periodic operators obtained linearizing (0.0.9) at each step of the Newton iteration,

$$\Pi_N \big( (\omega \cdot \partial_\varphi)^2 - \Delta + M_\sigma + \varepsilon b(\varphi, x) \big)_{|\mathcal{H}_N}\,, \tag{0.0.10}$$

where $\varphi \in \mathbb{T}^\nu$ ($\nu$ is the number of frequencies) and $\Pi_N$ denotes the projection on the finite dimensional subspace

$$\mathcal{H}_N := \Big\{ h = \sum_{|(\ell,j)| \leq N} h_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + jx)}\,,\ \ell \in \mathbb{Z}^\nu\,, j \in \mathbb{Z}^d \Big\}\,.$$

The matrix which represents (0.0.10) in the exponential basis is a perturbation of the diagonal matrix $\mathrm{Diag}(-(\omega\cdot\ell)^2 + |j|^2 + \sigma_j)$ with off-diagonal entries $\varepsilon(\widehat{b}_{\ell-\ell',j-j'})$ which decay exponentially to zero as $|(\ell-\ell', j-j')| \to +\infty$. The goal is to prove that such matrix is invertible, for most values of the parameters, and that its inverse has an exponential or Gevrey off-diagonal decay. It is not difficult to impose lower bounds for the eigenvalues of the self-adjoint operator (0.0.10) for most values of the parameters. These "first order Melnikov" non-resonance conditions are essentially the minimal assumptions for proving the persistence of quasi-periodic solutions of (0.0.9), and provide estimates of the inverse of the operator (0.0.10) in $L^2$ norm. In order to prove fast off-diagonal decay estimates for the inverse matrix, Bourgain's technique is a "multiscale" inductive analysis based on the repeated use of the "resolvent identity". An essential ingredient is that the "singular" sites

$$(\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d \quad \text{such that} \quad |-(\omega \cdot \ell)^2 + |j|^2 + \sigma_j| \leq 1 \tag{0.0.11}$$

are separated into clusters which are sufficiently distant from one another (otherwise the bounds on the inverse operators would not be good enough to allow the convergence of the Newton scheme). However, the information (0.0.11) about just the linear frequencies of (0.0.9) is not sufficient (unlike for time-periodic solutions [35]) and also finer properties of non-resonance at each scale along the induction are verified. We describe the multiscale approach in section 1.4 and we prove novel multiscale results in Chapter 4.

These techniques have been extended in the recent work of Wang [114] for the nonlinear Klein-Gordon equation

$$y_{tt} - \Delta y + y + y^{p+1} = 0\,, \quad p \in \mathbb{N}\,,\ x \in \mathbb{T}^d\,,$$

that, unlike (0.0.9), is parameter independent. A key step is to verify that suitable non-resonance conditions are fulfilled for most "initial data". We refer to [113] for a corresponding result for NLS.

Another stream of important results for multidimensional PDEs have been inaugurated in the breakthrough paper [58] of Eliasson-Kuksin for the NLS equation (0.0.8). In this paper the authors are able to block diagonalize, and reduce to constant coefficients, the quasi-periodic Hamiltonian operator obtained at each step of the iteration. This KAM reducibility approach extends the perturbative theory developed for $1d$-PDEs, by verifying the so called second order Melnikov non-resonance conditions. It allows to prove directly also the linear stability of the quasi-periodic solutions. Other results in this direction have been proved for the $2d$-cubic NLS by Geng-Xu-You [70], by Procesi-Procesi [106], [107] in any space dimension and arbitrary polynomial nonlinearities, by Geng-You [68] and Eliasson-Grébert-Kuksin [55] for beam equations. Unfortunately, the second order Melnikov conditions are strongly violated for nonlinear wave equations for which an analogous result does not hold. We describe the KAM reducibility approach with PDEs applications in section 1.3.

We now present more in detail the goal of this research Monograph. The main result is the existence of small amplitude time quasi-periodic solutions for autonomous nonlinear wave equations

$$u_{tt} - \Delta u + V(x)u + g(x, u) = 0\,, \quad x \in \mathbb{T}^d\,, \quad g(x, u) = a(x)u^3 + O(u^4)\,, \tag{0.0.12}$$

in *any* space dimension $d \geq 1$, where $V(x)$ is a smooth *multiplicative* potential such that $-\Delta + V(x) > 0$, and the nonlinearity $g(x, u)$ is $C^\infty$. Given a finite set $\mathbb{S} \subset \mathbb{N}$ (tangential sites) the quasi-periodic solutions $u(\omega t, x)$ that we construct have $|\mathbb{S}|$-independent frequencies $(\omega_j)_{j\in\mathbb{S}}$, and have the form

$$u(\omega t, x) = \sum_{j\in\mathbb{S}} \alpha_j \cos(\omega_j t)\Psi_j(x) + r(\omega t, x)\,, \quad \omega_j = \mu_j + O(|\alpha|)\,, \tag{0.0.13}$$

with $\alpha := (\alpha_j)_{j\in\mathbb{S}}$ and a remainder $r(\varphi, x)$ which is $o(|\alpha|)$-small in some Sobolev space. The solutions (0.0.13) are thus a small deformation of linear solutions (0.0.6), supported on the "tangential" space spanned by the eigenfunctions $(\Psi_j(x))_{j\in\mathbb{S}}$, with a much smaller component in the normal subspace. These quasi-periodic solutions of (0.0.12) exist for generic potentials $V(x)$, functions $a(x)$ and "most" small values of the amplitudes $(\alpha_j)_{j\in\mathbb{S}}$. The precise statement is given in Theorem 1.2.1 and Theorem 1.2.3.

The proof of this result requires various mathematical methods which this book aims to present in a systematic and self-contained way. A complete outline of the steps of proof is presented in section 1.5. Here we just mention that we shall use a Nash-Moser iterative scheme in scales of Sobolev spaces for the search of an invariant torus embedding supporting quasi-periodic solutions, with a frequency $\omega$ to be determined. One key step is to establish the existence of an approximate inverse for the operators obtained by linearizing the nonlinear wave equation at any approximate quasi-periodic solution $u(\omega t, x)$, and to prove that such approximate inverse satisfies tame estimates in Sobolev spaces, with loss of derivatives due to the small divisors. These are linear operators of the form

$$h \mapsto (\omega \cdot \partial_\varphi)^2 h - \Delta h + V(x)h + (\partial_u g)(x, u(\omega t, x))h$$

with coefficients depending on $x \in \mathbb{T}^d$ and $\varphi \in \mathbb{T}^{|\mathbb{S}|}$. The construction of an approximate inverse requires several steps. After writing the wave equation as a Hamiltonian system in infinite dimension, the first step is to use a symplectic change of variable to approximately decouple the tangential and normal components of the linearized operator. It is a rather general procedure for autonomous PDEs, which reduces the problem to the search of an approximate inverse for a quasi-periodic Hamiltonian linear operator acting in the subspace normal to the torus, see Chapter 6 and Appendix C.

In order to avoid the difficulty posed by the violation of the second order Melnikov non-resonance conditions required by a KAM reducibility scheme, we develop a multiscale inductive approach à la Bourgain, which is particularly delicate since the eigenfunctions $\Psi_j(x)$ of $-\Delta + V(x)$ defined in (0.0.5) are not localized near the exponentials. In particular the matrix elements $(\Psi_j, a(x)\Psi_{j'})_{L^2}$ representing the multiplication operator with respect to the basis of the eigenfunctions $\Psi_j(x)$ do not decay, in general, as $j - j' \to \infty$. In Chapter 4 we provide the complete proof of the multiscale proposition (which is fully self-contained together with the Appendix B) which we shall use in Chapters 9-10. These results extend the multiscale analysis developed for forced NLW and NLS in [22]-[23].

The presence of a multiplicative potential $V(x)$ in (0.0.12) makes also difficult to control the variations of the tangential and normal frequencies due to the effect of the nonlinearity $a(x)u^3 + O(u^4)$ with respect to parameters. In this Monograph, after a careful bifurcation analysis of the quasi-periodic solutions, we are able to use just the frequency length $|\omega|$ as an internal parameter to verify all the non-resonance conditions along the iteration. The frequency is constrained to a fixed direction, see (1.2.24)-(1.2.25). The measure estimates are obtained relying on positivity arguments for the variation of parameter dependent families of self-adjoint matrices, see section 4.8, that we verify of the linearized operators obtained along the iteration, see (8.1.8).

The genericity of the non-resonance and non-degeneracy conditions that we require on the potential $V(x)$ and the function $a(x)$ in the nonlinearity $a(x)u^3 + O(u^4)$, are finally verified in Chapter 12.

The techniques developed above for the NLW equation (0.0.12) would certainly apply to prove a corresponding result for nonlinear Schrödinger equations. However we have decided to focus on NLW because, as explained above, there are less results available. This context seems to make appear more evident the advantages of the present approach with respect to that of reducibility.

A feature of the Monograph is to present the proofs, techniques and ideas developed in a self-contained and expanded manner, with the hope to enhance further developments. We also aim to describe the connections of this result with previous works in the literature. The techniques developed in this Monograph have deep connections with those used in Anderson localization theory and we hope that the detailed presentation in this manuscript of all technical aspects of proof, will allow a deeper interchange between the scientific communities of Anderson-localization and “KAM for PDEs”.

Massimiliano Berti, Philippe Bolle,

*Keywords:* Nonlinear wave equation, KAM for PDEs, Quasi-Periodic Solutions, Small Divisors, invariant tori, Multiplicative potential, Nash-Moser Theory, Infinite Dimensional Hamiltonian Systems, Multiscale analysis.

*2000AMS subject classification:* 35Q55, 37K55, 37K50.

This research was partially supported by PRIN 2015 "Variational methods with applications to problems in mathematical physics and geometry" and by ANR-15-CE40-0001 "BEKAM (Beyond KAM theory)".

The authors thank Thomas Kappeler for several suggestions which improved the presentation of the manuscript.

# Contents

# Chapter 1

# Introduction

## 1.1 Main result and historical context

We consider autonomous nonlinear wave equations (NLW)

$$u_{tt} - \Delta u + V(x)u + g(x,u) = 0\,, \quad x \in \mathbb{T}^d := \mathbb{R}^d/(2\pi\mathbb{Z})^d\,, \tag{1.1.1}$$

in any space dimension $d \geq 1$, where $V(x) \in C^\infty(\mathbb{T}^d, \mathbb{R})$ is a real valued *multiplicative* potential and the nonlinearity $g \in C^\infty(\mathbb{T}^d \times \mathbb{R}, \mathbb{R})$ has the form

$$g(x,u) = a(x)u^3 + O(u^4) \tag{1.1.2}$$

with $a(x) \in C^\infty(\mathbb{T}^d, \mathbb{R})$. We require that

$$-\Delta + V(x) > \beta\,\mathrm{Id}\,, \quad \beta > 0\,. \tag{1.1.3}$$

Condition (1.1.3) is satisfied, in particular, if the potential $V(x) \geq 0$ and $V(x) \not\equiv 0$.

In this Monograph we prove the existence of small amplitude time quasi-periodic solutions of (1.1.1). We remind that a solution $u(t,x)$ of (1.1.1) is time quasi-periodic with frequency vector $\omega \in \mathbb{R}^\nu$, $\nu \in \mathbb{N}_+$, if it has the form

$$u(t,x) = U(\omega t, x)$$

where $U : \mathbb{T}^\nu \times \mathbb{T}^d \to \mathbb{R}$ is a continuous function and $\omega \in \mathbb{R}^\nu$ is a nonresonant vector, namely

$$\omega \cdot \ell \neq 0\,, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\,.$$

If $\nu = 1$ a solution of this form is time-periodic with period $2\pi/\omega$.

Small amplitude solutions of (1.1.1) will be close to solutions of the linear wave equation

$$u_{tt} - \Delta u + V(x)u = 0\,, \quad x \in \mathbb{T}^d\,. \tag{1.1.4}$$

The Sturm-Liouville operator $-\Delta + V(x)$ admits an $L^2$-orthonormal basis of eigenfunctions $\{\Psi_j\}_{j\in\mathbb{N}}$,

$$(-\Delta + V(x))\Psi_j(x) = \mu_j^2 \Psi_j(x)\,, \tag{1.1.5}$$

with positive eigenvalues

$$0 < \beta \le \mu_0^2 \le \mu_1^2 \le \ldots \le \mu_j^2 \le \ldots\,, \quad \mu_j > 0\,, \quad (\mu_j^2) \to +\infty\,,$$

written in increasing order and with multiplicities, and therefore the solutions of the linear wave equation (1.1.4) are given by the linear superpositions of *normal modes* oscillations,

$$\sum_{j\in\mathbb{N}} \alpha_j \cos(\mu_j t + \theta_j)\Psi_j(x)\,, \qquad \alpha_j\,, \theta_j \in \mathbb{R}\,. \tag{1.1.6}$$

All the solutions (1.1.6) of (1.1.4) are periodic, or quasi-periodic, or almost periodic in time, with linear frequencies of oscillations $\mu_j$, according to the resonance properties of $\mu_j$ (which depend on the potential $V(x)$) and how many normal mode amplitudes $\alpha_j$ are not zero. In particular, if $\alpha_j = 0$ for any index $j$ except a finite set $\mathbb{S}$ (tangential sites), and the frequency vector $\bar{\mu} := (\mu_j)_{j\in\mathbb{S}}$ is nonresonant, then the linear solutions (1.1.6) are quasi-periodic in time.

The main question we pose is the following:

- Do small amplitude quasi-periodic solutions of the nonlinear wave equation (1.1.1) exist?

The main result presented in this Monograph (Theorem 1.2.1) actually proves that *small amplitude quasi-periodic solutions* (1.1.6) *of the linear wave equation* (1.1.4)*, which are supported on finitely many indices $j \in \mathbb{S}$, persist, slightly deformed, as quasi-periodic solutions of the nonlinear wave equation* (1.1.1)*, with a frequency vector $\omega$ close to $\bar{\mu}$, for "generic" potentials $V(x)$ and coefficients $a(x)$ (Theorem 1.2.3) and "most" amplitudes* $(\alpha_j)_{j\in\mathbb{S}}$.

The potentials $V(x)$ and the functions $a(x)$ such that Theorem 1.2.1 holds are generic in a very strong sense; in particular they are $C^\infty$-dense, according to Definition 1.2.2, in the set

$$\big(\mathcal{P} \cap C^\infty(\mathbb{T}^d)\big) \times C^\infty(\mathbb{T}^d)$$

where $\mathcal{P} := \big\{V(x) \in H^s(\mathbb{T}^d) \ : \ -\Delta + V(x) > 0\big\}$, see (1.2.38).

Theorem 1.2.1 is a KAM (Kolmogorov-Arnold-Moser) type perturbative result. We construct recursively an embedded invariant torus which supports quasi-periodic solutions of (1.1.1) with frequency vector $\omega$ (to be determined), by a modified Nash-Moser iterative scheme for the search of zeros

$$\mathcal{F}(\lambda; i) = 0$$

of a nonlinear operator $\mathcal{F}$ acting on scales of Sobolev spaces of functions $i$, depending on a suitable parameter $\lambda$, see Chapter 5. As in a Newton scheme, the core of the problem consists in the analysis of the linearized operators

$$d_i\mathcal{F}(\lambda;\underline{i})$$

at any approximate solution $\underline{i}$ at each step of the iteration, and proving its approximate invertibility, for most values of the parameters, with quantitative tame estimates for the approximate inverse in high Sobolev norms. The approximate inverse will be unbounded, i.e. it loses derivatives, due to the presence of *small divisors*. As we shall describe in detail in section 1.5, the construction of an approximate inverse for the linearized operators obtained from (1.1.1) is a subtle problem due to complicated resonance phenomena between the frequency vector $\omega$ of the expected quasi-periodic solutions and the multiple normal mode frequencies of oscillations, shifted by the nonlinearity, and the fact that the normal mode eigenfunctions $\Psi_j(x)$ are not "localized close to the exponentials".

We now make a short historical introduction to KAM theory for partial differential equations, that we shall expand in section 1.3. As we already mentioned in the preface, in these small divisors problems for PDEs, as (1.1.1), the space dimension $d = 1$ or $d \geq 2$ makes a fundamental difference, due to the very different properties of the eigenvalues and eigenfunctions of the Sturm-Liouville operator $-\Delta + V(x)$ on $\mathbb{T}^d$ for $d = 1$ and $d \geq 2$.

The first KAM existence results of quasi-periodic solutions were proved by Kuksin [93], see also [95], and Wayne [115] for 1-$d$ wave and Schrödinger (NLS) equations on the interval $x \in [0, \pi]$ with Dirichlet boundary conditions and analytic nonlinearities, see (1.3.22)-(1.3.23). These pioneering theorems were limited to Dirichlet boundary conditions because the eigenvalues $\mu_j^2$ of the Sturm-Liouville operator $-\partial_{xx} + V(x)$ had to be simple. Indeed the KAM scheme in [95], [115], see also [103], reduces the linearized equations along the iteration to a diagonal form, with coefficients constant in time, requiring "second-order Melnikov" non-resonance conditions, which concern lower bounds for differences among the linear frequencies. In these papers the potential $V(x)$ is used as a parameter to impose non-resonance conditions. Once the linearized PDEs obtained along the iteration are reduced to diagonal, constant in time, form, it is easy to prove that the corresponding linear operators are invertible, for most values of the parameters, with good estimates of their inverses in high norms (with of course loss of derivatives). We refer to section 1.3 for a more detailed explanation of the KAM reducibility approach.

Subsequently these results have been extended by Pöschel [104] for parameter independent nonlinear Klein-Gordon equations like (1.3.31), and by Kuksin-Pöschel [96] for NLS equations like (1.3.30), using Birkhoff normal form techniques to verify (weak) non-resonance conditions among the perturbed frequencies, tuning the amplitudes of the solutions as parameters.

In the case $x \in \mathbb{T}$, the eigenvalues of the Sturm-Liouville operator $-\partial_{xx} + V(x)$ are asymptotically double, and therefore the previous second order Melnikov non-resonance conditions are violated. In this case the first existence results were obtained by Craig-Wayne [51] for time periodic solutions of analytic nonlinear Klein-Gordon equations (see also [49] and [19] for completely resonant wave equations), and then extended by Bourgain [34] for time quasi-periodic solutions. The proofs are based on a Lyapunov-Schmidt bifurcation approach and a Nash-Moser implicit function iterative scheme. The key point of these papers is to renounce to diagonalize the linearized equations at each step of the Nash-Moser iteration. The advantage is to require only minimal non-resonance conditions which are easily verified for PDEs also in presence of multiple frequencies (the second order Melnikov non-resonance conditions are not used). On the other hand, a difficulty of this approach is that, since the linearized equations obtained along the iteration are variable coefficients PDEs, it is hard to prove that the corresponding linear operators are invertible with estimates of their inverses in high norms, sufficient to imply the convergence of the iterative scheme. Relying on a "resolvent" type analysis inspired by the work of Frölich-Spencer [66] in the context of Anderson localization, Craig-Wayne [51] were able to solve this problem for time periodic solutions in $d = 1$, and Bourgain in [34] also for quasi-periodic solutions. Key properties of this approach are:

$(i)$ "separation properties" between singular sites, namely the Fourier indices $(\ell, j)$ of the small divisors $|(\omega \cdot \ell)^2 - j^2| \leq C$ in the case of (NLW);

$(ii)$ "localization" of the eigenfunctions of the Sturm-Liouville operator $-\partial_{xx}+V(x)$ with respect to the exponential basis $(e^{\mathrm{i}kx})_{k\in\mathbb{Z}}$, namely that the Fourier coefficients $(\hat{\Psi}_j)_k$ converge rapidly to zero when $||k| - j| \to \infty$. This property is always true if $d = 1$.

Property $(ii)$ implies that the matrix which represents, in the eigenfunction basis, the multiplication operator for an analytic (resp. Sobolev) function has an exponentially (resp. polynomially) fast decay off the diagonal. Then the "separation properties" $(i)$ imply a very "weak interaction" between the singular sites. If the singular sites were "too many" the inverse operator would be "too unbounded" to prevent the convergence of the iterative scheme. This approach is particularly inspiring in presence of multiple normal mode frequencies and it stands at the basis of the present Monograph. We describe it in more detail in section 1.4.

Later on, Chierchia-You [45] were able to extend the KAM reducibility approach to prove existence and stability of small amplitude quasi-periodic solutions of 1-$d$ NLW on $\mathbb{T}$ with an external potential. We also mention the KAM reducibility results in Berti-Biasco-Procesi [17]-[18] for 1-$d$ derivative wave equations.

In the case the space dimension $d$ is $\geq 2$ major difficulties are:

1. the eigenvalues $\mu_j^2$ of $-\Delta + V(x)$ in (1.1.5) may be highly degenerate, or not sufficiently separated from each other in a suitable quantitative sense, required by the perturbation theory developed for 1-$d$-PDEs;

2. the eigenfunctions $\Psi_j(x)$ of $-\Delta + V(x)$ may be not "localized" with respect to the exponentials, see [61].

As discussed in the preface, if $d \geq 2$, the first KAM existence result for nonlinear wave equations has been proved for time periodic solutions by Bourgain [35], see also the extensions in [21], [25], [71]. Concerning quasi-periodic solutions in $d \geq 2$, the first existence result was proved by Bourgain in Chapter 20 of [40], for wave type equations of the form

$$u_{tt} - \Delta u + M_\sigma u + \varepsilon F'(u) = 0$$

where $M_\sigma = \mathrm{Op}(\sigma_j)$ is a Fourier multiplier supported on finitely many sites $\mathbb{S} \subset \mathbb{Z}^2$, i.e. $\sigma_j = 0$, $\forall j \in \mathbb{Z}^d \setminus \mathbb{S}$. The $\sigma_j$, $j \in \mathbb{S}$, are used as a parameter, and $F$ is a polynomial nonlinearity. Notice that the linear equation $u_{tt} - \Delta u + M_\sigma u = 0$ is diagonal in the exponential basis $e^{\mathrm{i}j\cdot x}$, $j \in \mathbb{Z}^d$, unlike the linear wave equation (1.1.4). We also mention the paper by Wang [113] for the completely resonant NLS (1.3.39) and the Anderson localization result of Bourgain-Wang [42] for time quasi-periodic random linear Schrödinger and wave equations.

As already mentioned, a major difficulty of this approach is that the linearized equations obtained along the iteration are PDEs with variable coefficients. A key property which plays a fundamental role in [40] (as well as in previous papers as [37] for NLS) for proving estimates for the inverse of linear operators

$$\Pi_N \big( (\omega \cdot \partial_\varphi)^2 - \Delta + M_\sigma + \varepsilon b(\varphi, x) \big)_{|\mathcal{H}_N} ,$$

(see (0.0.10)) is that the matrix which represents the multiplication operator for a smooth function $b(x)$ in the exponential basis $\{e^{\mathrm{i}j\cdot x}\}$, $j \in \mathbb{Z}^d$, has a sufficiently fast off-diagonal decay. Indeed the multiplication operator is represented in Fourier space as a convolution operator with a Töplitz matrix $(\hat{b}_{j-j'})_{j,j' \in \mathbb{Z}^d}$, with entries given by the Fourier coefficients $\hat{b}_J$ of the function $b(x)$, constant on the diagonal $j - j' = J$. The smoother the function $b(x)$ is, the faster is the decay of $\hat{b}_{j-j'}$ as $|j - j'| \to +\infty$. We refer to section 1.4 for more explanations.

Weaker forms of this property, as for example those required in Berti-Corsi-Procesi [26], [31] may be sufficient for dealing with the eigenfunctions of $-\Delta$ on compact Lie groups. However, any possible off-diagonal decay-property may lack for the matrix elements $(\Psi_j, b(x)\Psi_{j'})_{L^2}$ representing the multiplication operator with respect to the basis of the eigenfunctions $\Psi_j(x)$ defined in (1.1.5) of $-\Delta + V(x)$ on $\mathbb{T}^d$, $d \geq 2$. This was proved by Feldman, Knörrer, Trubowitz in [61] and it is the difficulty mentioned in item 2. We remark that weak properties of localization have been proved by Wang [112] in $d = 2$ for potentials $V(x)$ which are trigonometric polynomials.

In the present Monograph we shall not use any kind of localizations properties of the eigenfunctions $\Psi_j(x)$, that actually might *not* be true. A major reason why we are able to avoid the use of such properties is that our Nash-Moser iterative scheme requires only very weak tame estimates for the approximate inverse of the linearized operators as (1.4.6) see the end of subsection 1.4.3. Such conditions are close to the optimal ones, as a famous counterexample of Lojiaciewitz-Zehnder in [87] shows.

The properties of the exponential basis $e^{\mathrm{i} j \cdot x}$, $j \in \mathbb{Z}^d$, play a key role also for developing the KAM perturbative diagonalization/reducibility techniques, and, indeed, no reducibility results are available so far for multidimensional PDEs in presence of a multiplicative potential which is not small. Concerning higher space dimensional PDEs we refer to the results in Eliasson-Kuksin [58] for the NLS equation (1.3.37) with a convolution potential on $\mathbb{T}^d$, used as a parameter, Geng-You [68] and Eliasson-Grébert-Kuksin [56] for beam equations with a constant mass potential, Procesi-Procesi [106] for the completely resonant NLS (1.3.39), Grébert-Paturel [75] for the Klein-Gordon equation (1.3.40) on $\mathbb{S}^d$ and Grébert-Paturel [76] for multidimensional harmonic oscillators.

On the other hand, no reducibility results for NLW on $\mathbb{T}^d$ are known so far. Actually a serious difficulty which appears is the following: the infinitely many second order Melnikov non-resonance conditions required by the KAM-diagonalization approach are strongly violated yet by the linear unperturbed frequencies of oscillations of the Klein-Gordon equation $u_{tt} - \Delta u + mu = 0$, see [55]. A key difference with respect to the Schrödinger equation is that the linear frequencies of the wave equations are $\sim |j|$, $j \in \mathbb{Z}^d$, while for NLS, and beam equation, are $\sim |j|^2$, respectively $\sim |j|^4$, and $|j|^2$, $|j|^4$ are integer. Also for the multidimensional harmonic oscillator the linear frequencies are, up to a translation, integer numbers. Although no reducibility results are known so far for NLW, a result of "almost" reducibility for linear quasi-periodically forced Klein-Gordon equations has been presented in [54], [55].

Existence of Sobolev quasi-periodic solutions for wave equations on $\mathbb{T}^d$ with a time-quasi periodic differentiable forcing nonlinearity

$$u_{tt} - \Delta u + V(x)u = \varepsilon f(\omega t, x, u), \quad x \in \mathbb{T}^d, \tag{1.1.7}$$

has been proved in Berti-Bolle [22] extending the multiscale approach of Bourgain [40]. The forcing frequency vector $\omega$, which in [22] is constrained to a fixed direction $\omega = \lambda \bar{\omega}$, $\lambda \in [1/2, 3/2]$, plays the role of an external parameter. In [26] a corresponding result has been extended for NLW on compact Lie groups, in [33] for Zoll manifolds, in [29] for general flat tori, and in [48] for forced Kirkhoff equations.

Existence of quasi-periodic solutions for autonomous non-linear Klein Gordon equations

$$u_{tt} - \Delta u + u + u^{p+1} + h.o.t. = 0, \quad p \in \mathbb{N}, \quad x \in \mathbb{T}^d, \tag{1.1.8}$$

have been recently presented by Wang [114], relying on a bifurcation analysis to study the modulation of the frequencies induced by the nonlinearity $u^{p+1}$, and multiscale methods of [40] for implementing a Nash-Moser iteration. The result proves the continuation of quasi-periodic solutions supported on "good" tangential sites.

The papers [22]-[23] for forced NLW and NLS are the closest background of the present Monograph. The passage to prove KAM results for autonomous nonlinear wave equations with a multiplicative potential as (1.1.1) is a non trivial task, since it requires a bifurcation analysis which distinguishes the tangential directions where the major part of the oscillation of the quasi-periodic solutions takes place, and the normal ones, see the form (1.2.33) of the quasi-periodic solutions proved in Theorem 1.2.1. When the multiplicative potential $V(x)$ changes, both the tangential and the normal frequencies vary simultaneously in an intricate way (unlike the case of the convolution potential). This makes difficult to verify the non-resonance conditions required by the Nash-Moser iteration. In particular, the choice of the parameters adopted in order to fulfill all these conditions is relevant. In this Monograph we choose any finite set $\mathbb{S} \subset \mathbb{N}$ of tangential sites, we fix the potential $V(x)$ and the function $a(x)$ appearing in the nonlinearity (1.1.2) (in such a way that generic non-resonance and non-degeneracy conditions hold, see Theorem 1.2.3) and then we prove, in Theorem 1.2.1, the existence of quasi-periodic solutions of (1.1.1) for most values of the one dimensional internal parameter $\lambda$ introduced in (1.2.24), which amounts just to a *time rescaling* of the frequency vector $\omega$. This also implies a density result for the frequencies of the quasi-periodic solutions close to the unperturbed vector $\bar{\mu}$. We shall explain more in detail the choice of this parameter in section 1.5.

## 1.2 Rigorous statement

In this section we state precisely the main result of this Monograph, which is Theorem 1.2.1.

Under the rescaling $u \mapsto \varepsilon u$, $\varepsilon > 0$, the equation (1.1.1) is transformed into the nonlinear wave equation

$$u_{tt} - \Delta u + V(x)u + \varepsilon^2 g(\varepsilon, x, u) = 0 \tag{1.2.1}$$

with the $C^\infty$ nonlinearity

$$g(\varepsilon, x, u) := \varepsilon^{-3} g(x, \varepsilon u) = a(x)u^3 + O(\varepsilon u^4). \tag{1.2.2}$$

We choose arbitrarily a finite set of indices $\mathbb{S} \subset \mathbb{N}$, called the "tangential sites". We denote by $|\mathbb{S}| \in \mathbb{N}$ the cardinality of $\mathbb{S}$ and we order the tangential sites by $\mathbb{S} = \{j_1, \ldots, j_{|\mathbb{S}|}\}$. We look for quasi-periodic solutions of (1.2.1) which are perturbations of normal modes oscillations supported on $j \in \mathbb{S}$. We denote by

$$\bar{\mu} := (\mu_j)_{j\in\mathbb{S}} = (\mu_{j_1}, \ldots, \mu_{j_{|\mathbb{S}|}}) \in \mathbb{R}^{|\mathbb{S}|}, \quad \mu_j > 0, \tag{1.2.3}$$

the frequency vector of the quasi-periodic solutions

$$\sum_{j\in\mathbb{S}} \mu_j^{-1/2} \sqrt{2\xi_j} \cos(\mu_j t) \Psi_j(x), \qquad \xi_j > 0, \tag{1.2.4}$$

of the linear wave equation (1.1.4). The components of $\bar{\mu}$ are called the unperturbed tangential frequencies. We shall call the indices in the complementary set $\mathbb{S}^c := \mathbb{N} \setminus \mathbb{S}$, the "normal" sites, and the corresponding $\mu_j$, $j \in \mathbb{S}^c$, the unperturbed "normal" frequencies.

Since (1.2.1) is an *autonomous* PDE, the frequency vector $\omega \in \mathbb{R}^{|\mathbb{S}|}$ of its expected quasi-periodic solutions $u(\omega t, x)$ is an unknown, that we introduce as an explicit parameter in the equation, looking for solutions $u(\varphi, x)$, $\varphi = (\varphi_1, \ldots, \varphi_{|\mathbb{S}|}) \in \mathbb{T}^{|\mathbb{S}|}$, of

$$(\omega \cdot \partial_\varphi)^2 u - \Delta u + V(x) u + \varepsilon^2 g(\varepsilon, x, u) = 0 \,. \tag{1.2.5}$$

The frequency vector $\omega \in \mathbb{R}^{|\mathbb{S}|}$ of the expected quasi-periodic solutions of (1.2.1) will be $O(\varepsilon^2)$-close to the unperturbed tangential frequency $\bar{\mu}$ in (1.2.3), see precisely (1.2.24)-(1.2.25).

Since the nonlinear wave equation (1.2.1) is time-reversible (see Appendix A), it makes sense to look for solutions of (1.2.1) which are even in $t$. Since (1.2.1) is autonomous, more general solutions are obtained from these even solutions by time translation. Thus we look for solutions $u(\varphi, x)$ of (1.2.5) *even* in $\varphi$. This induces a small simplification in the proof, see remark 5.1.1.

In order to prove, for $\varepsilon$ small enough, existence of solutions of (1.2.1) close to the solutions (1.2.4) of the linear wave equation (1.1.4), we first require non-resonance conditions for the unperturbed linear frequencies $\mu_j$, $j \in \mathbb{N}$, which will be verified by generic potentials $V(x)$, see Theorem 1.2.3.

**Diophantine and 1-th order Melnikov non-resonance conditions.** We assume that

- the tangential frequency vector $\bar{\mu}$ in (1.2.3) is Diophantine, i.e. for some constants $\gamma_0$, $\tau_0 > 0$,

$$|\bar{\mu} \cdot \ell| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}} \,, \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|} \setminus \{0\} \,, \qquad \langle \ell \rangle := \max\{1, |\ell|\} \,, \tag{1.2.6}$$

  where $|\ell| := \max\{|\ell_1|, \ldots, |\ell_{|\mathbb{S}|}|\}$. Notice that (1.2.6) implies, in particular, that the unperturbed tangential frequencies $\mu_j$, $j \in \mathbb{S}$, are simple.

- the unperturbed "first order Melnikov" non-resonance conditions hold:

$$|\bar{\mu} \cdot \ell + \mu_j| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}} \,, \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|} \,, \; j \notin \mathbb{S} \,. \tag{1.2.7}$$

The non-resonance conditions (1.2.6), (1.2.7) imply, in particular, that the linear equation (1.1.4) has no other quasi-periodic solutions with frequency $\bar{\mu}$, even in $t$, except the trivial ones (1.2.4).

In order to prove "separation properties" of the small divisors as required by the multiscale analysis that we perform in Chapter 4, we require, as in [22], that

- the tangential frequency vector $\bar{\mu}$ in (1.2.3) satisfies the quadratic Diophantine condition

$$\Big| n + \sum_{i,j\in\mathbb{S}, i<j} p_{ij}\mu_i\mu_j \Big| \geq \frac{\gamma_0}{\langle p\rangle^{\tau_0}}, \quad \forall (n,p) \in \mathbb{Z}\times\mathbb{Z}^{\frac{|\mathbb{S}|(|\mathbb{S}|+1)}{2}} \setminus \{0\}. \tag{1.2.8}$$

The non-resonance conditions (1.2.6), (1.2.7) and (1.2.8) amount to assumptions on the potential $V(x)$, which are "generic" in the sense of Kolmogorov measure, see [89] where (1.2.6), (1.2.7) are proved to hold for most potentials. Genericity results are stated in Theorem 1.2.3, proved in Chapter 12.

We underline that along the Monograph the constant $\gamma_0 > 0$ in (1.2.6), (1.2.7), (1.2.8) is regarded as fixed, and we shall often omit to track its dependence in the estimates.

**Birkhoff matrices.** We are interested in quasi-periodic solutions of (1.2.1) which bifurcate for small $\varepsilon > 0$ from a solution of the form (1.2.4) of the linear wave equation. In order to prove their existence, it is important to know precisely how the tangential and the normal frequencies change with respect to the unperturbed actions $(\xi_j)_{j\in\mathbb{S}}$, under the effect of the nonlinearity $\varepsilon^2 a(x)u^3 + O(\varepsilon^3 u^4)$. This is described in terms of the "Birkhoff" matrices

$$\mathscr{A} := (\mu_k^{-1} G_k^j \mu_j^{-1})_{j,k\in\mathbb{S}}, \quad \mathscr{B} := (\mu_j^{-1} G_j^k \mu_k^{-1})_{j\in\mathbb{S}^c, k\in\mathbb{S}}, \tag{1.2.9}$$

where

$$G_k^j := G_k^j(V,a) := \begin{cases} (3/2)(\Psi_j^2, a(x)\Psi_k^2)_{L^2}, & j\neq k, \\ (3/4)(\Psi_j^2, a(x)\Psi_j^2)_{L^2}, & j = k \end{cases} \tag{1.2.10}$$

and $\Psi_j(x)$ are the eigenfunctions of $-\Delta + V(x)$ introduced in (1.1.5). Notice that the matrix $(G_k^j)$ depends on the function $a(x)$ and the eigenfunctions $\Psi_j$, thus on the potential $V(x)$. The $|\mathbb{S}|\times|\mathbb{S}|$ symmetric matrix $\mathscr{A}$ is called the "twist"-matrix. The matrices $\mathscr{A}, \mathscr{B}$ describe the shift of the tangential and normal frequencies induced by the nonlinearity $a(x)u^3$ as they appear in the fourth order Birkhoff normal form of (1.1.1)-(1.1.2). Actually, we prove in section 5.2 that, up to terms $O(\varepsilon^4)$, the tangential frequency $\omega$ of a small amplitude quasi-periodic solution of (1.1.1)-(1.1.2) close to (1.2.4) is given by the action-to-frequency map

$$\xi \mapsto \bar{\mu} + \varepsilon^2 \mathscr{A}(\xi), \quad \xi \in \mathbb{R}_+^{|\mathbb{S}|}. \tag{1.2.11}$$

On the other hand the perturbed normal frequencies are shifted by the matrix $\mathscr{B}$ as described in Lemma 7.3.2. We assume that

- **(Twist condition)**

$$\det \mathscr{A} \neq 0, \tag{1.2.12}$$

and therefore the action-to-frequency map in (1.2.11) is invertible. The non-degeneracy, or "twist"-condition (1.2.12), is generically satisfied by choosing the potential $V(x)$ and the function $a(x)$, as stated in Theorem 1.2.3 (see in particular Corollary 12.1.10 and remark 12.1.11).

**Second order Melnikov non-resonance conditions.** We also assume second order Melnikov non-resonance conditions which concern only *finitely* many unperturbed normal frequencies. We have first to introduce an important decomposition of the normal indices $j \in \mathbb{S}^c$. Note that, since $\mu_j \to +\infty$, the indices $j \in \mathbb{S}^c$ such that $\mu_j - (\mathscr{B}\mathscr{A}^{-1}\bar{\mu})_j < 0$ are finitely many. Denoting

$$-\mathfrak{g} := \min \left\{ \mu_j - (\mathscr{B}\mathscr{A}^{-1}\bar{\mu})_j \, , j \in \mathbb{S}^c \right\} , \tag{1.2.13}$$

we split the normal indices as

$$\mathbb{S}^c = \mathbb{F} \cup \mathbb{G} \, , \quad \mathbb{G} := \mathbb{S}^c \setminus \mathbb{F} \, , \tag{1.2.14}$$

where

$$\begin{aligned} \mathbb{F} &:= \Big\{ j \in \mathbb{S}^c \, : \, |\mu_j - (\mathscr{B}\mathscr{A}^{-1}\bar{\mu})_j| \leq \mathfrak{g} \Big\} \, , \\ \mathbb{G} &:= \Big\{ j \in \mathbb{S}^c \, : \, \mu_j - (\mathscr{B}\mathscr{A}^{-1}\bar{\mu})_j > \mathfrak{g} \Big\} \, . \end{aligned} \tag{1.2.15}$$

The set $\mathbb{F}$ is always finite, and it is empty if $\mathfrak{g} < 0$. The relevance of the decomposition (1.2.14) of the normal sites, concerns the variation of the normal frequencies with respect to the length of the tangential frequency vector, as we describe in (1.5.24) below, see also Lemma 7.1.1. If all the $\mu_j - (\mathscr{B}\mathscr{A}^{-1}\bar{\mu})_j$ ($j \in \mathbb{S}^c$) were positive, then, by (1.5.24), one could directly rely on positivity arguments as in [22], [23], used in the forced case to verify the measure estimates. In general $\mathfrak{g} > 0$ and we shall be able to decouple, for most values of the parameter $\lambda$, the linearized operators obtained at each step of the nonlinear Nash-Moser iteration, acting in the normal subspace $H_{\mathbb{S}}^{\perp}$, along $H_{\mathbb{F}}$ and its orthogonal $H_{\mathbb{F}}^{\perp} = H_{\mathbb{G}}$. We discuss the relevance of this decomposition in section 1.5.

We assume the following

- unperturbed "second order Melnikov" non-resonance conditions:

$$|\bar{\mu} \cdot \ell + \mu_j - \mu_k| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}} \, , \; \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{F} \times \mathbb{S}^c \, , \quad (\ell, j, k) \neq (0, j, j) \, , \tag{1.2.16}$$

$$|\bar{\mu} \cdot \ell + \mu_j + \mu_k| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}} \, , \; \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{F} \times \mathbb{S}^c \, . \tag{1.2.17}$$

Note that (1.2.16) implies, in particular, that the finitely many normal frequencies $\mu_j, j \in \mathbb{F}$, are simple (clearly all the other eigenvalues $\mu_j$, $j \notin \mathbb{F}$, could be highly degenerate).

In order to verify a key positivity property for the variations of the restricted linearized operator with respect to $\lambda$ (Lemma 9.3.8), we assume further

- unperturbed "second order Melnikov" non-resonance conditions:

$$|\bar\mu \cdot \ell + \mu_j - \mu_k| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}}\,, \quad \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times (\mathbb{M} \setminus \mathbb{F}) \times \mathbb{S}^c\,, \quad (\ell, j, k) \neq (0, j, j)\,, \tag{1.2.18}$$

$$|\bar\mu \cdot \ell + \mu_j + \mu_k| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}}\,, \quad \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times (\mathbb{M} \setminus \mathbb{F}) \times \mathbb{S}^c\,, \tag{1.2.19}$$

  where

$$\mathbb{M} := \left\{ j \in \mathbb{S}^c \, : \, |j| \leq C_1 \right\} \tag{1.2.20}$$

  and the constant $C_1 := C_1(V, a) > 0$ is taken large enough such that $\mathbb{F} \subset \mathbb{M}$ and (7.1.8) holds. Notice that the set $\mathbb{M}$ depends on $(V, a)$, but can be chosen to be locally constant in $(V, a)$; it will be fixed in Lemma 7.1.1.

Clearly the conditions (1.2.16)-(1.2.17) and (1.2.18)-(1.2.19) could have been written together, requiring such conditions for $j \in \mathbb{M}$, without distinguishing the cases $j \in \mathbb{F}$ and $j \in \mathbb{M} \setminus \mathbb{F}$. However, for conceptual clarity, in view of their different role in the proof, we prefer to state them separately. The above conditions (1.2.16)-(1.2.19) on the unperturbed frequencies allow to perform one step of averaging and so to diagonalize, up to $O(\varepsilon^4)$, the normal frequencies supported on $\mathbb{M}$, see Proposition 7.3.1. This is the only step where conditions (1.2.18)-(1.2.19) play a role. Conditions (1.2.16)-(1.2.17) are used also in the splitting step of Chapter 9, see Lemma 9.3.3.

Conditions (1.2.16)-(1.2.19) depend on the potential $V(x)$ and also on $a(x)$, because the constant $C_1$ in (1.2.20) (hence the set $\mathbb{M}$) depends on $a(x)$, actually on $\|a\|_{L^\infty}$ and $\|\mathcal{A}^{-1}\|$. Given $(V_0, a_0)$ such that the matrix $\mathscr{A}$ defined in (1.2.9) is invertible and $s > d/2$, the set $\mathbb{M}$ can be chosen constant in some open neighborhood $U$ of $(V_0, a_0)$ for the $H^s$-norm. In $U$, conditions (1.2.16)-(1.2.19) are generic in $V(x)$, as it is proved in Chapter 12 (see Theorem 1.2.3).

**Non-degeneracy conditions.** We also require the following *finitely* many

- non-degeneracy conditions:

$$(\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_j) - (\mu_k - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_k) \neq 0\,, \quad \forall j, k \in \mathbb{F}\,,\ j \neq k\,, \tag{1.2.21}$$

$$(\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_j) + (\mu_k - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_k) \neq 0\,, \quad \forall j, k \in \mathbb{F}\,, \tag{1.2.22}$$

  where $\mathscr{A}$ and $\mathscr{B}$ are the Birkhoff matrices defined in (1.2.9).

Such assumptions are similar to the non-degeneracy conditions required for the continuation of elliptic tori for finite dimensional systems in [53], [102] and for PDEs in [96], [104], [16]. Notice that the finitely many non-degeneracy conditions (1.2.21) depend on the potential $V(x)$ and the nonlinearity $a(x)u^3$ and we prove in Theorem 1.2.3 that they are generic in

$(V, a)$.

**Parameter.** We now introduce the 1-dimensional parameter that we shall use to perform the measure estimates.

In view of (1.2.11) the frequency $\omega$ has to belong to the cone of the "admissible" frequencies $\bar{\mu} + \varepsilon^2 \mathscr{A}(\mathbb{R}_+^{|\mathbb{S}|})$, more precisely we require that $\omega$ belongs to the image

$$\mathcal{A} := \bar{\mu} + \varepsilon^2 \mathscr{A}\Big(\Big[\frac{1}{2}, 4\Big]^{|\mathbb{S}|}\Big) \subset \mathbb{R}^{|\mathbb{S}|} \tag{1.2.23}$$

of the compact set of actions $\xi \in [1/2, 4]^{|\mathbb{S}|}$ under the approximate action-to-frequency map (1.2.11). Then, in view of the method that we shall use for the measure estimates for the linearized operator, we look for quasi-periodic solutions with frequency vector

$$\omega = (1 + \varepsilon^2 \lambda)\bar{\omega}_\varepsilon \,, \quad \lambda \in \Lambda := [-\lambda_0, \lambda_0] \,, \tag{1.2.24}$$

constrained to a fixed admissible direction

$$\bar{\omega}_\varepsilon := \bar{\mu} + \varepsilon^2 \zeta \,, \quad \zeta \in \mathscr{A}([1, 2]^{|\mathbb{S}|}) \,, \tag{1.2.25}$$

(notice that in general we can not take $\bar{\omega}_\varepsilon = \bar{\mu}$, because $\zeta = 0$ might not belong to $\mathscr{A}([1, 2]^{|\mathbb{S}|})$). We fix $\zeta$ below so that the Diophantine conditions (1.2.29)-(1.2.30) hold.

In (1.2.24) there exists $\lambda_0 > 0$ small, independent of $\varepsilon > 0$ and of $\zeta \in \mathscr{A}([1, 2]^{|\mathbb{S}|})$, such that,

$$\forall \lambda \in \Lambda := [-\lambda_0, \lambda_0] \ , \quad \omega = (1 + \varepsilon^2 \lambda)\bar{\omega}_\varepsilon \in \mathcal{A} \tag{1.2.26}$$

are still admissible (see (1.2.23)) and, using (1.2.25),

$$\begin{aligned} (1 + \varepsilon^2 \lambda)\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2 \mathscr{A}(\xi) \quad &\Longleftrightarrow \\ \xi := \xi(\lambda) = (1 + \varepsilon^2 \lambda)\mathscr{A}^{-1}\zeta + \lambda \mathscr{A}^{-1}\bar{\mu} \,. \end{aligned} \tag{1.2.27}$$

We shall use the 1-dimensional "parameter" $\lambda \in \Lambda := [-\lambda_0, \lambda_0]$ in order to verify all the non-resonance conditions required for the frequency vector $\omega$ in the proof of Theorem 1.2.1.

For $\varepsilon$ small fixed, we take the vector $\zeta$ such that the direction $\bar{\omega}_\varepsilon$ in (1.2.25) still verifies Diophantine conditions like (1.2.6), (1.2.8) with the different exponents

$$\gamma_1 := \gamma_0/2 \,, \quad \tau_1 := 3\tau_0 + |\mathbb{S}|(|\mathbb{S}| + 1) + 5 > \tau_0 \,, \tag{1.2.28}$$

namely

$$|\bar{\omega}_\varepsilon \cdot \ell| \geq \frac{\gamma_1}{\langle \ell \rangle^{\tau_1}} \,, \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|} \setminus \{0\} \,, \tag{1.2.29}$$

$$\Big| n + \sum_{1 \leq i \leq j \leq |\mathbb{S}|} p_{ij} (\bar{\omega}_\varepsilon)_i (\bar{\omega}_\varepsilon)_j \Big| \geq \frac{\gamma_1}{\langle p \rangle^{\tau_1}} \,, \quad \forall (n, p) \in \mathbb{Z} \times \mathbb{Z}^{\frac{|\mathbb{S}|(|\mathbb{S}|+1)}{2}} \setminus \{0\} \,. \tag{1.2.30}$$

This is possible by Lemma 2.3.1. Actually the vector $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2\zeta$ satisfies (1.2.29)-(1.2.30) for all $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})$ except a small set of measure $O(\varepsilon)$. In (1.2.30), we denote, for $i = 1, \ldots, |\mathbb{S}|$, the $i$-component $(\bar{\omega}_\varepsilon)_i = \mu_{j_i} + \varepsilon^2\zeta_i$, where $j_1, \ldots, j_{|\mathbb{S}|}$ are the tangential sites ordered according to (1.2.3).

**Main result.** We may now rigorously state the main result of this Monograph, concerning existence of quasi-periodic solutions of the nonlinear wave equation (1.1.1). Let us define the Sobolev spaces

$$\begin{aligned}\mathcal{H}^s &:= \mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d; \mathbb{R}) \\ &:= \Big\{ u(\varphi, x) := \sum_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{Z}^d} u_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)} : \|u\|_s^2 := \sum_{i\in\mathbb{Z}^{|\mathbb{S}|+d}} |u_i|^2 \langle i\rangle^{2s} < \infty\,, \\ &\qquad u_{-i} = \bar{u}_i\,,\ \forall\, i := (\ell, j) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{Z}^d \Big\}\end{aligned} \tag{1.2.31}$$

where

$$\langle i\rangle := \max(|\ell|, |j|, 1)\,, \ |j| := \max\{|j_1|, \ldots, |j_d|\}\,.$$

We look for solutions of the equation (1.2.5) in $\mathcal{H}^s$ for some

$$s \geq s_0 > (|\mathbb{S}| + d)/2\,, \tag{1.2.32}$$

so that $\mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|+d}) \hookrightarrow C^0(\mathbb{T}^{|\mathbb{S}|+d})$ and $\mathcal{H}^s$ is an algebra.

**Theorem 1.2.1. (Quasi-periodic solutions for the nonlinear wave equation** (1.1.1)**).** *Fix finitely many tangential sites $\mathbb{S} \subset \mathbb{N}$. Take the multiplicative potential $V(x) \in C^\infty(\mathbb{T}^d, \mathbb{R})$ such that the positivity condition* (1.1.3) *holds, the unperturbed frequency vector $\bar{\mu} \in \mathbb{R}^{|\mathbb{S}|}$ in* (1.2.3) *satisfies the Diophantine conditions* (1.2.6), (1.2.8), *and such that the unperturbed first and second order Melnikov non-resonance conditions* (1.2.7), (1.2.16)-(1.2.19) *hold. Assume also the twist condition* (1.2.12) *and the finitely many non-degeneracy conditions* (1.2.21)-(1.2.22). *Fix $\bar{\omega}_\varepsilon := \bar{\mu} + \varepsilon^2\zeta$, $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})$, as in* (1.2.25) *such that the Diophantine conditions* (1.2.29)-(1.2.30) *hold. There exist $\bar{s} > s_0$ and a Cantor like set $\mathcal{G}_{\varepsilon,\zeta} \subset \Lambda$ (the set $\Lambda$ is fixed in* (1.2.26)*) with asymptotically full measure, i.e.*

$$|\Lambda \setminus \mathcal{G}_{\varepsilon,\zeta}| \to 0 \quad as \quad \varepsilon \to 0\,,$$

*such that, for all $\lambda \in \mathcal{G}_{\varepsilon,\zeta}$, there exists a solution $u_{\varepsilon,\lambda} \in C^\infty$ of* (1.2.5), *even in $\varphi$, with frequency vector*

$$\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon\,, \quad \bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2\zeta\,,$$

*of the form*

$$u_{\varepsilon,\lambda}(\varphi, x) = \sum_{j\in\mathbb{S}} \mu_j^{-1/2}\sqrt{2\xi_j}\cos(\varphi_j)\Psi_j(x) + r_\varepsilon(\varphi, x)\,, \tag{1.2.33}$$

*where* $\xi := \xi(\lambda) \in [1/2, 4]^{|\mathbb{S}|}$ *is given in* (1.2.27) *and* $\|r_\varepsilon\|_{\bar{s}} \to 0$ *as* $\varepsilon \to 0$. *As a consequence* $\varepsilon u_{\varepsilon,\lambda}(\omega t, x)$ *is a quasi-periodic solution of the nonlinear wave equation* (1.1.1) *with frequencies* $\omega = (1 + \varepsilon^2 \lambda)\bar{\omega}_\varepsilon$.

Theorem 1.2.1 is a direct consequence of Theorem 5.1.2.

Let us make some comments on the result.

1. **(Measure estimate of $\mathcal{G}_{\varepsilon,\zeta}$)** The speed of convergence of $|\Lambda \setminus \mathcal{G}_{\varepsilon,\zeta}|$ to 0 does not depend on $\zeta$. More precisely ($\gamma_0, \gamma_1, \tau_0, \tau_1$ being fixed) there is a map $\varepsilon \mapsto b(\varepsilon)$, satisfying $\lim_{\varepsilon \to 0} b(\varepsilon) = 0$, such that, for all $\zeta \in \mathscr{A}([1, 2]^{|\mathbb{S}|})$ so that the vector $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2 \zeta$ satisfies the Diophantine conditions (1.2.29)-(1.2.30), we have the measure estimate $|\Lambda \setminus \mathcal{G}_{\varepsilon,\zeta}| \leq b(\varepsilon)$.

2. **(Density)** Integrating in $\lambda$ along all possible admissible directions $\bar{\omega}_\varepsilon$ in (1.2.25), we deduce the existence of quasi-periodic solutions of (1.1.1) for a set of frequency vectors $\omega$ of positive measure. More precisely, defining the convex subsets of $\mathbb{R}^{|\mathbb{S}|}$,

$$\begin{aligned} &\mathcal{C}_2 := \bar{\mu} + \mathbb{R}_+ \mathcal{C}_1 \,, \\ &\mathcal{C}_1 := \mathscr{A}([1, 2]^{|\mathbb{S}|}) + \Lambda \bar{\mu} := \left\{ \zeta + \lambda \bar{\mu} \ ; \ \zeta \in \mathscr{A}([1, 2]^{|\mathbb{S}|}) \,, \ \lambda \in \Lambda \right\} , \end{aligned} \tag{1.2.34}$$

the set $\Omega$ of the frequency vectors $\omega$ of the quasi-periodic solutions of (1.1.1) provided by Theorem 1.2.1 has Lebesgue density 1 at $\bar{\mu}$ in $\mathcal{C}_2$, *i.e.*

$$\lim_{r \to 0+} \frac{|\Omega \cap \mathcal{C}_2 \cap B(\bar{\mu}, r)|}{|\mathcal{C}_2 \cap B(\bar{\mu}, r)|} = 1 \tag{1.2.35}$$

(see the proof below Theorem 5.1.2). Moreover, we restrict ourself to $\zeta \in \mathscr{A}([1, 2]^{|\mathbb{S}|})$ just to fix the ideas, and we could replace this condition by $\zeta \in \mathscr{A}([r, R]^{|\mathbb{S}|})$, for any $0 < r < R$ (at the cost of stronger smallness conditions for $\lambda_0$ and $\varepsilon$ if $r$ is small and $R$ is large). Therefore we could obtain a similar density result with $\mathcal{C}_2' := \bar{\mu} + \mathscr{A}(\mathbb{R}_+^{|\mathbb{S}|})$ instead of $\mathcal{C}_2$.

3. **(Regularity)** Theorem 1.2.1 also holds if the nonlinearity $g(x, u)$ and the potential $V(x)$ in (1.1.1) are of class $C^q$ for some $q$ large enough, proving the existence of a solution $u_{\varepsilon,\lambda}$ in $\mathcal{H}^{\bar{s}}$, see remark 11.2.9.

4. **(Lipschitz dependence)** The solution $u_{\varepsilon,\lambda}$ is a Lipschitz function of $\lambda \in \mathcal{G}_{\varepsilon,\zeta}$ with values in any $\mathcal{H}^s$, $s \geq \bar{s}$.

Theorem 1.2.3 below proves that, for any choice of finitely many tangential sites $\mathbb{S} \subset \mathbb{N}$, all the non-resonance and non-degeneracy assumptions required in Theorem 1.2.1 are *generically* verified varying the potential $V(x)$ and the function $a(x)$ present in the nonlinearity $g(x, u) = a(x)u^3 + O(u^4)$ in (1.1.2). In order to state a precise result we anticipate the following definition.

**Definition 1.2.2. ($C^\infty$-dense open)** *Given an open subset $\mathcal{U}$ of $H^s(\mathbb{T}^d)$ (resp. $H^s(\mathbb{T}^d)\times H^s(\mathbb{T}^d)$) a subset $\mathcal{V}$ of $\mathcal{U}$ will be called $C^\infty$-dense open in $\mathcal{U}$ if*

1. *$\mathcal{V}$ is open for the topology defined by the $H^s(\mathbb{T}^d)$-norm,*
2. *$\mathcal{V}$ is $C^\infty$-dense in $\mathcal{U}$, in the sense that, for any $w\in\mathcal{U}$, there is a sequence $(h_n)\in C^\infty(\mathbb{T}^d)$ (resp. $C^\infty(\mathbb{T}^d)\times C^\infty(\mathbb{T}^d)$) such that $w+h_n\in\mathcal{V}$, for all $n\in\mathbb{N}$, and $h_n\to 0$ in $H^r$ for any $r\geq 0$.*

Let $s>d/2$ and define the subset of potentials

$$\mathcal{P} := \big\{ V\in H^s(\mathbb{T}^d) \ : \ -\Delta + V(x) > 0 \big\} \tag{1.2.36}$$

which is open in $H^s(\mathbb{T}^d)$ and convex, thus connected. *Fixed* a finite subset $\mathbb{S}\subset\mathbb{N}$ of tangential sites, consider the set $\widetilde{\mathcal{G}}$ of potentials $V(x)$ and functions $a(x)$ such that the conditions required in Theorem 1.2.1 hold, namely

$$\begin{aligned}\widetilde{\mathcal{G}} := \Big\{ &(V(x),a(x))\in\mathcal{P}\times H^s(\mathbb{T}^d) \ : \ \text{there are } \tau_0,\gamma_0>0 \text{ such that} \\ &(1.2.6)\text{-}(1.2.8), (1.2.12), (1.2.16)\text{-}(1.2.19), (1.2.21)\text{-}(1.2.22) \text{ hold}\Big\}.\end{aligned} \tag{1.2.37}$$

Given a subspace $E$ of $L^2(\mathbb{T}^d)$ we denote by $E^{\perp_{L^2}}$ its orthogonal with respect to the $L^2$ scalar product.

**Theorem 1.2.3. (Genericity)** *Let $s>d/2$. The set*

$$\widetilde{\mathcal{G}}\bigcap\big(C^\infty(\mathbb{T}^d)\times C^\infty(\mathbb{T}^d)\big) \quad is \quad C^\infty\text{−dense } in \ \big(\mathcal{P}\cap C^\infty(\mathbb{T}^d)\big)\times C^\infty(\mathbb{T}^d) \tag{1.2.38}$$

*where $\mathcal{P}\subset H^s(\mathbb{T}^d)$ is the open and connected set of potentials $V(x)$ defined in (1.2.36).*

*More precisely, there is a $C^\infty$-dense open subset $\mathcal{G}$ of $\mathcal{P}\times H^s(\mathbb{T}^d)$ and a $|\mathbb{S}|$-dimensional linear subspace $E$ of $C^\infty(\mathbb{T}^d)$ such that, for all $v_2(x)\in E^{\perp_{L^2}}\cap H^s(\mathbb{T}^d)$, $a(x)\in H^s(\mathbb{T}^d)$, the Lebesgue measure (on the finite dimensional space $E\simeq\mathbb{R}^{|\mathbb{S}|}$)*

$$\big|\big\{ v_1\in E \ : \ (v_1+v_2,a)\in\mathcal{G}\setminus\widetilde{\mathcal{G}}\big\}\big| = 0\,. \tag{1.2.39}$$

Theorem 1.2.3 is proved in Chapter 12.

In order to introduce the reader to the topic, we first provide a non technical survey about the main methods and results in KAM theory for PDEs.

As already mentioned, in a Newton-Nash-Moser iterative scheme, a key step for the existence proof of quasi-periodic solutions consists in the analysis of the linearized operators obtained at each step of the iteration, and proving its approximate invertibility, for most values of suitable parameters, with quantitative estimates for the inverse in high norms. For achieving this task two main approaches have been developed:

1. the “reducibility” approach, that we describe in section 1.3;
2. the “multiscale” approach, presented in section 1.4.

In section 1.5, we shall provide a detailed account of the proof of Theorem 1.2.1.

## 1.3 The reducibility approach to KAM for PDEs

The goal of this section is to present the perturbative reducibility approach for a time quasi-periodic linear operator (subsection 1.3.1) and then describe its main applications to KAM theory for PDEs (subsection 1.3.2).

**Transformation laws.** Consider a quasi-periodically time dependent linear system

$$u_t + A(\omega t)u = 0\,, \quad u \in H\,, \tag{1.3.1}$$

where, for any $\varphi \in \mathbb{T}^\nu$, $\nu \in \mathbb{N}$, $A(\varphi)$ is a linear operator acting on a phase space $H$, which may be a finite or infinite dimensional Hilbert space with scalar product $\langle\, , \rangle$, and $\omega \in \mathbb{R}^\nu \backslash \{0\}$ is the frequency vector. We suppose that $\omega$ is a nonresonant vector, i.e. $\omega \cdot \ell \neq 0$, $\forall \ell \in \mathbb{Z}^\nu \setminus \{0\}$, thus the linear flow $\{\omega t\}_{t\in\mathbb{R}}$ densily fills the torus $\mathbb{T}^\nu$. Under a quasi-periodically time dependent transformation

$$u = \Phi(\omega t)[v] \tag{1.3.2}$$

where $\Phi(\varphi) : H \to H$, $\varphi \in \mathbb{T}^\nu$, are invertible linear operators of the phase space (or of dense subspaces), system (1.3.1) transforms into

$$v_t + B(\omega t)v = 0 \tag{1.3.3}$$

with the new linear operator

$$B(\varphi) = \Phi^{-1}(\varphi)(\omega \cdot \partial_\varphi \Phi)(\varphi) + \Phi(\varphi)^{-1}A(\varphi)\Phi(\varphi)\,. \tag{1.3.4}$$

**Remark 1.3.1.** *Suppose that $H$ is endowed with a symplectic form $\Omega$ defined by $\Omega(u, v) := \langle J^{-1}u, v\rangle$, $\forall u, v \in H$, where $J$ is an antisymmetric, non-degenerate operator. If $A(\omega t)$ is Hamiltonian, namely $A(\omega t) = JS(\omega t)$ where $S(\omega t)$ is a (possibly unbounded) self-adjoint operator, and $\Phi(\omega t)$ is symplectic, then the new operator $B(\omega t)$ is Hamiltonian as well, see Lemma 3.2.3.*

**Reducibility.** If the operator $B$ in (1.3.3) is a diagonal, time independent operator, i.e.

$$B(\omega t) = B = \mathrm{Diag}_j(b_j) \tag{1.3.5}$$

in a suitable basis of $H$, then (1.3.3) reduces to the decoupled scalar linear ordinary differential equations

$$\dot v_j + b_j v_j = 0 \tag{1.3.6}$$

where $(v_j)$ denote the coordinates of $v$ in the basis of eigenvectors of $B$. Then (1.3.3) is integrated in a straightforward way,

$$v_j(t) = e^{-b_j t} v_j(0)\,,$$

and all the solutions of system (1.3.1) are obtained via the change of variable (1.3.2). We say that (1.3.1) has been *reduced* to constant coefficients by the change of variable $\Phi$.

**Remark 1.3.2.** *If all the* $b_j$ *in* (1.3.6) *are purely imaginary, then the linear system* (1.3.1) *is stable (in the sense of Lyapunov), otherwise, it is unstable.*

We shall also say that system (1.3.1) is reducible if $B$ is a constant coefficient block-diagonal operator, i.e. $b_j$ in (1.3.5)-(1.3.6) are finite dimensional matrices, constant in time. The spectrum of each matrix $b_j$ determines the stability/instability properties of the system (1.3.1).

If $\omega \in \mathbb{R}$ (time-periodic forcing) and the phase space $H$ is finite dimensional, the classical Floquet theory proves that any time periodic linear system (1.3.1) is reducible, see e.g. [59], Chapter I. On the other hand, if $\omega \in \mathbb{R}^\nu$, $\nu \geq 2$, it is known that there exist pathological non reducible linear systems, see e.g. [44]-Chapter 1.

If $A(\omega t)$ is a small perturbation of a constant coefficient operator, perturbative algorithms for reducibility can be implemented. In the next subsection we describe this strategy in the simplest setting. This approach was systematically adopted by Moser [99] for developing finite dimensional KAM theory (in a much more general context).

### 1.3.1 Perturbative reducibility

Consider a quasi-periodic operator

$$\omega \cdot \partial_\varphi + A(\varphi) \qquad \text{where} \qquad A(\varphi) = D + R(\varphi)\,, \quad \varphi \in \mathbb{T}^\nu\,, \tag{1.3.7}$$

is a perturbation of a diagonal operator

$$D = \mathrm{Diag}(\mathrm{i}d_j)_{j\in\mathbb{Z}} = \mathrm{Op}(\mathrm{i}d_j)\,, \quad d_j \in \mathbb{R}\,, \tag{1.3.8}$$

where the eigenvalues $\mathrm{i}d_j$ are simple and constant in $\varphi$. The $\varphi$-dependent family of operators $R(\varphi)$ acting on $H$ is a small perturbation of $D$.

**Remark 1.3.3.** *We suppose that the* $d_j$ *are real because this is the common situation arising for PDEs, i.e.* $u = 0$ *is an elliptic equilibrium for the linear system* $u_t + Du = 0$*. On the other hand, if some* $\mathrm{Im}\, d_j \neq 0$*, then there are hyperbolic directions which do not create resonance phenomena, and perturbative reducibility theory is easier. For nonlinear systems, this case corresponds to the search of whiskered tori, see e.g. [73] for finite dimensional systems, and [65] for PDEs.*

We look for a transformation $\Phi(\varphi)$, $\varphi \in \mathbb{T}^\nu$, of the phase space $H$, as in (1.3.2) which removes from $R(\varphi)$ the angles $\varphi$ up to terms of size $\sim O(|R|^2)$. We present below only the algebraic aspect of the reducibility scheme, without specifying the norms.

For computational purposes, it is convenient to transform the linear system (1.3.1) under the flow $\Phi_F(\varphi, \tau)$ generated by an auxiliary linear equation

$$\partial_\tau \Phi_F(\varphi, \tau) = F(\varphi)\Phi_F(\varphi, \tau)\,, \quad \Phi_F(\varphi, 0) = \mathrm{Id}\,, \tag{1.3.9}$$

generated by a linear operator $F(\varphi)$ to be chosen (which could also be $\tau$-dependent). This amounts to computing the Lie derivative of $A(\varphi)$ in the direction of the vector field $F(\varphi)$. Notice that, if $F(\varphi)$ is bounded, then the flow (1.3.9) is well posed. This is always the case for a finite dimensional system, but it may be an issue for infinite dimensional systems.

Given a linear operator $A_0(\varphi)$, the conjugated operator under the flow $\Phi_F(\varphi,\tau)$ generated by (1.3.9),

$$A(\varphi,\tau) := \Phi_F(\varphi,\tau)A_0(\varphi)\Phi_F(\varphi,\tau)^{-1}\,,$$

satisfies the Heisenberg equation

$$\begin{cases} \partial_\tau A(\varphi,\tau) = [F(\varphi), A(\varphi,\tau)] \\ A(\varphi,\tau)_{|\tau=0} = A_0(\varphi) \end{cases} \tag{1.3.10}$$

where $[A,B] := A\circ B - B\circ A$ denotes the commutator between two linear operators $A, B$. Then, by a Taylor expansion, using (1.3.10), we obtain the formal Lie expansion

$$A(\varphi,\tau)_{|\tau=1} = A_0(\varphi) + \mathrm{Ad}_F A_0 + \frac{1}{2}\mathrm{Ad}_F^2 A_0 + \dots \tag{1.3.11}$$

where $\mathrm{Ad}_F[\,\cdot\,] := [F,\,\cdot\,]$. One may expect this expansion to be certainly convergent if $F$ and $A_0$ are bounded and $F$ is small, in suitable norms, because the adjoint action produces, in such a case, bounded operators $\mathrm{Ad}_F^k[A_0]$, $k\in\mathbb{N}$, with smaller and smaller size.

Conjugating (1.3.7) under the flow generated by (1.3.9) we then obtain an operator of the form

$$\omega\cdot\partial_\varphi + D - \omega\cdot\partial_\varphi F(\varphi) + [F(\varphi), D] + R(\varphi) + \text{ smaller terms}\dots. \tag{1.3.12}$$

We want to choose $F(\varphi)$ in such a way to solve the "homological" equation

$$-\omega\cdot\partial_\varphi F(\varphi) + R(\varphi) + [F(\varphi), D] = [R] \tag{1.3.13}$$

where

$$[R] := \mathrm{Diag}_j(\widehat{R}^j_j(0))\,, \quad \widehat{R}^j_j(0) := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} R^j_j(\varphi)\,d\varphi\,, \tag{1.3.14}$$

is the normal form part of the operator $R(\varphi)$, independent of $\varphi$, that we can not eliminate. Representing the linear operators $F(\varphi) = (F^j_k(\varphi))_{j,k\in\mathbb{Z}}$ and $R(\varphi) = (R^j_k(\varphi))_{j,k\in\mathbb{Z}}$ as matrices, and computing the commutator with the diagonal operator $D$ in (1.3.8) we obtain that (1.3.13) is represented as

$$-\omega\cdot\partial_\varphi F^j_k(\varphi) + R^j_k(\varphi) + \mathrm{i}(d_j - d_k)F^j_k(\varphi) = [R]^j_k\,,$$

and, performing the Fourier expansion in $\varphi$,

$$F^j_k(\varphi) = \sum_{\ell\in\mathbb{Z}^\nu}\widehat{F}^j_k(\ell)e^{\mathrm{i}\ell\cdot\varphi}\,, \quad R^j_k(\varphi) = \sum_{\ell\in\mathbb{Z}^\nu}\widehat{R}^j_k(\ell)e^{\mathrm{i}\ell\cdot\varphi}\,,$$

it reduces to the infinitely many scalar equations

$$-\mathrm{i}\omega\cdot\ell\,\widehat{F}_k^j(\ell)+\widehat{R}_k^j(\ell)+\mathrm{i}(d_j-d_k)\widehat{F}_k^j(\ell)=[R]_k^j\delta_{\ell,0}\,,\quad j,k\in\mathbb{Z}\,,\quad \ell\in\mathbb{Z}^\nu\,, \tag{1.3.15}$$

where $\delta_{\ell,0}:=1$ if $\ell=0$ and zero otherwise. Assuming the so called *second-order Melnikov non-resonance conditions*

$$|\omega\cdot\ell+d_j-d_k|\geq\frac{\gamma}{\langle\ell\rangle^\tau}\,,\quad \forall(\ell,j,k)\neq(0,j,j)\,, \tag{1.3.16}$$

for some $\gamma,\tau>0$, we can define the solution of the homological equations (1.3.15) (see (1.3.14))

$$\widehat{F}_k^j(\ell):=\begin{cases}\dfrac{-\widehat{R}_k^j(\ell)}{\mathrm{i}(-\omega\cdot\ell+d_j-d_k)} & \forall(\ell,j,k)\neq(0,j,j)\\ 0 & \forall(\ell,j,k)=(0,j,j)\,.\end{cases} \tag{1.3.17}$$

Therefore the transformed operator (1.3.12) becomes

$$\omega\cdot\partial_\varphi+D_++\text{smaller terms} \tag{1.3.18}$$

where

$$D_+:=D+[R]=(\mathrm{i}d_j+[R]_j^j)_{j\in\mathbb{Z}} \tag{1.3.19}$$

is the new diagonal operator, constant in $\varphi$. We can iterate this step to reduce also the small terms of order $O(|R|^2\gamma^{-1})$ which are left in (1.3.18), and so on. Notice that, if $R(\varphi)$ is a bounded operator and depends smoothly enough on $\varphi$, then $F(\varphi)$ defined in (1.3.17), with the denominators satisfying (1.3.16), is bounded as well and thus (1.3.9) certainly defines a flow by standard Banach space ODE techniques. On the other hand the loss of time derivatives induced on $F(\varphi)$ by the divisors in (1.3.16) can be recovered by a smoothing procedure in the angles $\varphi$, like a truncation in Fourier space.

**Remark 1.3.4.** *If the operator $A(\varphi)$ in* (1.3.7) *is Hamiltonian, as defined in remark 1.3.1 (with a symplectic form $J$ which commutes with $D$), then $F(\varphi)$ is Hamiltonian, its flow $\Phi_F(\varphi,\tau)$ is symplectic, and the new operator in* (1.3.18) *is Hamiltonian as well.*

In order to continue the iteration one also needs to impose non-resonance conditions as in (1.3.16) at each step and therefore we need information about the perturbed normal form $D_+$ in (1.3.19), in particular the asymptotic of $[R]_j^j$. If these steps work, then, after an infinite iteration, one could conjugate the quasi-periodic operator (1.3.7) to a diagonal, constant in $\varphi$, operator of the form

$$\omega\cdot\partial_\varphi+\mathrm{Diag}_j(\mathrm{i}d_j^\infty)\,,\quad \mathrm{i}d_j^\infty=\mathrm{i}d_j+[R]_j^j+\dots\,. \tag{1.3.20}$$

At this stage, imposing the first order Melnikov non-resonance conditions

$$|\omega \cdot \ell + d_j^\infty| \geq \frac{\gamma}{\langle \ell \rangle^\tau}, \quad \forall \ell, j,$$

the diagonal linear operator (1.3.20) is invertible with an inverse which loses $\tau$ time-derivatives. Verifying that all the changes of coordinates that have been constructed iteratively to conjugate (1.3.7) to (1.3.20) map spaces of high regularity in itselfs, this approach finally enables to prove the existence of an inverse of the initial quasi-periodic linear operator (1.3.7) which satisfies tame estimates in high norms (with loss of $\tau$ derivatives).

This is the essence of the Newton-Nash-Moser-KAM perturbative reducibility scheme, that has been used for proving KAM results for 1d NLW and NLS equations with Dirichlet boundary conditions in [93], [115], [103], as we shall describe in the next section.

The following questions arise naturally:

1. *What happens if the eigenvalues $d_j$ are multiple ?* This is the common situation for 1-$d$ PDEs with periodic boundary conditions or in higher space dimensions. In such a case it is conceivable to reduce $\omega \cdot \partial_\varphi + A(\varphi)$ to a block-diagonal normal form linear system of the form (1.3.20) where $d_\infty^j$ are finite dimensional matrices.

2. *What happens if the operator $R(\varphi)$ in* (1.3.7) *is unbounded ?* This is the common situation for PDEs with nonlinearities which contain derivatives. In such a case also the operator $F(\varphi)$ defined in (1.3.17) is unbounded and therefore (1.3.9) could not define a flow.

3. *What happens if, instead of the Melnikov non-resonance conditions* (1.3.16)*, we have only*
$$|\omega \cdot \ell + d_j - d_k| \geq \frac{\gamma}{\langle \ell \rangle^\tau \langle j \rangle^{\mathtt{d}} \langle k \rangle^{\mathtt{d}}}, \quad \forall (\ell, j, k) \neq (0, j, j), \tag{1.3.21}$$
*for some $\mathtt{d} > 0$, which induce a loss of space derivatives ?* This is the common situation when the dispersion relation $d_j \sim j^\alpha$, $\alpha < 1$, has a sublinear growth. Also in this situation the operator $F(\varphi)$ defined in (1.3.17) would be unbounded. This situation appears for example for pure gravity water waves equations.

We describe below some answers to the above questions.

### 1.3.2 Reducibility results

We now present the main results about KAM theory for PDEs based on the reducibility scheme described in the previous section.

### KAM for $1d$ NLW and NLS with Dirichlet boundary conditions

The iterative reducibility scheme outlined in section 1.3.1 has been effectively implemented by Kuksin [93] and Wayne [115] for proving existence of quasi-periodic solutions of 1-$d$ semilinear wave

$$y_{tt} - y_{xx} + V(x)y + \varepsilon f(x, y) = 0\,, \quad y(0) = y(\pi) = 0\,, \tag{1.3.22}$$

and Schrödinger equations

$$\mathrm{i}u_t - u_{xx} + V(x)u + \varepsilon f(|u|^2)u = 0\,, \quad u(0) = u(\pi) = 0\,, \tag{1.3.23}$$

with Dirichlet boundary conditions. These equations are regarded as a perturbation of the linear PDEs

$$y_{tt} - y_{xx} + V(x)y = 0\,, \quad \mathrm{i}u_t - u_{xx} + V(x)u = 0\,, \tag{1.3.24}$$

which depend on the potential $V(x)$, used as a parameter.

The linearized operators obtained at an approximate quasi-periodic solution are, for NLS,

$$h \mapsto \mathrm{i}\omega \cdot \partial_\varphi h - h_{xx} + V(x)h + \varepsilon q(\varphi, x)h + \varepsilon p(\varphi, x)\bar{h} \tag{1.3.25}$$

with $q(\varphi, x) \in \mathbb{R}$, $p(\varphi, x) \in \mathbb{C}$, and, for NLW,

$$y \mapsto (\omega \cdot \partial_\varphi)^2 y - y_{xx} + V(x)y + \varepsilon a(\varphi, x)y\,, \tag{1.3.26}$$

with $a(\varphi, x) \in \mathbb{R}$, that, in the complex variable

$$h = D_V^{\frac{1}{2}} y + \mathrm{i}D_V^{-\frac{1}{2}} y_t\,, \quad D_V := \sqrt{-\Delta + V(x)}\,,$$

assumes the form

$$h \mapsto \omega \cdot \partial_\varphi h + \mathrm{i}D_V h + \mathrm{i}\frac{\varepsilon}{2} D_V^{-\frac{1}{2}} a(\varphi, x) D_V^{-\frac{1}{2}}(h + \bar{h})\,. \tag{1.3.27}$$

Coupling these equations with their complex conjugated component, we have to invert the quasi-periodic operators, acting on $\begin{pmatrix} h \\ \bar{h} \end{pmatrix}$, given, for NLS, by

$$\omega \cdot \partial_\varphi + \begin{pmatrix} \mathrm{i}\partial_{xx} - \mathrm{i}V(x) & 0 \\ 0 & -\mathrm{i}\partial_{xx} + \mathrm{i}V(x) \end{pmatrix} - \varepsilon \begin{pmatrix} \mathrm{i}q(\varphi, x) & \mathrm{i}p(\varphi, x) \\ -\mathrm{i}\bar{p}(\varphi, x) & -\mathrm{i}q(\varphi, x) \end{pmatrix} \tag{1.3.28}$$

and, for NLW,

$$\omega \cdot \partial_\varphi + \begin{pmatrix} \mathrm{i}D_V & 0 \\ 0 & -\mathrm{i}D_V \end{pmatrix} + \mathrm{i}\frac{\varepsilon}{2} D_V^{-\frac{1}{2}} \begin{pmatrix} a(\varphi, x) & a(\varphi, x) \\ -a(\varphi, x) & -a(\varphi, x) \end{pmatrix} D_V^{-\frac{1}{2}} \tag{1.3.29}$$

which have the form (1.3.7) with an operator $R(\varphi)$ which is bounded. Actually notice that for NLW the perturbative term in (1.3.29) is also 1-smoothing. Moreover the eigenvalues $\mu_j^2$, $j \in \mathbb{N}\backslash\{0\}$, of the Sturm-Liouville operator $-\partial_{xx} + V(x)$ with Dirichlet boundary conditions are simple and the quasi-periodic operators (1.3.28) and (1.3.29) take the form (1.3.7)-(1.3.8), where the eigenvalues of $D$ are

$$\pm \mathrm{i}\mu_j^2\,,\ \mu_j^2 \sim j^2,\ \text{for NLS}\,, \quad \pm \mathrm{i}\mu_j\,,\ \mu_j \sim j\,,\ \text{for NLW}\ .$$

Then it is not hard to impose second order Melnikov non-resonance conditions as in (1.3.16). In view of these observations, it is possible to implement the KAM reducibility scheme presented above to prove the existence of quasi-periodic solutions for (1.3.22)-(1.3.23) (actually the KAM iteration in [93], [115] is a bit different but the previous argument catches its essence).

Later on these results have been extended in Kuksin-Pöschel [96] to parameter independent Schrödinger equations

$$\begin{cases} \mathrm{i}u_t = u_{xx} + f(|u|^2)u\,, \\ u(0) = u(\pi) = 0\,, \end{cases} \qquad \text{where} \quad f(0) = 0\,,\ f'(0) \neq 0\,, \tag{1.3.30}$$

and in Pöschel [104] to nonlinear Klein-Gordon equations

$$y_{tt} - y_{xx} + my = y^3 + \text{h.o.t}\,, \quad y(0) = y(\pi) = 0\,. \tag{1.3.31}$$

The main new difficulty of these equations is that the linear equations

$$\mathrm{i}u_t = u_{xx}\,, \quad y_{tt} - y_{xx} + my = 0\,,$$

have resonant invariant tori. Actually all the solutions of the first equation,

$$u(t,x) = \sum_{j\in\mathbb{Z}} u_j(0)e^{\mathrm{i}j^2t}e^{\mathrm{i}jx}\,, \tag{1.3.32}$$

are $2\pi$-periodic in time (for this reason (1.3.30) is called a completely resonant PDE) and the Klein-Gordon linear frequencies $\sqrt{j^2+m}$ may be resonant for several values of the mass $m$. The new key idea in [104], [96] is to compute precisely how the nonlinearity in (1.3.30)-(1.3.31) modulates the tangential and normal frequencies of the expected quasi-periodic solutions. In particular, a Birkhoff normal form analysis enables to prove that the tangential frequencies vary diffeomorphically with the "amplitudes" of the solutions. This non-degeneracy property allows then to prove that the Melnikov non-resonance conditions are satisfied for most amplitudes. We notice however the following difficulty for the equations (1.3.30)-(1.3.31) which is not present for (1.3.22)-(1.3.23): the frequency vector $\omega$ may satisfy only a Diophantine condition (1.2.6) with a constant $\gamma_0$ which tends to 0 as the solution tends to 0 (the linear frequencies in (1.3.32) are integers), and similarly for the second order Melnikov conditions (1.3.16). Notice that the remainders in (1.3.18) have size $O(|R|^2\gamma^{-1})$ and, as a consequence, careful estimates have to be performed to overcome this "singular" perturbation issue.

## Periodic boundary conditions $x \in \mathbb{T}$

The above results do not apply for periodic boundary conditions $x \in \mathbb{T}$ because two eigenvalues of $-\partial_{xx} + V(x)$ coincide (or are too close), and thus the second order Melnikov non resonance conditions (1.3.16) are violated. This is the first instance where the difficulty mentioned in item 1 appears.

Historically this difficulty was first solved by Craig-Wayne [51] and Bourgain [34] developing a multiscale approach, based on repeated use of the "resolvent identity" in the spirit of the work [66] by Frölich-Spencer for Anderson localization. This approach does not require the second order Melnikov non-resonance conditions. We describe it in section 1.4. Developments of this multiscale approach are the basis of the present Monograph.

The KAM reducibility approach was extended later by Chierchia-You [45] for semilinear wave equations like (1.3.22) with periodic boundary conditions. Because of the near resonance between pairs of frequencies, the linearized operators (1.3.26)-(1.3.27) are reduced to a diagonal system of $2 \times 2$ self-adjoint matrices, namely of the form (1.3.20) with $d_j^\infty \in \mathrm{Mat}(2 \times 2; \mathbb{C})$, by requiring at each step second-order Melnikov non-resonance conditions of the form

$$|\omega \cdot \ell + d_j - d_k| \geq \frac{\gamma}{\langle \ell \rangle^\tau}, \quad \forall (\ell, j, k) \neq (0, j, \pm j). \tag{1.3.33}$$

Notice that we do not require in (1.3.33) non-resonance conditions for $\ell = 0$ and $k = \pm j$. Since NLW is a second order equation, the nonlinear perturbative part of its Hamiltonian vector field is regularizing of order 1 (it gains one space derivative), see (1.3.29), and this is sufficient to prove that the perturbed frequencies of (1.3.29) satisfy an asymptotic estimate like

$$\mu_j(\varepsilon) = \mu_j + O(\varepsilon |j|^{-1}) = |j| + O(|j|^{-1})$$

as $|j| \to +\infty$, where $\mu_j^2$ denote the eigenvalues of the Sturm-Liouville operator $-\partial_{xx} + V(x)$. Thanks to this asymptotic expansion it is sufficient to impose, for each $\ell \in \mathbb{Z}^\nu$, only finitely many second order Melnikov non-resonance conditions as (1.3.33), by requiring first order Melnikov conditions like

$$|\omega \cdot \ell + h| \geq \gamma \langle \ell \rangle^{-\tau}, \tag{1.3.34}$$

for all $(\ell, h) \in (\mathbb{Z}^\nu \times \mathbb{Z}) \setminus (0,0)$. Indeed, if $|j|, |k| > C|\ell|^\tau \gamma^{-1}$, for an appropriate constant $C > 0$, we get

$$\begin{aligned} |\omega \cdot \ell + \mu_j(\varepsilon) - \mu_k(\varepsilon)| &\geq |\omega \cdot \ell + |j| - |k|| - O(1/\min(|k|, |j|)) \\ &\overset{(1.3.34)}{\geq} \gamma \langle \ell \rangle^{-\tau} - O(1/\min(|k|, |j|)) \geq \frac{\gamma}{2} \langle \ell \rangle^{-\tau} \end{aligned} \tag{1.3.35}$$

noting that $|j| - |k|$ is an integer. Moreover if $||k| - |j|| \geq C|\ell|$ for another appropriate constant $C > 0$ then $|\omega \cdot \ell + \mu_j(\varepsilon) - \mu_k(\varepsilon)| \geq |\ell|$. Hence, under (1.3.34), for $\ell$ given, the second order Melnikov conditions with time-index $\ell$ are automatically satisfied for all $(j, k)$ except a finite number.

**Remark 1.3.5.** *If the Hamiltonian nonlinearity does not depend on the space variable $x$, the equations* (1.3.30)-(1.3.31) *are invariant under space translations and therefore possess a prime integral by Noether Theorem. Geng-You [67], [69] were the first to exploit such conservation law, which is preserved along the KAM iteration, to fulfill the non-resonance conditions. The main observation is that such symmetry enables to prove that many monomials are a-priori never present along the KAM iteration. In particular, this symmetry removes the degeneracy produced by the multiple normal frequencies.*

For semilinear Schrödinger equations like (1.3.23), (1.3.30), the nonlinear vector field is not smoothing. Correspondingly, notice that in the linearized operator (1.3.28) the remainder $R(\varphi)$ is a matrix of multiplication operators. In such a case the basic perturbative estimate for the eigenvalues gives

$$\mu_j(\varepsilon) = \mu_j^2 + O(\varepsilon) \,,$$

which is not sufficient to verify second order Melnikov non resonance conditions like (1.3.33), in particular

$$|\omega \cdot \ell + \mu_j(\varepsilon) - \mu_{-j}(\varepsilon)| \geq \frac{\gamma}{\langle \ell \rangle^{\tau}} \,, \quad \forall \ell \in \mathbb{Z}^{\nu} \setminus \{0\} \,, \ j \in \mathbb{N} \,,$$

for most values of the parameters.

The first KAM reducibility result for NLS with $x \in \mathbb{T}$ has been proved by Eliasson-Kuksin in [58] as a particular case of a much more general result valid for tori $\mathbb{T}^d$ of any space dimension $d \geq 1$, that we discuss below. The key point is to extract, using the notion of Töplitz-Lipschitz matrices, the first order asymptotic expansion of the perturbed eigenvalues. For perturbations 1-dimensional Schrödinger equations another recent approach to obtain the improved asymptotics of the perturbed frequencies, i.e.

$$\mu_j(\varepsilon) = j^2 + c + O(\varepsilon/|j|) \,,$$

for some constant $c$ independent of $j$, which allows to verify the second order Melnikov non-resonance conditions (1.3.33), is developed in Berti-Kappeler-Montalto [28] via a regularization technique based on pseudo-differential ideas, that we explain below. The approach in [28] applies to semilinear perturbations (also $x$-dependent) of any large "finite gap" solution of

$$\mathrm{i} u_t = -\partial_{xx} u + |u|^2 u + \varepsilon f(x, u) \,, \quad x \in \mathbb{T} \,.$$

Let us explain the term "finite gap" solutions. The $1d$-cubic NLS

$$\mathrm{i} u_t = -\partial_{xx} u + |u|^2 u \,, \quad x \in \mathbb{T} \,, \tag{1.3.36}$$

possesses global analytic action-angle variables, in the form of Birkhoff coordinates, see [74], and the whole infinite dimensional phase space is foliated by quasi-periodic -called "finite gap" solutions- and almost-periodic solutions. The Birkhoff coordinates are a cartesian smooth version of the action-angle variables to avoid the singularity when one action component vanishes, i.e. close to the elliptic equilibrium $u = 0$. This situation generalizes what happens for a finite dimensional Hamiltonian system in $\mathbb{R}^{2n}$ which possesses $n$-independent prime integrals in involution. According to the celebrated Liouville-Arnold theorem (see e.g. [4]), in suitable local symplectic angle-action variables $(\theta, I) \in \mathbb{T}^n \times \mathbb{R}^n$, the integrable Hamiltonian $H(I)$ depends only on the actions and the dynamics is described by

$$\dot{\theta} = \partial_I H(I)\,, \quad \dot{I} = 0\,.$$

Thus the phase space is foliated by the invariant tori $\mathbb{T}^n \times \{\xi\}$, $\xi \in \mathbb{R}^n$, filled by the quasi-periodic solutions $\theta(t) = \theta_0 + \omega(\xi)t$, $I(t) = \xi$, with frequency vector $\omega(\xi) = (\partial_I H)(\xi)$. The analogous construction close to an elliptic equilibrium, where the action-angle variables become singular, is provided by the Rüssmann-Vey-Ito theorem [109], [111], [83], see [90] for an introduction.

Other integrable PDEs which possess Birkhoff coordinates are KdV [90] and mKdV [91], see Appendix A.4.

**Remark 1.3.6.** *The Birkhoff normal form construction of [96] discussed for the NLS equation* (1.3.30) *provides, close to* $u = 0$*, an approximation of the global Birkhoff coordinates of the* 1*d-cubic NLS.*

## Space multidimensional PDEs

For space multidimensional PDEs the reducibility approach has been first worked out for semilinear Schrödinger equations

$$-\mathrm{i}u_t = -\Delta u + V * u + \varepsilon \partial_{\bar{u}} F(x, u, \bar{u})\,, \quad x \in \mathbb{T}^d\,, \tag{1.3.37}$$

with a convolution potential by Eliasson-Kuksin [58], [57]. This is a much more difficult situation with respect to the 1-$d$-case because the eigenvalues of $-\Delta + V(x)$ appear in clusters of unbounded size. This is the difficulty mentioned in item 1. In such a case the reducibility result that one could look for is to block-diagonalize the quasi-periodic Schrödinger linear operator

$$h \mapsto \mathrm{i}\omega \cdot \partial_\varphi h - \Delta h + V * h + \varepsilon q(\varphi, x) h + \varepsilon p(\varphi, x) \bar{h}\,,$$

i.e. to obtain an operator as (1.3.20) with finite dimensional blocks $d_j^\infty$, which are self-adjoint matrices of increasing dimension as $j \to +\infty$. The convolution potential $V$ plays the role of "external parameters". Eliasson-Kuksin introduced in [58] the notion of Töplitz-Lipschitz matrices in order to extract asymptotic information on the eigenvalues, and so verify the second order Melnikov non resonance conditions. The quasi-periodic solutions of (1.3.37) obtained in [58] are linearly stable.

**Remark 1.3.7.** *The reducibility techniques in [58] enable to prove a stability result for all the solutions of the linear Schrödinger equation*

$$\mathrm{i}u_t = \Delta u + \varepsilon V(\omega t, x)\,, \quad x \in \mathbb{T}^d\,,$$

*with a small quasi-periodic analytic potential $V(\omega t, x)$. For all frequencies $\omega \in \mathbb{R}^\nu$, except a set of measure tending to $0$ as $\varepsilon \to 0$, the Sobolev norms of any solution $u(t,\cdot)$ satisfy*

$$\|u(t,\cdot)\|_{H^s(\mathbb{T}^d)} \sim \|u(0,\cdot)\|_{H^s(\mathbb{T}^d)}\,, \quad \forall t \in \mathbb{R}\,.$$

Subsequently for the cubic NLS equation

$$\mathrm{i}u_t = -\Delta u + |u|^2 u\,, \quad x \in \mathbb{T}^2\,, \tag{1.3.38}$$

which is parameter independent and completely resonant, Geng-Xu-You [70] proved a KAM result using a Birkhoff normal form analysis. We remark that the Birkhoff normal form of (1.3.38) is not-integrable, unlike in space dimension $d = 1$, causing additional difficulties with respect to [96].

For completely resonant NLS equations in any space dimension and a polynomial nonlinearity

$$\mathrm{i}u_t = -\Delta u + |u|^{2p} u\,, \quad p \in \mathbb{N}\,,\ x \in \mathbb{T}^d\,, \tag{1.3.39}$$

Procesi-Procesi [105] realized a systematic study of the resonant Birkhoff normal form and, using the notion of quasi-Töplitz matrices developed in Procesi-Xu [108], proved in [106], [107], the existence of reducible quasi-periodic solutions of (1.3.39).

**Remark 1.3.8.** *The resonant Birkhoff normal form of* (1.3.38) *is exploited in [46], [79], to construct chaotic orbits with a growth of the Sobolev norm. This "norm inflation" phenomenon is an analogue of the Arnold diffusion problem [3] for finite dimensional Hamiltonian systems. Similar results for the NLS equation* (1.3.39) *have been proved in [78].*

KAM results have been proved for parameter dependent beam equations by Geng-You [68] and, more recently, in Eliasson-Grébert-Kuksin [56] for multidimensional beam equations like

$$u_{tt} + \Delta^2 u + mu + \partial_u G(x, u) = 0\,, \quad x \in \mathbb{T}^d\,, \quad u \in \mathbb{R}\,.$$

We also mention the work [75] of Grébert-Paturel concerning the existence of reducible quasi-periodic solutions of Klein-Gordon equations on the sphere $\mathbb{S}^d$,

$$u_{tt} - \Delta u + mu + \delta M_\rho u + \varepsilon g(x, u) = 0\,, \quad x \in \mathbb{S}^d\,, \quad u \in \mathbb{R}\,, \tag{1.3.40}$$

where $\Delta$ is the Laplace-Beltrami operator and $M_\rho$ is a Fourier multiplier.

On the other hand, if $x \in \mathbb{T}^d$, the infinitely many second order Melnikov conditions, required to block-diagonalize the quasi-periodic linear wave operator

$$(\omega \cdot \partial_\varphi)^2 - \Delta + V * + \varepsilon a(\varphi, x) , \quad x \in \mathbb{T}^d ,$$

are violated for $d \geq 2$, and no reducibility results are available so far. Nevertheless, results of "almost" reducibility have been announced in Eliasson [54], Eliasson-Grébert-Kuksin [55].

Before concluding this subsection, we also mention the KAM result by Grébert-Thomann [77] for smoothing nonlinear perturbations of the 1-d harmonic oscillator and Grébert-Paturel [76] in higher space dimension.

### 1-d quasi and fully nonlinear PDEs

Another situation where the reducibility approach that we described in subsection 1.3.1 encounters a serious difficulty is when the non-diagonal remainder $R(\varphi)$ is unbounded. This is the difficulty mentioned in item 2. In such a case, the auxiliary vector field $F(\varphi)$ defined in (1.3.17) is unbounded as well. Therefore it could not define a flow, and the iterative reducibility scheme described in subsection 1.3.1 would formally produce remainders which accumulate more and more derivatives.

**KAM for semilinear PDEs with derivatives.** The first KAM results for PDEs with an *unbounded* nonlinearity have been proved by Kuksin [94] and, then, Kappeler-Pöschel [90], for perturbations of finite-gap solutions of

$$u_t + u_{xxx} + \partial_x u^2 + \varepsilon \partial_x (\partial_u f)(x, u) = 0 , \quad x \in \mathbb{T} . \tag{1.3.41}$$

The corresponding quasi-periodic linearized operator at an approximate quasi-periodic solution $u$ has the form

$$h \mapsto \omega \cdot \partial_\varphi h + \partial_{xxx} h + \partial_x (2uh) + \varepsilon \partial_x (ah) , \quad a := (\partial_{uu} f)(x, u) .$$

The key idea in [94] is to exploit the fact that the frequencies of KdV grow asymptotically as $\sim j^3$ as $j \to +\infty$, and therefore one can impose second order Melnikov non-resonance conditions like

$$|\omega \cdot \ell + j^3 - i^3| \geq \gamma (j^2 + i^2)/\langle \ell \rangle^\tau , \quad i \neq j ,$$

which gain 2 space derivatives (outside the diagonal $i = j$), sufficient to compensate the loss of one space derivative produced by the vector field $\varepsilon \partial_x (\partial_u f)(x, u)$. On the diagonal $\ell \neq 0, i = j$, one renounces to solve the homological equations, with the consequence that the KAM normal form is $\varphi$-dependent and one uses the so called Kuksin Lemma to invert the corresponding quasi-periodic scalar operator. Subsequently, developing an improved version of the Kuksin Lemma, Liu-Yuan in [86] proved KAM results for semilinear perturbations of Hamiltonian derivative NLS and Benjiamin-Ono equations and Zhang-Gao-Yuan [119] for the reversible derivative NLS equation

$$\mathrm{i}u_t + u_{xx} = |u_x|^2 u\,, \quad u(0) = u(\pi) = 0\,.$$

These PDEs are more difficult than KdV because the linear frequencies grow like $\sim j^2$ and not $\sim j^3$, and therefore one gains only 1 space derivative when solving the homological equations.

These methods do not apply for derivative wave equations where the dispersion relation is asymptotically linear. Such a case has been addressed more recently by Berti-Biasco-Procesi [17]-[18] who proved the existence and the stability of quasi-periodic solutions of autonomous derivative Klein-Gordon equations

$$y_{tt} - y_{xx} + my = g(x, y, y_x, y_t) \tag{1.3.42}$$

satisfying reversibility conditions which rule out nonlinearities like $y_t^3$, $y_x^3$, for which no periodic nor quasi-periodic solutions exist (with these nonlinearities all the solutions dissipate to zero). The key point in [17]-[18] was to adapt the notion of quasi-Töplitz vector field introduced in [108] to obtain the higher order *asymptotic expansion* of the perturbed normal frequencies

$$\mu_j(\varepsilon) = \sqrt{j^2 + m} + a_\pm + O(1/j)\,, \quad \text{as } j \to \pm\infty\,,$$

for suitable constants $a_\pm$ (of the size $a_\pm = O(\varepsilon)$ of the solution $y = O(\varepsilon)$). Thanks to this asymptotic expansion it is sufficient to verify, for each $\ell \in \mathbb{Z}^\nu$, that only finitely many second order Melnikov non-resonance conditions hold. Indeed, using an argument as in (1.3.35), infinitely many conditions in (1.3.33) are already verified by imposing only first order Melnikov conditions like

$$|\omega \cdot \ell + h| \geq \gamma \langle \ell \rangle^{-\tau}\,, \qquad |\omega \cdot \ell + (a_+ - a_-) + h| \geq \gamma \langle \ell \rangle^{-\tau}\,,$$

for all $(\ell, h) \in \mathbb{Z}^\nu \times \mathbb{Z}$ such that $\omega \cdot \ell + h$ and $\omega \cdot \ell + (a_+ - a_-) + h$ do not vanish identically.

### KAM for quasi-linear and fully nonlinear PDEs

All the above results still concern semi-linear perturbations, namely when the number of derivatives which are present in the nonlinearity is strictly lower than the order of the linear differential operator. The first existence results of quasi-periodic solutions for quasi-linear PDEs have been proved by Baldi-Berti-Montalto in [6] for fully nonlinear perturbations of the Airy equation

$$u_t + u_{xxx} + \varepsilon f(\omega t, x, u, u_x, u_{xx}, u_{xxx}) = 0\,, \quad x \in \mathbb{T}\,, \tag{1.3.43}$$

and in [8] for quasi-linear autonomous perturbed KdV equations

$$u_t + u_{xxx} + \partial_x u^2 + \mathcal{N}(x, u, u_x, u_{xx}, u_{xxx}) = 0 \tag{1.3.44}$$

where the Hamiltonian nonlinearity

$$\mathcal{N}(x,u,u_x,u_{xx},u_{xxx}) := -\partial_x\big[(\partial_u f)(x,u,u_x) - \partial_x((\partial_{u_x} f)(x,u,u_x))\big] \tag{1.3.45}$$

vanishes at the origin as $O(u^4)$, see [72] when $\mathcal{N}$ vanishes only quadratically.

The main new tool which has been introduced to solve this problem is a systematic use of pseudo-differential calculus. The key point is to *reduce* to constant coefficients the linear PDE

$$u_t + (1 + a_3(\omega t,x))u_{xxx} + a_2(\omega t,x)u_{xx} + a_1(\omega t,x)u_x + a_0(\omega t,x)u = 0 \tag{1.3.46}$$

which is obtained linearizing (1.3.43) at an approximate quasi-periodic solution $u(\omega t,x)$. The coefficients $a_i(\omega t,x) = O(\varepsilon)$, $i = 0,\dots,3$. Instead of trying to diminish the size of the variable-dependent terms in (1.3.46), as in the scheme outlined in subsection 1.3.1 –the big difficulty 2 would appear–, the aim is to conjugate (1.3.46) to a system like

$$u_t + m_3 u_{xxx} + m_1 u_x + \mathcal{R}_0(\omega t)u = 0 \tag{1.3.47}$$

where $m_3 = 1 + O(\varepsilon)$, $m_1 = O(\varepsilon)$ are constants and $\mathcal{R}_0(\omega t)$ is a zero order operator, still time dependent. To do this, (1.3.46) is conjugated with a time quasi-periodic change of variable (as in (1.3.2))

$$u = \Phi(\omega t)[v] = v(t, x + \beta(\omega t,x))\,, \tag{1.3.48}$$

induced by the composition with a diffeomorphism $x \mapsto x + \beta(\varphi,x)$ of $\mathbb{T}_x$ ( requiring $|\beta_x| < 1$). The conjugated system (1.3.3)-(1.3.4) is

$$v_t + \Phi^{-1}(\omega t)((1 + a_3(\omega t,x))(1 + \beta_x(\omega t,x))^3)v_{xxx}(t,x) + \text{lower order operators} = 0$$

and therefore one chooses a periodic function $\beta(\varphi,x)$ such that

$$(1 + a_3(\varphi,x))(1 + \beta_x(\varphi,x))^3 = m_3(\varphi)$$

is independent of $x$. Since $\beta_x(\varphi,x)$ has zero space average, this is possible with

$$m_3(\varphi) = \Big(\frac{1}{2\pi}\int_{\mathbb{T}} \frac{dx}{(1 + a_3(\varphi,x))^{\frac{1}{3}}}\Big)^{-3}.$$

The $\varphi$ dependence of $m_3(\varphi)$ can also be eliminated at the highest order using a quasi-periodic reparametrization of time and, using other pseudo-differential transformations, we can reduce also the lower order terms to constant coefficients obtaining (1.3.47). The reduction (1.3.47) implies the accurate asymptotic expansion of the perturbed frequencies

$$\mu_j(\varepsilon) = -\mathrm{i}m_3 j^3 + \mathrm{i}m_1 j + O(\varepsilon)$$

and therefore now it is possible to verify the second order Melnikov non-resonance conditions required by a KAM reducibility scheme (as outlined in subsection 1.3.1) to diagonalize $\mathcal{R}_0(\omega t)$, completing the reduction of (1.3.47), thus (1.3.46).

These techniques have been then employed by Feola-Procesi [62] for quasi-linear forced perturbations of Schrödinger equations and in [47], [63] for the search of analytic solutions of autonomous PDEs. These kind of ideas have been also successfully generalized for unbounded perturbations of harmonic oscillators by Bambusi [9], [10] and Bambusi-Montalto [13].

The KdV and the NLS equation are partial differential equations and the pseudo-differential tools required are essentially commutators of multiplication operators and Fourier multipliers. On the other hand, for the water waves equations, that we now present, the theory of pseudo-differential operators has to be used in full strength.

### Water waves equations

The water waves equations for a perfect, incompressible, inviscid, irrotational fluid occupying the time dependent region

$$\mathcal{D}_\eta := \big\{(x,y) \in \mathbb{T} \times \mathbb{R} \, : \, -h < y < \eta(t,x)\big\}\,, \quad \mathbb{T} := \mathbb{T}_x := \mathbb{R}/2\pi\mathbb{Z}\,, \tag{1.3.49}$$

under the action of gravity, and possible capillary forces at the free surface, are the Euler equations of hydrodynamics combined with conditions at the boundary of the fluid:

$$\begin{cases} \partial_t \Phi + \dfrac{1}{2}|\nabla\Phi|^2 + g\eta = \kappa \partial_x \Big( \dfrac{\eta_x}{\sqrt{1+\eta_x^2}} \Big) & \text{at } y = \eta(t,x) \\ \Delta \Phi = 0 & \text{in } \mathcal{D}_\eta \\ \partial_y \Phi = 0 & \text{at } y = -h \\ \partial_t \eta = \partial_y \Phi - \partial_x \eta \partial_x \Phi_x & \text{at } y = \eta(t,x) \end{cases} \tag{1.3.50}$$

where $g$ is the acceleration of gravity and $\kappa$ is the surface tension coefficient. The unknowns of the problem (1.3.50) are the free surface $y = \eta(t,x)$ and the velocity potential $\Phi : \mathcal{D}_\eta \to \mathbb{R}$, i.e. the irrotational velocity field of the fluid $v = \nabla_{x,y}\Phi$. The first equation in (1.3.50) is the Bernoulli condition according to which the jump of pressure across the free surface is proportional to the mean curvature. The second equation in (1.3.50) is the incompressibility property $\operatorname{div} v = 0$. The third equation expresses the impermeability of the bottom of the ocean. The last condition in (1.3.50) means that the fluid particles on the free surface $y = \eta(x,t)$ remain forever on it along the fluid evolution.

Following Zakharov [116] and Craig-Sulem [50], the evolution problem (1.3.50) may be written as an infinite-dimensional Hamiltonian system in the unknowns $(\eta(t,x), \psi(t,x))$ where $\psi(t,x) = \Phi(t,x,\eta(t,x))$ is, at each instant $t$, the trace at the free boundary of the velocity potential. Given $\eta(t,x)$ and $\psi(t,x)$ there is a unique solution $\Phi(t,x,y)$ of the elliptic problem

$$\begin{cases} \Delta\Phi = 0 & \text{in } \{-h < y < \eta(t,x)\} \\ \partial_y \Phi = 0 & \text{on } y = -h \\ \Phi = \psi & \text{on } \{y = \eta(t,x)\}\,. \end{cases}$$

System (1.3.50) is then equivalent to the Zakharov-Craig-Sulem system

$$\begin{cases} \partial_t \eta = G(\eta)\psi \\ \partial_t \psi = -g\eta - \dfrac{\psi_x^2}{2} + \dfrac{1}{2(1+\eta_x^2)}\big(G(\eta)\psi + \eta_x\psi_x\big)^2 + \kappa\partial_x\Big(\dfrac{\eta_x}{\sqrt{1+\eta_x^2}}\Big) \end{cases} \tag{1.3.51}$$

where $G(\eta) := G(\eta; h)$ is the Dirichlet-Neumann operator defined as

$$G(\eta)\psi := \big(\Phi_y - \eta_x \Phi_x\big)_{|y=\eta(t,x)} \tag{1.3.52}$$

which maps the Dirichlet datum $\psi$ into the (normalized) normal derivative $G(\eta)\psi$ at the top boundary. The operator $G(\eta)$ is linear in $\psi$, self-adjoint with respect to the $L^2$ scalar product, positive-semidefinite, and its kernel contains only the constant functions. The Dirichlet-Neumann operator depends smoothly with respect to the wave profile $\eta$, and it is a *pseudo-differential* operator with principal symbol $D\tanh(hD)$.

Furthermore the equations (1.3.51) are the Hamiltonian system

$$\partial_t \eta = \nabla_\psi H(\eta,\psi)\,, \quad \partial_t \psi = -\nabla_\eta H(\eta,\psi) \tag{1.3.53}$$

where $\nabla$ denotes the $L^2$-gradient, and the Hamiltonian

$$H(\eta,\psi) = \frac{1}{2}\int_{\mathbb{T}} \psi\, G(\eta,h)\psi\, dx + \frac{g}{2}\int_{\mathbb{T}} \eta^2\, dx + \kappa \int_{\mathbb{T}} \sqrt{1+\eta_x^2}\, dx \tag{1.3.54}$$

is the sum of the kinetic, potential and capillary energies expressed in terms of the variables $(\eta,\psi)$.

The water waves system (1.3.51)-(1.3.53) exhibits several symmetries. First of all, the mass

$$\int_{\mathbb{T}} \eta(x)\, dx$$

is a first integral of (1.3.51). Moreover (1.3.51) is invariant under spatial translations and Noether's theorem implies that the momentum

$$\int_{\mathbb{T}} \eta_x(x)\psi(x)\, dx$$

is a prime integral of (1.3.53). In addition, the subspace of functions that are even in $x$,

$$\eta(x) = \eta(-x) , \quad \psi(x) = \psi(-x) , \tag{1.3.55}$$

is invariant under (1.3.51). In this case also the velocity potential $\Phi(x, y)$ is even and $2\pi$-periodic in $x$ and so the $x$-component of the velocity field $v = (\Phi_x, \Phi_y)$ vanishes at $x = k\pi$, for all $k \in \mathbb{Z}$. Hence there is no flow of fluid through the lines $x = k\pi$, $k \in \mathbb{Z}$, and a solution of (1.3.51) satisfying (1.3.55) describes the motion of a liquid confined between two vertical walls.

We also notice that the water waves system (1.3.51)-(1.3.53) is reversible with respect to the involution $S : (\eta, \psi) \mapsto (\eta, -\psi)$, i.e. the Hamiltonian $H$ in (1.3.54) is even in $\psi$, see Appendix A. As a consequence it is natural to look for solutions of (1.3.51) satisfying

$$u(-t) = Su(t) , \quad i.e. \quad \eta(-t, x) = \eta(t, x) , \ \psi(-t, x) = -\psi(t, x) \quad \forall t, x \in \mathbb{R} . \tag{1.3.56}$$

Solutions of the water waves equations (1.3.51) satisfying (1.3.55) and (1.3.56) are called *standing water waves*.

The phase space of (1.3.51) is (a dense subspace) of

$$(\eta, \psi) \in H^1_0(\mathbb{T}) \times \dot{H}^1(\mathbb{T}) \qquad \text{where} \qquad \dot{H}^1(\mathbb{T}) := H^1(\mathbb{T})/_{\sim}$$

is the homogeneous space obtained by the equivalence relation $\psi_1(x) \sim \psi_2(x)$ if and only if $\psi_1(x) - \psi_2(x) = c$ is a constant, and $H^1_0(\mathbb{T})$ is the subspace of $H^1(\mathbb{T})$ of zero average functions. For simplicity of notation we denote the equivalence class $[\psi]$ by $\psi$. Note that the second equation in (1.3.51) is in $\dot{H}^1(\mathbb{T})$, as it is natural because only the gradient of the velocity potential has a physical meaning.

Linearizing (1.3.51) at the equilibrium $(\eta, \psi) = (0, 0)$ we get

$$\begin{cases} \partial_t \eta = G(0)\psi, \\ \partial_t \psi = -g\eta + \kappa \eta_{xx} \end{cases} \tag{1.3.57}$$

where $G(0) = D \tanh(hD)$ is the Dirichlet-Neumann operator at the flat surface $\eta = 0$. The linear frequencies of oscillations of (1.3.57) are

$$\omega_j = \sqrt{j \tanh(hj)(g + \kappa j^2)} , \quad j \in \mathbb{Z} , \tag{1.3.58}$$

which, in the phase space of even functions (1.3.55), are *simple*. Also notice that

$$\begin{aligned} &\text{if } \kappa > 0 \text{ (capillary} - \text{gravity waves)} &&\Longrightarrow \quad \omega_j \sim |j|^{3/2} \text{ as } |j| \to \infty , \\ &\text{if } \kappa = 0 \text{ (pure gravity waves)} &&\Longrightarrow \quad \omega_j \sim |j|^{1/2} \text{ as } |j| \to \infty . \end{aligned}$$

**KAM for water waves.** The first existence results of small amplitude time-periodic gravity standing wave solutions for bi-dimensional fluids has been proved by Plotinkov and

Toland [101] in finite depth and by Iooss, Plotnikov and Toland in [84] in infinite depth. More recently, the existence of time periodic gravity-capillary standing wave solutions in infinite depth has been proved by Alazard-Baldi [1].

The main result in [30] proves that most of the standing wave solutions of the linear system (1.3.57), which are Fourier supported on finitely many space Fourier modes, can be continued to standing wave solutions of the nonlinear water-waves system (1.3.51) for most values of the surface tension parameter $\kappa \in [\kappa_1, \kappa_2]$.

A key step is the reducibility to constant coefficients of the quasi-periodic operator $\mathcal{L}_\omega$ obtained linearizing (1.3.51) at a quasi-periodic approximate solution. After the introduction of a linearized Alinhac good unknown, and using in a systematic way pseudo-differential calculus, it is possible to transform $\mathcal{L}_\omega$ into a complex quasi-periodic linear operator of the form

$$(h, \bar{h}) \mapsto \big(\omega \cdot \partial_\varphi + \mathrm{i}\mathtt{m}_3 |D|^{\frac{1}{2}} (1 - \kappa \partial_{xx})^{\frac{1}{2}} + \mathrm{i}\mathtt{m}_1 |D|^{\frac{1}{2}} \big) h + \mathcal{R}(\varphi)[h, \bar{h}] \tag{1.3.59}$$

where $\mathtt{m}_3, \mathtt{m}_1 \in \mathbb{R}$ are constants satisfying $\mathtt{m}_3 \approx 1$, $\mathtt{m}_1 \approx 0$, and the remainder $\mathcal{R}(\varphi)$ is a small bounded operator. Then a KAM reducibility scheme completes the diagonalization of the linearized operator $\mathcal{L}_\omega$. The required second order Melnikov non-resonance conditions are fulfilled for most values of the surface tension parameter $\kappa$ generalizing ideas of degenerate KAM theory for PDEs [11], exploiting that the linear frequencies $\omega_j$ in (1.3.58) are analytic and non degenerate in $\kappa$, and the sharp asymptotic expansion of the perturbed frequencies obtained by the regularization procedure.

In the case of pure gravity water waves, i.e. $\kappa = 0$, the linear frequencies of oscillation are (see (1.3.58))

$$\omega_j := \omega_j(h) := \sqrt{gj \tanh(hj)}\,, \quad j \geq 1\,, \tag{1.3.60}$$

and three major further difficulties in proving the existence of time quasi-periodic solutions are:

(*i*) The nonlinear water waves system (1.3.51) (with $\kappa = 0$) is a singular perturbation of (1.3.57) (with $\kappa = 0$) in the sense that the quasi-periodic linearized operator assumes the form

$$\omega \cdot \partial_\varphi + \mathrm{i}|D|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(h|D|) + V(\varphi, x)\partial_x$$

and the term $V(\varphi, x)\partial_x$ is now a *singular* perturbation of the linear dispersion relation operator $\mathrm{i}|D|^{\frac{1}{2}} \tanh^{\frac{1}{2}}(h|D|)$ (on the contrary, for the gravity-capillary case the transport term $V(\varphi, x)\partial_x$ is a lower order perturbation of $|D|^{\frac{3}{2}}$, see (1.3.59)).

(*ii*) The dispersion relation (1.3.60) is sublinear, i.e. $\omega_j \sim \sqrt{j}$ for $j \to \infty$, and therefore it is only possible to impose second order Melnikov non-resonance conditions as in (1.3.21) which lose space derivatives.

(*iii*) The linear frequencies $\omega_j(h)$ in (1.3.60) vary with $h$ of just exponentially small quantities.

The main result in Baldi-Berti-Haus-Montalto [5] proves the existence of pure gravity standing water waves solutions.

The difficulty (i) is solved proving a straightening theorem for a quasi-periodic transport operator: there is a quasi-periodic change of variables of the form $x \mapsto x + \beta(\varphi, x)$ which conjugates

$$\omega \cdot \partial_\varphi + V(\varphi, x)\partial_x$$

to the constant coefficient vector field $\omega \cdot \partial_\varphi$, for $V(\varphi, x)$ small. This perturbative rectification result is a classical small divisor problem, solved for perturbations of a Diophantine vector field at the beginning of KAM theory, see e.g. [117, 118]. Notice that, despite the fact that $\omega \in \mathbb{R}^\nu$ is Diophantine, the constant vector field $\omega \cdot \partial_\varphi$ is resonant on the higher dimensional torus $\mathbb{T}^\nu_\varphi \times \mathbb{T}_x$. We exploit in a crucial way the *symmetry* induced by the *reversible* structure of the water waves equations, i.e. $V(\varphi, x)$ is odd in $\varphi$. This problem amounts to prove that all the solutions of the quasi periodically time-dependent scalar characteristic equation $\dot{x} = V(\omega t, x)$, $x \in \mathbb{T}$, are quasi-periodic in time with frequency $\omega$.

The difficulty (ii) is overcome performing a regularizing procedure which conjugates the linearized operator, obtained along the Nash-Moser iteration, to a diagonal, constant coefficient linear one, up to a sufficiently *smoothing* operator. In this way the subsequent KAM reducibility scheme converges also in presence of very weak Melnikov non-resonance conditions as in (1.3.21) which lose space derivatives. This regularization strategy is in principle applicable to a broad class of PDEs where the second order Melnikov non-resonance conditions lose space derivatives.

The difficulty (iii) is solved by an improvement of the degenerate KAM theory for PDEs in Bambusi-Berti-Magistrelli [11], which allows to prove that all the Melnikov non-resonance are fulfilled for most values of $h$.

**Remark 1.3.9.** *We can introduce the space wavelength $2\pi\varsigma$ as an internal free parameter in the water waves equations* (1.3.51)*. Rescaling properly time, space and amplitude of the solution $(\eta(t,x), \psi(t,x))$ we obtain system* (1.3.51) *where the gravity constant $g = 1$ and the depth parameter $h$ depends linearly on $\varsigma$. In this way [5] proves existence results for a fixed equation, i.e. a fixed depth $h$, for most values of the space wavelength $2\pi\varsigma$.*

## 1.4 The multiscale approach to KAM for PDEs

We present now some key ideas about another approach developed for analyzing linear quasi-periodic systems, in order to prove KAM results for PDEs, started with the seminal paper of Craig-Wayne [51] and strongly extended by Bourgain [34]-[40]. This set of ideas and techniques -referred as "multiscale analysis"- is at the basis of the present Monograph. For this reason we find convenient to present it in some detail.

For definiteness we consider a quasi-periodic linear wave operator

$$(\omega \cdot \partial_\varphi)^2 - \Delta + m + \varepsilon b(\varphi, x), \quad \varphi \in \mathbb{T}^\nu, \ x \in \mathbb{T}^d, \tag{1.4.1}$$

where $m > 0$ and $b(\varphi, x)$ is a smooth function, that is obtained linearizing a nonlinear wave equation at a smooth approximate quasi-periodic solution.

**Remark 1.4.1.** *The choice of the parameters in* (1.4.1) *makes a significant difference. If* (1.4.1) *arises linearizing a quasi-periodically forced NLW as* (1.1.7)*, the frequency vector* $\omega$ *can be regarded as a free parameter belonging to a subset of* $\mathbb{R}^\nu$*, independent of* $\varepsilon$*. On the other hand, if* (1.4.1) *arises linearizing an autonomous NLW like* (1.1.1)*, or* (1.1.8)*, the frequency* $\omega$ *and* $\varepsilon$ *are linked. In particular* $\omega$ *has to be* $\varepsilon^2$*-close to some frequency vector* $((|j_i|^2 + m)^{1/2})_{i=1,\ldots,\nu}$ *(actually* $\omega$ *has to belong to the region of "admissible" frequencies as in* (1.2.23)*). This difficulty is present in this Monograph as well as in [114].*

In the exponential basis $\{e^{\mathrm{i}(\ell \cdot \varphi + j \cdot x)}\}_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}^d}$ the linear operator (1.4.1) is represented by the self-adjoint matrix

$$\begin{aligned} & A := D + \varepsilon T, \\ & D := \mathrm{Diag}_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{Z}^d}\big( -(\omega\cdot\ell)^2 + |j|^2 + m \big), \quad T = (T_{\ell,j}^{\ell',j'}) := (\widehat{b}_{\ell-\ell',j-j'}), \end{aligned} \tag{1.4.2}$$

where $\widehat{b}_{\ell-\ell',j-j'}$ are the Fourier coefficients of the function $b(\varphi, x)$ that decay rapidly to zero as $|(\ell - \ell', j - j')| \to +\infty$, exponentially fast, if $b(\varphi, x)$ is analytic, polynomially, if $b(\varphi, x)$ is a Sobolev function. Notice that the matrix $T$ is Töplitz, namely it has constant entries on the diagonals $(\ell - \ell', j - j') = (L, J) \in \mathbb{Z}^\nu \times \mathbb{Z}^d$.

**Remark 1.4.2.** *The analytic/Gevrey setting has been considered in [34]-[40] and the Sobolev case in [21]-[26]. The Sobolev regularity of* $b(\varphi, x)$ *has to be large enough, see remark 1.4.6. In KAM for PDE applications this requires that the nonlinearity (thus the solution) is sufficiently many times differentiable. For finite dimensional Hamiltonian systems, it has been rigorously proved that, otherwise, all the invariant tori could be destroyed and only discontinuous Aubry-Mather invariant sets survive, see e.g. [82].*

The infinite set of the eigenvalues of the diagonal operator $D$,

$$-(\omega \cdot \ell)^2 + |j|^2 + m, \quad \ell \in \mathbb{Z}^\nu, \ j \in \mathbb{Z}^d,$$

accumulate to zero (small divisors) and therefore the matrix $T$ in (1.4.2), which represents in Fourier space the multiplication operator for the function $b(\varphi, x)$, is a "singular" perturbation of $D$. As a consequence is not obvious at all that the self-adjoint operator $D + \varepsilon T$ has still pure point spectrum with a basis of eigenfunctions with exponential/polinomial decay, for $\varepsilon$ small, for most values of the frequency vector $\omega \in \mathbb{R}^\nu$. This is the main problem addressed in Anderson localization theory.

**Remark 1.4.3.** *If a quasi-periodic operator* $\mathrm{i}\omega \cdot \partial_\varphi + S(\varphi)$ *where* $S(\varphi)$ *is self-adjoint, is reduced to constant coefficients as we discussed in section 1.3 with a quasi-periodic change of variables* $\Phi(\omega t)$ *which acts between Sobolev spaces, then it is pure point.*

Actually, for the convergence of a Nash-Moser scheme in applications to KAM for PDEs, it is sufficient to prove, for most values of the parameters, the invertibility of its finite dimensional restrictions,

$$\mathcal{L}_N := \Pi_N \big( (\omega \cdot \partial_\varphi)^2 - \Delta + m + \varepsilon b(\varphi, x) \big)_{|\mathcal{H}_N} , \tag{1.4.3}$$

for any $N$ large, where $\Pi_N$ denotes the projection on the finite dimensional subspace

$$\mathcal{H}_N := \Big\{ h(\varphi, x) = \sum_{|(\ell, j)| \leq N} h_{\ell, j} e^{\mathrm{i}(\ell \cdot \varphi + j \cdot x)} \Big\} , \tag{1.4.4}$$

and prove that the inverse satisfies, for some $\mu > 0$, $s_1 > 0$, tame estimates as

$$\| \mathcal{L}_N^{-1} h \|_s \leq C(s) N^\mu \big( \|h\|_s + \|b\|_s \|h\|_{s_1} \big) , \ \forall h \in \mathcal{H}_N , \ \forall s \geq s_1 , \tag{1.4.5}$$

where $\| \ \|_s$ denotes the Sobolev norm in (1.2.31). Notice that $\mu > 0$ represents a "loss of derivatives" due to the small divisors. Since the multiplication operator $h \mapsto bh$ satisfies a tame estimate like (1.4.5) with $\mu = 0$, the estimate (1.4.5) means that $\mathcal{L}_N^{-1}$ acts, on the Sobolev scale $\mathcal{H}^s$, somehow as an unbounded differential operator of order $\mu$.

Also weaker tame estimates as

$$\| \mathcal{L}_N^{-1} h \|_s \leq C(s) N^{\tau'} \big( (N^{\varsigma s} + \|b\|_s) \|h\|_{s_1} + N^{\varsigma s_1} \|h\|_s \big) , \ \forall h \in \mathcal{H}_N , \ \forall s \geq s_1 , \tag{1.4.6}$$

for $\varsigma < 1$, are sufficient for the convergence of the Nash-Moser scheme. Notice that such conditions are much weaker than (1.4.5) because the loss of derivatives $\tau' + \varsigma s$ increases with $s$. Conditions like (1.4.6) are essentially optimal for the convergence, compare with Lojasiewicz-Zehnder [87].

**Remark 1.4.4.** *It would be also sufficient to prove the existence of a right inverse of the operator* $[\mathcal{L}]_N^{2N} := \Pi_N \big( (\omega \cdot \partial_\varphi)^2 - \Delta + m + \varepsilon b(\varphi, x) \big)_{|\mathcal{H}_{2N}}$ *satisfying* (1.4.5) *or* (1.4.6)*, as we do in section 4.6. We remind that a linear operator acting between finite dimensional vector spaces admits a right inverse if it is surjective, see Definition 3.3.11.*

In the time periodic setting, i.e. $\nu = 1$, we achieve the stronger tame estimates (1.4.5) while, in the time quasi-periodic setting, i.e. $\nu \geq 2$, we prove the weaker tame estimates (1.4.6). Actually, in the present Monograph, as in [22]-[23], we shall prove that the approximate inverse satisfies tame estimates of the weaker form (1.4.6).

In order to achieve (1.4.5) or (1.4.6) there are two main steps:

1. ($L^2$-*estimates*) Impose lower bounds for the eigenvalues of the self-adjoint operator (1.4.3), for most values of the parameters. These "first order Melnikov" non-resonance conditions provide estimates of the inverse of $\mathcal{L}_N^{-1}$ in $L^2$-norm.

2. ($\mathcal{H}^s$-*estimates*) Prove the estimates (1.4.5) or (1.4.6) in high Sobolev norms. In the language of Anderson localization theory, this amounts to prove polynomially fast off-diagonal decay estimates for the inverse matrix $\mathcal{L}_N^{-1}$.

In the forced case when the frequency vector $\omega \in \mathbb{R}^\nu$ provides independent parameters, item 1 is not a too difficult task, using results about the eigenvalues of self-adjoint matrices depending on parameters, as Lemmata 4.8.1 and 4.8.2. On the other hand, in the autonomous case, for the difficulties discussed in remark 1.4.1, this is a more delicate task (that we address in this Monograph).

In the sequel we concentrate on the analysis of item 2. An essential ingredient is the decomposition into "singular" and "regular" sites, for some $\rho > 0$,

$$\begin{aligned} S &:= \Big\{ (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d \quad \text{such that} \quad |-(\omega \cdot \ell)^2 + |j|^2 + m| < \rho \Big\} \\ R &:= \Big\{ (\ell, j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d \quad \text{such that} \quad |-(\omega \cdot \ell)^2 + |j|^2 + m| \geq \rho \Big\} . \end{aligned} \tag{1.4.7}$$

It is clear, indeed, that, in order to achieve (1.4.5), conditions that limit the quantity of singular sites have to be fulfilled, otherwise the inverse operator $\mathcal{L}_N^{-1}$ would be "too big" in any sense and (1.4.6) would be violated.

**Remark 1.4.5.** *If in* (1.4.1) *the constant $m$ is replaced by a (not small) multiplicative potential $V(x)$ (this is the case considered in the present Monograph), it is natural to define the singular sites as $|-(\omega \cdot \ell)^2 + |j|^2 + m| < \Theta$ for some $\Theta$ large depending on $V(x)$.*

We first consider the easier case $\nu = 1$ (time-periodic solutions).

### 1.4.1 Time periodic case $\nu = 1$

Existence of time periodic solutions for NLW on $\mathbb{T}^d$ have been first obtained in [35]. In the exposition below we follow [21]. The following "separation properties" of the singular sites (1.4.7) are sufficient for proving (1.4.5): the singular sites $S$ in the box $[-N, N]^{1+d}$ are partitioned into disjoint clusters $\Omega_\alpha$,

$$S \cap [-N, N]^{1+d} = \bigcup_\alpha \Omega_\alpha \,, \tag{1.4.8}$$

satisfying

- **(H1) (Dyadic)** $M_\alpha \leq 2 m_\alpha$, $\forall \alpha$, where $M_\alpha := \max_{(\ell,j)\in\Omega_\alpha} |(\ell, j)|$ and $m_\alpha := \min_{(\ell,j)\in\Omega_\alpha} |(\ell, j)|$;

- **(H2) (Separation at infinity)** There is $\delta = \delta(d) > 0$ (independent of $N$) such that

$$\mathrm{d}(\Omega_\alpha, \Omega_\beta) := \min_{(\ell,j)\in\Omega_\alpha,(\ell',j')\in\Omega_\sigma} |(\ell, j) - (\ell', j')| \geq (M_\alpha + M_\sigma)^\delta , \quad \forall \alpha \neq \sigma . \tag{1.4.9}$$

Notice that in (H2) the clusters $\Omega_\alpha$ of singular sites are "separated at infinity", namely the distance between distinct clusters increases when the Fourier indices tend to infinity. A partition of the singular sites as (1.4.8), satisfying (H1)-(H2), has been proved in [21] assuming that $\omega^2$ is Diophantine.

**Remark 1.4.6.** *We require that the function* $b(\varphi, x)$ *in* (1.4.1) *has the same (high) regularity in* $(\varphi, x)$, *i.e.* $\|b\|_s < \infty$ *for some large Sobolev index* $s$, *because the clusters* $\Omega_\alpha$ *in* (1.4.9) *are separated in time-space Fourier indices only. This implies, in KAM applications, that the solutions that we obtain will have the same high Sobolev regularity in time and space.*

We have to solve the linear system

$$\mathcal{L}_N g = h , \quad g, h \in \mathcal{H}_N . \tag{1.4.10}$$

Given a subset $\Omega$ of $[-N, N]^{1+d} \subset \mathbb{Z} \times \mathbb{Z}^d$ we denote by $\mathcal{H}_\Omega$ the vector space spanned by $\{e^{\mathrm{i}(\ell\varphi+j\cdot x)}, (\ell, j) \in \Omega\}$ and by $\Pi_\Omega$ the corresponding orthogonal projector. With this notation the subspace $\mathcal{H}_N$ in (1.4.4) coincides with $\mathcal{H}_{[-N,N]^{1+d}}$. Given a linear operator $L$ of $\mathcal{H}_N$ and another subset $\Omega' \subset [-N, N]^{1+d}$ we denote $L^\Omega_{\Omega'} := \Pi_{\mathcal{H}_{\Omega'}} L_{|\mathcal{H}_\Omega}$. For simplicity we also set $L^\Omega_{\Omega'} := \Pi_{\mathcal{H}_{\Omega'}} (\mathcal{L}_N)_{|\mathcal{H}_\Omega}$.

According to the splitting of the indices $[-N, N]^{1+d} = S \cup R$, where for simplicity of notation we still denote by $S, R$ the sets of the singular and regular sites (1.4.7) intersected with the box $[-N, N]^{1+d}$, we decompose in orthogonal subspaces

$$\mathcal{H}_N = \mathcal{H}_S \oplus \mathcal{H}_R .$$

Writing the unique decomposition $g = g_S + g_R$, $g_R \in \mathcal{H}_R$, $g_S \in \mathcal{H}_S$, and similarly for $h$, the linear system (1.4.10) then amounts to

$$\begin{cases} L^R_R g_R + L^S_R g_S = h_R \\ L^R_S g_R + L^S_S g_S = h_S . \end{cases}$$

Notice that, by (1.4.2), the coupling terms $L^S_R = \varepsilon T^S_R$, $L^R_S = \varepsilon T^R_S$, have polynomial off-diagonal decay.

By standard perturbative arguments, the operator $L^R_R$, which is the restriction of $\mathcal{L}_N$ to the regular sites, is invertible for $\varepsilon\rho^{-1} \ll 1$, and therefore, solving the first equation in $g_R$, and inserting the result in the second one, we are reduced to solve

$$\big(L^S_S - L^R_S (L^R_R)^{-1} L^S_R\big) g_S = h_S - L^R_S (L^R_R)^{-1} h_R .$$

Thus the main task is now to invert the self-adjoint matrix

$$U := L_S^S - L_S^R (L_R^R)^{-1} L_R^S \,, \quad U : \mathcal{H}_S \to \mathcal{H}_S \,.$$

This reduction procedure is sometimes referred as a resolvent identity.

According to (1.4.8) we have the orthogonal decomposition $\mathcal{H}_S = \oplus_\alpha \mathcal{H}_{\Omega_\alpha}$ which induces a block decomposition for the operator

$$U = (U_{\Omega_\sigma}^{\Omega_\alpha})_{\alpha,\sigma} \,, \quad U_{\Omega_\sigma}^{\Omega_\alpha} := L_{\Omega_\sigma}^{\Omega_\alpha} - L_{\Omega_\sigma}^R (L_R^R)^{-1} L_R^{\Omega_\alpha} \,.$$

Then we decompose $U$ in block-diagonal and off-diagonal parts:

$$U = \mathcal{D} + \mathcal{R} \,, \quad \mathcal{D} := \mathrm{Diag}_\alpha U_{\Omega_\alpha}^{\Omega_\alpha} \,, \quad \mathcal{R} := (U_{\Omega_\sigma}^{\Omega_\alpha})_{\alpha \neq \sigma} \,. \tag{1.4.11}$$

Since the matrix $T$ has off-diagonal decay (see (1.4.2) and recall that the function $b(\varphi, x)$ is smooth enough) and the matrices with off-diagonal decay form an algebra (with interpolation estimates) it easily results an off-diagonal decay of the matrix $\mathcal{R}$ like

$$\|U_{\Omega_\sigma}^{\Omega_\alpha}\|_{\mathcal{L}(L^2)} \leq \frac{\varepsilon\, C(s)}{\mathrm{d}(\Omega_\alpha, \Omega_\sigma)^{s - \frac{1+d}{2}}} \,, \quad \alpha \neq \sigma \,. \tag{1.4.12}$$

In the simplest case that the $\omega$ are independent parameters (forced case), it is not too difficult to impose that each self-adjoint operator $U_{\Omega_\alpha}^{\Omega_\alpha}$ is invertible for most parameters and that

$$\|(U_{\Omega_\alpha}^{\Omega_\alpha})^{-1}\|_{\mathcal{L}(L^2)} \leq M_\alpha^\tau \,, \quad \forall \alpha \,,$$

for some $\tau$ large enough. Since, by (H1), each cluster $\Omega_\alpha$ is *dyadic*, this $L^2$-estimate implies also a $\mathcal{H}^s$-Sobolev bound: for all $h \in \mathcal{H}_{\Omega_\alpha}$,

$$\begin{aligned} \|(U_{\Omega_\alpha}^{\Omega_\alpha})^{-1} h\|_s \leq M_\alpha^s \|(U_{\Omega_\alpha}^{\Omega_\alpha})^{-1} h\|_{L^2} \leq M_\alpha^s M_\alpha^\tau \|h\|_{L^2} &\leq \frac{M_\alpha^s}{m_\alpha^s} M_\alpha^\tau \|h\|_s \\ &\leq 2^s N^\tau \|h\|_s \end{aligned}$$

by (H1) and $M_\alpha \leq N$. As a consequence the whole operator $\mathcal{D}$ defined in (1.4.11) is invertible with a Sobolev estimate

$$\|\mathcal{D}^{-1} h\|_s \leq C(s) N^\tau \|h\|_s \,, \quad \forall h \in \mathcal{H}_N \,.$$

Finally, using the off-diagonal decay (1.4.12) and the "separation at infinity" property (1.4.9), the operator $\mathcal{D}^{-1}\mathcal{R}$ is bounded in $L^2$, and it is easy to reproduce a Neumann-series argument to prove the invertibility of $U = \mathcal{D}(I + \mathcal{D}^{-1}\mathcal{R})$ with tame estimates for the inverse $U^{-1}$ in Sobolev norms, implying (1.4.5).

### 1.4.2 Quasi-periodic case $\nu \geq 2$

In the quasi-periodic setting, i.e. $\nu \geq 2$, the proof that the operator $\mathcal{L}_N$ defined in (1.4.3) is invertible and its inverse satisfies the tame estimates (1.4.6) is more difficult. Indeed the 'separation at infinity" property (H2) never holds in the quasi-periodic case, neither for finite dimensional systems. For example, the operator $\omega \cdot \partial_\varphi$, is represented in the Fourier basis as the diagonal matrix $\mathrm{Diag}_{\ell \in \mathbb{Z}^\nu}(\mathrm{i}\omega \cdot \ell)$. If the frequency vector $\omega \in \mathbb{R}^\nu$ is Diophantine, then the singular sites $\ell \in \mathbb{Z}^\nu$ such that

$$|\omega \cdot \ell| \leq \rho$$

are "uniformly distributed" in a neighborhood of the hyperplane $\omega \cdot \ell = 0$, with nearby indices at distance $O(\rho^{-\alpha})$ for some $\alpha > 0$. Therefore, unlike in the time periodic case, the decomposition (1.4.7) into singular and regular sites of the unperturbed linear operator is not sufficient, and a finer analysis has to be performed.

In the sequel we follow the exposition of [22]. Let

$$A = D + \varepsilon T\,, \quad D := \mathrm{Diag}_{(\ell,j) \in \mathbb{Z}^\nu \times \mathbb{Z}^d}(-(\omega \cdot \ell)^2 + |j|^2 + m)\,, \quad T := (\widehat{b}_{\ell - \ell', j - j'})\,,$$

be the self-adjoint matrix in (1.4.2) that represents the second order operator (1.4.1). We suppose (forced case) that $\omega$ and $\varepsilon$ are unrelated and that $\omega$ is constrained to a fixed Diophantine direction

$$\omega = \lambda \bar{\omega}\,, \quad \lambda \in \left[\frac{1}{2}, \frac{3}{2}\right], \quad |\bar{\omega} \cdot \ell| \geq \frac{\gamma_0}{|\ell|^{\tau_0}}\,,\ \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\,. \tag{1.4.13}$$

A way to overcome the difficulty mentioned above, concerning the lack of "separation properties at infinity" of the singular sites, is to implement inductive arguments which prove, for most parameters, off-diagonal decay estimates for the inverses $A_{N_n}^{-1}$ of the restrictions

$$A_{N_n} := \Pi_{N_n} A_{|\mathcal{H}_{N_n}}\,, \tag{1.4.14}$$

at an increasing sequence of scales

$$N_n = \left[N_0^{\chi^n}\right], \quad n \geq 0\,, \tag{1.4.15}$$

for some $N_0$ and $\chi$ large enough, relying on information about the invertibility and the off-diagonal decay of "most" inverses $A_{N_{n-1},\ell_0,j_0}^{-1}$ of submatrices

$$A_{N_{n-1},\ell_0,j_0} := A_{|\ell - \ell_0| \leq N_{n-1}, |j - j_0| \leq N_{n-1}}$$

of size $N_{n-1}$. This program motivates the name 'multiscale analysis".

In order to give a precise statement we first introduce decay norms. Given a matrix $M = (M_i^{i'})_{i' \in B, i \in C}$, where $B, C$ are subsets of $\mathbb{Z}^{\nu+d}$, we define its $s$-norm

$$|M|_s^2 := \sum_{n \in \mathbb{Z}^{\nu+d}} [M(n)]^2 \langle n \rangle^{2s} \quad \text{where} \quad \langle n \rangle := \max(1, |n|)\,,$$

and

$$[M(n)] := \begin{cases} \max_{i-i'=n, i \in C, i' \in B} |M_i^{i'}| & \text{if } \ n \in C - B \\ 0 & \text{if } \ n \notin C - B\,. \end{cases}$$

**Remark 1.4.7.** *The norm $|T|_s$ of the matrix which represents the multiplication operator for the function $b(\varphi, x)$ is equal to $|T|_s = \|b\|_s$. Product of matrices (when it makes sense) with finite $s$-norm satisfy algebra and interpolation inequalities, see Appendix B.1.*

We now outline how to prove that the finite dimensional matrices $A_{N_n}$ in (1.4.14) are invertible for "most" parameters $\lambda \in [1/2, 3/2]$ and satisfy, for all $n \geq 0$,

$$|A_{N_n}^{-1}|_s \leq C(s) N_n^{\tau'} \big( N_n^{\varsigma s} + \|b\|_s \big)\,, \quad \varsigma \in (0, 1/2)\,, \ \tau' > 0\,, \ \forall s > s_0\,. \tag{1.4.16}$$

Such estimates imply the off-diagonal decay of the entries of the inverse matrix

$$|(A_{N_n}^{-1})_{i'}^i| \leq C(s) \frac{N_n^{\tau'} (N_n^{\varsigma s} + \|b\|_s)}{\langle i - i' \rangle^s}\,, \quad |i|, |i'| \leq N_n\,,$$

and Sobolev tame estimates as (1.4.6) (by Lemma B.1.9), assuming that $\|b\|_{s_1} \leq C$.

In order to prove (1.4.16) at the initial scale $N_0$ we impose that, for most parameters, the eigenvalues of the diagonal matrix $D$ satisfy

$$|-(\omega \cdot \ell)^2 + |j|^2 + m| \geq N_0^{-\tau}\,, \quad \forall |(\ell, j)| \leq N_0\,,$$

and, then, for $\varepsilon$ small enough, we deduce, by a direct Neumann argument, the invertibility of $A_{N_0} = D_{N_0} + \varepsilon T_{N_0}$ and the decay of the inverse $A_{N_0}^{-1}$.

In order to proceed at the higher scales, we use a multiscale analysis.

**$L^2$-bounds.** The first step is to show that, for "most" parameters, the eigenvalues of $A_{N_n}$ are in modulus bounded from below by $O(N_n^{-\tau})$ and so the $L^2$-norm of the inverse satisfies

$$\|A_{N_n}^{-1}\|_0 = O(N_n^{\tau}) \qquad \text{where} \qquad \| \ \|_0 := \| \ \|_{\mathcal{L}(L^2)}\,. \tag{1.4.17}$$

The proof is based on eigenvalue variation arguments. Recalling (1.4.13), dividing $A_{N_n}$ by $\lambda^2$, and setting $\xi := 1/\lambda^2$, we observe that the derivative with respect to $\xi$ satisfies

$$\partial_\xi (\xi A_{N_n}) = \mathrm{Diag}_{|(\ell,j)| \leq N_n} (|j|^2 + m) + O\big(\varepsilon \|T\|_0 + \varepsilon \|\partial_\lambda T\|_0\big) \geq c > 0\,, \tag{1.4.18}$$

for $\varepsilon$ small, i.e. it is positive definite. So, the eigenvalues $\mu_{\ell,j}(\xi)$ of the self-adjoint matrix $\xi A_{N_n}$ satisfy (see Lemma 4.8.1)

$$\partial_\xi \mu_{\ell,j}(\xi) \geq c > 0 \,, \quad \forall |(\ell,j)| \leq N_n \,,$$

which implies (1.4.17) except in a set of $\lambda$'s of measure $O(N_n^{-\tau+d+\nu})$.

**Remark 1.4.8.** *Monotonicity arguments for proving lower bounds for the moduli of the eigenvalues of (huge) self-adjoint matrices have been also used in [58], [23], [22]. Notice that the eigenvalues could be degenerate for some values of the parameters. This approach to verify "large deviation estimates" is very robust and the measure estimates that we perform at each step of the iteration are not inductive, as those in [40].*

**Multiscale Step.** The bounds (1.4.16) for the decay norms of $A_{N_n}^{-1}$ follow by an inductive application of the multiscale step proposition B.2.4, that we now describe. A matrix $A \in \mathcal{M}_E^E$, $E \subset \mathbb{Z}^{\nu+d}$, with $\mathrm{diam}(E) \leq 4N$ is called $N$-GOOD if it is invertible and

$$|A^{-1}|_s \leq N^{\tau'+\varsigma s} \,, \quad \forall s \in [s_0, s_1] \,,$$

for some $s_1 := s_1(d,\nu)$ large. Otherwise we say that $A$ is $N$-bad.

The aim of the multiscale step is to deduce that a matrix $A \in \mathcal{M}_E^E$ with

$$\mathrm{diam}(E) \leq N' := N^{\chi} \quad \text{with} \quad \chi \gg 1 \,, \tag{1.4.19}$$

is $N'$-good, knowing

- **(H1) (Off-diagonal decay)** $|A - \mathrm{Diag}(A)|_{s_1} \leq \Upsilon$ where $\mathrm{Diag}(A) := (\delta_{i,i'} A_i^{i'})_{i,i' \in E}$;
- **(H2) ($L^2$-bound)** $\|A^{-1}\|_0 \leq (N')^{\tau}$ where we set $\| \ \|_0 := \| \ \|_{\mathcal{L}(L^2)}$;

and suitable assumptions concerning the $N$-dimensional submatrices along the diagonal of $A$. We define an index $i \in E$ to be $(A,N)$-REGULAR if there is $F \subset E$ containing $i$ such that $\mathrm{diam}(F) \leq 4N$, $\mathrm{d}(i, E \backslash F) \geq N/2$ and $A_F^F$ is $N$-good. A site $i$ is $(A,N)$-BAD if it is singular (i.e. $|A_i^i| < \rho$) and $(A,N)$-REGULAR.

We suppose the following "mild" separation properties for the $(A,N)$-bad sites.

- **(H3) (Separation properties)** There is a partition of the $(A,N)$-bad sites $B = \cup_\alpha \Omega_\alpha$ with
$$\mathrm{diam}(\Omega_\alpha) \leq N^{C_1} \,, \quad \mathrm{d}(\Omega_\alpha, \Omega_\beta) \geq N^2 \,, \ \forall \alpha \neq \beta \,, \tag{1.4.20}$$
for some $C_1 := C_1(d,\nu) \geq 2$.

The multiscale step proposition B.2.4 deduces that $A$ is $N'$-good and (1.4.16) holds at the new scale $N'$, by (H1)-(H2)-(H3), with suitable relations between the constants $\chi$, $C_1$, $\varsigma$, $s_1$, $\tau$. Roughly, the main conditions on the exponents are

$$C_1 < \varsigma\chi \qquad \text{and} \qquad s_1 \gg \chi\,\tau\,.$$

The first means that the size $N^{C_1}$ of any bad cluster $\Omega_\alpha$ is small with respect to the size $N' := N^\chi$ of the matrix $A$. The second means that the Sobolev regularity $s_1$ is large enough to "separate" the resonance effects of two nearby bad clusters $\Omega_\alpha$, $\Omega_\beta$.

**Separation properties.** We apply the multiscale step Proposition B.2.4 to the matrix $A_{N_{n+1}}$. The key property to verify is (H3). A first key ingredient is the following co-variance property: consider the family of infinite dimensional matrices

$$A(\theta) := D(\theta) + \varepsilon T\,, \quad D(\theta) := \operatorname{diag}_{(\ell,j)\in\mathbb{Z}^\nu\times\mathbb{Z}^d}\Big( - (\omega\cdot\ell+\theta)^2 + |j|^2 + m\Big)\,,$$

and its $(2N+1)^{\nu+d}$-restrictions $A_{N,\ell_0,j_0}(\theta) := A_{|\ell-\ell_0|\le N, |j-j_0|\le N}(\theta)$ centered at any $(\ell_0, j_0) \in \mathbb{Z}^\nu\times\mathbb{Z}^d$. Since the matrix $T$ in (1.4.2) is Töplitz we have the co-variance property

$$A_{N,\ell_0,j_0}(\theta) = A_{N,j_0,0}(\theta + \omega\cdot\ell_0)\,. \tag{1.4.21}$$

In order to deduce (H3), it is sufficient to prove the "separation properties" (1.4.20) for the $N_n$-BAD sites of $A$, namely the indices $(\ell_0, j_0)$ which are singular

$$|-(\omega\cdot\ell_0)^2 + |j_0|^2 + m| \le \rho\,, \tag{1.4.22}$$

and for which there exists a site $(\ell, j)$, with $|(\ell,j)-(\ell_0,j_0)| \le N$, such that $A_{N_n,\ell,j}$ is $N_n$-bad. Such separation properties are obtained for all the parameters $\lambda$ which are $N_n$-good, namely such that

$$\begin{aligned}\forall\, j_0\in\mathbb{Z}^d\,, \quad B_{N_n}(j_0;\lambda) &:= \Big\{\theta\in\mathbb{R} \,:\, A_{N_n,0,j_0}(\theta) \text{ is } N_n - \text{bad}\Big\} \\ &\subset \bigcup_{q=1,\ldots,N_n^{C(d,\nu)}} I_q \ \text{ where } I_q \text{ are intervals with } |I_q| \le N_n^{-\tau}\,.\end{aligned} \tag{1.4.23}$$

We first use the covariance property (1.4.21) and the "complexity" information (1.4.23) to bound the number of "bad" time-Fourier components. Indeed

$$A_{N_n,\ell_0,j_0} \text{ is } N_n\text{-bad} \iff A_{N_n,0,j_0}(\omega\cdot\ell_0) \text{ is } N_n\text{-bad} \iff \omega\cdot\ell_0 \in B_{N_n}(j_0;\lambda)\,.$$

Then, using that $\omega$ is Diophantine, the complexity bound (1.4.23) implies that, for each fixed $j_0$, there are at most $O(N_n^{C(d,\nu)})$ sites $(\ell_0, j_0)$ in the larger box $|\ell_0| \le N_{n+1}$, which are $N_n$-bad.

Next, we prove that a $N_n^2$-"chain" of singular sites, i.e. a sequence of distinct integer vectors $(\ell_1, j_1), \ldots, (\ell_L, j_L)$ satisfying (1.4.22) and $|\ell_{q+1} - \ell_q| + |j_{q+1} - j_q| \leq N_n^2$, for any $q = 1, \ldots, L$, which are also $N_n$-bad, has a "length" $L$ bounded by $L \leq N_n^{C_1(d,\nu)}$. This implies a partition of the $(A_{N_{n+1}}, N_n)$-bad sites as in (1.4.20) at order $N_n$. In this step we require that $\omega \in \mathbb{R}^\nu$ satisfies the quadratic Diophantine condition

$$\Big| n + \sum_{1 \leq i \leq j \leq \nu} \omega_i \omega_j p_{ij} \Big| \geq \frac{\gamma_2}{|p|^{\tau_2}}, \quad \forall p := (p_{ij}) \in \mathbb{Z}^{\frac{\nu(\nu+1)}{2}}, \ \forall n \in \mathbb{Z}, \ (p, n) \neq (0, 0), \tag{1.4.24}$$

for some positive $\gamma_2, \tau_2$.

**Remark 1.4.9.** *The singular sites* (1.4.22) *are integer vectors close to a "cone" and* (1.4.24) *can be seen as an irrationality condition on its slopes. For NLS,* (1.4.24) *is not required, because the corresponding singular sites* $(\ell, j)$ *satisfy* $|\omega \cdot \ell + |j|^2| \leq C$, *i.e. they are close to a paraboloid. We refer to [29] to avoid the use of* (1.4.24) *for NLW.*

**Measure and "complexity" estimates.** In order to conclude the inductive proof we have to verify that "most" parameters $\lambda$ are $N_n$-good, according to (1.4.23). We prove first that, except a set of measure $O(N_n^{-1})$, all parameters $\lambda \in [1/2, 3/2]$ are $N_n$-good in a $L^2$-sense, namely

$$\begin{aligned} \forall\, j_0 \in \mathbb{Z}^d, \quad B^0_{N_n}(j_0; \lambda) &:= \Big\{ \theta \in \mathbb{R} \, : \, \| A^{-1}_{N_n, 0, j_0}(\theta) \|_0 > N_n^\tau \Big\} \\ &\subset \bigcup_{q = 1, \ldots, N_n^{C(d,\nu)}} I_q \ \text{ where } I_q \text{ are intervals with } |I_q| \leq N_n^{-\tau}. \end{aligned} \tag{1.4.25}$$

The proof is again based on eigenvalue variation arguments as in (1.4.18), using that $-\Delta + m$ is positive definite.

Finally, the multiscale Proposition step B.2.4, and the fact that the separation properties of the $N_n$-bad sites of $A(\theta)$ hold uniformly in $\theta \in \mathbb{R}$, imply inductively that the parameters $\lambda$ which are $N_n$-good in $L^2$-sense, are actually $N_n$-good, i.e. (1.4.23) holds, concluding the inductive argument.

### 1.4.3 The multiscale analysis of Chapter 4

In this Monograph we consider *autonomous* nonlinear wave equations (1.1.1). As a consequence the analysis of the quasi-periodic linear operator

$$(\omega \cdot \partial_\varphi)^2 - \Delta + V(x) + \varepsilon^2 (\partial_u g)(x, u(\varphi, x)), \quad \varphi \in \mathbb{T}^\nu, \ x \in \mathbb{T}^d, \tag{1.4.26}$$

obtained linearizing (1.2.1) at an approximate quasi-periodic solution, acting in a subspace orthogonal to the unperturbed linear 'tangential" solutions (see (1.2.33))

$$\Big\{ \sum_{j \in \mathbb{S}} \alpha_j \cos(\varphi_j) \Psi_j(x), \ \alpha_j \in \mathbb{R} \Big\} \subset \mathrm{Ker}\big( (\bar\mu \cdot \partial_\varphi)^2 - \Delta + V(x) \big),$$

is a much more difficult task than in the previous section. First of all, in (1.4.26), as discussed in remark 1.4.1, the frequency vector $\omega$ and the small parameter $\varepsilon$ are linked. In particular $\omega$ has to be $\varepsilon^2$-close to the unperturbed frequency vector $\bar{\mu} = (\mu_j)_{j \in \mathbb{S}}$ in (1.2.3). More precisely $\omega$ varies approximately according to the frequency-to-action map (1.2.11), and thus it has to belong to the region of admissible frequencies (1.2.23). Moreover the multiplicative function

$$(\partial_u g)(x, u(\varphi, x)) = 3a(x)(u(\varphi, x))^2 + \dots$$

in (1.4.26) (recall the form of the nonlinearity (1.2.2)) depends itself on $\omega$, via the approximate quasi-periodic solution $u(\varphi, x)$, and, as the tangential frequency vector $\omega$ changes, also the normal frequencies undergo a significant modification. At least for finitely many modes, the shift of the normal frequencies due to the effect of $3a(x)(u(\varphi, x))^2$ can be approximately described in terms of the Birkhoff matrix $\mathscr{B}$ in (1.2.9) as $\mu_j + \varepsilon^2(\mathscr{B}\xi)_j$ (see (1.5.22)). Because of all these constraints, positivity properties like (1.4.18) in general fail, see (1.5.24) and remark 1.5.3. This implies difficulties for imposing non-resonance conditions and for proving suitable complexity bounds.

An additional difficulty in (1.4.26) is the presence of the multiplicative potential $V(x)$, which is not diagonal in the Fourier basis.

In this Monograph we shall still be able to use positivity arguments for imposing non-resonance conditions and proving complexity bounds along the multiscale analysis. This requires to write (1.2.1) as a first order Hamiltonian system and to perform a block-reduction of the corresponding first order quasi-periodic linear operator acting in the normal subspace. We refer to the next section 1.5 for an explanation about this procedure and here we limit ourselves to describe the resulting quasi-periodic operators that we shall analyze in Chapter 4 with multiscale techniques.

The multiscale Proposition 4.1.5 of Chapter 4 provides the existence of a right inverse of finite dimensional restrictions of self-adjoint linear operators of the form:

1. for $\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon$, $\lambda \in \widetilde{\Lambda} \subset \Lambda$,
$$\mathcal{L}_r = J\omega \cdot \partial_\varphi + D_V + r(\varepsilon, \lambda, \varphi) \quad \text{acting on } h \in (L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d, \mathbb{R}))^2 \,, \tag{1.4.27}$$
where $D_V := \sqrt{-\Delta + V(x)}$, $J$ is the symplectic matrix
$$J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix} ,$$
and $r(\varepsilon, \lambda, \varphi)$ is a self-adjoint operator with polynomial off-diagonal decay;

2. for $\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon$, $\lambda \in \widetilde{\Lambda} \subset \Lambda$,
$$\begin{gathered} \mathcal{L}_{r,\mu} = J\omega \cdot \partial_\varphi + D_V + \mu(\varepsilon, \lambda)\mathcal{J}\Pi_{\mathbb{G}} + r(\varepsilon, \lambda, \varphi) \\ \text{acting on } h \in (L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d, \mathbb{R}))^4 \,, \end{gathered} \tag{1.4.28}$$

where $\mathcal{J}h := JhJ$ and the left/right action of $J$ on $\mathbb{R}^4$ are defined in (4.1.3), (4.1.4), $\mu(\varepsilon, \lambda)$ is a scalar, $\Pi_{\mathbb{G}}$ is the projector on the infinite dimensional subspace

$$H_{\mathbb{G}} := H_{\mathbb{S}\cup\mathbb{F}}^{\perp}, \quad H_{\mathbb{S}\cup\mathbb{F}} := \Big\{ \sum_{j\in\mathbb{S}\cup\mathbb{F}} (Q_j, P_j)\Psi_j(x)\,,\ (Q_j, P_j) \in \mathbb{R}^2 \Big\}$$

(the set $\mathbb{G} \subset \mathbb{N}$ is defined in (1.2.15)) and $r(\varepsilon, \lambda, \varphi)$ is a self-adjoint operator with polynomial off-diagonal decay.

The operators (1.4.27), (1.4.28) are defined for all $\lambda$ belonging to a subset $\widetilde{\Lambda} \subset \Lambda$ (which may shrink during the Nash-Moser iteration).

Notice that (1.4.27), (1.4.28) are first order vector valued quasi-periodic operators, unlike (1.4.26) which is a second order scalar quasi-periodic operator (it acts in "configuration space"). Since (1.4.27) arises (essentially) by writing the linear wave operator (1.4.26) as a first order system as in (1.3.27), where the non-diagonal vector field is 1-smoothing, it is natural to require that the operators $r = r(\varepsilon, \lambda, \varphi)$ in (1.4.27), (1.4.28) satisfy the decay condition

$$|r|_{+,s_1} := |D_m^{\frac12} r D_m^{\frac12}|_{s_1} = O(\varepsilon^2) \quad \text{where} \quad D_m := \sqrt{-\Delta + m} \tag{1.4.29}$$

for some $m > 0$.

A key property of the operators $\mathcal{L}_r$ and $\mathcal{L}_{r,\mu}$, that we shall be able to verify at each step of the Nash-Moser iteration, is the monotonicity property in item 3 of Definition 4.1.2. For $\mathcal{L}_r$ it is

$$\mathfrak{d}_\lambda\Big(\frac{D_V + r}{1 + \varepsilon^2\lambda}\Big) \leq -c\,\varepsilon^2 \mathrm{Id}\,, \quad c > 0\,, \tag{1.4.30}$$

where we use the notation, given a family of linear self-adjoint operators $A(\lambda)$,

$$\mathfrak{d}_\lambda A(\lambda) \leq -c\,\mathrm{Id} \qquad \Longleftrightarrow \qquad \frac{A(\lambda_2) - A(\lambda_1)}{\lambda_2 - \lambda_1} \leq -c\,\mathrm{Id}\,, \quad \forall \lambda_1 \neq \lambda_2\,, \tag{1.4.31}$$

and $A \leq -c\,\mathrm{Id}$ means as usual $(Aw, w)_{L^2} \leq -c\|w\|_{L^2}^2$. The condition for $\mathcal{L}_{r,\mu}$ is similar. Such property allows to prove the measure estimates stated in Properties 1-3 of Proposition 4.1.5.

All the precise assumptions on the operators $\mathcal{L}_r$ in (1.4.27) and $\mathcal{L}_{r,\mu}$ in (1.4.28) are stated in Definition 4.1.2 (we neglect in (1.4.27), (1.4.28) the projector $\mathfrak{c}\Pi_{\mathbb{S}}$ which has a purely technical role, see remark 4.1.3).

**Remark 1.4.10.** *We shall use the results of the multiscale Proposition 4.1.5 about the approximate invertibility of the operator* $\mathcal{L}_r$ *in* (1.4.27) *in Chapter 10, and, for the operator* $\mathcal{L}_{r,\mu}$ *defined in* (1.4.28)*, in Chapter 9. In the next section 1.5 we shall describe their role in the proof of Theorem 1.2.1.*

The application of multiscale techniques to the operators (1.4.27), (1.4.28) is more involved than for the second order quasi-periodic operator (1.4.1) described in the previous section. For definiteness we consider $\mathcal{L}_{r,\mu}$. In the Fourier exponential basis, the operator $\mathcal{L}_{r,\mu}$ is represented by a self-adjoint matrix

$$\mathtt{A} = \mathtt{D}_\omega + \mathtt{T} \tag{1.4.32}$$

with a diagonal part (see (4.2.4))

$$\mathtt{D}_\omega = \mathrm{Diag}_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|+d}} \begin{pmatrix} \langle j\rangle_m - \mu & \mathrm{i}\omega\cdot\ell & 0 & 0 \\ -\mathrm{i}\omega\cdot\ell & \langle j\rangle_m - \mu & 0 & 0 \\ 0 & 0 & \langle j\rangle_m + \mu & \mathrm{i}\omega\cdot\ell \\ 0 & 0 & -\mathrm{i}\omega\cdot\ell & \langle j\rangle_m + \mu \end{pmatrix}, \tag{1.4.33}$$

where $\langle j\rangle_m := \sqrt{|j|^2+m}$ for some $m>0$ and $\mu := \mu(\varepsilon,\lambda)$. The matrix $\mathtt{T}$ in (1.4.32) is

$$\mathtt{T} := (\mathtt{T}_{\ell,j}^{\ell',j'})_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|+d},(\ell',j')\in\mathbb{Z}^{|\mathbb{S}|+d}}\,,$$
$$\mathtt{T}_{\ell,j}^{\ell',j'} := (D_V - D_m)_j^{j'} - \mu[\mathcal{J}\Pi_{\mathbb{S}\cup\mathbb{F}}]_j^{j'} + r_{\ell,j}^{\ell',j'} \in \mathrm{Mat}(4\times 4)\,.$$

The matrix $\mathtt{T}$ is *Töplitz* in $\ell$, namely $\mathtt{T}_{\ell,j}^{\ell',j'}$ depends only on the indices $\ell-\ell', j, j'$.

The infinitely many eigenvalues of the matrix $\mathtt{D}_\omega$ are

$$\sqrt{|j|^2+m} \pm \mu \pm \omega\cdot\ell\,, \quad j\in\mathbb{Z}^d\,,\ \ell\in\mathbb{Z}^{|\mathbb{S}|}\,.$$

By Proposition 3.4.1 and Lemma 3.3.8, the matrix $\mathtt{T}$ satisfies the off-diagonal decay, see (4.2.6),

$$|\mathtt{T}|_{+,s_1} := |D_m^{1/2}\mathtt{T}D_m^{1/2}|_{s_1} < +\infty\,. \tag{1.4.34}$$

We introduce the index $\mathfrak{a}\in\mathfrak{I} := \{1,2,3,4\}$ to distinguish each component in the $4\times 4$ matrix (1.4.33). Then the singular sites $(\ell,j,\mathfrak{a})\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{Z}^d\times\mathfrak{I}$ of $\mathcal{L}_{r,\mu}$ are those integer vectors such that, for some choice of the signs $\sigma_1(\mathfrak{a}),\sigma_2(\mathfrak{a})\in\{-1,1\}$,

$$\big|\sqrt{|j|^2+m} + \sigma_1(\mathfrak{a})\mu + \sigma_2(\mathfrak{a})\omega\cdot\ell\big| < \frac{\Theta}{\sqrt{|j|^2+m}}\,. \tag{1.4.35}$$

In (1.4.35) the constant $\Theta := \Theta(V)$ is chosen large enough depending on the multiplicative potential $V(x)$ (which is not small). Note that, if $(\ell,j,\mathfrak{a})$ is singular, then we recover the second order bound

$$\big||j|^2 + m - (\sigma_1(\mathfrak{a})\mu + \sigma_2(\mathfrak{a})\omega\cdot\ell)^2\big| \le C\Theta\,,$$

similarly to (1.4.22).

The invertibility of the restricted operator $\mathcal{L}_{r,\mu,N} := \Pi_N(\mathcal{L}_{r,\mu})_{|\mathcal{H}_N}$ and the proof of the off-diagonal decay estimates (4.1.23) is obtained in section 4.7 by an inductive application of the multiscale step Proposition 4.3.4, which is deduced by the corresponding Proposition B.2.4. On the other hand, section 4.6 contains the proof of the existence of a right inverse of $[\mathcal{L}_{r,\mu}]_N^{2N} := \Pi_N(\mathcal{L}_{r,\mu})_{|\mathcal{H}_{2N}}$ satisfying (4.1.20) at the small scales $N < N(\varepsilon)$ (notice that the size of $N$ depends on $\varepsilon$).

In view of the multiscale proof, we consider the family of operators

$$\mathcal{L}_{r,\mu}(\theta) := J\omega \cdot \partial_\varphi + D_V + \mathrm{i}\theta J + \mu \mathcal{J} \Pi_{\mathbb{G}} + r \,, \quad \theta \in \mathbb{R} \,, \tag{1.4.36}$$

which is represented, in Fourier basis, by a self-adjoint matrix denoted $\mathtt{A}(\theta)$.

The monotonicity assumption in item 3 of Definition 4.1.2, see (1.4.30), allows to obtain effective measure estimates and complexity bounds similar to (1.4.25). Notice however that by (1.4.30) the eigenvalues of $\mathcal{L}_{r,\mu,N}$ vary in $\lambda$ with only a $O(\varepsilon^2)$-speed, creating further difficulties with respect to the previous section.

The verification of assumption (H3) of the multiscale step Proposition 4.3.4, concerning separation properties of the $N$-bad sites, is a key part of the analysis. The new notion of $N$-bad site is introduced in Definition 4.4.2, according to the new definition of $N$-good matrix introduced in Definition 4.3.1 (the difference is due to the fact that the singular sites are defined as in (1.4.35) and (1.4.34) holds).

Then, in section 4.4, we prove that a $\Gamma$-chain of singular sites is not too long, under the bound (4.4.9) for its time components, see Lemma 4.4.8. Finally in Lemma 4.4.15 we are able to conclude, for the parameters $\lambda$ which are $N$-good, i.e.

$$\begin{gathered} \forall\, j_0 \in \mathbb{Z}^d \,, \ B_N(j_0;\lambda) := \Big\{ \theta \in \mathbb{R} \, : \, \mathtt{A}_{N,0,j_0}(\theta) \text{ is } N\text{-bad} \Big\} \subset \bigcup_{q=1,\dots,N^{C(d,|\mathbb{S}|,\tau_0)}} I_q \,, \\ \text{where } I_q \text{ are intervals with measure } |I_q| \leq N^{-\tau} \,, \end{gathered} \tag{1.4.37}$$

(Definition 4.4.4) an upper bound $L \leq N^{C_3}$ (see (4.4.40)) for the length $L$ of a $N^2$-chain of $N$-bad sites. This result implies the partition of the $N$-bad sites into clusters separated by $N^2$, according to the assumption (H3) of the multiscale step Proposition 4.3.4.

We remark that Lemma 4.4.15 requires a significant improvement with respect to the arguments in [22], described in the previous section: the exponent $\tau$ in (1.4.37) is large, but independent of $\chi$, which defines the new scale $N' = N^\chi$ in the multiscale step, see remark 4.4.6. This improvement is required by the fact that by (1.4.30) the eigenvalues of $\mathcal{L}_{r,\mu,N}$ vary in $\lambda$ with only a $O(\varepsilon^2)$-speed.

We finally comment why the multiscale analysis works also for operators $\mathcal{L}_r$, $\mathcal{L}_{r,\mu}$ of the form (1.4.27), (1.4.28), where $D_V$ is not a Fourier multiplier. The reason is similar to [23], [22]. In Lemma 4.7.3 we have to prove that *all* the parameters $\lambda \in \Lambda$ are $N$-good (Definition 4.4.4) at the small scales $N \leq N_0$. We proceed as follows. We regard the operator $\mathcal{L}_{r,\mu}(\theta)$ in (1.4.36) as a small perturbation of the operator

$$\mathcal{L}_{0,\mu}(\theta) = J\omega \cdot \partial_\varphi + D_V + \mathrm{i}\theta J + \mu \mathcal{J} \Pi_{\mathbb{G}}$$

which is $\varphi$-independent. Thus a lower bound on the modulus of the eigenvalues of its restriction to $\mathcal{H}_N$ implies an estimate in $L^2$ norm of its inverse, and, thanks to *separation properties of the eigenvalues* $|j|^2$ *of* $-\Delta$, also a bound of its $s$-decay norm as $O(N^{\tau'+\varsigma s})$. By a Neumann perturbative argument this bound also persists for $\Pi_N \mathcal{L}_{r,\mu}(\theta)_{|\mathcal{H}_N}$, taking $\varepsilon$ small enough, up to the scales $N \leq N_0$. The proof is done precisely in Lemmata 4.7.2-4.7.3. The set of $\theta$ such that the spectrum of $\Pi_N \mathcal{L}_{0,\mu}(\theta)_{|\mathcal{H}_N}$ is at a distance $O(N^{-\tau})$ from 0 is contained into a union of intervals like (1.4.37), implying the claimed complexity bounds. The proof at higher scales follows by the induction multiscale process.

## 1.5 Outline of proof of Theorem 1.2.1

In this section we present in detail the plan of proof of Theorem 1.2.1, which occupies Chapters 2-11. This section is a road map through the technical aspects of proof.

In Chapter 2 we first write the second order wave equation (1.2.1) as the first order Hamiltonian system (2.1.13)-(2.1.14),

$$\begin{cases} q_t = D_V p \\ p_t = -D_V q - \varepsilon^2 D_V^{-\frac{1}{2}} g(\varepsilon, x, D_V^{-\frac{1}{2}} q)\,, \end{cases} \tag{1.5.1}$$

where $D_V := \sqrt{-\Delta + V(x)}$ is defined spectrally in (2.1.11), and the variables $(q,p)$ belong to a dense subspace of

$$H = L^2(\mathbb{T}^d, \mathbb{R}) \times L^2(\mathbb{T}^d, \mathbb{R})\,.$$

We prove polynomial off-diagonal decay of $D_V$ in section 3.4.

Fixed finitely many tangential sites $\mathbb{S} \subset \mathbb{N}$, we decompose (see (2.2.2)) the canonical variables

$$(q,p) = \sum_{j\in\mathbb{S}} (q_j, p_j)\Psi_j(x) + (Q,P)\,, \quad (q_j,p_j) \in \mathbb{R}^2\,,$$

into "tangential" and "normal" components, where $\Psi_j(x)$ are the eigenfunctions of the Sturm-Liouville operator $-\Delta + V(x)$ defined in (1.1.5), and $(Q,P)$ belong to the subspace $H_{\mathbb{S}}^{\perp}$, called the normal subspace, $L^2$-orthogonal of

$$H_{\mathbb{S}} := \Big\{ \sum\nolimits_{j\in\mathbb{S}} (q_j,p_j)\Psi_j(x)\,,\, (q_j,p_j) \in \mathbb{R}^2 \Big\}\,. \tag{1.5.2}$$

The dynamics of (1.5.1) on the symplectic subspaces $H_{\mathbb{S}}$ and $H_{\mathbb{S}}^{\perp}$ is handled quite differently. On the subspace $H_{\mathbb{S}}$ we introduce action-angle variables $(\theta, I)$, by setting (see (2.2.3))

$$(q_j,p_j) := \sqrt{2I_j}\big(\cos\theta_j, -\sin\theta_j\big)\,, \quad \forall j \in \mathbb{S}\,.$$

In these new coordinates the solutions (1.2.4) of the linear wave equation (1.1.4) are described as the continuous family of quasi-periodic solutions

$$\theta(t) = \bar{\mu} t\,,\ I(t) = \xi\,,\ Q(t) = P(t) = 0\,, \quad \xi \in \mathbb{R}_+^{|\mathbb{S}|}\,, \tag{1.5.3}$$

with frequency vector $\bar{\mu} = (\mu_j)_{j\in\mathbb{S}}$, which is independent of the unperturbed actions $\xi \in \mathbb{R}_+^{|\mathbb{S}|}$. Notice that, by the assumption (1.2.6), the unperturbed tangential frequency vector $\bar{\mu}$ is Diophantine. Introducing the translated action variable $y$ by setting

$$I = \xi + y\,, \quad y \in \mathbb{R}^{|\mathbb{S}|}\,,$$

the quasi-periodic solutions (1.5.3) densely fill the invariant torus

$$\mathbb{T}_\theta^{|\mathbb{S}|} \times \{0\}_y \times \{(0,0)\}_{(Q,P)}\,.$$

In the variables $(\theta, y, Q, P)$ the Hamiltonian system (1.5.1) assumes the form

$$\begin{cases} \dot{\theta} - \bar{\mu} - \varepsilon^2 \partial_y R(\theta, y, Q, \xi) = 0 \\ \dot{y} + \varepsilon^2 \partial_\theta R(\theta, y, Q, \xi) = 0 \\ \big(\partial_t - J D_V\big)(Q,P) - \varepsilon^2\big(0, \nabla_Q R(\theta, y, Q, \xi)\big) = 0\,, \end{cases} \tag{1.5.4}$$

see (2.2.12), where $J$ is the symplectic matrix

$$J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}\,,$$

and $R(\theta, y, Q, \xi)$ is the Hamiltonian in (2.2.8).

The goal is now to look, for $\varepsilon$ sufficiently small, for quasi-periodic solutions

$$\big(\omega t + \vartheta(\omega t), y(\omega t), Q(\omega t), P(\omega t)\big) \tag{1.5.5}$$

of the nonlinear Hamiltonian system (1.5.4), with a frequency vector $\omega \in \mathbb{R}^{|\mathbb{S}|}$, $O(\varepsilon^2)$-close to $\bar{\mu}$, to be determined, and where the function

$$\varphi \mapsto (\vartheta(\varphi), y(\varphi), Q(\varphi), P(\varphi)) \in \mathbb{R}^{|\mathbb{S}|} \times \mathbb{R}^{|\mathbb{S}|} \times H_{\mathbb{S}}^{\perp}$$

is periodic in the variable $\varphi = (\varphi_j)_{j\in\mathbb{S}} \in \mathbb{T}^{|\mathbb{S}|}$, and close to $(0,0,0,0)$.

We shall be able to constrain the frequency vector $\omega$ in (1.5.5) to a fixed "admissible" direction $\bar{\omega}_\varepsilon$, as stated in (1.2.24)-(1.2.25), namely

$$\omega = (1 + \varepsilon^2 \lambda)\bar{\omega}_\varepsilon\,, \quad \lambda \in \Lambda = [-\lambda_0, \lambda_0]\,, \tag{1.5.6}$$

where $\bar{\omega}_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ satisfies the Diophantine conditions (1.2.29) and (1.2.30). In Lemma 2.3.1 we prove that these conditions are actually satisfied by "most" vectors $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2 \zeta$ close to $\bar{\mu}$. We shall use the 1-dimensional parameter $\lambda$, which corresponds to a time-rescaling, in order to impose all the non-resonance conditions required by our KAM construction, in particular along the multiscale analysis of the linearized operator. Notice that $\lambda$ has to be considered as an "internal" parameter of the wave equation (1.1.1).

**Remark 1.5.1.** *The existence of quasi-periodic solutions with tangential frequencies constrained along a fixed direction had been proved, for finite dimensional autonomous Hamiltonian systems, by Eliasson [53] and Bourgain [36], and for* 1*-d nonlinear autonomous wave and Schrödinger equations in [16]. Results in the easier case of quasi-periodically forced PDEs, where* $\omega$ *is an external parameter constrained to a fixed direction, have been obtained in [23] for NLS, [22] for NLW, [6] for KdV.*

The search of an embedded invariant torus

$$\varphi \mapsto i(\varphi) = (\varphi, 0, 0) + (\vartheta(\varphi), y(\varphi), (Q, P)(\varphi)) \,,$$

of the Hamiltonian system (1.5.4), supporting quasi-periodic solutions with frequency $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$ as in (1.5.6), amounts to solve the functional equation $\mathcal{F}(\lambda; i) = 0$ where $\mathcal{F}$ is the nonlinear operator defined in (5.1.2),

$$\mathcal{F}(\lambda; i) = \begin{pmatrix} \omega\cdot\partial_\varphi\vartheta(\varphi) + \omega - \bar{\mu} - \varepsilon^2(\partial_y R)(i(\varphi), \xi) \\ \omega\cdot\partial_\varphi y(\varphi) + \varepsilon^2(\partial_\theta R)(i(\varphi), \xi) \\ \omega\cdot\partial_\varphi(Q, P)(\varphi) - JD_V(Q, P)(\varphi) - \varepsilon^2\big(0, (\nabla_Q R)(i(\varphi), \xi)\big) \end{pmatrix} . \tag{1.5.7}$$

The operator $\mathcal{F}$ acts on the Sobolev scale of spaces

$$(\vartheta(\varphi), y(\varphi), (Q, P)(\varphi)) \in H^s_\varphi(\mathbb{T}^{|\mathbb{S}|}, \mathbb{R}^{|\mathbb{S}|}) \times H^s_\varphi(\mathbb{T}^{|\mathbb{S}|}, \mathbb{R}^{|\mathbb{S}|}) \times (\mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d, \mathbb{R}^2) \cap H^\perp_{\mathbb{S}}) \,,$$

for $s \geq s_0$, losing one $(\varphi, x)$-derivative due to the unbounded operators $\omega\cdot\partial_\varphi$ and $D_V$.

We have that

$$\mathcal{F}(\lambda; \varphi, 0, 0, 0) = O(\varepsilon^2) \,, \quad \forall \lambda \in \Lambda \,.$$

Then, in section 5.3, using the unperturbed first order Melnikov non-resonance condition (1.2.7), we are able to construct the first approximate solution $i_1(\varphi)$ such that

$$\mathcal{F}(\lambda; i_1(\varphi)) = O(\varepsilon^4) \,, \quad \forall \lambda \in \Lambda \,.$$

Next, in Theorem 5.1.2, which contains the core of the result, we construct, by means of a Nash-Moser implicit function iterative scheme, for "most" values of $\lambda \in \Lambda$, a solution $i_\infty(\lambda; \varphi)$ of the equation

$$\mathcal{F}(\lambda; i_\infty(\lambda; \varphi)) = 0 \,,$$

thus an invariant torus of (1.5.4) with frequency $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$. The iteration is performed in Chapter 11. In particular, in Theorem 11.2.1 we construct a sequence of approximate solutions $i_n(\lambda; \varphi)$ which converges to $i_\infty(\lambda; \varphi)$ for $\lambda$ in a set of large measure. The key point is to prove the approximate invertibility of the linearized operators

$$d_i\mathcal{F}(\lambda; i_n(\lambda; \varphi)) \,,$$

obtained at any approximate quasi-periodic solution $i_n := i_n(\lambda; \varphi)$ along the Nash-Moser iteration, for a large set of $\lambda$'s, together with suitable tame estimates for the approximate inverse in high Sobolev norms, with of course loss of derivatives, due to the small divisors. This is achieved by the analysis performed in Chapters 6-10.

In Chapter 6 we implement the general strategy proposed in [24] where, instead of (approximately) inverting $d_i\mathcal{F}(\lambda; i_n)$, where all the $(\vartheta, y, Q, P)$ components are coupled by the differential of the nonlinear term $(-\partial_y R, \partial_\theta R, 0, -\nabla_Q R)$ in (1.5.7), we introduce suitable symplectic coordinates

$$(\phi, \zeta, w) \in \mathbb{T}^{|\mathbb{S}|} \times \mathbb{R}^{|\mathbb{S}|} \times H_{\mathbb{S}}^{\perp}$$

in which it is sufficient to (approximately) invert the linear operator $\mathbb{D}(i_n)$ defined in (6.1.22). The advantage is that the components of the operator $\mathbb{D}(i_n)$ can be inverted in a triangular way, that is, first one inverts the operator in the tangential action component $\widehat{\zeta}$, then the normal one for $\widehat{w} \in H_{\mathbb{S}}^{\perp}$, and finally the operator for the tangential angles $\widehat{\phi}$. This construction is implemented in detail in Chapter 6, completed with the results reported in Appendix C.

**Remark 1.5.2.** *The above decomposition is deeply related to the Nash-Moser approach for isotropic tori of finite dimensional Hamiltonian systems in Herman-Fejoz [64]. A related construction for reversible PDEs is performed in [47].*

After this transformation, the main issue is reduced to prove the approximate invertibility of the linear operator $\mathcal{L}_\omega(i_n)$ defined in (6.1.23) which acts on functions $h : \mathbb{T}^{|\mathbb{S}|} \to H_{\mathbb{S}}^{\perp}$ with values in the normal subspace $H_{\mathbb{S}}^{\perp}$. As proved in Lemma 6.1.2, this operator has the form

$$\mathcal{L}_\omega(i_n) = \omega \cdot \partial_\varphi - J(D_V + \varepsilon^2 \mathtt{B}(\varphi) + \mathtt{r}_\varepsilon(\varphi)), \quad D_V = \sqrt{-\Delta + V(x)}, \tag{1.5.8}$$

where $\mathtt{B}(\varphi)$ is the self-adjoint operator defined in (6.1.25) and $\mathtt{r}_\varepsilon(\varphi)$ is a self-adjoint remainder of size $O(\varepsilon^4)$, more precisely it satisfies the quantitative bounds (6.1.26)-(6.1.27). Notice that $J(D_V + \varepsilon^2 \mathtt{B}(\varphi) + \mathtt{r}_\varepsilon(\varphi))w$ is a linear Hamiltonian vector field and that the corresponding quasi-periodic operator (1.5.8) is a small deformation of the operator obtained linearizing the normal component of $\mathcal{F}$, defined in (1.5.7), at an approximate solution. A quasi-periodic operator as (1.5.8) is called Hamiltonian, see Definition 3.2.1.

Chapters 7-10 are devoted to prove the existence of an approximate right inverse of $\mathcal{L}_\omega(i_n)$, as stated in Proposition 11.1.1, for "most" values of $\lambda \in \Lambda$. This is obtained in several steps.

**Remark 1.5.3.** *We cannot directly apply to the operator* $\frac{1}{1+\varepsilon^2\lambda}\mathcal{L}_\omega(i_n)$ *the approach developed for the quasi-periodically forced NLW and NLS in [23], [22], [26], described in section 1.4.2. Actually, since* $\partial_\lambda \mathtt{B}(\varphi) = O(1)$ *the operator* $\partial_\lambda(\frac{1}{1+\varepsilon^2\lambda}\mathcal{L}_\omega(i_n))$ *is not positive or* *negative definite, posing a serious difficulty for verifying that the eigenvalues of its finite dimensional restrictions are in modulus bounded away from zero for most values of the parameter $\lambda$. In order to overcome this problem we perform the splitting of the normal subspace that we describe below.*

In Chapter 7, using the unperturbed second order Melnikov conditions (1.2.16)-(1.2.19), we first apply an averaging procedure to conjugate $\frac{1}{1+\varepsilon^2\lambda}\mathcal{L}_\omega(i_n)$ to the Hamiltonian operator (7.3.4)-(7.3.5) which has the form

$$\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(\mathtt{A}_0 + \varrho^+) \tag{1.5.9}$$

where $\mathtt{A}_0$ is split admissible according to Definition 8.1.1 (see Lemma 8.1.2) and the coupling operator $\varrho^+$ has size $O(\varepsilon^4)$ in the $|\ |_{+,s_1}$-norm defined in (1.4.29) (more precisely it satisfies $|\varrho^+|_{\mathrm{Lip},+,s_1} = O(\varepsilon^4)$ where the norm $|\ |_{\mathrm{Lip},+,s_1}$ is introduced in Definition 3.3.4). The main feature of a split-admissible operator is to be self-adjoint and block diagonal with respect to the orthogonal splitting

$$\begin{aligned} H_{\mathbb{S}}^\perp &= H_{\mathbb{F}} \oplus H_{\mathbb{G}}\,, \\ H_{\mathbb{F}} &:= \Big\{ \sum\nolimits_{j\in\mathbb{F}} (q_j,p_j)\Psi_j(x)\,,\ (q_j,p_j)\in\mathbb{R}^2 \Big\}\,, \\ H_{\mathbb{G}} &:= \Big\{ \sum\nolimits_{j\in\mathbb{G}} (q_j,p_j)\Psi_j(x)\,,\ (q_j,p_j)\in\mathbb{R}^2 \Big\}\,, \end{aligned} \tag{1.5.10}$$

i.e. of the form

$$\mathtt{A}_0 = \begin{pmatrix} D_0(\varepsilon,\lambda) & 0 \\ 0 & V_0(\varepsilon,\lambda,\varphi) \end{pmatrix}\,, \tag{1.5.11}$$

and, in the basis of the eigenfunctions $\{(\Psi_j,0), (0,\Psi_j)\}_{j\in\mathbb{F}}$, $D_0(\varepsilon,\lambda)$ is a diagonal operator

$$D_0(\varepsilon,\lambda) = \mathrm{Diag}_{j\in\mathbb{F}}\, \mu_j(\varepsilon,\lambda)\mathrm{Id}_2\,, \quad \mu_j(\varepsilon,\lambda)\in\mathbb{R}\,, \tag{1.5.12}$$

with eigenvalues $\mu_j(\varepsilon,\lambda) = \mu_j + O(\varepsilon^2)$ (the $\mu_j$ are defined in (1.1.5)) satisfying the non-degeneracy conditions (8.1.5)-(8.1.7) and, for all $j\in\mathbb{F}$, the monotonicity property

$$\begin{cases} \mathfrak{d}_\lambda\big(V_0(\varepsilon,\lambda) + \mu_j(\varepsilon,\lambda)\mathrm{Id}\big) \le -c_1\varepsilon^2 \\ \mathfrak{d}_\lambda\big(V_0(\varepsilon,\lambda) - \mu_j(\varepsilon,\lambda)\mathrm{Id}\big) \le -c_1\varepsilon^2 \end{cases} \tag{1.5.13}$$

where we use the notation (1.4.31).

The subsets $\mathbb{F}$ and $\mathbb{G}$ of the normal sites $\mathbb{S}^c$ in (1.5.10) are defined in (1.2.15). Recall that

$$\mathbb{F}\cup\mathbb{G} = \mathbb{S}^c\,, \quad \mathbb{F}\cap\mathbb{G} = \emptyset\,.$$

It is for proving Lemma 8.1.2, i.e. that $\mathtt{A}_0$ is a split-admissible operator, that we need the precise knowledge of how the tangential and normal frequencies are shifted by the nonlinearity $a(x)u^3 + O(u^4)$ (via the Birkhoff matrices $\mathscr{A}, \mathscr{B}$ in (1.2.9)), we use the additional second order Melnikov non-resonance conditions (1.2.18)-(1.2.19), the specific definition (1.2.15) of the subsets $\mathbb{F}$ and $\mathbb{G}$, and the non-degeneracy conditions (1.2.21)-(1.2.22). More precisely we use (1.2.18)-(1.2.19) and (1.2.15) to prove the monotonicity property (1.5.13), and the non-degeneracy conditions (1.2.21)-(1.2.22) to prove (8.1.5)-(8.1.7). The properties (8.1.5)-(8.1.8) allow to prove that the non-resonance conditions required along the multi-scale analysis of the linearized operator are fulfilled for a large set of values of the parameter $\lambda$. We comment about this issue below, around (1.5.27).

The quasi-periodic Hamiltonian operator $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(\mathtt{A}_0 + \varrho^+)$ in (1.5.9) is in a suitable form to apply Proposition 8.2.1, in order to prove that it admits an approximate right inverse. Proposition 8.2.1 is proved in Chapters 9 and 10.

In Chapter 9 we block-diagonalize $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(\mathtt{A}_0 + \varrho^+)$, according to the splitting $H_{\mathbb{S}}^\perp = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, up to a very small coupling term, see Corollary 9.1.2. More precisely, we conjugate, via symplectic transformations, the Hamiltonian operator $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(\mathtt{A}_0 + \varrho^+)$ in (1.5.9), where the block non-diagonal term $\varrho^+$ has size $O(\varepsilon^4)$ in $| \ |_{+,s_1}$ norm, to the quasi-periodic Hamiltonian operator in (9.1.34),

$$\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_{\mathtt{n}} + \rho_{\mathtt{n}}) \,, \tag{1.5.14}$$

where the operator $A_{\mathtt{n}}$ is split admissible, thus of the form (1.5.11), with

$$\begin{gathered} \mathtt{A}_{\mathtt{n}} = \begin{pmatrix} D_{\mathtt{n}}(\varepsilon, \lambda) & 0 \\ 0 & V_{\mathtt{n}}(\varepsilon, \lambda, \varphi) \end{pmatrix} , \\ \|D_{\mathtt{n}} - D_0\|_0 = o(\varepsilon^2) \,, \ \|V_{\mathtt{n}} - V_0\|_0 = o(\varepsilon^2) \,, \end{gathered} \tag{1.5.15}$$

(the norm $\| \ \|_0 = \| \ \|_{\mathcal{L}(L^2)}$ is the operatorial $L^2$ norm) and the block non-diagonal self-adjoint term $\rho_{\mathtt{n}}$ is super-exponentially small, i.e. satisfies (9.1.33),

$$|\rho_{\mathtt{n}}|_{+,s_1} = O\big((\varepsilon^3)^{-(\frac{3}{2})^{\mathtt{n}-1}}\big) \,. \tag{1.5.16}$$

The proof of the splitting Corollary 9.1.2 is based on an iterative application of the 'splitting step" Proposition 9.1.1 which block diagonalizes a Hamiltonian operator of the form

$$\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho) \quad \text{where} \quad \begin{gathered} \rho(\varphi) = \begin{pmatrix} \rho_1(\varphi) & \rho_2(\varphi)^* \\ \rho_2(\varphi) & 0 \end{pmatrix} \in \mathcal{L}(H_{\mathbb{S}}^\perp) \,, \\ \rho_1(\varphi) = \rho_1^*(\varphi) \in \mathcal{L}(H_{\mathbb{F}}) \,, \ \rho_2(\varphi) \in \mathcal{L}(H_{\mathbb{F}}, H_{\mathbb{G}}) \,, \end{gathered}$$

into a new Hamiltonian operator $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0^+ + \rho^+)$ where $A_0^+$ is block-diagonal with respect to $H_{\mathbb{F}} \oplus H_{\mathbb{G}}$ and the coupling remainder $\rho^+$ is much smaller than the previous one, i.e. $\rho^+ = O(\rho^{3/2})$ (compare the size of $\rho$ in (9.1.2) and that of $\rho^+$ in (9.1.11)). The iteration is based on a super-convergent Nash-Moser scheme, to compensate the loss of derivatives due to the small divisors.

In this decoupling procedure (Proposition 9.1.1) a central role is played by the possibility of solving approximately the homological equations, see (9.2.5)-(9.2.6),

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi d + D_0 d + JdJD_0 = J\rho_1 \,, \tag{1.5.17}$$
$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi a - JV_0 Ja + JaJD_0 = J\rho_2 \,. \tag{1.5.18}$$

where $d(\varphi) \in \mathcal{L}(H_{\mathbb{F}})$ is self-adjoint, and $a(\varphi) \in \mathcal{L}(H_{\mathbb{F}}, H_{\mathbb{G}})$, for all $\varphi \in \mathbb{T}^{|\mathbb{S}|}$.

For solving the homological equation (1.5.17) (Lemma 9.3.2), we need the second order Melnikov non-resonance conditions (9.3.6), which concern only the finitely many normal modes in $\mathbb{F}$. Then we use the non-degeneracy properties (8.1.5)-(8.1.7) to prove that they are fulfilled for most values of $\lambda$'s (see the measure estimate of Lemma 9.3.3).

On the other hand we use the monotonicity property (1.5.13) to solve (approximately) the homological equation (1.5.18) for most values of $\lambda$'s. This is a difficult step where we use the multiscale techniques of Chapter 4. Let us try to indicate some key aspects of our approach. We have to solve, approximately, each equation

$$T_j(a^j) = J\rho_2^j \,, \quad \forall j \in \mathbb{F} \,,$$

where

$$a^j(\varphi) := (a(\varphi))_{|H_j} \,, \quad \rho_2^j(\varphi) := (\rho_2(\varphi))_{|H_j} \in \mathcal{L}(H_j, H_{\mathbb{G}}) \,,$$
$$\forall \varphi \in \mathbb{T}^{|\mathbb{S}|} \qquad H_j := \big\{ (q_j, p_j)\Psi_j(x) \,, \, (q_j, p_j) \in \mathbb{R}^2 \big\} \,,$$

and $T_j$ is the linear operator (see (9.3.33))

$$T_j(a^j) := J\bar{\omega}_\varepsilon \cdot \partial_\varphi a^j - JV_0 Ja^j + \mu_j(\varepsilon, \lambda) Ja^j J \,. \tag{1.5.19}$$

We solve approximately (1.5.19) by applying the multiscale Proposition 4.1.5 to the extended operator (see (9.3.39)),

$$(1+\varepsilon^2\lambda)T_j^\sharp \,, \quad T_j^\sharp := \begin{pmatrix} J\bar{\omega}_\varepsilon \cdot \partial_\varphi + \dfrac{D_V}{1+\varepsilon^2\lambda} + \dfrac{\mathfrak{c}}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}} & 0 \\ 0 & T_j \end{pmatrix} \,, \tag{1.5.20}$$

which acts on $\varphi$-dependent functions with values in $\mathcal{L}(H_j, H) = \mathcal{L}(H_j, H_{\mathbb{S}\cup\mathbb{F}}) \oplus \mathcal{L}(H_j, H_{\mathbb{G}})$, see (9.3.36). The operator $\Pi_{\mathbb{S}}$ is the $L^2$ projector on the subspace $H_{\mathbb{S}}$ in (1.5.2) and $\mathfrak{c}$ is a positive constant. Identifying (see (3.2.6))

$$\mathcal{L}(H_j, H) \sim H \times H \sim (L^2(\mathbb{T}^d, \mathbb{R}))^4$$

the operator $T_j^\sharp$ can be regarded to act on (dense subspaces of) the whole $(L^2(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d,\mathbb{R}))^4$ and not only in subspaces of $\varphi$-dependent functions with values in the normal subspace $H_{\mathbb{S}}^\perp \times H_{\mathbb{S}}^\perp$. The self-adjoint operator $T_j^\sharp$ has the form (9.3.37), thus as in Definition 4.1.2-(ii). We can apply the multiscale Proposition 4.1.5 to $T_j^\sharp$. The monotonicity property (1.5.13) implies, in Lemma 9.3.8, the sign condition $\mathfrak{d}_\lambda T_j^\sharp \leq -c\varepsilon^2$ required in Definition 4.1.2-item 3.

Notice that, by (1.5.20), the subspaces $H_{\mathbb{S}}$ and $H_{\mathbb{S}}^\perp$ are *invariant* for the action of the extended operator $T_j^\sharp$ and therefore estimates for the approximate inverse of $T_j^\sharp$ (obtained by the multiscale analysis with the exponential basis) provide also estimates for the approximate inverse of $T_j$ and thus for the approximate solution of (1.5.19).

We now explain the relevance of the decomposition $\mathbb{S}^c = \mathbb{F}\cup\mathbb{G}$ of the normal sites introduced in (1.2.14)-(1.2.15), and why we are able to obtain, in the splitting Corollary 9.1.2, a block-diagonal operator as in (1.5.14)-(1.5.15), with an error like (1.5.16), with respect to the splitting (1.5.10). The decomposition $\mathbb{S}^c = \mathbb{F}\cup\mathbb{G}$ of the normal sites is important for the proof of the 'splitting step" Proposition 9.1.1.

As we prove in section 5.2, under the effect of the nonlinearity $a(x)u^3 + O(u^4)$, the tangential frequency vector of the expected quasi-periodic solutions of (1.5.4) is, up to terms of $O(\varepsilon^4)$,

$$\bar{\mu} + \varepsilon^2 \mathscr{A}\xi \tag{1.5.21}$$

where $\mathscr{A}$ is the twist matrix defined in (1.2.9). Moreover, as proved in Lemma 7.3.2, the perturbed normal frequencies of the Hamiltonian linear operator (1.5.9), with normal indices in a large finite set $\mathbb{M}\subset\mathbb{S}^c$, admit the expansion, up to $O(\varepsilon^4)$,

$$\mu_j + \varepsilon^2(\mathscr{B}\xi)_j\,, \quad \forall j\in\mathbb{M}\,, \tag{1.5.22}$$

where $\mathscr{B}$ is the Birkhoff matrix defined in (1.2.9). Thus the Hamiltonian linear operator (1.5.9), restricted to $H_{\mathbb{M}}$, is in diagonal form, up to terms $O(\varepsilon^4)$. The set $\mathbb{M}$ contains $\mathbb{F}$ and it is fixed in Lemma 7.1.1 large enough so that the sign condition (7.1.9) holds. Notice that, in order to get the expansion (1.5.22) for all the indices $j\in\mathbb{M}$, and not just in $\mathbb{F}$, we have assumed the further second order Melnikov conditions (1.2.18)-(1.2.19), and performed the averaging Proposition 7.3.1. Expressing $\xi$ in terms of $\omega$ by inverting the relation

$$\omega = \bar{\mu} + \varepsilon^2\mathscr{A}\xi\,, \quad \omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon\,,$$

(the twist matrix $\mathscr{A}$ is invertible by (1.2.12)), the shifted normal frequencies (1.5.22) become

$$\begin{aligned}&\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar{\mu}]_j + [\mathscr{B}\mathscr{A}^{-1}\omega]_j = \\ &\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar{\mu}]_j + (1+\varepsilon^2\lambda)[\mathscr{B}\mathscr{A}^{-1}\bar{\omega}_\varepsilon]_j =: \Omega_j(\varepsilon,\lambda)\,.\end{aligned} \tag{1.5.23}$$

Then we divide (1.5.23) by $1+\varepsilon^2\lambda$ and we consider the derivative

$$\frac{d}{d\lambda}\frac{\Omega_j(\varepsilon,\lambda)}{1+\varepsilon^2\lambda} = \frac{-\varepsilon^2}{(1+\varepsilon^2\lambda)^2}(\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar{\mu}]_j)\,, \quad \forall j\in\mathbb{M}\,. \tag{1.5.24}$$

In order to decouple the Hamiltonian linear operator (1.5.9) with respect to $H_{\mathbb{F}}$ and $H_{\mathbb{G}}$, the first naive idea suggests to impose second order Melnikov non-resonance conditions like

$$|\omega \cdot \ell + \Omega_j(\varepsilon, \lambda) \pm \Omega_k(\varepsilon, \lambda)| \geq \frac{\gamma}{\langle \ell \rangle^\tau}, \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|}, \ j \in \mathbb{G}, \ k \in \mathbb{F}, \tag{1.5.25}$$

or equivalently,

$$|f_{\ell,j,k}(\lambda)| \geq \frac{\gamma}{\langle \ell \rangle^\tau} \quad \text{where} \quad f_{\ell,j,k}(\lambda) := \bar{\omega}_\varepsilon \cdot \ell + \frac{\Omega_j(\varepsilon, \lambda)}{1+\varepsilon^2\lambda} \pm \frac{\Omega_k(\varepsilon, \lambda)}{1+\varepsilon^2\lambda}. \tag{1.5.26}$$

Thanks to (1.5.24), for $j \in \mathbb{G}$, $k \in \mathbb{F}$, by the definition of $\mathbb{F}$ and $\mathbb{G}$ in (1.2.15),

$$\begin{aligned} \partial_\lambda f_{\ell,j,k}(\lambda) &= \frac{-\varepsilon^2}{(1+\varepsilon^2\lambda)^2}(\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar{\mu}]_j) \mp \frac{\varepsilon^2}{(1+\varepsilon^2\lambda)^2}(\mu_k - [\mathscr{B}\mathscr{A}^{-1}\bar{\mu}]_k) \\ &\leq -\frac{\varepsilon^2}{1+\varepsilon^2\lambda}\mathfrak{g} \end{aligned} \tag{1.5.27}$$

is negative for all $\lambda$, allowing to prove that (1.5.26) is fulfilled for most values of $\lambda$. This explains the relevance of the splitting (1.2.14)-(1.2.15) on the normal indices $\mathbb{S}^c$.

Actually $\mathbb{G}$ is an infinite set and we can not impose (1.5.25) for all the $j \in \mathbb{G}$, since we have only the expansion (1.5.22) for finitely many $j \in \mathbb{M}$. The homological equation to be solved, in order to decouple $H_{\mathbb{F}}$ and $H_{\mathbb{G}}$, is indeed (1.5.18), i.e. (1.5.19).

On the other hand we take $\mathbb{M}$ large enough as in Lemma 7.1.1 in order to prove that the operator $\mathtt{A}_0$ in (1.5.9) is split-admissible, see Lemma 8.1.2, and so the operator $A_{\mathtt{n}}$ in (1.5.14)-(1.5.15). Indeed, for $\mathbb{M}$ large enough, the infinite dimensional operator

$$\partial_\lambda {\mathtt{A}_0}_{\mathbb{M}^c}^{\mathbb{M}^c} = \partial_\lambda \mathtt{A}_{\mathbb{M}^c}^{\mathbb{M}^c}, \quad \text{where} \quad \mathtt{A} := \frac{D_V + \varepsilon^2 \mathtt{B}(\varphi)}{1+\varepsilon^2\lambda} \quad (\text{see } (1.5.8)),$$

is strongly negative definite (see (7.1.9)) and, jointly with (1.5.24), this allows to prove the sign conditions (1.5.13). As already mentioned, this property allows to use monotonicity arguments for families of self-adjoint matrices to verify that the non-resonance conditions required to solve the homological equation (1.5.19) are fulfilled for most values of the parameter $\lambda$.

Once the linear operator (1.5.9) has been approximately block-diagonalized, according to the decomposition $H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, obtaining, as in (1.5.14), (1.5.15), the quasi-periodic Hamiltonian operator

$$\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_{\mathtt{n}} + \rho_{\mathtt{n}}), \tag{1.5.28}$$

where $\rho_{\mathtt{n}}$ is a very small coupling term according to (1.5.16), we prove in Chapter 10 that it admits an approximate inverse, for most values of $\lambda$'s, applying once more the multiscale

Proposition 4.1.5. More precisely we first find in section 10.2 an approximate inverse of the block-diagonal operator

$$\bar\omega_\varepsilon \cdot \partial_\varphi - JA_{\mathtt n}\,, \quad A_{\mathtt n} = \begin{pmatrix} D_{\mathtt n}(\varepsilon,\lambda) & 0 \\ 0 & V_{\mathtt n}(\varepsilon,\lambda,\varphi) \end{pmatrix}\,,$$

applying the multiscale Proposition 4.1.5 to an extension of the self-adjoint operator

$$J\bar\omega_\varepsilon \cdot \partial_\varphi + V_{\mathtt n}(\varepsilon,\lambda,\varphi)$$

acting on a dense subspace of the whole $H$, and not just $H_{\mathbb S}^\perp$, see (10.2.26)-(10.2.27). The extended operator has the form $\mathcal L_r$ in (1.4.27), see Definition 4.1.2-(i), and satisfies the sign condition

$$\mathfrak{d}_\lambda(\mathcal L_r(1+\varepsilon^2\lambda)^{-1}) \leq -c\varepsilon^2\,.$$

This allows to verify lower bounds for the moduli of the eigenvalues of finite dimensional restrictions of $\mathcal L_r$, for most values of $\lambda$'s. These lower bounds amount to first order Melnikov type non-resonance conditions.

Finally, once we have constructed an approximate inverse of the block-diagonal operator $\bar\omega_\varepsilon \cdot \partial_\varphi - JA_{\mathtt n}$, we obtain an approximate inverse of $\bar\omega_\varepsilon \cdot \partial_\varphi - J(A_{\mathtt n} + \rho_{\mathtt n})$ in (1.5.28), taking into account the small residual coupling term $\rho_{\mathtt n}$, which satisfies (1.5.16), by a Neumann perturbative argument, see section 10.3.

In conclusion, after all this analysis, going back to the original coordinates, we finally prove, in Proposition 11.1.1, the existence of an approximate right inverse of the quasi-periodic Hamiltonian operator $\mathcal L_\omega(i_n)$ in (1.5.8), and thus of $d_i\mathcal F(\lambda; i_n)$, for most values of $\lambda$'s, satisfying tame estimates. This enables to implement a Nash-Moser iterative scheme (Chapter 11) which proves Theorem 5.1.2 and therefore Theorem 1.2.1.

**Sobolev regularity thresholds.** Along the Monograph we shall use four Sobolev indices

$$s_0 \ll s_1 \ll s_2 \ll s_3\,.$$

The first index

$$s_0 > (|\mathbb S| + d)/2$$

is fixed in (1.2.32) to have the algebra and interpolation properties for all the Sobolev spaces $\mathcal H^s$, $s \geq s_0$, defined in (1.2.31). We prove these properties in section 3.5. The index

$$s_1 \gg s_0$$

is the one required for the multiscale Proposition 4.1.5 (see hypothesis 1 in Definition 4.1.2) to have the sufficient off-diagonal decay to apply the multiscale step Proposition 4.3.4, see in particular (4.3.6) and assumption $(H1)$. It is the Sobolev threshold which defines the "good" matrices in Definition 4.3.1. Another condition of the type $s_1 \gg s_0$ appears in the proof of Lemma 10.2.6.

Then the Sobolev index

$$s_2 \gg s_1$$

is used in the splitting step Proposition 9.1.1, see (8.2.1). This proposition is based on a Nash-Moser iterative scheme where $s_2$ represents a 'high-norm", see (8.2.2).

The largest index

$$s_3 \gg s_2$$

is finally used for the convergence of the Nash-Moser nonlinear iteration in Chapter 11. Notice that in Theorem 11.2.1 the divergence of the approximate solutions $i_n$ in the high norm $\| \ \|_{\mathrm{Lip},s_3}$ is under control. We require in particular (8.2.1) and, in section 11.2, also stronger largeness conditions for $s_3 - s_2$. Along the iteration we shall verify (impose) that the Sobolev norms of the approximate quasi-periodic solutions remain bounded in $\| \ \|_{\mathrm{Lip},s_1}$ and $\| \ \|_{\mathrm{Lip},s_2}$ norms (so that the assumptions of Propositions 4.1.5, 8.2.1 and 9.1.1 are fulfilled).

## 1.6 Basic Notation

For $s \in \mathbb{R}$ we denote the Sobolev spaces

$$\begin{aligned}\mathcal{H}^s &:= \mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d; \mathbb{C}^r) \\ &:= \Big\{ u(\varphi, x) := \sum_{(\ell,j) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{Z}^d} u_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)} : \|u\|_s^2 := \sum_{i \in \mathbb{Z}^{|\mathbb{S}|+d}} |u_i|^2 \langle i \rangle^{2s}, \\ &\qquad i := (\ell, j), \ \langle i \rangle := \max(|\ell|, |j|, 1), \ |j| := \max\{|j_1|, \ldots, |j_d|\} \Big\}\end{aligned}$$

and we use the same notation $\mathcal{H}^s$ also for the subspace of real valued functions. Moreover we denote by $H^s := H^s_x$ the Sobolev space of functions $u(x)$ in $H^s(\mathbb{T}^d, \mathbb{C})$ and $H^s_\varphi$ the Sobolev space of functions $u(\varphi)$ in $H^s(\mathbb{T}^{|\mathbb{S}|}, \mathbb{C})$. We denote by $b := |\mathbb{S}| + d$.

Let $E$ be a Banach space. Given a continuous map $u : \mathbb{T}^{|\mathbb{S}|} \to E$, $\varphi \mapsto u(\varphi)$, we denote by $\widehat{u}(\ell) \in E$, $\ell \in \mathbb{Z}^{|\mathbb{S}|}$, its Fourier coefficients

$$\widehat{u}(\ell) := \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} u(\varphi) e^{-\mathrm{i}\ell\cdot\varphi} \, d\varphi,$$

and its average

$$\langle u \rangle := \widehat{u}(0) := \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} u(\varphi) \, d\varphi$$

Given an irrational vector $\omega \in \mathbb{R}^{|\mathbb{S}|}$, i.e. $\omega \cdot \ell \neq 0$, $\forall \ell \in \mathbb{Z}^{|\mathbb{S}|} \setminus \{0\}$, and a function $g(\varphi) \in \mathbb{R}^{|\mathbb{S}|}$ with zero average, we define the solution $h(\varphi)$ of $\omega \cdot \partial_\varphi h = g$, with zero average,

$$h(\varphi) = (\omega \cdot \partial_\varphi)^{-1} g := \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|} \setminus \{0\}} \frac{\widehat{g}(\ell)}{\mathrm{i}\omega \cdot \ell} e^{\mathrm{i}\ell\cdot\varphi}. \tag{1.6.1}$$

Let $E$ be a Banach space with norm $\| \ \|_E$. Given a function $f : \Lambda := [-\lambda_0, \lambda_0] \subset \mathbb{R} \to E$ we define its Lipschitz norm

$$\begin{aligned}\|f\|_{\mathrm{Lip}} &:= \|f\|_{\mathrm{Lip},\Lambda} := \|f\|_{\mathrm{Lip},E} := \sup_{\lambda \in \Lambda} \|f\|_E + |f|_{\mathrm{lip}}, \\ |f|_{\mathrm{lip}} &:= |f|_{\mathrm{lip},\Lambda} := |f|_{\mathrm{lip},E} := \sup_{\lambda_1,\lambda_2 \in \Lambda, \lambda_1 \neq \lambda_2} \frac{\|f(\lambda_2) - f(\lambda_1)\|_E}{|\lambda_2 - \lambda_1|}.\end{aligned} \tag{1.6.2}$$

If a function $f : \widetilde{\Lambda} \subset \Lambda \to E$ is defined only on a subset $\widetilde{\Lambda}$ of $\Lambda$ we shall still denote by $\|f\|_{\mathrm{Lip}} := \|f\|_{\mathrm{Lip},\widetilde{\Lambda}} := \|f\|_{\mathrm{Lip},E}$ the norm in (1.6.2) where the sup-norm and the Lipschitz seminorm are intended in $\widetilde{\Lambda}$, without specifying explicitly the domain $\widetilde{\Lambda}$.

If the Banach space $E$ is the Sobolev space $\mathcal{H}^s$ then we denote more simply $\| \ \|_{\mathrm{Lip},\mathcal{H}^s} = \| \ \|_{\mathrm{Lip},s}$. If $E = \mathbb{R}$ then $\| \ \|_{\mathrm{Lip},\mathbb{R}} = \| \ \|_{\mathrm{Lip}}$.

If $A(\lambda)$ is a function, operator, ..., which depends on a parameter $\lambda$, we shall use the following notation for the partial quotient

$$\frac{\Delta A}{\Delta\lambda} := \frac{A(\lambda_2) - A(\lambda_1)}{\lambda_2 - \lambda_1}, \quad \forall \lambda_1 \neq \lambda_2 . \tag{1.6.3}$$

Given a family of functions, or linear self-adjoint operators $A(\lambda)$ on a Hilbert space $H$, defined for all $\lambda \in \widetilde{\Lambda}$, we shall use the notation

$$\mathfrak{d}_\lambda A(\lambda) \geq \beta \mathrm{Id} \qquad \Longleftrightarrow \qquad \frac{\Delta A}{\Delta\lambda} \geq \beta \mathrm{Id}, \quad \forall \lambda_1, \lambda_2 \in \widetilde{\Lambda}, \ \lambda_1 \neq \lambda_2 , \tag{1.6.4}$$

where, for a self-adjoint operator, $A \geq \beta\mathrm{Id}$ means as usual $(Aw, w)_H \geq \beta \|w\|_H^2$, $\forall w \in H$.

Given linear operators $A, B$ we denote their commutator by

$$\mathrm{Ad}_A B := [A, B] := AB - BA . \tag{1.6.5}$$

We define by $D_V := \sqrt{-\Delta + V(x)}$ and $D_m := \sqrt{-\Delta + m}$ for some $m > 0$.

- Given $x \in \mathbb{R}$ we denote by $\lceil x \rceil$ the smallest integer greater or equal to $x$, and by $[x]$ the integer part of $x$, i.e. the greatest integer smaller or equal to $x$;
- Given $L \in \mathbb{N}$, we denote by $[\![0, L]\!]$ the integers in the interval $[0, L]$;
- We use the notation $a \lesssim_s b$ to mean $a \leq C(s) b$ for some positive constant $C(s)$, and $a \sim_s b$ means that $C_1(s) b \leq a \leq C_2(s) b$ for positive constants $C_1(s), C_2(s)$;
- Given functions $a, b : (0, \varepsilon_0) \to \mathbb{R}$ we write

$$a(\varepsilon) \ll b(\varepsilon) \qquad \Longleftrightarrow \qquad \lim_{\varepsilon \to 0} \frac{a(\varepsilon)}{b(\varepsilon)} = 0 . \tag{1.6.6}$$

In the Monograph we denote by $\mathbb{S}$, $\mathbb{F}$, $\mathbb{G}$, $\mathbb{M}$ subsets of the natural numbers $\mathbb{N}$, with

$$\mathbb{N} = \mathbb{S} \cup \mathbb{S}^c , \quad \mathbb{F} \cup \mathbb{G} = \mathbb{S}^c , \quad \mathbb{F} \cap \mathbb{G} = \emptyset , \quad \mathbb{F} \subset \mathbb{M} .$$

We refer to Chapter 3 for the detailed notation of operators, matrices, decay norms, ...

# Chapter 2

# Hamiltonian formulation

In this Chapter we write the nonlinear wave equation (1.2.1) as a first order Hamiltonian system in action-angle and normal variables. These coordinates are convenient for the proof of Theorem 1.2.1. In section 2.3 we provide estimates about the measure of the admissible Diophantine directions $\bar{\omega}_\varepsilon$ of the frequency vector $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$ of the quasi-periodic solutions of Theorem 1.2.1.

## 2.1 Hamiltonian form of NLW

We write the second order nonlinear wave equation (1.2.1) as the first order system

$$\begin{cases} u_t = v \\ v_t = \Delta u - V(x)u - \varepsilon^2 g(\varepsilon, x, u) \end{cases} \tag{2.1.1}$$

which is the Hamiltonian system

$$\partial_t(u,v) = X_H(u,v)\,, \qquad X_H(u,v) := J\nabla_{L^2} H(u,v)\,, \tag{2.1.2}$$

where

$$J := \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}, \tag{2.1.3}$$

is the symplectic matrix, and $H$ is the Hamiltonian

$$H(u,v) := \int_{\mathbb{T}^d} \frac{v^2}{2} + \frac{(\nabla u)^2}{2} + V(x)\frac{u^2}{2} + \varepsilon^2 G(\varepsilon, x, u)\, dx$$

with primitive

$$G(\varepsilon, x, u) := \int_0^u g(\varepsilon, x, s)\, ds\,, \quad (\partial_u G)(\varepsilon, x, u) = g(\varepsilon, x, u)\,. \tag{2.1.4}$$

For the sequel it is convenient to highlight also the fourth order term of the nonlinearity $g(x,u)$ in (1.1.2), i.e. writing

$$g(x,u) = a(x)u^3 + a_4(x)u^4 + g_{\geq 5}(x,u)\,, \quad g_{\geq 5}(x,u) = O(u^5)\,, \tag{2.1.5}$$

so that the rescaled nonlinearity $g(\varepsilon, x, u)$ in (1.2.2) has the expansion

$$\begin{aligned} g(\varepsilon,x,u) := \varepsilon^{-3} g(x,\varepsilon u) = a(x)u^3 + \varepsilon a_4(x)u^4 + \varepsilon^2 \mathfrak{r}(\varepsilon,x,u)\,, \\ \mathfrak{r}(\varepsilon,x,u) := \varepsilon^{-5} g_{\geq 5}(x,\varepsilon u)\,, \end{aligned} \tag{2.1.6}$$

and its primitive in (2.1.4) has the form

$$G(\varepsilon,x,u) = \frac{1}{4}a(x)u^4 + \frac{\varepsilon}{5}a_4(x)u^5 + \varepsilon^2 \mathfrak{R}(\varepsilon,x,u) \tag{2.1.7}$$

where $(\partial_u \mathfrak{R})(\varepsilon,x,u) = \mathfrak{r}(\varepsilon,x,u)$.

The phase space of (2.1.2) is a dense subspace of the real Hilbert space

$$L^2(\mathbb{T}^d) \times L^2(\mathbb{T}^d)\,, \quad L^2(\mathbb{T}^d) := L^2(\mathbb{T}^d, \mathbb{R})\,,$$

endowed with the standard constant symplectic 2-form

$$\Omega((u,v),(u',v')) := (J(u,v),(u',v'))_{L^2 \times L^2} = (v,u')_{L^2} - (u,v')_{L^2}\,. \tag{2.1.8}$$

Notice that the Hamiltonian vector field $X_H$ is characterized by the relation

$$\Omega(X_H, \cdot) = -dH\,.$$

System (2.1.1) is reversible with respect to the involution

$$S(u,v) := (u,-v)\,, \quad S^2 = \mathrm{Id}\,, \tag{2.1.9}$$

namely (see Appendix A.1)

$$X_H \circ S = -S \circ X_H\,, \qquad \text{equivalently} \qquad H \circ S = H\,.$$

Notice that the above equivalence is due to the fact that the involution $S$ is antisymplectic, namely the pull-back

$$S^* \Omega = -\Omega\,.$$

Let us write (2.1.1) in a more symmetric form. Under the symplectic transformation

$$q = D_V^{\frac{1}{2}} u\,, \quad p = D_V^{-\frac{1}{2}} v\,, \tag{2.1.10}$$

where $D_V := \sqrt{-\Delta + V(x)}$ is the linear (unbounded) operator defined on the orthonormal basis $\{\Psi_j(x), j \in \mathbb{N}\}$ of $L^2(\mathbb{T}^d)$ formed by the eigenfunctions of $-\Delta + V(x)$ in (1.1.5), by setting

$$D_V \Psi_j := \sqrt{-\Delta + V(x)}\Psi_j := \mu_j \Psi_j \,, \quad \forall j \in \mathbb{N} \,, \tag{2.1.11}$$

the Hamiltonian system (2.1.1) becomes

$$\begin{cases} q_t = D_V p \\ p_t = -D_V q - \varepsilon^2 D_V^{-\frac{1}{2}} g(\varepsilon, x, D_V^{-\frac{1}{2}} q) \,, \end{cases} \tag{2.1.12}$$

which is the Hamiltonian system

$$\partial_t (q,p) = J \nabla_{L^2} \mathtt{K}(q,p) \tag{2.1.13}$$

with transformed Hamiltonian

$$\mathtt{K}(q,p) := \int_{\mathbb{T}^d} \frac{(D_V^{\frac{1}{2}} p)^2}{2} + \frac{(D_V^{\frac{1}{2}} q)^2}{2} + \varepsilon^2 G(\varepsilon, x, D_V^{-\frac{1}{2}} q)\, dx \,. \tag{2.1.14}$$

This Hamiltonian system is still reversible with respect to $S$ defined in (2.1.9), i.e. $\mathtt{K} \circ S = \mathtt{K}$, i.e. $\mathtt{K}$ is even in $p$.

## 2.2 Action-angle and "normal" variables

We look for quasi-periodic solutions of the Hamiltonian system (2.1.13) which are mainly Fourier supported on the tangential sites $\mathbb{S} \subset \mathbb{N}$. Thus we decompose

$$\begin{aligned} L^2(\mathbb{T}^d, \mathbb{R}) \times L^2(\mathbb{T}^d, \mathbb{R}) &= H_{\mathbb{S}} \oplus H_{\mathbb{S}}^{\perp} \,, \\ H_{\mathbb{S}} &:= \Big\{ (q(x), p(x)) = \sum_{j \in \mathbb{S}} (q_j, p_j) \Psi_j(x) \,, \quad (q_j, p_j) \in \mathbb{R}^2 \Big\} \end{aligned} \tag{2.2.1}$$

splitting the canonical variables $(q,p)$ into *tangential* and *normal* components

$$(q,p) = \sum_{j \in \mathbb{S}} (q_j, p_j) \Psi_j(x) + (Q,P) \tag{2.2.2}$$

where $(q_j, p_j) \in \mathbb{R}^2$ and $(Q,P) \in H_{\mathbb{S}}^{\perp}$. Then we introduce usual action-angle variables on the tangential sites by setting

$$(q_j, p_j) := \sqrt{2 I_j} \big( \cos\theta_j, -\sin\theta_j \big) \,, \quad \forall j \in \mathbb{S} \,. \tag{2.2.3}$$

The symplectic form (2.1.8) then becomes (recall that $\Psi_j(x)$ are $L^2$-orthonormal)

$$\mathcal{W} := (dI \wedge d\theta) \,\oplus\, \Omega \tag{2.2.4}$$

where, for simplicity, we still denote by $\Omega := \Omega_{|H_{\mathbb{S}}^{\perp}}$ the restriction of the 2-form $\Omega$, defined in (2.1.8), to the symplectic subspace $H_{\mathbb{S}}^{\perp}$.

By (2.2.2), (2.2.3), the Hamiltonian in (2.1.14) then becomes, recalling that $\bar{\mu} \in \mathbb{R}^{|\mathbb{S}|}$ is the unperturbed tangential frequency vector defined in (1.2.3),

$$\begin{aligned} K(\theta, I, Q, P) = \bar{\mu} \cdot I + \frac{1}{2} \int_{\mathbb{T}^d} (D_V^{\frac{1}{2}} Q)^2 + (D_V^{\frac{1}{2}} P)^2 \, dx \\ + \varepsilon^2 \int_{\mathbb{T}^d} G\Big(\varepsilon, x, \sum_{j \in \mathbb{S}} \mu_j^{-\frac{1}{2}} \sqrt{2 I_j} \cos\theta_j \Psi_j(x) + D_V^{-\frac{1}{2}} Q\Big) dx \, . \end{aligned} \tag{2.2.5}$$

For $\varepsilon = 0$, the Hamiltonian system generated by (2.2.5) admits the continuous family of quasi-periodic solutions

$$\theta(t) = \theta_0 + \bar{\mu} t \, , \ I(t) = \xi \, , \ Q(t) = P(t) = 0 \, ,$$

parametrized by the unperturbed tangential "actions" $\xi := (\xi_j)_{j \in \mathbb{S}}$, $\xi_j > 0$. The aim is to prove their persistence, being just slightly deformed, for $\varepsilon$ small enough, for "most" values of $\xi$, and with a frequency close to $\bar{\mu}$.

Then we introduce nearby coordinates by the symplectic transformation

$$I = \xi + y \, , \tag{2.2.6}$$

and, substituting in (2.2.5), we are reduced to study the parameter dependent family of Hamiltonians (that for simplicity we denote with the same letter $K$)

$$K(\theta, y, Q, P, \xi) = c + \bar{\mu} \cdot y + \frac{1}{2}(D_V Q, Q)_{L^2} + \frac{1}{2}(D_V P, P)_{L^2} + \varepsilon^2 R(\theta, y, Q, \xi) \tag{2.2.7}$$

where

$$R(\theta, y, Q, \xi) := \int_{\mathbb{T}^d} G\big(\varepsilon, x, v(\theta, y, \xi) + D_V^{-\frac{1}{2}} Q\big) \, dx \tag{2.2.8}$$

and

$$v(\theta, y, \xi) := \sum_{j \in \mathbb{S}} \mu_j^{-\frac{1}{2}} \sqrt{2(\xi_j + y_j)} \cos\theta_j \Psi_j(x) \, . \tag{2.2.9}$$

The phase space of (2.2.7) is now

$$\mathbb{T}^{|\mathbb{S}|} \times \mathbb{R}^{|\mathbb{S}|} \times H_{\mathbb{S}}^{\perp} \ni (\theta, y, z) \, , \quad z := (Q, P) \in H_{\mathbb{S}}^{\perp} \, ,$$

endowed with the symplectic structure (see (2.2.4))

$$\mathcal{W} := (dy \wedge d\theta) \, \oplus \, \Omega \, , \tag{2.2.10}$$

so that the Hamilton equations generated by (2.2.7) have the form

$$\begin{cases} \dot{\theta} = \partial_y K(\theta, y, z, \xi) \\ \dot{y} = -\partial_\theta K(\theta, y, z, \xi) \\ \dot{z} = J\nabla_z K(\theta, y, z, \xi) \end{cases} \tag{2.2.11}$$

or, in expanded form,

$$\begin{cases} \dot{\theta} - \bar{\mu} - \varepsilon^2 \partial_y R(\theta, y, Q, \xi) = 0 \\ \dot{y} + \varepsilon^2 \partial_\theta R(\theta, y, Q, \xi) = 0 \\ \big(\partial_t - JD_V\big)(Q, P) + \varepsilon^2 \big(0, \nabla_Q R(\theta, y, Q, \xi)\big) = 0 \,. \end{cases} \tag{2.2.12}$$

By (2.2.8) and (2.1.4), we have that $\partial_y R := (\partial_{y_m} R)_{m=1,\dots,|\mathbb{S}|} \in \mathbb{R}^{|\mathbb{S}|}$ has components

$$\partial_{y_m} R(\theta, y, Q, \xi) = \int_{\mathbb{T}^d} g(\varepsilon, x, v(\theta, y, \xi) + D_V^{-\frac{1}{2}} Q) \frac{\mu_m^{-\frac{1}{2}}}{\sqrt{2(\xi_m + y_m)}} \cos(\theta_m) \Psi_m(x)\, dx \tag{2.2.13}$$

and

$$\nabla_Q R(\theta, y, Q, \xi) = D_V^{-\frac{1}{2}} g\big(\varepsilon, x, v(\theta, y, \xi) + D_V^{-\frac{1}{2}} Q\big) \,. \tag{2.2.14}$$

We note that the 2-form $\mathcal{W}$ in (2.2.10) is exact, i.e. $\mathcal{W} = d\varkappa$ where $\varkappa$ is the Liouville 1-form

$$\varkappa_{(\theta,y,z)}[\widehat{\theta}, \widehat{y}, \widehat{z}] = \sum_{j=1,\dots,|\mathbb{S}|} y_j \widehat{\theta}_j + \frac{1}{2}(Jz, \widehat{z})_{L^2_x} \,. \tag{2.2.15}$$

The Hamiltonian system (2.2.11) is reversible with respect to the involution

$$\tilde{S}(\theta, y, Q, P) := (-\theta, y, Q, -P) \tag{2.2.16}$$

which is nothing but (2.1.9) in the variables (2.2.2)-(2.2.3). This means that

$$K \circ \tilde{S} = K \,, \quad K(-\theta, y, Q, -P) = K(\theta, y, Q, P) \,,$$

and the Hamiltonian $R$ in (2.2.8) satisfies

$$R(-\theta, y, Q, \xi) = R(\theta, y, Q, \xi) \,. \tag{2.2.17}$$

## 2.3 Admissible Diophantine directions $\bar{\omega}_\varepsilon$

As explained in the introduction, we look for quasi-periodic solutions of (2.2.11)-(2.2.12) with frequency vector

$$\omega = (1 + \varepsilon^2 \lambda) \bar{\omega}_\varepsilon$$

restricted to a fixed line, spanned by $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2 \zeta$ (see (1.2.24)-(1.2.25)), which has to satisfy the Diophantine conditions (1.2.29)-(1.2.30). We now prove that for "most" vector $\zeta \in \mathbb{R}^{|\mathbb{S}|}$ these conditions are satisfied.

**Lemma 2.3.1.** *Assume* (1.2.6) *and* (1.2.8). *Then there exists a subset* $B_\varepsilon \subset \mathscr{A}([1,2]^{|\mathbb{S}|})$ *(where* $\mathscr{A}$ *is the invertible twist matrix in* (1.2.9)*) with measure* $|B_\varepsilon| \leq \varepsilon$*, such that all the vectors*

$$\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2 \zeta \,, \quad \zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|}) \setminus B_\varepsilon \,,$$

*satisfy the Diophantine conditions* (1.2.29)-(1.2.30) *with* $\gamma_1, \tau_1$ *defined in* (1.2.28).

Proof. We first verify that most vectors $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2\zeta$, $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})$, satisfy (1.2.29). Since $\bar{\mu}$ is $(\gamma_0, \tau_0)$-Diophantine, i.e. (1.2.6) holds, then, for all $\ell \in \mathbb{Z}^{|\mathbb{S}|} \setminus \{0\}$,

$$|(\bar{\mu} + \varepsilon^2\zeta) \cdot \ell| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}} - \varepsilon^2 C|\ell| \geq \frac{\gamma_0}{2\langle \ell \rangle^{\tau_0}} \,, \quad \forall |\ell| \leq \Big( \frac{\gamma_0}{2C\varepsilon^2} \Big)^{1/(\tau_0+1)} \,.$$

Then it remains to estimate the measure of

$$\mathcal{B}_\varepsilon := \bigcup\nolimits_{|\ell| > \left( \frac{\gamma_0}{2C\varepsilon^2} \right)^{1/(\tau_0+1)}} \mathcal{R}_\ell \tag{2.3.1}$$

where

$$\mathcal{R}_\ell := \Big\{ \zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|}) \ : \ |(\bar{\mu} + \varepsilon^2\zeta) \cdot \ell| \leq \frac{\gamma_1}{\langle \ell \rangle^{\tau_1}} \,, \ \gamma_1 = \frac{\gamma_0}{2} \Big\} \,.$$

Since the derivative $\dfrac{\ell}{|\ell|} \cdot \partial_\zeta\big((\bar{\mu} + \varepsilon^2\zeta) \cdot \ell\big) = \varepsilon^2|\ell|$, then the measure $|\mathcal{R}_\ell| \lesssim \dfrac{\gamma_0}{\varepsilon^2 \langle \ell \rangle^{\tau_1+1}}$. Therefore

$$|\mathcal{B}_\varepsilon| \lesssim \frac{\gamma_0}{\varepsilon^2} \sum\nolimits_{|\ell| > \left( \frac{\gamma_0}{2C\varepsilon^2} \right)^{1/(\tau_0+1)}} \frac{1}{\langle \ell \rangle^{\tau_1+1}} \leq C(\gamma_0) \varepsilon^{2\frac{(\tau_1+1-|\mathbb{S}|)}{\tau_0+1} - 2} \leq \varepsilon$$

by (1.2.28).

We now consider the quadratic Diophantine condition (1.2.30). Let $M := M_p$ be the $(|\mathbb{S}| \times |\mathbb{S}|)-$*symmetric* matrix such that

$$\sum_{1 \leq i \leq j \leq |\mathbb{S}|} \omega_i \omega_j p_{ij} = M\omega \cdot \omega \,, \quad \forall \omega \in \mathbb{R}^{|\mathbb{S}|} \,.$$

The symmetric matrix $M$ has coefficients

$$M_{ij} := \frac{p_{ij}}{2}(1 + \delta_{ij}) \,, \ \forall 1 \leq i \leq j \leq |\mathbb{S}| \,, \quad \text{and} \quad M_{ij} = M_{ji} \,. \tag{2.3.2}$$

We want to prove that for most $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})$ the vector $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2\zeta$ satisfies the non-resonance condition

$$\big| n + M\bar{\omega}_\varepsilon \cdot \bar{\omega}_\varepsilon \big| \geq \gamma_1 \langle p \rangle^{-\tau_1} \,, \quad \forall (n,p) \in \mathbb{Z} \times \mathbb{Z}^{\frac{|\mathbb{S}|(|\mathbb{S}|+1)}{2}} \setminus \{(0,0)\} \,.$$

Then we write

$$n + M\bar{\omega}_\varepsilon \cdot \bar{\omega}_\varepsilon = n + M\bar{\mu} \cdot \bar{\mu} + 2\varepsilon^2 M\zeta \cdot \bar{\mu} + \varepsilon^4 M\zeta \cdot \zeta \,. \tag{2.3.3}$$

We first note that, by (1.2.8) and $|M_{ij}| \lesssim \langle p \rangle$ by (2.3.2), then, for all $|\zeta| \le 1$,

$$\left| n + M\bar{\mu} \cdot \bar{\mu} + 2\varepsilon^2 M\zeta \cdot \bar{\mu} + \varepsilon^4 M\zeta \cdot \zeta \right| \ge |n + M\bar{\mu} \cdot \bar{\mu}| - \varepsilon^2 C \langle p \rangle \ge \frac{\gamma_0/2}{\langle p \rangle^{\tau_0}}$$

if $\left(\frac{\gamma_0}{2\varepsilon^2 C}\right)^{1/(\tau_0+1)} \ge |p|$. Thus we erase values of $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})$ only when

$$|p| > \left(\frac{\gamma_0}{2\varepsilon^2 C}\right)^{1/(\tau_0+1)} . \tag{2.3.4}$$

Since $M$ is symmetric there is an orthonormal basis $V := (v_1, \dots, v_k)$ of eigenvectors of $M$ with real eigenvalues $\lambda_k := \lambda_k(p)$, i.e. $Mv_k = \lambda_k v_k$. Under the isometric change of variables $\zeta = Vy$ we have to estimate

$$\begin{aligned} |\mathcal{R}_{n,p}| = \Big| \Big\{ & y \in \mathbb{R}^{|\mathbb{S}|} \,,\ |y| \le 1 \ : \\ & \Big| n + M\bar{\mu} \cdot \bar{\mu} + 2\varepsilon^2 Vy \cdot M\bar{\mu} + \varepsilon^4 \sum_{1 \le i \le |\mathbb{S}|} \lambda_k y_k^2 \Big| < \frac{\gamma_0}{2\langle p \rangle^{\tau_1}} \Big\} \Big| \,. \end{aligned} \tag{2.3.5}$$

Since $M^2 v_k = \lambda_k^2 v_k$, $\forall k = 1, \dots, |\mathbb{S}|$, we get

$$\sum_{k=1}^{|\mathbb{S}|} \lambda_k^2 = \mathrm{Tr}(M^2) = \sum_{i,j=1}^{|\mathbb{S}|} M_{ij}^2 \overset{(2.3.2)}{\ge} \frac{|p|^2}{2} \,.$$

Hence there is an index $k_0 \in \{1, \dots, |\mathbb{S}|\}$ such that $|\lambda_{k_0}| \ge |p|/\sqrt{2|\mathbb{S}|}$ and the derivative

$$\begin{aligned} \Big| \partial^2_{y_{k_0}} \Big( n + M\bar{\mu} \cdot \bar{\mu} + 2\varepsilon^2 Vy \cdot M\bar{\mu} + \varepsilon^4 \sum_{1 \le i \le |\mathbb{S}|} \lambda_k y_k^2 \Big) \Big| &= \varepsilon^4 |2\lambda_{k_0}| \\ &\ge \varepsilon^4 \sqrt{2}\, |p| / \sqrt{|\mathbb{S}|} \,. \end{aligned} \tag{2.3.6}$$

As a consequence of (2.3.5) and (2.3.6) we deduce the measure estimate

$$|\mathcal{R}_{n,p}| \lesssim \varepsilon^{-2} \sqrt{\frac{\gamma_0}{\langle p \rangle^{\tau_1+1}}} \,.$$

Recalling (2.3.4), and since $\mathcal{R}_{n,p} = \emptyset$ if $|n| \ge C\langle p \rangle$, we have

$$\begin{aligned} \Big| \bigcup_{n,p \in \mathbb{Z}^{\frac{|\mathbb{S}|(|\mathbb{S}|+1)}{2}} \setminus \{0\}} \mathcal{R}_{n,p} \Big| \lesssim \sum_{|p| > \left(\frac{\gamma_0}{2\varepsilon^2 C}\right)^{1/(\tau_0+1)}} \varepsilon^{-2} \langle p \rangle \sqrt{\frac{\gamma_0}{\langle p \rangle^{\tau_1+1}}} &\lesssim_{\gamma_0} \varepsilon^{\left[\frac{\tau_1 - 1}{\tau_0+1} - \frac{|\mathbb{S}|(|\mathbb{S}|+1)}{\tau_0+1} - 2\right]} \\ &\le \varepsilon \end{aligned}$$

for $\varepsilon$ small, by (1.2.28). ■

**Remark 2.3.2.** *The measure of the set* $|B_\varepsilon| \leq \varepsilon$ *is smaller than* $\varepsilon^p$*, for any* $p$*, at the expense of taking a larger Diophantine exponent* $\tau_1$*. We have written* $|B_\varepsilon| \leq \varepsilon$ *for definiteness.*

We finally notice that, for $\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon$ with a Diophantine vector $\bar{\omega}_\varepsilon$ satisfying (1.2.29), for any zero average function $g(\varphi)$ we have that the function $(\omega \cdot \partial_\varphi)^{-1}g$, defined in (1.6.1), satisfies

$$\|(\omega \cdot \partial_\varphi)^{-1}g\|_{\mathrm{Lip},s} \leq C\gamma_1^{-1}\|g\|_{\mathrm{Lip},s+\tau_1} \,. \tag{2.3.7}$$

# Chapter 3

# Functional setting

In this Chapter we collect all the properties of the phase spaces, linear operators, norms, interpolation inequalities used through the Monograph. Of particular importance for proving Theorem 1.2.1 is the result of section 3.4.

## 3.1 Phase space and basis

The phase space of the nonlinear wave equation (2.1.12) is a dense subspace of the real Hilbert space

$$H := L^2(\mathbb{T}^d) \times L^2(\mathbb{T}^d) \,, \quad L^2(\mathbb{T}^d) := L^2(\mathbb{T}^d, \mathbb{R}) \,. \tag{3.1.1}$$

In the Monograph we shall denote for convenience an element of $H$ either as $h = (h^{(1)}, h^{(2)})$ a row, either as a column $h = \begin{pmatrix} h^{(1)} \\ h^{(2)} \end{pmatrix}$ with components $h^{(l)} \in L^2(\mathbb{T}^d)$, $l = 1, 2$.

In the exponential basis any function of $H$ can be decomposed as

$$(q(x), p(x)) = \sum_{j \in \mathbb{Z}^d} (q_j, p_j) e^{\mathrm{i} j \cdot x} \,, \quad q_{-j} = \bar{q}_j \,, \; p_{-j} = \bar{p}_j \,. \tag{3.1.2}$$

We will also use the orthonormal basis

$$\{\Psi_j(x), j \in \mathbb{N}\}$$

of $L^2(\mathbb{T}^d)$ formed by the eigenfunctions of $-\Delta + V(x)$ defined in (1.1.5) with eigenvalues $\mu_j^2$. We then consider the Hilbert spaces

$$\mathtt{H}^s_x := \Big\{ u := \sum_{j \in \mathbb{N}} u_j \Psi_j \, : \, \|u\|^2_{\mathtt{H}^s_x} := \big( (-\Delta + V(x))^s u, u \big)_{L^2} = \sum_{j \in \mathbb{N}} \mu_j^{2s} |u_j|^2 < \infty \Big\} \,. \tag{3.1.3}$$

Clearly $\mathtt{H}^0_x = L^2$. Actually, for any $s \geq 0$, the "spectral" norm $\|u\|_{\mathtt{H}^s_x}$ is equivalent to the usual Sobolev norm

$$\|u\|_{\mathtt{H}^s_x} \simeq_s \|u\|_{H^s_x} := \big((-\Delta)^s u, u\big)_{L^2} + (u,u)_{L^2}\big)^{\frac{1}{2}} . \tag{3.1.4}$$

For $s \in \mathbb{N}$, the equivalence (3.1.4) can be directly proved noting that $V(x)$ is a lower order perturbation of the Laplacian $-\Delta$. Then, for $s \in \mathbb{R} \setminus \mathbb{N}$, $[s] < s < [s]+1$, the equivalence (3.1.4) follows by the classical interpolation result stating that the Hilbert space $\mathtt{H}^s_x$ in (3.1.3), respectively the Sobolev space $H^s_x$, is the interpolation space between $\mathtt{H}^{[s]}_x$ and $\mathtt{H}^{[s]+1}_x$, respectively between $H^{[s]}_x$ and $H^{[s]+1}_x$, and the equivalence (3.1.4) for integers $s$.

**Tangential and normal subspaces.** Given the finite set $\mathbb{S} \subset \mathbb{N}$, we consider the $L^2$-orthogonal decomposition of the phase space in tangential and normal subspaces as in (2.2.1),

$$H = H_{\mathbb{S}} \oplus H_{\mathbb{S}}^{\perp} ,$$

where

$$H_{\mathbb{S}} = \Big\{ (q(x), p(x)) = \sum_{j \in \mathbb{S}} (q_j, p_j) \Psi_j(x) , \quad (q_j, p_j) \in \mathbb{R}^2 \Big\} .$$

In addition, recalling the disjoint splitting

$$\mathbb{N} = \mathbb{S} \cup \mathbb{F} \cup \mathbb{G}$$

where $\mathbb{F}, \mathbb{G} \subset \mathbb{N}$ are defined in (1.2.14)-(1.2.15), we further decompose the normal subspace $H_{\mathbb{S}}^{\perp}$ as

$$H_{\mathbb{S}}^{\perp} = H_{\mathbb{F}} \oplus H_{\mathbb{G}} \tag{3.1.5}$$

where

$$\begin{aligned} H_{\mathbb{G}} &:= \Big\{ (Q(x), P(x)) \in H \ : \ (Q,P) \perp H_{\mathbb{S}} , \ (Q,P) \perp H_{\mathbb{F}} \Big\} , \\ H_{\mathbb{F}} &:= \Big\{ (Q(x), P(x)) \in H \ : \ (Q,P) \perp H_{\mathbb{S}} , \ (Q,P) \perp H_{\mathbb{G}} \Big\} . \end{aligned} \tag{3.1.6}$$

Thus

$$H = H_{\mathbb{S}} \oplus H_{\mathbb{S}}^{\perp} = H_{\mathbb{S}} \oplus H_{\mathbb{F}} \oplus H_{\mathbb{G}} = H_{\mathbb{S} \cup \mathbb{F}} \oplus H_{\mathbb{G}} . \tag{3.1.7}$$

Accordingly we denote by $\Pi_{\mathbb{S}}$, $\Pi_{\mathbb{F}}$, $\Pi_{\mathbb{G}}$, $\Pi_{\mathbb{S} \cup \mathbb{F}}$, the orthogonal $L^2$-projectors on $H_{\mathbb{S}}$, $H_{\mathbb{F}}$, $H_{\mathbb{G}}$, $H_{\mathbb{S} \cup \mathbb{F}}$. We define

$$\Pi_{\mathbb{S}}^{\perp} := \mathrm{Id} - \Pi_{\mathbb{S}} = \Pi_{\mathbb{N} \setminus \mathbb{S}}$$

and similarly for the other subspaces.

Decomposing the finite dimensional space

$$H_{\mathbb{F}} = \oplus_{j \in \mathbb{F}} H_j , \tag{3.1.8}$$

we denote by $\Pi_j$ the $L^2$-projectors onto $H_j$. In each real subspace $H_j$, $j \in \mathbb{F}$, we take the basis $(\Psi_j(x), 0), (0, \Psi_j(x))$, namely we represent

$$H_j = \Big\{ q(\Psi_j(x), 0) + p(0, \Psi_j(x)) \, , \ q, p \in \mathbb{R} \Big\} . \tag{3.1.9}$$

Thus $H_j$ is isometrically isomorphic to $\mathbb{R}^2$.

**Symplectic operator $J$.** We define the linear *symplectic* operator $J \in \mathcal{L}(H)$ as

$$J \begin{pmatrix} Q(x) \\ P(x) \end{pmatrix} := \begin{pmatrix} P(x) \\ -Q(x) \end{pmatrix} , \quad J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix} . \tag{3.1.10}$$

The symplectic operator $J$ leaves invariant the symplectic subspaces $H_{\mathbb{S}}$, $H_{\mathbb{S}}^\perp$, $H_{\mathbb{F}}$, $H_{\mathbb{G}}$. For simplicity of notation we shall still denote by $J$ the restriction of the symplectic operator

$$J := J_{|H_{\mathbb{S}}} \in \mathcal{L}(H_{\mathbb{S}}) \, , \ J := J_{|H_{\mathbb{S}}^\perp} \in \mathcal{L}(H_{\mathbb{S}}^\perp) \, , \ J := J_{|H_{\mathbb{F}}} \in \mathcal{L}(H_{\mathbb{F}}) \, , \ J := J_{|H_{\mathbb{G}}} \in \mathcal{L}(H_{\mathbb{G}}) \, .$$

The symplectic operator $J$ is represented, in the basis of the exponentials (see (3.1.2))

$$\big\{ (e^{\mathrm{i} j \cdot x}, 0), (0, e^{\mathrm{i} j \cdot x}) \, , j \in \mathbb{Z}^d \big\} \, ,$$

by the matrix

$$J = \mathrm{Diag}_{j \in \mathbb{Z}^d} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} .$$

Note also that the symplectic operator $J$ leaves invariant each subspace $H_j$, $j \in \mathbb{F}$, and it is represented, in the eigenfunction basis (3.1.9), by the same symplectic matrix $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

## 3.2 Linear operators and matrix representation

According to the decomposition $H_{\mathbb{S}}^\perp = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$ in (3.1.5) a linear operator $A$ of $H_{\mathbb{S}}^\perp$ can be represented by a matrix of operators as

$$\begin{aligned} & \begin{pmatrix} A_{\mathbb{F}}^{\mathbb{F}} & A_{\mathbb{F}}^{\mathbb{G}} \\ A_{\mathbb{G}}^{\mathbb{F}} & A_{\mathbb{G}}^{\mathbb{G}} \end{pmatrix} , \\ & A_{\mathbb{F}}^{\mathbb{F}} := \Pi_{\mathbb{F}} A_{|H_{\mathbb{F}}} \, , \ A_{\mathbb{F}}^{\mathbb{G}} := \Pi_{\mathbb{F}} A_{|H_{\mathbb{G}}} \, , \ A_{\mathbb{G}}^{\mathbb{F}} := \Pi_{\mathbb{G}} A_{|H_{\mathbb{F}}} \, , \ A_{\mathbb{G}}^{\mathbb{G}} := \Pi_{\mathbb{G}} A_{|H_{\mathbb{G}}} \, . \end{aligned} \tag{3.2.1}$$

Moreover the decomposition (3.1.8) induces the splitting

$$\mathcal{L}(H_{\mathbb{F}}, H) = \oplus_{j \in \mathbb{F}} \mathcal{L}(H_j, H) \, , \tag{3.2.2}$$

namely a linear operator $A \in \mathcal{L}(H_{\mathbb{F}}, H)$ can be written as

$$A = (a^j)_{j \in \mathbb{F}} \, , \quad a^j := A_{|H_j} \in \mathcal{L}(H_j, H) \, . \tag{3.2.3}$$

In each space $\mathcal{L}(H_j, H)$ we define the scalar product

$$\langle a, b\rangle_0 := \mathrm{Tr}(b^* a)\,, \quad a, b \in \mathcal{L}(H_j, H)\,, \tag{3.2.4}$$

where $b^* \in \mathcal{L}(H, H_j)$ denotes the adjoint of $b$ with respect to the scalar product in $H$. Note that $b^* a \in \mathcal{L}(H_j, H_j)$ is represented, in the basis (3.1.9), by the $2 \times 2$ real matrix whose elements are $L^2$ scalar products

$$\begin{pmatrix} (b(\Psi_j, 0), a(\Psi_j, 0))_H & (b(\Psi_j, 0), a(0, \Psi_j))_H \\ (b(0, \Psi_j), a(\Psi_j, 0))_H & (b(0, \Psi_j), a(0, \Psi_j))_H \end{pmatrix}. \tag{3.2.5}$$

Using the basis (3.1.9), the space of linear operators $\mathcal{L}(H_j, H)$ can be identified with $H \times H$,

$$\mathcal{L}(H_j, H) \simeq H \times H = (L^2(\mathbb{T}^d, \mathbb{R}))^4\,, \tag{3.2.6}$$

identifying $a \in \mathcal{L}(H_j, H)$ with the vector

$$\begin{gathered} (a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}) \in H \times H \\ a(\Psi_j, 0) =: (a^{(1)}, a^{(2)}) \in H\,, \quad a(0, \Psi_j) =: (a^{(3)}, a^{(4)}) \in H\,, \end{gathered} \tag{3.2.7}$$

so that

$$a\big(q(\Psi_j, 0) + p(0, \Psi_j)\big) = q(a^{(1)}, a^{(2)}) + p(a^{(3)}, a^{(4)})\,, \quad \forall (q, p) \in \mathbb{R}^2\,. \tag{3.2.8}$$

With this identification and (3.2.5) the scalar product (3.2.4) takes the form

$$\langle a, b\rangle_0 := \mathrm{Tr}(b^* a) = \sum_{l=1}^{4} (a^{(l)}, b^{(l)})_{L^2} \tag{3.2.9}$$

and the induced norm

$$\langle a, a\rangle_0 = \mathrm{Tr}(a^* a) = \sum_{l=1}^{4} \|a^{(l)}\|^2_{L^2(\mathbb{T}^d)}\,. \tag{3.2.10}$$

By the identification (3.2.7) and taking in $H$ the exponential basis (3.1.2), a linear operator $a \in \mathcal{L}(H_j, H)$ can be also identified with the sequence of $2 \times 2$ matrices $(a_k^j)_{k \in \mathbb{Z}^d}$,

$$a_k^j := \begin{pmatrix} \widehat{a}_k^{(1)} & \widehat{a}_k^{(3)} \\ \widehat{a}_k^{(2)} & \widehat{a}_k^{(4)} \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})\,, \quad a^{(l)} = \sum_{k \in \mathbb{Z}^d} \widehat{a}_k^{(l)} e^{\mathrm{i} k \cdot x}\,, \quad l = 1, 2, 3, 4\,,$$

so that, by (3.2.8),

$$a\big(q(\Psi_j, 0) + p(0, \Psi_j)\big) = \sum_{k \in \mathbb{Z}^d} a_k^j (q, p) e^{\mathrm{i} k x}\,, \quad \forall (q, p) \in \mathbb{R}^2\,.$$

Similarly, using the basis $\{\Psi_j(x)\}_{j\in\mathbb{N}}$ a linear operator $a \in \mathcal{L}(H_j, H)$ can be also identified with the sequence of $2\times 2$ matrices $(\mathtt{a}_k^j)_{k\in\mathbb{N}}$,

$$\mathtt{a}_k^j := \begin{pmatrix} \mathtt{a}_k^{(1)} & \mathtt{a}_k^{(3)} \\ \mathtt{a}_k^{(2)} & \mathtt{a}_k^{(4)} \end{pmatrix}, \quad \mathtt{a}_k^{(l)} := (\Psi_k, a^{(l)})_{L^2}, \quad l = 1, 2, 3, 4, \tag{3.2.11}$$

so that, by (3.2.8),

$$a\big(q(\Psi_j, 0) + p(0, \Psi_j)\big) = \sum_{k\in\mathbb{N}} \mathtt{a}_k^j(q, p)\Psi_k(x), \quad \forall (q, p) \in \mathbb{R}^2. \tag{3.2.12}$$

In addition, if $a \in \mathcal{L}(H_j, H)$ is identified with the vector $(a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}) \in H \times H$ as in (3.2.7), then, $Ja \in \mathcal{L}(H_j, H)$ where $J$ is the symplectic operator in (3.1.10), can be identified with

$$Ja = (a^{(2)}, -a^{(1)}, a^{(4)}, -a^{(3)}), \tag{3.2.13}$$

and $aJ \in \mathcal{L}(H_j, H)$ with

$$aJ = (a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)})J = (-a^{(3)}, -a^{(4)}, a^{(1)}, a^{(2)}). \tag{3.2.14}$$

Each space $\mathcal{L}(H_j, H)$ admits the orthogonal decomposition

$$\mathcal{L}(H_j, H) = \mathcal{L}(H_j, H_{\mathbb{S}\cup\mathbb{F}}) \oplus \mathcal{L}(H_j, H_{\mathbb{G}}), \tag{3.2.15}$$

defined, for any $a^j \in \mathcal{L}(H_j, H)$, by

$$a^j = \Pi_{\mathbb{S}\cup\mathbb{F}} a^j + \Pi_{\mathbb{G}} a^j, \quad \Pi_{\mathbb{S}\cup\mathbb{F}} a^j \in \mathcal{L}(H_j, H_{\mathbb{S}\cup\mathbb{F}}),\ \Pi_{\mathbb{G}} a^j \in \mathcal{L}(H_j, H_{\mathbb{G}}) \tag{3.2.16}$$

where $\Pi_{\mathbb{S}\cup\mathbb{F}}$ and $\Pi_{\mathbb{G}}$ are the $L^2$-orthogonal projectors respectively on the subspaces $H_{\mathbb{S}\cup\mathbb{F}} = H_{\mathbb{F}} \oplus H_{\mathbb{S}}$ and $H_{\mathbb{G}}$, see (3.1.5).

A (possibly unbounded) linear operator $A$ acting on the Hilbert space

$$(i)\ \mathbf{H} := H = (L^2(\mathbb{T}^d))^2, \qquad (ii)\ \mathbf{H} := H \times H = (L^2(\mathbb{T}^d))^4 \tag{3.2.17}$$

can be represented in the Fourier basis of $L^2(\mathbb{T}^d)$ by a matrix $(A_{j'}^j)_{j,j'\in\mathbb{Z}^d}$ with $A_{j'}^j \in \mathrm{Mat}_{2\times 2}(\mathbb{C})$ in case $(i)$, respectively $\mathrm{Mat}_{4\times 4}(\mathbb{C})$ in case $(ii)$, by the relation

$$A\Big(\sum_{j\in\mathbb{Z}^d} h_j e^{\mathrm{i}j\cdot x}\Big) = \sum_{j'\in\mathbb{Z}^d} \Big(\sum_{j\in\mathbb{Z}^d} A_{j'}^j h_j\Big) e^{\mathrm{i}j'\cdot x}$$

with $h_j \in \mathbb{C}^2$ in case (i), respectively $h_j \in \mathbb{C}^4$ in case (ii).

We decompose the space of $2\times 2$-real matrices

$$\mathrm{Mat}_2(\mathbb{R}) = M_+ \oplus M_- \tag{3.2.18}$$

where $M_+$, respectively $M_-$, is the subspace of the $2 \times 2$-matrices which commute, respectively anti-commute, with the symplectic matrix $J$. A basis of $M_+$ is formed by

$$M_1 := J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad M_2 := \mathrm{Id}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.2.19}$$

and a basis of $M_-$ is formed by

$$M_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad M_4 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.2.20}$$

Notice that the matrices $\{M_1, M_2, M_3, M_4\}$ form also a basis for the $2 \times 2$-complex matrices $\mathrm{Mat}_2(\mathbb{C})$.

We shall denote by $\pi^+$, $\pi^-$ the projectors on $M^+$, respectively $M^-$. We shall use that, for a $2 \times 2$ real symmetric matrix

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \qquad \pi_+(M) = \frac{a+c}{2} \mathrm{Id}_2 = \frac{\mathrm{Tr}(M)}{2} \mathrm{Id}_2 . \tag{3.2.21}$$

**$\varphi$-dependent families of functions and operators.** In this Monograph we often identify a function

$$h \in L^2(\mathbb{T}^{|\mathbb{S}|}, L^2(\mathbb{T}^d)), \quad h : \varphi \mapsto h(\varphi) \in L^2(\mathbb{T}^d),$$

with the function $h(\varphi, \cdot)(x) = h(\varphi, x)$ of time-space, i.e.

$$L^2(\mathbb{T}^{|\mathbb{S}|}, L^2(\mathbb{T}^d)) \equiv L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d).$$

Correspondingly, we regard a $\varphi$-dependent family of (possibly unbounded) operators

$$A : \mathbb{T}^{|\mathbb{S}|} \to \mathcal{L}(H_1, H_2), \quad \varphi \mapsto A(\varphi) \in \mathcal{L}(H_1, H_2),$$

acting between Hilbert spaces $H_1, H_2$, as an operator $A$ which acts on functions $h(\varphi, x) \in L^2(\mathbb{T}^{|\mathbb{S}|}, H_1)$ of space-time. When $H_1 = H_2 = L^2(\mathbb{T}^d)$ we regard $A$ as the linear operator $A : L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d) \to L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d)$ defined by

$$(Ah)(\varphi, x) := (A(\varphi) h(\varphi, \cdot))(x).$$

For simplicity of notation we still denote such operator by $A$.

Given $a, b : \mathbb{T}^{|\mathbb{S}|} \to \mathcal{L}(H_j, H)$ we define the scalar product

$$\langle a, b \rangle_0 := \int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}\big(b^*(\varphi) a(\varphi)\big)\, d\varphi, \tag{3.2.22}$$

that, with a slight abuse of notation, we denote with the same symbol (3.2.4). Using (3.2.7) we may identify $a, b : \mathbb{T}^{|\mathbb{S}|} \to \mathcal{L}(H_j, H)$ with $a, b : \mathbb{T}^{|\mathbb{S}|} \to H \times H$ defined by

$$a(\varphi) = (a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)})(\varphi), \quad b(\varphi) = (b^{(1)}, b^{(2)}, b^{(3)}, b^{(4)})(\varphi),$$

and, by (3.2.22) and (3.2.9), (3.2.10),

$$\langle a, b\rangle_0 = \int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}(a(\varphi)^* b(\varphi))\, d\varphi = \sum_{l=1}^{4} (a^{(l)}, b^{(l)})_{L^2(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d)}\,, \tag{3.2.23}$$

$$\|a\|_0^2 = \langle a, a\rangle_0 = \int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}(a(\varphi)^* a(\varphi))\, d\varphi = \sum_{l=1}^{4} \|a^{(l)}\|^2_{L^2(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d)}\,. \tag{3.2.24}$$

We identify a $\varphi$-dependent family of linear operators $A(\varphi)$ acting on $\varphi$-dependent family of functions with values in the Hilbert space $\mathbf{H}$ defined in (3.2.17), with the infinite dimensional matrix $(A^{\ell,j}_{\ell',j'})_{(\ell,j),(\ell',j')\in\mathbb{Z}^{|\mathbb{S}|+d}}$, with $A^{\ell,j}_{\ell',j'} \in \mathrm{Mat}_{2\times 2}(\mathbb{C})$ in case $(i)$, respectively $\mathrm{Mat}_{4\times 4}(\mathbb{C})$ in case $(ii)$, defined by the relation

$$A\Big(\sum_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|+d}} h_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)}\Big) = \sum_{(\ell',j')\in\mathbb{Z}^{|\mathbb{S}|+d}} \Big(\sum_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|+d}} A^{\ell,j}_{\ell',j'} h_{\ell,j}\Big) e^{\mathrm{i}(\ell'\cdot\varphi + j'\cdot x)} \tag{3.2.25}$$

where $h_{\ell,j} \in \mathbb{C}^2$ in case (3.2.17)-$(i)$, respectively $h_{\ell,j} \in \mathbb{C}^4$ in case (3.2.17)-$(ii)$.

Taking in $(L^2(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d))^2$ the basis $\{e^{\mathrm{i}\ell\cdot\varphi}(\Psi_j(x), 0), e^{\mathrm{i}\ell\cdot\varphi}(0, \Psi_j(x))\}_{j\in\mathbb{N}}$ we identify $A$ with the matrix $(A^{\ell,j}_{\ell',j'})_{(\ell,j),(\ell',j')\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{N}}$ by the relation

$$A\Big(\sum_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{N}} h_{\ell,j} e^{\mathrm{i}\ell\cdot\varphi}\Psi_j(x)\Big) = \sum_{(\ell',j')\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{N}} \Big(\sum_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{N}} A^{\ell,j}_{\ell',j'} h_{\ell,j}\Big) e^{\mathrm{i}\ell'\cdot\varphi}\Psi_{j'}(x) \tag{3.2.26}$$

where $h_{\ell,j} \in \mathbb{C}^2$.

**Normal form.** We consider a $\varphi$-dependent family of linear operators $A(\varphi)$ acting in $H_{\mathbb{S}}^{\perp}$. According to the decomposition $H_{\mathbb{S}}^{\perp} = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, each $A(\varphi)$ can be represented by a matrix as in (3.2.1)

$$A(\varphi) = \begin{pmatrix} [A(\varphi)]_{\mathbb{F}}^{\mathbb{F}} & [A(\varphi)]_{\mathbb{F}}^{\mathbb{G}} \\ [A(\varphi)]_{\mathbb{G}}^{\mathbb{F}} & [A(\varphi)]_{\mathbb{G}}^{\mathbb{G}} \end{pmatrix}. \tag{3.2.27}$$

We denote by $\Pi_{\mathtt{D}} A$ the operator of $H_{\mathbb{S}}^{\perp}$ represented by the matrix

$$\Pi_{\mathtt{D}} A(\varphi) = \begin{pmatrix} D_+(A_{\mathbb{F}}^{\mathbb{F}}) & 0 \\ 0 & [A(\varphi)]_{\mathbb{G}}^{\mathbb{G}} \end{pmatrix} \tag{3.2.28}$$

where, in the basis $\{(\Psi_j, 0), (0, \Psi_j)\}_{j\in\mathbb{F}}$ of $H_{\mathbb{F}}$,

$$D_+(A_{\mathbb{F}}^{\mathbb{F}}) := \mathrm{Diag}_{j\in\mathbb{F}}\big(\pi_+[\widehat{A}_j^j(0)]\big) \tag{3.2.29}$$

and $\pi_+ : \mathrm{Mat}_2(\mathbb{R}) \to M_+$ denotes the projector on $M_+$, see (3.2.18), and $\widehat{A}_j^j(0)$ is the $\varphi$-average

$$\widehat{A}_j^j(0) := \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} A_j^j(\varphi)\, d\varphi\,.$$

We define also

$$\Pi_{\mathtt{0}} A := A - \Pi_{\mathtt{D}} A \,. \tag{3.2.30}$$

We remark that, along the Monograph, the functions and the operators may depend on a one-dimensional parameter $\lambda \in \widetilde{\Lambda} \subset \Lambda$ in a Lipschitz way, with norm defined as in (1.6.2).

**Hamiltonian and symplectic operators.** Along the paper we shall preserve the Hamiltonian structure of the vector fields.

**Definition 3.2.1.** *A $\varphi$-dependent family of linear operators $X(\varphi) : \mathcal{D}(X) \subset H \to H$, defined on a dense subspace $\mathcal{D}(X)$ of $H$ independent of $\varphi \in \mathbb{T}^{|\mathbb{S}|}$, is* HAMILTONIAN *if*

$$X(\varphi) = JA(\varphi)$$

*for some real linear operator $A(\varphi)$ which is self-adjoint with respect to the $L^2$ scalar product. We also say that $\omega \cdot \partial_\varphi - JA(\varphi)$ is Hamiltonian.*

In the Monograph we mean that $A$ is self-adjoint if its domain of definition $\mathcal{D}(A)$ is dense in $H$, and $(Ah, k) = (h, Ak)$, for all $h, k \in \mathcal{D}(A)$.

**Definition 3.2.2.** *A $\varphi$-dependent family of linear operators $\Phi(\varphi) : H \to H$, $\forall \varphi \in \mathbb{T}^{|\mathbb{S}|}$, is* SYMPLECTIC *if*

$$\Omega(\Phi(\varphi)u, \Phi(\varphi)v) = \Omega(u, v) \,, \quad \forall u, v \in H \,, \tag{3.2.31}$$

*where the symplectic 2-form $\Omega$ is defined in* (2.1.8). *Equivalently*

$$\Phi^*(\varphi) J \Phi(\varphi) = J \,, \quad \forall \varphi \in \mathbb{T}^{|\mathbb{S}|} \,.$$

A Hamiltonian operator transforms into an Hamiltonian one under a symplectic transformation.

**Lemma 3.2.3.** *Let $\Phi(\varphi)$, $\varphi \in \mathbb{T}^{|\mathbb{S}|}$, be a family of linear symplectic transformations and $A^*(\varphi) = A(\varphi)$, for all $\varphi \in \mathbb{T}^{|\mathbb{S}|}$. Then*

$$\Phi^{-1}(\varphi)\big(\omega \cdot \partial_\varphi - JA(\varphi)\big)\Phi(\varphi) = \omega \cdot \partial_\varphi - JA_+(\varphi)$$

*where $A_+(\varphi)$ is self-adjoint. Thus $\omega \cdot \partial_\varphi - JA_+(\varphi)$ is Hamiltonian.*

PROOF. We have that

$$\begin{aligned}\Phi^{-1}(\varphi)\big(\omega \cdot \partial_\varphi - JA(\varphi)\big)\Phi(\varphi) &= \omega \cdot \partial_\varphi + \Phi^{-1}(\varphi)(\omega \cdot \partial_\varphi \Phi)(\varphi) - \Phi^{-1}(\varphi) JA(\varphi)\Phi(\varphi) \\ &= \omega \cdot \partial_\varphi - JA_+(\varphi)\end{aligned}$$

with

$$\begin{aligned}A_+(\varphi) &= J\Phi^{-1}(\varphi)\,(\omega \cdot \partial_\varphi \Phi)(\varphi) - J\Phi^{-1}(\varphi) JA(\varphi)\Phi(\varphi) \\ &= J\Phi^{-1}(\varphi)\,(\omega \cdot \partial_\varphi \Phi)(\varphi) + \Phi^*(\varphi) A(\varphi)\Phi(\varphi)\end{aligned} \tag{3.2.32}$$

using that $\Phi(\varphi)$ is symplectic. Since $A(\varphi)$ is self-adjoint, the last operator in (3.2.32) is clearly self-adjoint. In order to prove that also the first operator in (3.2.32) is self-adjoint we notice that, since $\Phi(\varphi)$ is symplectic,

$$\Phi^*(\varphi)J(\omega\cdot\partial_\varphi\Phi)(\varphi)+(\omega\cdot\partial_\varphi\Phi)^*(\varphi)J\Phi(\varphi)=0\,. \tag{3.2.33}$$

Thus, using that $\Phi(\varphi)$ is symplectic,

$$\begin{aligned}\big(J\Phi^{-1}(\varphi)(\omega\cdot\partial_\varphi\Phi)(\varphi)\big)^* &= -(\omega\cdot\partial_\varphi\Phi)^*(\varphi)\,(\Phi^*)^{-1}(\varphi)J\\ &= -(\omega\cdot\partial_\varphi\Phi)^*(\varphi)\,J\Phi(\varphi)\\ &\overset{(3.2.33)}{=} \Phi^*(\varphi)J(\omega\cdot\partial_\varphi\Phi)(\varphi)=J\Phi^{-1}(\varphi)(\omega\cdot\partial_\varphi\Phi)(\varphi)\,.\end{aligned}$$

We have proved that $A^+(\varphi)$ is self-adjoint. ■

## 3.3 Decay norms

Let $b:=|\mathbb{S}|+d$. For $B\subset\mathbb{Z}^b$ we introduce the subspace

$$\mathcal{H}^s_B:=\Big\{u=\sum_{i\in\mathbb{Z}^b}u_ie_i\in\mathcal{H}^s\ :\ u_i\in\mathbb{C}^r\,,\ u_i=0 \text{ if } i\notin B\Big\} \tag{3.3.1}$$

where $e_i:=e^{\mathrm{i}(\ell\cdot\varphi+j\cdot x)}$, $i=(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{Z}^d$, and $\mathcal{H}^s$ is the Sobolev space

$$\mathcal{H}^s:=\mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d;\mathbb{C}^r):=\Big\{u=\sum_{i\in\mathbb{Z}^b}u_ie_i\ :\ \|u\|_s^2:=\sum_{i\in\mathbb{Z}^b}|u_i|^2\langle i\rangle^{2s}\Big\}\,. \tag{3.3.2}$$

Clearly $\mathcal{H}^0=L^2(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d;\mathbb{C}^r)$, and, for $s>b/2$, we have the continuous embedding $\mathcal{H}^s\subset C^0(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d;\mathbb{C}^r)$.

For a Lipschitz family of functions $f:\Lambda\mapsto\mathcal{H}^s$, $\lambda\mapsto f(\lambda)$, we define, as in (1.6.2),

$$\|f\|_{\mathrm{Lip},s}:=\sup_{\lambda\in\Lambda}\|f\|_s+\sup_{\lambda_1,\lambda_2\in\Lambda,\lambda_1\neq\lambda_2}\frac{\|f(\lambda_2)-f(\lambda_1)\|_s}{|\lambda_2-\lambda_1|}\,. \tag{3.3.3}$$

**Remark 3.3.1.** *In Chapter 4 we shall distinguish the components of the vector* $u_i=(u_{i,\mathfrak{a}})_{\mathfrak{a}\in\mathfrak{I}}\in\mathbb{C}^r$ *where* $\mathfrak{I}=\{1,2\}$ *if* $r=2$, *and* $\mathfrak{I}=\{1,2,3,4\}$ *if* $r=4$. *In this case we also write an element of* $\mathcal{H}^s$ *as*

$$u=\sum_{i,\mathfrak{a}\in\mathbb{Z}^b\times\mathfrak{I}}u_{i,\mathfrak{a}}e_{i,\mathfrak{a}}\,,\ u_{i,\mathfrak{a}}\in\mathbb{C}\,,\quad e_{i,\mathfrak{a}}:=e_{\mathfrak{a}}e_i\,,$$

*where* $e_{\mathfrak{a}}:=(0,\dots,\underbrace{1}_{\mathfrak{a}-th},\dots,0)$, $\mathfrak{a}\in\mathfrak{I}$, *denotes the canonical basis of* $\mathbb{C}^r$.

When $B$ is finite, the space $\mathcal{H}^s_B$ does not depend on $s$ and will be denoted $\mathcal{H}_B$. For $B, C \subset \mathbb{Z}^b$ finite, we identify the space $\mathcal{L}^B_C$ of the linear maps $L : \mathcal{H}_B \to \mathcal{H}_C$ with the space of matrices

$$\mathcal{M}^B_C := \Big\{ M = (M^{i'}_i)_{i' \in B, i \in C}\,,\ M^{i'}_i \in \mathrm{Mat}(r \times r; \mathbb{C}) \Big\} \tag{3.3.4}$$

identifying $L$ with the matrix $M$ with entries

$$M^{i'}_i = (M^{i',\mathfrak{a}'}_{i,\mathfrak{a}})_{\mathfrak{a},\mathfrak{a}' \in \mathfrak{I}} \in \mathrm{Mat}(r \times r, \mathbb{C})\,, \quad M^{i',\mathfrak{a}'}_{i,\mathfrak{a}} := (L e_{i',\mathfrak{a}'}, e_{i,\mathfrak{a}})_0\,,$$

where $(\ ,\ )_0 := (2\pi)^{-b}(\ ,\ )_{L^2}$ denotes the normalized $L^2$-scalar product.

Following [23] we shall use $s$-decay norms which quantify the polynomial decay off the diagonal of the matrix entries.

**Definition 3.3.2. ($s$-norm)** *The $s$-norm of a matrix $M \in \mathcal{M}^B_C$ is defined by*

$$|M|^2_s := \sum_{n \in \mathbb{Z}^b} [M(n)]^2 \langle n \rangle^{2s}$$

*where* $\langle n \rangle := \max(|n|, 1)$ *(see* (3.3.2)*),*

$$[M(n)] := \begin{cases} \sup_{i - i' = n} |M^{i'}_i| & \text{if } \ n \in C - B \\ 0 & \text{if } \ n \notin C - B\,, \end{cases}$$

*where* $|\ |$ *denotes a norm of the matrices* $\mathrm{Mat}(r \times r, \mathbb{C})$.

We shall use the above definition also if $B$ or $C$ are not finite (with the difference that $|M|_s$ may be infinite).

The $s$-norm is modeled on matrices which represent the multiplication operator. The (Töplitz) matrix $T$ which represents the multiplication operator

$$M_g : \mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d; \mathbb{C}) \to \mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d; \mathbb{C})\,, \quad h \mapsto gh\,,$$

by a function $g \in \mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d; \mathbb{C})$, $s \geq s_0$, satisfies

$$|T|_s \sim \|g\|_s\,, \quad |T|_{\mathrm{Lip},s} \sim \|g\|_{\mathrm{Lip},s}\,. \tag{3.3.5}$$

The $s$-norm satisfies algebra and interpolation inequalities and control the higher Sobolev norms as in (3.3.7) below: as proved in [23], for all $s \geq s_0 > b/2$

$$|AB|_s \lesssim_s |A|_{s_0} |B|_s + |A|_s |B|_{s_0}\,, \tag{3.3.6}$$

and for any subset $B, C \subset \mathbb{Z}^b$, $\forall M \in \mathcal{M}^B_C$, $w \in \mathcal{H}_B$ we have

$$\|Mw\|_s \lesssim_s |M|_{s_0} \|w\|_s + |M|_s \|w\|_{s_0}\,, \tag{3.3.7}$$

$$\|Mw\|_{\mathrm{Lip},s} \lesssim_s |M|_{\mathrm{Lip},s_0} \|w\|_{\mathrm{Lip},s} + |M|_{\mathrm{Lip},s} \|w\|_{\mathrm{Lip},s_0}\,. \tag{3.3.8}$$

The above inequalities can be easily obtained from the definition of the norms $| \ |_s$ and the functional interpolation inequality

$$\forall u, v \in \mathcal{H}^s\,, \ \|uv\|_s \lesssim_s \|u\|_{s_0}\|v\|_s + \|u\|_s\|v\|_{s_0}\,. \tag{3.3.9}$$

Actually, (3.3.9) can be slightly improved to obtain (see Lemma 3.5.1)

$$\|uv\|_s \leq C_0\|u\|_{s_0}\|v\|_s + C(s)\|u\|_s\|v\|_{s_0} \tag{3.3.10}$$

where only the second constant may depend on $s$ and $C_0$ depends only on $s_0$ (we recall that $s_0$ is fixed once for all). From (3.3.10) can be derived the following slight improvements of (3.3.6) and (3.3.8), which will be used in Chapter 10:

$$|AB|_s \leq C_0|A|_{s_0}|B|_s + C(s)|A|_s|B|_{s_0} \tag{3.3.11}$$

and

$$\|Mw\|_{\mathrm{Lip},s} \leq C_0|M|_{\mathrm{Lip},s_0}\|w\|_{\mathrm{Lip},s} + C(s)|M|_{\mathrm{Lip},s}\|w\|_{\mathrm{Lip},s_0}\,. \tag{3.3.12}$$

We also notice that, denoting by $A^*$ the adjoint matrix of $A$, we have

$$|A|_s = |A^*|_s\,, \quad |A|_{\mathrm{Lip},s} = |A^*|_{\mathrm{Lip},s}\,. \tag{3.3.13}$$

The following lemma is the analogue of the smoothing properties of the projection operators. See Lemma 3.6 of [23], and Lemma B.1.10.

**Lemma 3.3.3. (Smoothing)** *Let $M \in \mathcal{M}_C^B$. Then, $\forall s' \geq s \geq 0$,*

$$M_i^{i'} = 0\,, \ \forall |i - i'| < N \quad \Longrightarrow \quad |M|_s \leq N^{-(s'-s)}|M|_{s'}\,, \tag{3.3.14}$$

*and similarly for the Lipschitz norm $| \ |_{\mathrm{Lip},s}$.*

We now define the decay norm for an operator $A : E \to F$ defined on a closed subspace $E \subset L^2$ with range in a closed subset of $L^2$.

**Definition 3.3.4.** *Let $E, F$ be closed subspaces of $L^2 \equiv L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d, \mathbb{C}^r)$. Given a linear operator $A : E \to F$, we extend it to a linear operator $\widetilde{A} : L^2 \to L^2$ acting on the whole $L^2 = E \oplus E^\perp$, with image in $F$, by defining*

$$\widetilde{A}_{|E^\perp} := 0\,. \tag{3.3.15}$$

*Then, for $s \geq 0$ we define the (possibly infinite) $s$-decay norm*

$$|A|_s := |\widetilde{A}|_s\,, \tag{3.3.16}$$

*and*

$$|A|_{+,s} := |D_m^{\frac{1}{2}}\widetilde{A}D_m^{\frac{1}{2}}|_s \tag{3.3.17}$$

*where $D_m$ is the Fourier multiplier operator*

$$D_m := \sqrt{-\Delta + m}\,, \quad D_m(e^{\mathrm{i}j\cdot x}) := \sqrt{|j|^2 + m}\, e^{\mathrm{i}j\cdot x}\,, \quad j \in \mathbb{Z}^d\,, \tag{3.3.18}$$

*and $m > 0$ is a positive constant.*

*For a Lipschitz family of operators $A(\lambda) : E \to E$, $\lambda \in \Lambda$, we associate the Lipschitz norms $|A|_{\mathrm{Lip},s}$, $|A|_{\mathrm{Lip},+,s}$, accordingly to* (1.6.2).

The norm (3.3.17) is stronger than (3.3.16), actually, since $1 \lesssim_m (|j|^2+m)^{\frac14}(|j'|^2+m)^{\frac14}$, we have

$$|A|_s \lesssim_m |A|_{+,s}\,, \quad |A|_{\mathrm{Lip},s} \lesssim_m |A|_{\mathrm{Lip},+,s}\,. \tag{3.3.19}$$

**Lemma 3.3.5. (Tame estimates for composition)** *Let $A, B : E \to E$ be linear operators acting on a closed subspace $E \subset L^2$. Then, the following tame estimates hold: for all $s \geq s_0 > (|\mathbb{S}| + d)/2$,*

$$|AB|_s \lesssim_s |A|_{s_0}|B|_s + |A|_s|B|_{s_0} \tag{3.3.20}$$

$$|AB|_{\mathrm{Lip},s} \lesssim_s |A|_{\mathrm{Lip},s_0}|B|_{\mathrm{Lip},s} + |A|_{\mathrm{Lip},s}|B|_{\mathrm{Lip},s_0}\,, \tag{3.3.21}$$

*more precisely*

$$|AB|_{\mathrm{Lip},s} \leq C_0|A|_{\mathrm{Lip},s_0}|B|_{\mathrm{Lip},s} + C(s)|A|_{\mathrm{Lip},s}|B|_{\mathrm{Lip},s_0} \tag{3.3.22}$$

*and*

$$|AB|_{+,s} + |BA|_{+,s} \lesssim_s |A|_{s_0+\frac12}|B|_{+,s} + |A|_{s+\frac12}|B|_{+,s_0} \tag{3.3.23}$$

$$|AB|_{\mathrm{Lip},+,s} + |BA|_{\mathrm{Lip},+,s} \lesssim_s |A|_{\mathrm{Lip},s_0+\frac12}|B|_{\mathrm{Lip},+,s} + |A|_{\mathrm{Lip},s+\frac12}|B|_{\mathrm{Lip},+,s_0}\,. \tag{3.3.24}$$

*Notice that in* (3.3.23)*, resp.* (3.3.24)*, the operator $A$ is estimated in $|\ |_{s+\frac12}$ norm, resp. $|\ |_{\mathrm{Lip},s+\frac12}$, and not in $|\ |_{+,s}$, resp. $|\ |_{\mathrm{Lip},+,s}$.*

Proof. Notice first that the operation introduced in (3.3.15) of extension of an operator commutes with the composition: if $A, B : E \to E$ are linear operators acting in $E$, then $\widetilde{AB} = \widetilde{A}\,\widetilde{B}$. Thus (3.3.20)-(3.3.21) and (3.3.22) follow by the interpolation inequalities (3.3.6) and (3.3.11).

In order to prove (3.3.23)-(3.3.24) we first show that, given a linear operator acting on the whole $L^2$,

$$|D_m^{-\frac12}\, A\, D_m^{\frac12}|_s\,,\ |D_m^{\frac12}\, A\, D_m^{-\frac12}|_s \lesssim_s |A|_{s+\frac12}\,. \tag{3.3.25}$$

We prove (3.3.25) for

$$D_m^{\frac12}\, A\, D_m^{-\frac12} = A + [D_m^{\frac12}, A]D_m^{-\frac12}\,.$$

Since $|D_m^{-\frac{1}{2}}|_s \leq C(m)$, $\forall s$, it is sufficient to prove that $|[D_m^{\frac{1}{2}}, A]|_s \lesssim_s |A|_{s+\frac{1}{2}}$. Since, $\forall i, j \in \mathbb{Z}^d$,

$$\begin{aligned} &\big|(|j|^2+m)^{1/4} - (|i|^2+m)^{1/4}\big| = \\ &\frac{\big||j|-|i|\big|(|j|+|i|)}{\big((|j|^2+m)^{1/4} + (|i|^2+m)^{1/4}\big)\big((|j|^2+m)^{\frac{1}{2}} + (|i|^2+m)^{\frac{1}{2}}\big)} \\ &\leq \frac{|j-i|}{\big((|j|^2+m)^{1/4} + (|i|^2+m)^{1/4}\big)} \leq \sqrt{|j-i|}\,, \end{aligned}$$

the matrix elements of the commutator $[D_m^{\frac{1}{2}}, A]_j^i$ satisfy

$$|[D_m^{\frac{1}{2}}, A]_j^i| = |A_j^i|\big|(|j|^2+m)^{1/4} - (|i|^2+m)^{1/4}\big| \leq |A_j^i|\sqrt{|j-i|}$$

and therefore (3.3.25) follows.

We now prove the estimates (3.3.23)-(3.3.24). We consider the extended operator $\widetilde{AB} = \widetilde{A}\widetilde{B}$ and write

$$D_m^{\frac{1}{2}}\widetilde{AB}D_m^{\frac{1}{2}} = \big(D_m^{\frac{1}{2}}\widetilde{A}D_m^{-\frac{1}{2}}\big)\big(D_m^{\frac{1}{2}}\widetilde{B}D_m^{\frac{1}{2}}\big)\,.$$

Hence (3.3.23)-(3.3.24) follow, recalling (3.3.17), by (3.3.6) and (3.3.25). ■

By iterating (3.3.21) we deduce that, there exists $C(s) \geq 1$, non decreasing in $s \geq s_0$, such that

$$|A^k|_{\mathrm{Lip},s} \leq (C(s))^k |A|_{\mathrm{Lip},s_0}^{k-1} |A|_{\mathrm{Lip},s}\,, \quad \forall k \geq 1\,. \tag{3.3.26}$$

**Lemma 3.3.6.** *Let $A : E \to E$ be a linear operator acting on a closed subspace $E \subset L^2$. Then its operatorial norm satisfies*

$$\|A\|_0 \lesssim_{s_0} |A|_{s_0} \lesssim_{s_0} |A|_{+,s_0}\,. \tag{3.3.27}$$

Proof. Let $\widetilde{A}$ be the extended operator in $L^2$ defined in (3.3.15). Then, using Lemma 3.8 in [23] (see Lemma B.1.12), and (3.3.16), we get $\|A\|_0 \leq \|\widetilde{A}\|_0 \lesssim_{s_0} |\widetilde{A}|_{s_0} = |A|_{s_0} \lesssim_{s_0} |A|_{+s_0}$ by (3.3.19). ■

We shall also use the following elementary inequality: given a matrix $A \in \mathcal{M}_C^B$ where $B$ and $C$ are included in $[-N, N]^b$, then

$$|A|_s \lesssim N|D_m^{-\frac{1}{2}}AD_m^{-\frac{1}{2}}|_s\,. \tag{3.3.28}$$

We shall use several times the following simple lemma.

**Lemma 3.3.7.** *Let $g(\lambda,\varphi,x), \chi(\lambda,\varphi,x)$ be a Lipschitz family of functions in $\mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d,\mathbb{C})$ for $s\geq s_0$. Then the operator $L$ defined by*

$$L[h](\varphi,x) := \big(h(\varphi,\cdot),g(\varphi,\cdot)\big)_{L^2_x}\chi(\varphi,x)\,, \quad \forall h\in\mathcal{H}^0(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d,\mathbb{C})\,,$$

*satisfies*

$$|L|_{\mathrm{Lip},s} \lesssim_s \|g\|_{\mathrm{Lip},s_0}\|\chi\|_{\mathrm{Lip},s} + \|g\|_{\mathrm{Lip},s}\|\chi\|_{\mathrm{Lip},s_0}\,. \tag{3.3.29}$$

Proof. We write $L = M_\chi P_0 M_{\overline{g}}$ as the composition of the multiplication operators $M_\chi, M_{\overline{g}}$ for the functions $\chi,\overline{g}$ respectively, and the mean value projector $P_0$ defined as

$$P_0 h(\varphi) := \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} h(\varphi,x)\,dx\,, \quad \forall h\in\mathcal{H}^0\,.$$

For $i=(j,\ell)$, $i'=(j',\ell')\in\mathbb{Z}^d\times\mathbb{Z}^{|\mathbb{S}|}$, its entries are $(P_0)_i^{i'} = \delta_j^0\delta_0^{j'}\delta_\ell^{\ell'}$, and therefore

$$|P_0|_{\mathrm{Lip},s} = |P_0|_s \leq 1\,, \quad \forall s\,. \tag{3.3.30}$$

We derive (3.3.29) by (3.3.5), (3.3.30) and the tame estimates (3.3.6) for the composition of operators. ■

Now, given a finite set $\mathbb{M}\subset\mathbb{N}$, we estimate the $s$-decay norm of the $L^2$-orthogonal projector $\Pi_{\mathbb{M}}$ on the subspace of $L^2(\mathbb{T}^d,\mathbb{R})\times L^2(\mathbb{T}^d,\mathbb{R})$ defined by

$$H_{\mathbb{M}} := \Big\{(q(x),p(x)) := \sum_{j\in\mathbb{M}}(q_j,p_j)\Psi_j(x)\,, \quad q_j,p_j\in\mathbb{R}\Big\}\,. \tag{3.3.31}$$

In the next lemma we regard $\Pi_{\mathbb{M}}$ as an operator acting on functions $h(\varphi,x)$.

**Lemma 3.3.8. (Off-diagonal decay of $\Pi_{\mathbb{M}}$)** *Let $\mathbb{M}$ be a finite subset of $\mathbb{N}$. Then, for all $s\geq 0$, there is a constant $C(s):=C(s,\mathbb{M})>0$ such that*

$$|\Pi_{\mathbb{M}}|_{+,s} = |\Pi_{\mathbb{M}}|_{\mathrm{Lip},+,s} \leq C(s)\,. \tag{3.3.32}$$

*In addition, setting $\Pi_{\mathbb{M}}^\perp := \mathrm{Id}-\Pi_{\mathbb{M}} = \Pi_{\mathbb{M}^c}$, we have*

$$|\Pi_{\mathbb{M}}|_{\mathrm{Lip},s} \leq C(s)\,, \qquad |\Pi_{\mathbb{M}^c}|_{\mathrm{Lip},s} = |\Pi_{\mathbb{M}}^\perp|_{\mathrm{Lip},s} \leq C(s)\,. \tag{3.3.33}$$

Proof. To prove (3.3.32) we have to estimate $|\Pi_{\mathbb{M}}^\perp|_{\mathrm{Lip},+,s} = |D_m^{1/2}\Pi_{\mathbb{M}}^\perp D_m^{1/2}|_{\mathrm{Lip},s}$. For any $h=(h^{(1)},h^{(2)})\in\big(L^2(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d)\big)^2$ we have

$$(D_m^{\frac12}\Pi_{\mathbb{M}}^\perp D_m^{\frac12}h)(\varphi,x) = \sum_{j\in\mathbb{M}}\begin{pmatrix}(h^{(1)}(\varphi,\cdot),D_m^{\frac12}\Psi_j)_{L^2_x}\\ (h^{(2)}(\varphi,\cdot),D_m^{\frac12}\Psi_j)_{L^2_x}\end{pmatrix} D_m^{\frac12}\Psi_j\,. \tag{3.3.34}$$

Now, using Lemma 3.3.7, the fact that each $\Psi_j(x)$ is in $C^\infty$ and $\mathbb{M}$ is finite, we deduce, by (3.3.34), the estimate (3.3.32). The first estimate in (3.3.33) is trivial because $|\Pi_{\mathbb{M}}|_{\mathrm{Lip},s}\lesssim|\Pi_{\mathbb{M}}|_{\mathrm{Lip},+,s}$. The second estimate in (3.3.33) follows by $\Pi_{\mathbb{M}}+\Pi_{\mathbb{M}}^\perp=\mathrm{Id}$ and $|\mathrm{Id}|_{\mathrm{Lip},s}=1$. ■

**Lemma 3.3.9.** *Given a $\varphi$-dependent family of linear operators $A(\varphi)$ acting in $H_{\mathbb{S}}^{\perp}$, the operators $\Pi_{\mathtt{D}} A$ and $\Pi_{\mathtt{O}} A$ defined respectively in* (3.2.28) *and* (3.2.30) *satisfy*

$$|\Pi_{\mathtt{D}} A|_{\mathrm{Lip},+,s} + |\Pi_{\mathtt{O}} A|_{\mathrm{Lip},+,s} \le C(s) |A|_{\mathrm{Lip},+,s} \, . \tag{3.3.35}$$

PROOF. We consider the extension of the operator $\Pi_{\mathtt{D}} A$ defined in (3.2.28)-(3.2.29), acting on $\big(L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d)\big)^2$, defined as $A_1 + A_2$ where the operators $A_1$ and $A_2$, are

$$A_1 := \Pi_{\mathbb{G}} A \Pi_{\mathbb{G}} \, , \tag{3.3.36}$$

and, for any $h = (h^{(1)}, h^{(2)}) \in \big(L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d)\big)^2$,

$$(A_2 h)(\varphi, x) := \sum_{j \in \mathbb{F}} \pi_+ \widehat{A}_j^j(0) \begin{pmatrix} (h^{(1)}(\varphi, \cdot), \Psi_j)_{L^2_x} \\ (h^{(2)}(\varphi, \cdot), \Psi_j)_{L^2_x} \end{pmatrix} \Psi_j(x) \, . \tag{3.3.37}$$

We remark that, for $j \in \mathbb{F}$, $\widehat{A}_j^j(0)$ is the $2 \times 2$ matrix

$$\widehat{A}_j^j(0) = \begin{pmatrix} (A(\Psi_j, 0), (\Psi_j, 0))_0 & (A(0, \Psi_j), (\Psi_j, 0))_0 \\ (A(\Psi_j, 0), (0, \Psi_j))_0 & (A(0, \Psi_j), (0, \Psi_j))_0 \end{pmatrix} \tag{3.3.38}$$

where $( \ , \ )_0$ is the (normalized) $L^2$ scalar product in $\big(L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d)\big)^2$.

By Definition 3.3.4, we have $|\Pi_{\mathtt{D}} A|_{\mathrm{Lip},+,s} = |A_1 + A_2|_{\mathrm{Lip},+,s}$. By (3.3.36), (3.3.24) and (3.3.33) (with $\mathbb{M}^c = \mathbb{G}$) we get

$$|A_1|_{\mathrm{Lip},+,s} \lesssim_s |\Pi_{\mathbb{G}}|^2_{\mathrm{Lip}, s+\frac{1}{2}} |A|_{\mathrm{Lip},+,s} \lesssim_s |A|_{\mathrm{Lip},+,s} \, . \tag{3.3.39}$$

For $A_2$, we apply Lemma 3.3.7. Using that, by (3.3.37),

$$(D_m^{\frac{1}{2}} A_2 D_m^{\frac{1}{2}} h)(\varphi, x) = \sum_{j \in \mathbb{F}} \pi_+ \widehat{A}_j^j(0) \begin{pmatrix} (h^{(1)}(\varphi, \cdot), D_m^{\frac{1}{2}} \Psi_j)_{L^2_x} \\ (h^{(2)}(\varphi, \cdot), D_m^{\frac{1}{2}} \Psi_j)_{L^2_x} \end{pmatrix} D_m^{\frac{1}{2}} \Psi_j$$

with $\widehat{A}_j^j(0)$ given in (3.3.38), that $\Psi_j \in C^\infty(\mathbb{T}^d)$ and $\mathbb{F}$ is finite, we deduce by (3.3.29) that

$$|A_2|_{\mathrm{Lip},+,s} = |D_m^{\frac{1}{2}} A_2 D_m^{\frac{1}{2}}|_{\mathrm{Lip},s} \lesssim_s \max_{j \in \mathbb{F}} \|A\|_{\mathrm{Lip},0} \, . \tag{3.3.40}$$

Finally (3.3.39), (3.3.40) imply

$$|\Pi_{\mathtt{D}} A|_{\mathrm{Lip},+,s} \le |A_1|_{\mathrm{Lip},+,s} + |A_2|_{\mathrm{Lip},+,s} \le C(s) |A|_{\mathrm{Lip},+,s}$$

and

$$|\Pi_{\mathtt{O}} A|_{\mathrm{Lip},+,s} = |A - \Pi_{\mathtt{D}} A|_{\mathrm{Lip},+,s} \le |A|_{\mathrm{Lip},+,s} + |\Pi_{\mathtt{D}} A|_{\mathrm{Lip},+,s} \le C(s) |A|_{\mathrm{Lip},+,s} \, .$$

Thus (3.3.35) is proved. ■

We now mention some norm equivalences and estimates that shall be used in the Monograph.

Let us consider a $\varphi$-dependent family of operators $\rho(\varphi) \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$, that, according to the splitting (3.2.1), have the form

$$\rho(\varphi) = \begin{pmatrix} \rho_1(\varphi) & \rho_2(\varphi)^* \\ \rho_2(\varphi) & 0 \end{pmatrix} \in \mathcal{L}(H_{\mathbb{S}}^{\perp}) , \quad \rho_1(\varphi) \in \mathcal{L}(H_{\mathbb{F}}) , \ \rho_2(\varphi) \in \mathcal{L}(H_{\mathbb{F}}, H_{\mathbb{G}}) .$$

Recalling (3.2.3), we can identify each $\rho_l(\varphi)$, $l = 1, 2$, with

$$\rho_l(\varphi) = (\rho_{l,j}(\varphi))_{j \in \mathbb{F}} , \text{ where } \rho_{l,j}(\varphi) := (\rho_l)_{|H_j} \in \begin{cases} \mathcal{L}(H_j, H_{\mathbb{F}}) \text{ if } l = 1 \\ \mathcal{L}(H_j, H_{\mathbb{G}}) \text{ if } l = 2 . \end{cases}$$

Moreover, according to (3.2.7), we can identify each operator $\rho_{l,j}(\varphi)$ with the vector

$$(\rho_{l,j}^{(1)}(\varphi), \rho_{l,j}^{(2)}(\varphi), \rho_{l,j}^{(3)}(\varphi), \rho_{l,j}^{(4)}(\varphi)) \in \begin{cases} H_{\mathbb{F}} \times H_{\mathbb{F}} & \text{if } \ l = 1 \\ H_{\mathbb{G}} \times H_{\mathbb{G}} & \text{if } \ l = 2 , \end{cases} \tag{3.3.41}$$

where

$$\begin{pmatrix} \rho_{l,j}^{(1)} \\ \rho_{l,j}^{(2)} \end{pmatrix} := \rho_l \begin{pmatrix} \Psi_j \\ 0 \end{pmatrix} , \ \begin{pmatrix} \rho_{l,j}^{(3)} \\ \rho_{l,j}^{(4)} \end{pmatrix} := \rho_l \begin{pmatrix} 0 \\ \Psi_j \end{pmatrix} . \tag{3.3.42}$$

We define the Sobolev norms of each $\rho_l(\varphi)$, $l = 1, 2$, as

$$\|\rho_l\|_s^2 = \max_{j \in \mathbb{F}} \Big( \sum_{k=1}^4 \|\rho_{l,j}^{(k)}\|_s^2 \Big) , \tag{3.3.43}$$

and $\|\rho_l\|_{\mathrm{Lip},s}$ according to (1.6.2).

**Lemma 3.3.10.** *We have*

$$|\rho_1|_{\mathrm{Lip},+,s} \sim_s |\rho_1|_{\mathrm{Lip},s} \sim_s \|\rho_1\|_{\mathrm{Lip},s} , \tag{3.3.44}$$

$$|\rho_2|_{\mathrm{Lip},+,s} \lesssim_s |\rho_2|_{\mathrm{Lip},s+\frac{1}{2}} , \ |\rho_2|_{\mathrm{Lip},s} \sim_s \|\rho_2\|_{\mathrm{Lip},s} . \tag{3.3.45}$$

Proof.

**Step1.** We first justify that, for $l = 1, 2$,

$$|\rho_l|_{\mathrm{Lip},s} \sim_s \|\rho_l\|_{\mathrm{Lip},s} . \tag{3.3.46}$$

By Definition 3.3.4, the $s$-norm $|\rho_l|_{\mathrm{Lip},s} = |\rho_l \Pi_{\mathbb{F}}|_{\mathrm{Lip},s}$. Now, recalling (3.3.42), the extended operator $\rho_l \Pi_{\mathbb{F}}$ has the form, for any $h(\varphi) = (h^{(1)}(\varphi), h^{(2)}(\varphi)) \in H$,

$$(\rho_l \Pi_{\mathbb{F}} h)(\varphi) = \sum_{j\in\mathbb{F}} (h^{(1)}(\varphi), \Psi_j)_{L^2_x} \begin{pmatrix} \rho^{(1)}_{l,j}(\varphi) \\ \rho^{(2)}_{l,j}(\varphi) \end{pmatrix} + (h^{(2)}(\varphi), \Psi_j)_{L^2_x} \begin{pmatrix} \rho^{(3)}_{l,j}(\varphi) \\ \rho^{(4)}_{l,j}(\varphi) \end{pmatrix} . \tag{3.3.47}$$

Using that $\Psi_j \in C^\infty(\mathbb{T}^d)$, for all $j \in \mathbb{F}$, and that $\mathbb{F}$ is finite, we derive, from (3.3.47) and Lemma 3.3.7, the estimate $|\rho|_{\mathrm{Lip},s} \lesssim_s \|\rho_l\|_{\mathrm{Lip},s}$.

The reverse inequality $\|\rho_l\|_{\mathrm{Lip},s} \lesssim_s |\rho_l|_{\mathrm{Lip},s}$ follows by (3.3.42), (3.3.8), and that fact that $\Psi_j \in C^\infty(\mathbb{T}^d)$. This concludes the proof of (3.3.46) and so of the second equivalences in (3.3.44)-(3.3.45).

**Step 2.** We now prove that

$$|\rho_1|_{\mathrm{Lip},+,s} \lesssim_s \|\rho_1\|_{\mathrm{Lip},s} . \tag{3.3.48}$$

By Definition 3.3.4, the norm $|\rho_1|_{\mathrm{Lip},+,s} = |D_m^{\frac12} \rho_1 \Pi_{\mathbb{F}} D_m^{\frac12}|_{\mathrm{Lip},s}$ and, by (3.3.47),

$$\begin{aligned} (D_m^{\frac12} \rho_1 \Pi_{\mathbb{F}} D_m^{\frac12} h)(\varphi) &= \sum_{j\in\mathbb{F}} (h^{(1)}(\varphi), D_m^{\frac12}\Psi_j)_{L^2_x} \begin{pmatrix} D_m^{\frac12}\rho^{(1)}_{1,j}(\varphi) \\ D_m^{\frac12}\rho^{(2)}_{1,j}(\varphi) \end{pmatrix} \\ &\quad + \sum_{j\in\mathbb{F}} (h^{(2)}(\varphi), D_m^{\frac12}\Psi_j)_{L^2_x} \begin{pmatrix} D_m^{\frac12}\rho^{(3)}_{1,j}(\varphi) \\ D_m^{\frac12}\rho^{(4)}_{1,j}(\varphi) \end{pmatrix} . \end{aligned} \tag{3.3.49}$$

Now, applying Lemma 3.3.7, we deduce by (3.3.49) and the fact that $\Psi_j$ are $C^\infty$ (and independent on $\lambda$) and $\mathbb{F}$ is finite, that

$$|\rho_1|_{\mathrm{Lip},+,s} = |D_m^{\frac12} \rho_1 \Pi_{\mathbb{F}} D_m^{\frac12}|_{\mathrm{Lip},s} \lesssim_s \max_{j\in\mathbb{F}, k=1,\dots,4} \|D_m^{\frac12} \rho^{(k)}_{1,j}\|_{\mathrm{Lip},s} \tag{3.3.50}$$

Now, by (3.3.41), each $\rho^{(k)}_{1,j}$, $k = 1, \dots, 4$, is a function of the form $u(\varphi) = \sum_{j\in\mathbb{F}} u_j(\varphi)\Psi_j$. We claim that, for functions of this form, $\|D_m^{\frac12} u\|_{\mathrm{Lip},s} \sim_s \|u\|_{\mathrm{Lip},s}$. Indeed

$$\|u\|_{\mathrm{Lip},s} \sim_s \sum_{j\in\mathbb{F}} \|u_j\|_{\mathrm{Lip},H^s(\mathbb{T}^{|\mathbb{S}|})} ,$$

because $\mathbb{F}$ is finite and $\Psi_j$ is $C^\infty$ (and independent on $\lambda$). Similarly,

$$\|D_m^{\frac12} u\|_{\mathrm{Lip},s} = \Big\| \sum_{j\in\mathbb{F}} u_j(\varphi) D_m^{\frac12} \Psi_j \Big\|_{\mathrm{Lip},s} \sim_s \sum_{j\in\mathbb{F}} \|u_j\|_{\mathrm{Lip},H^s(\mathbb{T}^{|\mathbb{S}|})} ,$$

and the claim follows. Applying this property to $u = \rho^{(k)}_{1,j}$ we deduce, by (3.3.50), that

$$|\rho_1|_{\mathrm{Lip},+,s} \lesssim_s \max_{j\in\mathbb{F}, k=1,\dots,4} \|\rho^{(k)}_{1,j}\|_{\mathrm{Lip},s}$$

and thus (3.3.48). Finally, by (3.3.46) and (3.3.19) we have $\|\rho_1\|_{\mathrm{Lip},s} \lesssim_s |\rho_1|_{\mathrm{Lip},+,s}$ and thus we deduce the first equivalence in (3.3.44).

**Step 3.** We finally check

$$|\rho_2|_{\mathrm{Lip},+,s} \lesssim_s |\rho_2|_{\mathrm{Lip},s+\frac{1}{2}}\,, \tag{3.3.51}$$

which is the first estimate in (3.3.44). We have

$$|\rho_2|_{\mathrm{Lip},+,s} = |\rho_2 \Pi_{\mathbb{F}}|_{\mathrm{Lip},+,s} \overset{(3.3.24)}{\lesssim_s} |\rho_2|_{\mathrm{Lip},s+\frac{1}{2}} |\Pi_{\mathbb{F}}|_{\mathrm{Lip},+,s} \overset{(3.3.32)}{\lesssim_s} |\rho_2|_{\mathrm{Lip},s+\frac{1}{2}}$$

(applied with $\mathbb{M} = \mathbb{F}$). ■

We also remind a standard perturbation lemma for operators which admit a right inverse.

**Definition 3.3.11. (Right Inverse)** *A matrix $M \in \mathcal{M}_C^B$ has a right inverse, that we denote by $M^{-1} \in \mathcal{M}_B^C$, if $MM^{-1} = \mathrm{Id}_C$.*

Note that $M$ has a right inverse if and only if $M$ (considered as a linear map) is surjective. The following lemma is proved as in Lemma 3.9 of [23] (see also Lemma B.1.14) by a Neumann series argument.

**Lemma 3.3.12.** *There is a constant $c_0 > 0$ such that, for any $C, B \subset \mathbb{Z}^b$, for any $M \in \mathcal{M}_C^B$ having a right inverse $M^{-1} \in \mathcal{M}_B^C$, for any $P$ in $\mathcal{M}_C^B$ with $|M^{-1}|_{s_0} |P|_{s_0} \leq c_0$ the matrix $M + P$ has a right inverse that satisfies*

$$\begin{aligned} &|(M+P)^{-1}|_{s_0} \leq 2|M^{-1}|_{s_0} \\ &|(M+P)^{-1}|_s \lesssim_s |M^{-1}|_s + |M^{-1}|_{s_0}^2 |P|_s\,, \ \forall s \geq s_0\,. \end{aligned} \tag{3.3.52}$$

Finally, we report the following lemma (related to Lemma 2.1 of [25]) which will be used in Chapter 10.

**Lemma 3.3.13.** *Let $E$ be a closed subspace of $\mathcal{H}^0 = L^2(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d)$ and let $E^s := E \cap \mathcal{H}^s$ for $s \geq s_0$. Let $R : E \to E$ be a linear operator, satisfying, for some $s \geq s_1 \geq s_0$, $\alpha \geq 0$,*

$$\forall v \in E^{s_1}\,, \ \|Rv\|_{s_1} \leq \frac{1}{2}\|v\|_{s_1} \tag{3.3.53}$$

$$\forall v \in E^{s}\,, \ \|Rv\|_{s} \leq \frac{1}{2}\|v\|_{s} + \alpha\|v\|_{s_1}\,. \tag{3.3.54}$$

*Then $\mathrm{Id} + R$ is invertible as an operator of $E^{s_1}$, and*

$$\forall v \in E^{s_1}\,, \ \|(\mathrm{Id}+R)^{-1}v\|_{s_1} \leq 2\|v\|_{s_1} \tag{3.3.55}$$

$$\forall v \in E^{s}\,, \ \|(\mathrm{Id}+R)^{-1}v\|_{s} \leq 2\|v\|_{s} + 4\alpha\|v\|_{s_1}\,. \tag{3.3.56}$$

*Moreover, assume that $R$ depends on the parameter $\lambda \in \widetilde{\Lambda}$ and satisfies also*

$$\forall v(\lambda) \in E^{s_1}\,, \;\; \|Rv\|_{\mathrm{Lip},s_1} \leq \frac{1}{2}\|v\|_{\mathrm{Lip},s_1} \tag{3.3.57}$$

$$\forall v(\lambda) \in E^{s}\,, \;\; \|Rv\|_{\mathrm{Lip},s} \leq \frac{1}{2}\|v\|_{\mathrm{Lip},s} + \alpha\|v\|_{\mathrm{Lip},s_1}\,. \tag{3.3.58}$$

*Then*

$$\forall v(\lambda) \in E^{s_1}\,, \;\; \|(\mathrm{Id}+R)^{-1}v\|_{\mathrm{Lip},s_1} \leq 2\|v\|_{\mathrm{Lip},s_1} \tag{3.3.59}$$

$$\forall v(\lambda) \in E^{s}\,, \;\; \|(\mathrm{Id}+R)^{-1}v\|_{\mathrm{Lip},s} \leq 2\|v\|_{\mathrm{Lip},s} + 4\alpha\|v\|_{\mathrm{Lip},s_1}\,. \tag{3.3.60}$$

PROOF. Since $E$ is a closed subspace of $\mathcal{H}^0$, each space $E^s = E \cap \mathcal{H}^s$, $s \geq 0$, is complete. By (3.3.53) the operator $\mathrm{Id} + R$ is invertible in $E^{s_1}$ and (3.3.55) holds. In order to prove (3.3.56), let $k = (\mathrm{Id}+R)^{-1}v$ so that $k = v - Rk$. Then

$$\begin{aligned} \|k\|_s \leq \|v\|_s + \|Rk\|_s &\overset{(3.3.54)}{\leq} \|v\|_s + \frac{1}{2}\|k\|_s + \alpha\|k\|_{s_1} \\ &\overset{(3.3.55)}{\leq} \|v\|_s + \frac{1}{2}\|k\|_s + \alpha 2\|v\|_{s_1} \end{aligned}$$

and we deduce $\|k\|_s \leq 2(\|v\|_s + \alpha 2\|v\|_{s_1})$ which is (3.3.56). The proof of (3.3.59) and (3.3.60) when we assume (3.3.57) and (3.3.58) follows the same line. ■

## 3.4 Off-diagonal decay of $\sqrt{-\Delta + V(x)}$

The goal of this section is to provide a direct proof of the fact that the matrix which represents the operator

$$D_V = \sqrt{-\Delta + V(x)}\,,$$

defined in (2.1.11), in the exponential basis $\{e^{\mathrm{i}j\cdot x}\}_{j\in\mathbb{Z}^d}$ has off-diagonal decay. This is required by the multiscale analysis performed in Chapter 4. We compare $D_V$ to the Fourier multiplier $D_m = \sqrt{-\Delta + m}$ defined in (3.3.18).

**Proposition 3.4.1. (Off-diagonal decay of $D_V - D_m$)** *There exists a positive constant $\Upsilon_s := \Upsilon_s(\|V\|_{C^{n_s}})$, where $n_s := \lceil s + \frac{d}{2}\rceil + 1 \in \mathbb{N}$, such that*

$$\left|D_V - D_m\right|_{+,s} = \left|D_m^{\frac{1}{2}}(D_V - D_m)D_m^{\frac{1}{2}}\right|_s \leq \Upsilon_s\,. \tag{3.4.1}$$

The rest of this section is devoted to the proof of Proposition 3.4.1. In order to prove that the matrix $(A_i^j)_{i,j\in\mathbb{Z}^d}$ which represents, in the exponential basis $\{e^{\mathrm{i}j\cdot x}\}_{j\in\mathbb{Z}^d}$ a linear operator $A$ acting on a dense subspace of $L^2(\mathbb{T}^d)$, has polynomial off-diagonal decay, we shall use the following criterium. Since the operator $\mathrm{Ad}^n_{\partial_{x_k}} A$, where

$$\mathrm{Ad}_{\partial_{x_k}} := [\partial_{x_k}, \cdot\,]$$

denotes the commutator with the partial derivative $\partial_{x_k}$, is represented by the matrix

$$\big(\mathrm{i}^n (i_k - j_k)^n A_i^j\big)_{i,j\in\mathbb{Z}^d}\,,$$

it is sufficient that $A$ and the operators $\mathrm{Ad}^n_{\partial_{x_k}} A$, $k = 1, \ldots, d$, for $n$ large enough, extend to bounded operators in $L^2(\mathbb{T}^d)$.

We shall use several times the following abstract lemma.

**Lemma 3.4.2.** *Let $(H, \langle\ ,\ \rangle)$ be a separable Hilbert space with norm $\|u\| := \langle u, u\rangle^{\frac12}$. Let $B : \mathcal{D}(B) \subset H \to H$ be an unbounded symmetric operator with a dense domain of definition $\mathcal{D}(B) \subset H$, satisfying:*

*(H1) There is $\beta > 0$ such that $\langle Bu, u\rangle \geq \beta\|u\|^2$, $\forall u \in \mathcal{D}(B)$.*

*(H2) $B$ is invertible and $B^{-1} \in \mathcal{L}(H)$ is a compact operator.*

*Let $A : \mathcal{D}(A) \subset H \to H$ be a symmetric linear operator, such that:*

*(H3) $\mathcal{D}(A)$, respectively in addition $\mathcal{D}(B^{\frac12} A B^{\frac12})$, contains $\mathcal{D}(B^p)$ for some $p \geq 1$.*

*Moreover we assume that:*

*(H4) There is $\rho \geq 0$ such that $|\langle Au, Bu\rangle| \leq \rho\|u\|^2$, $\forall u \in \mathcal{D}(A) \cap \mathcal{D}(B)$.*

*Then $A$, respectively $B^{\frac12} A B^{\frac12}$, can be extended to a bounded operator of $H$ (still denoted by $A$, respectively $B^{\frac12} A B^{\frac12}$) satisfying*

$$\|A\|_{\mathcal{L}(H)} \leq \rho/\beta\,, \quad \text{respectively } \|B^{\frac12} A B^{\frac12}\|_{\mathcal{L}(H)} \leq \rho\,. \tag{3.4.2}$$

Proof. The operator $B^{-1} \in \mathcal{L}(H)$ is compact and symmetric, and therefore there is an orthonormal basis $(\psi_k)_{k\geq 1}$ of $H$ of eigenfunctions of $B^{-1}$, i.e. $B^{-1}\psi_k = \lambda_k \psi_k$, with eigenvalues $\lambda_k \in \mathbb{R}\backslash\{0\}$, $(\lambda_k) \to 0$. Each $\psi_k$ is an eigenfunction of $B$, i.e.

$$B\psi_k = \nu_k \psi_k \quad \text{with eigenvalue} \quad \nu_k := \lambda_k^{-1}\,, \quad (\nu_k) \to \infty\,.$$

By assumption $(H1)$, each $\nu_k \geq \beta > 0$. Clearly each eigenfunction $\psi_k$ belongs to the domain $\mathcal{D}(B^p)$ of $B^p$ for any $p \geq 1$.

For any $N \geq 1$, we consider the $N$-dimensional subspace $E_N := \mathrm{Span}(\psi_1, \ldots, \psi_N)$ of $H$, and we denote by $\Pi_N$ the corresponding orthogonal projector on $E_N$. We have that

$E_N \subset \mathcal{D}(B^p)$ for any $p \geq 1$, and therefore assumption $(H3)$ implies that $E_N \subset \mathcal{D}(A)$.

PROOF OF $\|A\|_{\mathcal{L}(H)} \leq \rho/\beta$. The operator $A_N := \Pi_N A_{|E_N}$ is a symmetric operator on the finite dimensional Hilbert space $(E_N, \langle\ ,\ \rangle)$ and

$$\|A_N\|_{\mathcal{L}(E_N)} = \max\{|\lambda|\,;\ \lambda \text{ eigenvalue of } A_N\}\,. \tag{3.4.3}$$

Let $\lambda$ be an eigenvalue of $A_N$ and $u \in E_N \backslash \{0\}$ be an associated eigenvector, i.e. $A_N u = \lambda u$. Since $B(E_N) \subset E_N$, the vector $Bu$ is in $E_N$, and we have

$$\lambda\langle u, Bu\rangle = \langle A_N u, Bu\rangle = \langle \Pi_N A u, Bu\rangle = \langle Au, Bu\rangle\,. \tag{3.4.4}$$

Since $\langle u, Bu\rangle$ is positive by $(H1)$, by (3.4.4) and using assumption $(H4)$ (notice that $u$ is in $E_N \subset \mathcal{D}(A) \cap \mathcal{D}(B)$), we get

$$|\lambda|\langle u, Bu\rangle = |\langle Au, Bu\rangle| \leq \rho\|u\|^2\,,$$

and, by assumption $(H1)$, we deduce that $|\lambda| \leq \rho/\beta$. By (3.4.3) we conclude that, for any $N \geq 1$,

$$\|A_N\|_{\mathcal{L}(E_N)} \leq \rho/\beta\,. \tag{3.4.5}$$

Defining the subspace $E := \bigcup_{N\geq 1} E_N$ of $\mathcal{D}(A)$, we deduce by (3.4.5) that

$$\forall (u, v) \in E \times E\,,\ \langle Au, v\rangle \leq \rho\beta^{-1}\|u\|\|v\|\,. \tag{3.4.6}$$

Moreover, since $E$ is dense in $H$ (the $(\psi_k)_{k\geq 1}$ are an orthonormal basis of $H$), the inequality (3.4.6) holds for all $(u, v) \in E \times H$, in particular for all $(u, v) \in E \times \mathcal{D}(A)$. Therefore, since $A$ is symmetric, we obtain that

$$\forall u \in E\,, \forall v \in \mathcal{D}(A)\,,\ \langle u, Av\rangle \leq \rho\beta^{-1}\|u\|\|v\|\,. \tag{3.4.7}$$

By the density of $E$ in $H$, the inequality (3.4.7) holds for all $(u, v) \in H \times \mathcal{D}(A)$, and we conclude that

$$\forall v \in \mathcal{D}(A)\,, \quad \|Av\| \leq \rho\beta^{-1}\|v\|\,. \tag{3.4.8}$$

By continuity and (3.4.8), the operator $A$ can be extended to a bounded operator on the closure $\overline{\mathcal{D}(A)} = H$ (that we still denote by $A$) with operatorial norm $\|A\|_{\mathcal{L}(H)} \leq \rho/\beta$, proving the first estimate in (3.4.2).

PROOF OF $\|B^{\frac{1}{2}} A B^{\frac{1}{2}}\|_{\mathcal{L}(H)} \leq \rho$. The linear operator $B^{\frac{1}{2}}$ is defined on the basis $(\psi_k)_{k\geq 1}$ by setting

$$B^{\frac{1}{2}}\psi_k := \sqrt{\nu_k}\psi_k\,.$$

Notice that, since in assumption $(H3)$ we also require that, for some $p \geq 1$, we have $\mathcal{D}(B^p) \subseteq \mathcal{D}(B^{\frac{1}{2}} A B^{\frac{1}{2}})$, we have the inclusion $E_N \subset \mathcal{D}(B^{\frac{1}{2}} A B^{\frac{1}{2}})$. Since $B^{\frac{1}{2}}(E_N) \subset E_N$, and the operators $B^{\frac{1}{2}}$, $A$ are symmetric on $E_N \subset \mathcal{D}(A)$, then $A'_N := B^{\frac{1}{2}}\Pi_N AB^{\frac{1}{2}}_{|E_N}$ is a symmetric operator of $E_N$. Let $\lambda'$ be an eigenvalue of $A'_N$ and $u' \in E_N\backslash\{0\}$ be an associated eigenvector, i.e. $A'_N u' = \lambda' u'$. Using that $y := B^{\frac{1}{2}}u' \in E_N$ and $By \in E_N$, we obtain (recall that $A_N = \Pi_N A_{|E_N}$)

$$\begin{aligned}\lambda'\langle u', Bu'\rangle = \langle B^{\frac{1}{2}}A_N B^{\frac{1}{2}}u', Bu'\rangle &= \langle B^{\frac{1}{2}}A_N y, B^{\frac{1}{2}}y\rangle \\ &= \langle A_N y, By\rangle = \langle Ay, By\rangle \,. \end{aligned} \tag{3.4.9}$$

Since $\langle u', Bu'\rangle$ is positive by $(H1)$, by (3.4.9) and assumption $(H4)$ (notice that $u'$ is in $E_N \subset \mathcal{D}(A) \cap \mathcal{D}(B)$), we get

$$|\lambda'|\langle u', Bu'\rangle \leq \rho\|y\|^2 = \rho\langle B^{\frac{1}{2}}u', B^{\frac{1}{2}}u'\rangle = \rho\langle u', Bu'\rangle \,. \tag{3.4.10}$$

Since $\langle u', Bu'\rangle > 0$ by $(H1)$, we deduce by (3.4.10) that $|\lambda'| \leq \rho$ and thus $\|A'_N\|_{\mathcal{L}(E_N)} \leq \rho$. Since $B^{\frac{1}{2}}$ and $\Pi_N$ commute and $E_N \subset \mathcal{D}(B^{\frac{1}{2}}AB^{\frac{1}{2}})$, we have

$$A'_N = \Pi_N(A')_{|E_N}\,, \quad A' := B^{\frac{1}{2}}AB^{\frac{1}{2}}\,,$$

and arguing by density as to prove (3.4.8) for $A$, we deduce that for all $v \in \mathcal{D}(A')$, $\|A'v\| \leq \rho\|v\|$. Hence $A'$ can be extended to a bounded linear operator of $H$ with norm $\|A'\|_{\mathcal{L}(H)} \leq \rho$. This proves the second estimate in (3.4.2). ■

In the sequel we shall apply Lemma 3.4.2 with Hilbert space $H = L^2(\mathbb{T}^d)$ and an operator $B \in \{B_1, B_2, B_3\}$ among

$$B_1 := D_m\,, \quad B_2 := D_V\,, \quad B_3 := D_m + D_V\,, \tag{3.4.11}$$

with dense domain

$$\mathcal{D}(B_i) := H^1(\mathbb{T}^d)\,.$$

Notice that each operator $B_i$, $i = 1, 2, 3$, satisfies assumption $(H1)$ (recall (1.1.3)) and $B_i^{-1}$ sends continuously $L^2(\mathbb{T}^d)$ into $H^1(\mathbb{T}^d)$ (note that $\|D_V u\|_{L^2} \simeq \|u\|_{H^1_x}$ by (3.1.3)-(3.1.4)). Moreover, since $H^1(\mathbb{T}^d)$ is compactly embedded into $L^2(\mathbb{T}^d)$, each $B_i^{-1}$ is a compact operator of $H = L^2(\mathbb{T}^d)$, and hence also assumption $(H2)$ holds.

**Lemma 3.4.3.** **($L^2$-bounds of $D_V - D_m$)** *Consider the linear operators*

$$A_0 := D_V - D_m, \tag{3.4.12}$$

$$A'_0 := D_m^{\frac{1}{2}}(D_V - D_m)D_m^{\frac{1}{2}} \tag{3.4.13}$$

*with domains $\mathcal{D}(A_0) := H^1(\mathbb{T}^d)$, $\mathcal{D}(A'_0) := H^2(\mathbb{T}^d)$. Then $A_0$ and $A'_0$ can be extended to bounded linear operators of $L^2(\mathbb{T}^d)$ satisfying*

$$\|A_0\|_{\mathcal{L}(L^2)}, \|A'_0\|_{\mathcal{L}(L^2)} \leq C(\|V\|_{L^\infty})\,. \tag{3.4.14}$$

PROOF. Since $D_m^2 = -\Delta + m.$ and $D_V^2 = -\Delta + V(x)$, we have

$$D_V(D_V - D_m) + (D_V - D_m)D_m = D_V^2 - D_m^2 = \mathrm{Op}(V(x) - m)\,,$$

where $\mathrm{Op}(a)$ denotes the multiplication operator by the function $a(x)$. Hence, for all $u \in H^2(\mathbb{T}^d)$,

$$\begin{aligned}\big|(D_V(D_V - D_m)u, u)_{L^2} + ((D_V - D_m)D_m u, u)_{L^2}\big| &= |((V(x) - m)u, u)_{L^2}| \\ &\le (\|V\|_{L^\infty} + m)\|u\|_{L^2}^2\,,\end{aligned}$$

which gives, by the symmetry of $D_V$ and $D_V - D_m$ that

$$\forall u \in H^2(\mathbb{T}^d)\,,\ |((D_V - D_m)u, (D_V + D_m)u)_{L^2}| \le (\|V\|_{L^\infty} + m)\|u\|_{L^2}^2\,. \tag{3.4.15}$$

Actually (3.4.15) holds for all $u \in H^1(\mathbb{T}^d)$, by the density of $H^2(\mathbb{T}^d)$ in $H^1(\mathbb{T}^d)$ and the fact that $D_V$ and $D_m$ send continuously $H^1(\mathbb{T}^d)$ into $L^2(\mathbb{T}^d)$: setting $A_0 = D_V - D_m$ as in (3.4.12) and $B_3 = D_V + D_m$ as in (3.4.11), we have

$$\forall u \in H^1(\mathbb{T}^d)\,,\ |(A_0 u, B_3 u)_{L^2}| \le (\|V\|_{L^\infty} + m)\|u\|_{L^2}^2\,. \tag{3.4.16}$$

Thus the assumption $(H4)$ of Lemma 3.4.2 holds with $A = A_0$, $B = B_3$, Hilbert space $H = L^2(\mathbb{T}^d)$ and $\rho = \|V\|_{L^\infty} + m$. Applying Lemma 3.4.2 to $A_0$ and $B_3$ (also assumption $(H3)$ holds since $\mathcal{D}(A_0) = \mathcal{D}(B_3) = H^1(\mathbb{T}^d)$) we conclude that $A_0$ can be extended to a bounded operator of $L^2(\mathbb{T}^d)$ with norm

$$\|A_0\|_{\mathcal{L}(L^2)} \le C(\|V\|_{L^\infty}) \tag{3.4.17}$$

(depending also on the constants $m$ and $\beta > 0$ in (1.1.3)). This proves the first bound in (3.4.14). Then, recalling the definitions of $A_0, B_1, B_3$ in (3.4.12), (3.4.11), using (3.4.16) and (3.4.17), we also deduce that

$$\forall u \in H^1(\mathbb{T}^d)\,,\ |(A_0 u, B_1 u)_{L^2}| = \frac{1}{2}|(A_0 u, (B_3 - A_0)u)_{L^2}| \le C'(\|V\|_{L^\infty})\|u\|_{L^2}^2\,.$$

Thus the assumption $(H4)$ of Lemma 3.4.2 holds with $A = A_0$, $B = B_1$. Also assumption $(H3)$ holds since $\mathcal{D}(A_0) = \mathcal{D}(B_1) = H^1(\mathbb{T}^d)$ and $\mathcal{D}(B_1^{\frac12} A_0 B_1^{\frac12}) \supset H^2(\mathbb{T}^d) = \mathcal{D}(B_1^2)$. Therefore Lemma 3.4.2 implies that $A_0' = B_1^{\frac12} A_0 B_1^{\frac12}$ can be extended to a bounded operator of $L^2(\mathbb{T}^d)$, with operatorial norm $\|A_0'\|_{\mathcal{L}(L^2)} \le C(\|V\|_{L^\infty})$. This proves the second bound in (3.4.14). ∎

Let $[A_0']_i^j = (|i|^2 + m)^{1/4}[D_V - D_m]_i^j(|j|^2 + m)^{1/4}$, $i, j \in \mathbb{Z}^d$, denote the elements of the matrix representing the operator $A_0' = D_m^{\frac12}(D_V - D_m)D_m^{\frac12}$ in the exponential basis. By Lemma 3.4.3 we have that

$$|[A_0']_i^j| \le \|A_0'\|_{\mathcal{L}(L^2)} \le C(\|V\|_{L^\infty})\,, \quad \forall i, j \in \mathbb{Z}^d\,. \tag{3.4.18}$$

In order to prove also a polynomial off-diagonal decay for $[A_0']_i^j$, $i \neq j$, we notice that, for $n \geq 1$,

$$\mathrm{Ad}^n_{\partial_{x_k}} A_0' \quad \text{is represented by the matrix} \quad \Big( \mathrm{i}^n (i_k - j_k)^n [A_0']_i^j \Big)_{i,j \in \mathbb{Z}^d} \tag{3.4.19}$$

and then prove that $\mathrm{Ad}^n_{\partial_{x_k}} A_0'$ extends to a bounded operator in $L^2(\mathbb{T}^d)$ .

**Lemma 3.4.4.** **($L^2$-bounds of $\mathrm{Ad}^n_{\partial_{x_k}} D_V$)** *For any $n \geq 1$, $k = 1, \ldots, d$ the operators $\mathrm{Ad}^n_{\partial_{x_k}} D_V$ and $D_m^{\frac12}(\mathrm{Ad}^n_{\partial_{x_k}} D_V) D_m^{\frac12}$ can be extended to bounded operators of $L^2(\mathbb{T}^d)$: there exist positive constants $C_n$ and $C_n'$ depending on $\|V\|_{C^n}$ such that*

$$\big\| \mathrm{Ad}^n_{\partial_{x_k}} D_V \big\|_{\mathcal{L}(L^2)} \leq C_n \,, \tag{3.4.20}$$

$$\big\| D_m^{\frac12} (\mathrm{Ad}^n_{\partial_{x_k}} D_V) D_m^{\frac12} \big\|_{\mathcal{L}(L^2)} \leq C_n' \,. \tag{3.4.21}$$

Proof. We shall use the following algebraic formulas: given linear operators $L_1$, $L_2$ we have

$$\mathrm{Ad}_{\partial_{x_k}} (L_1 L_2) = (\mathrm{Ad}_{\partial_{x_k}} L_1) L_2 + L_1 (\mathrm{Ad}_{\partial_{x_k}} L_2) \,, \tag{3.4.22}$$

$$\mathrm{Ad}^n_{\partial_{x_k}} (L_1 L_2) = \sum_{n_1 = 0}^{n} \binom{n}{n_1} \big(\mathrm{Ad}^{n_1}_{\partial_{x_k}} L_1\big)\big(\mathrm{Ad}^{n-n_1}_{\partial_{x_k}} L_2\big) \,. \tag{3.4.23}$$

We split the proof in two steps.

**1st step.** We prove by iteration that, for all $n \geq 1$, there are constants $C_n, C_n'' > 0$ such that

$$\| (\mathrm{Ad}^n_{\partial_{x_k}} D_V) D_V + D_V (\mathrm{Ad}^n_{\partial_{x_k}} D_V) \|_{\mathcal{L}(L^2)} \leq C_n'' \tag{3.4.24}$$

$$\| \mathrm{Ad}^n_{\partial_{x_k}} D_V \|_{\mathcal{L}(L^2)} \leq C_n \,. \tag{3.4.25}$$

Clearly the estimates (3.4.25) are (3.4.20).

Initialization: proof of (3.4.24)-(3.4.25) for $n = 1$. Applying (3.4.22) with $L_1 = L_2 = D_V$ and since $D_V^2 = -\Delta + V(x)$, we get

$$(\mathrm{Ad}_{\partial_{x_k}} D_V) D_V + D_V (\mathrm{Ad}_{\partial_{x_k}} D_V) = \mathrm{Ad}_{\partial_{x_k}} (-\Delta + V(x)) = \mathrm{Op}(V_{x_k}(x)) \tag{3.4.26}$$

where $\mathrm{Op}(V_{x_k})$ is the multiplication operator by the function $V_{x_k}(x) := (\partial_{x_k} V)(x)$. Hence (3.4.24) for $n = 1$ holds with $C_1'' = \|V\|_{C^1}$. In order to prove (3.4.25) for $n = 1$ we apply Lemma 3.4.2 to the operators $A_1 := \mathrm{Ad}_{\partial_{x_k}} D_V$ and $B_2 = D_V$. Assumption $(H3)$ holds because $\mathcal{D}(A_1) = H^2(\mathbb{T}^d) = \mathcal{D}(B_2^2)$. Note that, because of the $L^2$-antisymmetry of the operator $\partial_{x_k}$, if $L$ is $L^2$-symmetric then so is $\mathrm{Ad}_{\partial_{x_k}} L$. Hence $A_1 = \mathrm{Ad}_{\partial_{x_k}} D_V$ is symmetric. Also assumption $(H4)$ holds because

$$\begin{aligned}\forall u \in H^2(\mathbb{T}^d)\,,\ |(A_1 u, B_2 u)_{L^2}| &= |(\mathrm{Ad}_{\partial_{x_k}} D_V u, D_V u)_{L^2}| \\ &= \frac{1}{2}\big|\big((\mathrm{Ad}_{\partial_{x_k}} D_V) D_V + D_V (\mathrm{Ad}_{\partial_{x_k}} D_V) u, u\big)_{L^2}\big| \\ &\leq \frac{C_1''}{2}\|u\|_{L^2}^2\,. \end{aligned} \tag{3.4.27}$$

by (3.4.24) for $n = 1$. Therefore Lemma 3.4.2 implies that (3.4.25) holds for $n = 1$, for some constant $C_1$ depending on $\|V\|_{C^1}$.

Iteration: proof of (3.4.24)-(3.4.25) for $n > 1$. Assume by induction that (3.4.24)-(3.4.25) have been proved up to rank $n-1$. We now prove them at rank $n$. Applying (3.4.23) with $L_1 = L_2 = D_V$ and since $D_V^2 = -\Delta + V(x)$, we get

$$\begin{aligned}&\big(\mathrm{Ad}^n_{\partial_{x_k}} D_V\big) D_V + D_V\big(\mathrm{Ad}^n_{\partial_{x_k}} D_V\big) + \sum_{n_1=1}^{n-1}\binom{n}{n_1}\big(\mathrm{Ad}^{n_1}_{\partial_{x_k}} D_V\big)\big(\mathrm{Ad}^{n-n_1}_{\partial_{x_k}} D_V\big) \\ &= \mathrm{Ad}^n_{\partial_{x_k}}(-\Delta + V(x)) = \mathrm{Op}((\partial^n_{x_k} V))\,. \end{aligned} \tag{3.4.28}$$

Let $T_n := \big(\mathrm{Ad}^n_{\partial_{x_k}} D_V\big) D_V + D_V\big(\mathrm{Ad}^n_{\partial_{x_k}} D_V\big)$. By (3.4.28) and using the inductive assumption (3.4.25) at rank $n-1$, we obtain

$$\|T_n\|_{\mathcal{L}(L^2(\mathbb{T}^d))} \leq \|V\|_{C^n} + \sum_{n_1=1}^{n-1}\binom{n}{n_1}\|\mathrm{Ad}^{n_1}_{\partial_{x_k}} D_V\|_{\mathcal{L}(L^2)}\|\mathrm{Ad}^{n-n_1}_{\partial_{x_k}} D_V\|_{\mathcal{L}(L^2)} \leq C_n'' \tag{3.4.29}$$

where the constant $C_n''$ depends on $\|V\|_{C^n}$. We have that

$$\forall u \in H^{n+1}(\mathbb{T}^d)\,,\ |(\mathrm{Ad}^n_{\partial_{x_k}} D_V u, D_V u)_{L^2}| = \frac{1}{2}|(T_n u, u)_{L^2}| \overset{(3.4.29)}{\leq} \frac{C_n''}{2}\|u\|_{L^2}^2\,. \tag{3.4.30}$$

We now apply Lemma 3.4.2 to $A_n := \mathrm{Ad}^n_{\partial_{x_k}} D_V$ and $B_2 = D_V$. Assumption $(H3)$ holds because $\mathcal{D}(A_n) = H^{n+1}(\mathbb{T}^d) = \mathcal{D}(B_2^{n+1})$ by (3.1.4). Assumption $(H4)$ holds by (3.4.30). Therefore the first inequality in (3.4.2) implies that $\|A_n\|_{\mathcal{L}(L^2(\mathbb{T}^d))} \leq C_n(\|V\|_{C^n})$. This proves (3.4.25) at rank $n$.

**2nd step.** Proof of (3.4.21) for any $n \geq 1$. In order to prove that the operator $D_m^{\frac{1}{2}}\big(\mathrm{Ad}^n_{\partial_{x_k}} D_V\big) D_m^{\frac{1}{2}}$ extends to a bounded operator of $L^2(\mathbb{T}^d)$ we use Lemma 3.4.2 with $A = A_n = \mathrm{Ad}^n_{\partial_{x_k}} D_V$ and $B = D_m$. Assumption $(H3)$ holds because $\mathcal{D}(A_n) = H^{n+1}(\mathbb{T}^d) = \mathcal{D}(D_m^{n+1})$ and $\mathcal{D}(D_m^{\frac{1}{2}} A_n D_m^{\frac{1}{2}}) \supset H^{n+2}(\mathbb{T}^d) = \mathcal{D}(D_m^{n+2})$. Also Assumption (H4) holds because, by (3.4.30), and the fact that $A_n$ and $D_m - D_V$ are bounded operators of $L^2(\mathbb{T}^d)$ by (3.4.25) and Lemma 3.4.3, we get

$$\forall u \in H^{n+1}(\mathbb{T}^d)\,,\ |(A_n u, D_m u)_{L^2}| \le |(A_n u, D_V u)_{L^2}| + |(A_n u, D_m u - D_V u)_{L^2}| \\ \le \rho_n \|u\|_{L^2}^2\,,$$

where $\rho_n$ depends on $\|V\|_{C^n}$. Then the second inequality in (3.4.2) implies (3.4.21). ∎

**Proof of Proposition 3.4.1 concluded.** Recalling (3.4.18) the entries of the matrix $([A_0']_i^j)_{i,j\in\mathbb{Z}^d}$, which represents the operator $A_0' := D_m^{\frac12}(D_V - D_m)D_m^{\frac12}$ in the exponential basis, are bounded. Furthermore, since each $\partial_{x_k}$, $k = 1,\dots,d$, commutes with $D_m$ and $D_m^{\frac12}$, we have that, for any $n \ge 1$,

$$\mathrm{Ad}^n_{\partial_{x_k}} A_0' = D_m^{\frac12}(\mathrm{Ad}^n_{\partial_{x_k}} D_V)D_m^{\frac12}\,.$$

By Lemma 3.4.4, this operator can be extended to a bounded operator on $L^2(\mathbb{T}^d)$ satisfying (see (3.4.21))

$$\|\mathrm{Ad}^n_{\partial_{x_k}} A_0'\|_{\mathcal{L}(L^2(\mathbb{T}^d))} \le C_n'(\|V\|_{C^n})\,. \tag{3.4.31}$$

By (3.4.18), (3.4.19), (3.4.31) we deduce that

$$\forall n \ge 0\,,\ \forall k = 1,\dots,d\,,\ \forall (i,j) \in \mathbb{Z}^d\times\mathbb{Z}^d\,,\\ |i_k - j_k|^n |[A_0']_i^j| \le \|\mathrm{Ad}^n_{\partial_{x_k}} A_0'\|_{\mathcal{L}(L^2(\mathbb{T}^d))} \le C_n'(\|V\|_{C^n})$$

with the convention $\mathrm{Ad}^0_{\partial_{x_k}} := \mathrm{Id}$. Hence, for $s \ge 0$ and $n_s := \lceil s + \frac{d}{2}\rceil + 1 \ge s + \frac{d}{2} + 1$, we deduce that the $s$-decay norm of $A_0'$ satisfies

$$|A_0'|_s^2 \le K_0^2 + \sum_{\ell\in\mathbb{Z}^d\setminus\{0\}} \frac{K_{n_s}^2}{|\ell|^{2n_s}}|\ell|^{2s} \le K_{n_s}^2 \sum_{\ell\in\mathbb{Z}^d} \frac{1}{\langle\ell\rangle^{d+2}} =: \Upsilon_s^2$$

for some constant $\Upsilon_s$ depending on $s$ and $\|V\|_{C^{n_s}}$ only. The proof of Proposition 3.4.1 is complete. ∎

With the same methods we also obtain the following propositions.

**Proposition 3.4.5.** $|D_V^{\frac12} - D_m^{\frac12}|_s \le C(s, \|V\|_{C^{n_s}})$ *where* $n_s := \lceil s + \frac{d}{2}\rceil + 1$.

Proof. Since the proof closely follows Lemmata 3.4.3 and 3.4.4 we shall indicate just the main steps. One first proves that $D_V^{\frac12} - D_m^{\frac12}$ can be extended to a bounded operator of $L^2$, arguing as in Lemma 3.4.3. Writing

$$D_V^{\frac12}(D_V^{\frac12} - D_m^{\frac12}) + (D_V^{\frac12} - D_m^{\frac12})D_m^{\frac12} = D_V - D_m\,,$$

using the symmetry of $D_V^{\frac{1}{2}}$ and $D_V^{\frac{1}{2}} - D_m^{\frac{1}{2}}$, that $D_V - D_m$ is bounded on $L^2$ by Lemma 3.4.3 (see (3.4.14)), and the density of $H^1$ in $H^{\frac{1}{2}}$, we deduce that

$$\forall u \in H^{\frac{1}{2}}(\mathbb{T}^d)\,, \quad |((D_V^{\frac{1}{2}} - D_m^{\frac{1}{2}})u, (D_V^{\frac{1}{2}} + D_m^{\frac{1}{2}})u)_{L^2}| \leq C(\|V\|_{L^\infty})\|u\|_{L^2}^2\,.$$

Then, arguing as in Lemma 3.4.3, Lemma 3.4.2 implies that $D_V^{\frac{1}{2}} - D_m^{\frac{1}{2}}$ can be extended to a bounded operator of $L^2$ satisfying

$$\|D_V^{\frac{1}{2}} - D_m^{\frac{1}{2}}\|_{\mathcal{L}(L^2)} \leq C(\|V\|_{L^\infty})\,. \tag{3.4.32}$$

Next one proves that, for all $n \geq 1$, $k = 1, \ldots, d$, the operators $\mathrm{Ad}^n_{\partial_{x_k}} D_V^{\frac{1}{2}}$ can be extended to bounded operators on $L^2$ satisfying

$$\big\|\mathrm{Ad}^n_{\partial_{x_k}} D_V^{\frac{1}{2}}\big\|_{\mathcal{L}(L^2)} \leq C_n(\|V\|_{C^n})\,. \tag{3.4.33}$$

It is sufficient to argue as in Lemma 3.4.4, applying (3.4.22)-(3.4.23) to $L_1 = L_2 = D_V^{\frac{1}{2}}$ and $D_V^{\frac{1}{2}} D_V^{\frac{1}{2}} = D_V$, and using that $\mathrm{Ad}^n_{\partial_{x_k}} D_V$ are bounded operators of $L^2$ satisfying (3.4.20). The Proposition follows by (3.4.32) and (3.4.33) as in the conclusion of the proof of Proposition 3.4.1. ∎

**Proposition 3.4.6.** *$|D_m^{\frac{1}{2}} D_V^{-\frac{1}{2}}|_s$, $|D_V^{-\frac{1}{2}} D_m^{\frac{1}{2}}|_s \leq C(s, \|V\|_{C^{n_s}})$.*

Proof. Writing

$$D_V^{-\frac{1}{2}} D_m^{\frac{1}{2}} = \mathrm{Id} - D_V^{-\frac{1}{2}}(D_V^{\frac{1}{2}} - D_m^{\frac{1}{2}})\,,$$
$$D_m^{\frac{1}{2}} D_V^{-\frac{1}{2}} = \mathrm{Id} - (D_V^{\frac{1}{2}} - D_m^{\frac{1}{2}}) D_V^{-\frac{1}{2}}\,,$$

Proposition 3.4.6 follows by Proposition 3.4.5, (3.3.6), and

$$|D_V^{-\frac{1}{2}}|_s \leq C(s, \|V\|_{C^{n_s}})\,. \tag{3.4.34}$$

The estimate (3.4.34) follows, as in the conclusion of the proof of Proposition 3.4.1, by the fact that $D_V^{-\frac{1}{2}}$ is a bounded operator of $L^2$ and

$$\big\|\mathrm{Ad}^n_{\partial_{x_k}} D_V^{-\frac{1}{2}}\big\|_{\mathcal{L}(L^2)} \leq C_n(\|V\|_{C^n})\,, \quad \forall n \geq 1\,,\ k = 1, \ldots, d\,. \tag{3.4.35}$$

The estimates (3.4.35) can be proved applying (3.4.22)-(3.4.23) to $L_1 = L_2 = D_V^{\frac{1}{2}}$ and $D_V^{\frac{1}{2}} D_V^{-\frac{1}{2}} = \mathrm{Id}$. For example, applying (3.4.22), we get

$$\big(\mathrm{Ad}_{\partial_{x_k}} D_V^{-\frac{1}{2}}\big) = -D_V^{-\frac{1}{2}}\big(\mathrm{Ad}_{\partial_{x_k}} D_V^{\frac{1}{2}}\big) D_V^{-\frac{1}{2}}$$

which is a bounded operator on $L^2$ satisfying (3.4.33), and $D_V^{-\frac{1}{2}} \in \mathcal{L}(L^2)$. By iteration, the estimate (3.4.35) follows for any $n \geq 1$. ∎

## 3.5 Interpolation inequalities

We conclude this Chapter with useful interpolation inequalities for the Sobolev spaces $\mathcal{H}^s$ defined in (3.3.2).

For any $s \geq s_0 > (|\mathbb{S}| + d)/2$, we have the tame estimate for the product

$$\|fg\|_{\mathrm{Lip},s} \lesssim_s \|f\|_{\mathrm{Lip},s}\|g\|_{\mathrm{Lip},s_0} + \|f\|_{\mathrm{Lip},s_0}\|g\|_{\mathrm{Lip},s}\,. \tag{3.5.1}$$

Actually we directly prove the improved tame estimate (3.5.2) below, used in [23], [25].

**Lemma 3.5.1.** *Let $s \geq s_0 > (|\mathbb{S}| + d)/2$. Then*

$$\|uv\|_{\mathrm{Lip},s} \leq C_0\|u\|_{\mathrm{Lip},s}\|v\|_{\mathrm{Lip},s_0} + C(s)\|u\|_{\mathrm{Lip},s_0}\|v\|_{\mathrm{Lip},s} \tag{3.5.2}$$

*where the constant $C_0 > 0$ is independent of $s \geq s_0$.*

PROOF. Denoting $y := (\varphi, x) \in \mathbb{T}^b$, $b = |\mathbb{S}| + d$, we expand in Fourier series

$$u(y) = \sum_{m\in\mathbb{Z}^b} u_m\, e^{\mathrm{i}m\cdot y}\,, \qquad v(y) = \sum_{m\in\mathbb{Z}^b} v_m\, e^{\mathrm{i}m\cdot y}\,.$$

Thus

$$\|u\,v\|_s^2 = \sum_{m\in\mathbb{Z}^b}\Big|\sum_{k\in\mathbb{Z}^b} u_k v_{m-k}\Big|^2\langle m\rangle^{2s} \leq \sum_{m\in\mathbb{Z}^b}\Big(\sum_{k\in\mathbb{Z}^b}|u_k||v_{m-k}|\Big)^2\langle m\rangle^{2s} \leq S_1 + S_2 \tag{3.5.3}$$

where

$$S_1 := 2\sum_{m\in\mathbb{Z}^b}\Big(\sum_{|m|<|k|2^{1/s}}|u_k||v_{m-k}|\Big)^2\langle m\rangle^{2s}$$
$$S_2 := 2\sum_{m\in\mathbb{Z}^b}\Big(\sum_{|m|\geq|k|2^{1/s}}|u_k||v_{m-k}|\Big)^2\langle m\rangle^{2s}\,.$$

The indices in the sum $S_1$ are restricted to $|m| < |k|2^{1/s}$, thus $(\langle m\rangle/\langle k\rangle)^s \leq 2$, and, using Cauchy-Schwarz inequality, we deduce

$$\begin{aligned} S_1 &= 2\sum_{m\in\mathbb{Z}^b}\sum_{|m|<|k|2^{1/s}}\Big(|u_k|\langle k\rangle^s|v_{m-k}|\langle m-k\rangle^{s_0}\frac{\langle m\rangle^s}{\langle k\rangle^s\langle m-k\rangle^{s_0}}\Big)^2 \\ &\leq 2\sum_{m\in\mathbb{Z}^b}\Big(\sum_{k\in\mathbb{Z}^b}|u_k|^2\langle k\rangle^{2s}|v_{m-k}|^2\langle m-k\rangle^{2s_0}4\Big)\Big(\sum_{k\in\mathbb{Z}^b}\frac{1}{\langle m-k\rangle^{2s_0}}\Big) \\ &\leq C(s_0)\|u\|_s^2\|v\|_{s_0}^2\,. \end{aligned} \tag{3.5.4}$$

On the other hand, the indices in the sum $S_2$ are restricted to $|k| \leq |m|2^{-1/s}$, and therefore

$$|m-k| \geq |m| - |k| \geq |m|(1-2^{-1/s})$$

and $(\langle m\rangle/\langle m-k\rangle)^s \leq c(s)$. As a consequence

$$\begin{aligned} S_2 &= 2\sum_{m\in\mathbb{Z}^b} \sum_{|k|\leq |m|2^{-1/s}} \Big(|u_k|\langle k\rangle^{s_0}|v_{m-k}|\langle m-k\rangle^s \frac{\langle m\rangle^s}{\langle k\rangle^{s_0}\langle m-k\rangle^s}\Big)^2 \\ &\leq 2\sum_{m\in\mathbb{Z}^b} \Big(\sum_{k\in\mathbb{Z}^b} |u_k|^2\langle k\rangle^{2s_0}|v_{m-k}|^2\langle m-k\rangle^{2s}c(s)^2\Big)\Big(\sum_{k\in\mathbb{Z}^b}\frac{1}{\langle k\rangle^{2s_0}}\Big) \\ &\leq C(s)\|u\|_{s_0}^2\|v\|_s^2 . \end{aligned} \tag{3.5.5}$$

By (3.5.3) and the estimates (3.5.4)-(3.5.5) we deduce

$$\|u\,v\|_s \leq C(s_0)\|u\|_s\|v\|_{s_0} + C(s)\|u\|_{s_0}\|v\|_s .$$

Recalling (3.3.3), the estimate (3.5.2) follows. ∎

In the case when $0 \leq s \leq s_0$, the estimate of $\|uv\|_s$ can be simplified.

**Lemma 3.5.2.** *Let* $s_0 > (|\mathbb{S}|+d)/2$ *and* $0 \leq s \leq s_0$. *Then*

$$\|uv\|_{\mathrm{Lip},s} \leq C(s_0)\|u\|_{\mathrm{Lip},s_0}\|v\|_{\mathrm{Lip},s} . \tag{3.5.6}$$

Proof. Denoting $b = (|\mathbb{S}|+d)/2$, we have, as in the proof of Lemma 3.5.1,

$$\begin{aligned} \|uv\|_s^2 &\leq \sum_{m\in\mathbb{Z}^b} \Big(\sum_{k\in\mathbb{Z}^b} \langle m\rangle^s|u_k||v_{m-k}|\Big)^2 \\ &\leq \sum_{m\in\mathbb{Z}^b} \Big(\sum_{k\in\mathbb{Z}^b} |u_k|\langle k\rangle^{s_0}|v_{m-k}|\langle m-k\rangle^s \frac{\langle m\rangle^s}{\langle k\rangle^{s_0}\langle m-k\rangle^s}\Big)^2 \\ &\leq \sum_{m\in\mathbb{Z}^b} \Big(\sum_{k\in\mathbb{Z}^b} |u_k|^2\langle k\rangle^{2s_0}|v_{m-k}|^2\langle m-k\rangle^{2s}\Big)\Big(\sum_{k\in\mathbb{Z}^b}\frac{\langle m\rangle^{2s}}{\langle k\rangle^{2s_0}\langle m-k\rangle^{2s}}\Big) \end{aligned} \tag{3.5.7}$$

by the Cauchy-Schwarz inequality. We now use the following inequality: for $0 \leq s \leq s_0$,

$$\forall m \in \mathbb{Z}^b\,,\ \forall k \in \mathbb{Z}^b\,,\ \frac{\langle m\rangle^s}{\langle m-k\rangle^s\langle k\rangle^{s_0}} \leq 2^s\Big(\frac{1}{\langle k\rangle^{s_0}} + \frac{1}{\langle m-k\rangle^{s_0}}\Big) . \tag{3.5.8}$$

To prove (3.5.8), we distinguish two cases: if $\langle m-k\rangle \geq \langle m\rangle/2$ then (3.5.8) is trivial. If $\langle m-k\rangle < \langle m\rangle/2$, then $\langle k\rangle > \langle m\rangle/2 > \langle m-k\rangle$ and, since $s_0 - s \geq 0$,

$$\langle m\rangle^s \leq 2^s\langle k\rangle^s = 2^s\frac{\langle k\rangle^{s_0}}{\langle k\rangle^{s_0-s}} \leq 2^s\frac{\langle k\rangle^{s_0}}{\langle m-k\rangle^{s_0-s}} ,$$

which implies (3.5.8). By (3.5.8) we deduce that, for any $m \in \mathbb{Z}^b$,

$$\sum_{k \in \mathbb{Z}^b} \frac{\langle m \rangle^{2s}}{\langle k \rangle^{2s_0} \langle m-k \rangle^{2s}} \lesssim_s \sum_{k \in \mathbb{Z}^b} \frac{1}{\langle k \rangle^{2s_0}} + \sum_{k \in \mathbb{Z}^b} \frac{1}{\langle m-k \rangle^{2s_0}} \leq C(s_0)$$

and therefore, by (3.5.7), we obtain

$$\|uv\|_s \leq C(s_0) \|u\|_{s_0} \|v\|_s \, . \tag{3.5.9}$$

Recalling (3.3.3), the estimate (3.5.6) is a consequence of (3.5.9). ■

As in any scale of Sobolev spaces with smoothing operators, the Sobolev norms $\| \ \|_s$ defined in (3.3.2) admit an interpolation estimate.

**Lemma 3.5.3.** *For any $s_1 < s_2$, $s_1, s_2 \in \mathbb{R}$, and $\theta \in [0,1]$, we have*

$$\|h\|_{\mathrm{Lip},s} \leq 2 \|h\|^{\theta}_{\mathrm{Lip},s_1} \|h\|^{1-\theta}_{\mathrm{Lip},s_2} \, , \quad s := \theta s_1 + (1-\theta) s_2 \, . \tag{3.5.10}$$

PROOF. Recalling the definition of the Sobolev norm in (3.3.2), we deduce, by Hölder inequality,

$$\begin{aligned} \|h\|_s^2 = \sum_{i \in \mathbb{Z}^b} |h_i|^2 \langle i \rangle^{2s} &= \sum_{i \in \mathbb{Z}^b} (|h_i|^2 \langle i \rangle^{2s_1})^{\theta} (|h_i|^2 \langle i \rangle^{2s_2})^{1-\theta} \\ &\leq \Big( \sum_{i \in \mathbb{Z}^b} |h_i|^2 \langle i \rangle^{2s_1} \Big)^{\theta} \Big( \sum_{i \in \mathbb{Z}^b} |h_i|^2 \langle i \rangle^{2s_2} \Big)^{1-\theta} \\ &= \|h\|_{s_1}^{2\theta} \|h\|_{s_2}^{2(1-\theta)} \, . \end{aligned}$$

Thus $\|h\|_s \leq \|h\|^{\theta}_{s_1} \|h\|^{1-\theta}_{s_2}$, and, for a Lipschitz family of Sobolev functions, see (3.3.3), the inequality (3.5.10) follows. ■

As a corollary we deduce the following interpolation inequality.

**Lemma 3.5.4.** *Let $\alpha \leq a \leq b \leq \beta$ such that $a + b = \alpha + \beta$. Then*

$$\|h\|_{\mathrm{Lip},a} \|h\|_{\mathrm{Lip},b} \leq 4 \|h\|_{\mathrm{Lip},\alpha} \|h\|_{\mathrm{Lip},\beta} \, . \tag{3.5.11}$$

*Proof.* Write $a = \theta\alpha + (1-\theta)\beta$ and $b = \mu\alpha + (1-\mu)\beta$, where

$$\theta = \frac{a - \beta}{\alpha - \beta}, \quad \mu = \frac{b - \beta}{\alpha - \beta}, \quad \theta + \mu = 1 \ .$$

By the interpolation Lemma 3.5.3, we get

$$\|h\|_{\mathrm{Lip},a} \leq 2 \|h\|^{\theta}_{\mathrm{Lip},\alpha} \|h\|^{1-\theta}_{\mathrm{Lip},\beta}, \quad \|h\|_{\mathrm{Lip},b} \leq 2 \|h\|^{\mu}_{\mathrm{Lip},\alpha} \|h\|^{1-\mu}_{\mathrm{Lip},\beta} \, ,$$

and (3.5.11) follows multiplying these inequalities. □

We finally recall the following Moser tame estimates for the composition operator

$$(u_1, \dots, u_p) \mapsto \mathtt{f}(u_1, \dots, u_p)(\varphi, x) := f(\varphi, x, u_1(\varphi, x), \dots, u_p(\varphi, x)) \tag{3.5.12}$$

induced by a smooth function $f$.

**Lemma 3.5.5. (Composition operator)** *Let $f \in \mathcal{C}^\infty(\mathbb{T}^{|\mathbb{S}|} \times \mathbb{T}^d \times \mathbb{R}^p, \mathbb{R})$. Fix $s_0 > (d + |\mathbb{S}|)/2$, $s_0 \in \mathbb{N}$. Given real valued functions $u_i$, $1 \le i \le p$, satisfying $\|u_i\|_{\mathrm{Lip}, s_0} \le 1$, then, $\forall s \ge s_0$,*

$$\|\mathtt{f}(u_1, \dots, u_p)\|_{\mathrm{Lip}, s} \le C(s, f)\Big(1 + \sum_{i=1}^p \|u_i\|_{\mathrm{Lip}, s}\Big). \tag{3.5.13}$$

*Assuming also that $\|v_i\|_{\mathrm{Lip}, s_0} \le 1$, $1 \le i \le p$, then*

$$\begin{aligned} \|\mathtt{f}(v_1, \dots, v_p) - \mathtt{f}(u_1, \dots, u_p)\|_{\mathrm{Lip}, s} \lesssim_{s,f} & \sum_{i=1}^p \|v_i - u_i\|_{\mathrm{Lip}, s} + \\ & \Big(\sum_{i=1}^p \|u_i\|_{\mathrm{Lip}, s} + \|v_i\|_{\mathrm{Lip}, s}\Big) \sum_{i=1}^p \|v_i - u_i\|_{\mathrm{Lip}, s_0}. \end{aligned} \tag{3.5.14}$$

PROOF. Let $y := (\varphi, x) \in \mathbb{T}^b$, $b = |\mathbb{S}| + d$. For simplicity of notation we consider only the case $p = 1$ and denote $(u_1, \dots, u_p) = u_1 = u$.

STEP 1. For any $s \ge 0$, for any function $f \in C^\infty$, for all $\|u\|_{s_0} \le 1$,

$$\|\mathtt{f}(u)\|_s \le C(s, f)\big(1 + \|u\|_s\big). \tag{3.5.15}$$

This estimate was proved by Moser in [98]. We propose here a different proof, following [20]. Note that it is enough to prove (3.5.15) for $u \in C^\infty(\mathbb{T}^b)$.

*Initialization:* (3.5.15) *holds for any $s \in [0, s_0]$.* For a multi-index $\alpha = (\alpha_1, \dots, \alpha_b) \in \mathbb{N}^b$, we denote $\partial_y^\alpha = \partial_{y_1}^{\alpha_1} \dots \partial_{y_b}^{\alpha_b}$ and we set $|\alpha| := \alpha_1 + \dots + \alpha_b$. Recalling that $s_0$ is an integer, by the formula for the derivative of a composition of functions, we estimate the Sobolev norm $\|\mathtt{f}(u)\|_{s_0}$ by

$$C(s_0) \max\big\{\|(\partial_y^\beta \partial_u^q f)(\cdot, u)(\partial_y^{\alpha^{(1)}} u) \dots (\partial_y^{\alpha^{(q)}} u)\|_0 \,;\, 0 \le q \le s_0\,,\, |\alpha^{(1)} + \dots + \alpha^{(q)} + \beta| \le s_0\big\}$$

where $\alpha^{(1)}, \dots, \alpha^{(q)}, \beta \in \mathbb{N}^b$. By the Sobolev embedding $\|u\|_{L^\infty(\mathbb{T}^b)} \lesssim_{s_0} \|u\|_{s_0}$ we have

$$\|(\partial_y^\beta \partial_u^q f)(y, u(y))\|_{L^\infty(\mathbb{T}^b)} \le C(f, \|u\|_{s_0}).$$

Hence, in order to prove the bound $\|\mathtt{f}(u)\|_{s_0} \lesssim_f 1 + \|u\|_{s_0}$, it is enough to check that, for any $0 \le q \le s_0$, $|\alpha^{(1)} + \dots + \alpha^{(q)}| \le s_0$, the function

$$v := (\partial_y^{\alpha^{(1)}} u) \dots (\partial_y^{\alpha^{(q)}} u) \quad \text{satisfies} \quad \|v\|_0 \le C(s_0)\|u\|_{s_0}^q. \tag{3.5.16}$$

Expanding in Fourier $u(y) = \sum_{k \in \mathbb{Z}^b} u_k e^{\mathrm{i} k \cdot y}$, we have

$$\|v\|_0^2 = \sum_{k \in \mathbb{Z}^b} \Big| \sum_{k_1 + \ldots + k_q = k} (\mathrm{i}k_1)^{\alpha^{(1)}} \ldots (\mathrm{i}k_q)^{\alpha^{(q)}} u_{k_1} \ldots u_{k_q} \Big|^2 , \tag{3.5.17}$$

where, for $w \in \mathbb{C}^b$ and $\alpha \in \mathbb{N}^b$, we use the multi-index notation $w^\alpha := \prod_{l=1}^{b} w_l^{\alpha_l}$. Since $|\alpha^{(1)}| + \ldots + |\alpha^{(1)}| \le s_0$, the iterated Young inequality implies

$$|(\mathrm{i}k_1)^{\alpha^{(1)}} \ldots (\mathrm{i}k_q)^{\alpha^{(q)}}| \le |k_1|^{|\alpha^{(1)}|} \ldots |k_q|^{|\alpha^{(q)}|} \le \langle k_1 \rangle^{s_0} + \ldots + \langle k_q \rangle^{s_0} ,$$

and therefore by (3.5.17) we have

$$\|v\|_0^2 \lesssim_q \sum_{i=1}^{q} S_i \quad \text{with} \quad S_i := \sum_{k \in \mathbb{Z}^b} \Big( \sum_{k_1 + \ldots + k_q = k} \langle k_i \rangle^{s_0} |u_{k_1}| \ldots |u_{k_q}| \Big)^2 .$$

An application of the Cauchy-Schwarz inequality, using

$$\sum_{k_2, \ldots, k_q \in \mathbb{Z}^b} \langle k_2 \rangle^{-2s_0} \ldots \langle k_q \rangle^{-2s_0} < \infty ,$$

provides the bound $S_1 \lesssim_{s_0} \|u\|_{s_0}^{2q}$. The bounds for the other $S_i$ are obtained similarly. This proves (3.5.16) and therefore (3.5.15) for $s = s_0$. As a consequence, (3.5.15) trivially holds also for any $s \in [0, s_0]$: in fact, for all $\|u\|_{s_0} \le 1$, we have

$$\|\mathtt{f}(u)\|_s \le \|\mathtt{f}(u)\|_{s_0} \le C(f) \le C(f)(1 + \|u\|_s) .$$

*Induction:* (3.5.15) *holds for* $s > s_0$. We proceed by induction. Given some integer $k \ge s_0$, we assume that (3.5.15) holds for any $s \in [0, k]$, and for any function $f \in C^\infty$. We are going to prove (3.5.15) for any $s \in ]k, k+1]$. Recall that

$$\|u\|_s^2 = \sum_{k \in \mathbb{Z}^b} |u_k|^2 \langle k \rangle^{2s} \sim_s \|u\|_{L^2}^2 + \max_{i=1,\ldots,b} \|\partial_{y_i} u\|_{s-1}^2 . \tag{3.5.18}$$

Then

$$\begin{aligned} \|\mathtt{f}(u)\|_s & \overset{(3.5.18)}{\sim_s} \|f(y, u(y))\|_{L^2} + \max_{i=1,\ldots,b} \|\partial_{y_i}(f(y, u(y)))\|_{s-1} \\ & \lesssim_{s,f} 1 + \max_{i=1,\ldots,b} \big( \|(\partial_{y_i} f)(y, u(y))\|_{s-1} + \|(\partial_u f)(y, u(y))(\partial_{y_i} u)(y)\|_{s-1} \big) . \end{aligned} \tag{3.5.19}$$

By the inductive assumption,

$$\|(\partial_{y_i} f)(y, u(y))\|_{s-1} \le C(f)(1 + \|u\|_{s-1}) . \tag{3.5.20}$$

To estimate $\|(\partial_u f)(y, u(y))(\partial_{y_i} u)(y)\|_{s-1}$, we distinguish two cases:

- 1st case: $k = s_0$. Thus $s \in (s_0, s_0+1]$. Since $s-1 \in (s_0-1, s_0]$ we apply Lemma 3.5.2 obtaining

$$\|(\partial_u f)(y,u(y))(\partial_{y_i} u)(y)\|_{s-1} \lesssim_{s_0} \|(\partial_u f)(y,u(y))\|_{s_0} \|\partial_{y_i} u\|_{s-1} \lesssim_f \|u\|_s \,. \tag{3.5.21}$$

- 2nd case: $k \geq s_0+1$. For any $s \in (k, k+1]$ we have $s-1 \geq s_0$, and, by Lemma 3.5.1 and the inductive assumption, we obtain

$$\|(\partial_u f)(y,u(y))(\partial_{y_i} u)(y)\|_{s-1} \lesssim_{s,f} (1+\|u\|_{s-1})\|u\|_{s_0+1} + \|u\|_s \,.$$

By the interpolation inequality $\|u\|_{s-1}\|u\|_{s_0+1} \leq \|u\|_s \|u\|_{s_0}$ (Lemma 3.5.4), and $s \geq s_0+1$, we conclude that

$$\|(\partial_u f)(y,u(y))(\partial_{y_i} u)(y)\|_{s-1} \lesssim_{s,f} \|u\|_s \,. \tag{3.5.22}$$

Finally, by (3.5.19), (3.5.20), (3.5.21), (3.5.22), the estimate (3.5.15) holds for all $s \in ]k, k+1]$. This concludes the iteration and the proof of (3.5.15).

STEP 2. PROOF OF (3.5.13). In order to prove the Lipschitz estimate we write

$$\begin{aligned} \mathtt{f}(v)(y) - \mathtt{f}(u)(y) &= f(y, v(y)) - f(y, u(y)) \\ &= \int_0^1 (\partial_u f)(y, u(y) + \tau(v-u)(y))(v-u)(y) d\tau \,. \end{aligned} \tag{3.5.23}$$

Then

$$\begin{aligned} \|\mathtt{f}(v) - \mathtt{f}(u)\|_s &\leq \int_0^1 \|(\partial_u f)(y, u + \tau(v-u))(v-u)\|_s d\tau \qquad (3.5.24) \\ &\overset{(3.5.1)}{\lesssim_s} \int_0^1 \|(\partial_u f)(y, u+\tau(v-u))\|_s \|v-u\|_{s_0} + \|(\partial_u f)(y, u+\tau(v-u))\|_{s_0} \|v-u\|_s d\tau \,. \end{aligned}$$

Specializing (3.5.24) for $v = u(\lambda_2)$ and $u = u(\lambda_1)$, using (3.5.15) and $\|u\|_{\mathrm{Lip}, s_0} \leq 1$, we deduce

$$\|\mathtt{f}(u(\lambda_2)) - \mathtt{f}(u(\lambda_1))\|_s \lesssim_{s,f} \|u\|_{\mathrm{Lip},s} |\lambda_2 - \lambda_1| \,, \quad \forall \lambda_1, \lambda_2 \in \Lambda \,. \tag{3.5.25}$$

The estimates (3.5.15) and (3.5.25) imply (3.5.13).

STEP 3. PROOF OF (3.5.14). By (3.5.23),

$$\begin{aligned} \|\mathtt{f}(v) - \mathtt{f}(u)\|_{\mathrm{Lip},s} &\leq \int_0^1 \|(\partial_u f)(y, u+\tau(v-u))(v-u)\|_{\mathrm{Lip},s} d\tau \\ &\overset{(3.5.1)}{\lesssim_s} \int_0^1 \|(\partial_u f)(y, u+\tau(v-u))\|_{\mathrm{Lip},s} \|v-u\|_{\mathrm{Lip},s_0} \\ &\qquad + \|(\partial_u f)(y, u+\tau(v-u))\|_{\mathrm{Lip},s_0} \|v-u\|_{\mathrm{Lip},s} \, d\tau \end{aligned}$$

and, using (3.5.13) and $\|u\|_{\mathrm{Lip},s_0}, \|v\|_{\mathrm{Lip},s_0} \leq 1$, we deduce (3.5.14) for $p=1$.

Estimates (3.5.13) and (3.5.14) for $p \geq 2$ can be obtained exactly in the same way. ■

# Chapter 4

# Multiscale Analysis

The main result of this Chapter is the abstract multiscale Proposition 4.1.5, which provides invertibility properties of finite dimensional restrictions of the Hamiltonian operator $\mathcal{L}_{r,\mu}$ defined in (4.1.9) for a large set of parameters $\lambda \in \Lambda$. This multiscale Proposition 4.1.5 will be used in Chapters 9 and 10.

## 4.1 Multiscale proposition

Let $H := L^2(\mathbb{T}^d, \mathbb{R}) \times L^2(\mathbb{T}^d, \mathbb{R})$ and consider the Hilbert spaces

$$(i)\ \mathbf{H} := H\,, \qquad (ii)\ \mathbf{H} := H \times H\,. \tag{4.1.1}$$

Any vector $h \in \mathbf{H}$ can be written as

$$(i)\ h = (h_1, h_2)\,, \qquad (ii)\ h = (h_1, h_2, h_3, h_4)\,, \qquad h_i \in L^2(\mathbb{T}^d, \mathbb{R})\,.$$

On $\mathbf{H}$, we define the linear operator $J$ as

$$J(h_1, h_2) := (h_2, -h_1) \qquad \text{in case}\,(i) \tag{4.1.2}$$

$$J(h_1, h_2, h_3, h_4) := (h_2, -h_1, h_4, -h_3) \qquad \text{in case}\,(ii)\,. \tag{4.1.3}$$

Moreover, in case $(ii)$, we also define the right action of $J$ on $\mathbf{H}$ as

$$hJ = (h_1, h_2, h_3, h_4)J := (-h_3, -h_4, h_1, h_2)\,. \tag{4.1.4}$$

**Remark 4.1.1.** *The motivation for defining* (4.1.3), (4.1.4) *is the following. Identifying* $a \in \mathcal{L}(H_j, H)$ *with* $(a^{(1)}, a^{(2)}, a^{(3)}, a^{(4)}) \in H \times H$ *as in* (3.2.6)-(3.2.7)*, the operators* $Ja, aJ \in \mathcal{L}(H_j, H)$*, where* $J$ *is the symplectic operator in* (3.1.10)*, i.e.* (4.1.2)*, are identified with the vectors in* $H \times H$ *in* (3.2.13)*,* (3.2.14)*, i.e.* (4.1.3)*,* (4.1.4)*.*

We denote by $\Pi_{\mathbb{S}\cup\mathbb{F}}$ the $L^2$-orthogonal projection on the subspace $H_{\mathbb{S}\cup\mathbb{F}}$ in $H$ defined in (3.3.31) or the analogous one in $H \times H$. Note that $\Pi_{\mathbb{S}\cup\mathbb{F}}$ commutes with the left (and right in case (4.1.4)) action of $J$. We also denote

$$\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} := \mathrm{Id} - \Pi_{\mathbb{S}\cup\mathbb{F}} = \Pi_{\mathbb{G}} ,$$

see (1.2.14). We fix a constant $\mathfrak{c} \in (0,1)$ such that

$$|\bar{\mu} \cdot \ell + \mu_j + \mathfrak{c}| \geq \frac{\gamma_0}{\langle \ell \rangle^{\tau_0}} , \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|} , \ j \in \mathbb{S} . \tag{4.1.5}$$

By standard arguments, condition (4.1.5) is fulfilled by all $\mathfrak{c} \in (0,1)$ except a set of measure $O(\gamma_0)$. We explain the purely technical role of the term $\mathfrak{c}\Pi_{\mathbb{S}}$ in (4.1.6)-(4.1.7) below, in remarks 4.1.3 and 4.6.3 .

**Definition 4.1.2.** *Given positive constants $C_1, c_2 > 0$, we define the class $\mathfrak{C}(C_1, c_2)$ of $L^2$-self-adjoint operators acting on $\mathbf{H}$, of the form, according to the cases (i)-(ii) in* (4.1.1),

$$(i) \qquad X_r := X_r(\varepsilon, \lambda, \varphi) = D_V + \mathfrak{c}\Pi_{\mathbb{S}} + r(\varepsilon, \lambda, \varphi) \tag{4.1.6}$$

$$(ii) \qquad X_{r,\mu} := X_{r,\mu}(\varepsilon, \lambda, \varphi) = D_V + \mathfrak{c}\Pi_{\mathbb{S}} + \mu(\varepsilon, \lambda)\mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} + r(\varepsilon, \lambda, \varphi) \tag{4.1.7}$$

*defined for $\lambda \in \widetilde{\Lambda} \subset \Lambda$, where $D_V := \sqrt{-\Delta + V(x)}$, $\mu(\varepsilon, \lambda) \in \mathbb{R}$, $\mathcal{J}$ is the self-adjoint operator*

$$\mathcal{J} : H \times H \to H \times H , \qquad h \mapsto \mathcal{J}h := JhJ , \tag{4.1.8}$$

*(recall* (4.1.3)-(4.1.4)*) and such that*

1. *$|r|_{\mathrm{Lip},+,s_1} \leq C_1 \varepsilon^2$, for some $s_1 > s_0$,*
2. *$|\mu(\varepsilon, \lambda) - \mu_k|_{\mathrm{Lip}} \leq C_1 \varepsilon^2$ for some $k \in \mathbb{F}$ (set defined in* (1.2.15)*),*
3. *$\mathfrak{d}_\lambda \Big( \frac{X_{r,\mu}}{1 + \varepsilon^2 \lambda} \Big) \leq -c_2 \varepsilon^2 \mathrm{Id}$, see the notation* (1.6.4).

*We assume that the non-resonance conditions* (1.2.7), (1.2.16)-(1.2.17), (4.1.5) *hold.*

*For simplicity of notation, in the sequel we shall denote by $X_{r,\mu}$ also the operator $X_r$ in* (4.1.6), *understanding that $X_r = X_{r,\mu}$ does not depend on $\mu$, i.e. $\mathcal{J} = 0$.*

Notice that the operator $\mathcal{J}$ defined in (4.1.8) and $\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} = \Pi_{\mathbb{G}}$ commute.

**Remark 4.1.3.** *The form of the operators $X_r$, $X_{r,\mu}$ in* (4.1.6), (4.1.7) *is motivated by the application of the multiscale Proposition 4.1.5 to the operator $\mathcal{L}_r$ in* (10.2.30) *acting on $H$, and the operator $\mathcal{L}_{r,\mu}$ in* (9.3.44) *acting on $H \times H$. We add the term $\mathfrak{c}\Pi_{\mathbb{S}}$ in* (4.1.6), (4.1.7) *as a purely technical trick to prove Lemma 4.6.2, see remark 4.6.3.*

In the next proposition we prove invertibility properties of finite dimensional restrictions of the operator

$$\mathcal{L}_{r,\mu} := J\omega \cdot \partial_\varphi + X_{r,\mu}(\varepsilon, \lambda)\,, \quad \omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon\,, \tag{4.1.9}$$

for a large set of $\lambda \in \Lambda$. For $N \in \mathbb{N}$ we define the subspace of trigonometric polynomials

$$\begin{aligned} &\mathcal{H}_N := \Big\{ u(\varphi, x) = \sum_{|(\ell,j)|\le N} u_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)}\,,\ u_{\ell,j} \in \mathbb{C}^r \Big\} \\ &\text{where} \quad r := \begin{cases} 2 \text{ in case (4.1.1)-}(i) \\ 4 \text{ in case (4.1.1)-}(ii) \end{cases}, \end{aligned} \tag{4.1.10}$$

and we denote by $\varPi_N$ the corresponding $L^2$-projector:

$$u(\varphi, x) = \sum_{(\ell,j)\in\mathbb{Z}^{|\mathbb{S}|}\times\mathbb{Z}^d} u_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)} \quad \mapsto \quad \varPi_N u := \sum_{|(\ell,j)|\le N} u_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)}\,. \tag{4.1.11}$$

The projectors $\varPi_N$ satisfy the usual smoothing estimates in Sobolev spaces: for any $s, \beta \ge 0$,

$$\|\varPi_N u\|_{s+\beta} \le N^\beta \|u\|_s\,, \quad \|\varPi_N^\perp u\|_s \le N^{-\beta}\|u\|_{s+\beta} \tag{4.1.12}$$

$$\|\varPi_N u\|_{\mathrm{Lip},s+\beta} \le N^\beta \|u\|_{\mathrm{Lip},s}\,, \quad \|\varPi_N^\perp u\|_{\mathrm{Lip},s} \le N^{-\beta}\|u\|_{\mathrm{Lip},s+\beta}\,. \tag{4.1.13}$$

We shall require that $\omega$ satisfies the following quadratic Diophantine non-resonance condition.

**Definition 4.1.4.** $(\mathbf{NR})_{\gamma,\tau}$ *Given $\gamma \in (0,1)$, $\tau > 0$, a vector $\omega \in \mathbb{R}^{|\mathbb{S}|}$ is $(\mathbf{NR})_{\gamma,\tau}$ non-resonant, if, for any non zero polynomial $P(X) \in \mathbb{Z}[X_1, \dots, X_{|\mathbb{S}|}]$ of the form*

$$P(X) = n + \sum_{1\le i\le j\le |\mathbb{S}|} p_{ij} X_i X_j\,, \quad n, p_{ij} \in \mathbb{Z}\,, \tag{4.1.14}$$

*we have*

$$|P(\omega)| \ge \gamma \langle p\rangle^{-\tau}\,, \quad \langle p\rangle := \max_{i,j=1,\dots,|\mathbb{S}|}\{1, |p_{ij}|\}\,. \tag{4.1.15}$$

The main result of this section is the following proposition. Set

$$\varsigma := 1/10\,. \tag{4.1.16}$$

For simplicity of notation, in the next proposition $\mathcal{L}_{r,\mu}$ also denotes $J\omega\cdot\partial_\varphi + X_r$, understanding that $X_r = X_{r,\mu}$ in (4.1.6) does not depend on $\mu$.

**Proposition 4.1.5. (Multiscale)** *Let $\bar\omega_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ be $(\gamma_1, \tau_1)$-Diophantine and satisfy property* $(\mathbf{NR})_{\gamma_1,\tau_1}$ *in Definition 4.1.4 with $\gamma_1, \tau_1$ defined in (1.2.28). Then there are $\varepsilon_0 > 0$, $\tau' > 0$, $\bar s_1 > s_0$, $\bar N \in \mathbb{N}$ (not depending on $X_{r,\mu}$ but possibly on the constants $C_1$, $c_2$, $\gamma_0$, $\tau_0$, $\gamma_1$, $\tau_1$) and $\tau_0' > 0$ (depending only on $\tau_0$) such that the following holds:*

*assume $s_1 \geq \bar s_1$ and take an operator $X_{r,\mu}$ in $\mathfrak{C}(C_1, c_2)$ as in Definition 4.1.2, which is defined for all $\lambda \in \widetilde\Lambda$. Then for any $\varepsilon \in (0, \varepsilon_0)$, there are $N(\varepsilon) \in \mathbb{N}$, closed subsets $\Lambda(\varepsilon; \eta, X_{r,\mu}) \subset \widetilde\Lambda$, $\eta \in [1/2, 1]$, satisfying*

1. *$\Lambda(\varepsilon; \eta, X_{r,\mu}) \subseteq \Lambda(\varepsilon; \eta', X_{r,\mu})$, for all $1/2 \leq \eta \leq \eta' \leq 1$;*
2. *the complementary set $\Lambda(\varepsilon; 1/2, X_{r,\mu})^c := \Lambda \setminus \Lambda(\varepsilon; 1/2, X_{r,\mu})$ satisfies*
$$|\Lambda(\varepsilon; 1/2, X_{r,\mu})^c \cap \widetilde\Lambda| \lesssim \varepsilon\,; \tag{4.1.17}$$
3. *if $|r' - r|_{+,s_1} + |\mu' - \mu| \leq \delta \leq \varepsilon^2$, then, for $(1/2) + \sqrt{\delta} \leq \eta \leq 1$,*
$$|\Lambda(\varepsilon; \eta, X_{r',\mu'})^c \cap \Lambda(\varepsilon; \eta - \sqrt{\delta}, X_{r,\mu}) \cap \widetilde\Lambda'| \lesssim \delta^\alpha\,; \tag{4.1.18}$$

*such that,*

1. *$\forall \bar N \leq N < N(\varepsilon)$, $\lambda \in \widetilde\Lambda$, the operator*
$$[\mathcal{L}_{r,\mu}]_N^{2N} := \Pi_N (\mathcal{L}_{r,\mu})_{|\mathcal{H}_{2N}} \tag{4.1.19}$$
*has a right inverse $\big([\mathcal{L}_{r,\mu}]_N^{2N}\big)^{-1} : \mathcal{H}_N \to \mathcal{H}_{2N}$ satisfying, for all $s \geq s_0$,*
$$\Big|\Big(\frac{[\mathcal{L}_{r,\mu}]_N^{2N}}{1 + \varepsilon^2\lambda}\Big)^{-1}\Big|_{\mathrm{Lip},s} \leq C(s) N^{\tau_0'+1}\big(N^{\varsigma s} + |r|_{\mathrm{Lip},+,s}\big)\,. \tag{4.1.20}$$
*Moreover, for all $\lambda \in \widetilde\Lambda \cap \widetilde\Lambda'$, we have*
$$\big|\big([\mathcal{L}_{r,\mu}]_N^{2N}\big)^{-1} - \big([\mathcal{L}_{r',\mu'}]_N^{2N}\big)^{-1}\big|_{s_1} \lesssim_{s_1} N^{2(\tau_0'+\varsigma s_1)+1}\big(|\mu - \mu'|N^2 + |r - r'|_{+,s_1}\big) \tag{4.1.21}$$
*where $\mu$, $\mu'$, $r$, $r'$ are evaluated at fixed $\lambda$.*
2. *$\forall N \geq N(\varepsilon)$, $\lambda \in \Lambda(\varepsilon; 1, X_{r,\mu})$, the operator*
$$\mathcal{L}_{r,\mu,N} := \Pi_N\big(J\omega \cdot \partial_\varphi + X_{r,\mu}(\varepsilon, \lambda)\big)_{|\mathcal{H}_N}\,, \quad \omega = (1 + \varepsilon^2\lambda)\bar\omega_\varepsilon\,, \tag{4.1.22}$$
*is invertible and, for all $s \geq s_0$,*
$$\Big|\Big(\frac{\mathcal{L}_{r,\mu,N}}{1 + \varepsilon^2\lambda}\Big)^{-1}\Big|_{\mathrm{Lip},s} \leq C(s) N^{2(\tau'+\varsigma s_1)+3}\big(N^{\varsigma(s-s_1)} + |r|_{\mathrm{Lip},+,s}\big)\,. \tag{4.1.23}$$
*Moreover for all $\lambda \in \widetilde\Lambda \cap \widetilde\Lambda'$,*
$$\big|\mathcal{L}_{r,\mu,N}^{-1} - \mathcal{L}_{r',\mu',N}^{-1}\big|_{s_1} \leq C(s_1) N^{2(\tau'+\varsigma s_1)+1}\big(|\mu - \mu'|N^2 + |r - r'|_{+,s_1}\big) \tag{4.1.24}$$
*where $\mu$, $\mu'$, $r$, $r'$ are evaluated at fixed $\lambda$.*

**Remark 4.1.6.** *The measure of the set* $\Lambda(\varepsilon; 1/2, X_{r,\mu})^c$ *is smaller than* $\varepsilon^p$*, for any* $p$*, at the expense of taking a larger constant* $\tau'$*, see remark 4.8.17. We have written* $\varepsilon$ *in* (4.1.17) *for definiteness.*

**Remark 4.1.7.** *Properties 1-3 for the sets* $\Lambda(\varepsilon; \eta, X_{r,\mu})$ *are stable under finite intersection: if*

$$(\Lambda^{(1)}(\varepsilon; \eta, X_{r,\mu})), \ldots, (\Lambda^{(p)}(\varepsilon; \eta, X_{r,\mu}))$$

*are families of closed subsets of* $\Lambda$ *satisfying 1-3, then the family* $(\bigcap_{1\le k\le p} \Lambda^{(k)}(\varepsilon; \eta, X_{r,\mu}))$ *still satisfies these properties.*

The proof of Proposition 4.1.5 is based on the multiscale analysis of the papers [22], [23], [26] for quasi-periodically forced nonlinear wave and Schrödinger equations, but it is more complicated in the present autonomous setting. The proof of the multiscale Proposition 4.1.5 is given in the next sections 4.2-4.8.

## 4.2 Matrix representation

We decompose the operator $\mathcal{L}_{r,\mu}$ in (4.1.9), with $X_{r,\mu}$ as in (4.1.7) or (4.1.6) (in such a case we mean that $\mathcal{J} = 0$), as

$$\begin{aligned} &\mathcal{L}_{r,\mu} = \mathcal{D}_\omega + \mathcal{T}\,, \\ &\mathcal{D}_\omega := J\omega\cdot\partial_\varphi + D_m + \mu\mathcal{J}\,, \\ &\mathcal{T} := D_V - D_m - \mu\mathcal{J}\Pi_{\mathbb{S}\cup\mathbb{F}} + r + \mathfrak{c}\Pi_{\mathbb{S}} \end{aligned} \tag{4.2.1}$$

where $D_V := \sqrt{-\Delta + V(x)}$ is defined in (2.1.11), $D_m := \sqrt{-\Delta + m}$ in (3.3.18), and $\omega = (1+\varepsilon^2\lambda)\bar\omega_\varepsilon$. By Proposition 3.4.1 the matrix which represents the operator $D_V - D_m$ in the exponential basis has off-diagonal decay.

In what follows we identify a linear operator $A(\varphi)$, $\varphi \in \mathbb{T}^{|\mathbb{S}|}$, acting on functions $h(\varphi, x)$, with the infinite dimensional matrix $(A^{\ell,j}_{\ell',j'})_{\{(\ell,j),(\ell',j')\in\mathbb{Z}^b\}}$ of $2\times 2$ matrices $A^{\ell,j}_{\ell',j'}$ in case (4.1.1)-($i$), respectively $4\times 4$ in case (4.1.1)-($ii$), defined by the relation (3.2.25). In this way, the operator $\mathcal{L}_{r,\mu}$ in (4.2.1) is represented by the infinite dimensional Hermitian matrix

$$\mathtt{A}(\varepsilon,\lambda) := \mathtt{A}(\varepsilon,\lambda; r) := \mathtt{D}_\omega + \mathtt{T}\,, \tag{4.2.2}$$

of $2\times 2$ matrices in case ($i$), resp. $4\times 4$ matrices in case ($ii$), where the diagonal part is, in case ($i$),

$$\begin{aligned} &\mathtt{D}_\omega = \mathrm{Diag}_{i\in\mathbb{Z}^b}\begin{pmatrix} \langle j\rangle_m & \mathrm{i}\omega\cdot\ell \\ -\mathrm{i}\omega\cdot\ell & \langle j\rangle_m \end{pmatrix}, \\ &i := (\ell, j) \in \mathbb{Z}^b := \mathbb{Z}^{|\mathbb{S}|}\times\mathbb{Z}^d\,, \quad \langle j\rangle_m := \sqrt{|j|^2 + m}\,. \end{aligned} \tag{4.2.3}$$

In case $(ii)$, recalling the definition of $\mathcal{J}$ in (4.1.8), of the left and right action of $J$ in (4.1.3)-(4.1.4), and choosing the basis of $\mathbb{C}^4$

$$\{f_1, f_2, f_3, f_4\} := \{e_3 - e_2, e_1 + e_4, e_1 - e_4, e_2 + e_3\}\,,$$
$$e_{\mathfrak{a}} := (0, \ldots, \underbrace{1}_{\mathfrak{a}-th}, \ldots, 0) \in \mathbb{C}^4\,,$$

we have

$$\mathtt{D}_\omega = \mathrm{Diag}_{(\ell,j)\in\mathbb{Z}^b} \begin{pmatrix} \langle j\rangle_m - \mu & \mathrm{i}\omega\cdot\ell & 0 & 0 \\ -\mathrm{i}\omega\cdot\ell & \langle j\rangle_m - \mu & 0 & 0 \\ 0 & 0 & \langle j\rangle_m + \mu & \mathrm{i}\omega\cdot\ell \\ 0 & 0 & -\mathrm{i}\omega\cdot\ell & \langle j\rangle_m + \mu \end{pmatrix}. \tag{4.2.4}$$

The off-diagonal matrix

$$\mathtt{T} := (\mathtt{T}_i^{i'})_{i\in\mathbb{Z}^b, i'\in\mathbb{Z}^b}\,, \quad b = |\mathbb{S}| + d\,, \quad \mathtt{T}_i^{i'} := (D_V - D_m)_j^{j'} - \mu[\mathcal{J}\Pi_{\mathbb{S}\cup\mathbb{F}}]_j^{j'} + r_i^{i'}\,, \tag{4.2.5}$$

where $\mathtt{T}_i^{i'}$ are $2\times 2$, resp. $4\times 4$ matrices, satisfies, by (3.4.1), Lemma 3.3.8, property 1 of Definition 4.1.2,

$$|\mathtt{T}|_{+,s_1} \le C(s_1)\,. \tag{4.2.6}$$

Note that $(\mathtt{T}_i^{i'})^* = \mathtt{T}_{i'}^{i}$. Moreover, since the operator $\mathcal{T} = \mathcal{T}(\varphi)$ in (4.2.1) is a $\varphi$-dependent family of operators acting on $\mathbf{H}$ (as $r$ defined in (4.1.6)-(4.1.7) which is the only $\varphi$-dependent operator), the matrix $\mathtt{T}$ is *Töplitz* in $\ell$, namely $\mathtt{T}_i^{i'} = \mathtt{T}_{\ell,j}^{\ell',j'}$ depends only on the indices $\ell - \ell', j, j'$. We introduce a further index $\mathfrak{a} \in \mathfrak{I}$ where

$$\mathfrak{I} := \{1, 2\} \text{ in case (4.1.1)-}(i)\,, \qquad \mathfrak{I} := \{1, 2, 3, 4\} \text{ in case (4.1.1)-}(ii)\,, \tag{4.2.7}$$

to distinguish the matrix elements of each $2\times 2$, resp. $4\times 4$, matrix

$$\mathtt{T}_i^{i'} := \big(\mathtt{T}_{i,\mathfrak{a}}^{i',\mathfrak{a}'}\big)_{\mathfrak{a},\mathfrak{a}'\in\mathfrak{I}}\,.$$

Under the unitary change of variable (basis of eigenvectors)

$$U := \mathrm{Diag}_{(\ell,j)\in\mathbb{Z}^b} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix} \tag{4.2.8}$$

the matrix $\mathtt{D}_\omega$ in (4.2.3) becomes completely diagonal

$$D_\omega := U^{-1}\mathtt{D}_\omega U = \mathrm{Diag}_{(\ell,j)\in\mathbb{Z}^b} \begin{pmatrix} \langle j\rangle_m - \omega\cdot\ell & 0 \\ 0 & \langle j\rangle_m + \omega\cdot\ell \end{pmatrix} \tag{4.2.9}$$

and, under the unitary transformation

$$U := \mathrm{Diag}_{(\ell,j)\in\mathbb{Z}^b} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ \mathrm{i} & -\mathrm{i} & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & \mathrm{i} & -\mathrm{i} \end{pmatrix}, \tag{4.2.10}$$

the matrix in (4.2.4) becomes completely diagonal

$$D_\omega := U^{-1}\mathtt{D}_\omega U = \mathrm{Diag}_{(\ell,j)\in\mathbb{Z}^b} \begin{pmatrix} \langle j\rangle_m - \mu - \omega\cdot\ell & 0 & 0 & 0 \\ 0 & \langle j\rangle_m - \mu + \omega\cdot\ell & 0 & 0 \\ 0 & 0 & \langle j\rangle_m + \mu - \omega\cdot\ell & 0 \\ 0 & 0 & 0 & \langle j\rangle_m + \mu + \omega\cdot\ell \end{pmatrix}. \tag{4.2.11}$$

Under the unitary change of variable $U$ in (4.2.8) in case (4.1.1)-($i$), (4.2.10) in case (4.1.1)-($ii$), the hermitian matrix $\mathtt{A}$ in (4.2.2) transforms in the hermitian matrix

$$A(\varepsilon,\lambda) := A := U^{-1}\mathtt{A}U = D_\omega + T\,, \qquad T_i^{i'} := U^{-1}\mathtt{T}_i^{i'}U\,, \tag{4.2.12}$$

where the off-diagonal term $T$ satisfies, by (4.2.6),

$$|T|_{+,s_1} \le C(s_1)\,. \tag{4.2.13}$$

We introduce the one-parameter family of infinite dimensional matrices

$$A(\varepsilon,\lambda,\theta) := A(\varepsilon,\lambda) + \theta Y := D_\omega + \theta\, Y + T\,, \quad \theta\in\mathbb{R}\,, \tag{4.2.14}$$

where

$$\begin{aligned} Y &:= \mathrm{Diag}_{i\in\mathbb{Z}^b}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \text{ in case } (4.1.1)-(i)\,, \\ Y &:= \mathrm{Diag}_{i\in\mathbb{Z}^b}\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ in case } (4.1.1)-(ii)\,. \end{aligned} \tag{4.2.15}$$

The reason for adding $\theta Y$ is that, translating the time Fourier indices

$$(\ell,j) \mapsto (\ell+\ell_0, j)$$

in $A(\varepsilon,\lambda)$, gives $A(\varepsilon,\lambda,\theta)$ with $\theta = \omega\cdot\ell_0$. Note that the matrix $T$ remains unchanged under translation because it is Töplitz with respect to $\ell$.

**Remark 4.2.1.** *The matrix $A(\varepsilon,\lambda,\theta) := A(\varepsilon,\lambda,\theta;r)$ in (4.2.14) represents the $L^2$-self-adjoint operator*

$$\mathcal{L}_{r,\mu}(\theta) := J\omega\cdot\partial_\varphi + \mathrm{i}\theta J + D_V + \mu\mathcal{J}\Pi^\perp_{\mathbb{S}\cup\mathbb{F}} + \mathfrak{c}\Pi_{\mathbb{S}} + r\,. \tag{4.2.16}$$

*In section 4.7 we shall denote by $\breve{A}(\varepsilon,\lambda,\theta) := A(\varepsilon,\lambda,\theta;0)$ the matrix which represents*

$$\mathcal{L}_{0,\mu}(\theta) := J\omega\cdot\partial_\varphi + \mathrm{i}\theta J + D_V + \mu\mathcal{J}\Pi^\perp_{\mathbb{S}\cup\mathbb{F}} + \mathfrak{c}\Pi_{\mathbb{S}}\,. \tag{4.2.17}$$

*Notice that $\mathcal{L}_{0,\mu}(\theta)$ is independent of $\varphi$, thus $\breve{A}(\varepsilon,\lambda,\theta)$ is diagonal in $\ell\in\mathbb{Z}^{|\mathbb{S}|}$.*

The eigenvalues of the $2 \times 2$ matrix $D_\omega + \theta Y$, with $D_\omega$ in (4.2.9) and $Y$ defined in (4.2.15)-case $(i)$, resp. $4 \times 4$ matrix with $D_\omega$ in (4.2.11) and $Y$ defined in (4.2.15)-case $(ii)$, are

$$d_{i,\mathfrak{a}}(\theta) = \begin{cases} \langle j \rangle_m - (\omega \cdot \ell + \theta) & \text{if } \mathfrak{a} = 1 \\ \langle j \rangle_m + (\omega \cdot \ell + \theta) & \text{if } \mathfrak{a} = 2 \,, \end{cases} \tag{4.2.18}$$

and, in the second case,

$$d_{i,\mathfrak{a}}(\theta) = \begin{cases} \langle j \rangle_m - \mu - (\omega \cdot \ell + \theta) & \text{if } \mathfrak{a} = 1 \\ \langle j \rangle_m - \mu + (\omega \cdot \ell + \theta) & \text{if } \mathfrak{a} = 2 \\ \langle j \rangle_m + \mu - (\omega \cdot \ell + \theta) & \text{if } \mathfrak{a} = 3 \\ \langle j \rangle_m + \mu + (\omega \cdot \ell + \theta) & \text{if } \mathfrak{a} = 4 \,. \end{cases} \tag{4.2.19}$$

The main goal of the following sections is to prove polynomial off-diagonal decay for the inverse of the $|\mathfrak{I}|(2N+1)^b$-dimensional (where $|\mathfrak{I}| = 2$, resp. 4, in case (4.1.1)-$(i)$, resp. (4.1.1)-$(ii)$) sub-matrices of $A(\varepsilon, \lambda, \theta)$ centered at $(\ell_0, j_0)$ denoted by

$$A_{N,\ell_0,j_0}(\varepsilon, \lambda, \theta) := A_{|\ell-\ell_0| \le N, |j-j_0| \le N}(\varepsilon, \lambda, \theta) \tag{4.2.20}$$

where

$$|\ell| := \max\{|\ell_1|, \ldots, |\ell_{|\mathbb{S}|}|\} \,, \quad |j| := \max\{|j_1|, \ldots, |j_d|\} \,. \tag{4.2.21}$$

If $\ell_0 = 0$ we use the simpler notation

$$A_{N,j_0}(\varepsilon, \lambda, \theta) := A_{N,0,j_0}(\varepsilon, \lambda, \theta) \,. \tag{4.2.22}$$

If also $j_0 = 0$, we simply write

$$A_N(\varepsilon, \lambda, \theta) := A_{N,0}(\varepsilon, \lambda, \theta) \,, \tag{4.2.23}$$

and, for $\theta = 0$, we denote

$$A_{N,j_0}(\varepsilon, \lambda) := A_{N,j_0}(\varepsilon, \lambda, 0) \,, \quad A_N(\varepsilon, \lambda) := A_{N,0}(\varepsilon, \lambda, 0) \,. \tag{4.2.24}$$

**Remark 4.2.2.** *The matrix $A_N(\varepsilon, \lambda)$ in* (4.2.24) *represents the truncated self-adjoint operator*

$$\Pi_N(\mathcal{L}_{r,\mu})_{|\mathcal{H}_N} = \Pi_N(J\omega \cdot \partial_\varphi + X_{r,\mu})_{|\mathcal{H}_N} = \Pi_N(J\omega \cdot \partial_\varphi + D_V + \mu \mathcal{J} \Pi^\perp_{\mathbb{S} \cup \mathbb{F}} + r)_{|\mathcal{H}_N}$$

*where $\mathcal{H}_N$ is defined in* (4.1.10) *and $\Pi_N$ in* (4.1.11).

We have the following crucial *covariance* property

$$A_{N,\ell_1,j_1}(\varepsilon, \lambda, \theta) = A_{N,j_1}(\varepsilon, \lambda, \theta + \omega \cdot \ell_1) \,. \tag{4.2.25}$$

## 4.3 Multiscale step

The main result of this section is the multiscale step Proposition 4.3.4 which is a variant of that proved in [23]. The constant $\varsigma \in (0,1)$ is fixed and $\tau' > 0$, $\Theta \geq 1$ are real parameters, on which we shall impose some condition in Proposition 4.3.4.

Given $\Omega, \Omega' \subset E \subset \mathbb{Z}^b \times \mathfrak{I}$, where $\mathfrak{I}$ is defined in (4.2.7), we define

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

where, for $k = (i, \mathfrak{a})$, $k' := (i', \mathfrak{a}') \in \mathbb{Z}^b \times \mathfrak{I}$, we set

$$|k - k'| := \begin{cases} 1 & \text{if } i = i' \,, \mathfrak{a} \neq \mathfrak{a}' \,, \\ 0 & \text{if } i = i' \,, \mathfrak{a} = \mathfrak{a}' \,, \\ |i - i'| & \text{if } i \neq i' \,. \end{cases}$$

**Notation:** Given a matrix $A \in \mathcal{M}_E^E$, when writing the matrix $D_m^{1/2} A D_m^{1/2} \in \mathcal{M}_E^E$, we understand that we apply the diagonal matrix $D_m^{1/2} : E \to E$ to the right/left of $A$.

**Definition 4.3.1. ($N$-good/bad matrix)** *The matrix $A \in \mathcal{M}_E^E$, with $E \subset \mathbb{Z}^b \times \mathfrak{I}$, $\mathrm{diam}(E) \leq 4N$, is $N$-good if $A$ is invertible and*

$$\forall s \in [s_0, s_1] \,, \quad |D_m^{-1/2} A^{-1} D_m^{-1/2}|_s \leq N^{\tau' + \varsigma s} \,. \tag{4.3.1}$$

*Otherwise $A$ is $N$-bad.*

The above definition is different with respect to that of [23]: the matrix $A$ is $N$-good according to Definition 4.3.1 if and only if $D_m^{1/2} A D_m^{1/2}$ is $N$-good according to Definition B.2.1.

**Definition 4.3.2. (Regular/Singular sites)** *The index $k := (i, \mathfrak{a}) = (\ell, j, \mathfrak{a}) \in \mathbb{Z}^b \times \mathfrak{I}$ (where $\mathfrak{I}$ is defined in (4.2.7)) is* REGULAR *for $A$ if*

$$|A_k^k| \geq \Theta \langle j \rangle^{-1} \,.$$

*Otherwise $k$ is* SINGULAR.

Also the above definition is different with respect to that of [23]: the index $k$ is regular for $A$ according to Definition 4.3.2 if and only if $k$ is regular for $D_m^{1/2} A D_m^{1/2}$ according to Definition B.2.2 with $\Theta$ replaced by $c(m)\Theta$ (because of the equivalence $(|j|^2 + m)^{1/2} \sim_m \langle j \rangle$).

The constant $\Theta := \Theta(V)$ will be chosen large enough depending on the potential $V(x)$ in order to apply the multiscale proposition (as in [23], [22]).

**Definition 4.3.3. ($(A, N)$-good/bad site)** *For $A \in \mathcal{M}_E^E$, we say that $k \in E \subset \mathbb{Z}^b \times \mathfrak{I}$ is*

- $(A, N)$-REGULAR *if there is* $F \subset E$ *such that* $\mathrm{diam}(F) \le 4N$, $\mathrm{d}(k, E \backslash F) \ge N$ *and* $A_F^F$ *is* $N$*-good.*
- $(A, N)$-GOOD *if it is regular for* $A$ *or* $(A, N)$*-regular. Otherwise we say that* $k$ *is* $(A, N)$-BAD.

Note that a site $k$ is $(A, N)$-GOOD according to Definition 4.3.3 if and only if $k$ is $(D_m^{1/2} A D_m^{1/2}, N)$-GOOD according to [23].

Let us consider the new larger scale

$$N' = N^{\chi} \tag{4.3.2}$$

with $\chi > 1$.

For a matrix $A \in \mathcal{M}_E^E$ we define $\mathrm{Diag}(A) := (\delta_{kk'} A_k^{k'})_{k,k' \in E}$.

**Proposition 4.3.4. (Multiscale step)** *Assume*

$$\varsigma \in (0, 1/2)\,, \;\; \tau' > 2\tau + b + 1\,, \;\; C_1 \ge 2\,, \tag{4.3.3}$$

*and, setting* $\kappa := \tau' + b + s_0$,

$$\chi(\tau' - 2\tau - b) > 3(\kappa + (s_0 + b)C_1)\,, \tag{4.3.4}$$

$$\chi\varsigma > C_1\,, \tag{4.3.5}$$

$$s_1 > 3\kappa + \chi(\tau + b) + C_1 s_0\,. \tag{4.3.6}$$

*For any given* $\Upsilon > 0$*, there exist* $\Theta := \Theta(\Upsilon, s_1) > 0$ *large enough (appearing in Definition 4.3.2), and* $N_0(\Upsilon, \Theta, s_1) \in \mathbb{N}$ *such that:*
$\forall N \ge N_0(\Upsilon, \Theta, s_1)$, $\forall E \subset \mathbb{Z}^b \times \mathfrak{I}$ *with* $\mathrm{diam}(E) \le 4N' = 4N^{\chi}$ *(see (4.3.2)), if* $A \in \mathcal{M}_E^E$ *satisfies*

- **(H1) (Off-diagonal decay)** $|A - \mathrm{Diag}(A)|_{+,s_1} \le \Upsilon$
- **(H2) ($L^2$-bound)** $\|D_m^{-1/2} A^{-1} D_m^{-1/2}\|_0 \le (N')^{\tau}$
- **(H3) (Separation properties)** *There is a partition of the* $(A, N)$*-bad sites* $B = \cup_\alpha \Omega_\alpha$ *with*

$$\mathrm{diam}(\Omega_\alpha) \le N^{C_1}\,, \quad \mathrm{d}(\Omega_\alpha, \Omega_\beta) \ge N^2\,, \;\; \forall \alpha \ne \beta\,, \tag{4.3.7}$$

*then* $A$ *is* $N'$*-good. More precisely*

$$\forall s \in [s_0, s_1]\,, \quad |D_m^{-1/2} A^{-1} D_m^{-1/2}|_s \le \frac{1}{4}(N')^{\tau'} \big((N')^{\varsigma s} + |A - \mathrm{Diag}(A)|_{+,s}\big)\,, \tag{4.3.8}$$

*and, for all* $s \ge s_1$,

$$|D_m^{-1/2} A^{-1} D_m^{-1/2}|_s \le C(s)(N')^{\tau'} \big((N')^{\varsigma s} + |A - \mathrm{Diag}(A)|_{+,s}\big)\,. \tag{4.3.9}$$

**Remark 4.3.5.** *The main difference with respect to the multiscale Proposition 4.1 in [23] is that, since the Definition 4.3.2 of regular sites is weaker than that in [23], we require the stronger assumption (H1) concerning the off-diagonal decay of A in* $| \ |_{+,s_1}$ *norm defined in* (3.3.17)*, while in [23] we only require the off-diagonal decay of A in* $| \ |_{s_1}$ *norm. Another difference is that we prove* (4.3.8) *(with the constant* $1/4$*) for* $s \in [s_0, s_1]$*, and not in a larger interval* $[s_0, S]$ *for some* $S \geq s_1$*. For larger* $s \geq s_1$ *we prove* (4.3.9) *with* $C(s)$*.*

Proof of Proposition 4.3.4. The multiscale step Proposition 4.3.4 follows by Proposition 4.1 in [23], that we report in the Appendix B.2, see Proposition B.2.4. Set

$$\mathcal{T} := A - \mathrm{Diag}(A)\,, \qquad |\mathcal{T}|_{+,s_1} \stackrel{(H1)}{\leq} \Upsilon\,, \tag{4.3.10}$$

and consider the matrix

$$A_+ := D_m^{1/2} A D_m^{1/2} = \mathrm{Diag}(A_+) + T \tag{4.3.11}$$

where

$$\mathrm{Diag}(A_+) := D_m^{1/2}\mathrm{Diag}(A) D_m^{1/2}\,, \quad T := D_m^{1/2} \mathcal{T} D_m^{1/2}\,. \tag{4.3.12}$$

We apply the multiscale Proposition B.2.4 to the matrix $A_+$. By (H2) the matrix $A_+$ is invertible and

$$\|A_+^{-1}\|_0 \stackrel{(4.3.11)}{=} \|D_m^{-1/2} A^{-1} D_m^{-1/2}\|_0 \stackrel{(H2)}{\leq} (N')^{\tau}\,.$$

Moreover, the decay norm

$$|T|_{s_1} \stackrel{(4.3.12)}{=} |D_m^{1/2} \mathcal{T} D_m^{1/2}|_{s_1} \stackrel{(3.3.17)}{=} |\mathcal{T}|_{+,s_1} \stackrel{(4.3.10)}{\leq} \Upsilon\,.$$

Finally the $(A_+, N)$-bad sites according to Definition B.2.3 coincide with the $(A, N)$-bad sites according to Definition 4.3.3 (with a $\Theta$ replaced by $c(m)\Theta$). Hence by (H3) also the separation properties required to apply Proposition B.2.4 hold, and we deduce that

$$\forall s \in [s_0, s_1]\,, \quad |A_+^{-1}|_s \leq \frac{1}{4}(N')^{\tau'}\big((N')^{\varsigma s} + |A_+ - \mathrm{Diag}(A_+)|_s\big)\,,$$

that, recalling (4.3.11), (4.3.10), (4.3.12), implies (4.3.8). The more general estimate (4.3.9) follows by (B.2.8). ■

## 4.4 Separation properties of bad sites

The aim of this section is to verify the separation properties of the bad sites required in the multiscale step Proposition 4.3.4.

Let $A := A(\varepsilon, \lambda, \theta)$ be the infinite dimensional matrix defined in (4.2.14). Given $N \in \mathbb{N}$ and $i = (\ell_0, j_0)$, recall that the submatrix $A_{N,i}$ is defined in (4.2.20).

**Definition 4.4.1. ($N$-regular/singular site)** *A site $k := (i, \mathfrak{a}) \in \mathbb{Z}^b \times \mathfrak{I}$ is:*

- $N$-REGULAR *if $A_{N,i}$ is $N$-good (Definition 4.3.1).*
- $N$-SINGULAR *if $A_{N,i}$ is $N$-bad (Definition 4.3.1).*

We also define the $N$-good/bad sites of $A$.

**Definition 4.4.2. ($N$-good/bad site)** *A site $k := (i, \mathfrak{a}) \in \mathbb{Z}^b \times \mathfrak{I}$ is:*

- $N$-GOOD *if*

$$k \, is \, regular \, (Def. \, 4.3.2) \text{ or } all \, the \, sites \, k' \, with \, \mathrm{d}(k', k) \leq N \, are \, N-regular \, . \quad (4.4.1)$$

- $N$-BAD *if*

$$k \, is \, singular \, (Def. \, 4.3.2) \text{ and } \exists \, k' \, with \, \mathrm{d}(k', k) \leq N \, , \, k' \text{ is } N-singular \, . \quad (4.4.2)$$

**Remark 4.4.3.** *A site $k$ which is $N$-good according to Definition 4.4.2, is $(A_E^E, N)$-good according to Definition 4.3.3, for any set $E = E_0 \times \mathfrak{I}$ containing $k$ where $E_0 \subset \mathbb{Z}^b$ is a product of intervals of length $\geq N$.*

Let

$$B_N(j_0; \lambda) := \Big\{ \theta \in \mathbb{R} \, : \, A_{N,j_0}(\varepsilon, \lambda, \theta) \text{ is } N - bad \Big\} \, . \quad (4.4.3)$$

**Definition 4.4.4. ($N$-good/bad parameters)** *A parameter $\lambda \in \Lambda$ is $N$-good for $A$ if*

$$\forall \, j_0 \in \mathbb{Z}^d \, , \quad B_N(j_0; \lambda) \subset \bigcup_{q=1,\ldots,N^{\alpha-d-|\mathbb{S}|}} I_q \, , \quad \alpha := 3d + 2|\mathbb{S}| + 4 + 3\tau_0 \, , \quad (4.4.4)$$

*where $I_q$ are intervals with measure $|I_q| \leq N^{-\tau}$. Otherwise, we say that $\lambda$ is $N$-bad. We define the set of $N$-good parameters*

$$\mathcal{G}_N := \Big\{ \lambda \in \widetilde{\Lambda} \, : \, \lambda \;\; \text{is} \;\; N - \text{good for } A(\varepsilon, \lambda) \Big\} \, . \quad (4.4.5)$$

The main result of this section is Proposition 4.4.5 which enables to verify the assumption (H3) of Proposition 4.3.4 for the submatrices $A_{N',j_0}(\varepsilon, \lambda, \theta)$.

**Proposition 4.4.5. (Separation properties of $N$-bad sites)** *Let $\tau_1, \gamma_1, \tau_2, \gamma_2$ be fixed as in* (1.2.28) *and* (4.5.9)*, depending on the parameters $\tau_0, \gamma_0$ which appear in properties* (1.2.6)-(1.2.8), (1.2.16)-(1.2.19)*. Then there exist $C_1 := C_1(d, |\mathbb{S}|, \tau_0) \geq 2$, $\tau^\star := \tau^\star(d, |\mathbb{S}|, \tau_0)$, and $\bar{N} := \bar{N}(|\mathbb{S}|, d, \gamma_0, \tau_0, m, \Theta)$ such that, if $N \geq \bar{N}$ and*

- **(i)** *$\lambda$ is $N$-good for $A$,*

- **(ii)** $\tau \geq \tau^\star$,
- **(iii)** $\bar{\omega}_\varepsilon$ *satisfies* (1.2.29) *and* $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$ *satisfies* $(\mathbf{NR})_{\gamma_2,\tau_2}$ *(see Definition 4.1.4),*

*then,* $\forall \theta \in \mathbb{R}$*, the* $N$*-bad sites* $k = (\ell, j, \mathfrak{a}) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{Z}^d \times \mathfrak{I}$ *of* $A(\varepsilon, \lambda, \theta)$ *admit a partition* $\cup_\alpha \Omega_\alpha$ *in disjoint clusters satisfying*

$$\mathrm{diam}(\Omega_\alpha) \leq N^{C_1(d,|\mathbb{S}|,\tau_0)}, \quad \mathrm{d}(\Omega_\alpha, \Omega_\beta) > N^2, \ \forall \alpha \neq \beta. \tag{4.4.6}$$

We underline that the estimates (4.4.6) are *uniform* in $\theta$.

**Remark 4.4.6.** *Hypothesis (ii) in Proposition 4.4.5 just requires that the constant* $\tau$ *is larger than some* $\tau^\star(d, |\mathbb{S}|, \tau_0)$*. This is important in the present autonomous setting for the choice of the constants in section 4.5. On the contrary in the corresponding propositions in the papers [22], [23], the constant* $\tau$ *was required to be large with the exponent* $\chi$ *in* (4.3.2).

The rest of this section is devoted to the proof of Proposition 4.4.5. In some parts of the proof, we may point out the dependence of some constants on parameters such as $\tau_2, \tau_1$, which, by (1.2.28) and (4.5.9), amounts to a dependence on $\tau_0$.

**Definition 4.4.7. (Γ-chain)** *A sequence* $k_0, \ldots, k_L \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{Z}^d \times \mathfrak{I}$ *of distinct integer vectors satisfying*

$$|k_{q+1} - k_q| \leq \Gamma, \quad \forall q = 0, \ldots, L-1,$$

*for some* $\Gamma \geq 2$*, is called a* $\Gamma$*-chain of length* $L$*.*

We want to prove an upper bound for the length of a $\Gamma$-chain of $\theta$-singular integer vectors, i.e. satisfying (4.4.8) below. In the next lemma we obtain the upper bound (4.4.10) under the assumption (4.4.9). Define the functions of signs

$$\begin{aligned} &\sigma_1, \sigma_2 : \mathfrak{I} \to \{-1, +1\} \\ &\sigma_1(1) := \sigma_1(2) := \sigma_2(1) := \sigma_2(3) := -1, \\ &\sigma_1(3) := \sigma_1(4) := \sigma_2(2) := \sigma_2(4) := 1. \end{aligned} \tag{4.4.7}$$

**Lemma 4.4.8. (Length of Γ-chain of θ-singular sites)** *Assume that* $\omega$ *satisfies the non-resonance condition* $(\mathbf{NR})_{\gamma_2,\tau_2}$ *(Definition 4.1.4). For* $\Gamma \geq \bar{\Gamma}(d, m, \Theta, \gamma_2, \tau_2)$ *large enough, consider a* $\Gamma$*-chain* $(\ell_q, j_q, \mathfrak{a}_q)_{q=0,\ldots,L} \subset \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{Z}^d \times \mathfrak{I}$ *of singular sites for the matrix* $A(\varepsilon, \lambda, \theta)$*, namely*

$$\forall q = 0, \ldots, L, \quad \big|\langle j_q \rangle_m \big(\langle j_q \rangle_m + \sigma_1(\mathfrak{a}_q)\mu + \sigma_2(\mathfrak{a}_q)(\omega \cdot \ell_q + \theta)\big)\big| < \Theta, \tag{4.4.8}$$

*with* $\sigma_1, \sigma_2$ *defined in* (4.4.7)*, such that,* $\forall \tilde{\jmath} \in \mathbb{Z}^d$*, the cardinality*

$$|\{(\ell_q, j_q, \mathfrak{a}_q)_{q=0,\ldots,L} \ : \ j_q = \tilde{\jmath}\}| \leq K. \tag{4.4.9}$$

*Then there is* $C_2 := C_2(d, \tau_2) > 0$ *such that its length is bounded by*

$$L \leq (\Gamma K)^{C_2} . \tag{4.4.10}$$

*Moreover, if* $\ell$ *is fixed (i.e.* $\ell_q = \ell$, $\forall q = 0, \dots, L$*) the same result holds without assuming that* $\omega$ *satisfies* $(\mathbf{NR})_{\gamma_2,\tau_2}$.

The proof of Lemma 4.4.8 is a variant of Lemma 4.2 in [22]. We split it in several steps.

First note that it is sufficient to bound the length of a $\Gamma$-chain of singular sites when $\theta = 0$. Indeed, suppose first that $\theta = \omega \cdot \bar{\ell}$ for some $\bar{\ell} \in \mathbb{Z}^{|\mathbb{S}|}$. For a $\Gamma$-chain of $\theta$-singular sites $(\ell_q, j_q, \mathfrak{a}_q)_{q=0,\dots,L}$, see (4.4.8), the translated $\Gamma$-chain $(\ell_q + \bar{\ell}, j_q, \mathfrak{a}_q)_{q=0,\dots,L}$, is formed by 0-singular sites, namely

$$\big|\langle j_q \rangle_m \big(\langle j_q \rangle_m + \sigma_1(\mathfrak{a}_q)\mu + \sigma_2(\mathfrak{a}_q)(\omega \cdot (\ell_q + \bar{\ell}))\big)\big| < \Theta .$$

For any $\theta \in \mathbb{R}$, we consider an approximating sequence $\omega \cdot \bar{\ell}_n \to \theta$, $\bar{\ell}_n \in \mathbb{Z}^{|\mathbb{S}|}$. Indeed, by Assumption (1.2.29), $\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon$ is not colinear to an integer vector, and therefore the set $\{\omega \cdot \ell \ , \ \ell \in \mathbb{Z}^{|\mathbb{S}|}\}$ is dense in $\mathbb{R}$. A $\Gamma$-chain of $\theta$-singular sites (see (4.4.8)), is, for $n$ large enough, also a $\Gamma$-chain of $\omega \cdot \bar{\ell}_n$-singular sites. Then we bound its length arguing as in the above case.

We first prove Lemma 4.4.8 in a particular case.

**Lemma 4.4.9.** *Assume that* $\omega$ *satisfies* $(\mathbf{NR})_{\gamma_2,\tau_2}$ *(Definition 4.1.4). Let* $(\ell_q, j_q, \mathfrak{a}_q)_{q=0,\dots,L}$ *be a* $\Gamma$*-chain of integer vectors of* $\mathbb{Z}^{|\mathbb{S}|} \times \mathbb{Z}^d \times \mathfrak{I}$ *satisfying,* $\forall q = 0, \dots, L$,

$$\Big|\sqrt{|j_q|^2 + m} + \sigma_2(\mathfrak{a}_q)\,\omega \cdot \ell_q\Big| \leq \frac{\Theta}{\langle j_q \rangle} , \qquad \textit{in case (i) of Def.}\ 4.1.2 , \tag{4.4.11}$$

$$\Big|\sqrt{|j_q|^2 + m} + \sigma_1(\mathfrak{a}_q)\mu + \sigma_2(\mathfrak{a}_q)\,\omega \cdot \ell_q\Big| \leq \frac{\Theta}{\langle j_q \rangle} , \quad \textit{in case (ii) of Def.}\ 4.1.2 , \tag{4.4.12}$$

*where* $\sigma_1, \sigma_2$ *are defined in* (4.4.7)*. Suppose, in case* (4.4.12)*, that the product of the signs* $\sigma_1(\mathfrak{a}_q)\sigma_2(\mathfrak{a}_q)$ *is the same for any* $q \in [\![0, L]\!]$*. Then, for some constant* $C_1 := C_1(d, \tau_2)$ *and* $C := C(m, \Theta, \gamma_2, d, \tau_2)$*, its length* $L$ *is bounded by*

$$L \leq C(\Gamma K)^{C_1} \tag{4.4.13}$$

*where* $K$ *is defined in* (4.4.9)*.*

*Moreover, if* $\ell$ *is fixed (i.e.* $\ell_q = \ell$, $\forall q$*), the lemma holds without assuming that* $\omega$ *satisfies* $(\mathbf{NR})_{\gamma_2,\tau_2}$.

Proof. We make the proof when (4.4.12) holds, since (4.4.11) is a particular case of (4.4.12) setting $\mu = 0$ (notice that, in the case (4.4.11), the conclusion of the lemma follows without conditions on the signs $\sigma_2(\mathfrak{a}_q)$, see remark 4.4.11).

We introduce the quadratic form $Q : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ defined by

$$Q(x, y) := -x^2 + |y|^2 \tag{4.4.14}$$

and the associated bilinear symmetric form $\Phi : (\mathbb{R} \times \mathbb{R}^d)^2 \to \mathbb{R}$ defined by

$$\Phi\big((x, y), (x', y')\big) := -xx' + y \cdot y' . \tag{4.4.15}$$

Note that $\Phi$ is the sum of the bilinear forms

$$\begin{gathered} \Phi = -\Phi_1 + \Phi_2 \\ \Phi_1\big((x, y), (x', y')\big) := xx' , \quad \Phi_2\big((x, y), (x', y')\big) := y \cdot y' . \end{gathered} \tag{4.4.16}$$

**Lemma 4.4.10.** *For all $q, q_0 \in [\![0, L]\!]$,*

$$\big|\Phi\big((x_{q_0}, j_{q_0}), (x_q - x_{q_0}, j_q - j_{q_0})\big)\big| \leq C\Gamma^2 |q - q_0|^2 + C(\Theta) , \tag{4.4.17}$$

*where $x_q := \omega \cdot \ell_q$.*

Proof. Set for brevity $\sigma_{1,q} := \sigma_1(\mathfrak{a}_q)$ and $\sigma_{2,q} := \sigma_2(\mathfrak{a}_q)$. First note that by (4.4.12) we have

$$\big||j_q|^2 + m - (\sigma_{1,q}\mu + \sigma_{2,q}\omega \cdot \ell_q)^2\big| \leq C\Theta .$$

Therefore

$$\big| - (\omega \cdot \ell_q)^2 + |j_q|^2 - 2\sigma_{1,q}\, \sigma_{2,q}\, \mu\, \omega \cdot \ell_q\big| \leq C' , \quad C' := C\Theta + m + \mu^2 ,$$

and so, recalling (4.4.14), for all $q = 0, \dots, L$,

$$|Q(x_q, j_q) - 2\sigma_{1,q}\sigma_{2,q}\, \mu\, x_q| \leq C' \qquad \text{where} \qquad x_q := \omega \cdot \ell_q . \tag{4.4.18}$$

By the hypothesis of Lemma 4.4.9,

$$\sigma_{1,q}\sigma_{2,q} = \sigma_{1,q_0}\sigma_{2,q_0} , \quad \forall q, q_0 \in [\![0, L]\!] ,$$

and, by bilinearity, we get

$$\begin{aligned} Q(x_q, j_q) - 2\sigma_{1,q}\sigma_{2,q}\mu x_q = {} & Q(x_{q_0}, j_{q_0}) \\ & + 2\Phi\big((x_{q_0}, j_{q_0}), (x_q - x_{q_0}, j_q - j_{q_0})\big) + Q(x_q - x_{q_0}, j_q - j_{q_0}) \\ & - 2\sigma_{1,q_0}\sigma_{2,q_0}\mu x_{q_0} - 2\sigma_{1,q}\sigma_{2,q}\mu(x_q - x_{q_0}) . \end{aligned} \tag{4.4.19}$$

Recalling (4.4.14) and the Definition 4.4.7 of $\Gamma$-chain we have, $\forall q, q_0 \in [\![0, L]\!]$,

$$\big|Q(x_q - x_{q_0}, j_q - j_{q_0}) - 2\sigma_{1,q}\sigma_{2,q}\mu(x_q - x_{q_0})\big| \leq C\Gamma^2 |q - q_0|^2 . \tag{4.4.20}$$

Hence (4.4.19), (4.4.18), (4.4.20) imply (4.4.17). ∎

**Remark 4.4.11.** *In the case* (4.4.11)*, the conclusion of Lemma 4.4.10 follows without conditions on the signs $\sigma_2(\mathfrak{a}_q)$. This is why Lemma 4.4.9 holds without condition on the signs $\sigma_2(\mathfrak{a}_q)$.*

**Proof of Lemma 4.4.9 continued.** We introduce the subspace of $\mathbb{R}^{d+1}$

$$\begin{aligned} G &:= \mathrm{Span}_{\mathbb{R}}\Big\{(x_q - x_{q'}, j_q - j_{q'}) \,:\, 0 \le q, q' \le L\Big\} \\ &= \mathrm{Span}_{\mathbb{R}}\Big\{(x_q - x_{q_0}, j_q - j_{q_0}) \,:\, 0 \le q \le L\Big\} \end{aligned} \tag{4.4.21}$$

and we call $g \le d+1$ the dimension of $G$. Introducing a small parameter $\delta > 0$, to be specified later (see (4.4.38)), we distinguish two cases.

**Case I**. $\forall q_0 \in [\![0, L]\!]$,

$$\mathrm{Span}_{\mathbb{R}}\big\{(x_q - x_{q_0}, j_q - j_{q_0}) \,:\, |q - q_0| \le L^\delta\,,\ q \in [\![0, L]\!]\,\big\} = G\,. \tag{4.4.22}$$

We select a basis of $G \subset \mathbb{R}^{d+1}$ from $(x_q - x_{q_0}, j_q - j_{q_0})$ with $|q - q_0| \le L^\delta$, say

$$f_s := (x_{q_s} - x_{q_0}, j_{q_s} - j_{q_0}) = (\omega \cdot \Delta_s \ell, \Delta_s j)\,, \quad s = 1, \ldots, g\,, \tag{4.4.23}$$

where $(\Delta_s \ell, \Delta_s j) := (\ell_{q_s} - \ell_{q_0}, j_{q_s} - j_{q_0})$ satisfies, by the Definition 4.4.7 of $\Gamma$-chain,

$$|(\Delta_s \ell, \Delta_s j)| \le C\Gamma |q_s - q_0| \le C\Gamma L^\delta\,. \tag{4.4.24}$$

Hence

$$|f_s| \le C\,\Gamma L^\delta\,, \qquad \forall s = 1, \ldots, g\,. \tag{4.4.25}$$

Then, in order to derive from (4.4.17) a bound on $(x_{q_0}, j_{q_0})$ or its projection onto $G$, we need a nondegeneracy property for $Q_{|G}$. The following lemma states it.

**Lemma 4.4.12.** *Assume that $\omega$ satisfies* $(\mathbf{NR})_{\gamma_2,\tau_2}$*. Then the matrix*

$$\Omega := (\Omega_s^{s'})_{s,s'=1}^g\,, \quad \Omega_s^{s'} := \Phi(f_{s'}, f_s)\,, \tag{4.4.26}$$

*is invertible and*

$$|(\Omega^{-1})_s^{s'}| \le C(\Gamma L^\delta)^{C_3(d,\tau_2)}\,, \quad \forall s, s' = 1, \ldots, g\,, \tag{4.4.27}$$

*where the multiplicative constant $C$ depends on $\gamma_2$.*

Proof. According to the splitting (4.4.16) we write $\Omega$ as

$$\Omega := \big(-\Phi_1(f_{s'}, f_s) + \Phi_2(f_{s'}, f_s)\big)_{s,s'=1,\ldots,g} = -S + R \tag{4.4.28}$$

where, by (4.4.23),

$$\begin{aligned} S_s^{s'} &:= \Phi_1(f_{s'}, f_s) = (\omega \cdot \Delta_{s'} \ell)(\omega \cdot \Delta_s \ell)\,, \\ R_s^{s'} &:= \Phi_2(f_{s'}, f_s) = \Delta_{s'} j \cdot \Delta_s j\,. \end{aligned} \tag{4.4.29}$$

The matrix $R = (R_1, \dots, R_g)$ has integer entries (the $R_i \in \mathbb{Z}^g$ denote the columns). The matrix $S := (S_1, \dots, S_g)$ has rank 1 since all its columns $S_s \in \mathbb{R}^g$ are colinear:

$$S_s = (\omega \cdot \Delta_s \ell)(\omega \cdot \Delta_1 \ell, \dots, \omega \cdot \Delta_g \ell)^\top , \quad s = 1, \dots g \, . \tag{4.4.30}$$

We develop the determinant

$$\begin{aligned} P(\omega) \; &:= \; \det \Omega \overset{(4.4.28)}{=} \det(-S + R) \\ &= \; \det(R) - \det(S_1, R_2, \dots, R_g) - \dots - \det(R_1, \dots, R_{g-1}, S_g) \\ &= \; \det(R) - \sum_{1 \le s \le g} (-1)^{g+s} S_s \cdot (R_1 \wedge \dots \wedge R_{s-1} \wedge R_{s+1} \wedge \dots \wedge R_g) \end{aligned} \tag{4.4.31}$$

using that the determinant of matrices with 2 columns $S_i$, $S_j$, $i \neq j$, is zero. By (4.4.30), the expression in (4.4.31) is a polynomial in $\omega$ of degree 2 of the form (4.1.14) with coefficients

$$|(n,p)| \overset{(4.4.29),(4.4.24)}{\le} C(\Gamma L^\delta)^{2s} \le C(\Gamma L^\delta)^{2(d+1)} \, . \tag{4.4.32}$$

If $P \neq 0$ then the non-resonance condition $(\mathbf{NR})_{\gamma_2,\tau_2}$ implies

$$|\det \Omega| = |P(\omega)| \overset{(4.1.15)}{\ge} \frac{\gamma_2}{\langle p \rangle^{\tau_2}} \overset{(4.4.32)}{\ge} \frac{\gamma_2}{C(\Gamma L^\delta)^{2\tau_2(d+1)}} \, . \tag{4.4.33}$$

In order to conclude the proof of the lemma, we have to show that $P(\omega)$ is not identically zero in $\omega$. We have that

$$P(\mathrm{i}\omega) = \det\big(\Phi_1(f_{s'}, f_s) + \Phi_2(f_{s'}, f_s)\big)_{s,s'=1,\dots g} = \det(f_{s'} \cdot f_s)_{s,s'=1,\dots g} > 0$$

because $(f_s)_{1 \le s \le g}$ is a basis of $G$. Thus $P$ is not the zero polynomial.

By (4.4.33), the Cramer rule, and (4.4.25) we deduce (4.4.27). ∎

**Remark 4.4.13.** *As recently proved in [29] the same result holds also assuming just that $\omega$ is Diophantine, instead of the quadratic non-resonance condition* $(\mathbf{NR})_{\gamma_2,\tau_2}$.

**Proof of Lemma 4.4.9 continued.** We introduce

$$G^{\perp \Phi} := \Big\{ z \in \mathbb{R}^{d+1} \; : \; \Phi(z, f) = 0 \, , \; \forall f \in G \Big\} \, .$$

Since $\Omega$ is invertible (Lemma 4.4.12), $\Phi_{|G}$ is nondegenerate, hence

$$\mathbb{R}^{d+1} = G \oplus G^{\perp \Phi}$$

and we denote by $P_G : \mathbb{R}^{d+1} \to G$ the corresponding projector onto $G$.

We are going to estimate

$$P_G(x_{q_0}, j_{q_0}) = \sum_{s'=1}^{g} a_{s'} f_{s'} \,. \tag{4.4.34}$$

For all $s = 1, \ldots, g$, and since $f_s \in G$, we have

$$\Phi\big((x_{q_0}, j_{q_0}), f_s\big) = \Phi\big(P_G(x_{q_0}, j_{q_0}), f_s\big) \overset{(4.4.34)}{=} \sum_{s'=1}^{g} a_{s'} \Phi(f_{s'}, f_s)$$

that we write as the linear system

$$\Omega a = b \,, \qquad a := \begin{pmatrix} a_1 \\ \vdots \\ a_g \end{pmatrix} , \quad b := \begin{pmatrix} \Phi\Big((x_{q_0}, j_{q_0}), f_1\Big) \\ \vdots \\ \Phi\Big((x_{q_0}, j_{q_0}), f_g\Big) \end{pmatrix} \tag{4.4.35}$$

with $\Omega$ defined in (4.4.26).

**Lemma 4.4.14.** *For all $q_0 \in [\![0, L]\!]$ we have*

$$|P_G(x_{q_0}, j_{q_0})| \leq C(\Gamma L^\delta)^{C_4(d,\tau_2)} \,, \tag{4.4.36}$$

*where $C$ depends on $\gamma_2, \Theta$.*

Proof. We have

$$|b| \overset{(4.4.35)}{\lesssim} \max_{1 \leq s \leq g} \Big|\Phi\Big((x_{q_0}, j_{q_0}), f_s\Big)\Big| \overset{(4.4.23),(4.4.17)}{\leq} C(\Theta)\Gamma^2 \max_{1\leq s \leq g} |q_s - q_0|^2 \leq C(\Theta)(\Gamma L^\delta)^2 \,,$$

recalling that, by (4.4.22), the indices $(q_s)_{1\leq s\leq g}$ were selected such that $|q_s - q_0| \leq L^\delta$. Hence, by (4.4.35) and (4.4.27),

$$|a| = |\Omega^{-1} b| \leq C(\gamma_2, \Theta)(\Gamma L^\delta)^C \tag{4.4.37}$$

for some constant $C := C(d, \tau_2)$. We deduce (4.4.36) by (4.4.34), (4.4.37) and (4.4.25). ■

We now complete the proof of Lemma 4.4.9 when case I holds. As a consequence of Lemma 4.4.14, for all $q_1, q_2 \in [\![0, L]\!]$, we have

$$|(x_{q_1}, j_{q_1}) - (x_{q_2}, j_{q_2})| = \big|P_G\big((x_{q_1}, j_{q_1}) - (x_{q_2}, j_{q_2})\big)\big| \leq C(\Gamma L^\delta)^{C_4(d,\tau_2)}$$

where $C$ depends on $\gamma_2, \Theta$. Therefore, for all $q_1, q_2 \in [\![0, L]\!]$, we have $|j_{q_1} - j_{q_2}| \leq C(\Gamma L^\delta)^{C_4(d,\tau_2)}$, and so

$$\operatorname{diam}\{j_q \ ; \ 0 \leq q \leq L\} \leq C(\Gamma L^\delta)^{C_4(d,\tau_2)} \,.$$

Since all the $j_q$ are in $\mathbb{Z}^d$, their number (counted without multiplicity) does not exceed $C(\Gamma L^\delta)^{C_4(d,\tau_2)d}$, for some other constant $C$ which depends on $\gamma_2, \tau_2, d$. Thus we have obtained the bound

$$\sharp\{j_q \ : \ 0 \leq q \leq L\} \leq C(\Gamma L^\delta)^{C_4(d,\tau_2)d} .$$

By assumption (4.4.9), for each $q_0 \in [\![0, L]\!]$, the number of $q \in [\![0, L]\!]$ such that $j_q = j_{q_0}$ is at most $K$, and so

$$L \leq C(\Gamma L^\delta)^{C_5(d,\tau_2)} K \,, \qquad C_5(d, \tau_2) := C_4(d, \tau_2)d \,.$$

Choosing $\delta > 0$ such that

$$\delta C_5(d, \tau_2) < 1/2 \,, \tag{4.4.38}$$

we get $L \leq C(\Gamma^{C_5(d,\tau_2)} K)^2$, for some multiplicative constant $C$ that may depend on $\gamma_2, \Theta, d$. This proves (4.4.13).

**Case II.** There is $q_0 \in [\![0, L]\!]$ such that

$$\dim \mathrm{Span}_{\mathbb{R}} \big\{ (x_q - x_{q_0}, j_q - j_{q_0}) \ : \ |q - q_0| \leq L^\delta \,, \ q \in [\![0, L]\!] \,\big\} \leq g - 1 \,,$$

namely all the vectors $(x_q, j_q)$ stay in a affine subspace of dimension less than $g - 1$. Then we repeat on the sub-chain $(\ell_q, j_q)$, $|q - q_0| \leq L^\delta$, the argument of case I, to obtain a bound for $L^\delta$ (and hence for $L$).

Applying at most $(d + 1)$-times the above procedure, we obtain a bound for $L$ of the form $L \leq C(\Gamma K)^{C(d,\tau_2)}$. This concludes the proof of (4.4.13) of Lemma 4.4.9.

To prove the last statement of Lemma 4.4.9, notice that, if $\ell$ is fixed, then (4.4.17) reduces just to $|j_q \cdot (j_q - j_{q_0})| \leq C\Gamma^2 |q - q_0|^2$ and the conclusion of the lemma follows as above, see also Lemma 5.2 in [23] (actually it is the same argument for NLS in [37]). ■

**Proof of Lemma 4.4.8.** Consider a general $\Gamma$-chain $(k_q)_{q=0,\ldots,L} = (\ell_q, j_q, \mathfrak{a}_q)_{q=0,\ldots,L}$ of singular sites satisfying (4.4.12). To fix ideas assume $\sigma_1(\mathfrak{a}_0)\sigma_2(\mathfrak{a}_0) > 0$. Then the integer vectors $k_q$ along the chain with the same sign $\sigma_1(\mathfrak{a}_q)\sigma_2(\mathfrak{a}_q) > 0$, say $k_{q_m}$, satisfy

$$q_{m+1} - q_m \leq C(\Gamma K)^{C_1} \,,$$

by Lemma 4.4.9 applied to each subchain of consecutive indices with $\sigma_1(\mathfrak{a}_q)\sigma_2(\mathfrak{a}_q) < 0$. Hence we deduce that all such $k_{q_m}$ form a $\Gamma' := C\Gamma(\Gamma K)^{C_1}$-chain and all of them have the same sign $\sigma_1(\mathfrak{a}_{q_m})\sigma_2(\mathfrak{a}_{q_m}) > 0$. It follows, again by Lemma 4.4.9, that their length is bounded by

$$C(\Gamma' K)^{C_1} = C\big(C\Gamma(\Gamma K)^{C_1} K\big)^{C_1} = C'(\Gamma K)^{C_1(C_1+1)} \,.$$

Hence the length of the original $\Gamma$-chain $(k_q)_{q=0,\ldots,L} = (\ell_q, j_q, \mathfrak{a}_q)_{q=0,\ldots,L}$ satisfies

$$L \leq C(\Gamma K)^{C_1} C'(\Gamma K)^{C_1(C_1+1)} \leq (\Gamma K)^{C_2}$$

where $C_2 := C_2(d, \tau_2) = C_1(C_1 + 2) + 1$, and provided $\Gamma$ is large enough, depending on $m, \Theta, \gamma_2, \tau_2, d$. This proves (4.4.10).

Finally, the last statement of Lemma 4.4.8 follows as well by the last statement of Lemma 4.4.9. ■

We fix

$$\tau^\star := \tau_1(2 + C_2(3 + \alpha)) + 1 \tag{4.4.39}$$

where $\tau_1$ is the Diophantine exponent of $\bar{\omega}_\varepsilon$ in (1.2.29), $C_2$ is the constant defined in Lemma 4.4.8, and $\alpha$ is defined in (4.4.4).

The next lemma proves an upper bound for the length of a chain of $N$-bad-sites.

**Lemma 4.4.15. (Length of $\Gamma$-chain of $N$-bad sites)** *Assume (i)-(iii) of Proposition 4.4.5 with $\tau^\star$ defined in* (4.4.39)*. Then, for N large enough (depending on $m, \Theta, \gamma_2, \tau_2, \gamma_1, \tau_1$), any $N^2$-chain $(\ell_q, j_q, \mathfrak{a}_q)_{q=0,\ldots,L}$ of $N$-bad sites of $A(\varepsilon, \lambda, \theta)$ (see Definition 4.4.2) has length*

$$L \leq N^{(3+\alpha)C_2} \tag{4.4.40}$$

*where $C_2$ is defined in Lemma 4.4.8 and $\alpha$ in* (4.4.4)*.*

Proof. Arguing by contradiction we assume that

$$L > l := N^{(3+\alpha)C_2} . \tag{4.4.41}$$

We consider the subchain $(\ell_q, j_q, \mathfrak{a}_q)_{q=0,\ldots,l}$ of $N$-bad sites, which, recalling (4.4.2), are singular sites. Then, for $N$ large enough, the assumption (4.4.9) of Lemma 4.4.8 with $\Gamma = N^2$, and $L$ replaced by $l$, can not hold with $K < N^{1+\alpha}$, otherwise (4.4.10) would imply $l < (N^2 N^{1+\alpha})^{C_2} = N^{(3+\alpha)C_2}$, contradicting (4.4.41). As a consequence there exists $\tilde{\jmath} \in \mathbb{Z}^d$, and distinct indices $q_i \in [\![0, l]\!]$, $i = 0, \ldots, M := [N^{\alpha+1}/4]$, such that

$$0 \leq q_i \leq l \qquad \text{and} \qquad j_{q_i} = \tilde{\jmath}\,,\ \forall i = 0, \ldots, M := \Big[\frac{N^{\alpha+1}}{4}\Big] . \tag{4.4.42}$$

Since the sites $(\ell_{q_i}, \tilde{\jmath}, \mathfrak{a}_{q_i})_{i=0,\ldots,M}$ belong to a $N^2$-chain of length $l$, the diameter of the set $E := \{\ell_{q_i}, \mathfrak{a}_{q_i}\}_{i=0,\ldots,M} \subset \mathbb{Z}^{|\mathbb{S}|} \times \mathfrak{I}$ satisfies

$$\mathrm{diam}(E) \leq N^2 l .$$

Moreover, each site $(\ell_{q_i}, \tilde{\jmath}, \mathfrak{a}_{q_i})$ is $N$-bad, and therefore, recalling (4.4.2) it is in a $N$-neighborhood of some $N$-singular site. Let $\mathfrak{K}$ be the number of $N$-singular sites $(\ell, j, \mathfrak{a})$ such that $|j - \tilde{\jmath}| \leq N$ and $\mathtt{d}(E, (\ell, \mathfrak{a})) \leq N$. Then the cardinality $|E| \leq CN^{|\mathbb{S}|}\mathfrak{K}$. Moreover there is $\tilde{\jmath}' \in \mathbb{Z}^d$ with $|\tilde{\jmath}' - \tilde{\jmath}| \leq N$ such that there are at least $\mathfrak{K}/(CN^d)$ $N$-singular sites $(\ell, \tilde{\jmath}', \mathfrak{a})$ with distance $\mathtt{d}(E, (\ell, \mathfrak{a})) \leq N$. Let

$$E' := \Big\{(\ell', \mathfrak{a}') \in \mathbb{Z}^{|\mathbb{S}|} \times \mathfrak{I} \ : \ (\ell', \tilde{\jmath}', \mathfrak{a}') \text{ is } N\text{−singular and } \mathtt{d}(E, (\ell', \mathfrak{a}')) \leq N\Big\} .$$

By what preceeds the cardinality

$$
\begin{aligned}
&|E'| \geq \frac{|E|}{CN^{d+|\mathbb{S}|}} \geq \frac{N^{\alpha+1-d-|\mathbb{S}|}}{C} > N^{\alpha-d-|\mathbb{S}|}\,, \\
&\text{and} \quad \mathrm{diam}(E') \leq N^2 l + 2N\,.
\end{aligned} \tag{4.4.43}
$$

Recalling Definition 4.4.1 and the covariance property (4.2.25),

$$
\forall (\ell', \mathfrak{a}') \in E'\,, \quad A_{N,\ell',\tilde{\jmath}'}(\varepsilon,\lambda) = A_{N,\tilde{\jmath}'}(\varepsilon,\lambda,\omega\cdot\ell') \ \text{ is } \ N\text{−bad}\,,
$$

and, recalling (4.4.3),

$$
\forall (\ell', \mathfrak{a}') \in E'\,, \quad \omega\cdot\ell' \in B_N(\tilde{\jmath}';\lambda)\,. \tag{4.4.44}
$$

Since $\lambda$ in $N$-good (Definition 4.4.4), (4.4.4) holds and therefore

$$
B_N(\tilde{\jmath}';\lambda) \subset \bigcup_{q=0,\ldots,N^{\alpha-d-|\mathbb{S}|}} I_q \quad \text{where } I_q \text{ are intervals with measure } |I_q| \leq N^{-\tau}\,. \tag{4.4.45}
$$

By (4.4.44), (4.4.45) and since, by (4.4.43), the cardinality $|E'| \geq N^{\alpha-d-|\mathbb{S}|}$, there are two distinct integer vectors $\ell_1$, $\ell_2 \in E'$ such that $\omega\cdot\ell_1\,, \omega\cdot\ell_2$ belong to the same interval $I_q$. Therefore

$$
|\omega\cdot(\ell_1-\ell_2)| = |\omega\cdot\ell_1 - \omega\cdot\ell_2| \leq |I_q| \leq N^{-\tau}\,. \tag{4.4.46}
$$

Moreover, since by (1.2.29) the frequency vectors $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$, $\forall\lambda\in\Lambda$, are Diophantine, namely

$$
|\omega\cdot\ell| \geq \frac{\gamma_1}{2|\ell|^{\tau_1}}\,, \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|}\setminus\{0\}\,,
$$

we also deduce

$$
\begin{aligned}
|\omega\cdot(\ell_1-\ell_2)| &\geq \frac{\gamma_1}{|\ell_1-\ell_2|^{\tau_1}} \geq \frac{\gamma_1}{(\mathrm{diam}(E'))^{\tau_1}} \\
&\overset{(4.4.43)}{\geq} \frac{\gamma_1}{(N^2 l + N)^{\tau_1}} \\
&\overset{(4.4.41)}{\geq} \frac{\gamma_1}{(2N^{2+(3+\alpha)C_2})^{\tau_1}}\,.
\end{aligned} \tag{4.4.47}
$$

The conditions (4.4.46)-(4.4.47) contradict, for $N$ large enough, the assumption that $\tau \geq \tau^\star$ where $\tau^\star$ is defined in (4.4.39). ■

Proof of Proposition 4.4.5 completed. We introduce the following equivalence relation in the set

$$
\mathcal{S}_N := \Big\{ k = (\ell, j, \mathfrak{a}) \in \mathbb{Z}^{|\mathbb{S}|}\times\mathbb{Z}^d\times\mathfrak{I} \,:\, k \text{ is } N\text{- bad for } A(\varepsilon,\lambda,\theta) \Big\}.
$$

**Definition 4.4.16.** *We say that* $x \equiv y$ *if there is a* $N^2$*-chain* $\{k_q\}_{q=0,\dots,L}$ *in* $\mathcal{S}_N$ *connecting* $x$ *to* $y$*, namely* $k_0 = x$, $k_L = y$.

This equivalence relation induces a partition of the $N$-bad sites of $A(\varepsilon, \lambda, \theta)$, in disjoint equivalent classes $\cup_\alpha \Omega_\alpha$, satisfying, by Lemma 4.4.15,

$$\mathrm{d}(\Omega_\alpha, \Omega_\beta) > N^2 , \quad \mathrm{diam}(\Omega_\alpha) \leq N^2 N^{(3+\alpha)C_2} = N^{C_1} \tag{4.4.48}$$

with $C_1 := C_1(d, |\mathbb{S}|, \tau_0) := 2 + (3+\alpha)C_2$. This proves (4.4.6).

## 4.5 Definition of the sets $\Lambda(\varepsilon; \eta, X_{r,\mu})$

In order to define the sets $\Lambda(\varepsilon; \eta, X_{r,\mu}) \subset \widetilde{\Lambda}$ appearing in the statement of Proposition 4.1.5, we first fix the values of some constants.

1. **Choice of $\tau$.** First we fix $\tau$ satisfying

$$\tau > \max \left\{ \tau^\star, 9d + 8|\mathbb{S}| + 5s_0 + 5 \right\} , \tag{4.5.1}$$

   where the constant $\tau^\star$ is defined in (4.4.39). Thus (4.5.1) implies hypothesis $(ii)$ of Proposition 4.4.5 about the separation properties of the bad sites. The second condition on $\tau$ arises in the measure estimates of section 4.8 (see Lemma 4.8.16). Moreover the condition (4.5.1) on $\tau$ is also used in the proof of Proposition 4.6.1, see (4.6.47).

2. **Choice of $\bar{\chi}$.** Then we choose a constant $\bar{\chi}$ such that

$$\bar{\chi}\varsigma > C_1 , \quad \bar{\chi} > \tau + s_0 + d , \tag{4.5.2}$$

   where $\varsigma := 1/10$ is fixed as in (4.1.16) and the constant $C_1 \geq 2$ is defined in Proposition 4.4.5. The constant $\bar{\chi}$ is the exponent which enters in the definition in (4.5.11) of the scales $N_{k+1} = [N_k^{\bar{\chi}}]$ along the multiscale analysis. Notice that the first inequality in (4.5.2) is condition (4.3.5) in the multiscale step Proposition 4.3.4. The second condition on $\bar{\chi}$ arises in the measure estimates of section 4.8 (see Lemma 4.8.13).

3. **Choice of $\tau'$.** Subsequently we choose $\tau'$ large enough so that the inequalities (4.3.3)-(4.3.4) hold for all $\chi \in [\bar{\chi}, \bar{\chi}^2]$. This is used in the multiscale argument in the proof of Proposition 4.7.6. We also take

$$\tau' > \tilde{\tau}' + (|\mathbb{S}|/2) \tag{4.5.3}$$

   where $\tilde{\tau}' > \tau$ is the constant provided by Lemma 4.6.5 associated to $\tilde{\tau} = \tau$.

4. **Choice of $s_1$.** Finally we choose the Sobolev index $s_1$ large enough so that (4.3.6) holds for all $\chi \in [\bar{\chi}, \bar{\chi}^2]$.

We define the set of $L^2$-$(N,\eta)$-good/bad parameters.

**Definition 4.5.1. ($L^2$-$(N,\eta)$-good/bad parameters)** *Given $N \in \mathbb{N}$, $\eta \in (0,1]$, let*

$$\begin{aligned} B_N^0(j_0;\lambda,\eta) &:= \Big\{\theta \in \mathbb{R} \,:\, \|D_m^{-1/2} A_{N,j_0}^{-1}(\varepsilon,\lambda,\theta) D_m^{-1/2}\|_0 > \eta N^\tau \Big\} \\ &= \Big\{\theta \in \mathbb{R} \,:\, \exists \text{ an eigenvalue of } D_m^{1/2} A_{N,j_0}(\varepsilon,\lambda,\theta) D_m^{1/2} \\ &\qquad\qquad \text{with modulus less than } \eta^{-1} N^{-\tau} \Big\} \end{aligned} \tag{4.5.4}$$

*where $\| \ \|_0$ is the operatorial $L^2$-norm, and define the set of $L^2$-$(N,\eta)$-good parameters*

$$\begin{aligned} \mathcal{G}_{N,\eta}^0 := \Big\{ \lambda \in \widetilde{\Lambda} \,:\, \forall\, j_0 \in \mathbb{Z}^d\,, \quad B_N^0(j_0;\lambda,\eta) \subset \bigcup_{q=1,\dots,N^{2d+|\mathbb{S}|+4+3\tau_0}} I_q \\ \text{where } I_q \text{ are intervals with measure } |I_q| \leq N^{-\tau} \Big\}\,. \end{aligned} \tag{4.5.5}$$

*Otherwise we say that $\lambda$ is $L^2$-$(N,\eta)$-bad.*

Given $N \in \mathbb{N}$, $\eta \in (0,1]$, we also define

$$\mathtt{G}_{N,\eta}^0 := \Big\{ \lambda \in \widetilde{\Lambda} \,:\, \|D_m^{-1/2} A_N^{-1}(\varepsilon,\lambda) D_m^{-1/2}\|_0 \leq \eta N^\tau \Big\}\,. \tag{4.5.6}$$

Notice that the sets $\mathcal{G}_{N,\eta}^0$, $\mathtt{G}_{N,\eta}^0$ are increasing in $\eta$, namely

$$\eta < \eta' \quad \Rightarrow \quad \mathcal{G}_{N,\eta}^0 \subset \mathcal{G}_{N,\eta'}^0\,, \quad \mathtt{G}_{N,\eta}^0 \subset \mathtt{G}_{N,\eta'}^0\,. \tag{4.5.7}$$

We also define the set

$$\tilde{\mathcal{G}} := \Big\{ \lambda \in \Lambda \,:\, \omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon \quad \text{satisfies} \quad (\mathbf{NR})_{\gamma_2,\tau_2} \Big\} \tag{4.5.8}$$

(recall Definition 4.1.4) with

$$\gamma_2 := \frac{\gamma_1}{2} = \frac{\gamma_0}{4}\,, \quad \tau_2 := \frac{|\mathbb{S}|(|\mathbb{S}|-1)}{2} + 2(\tau_1+2)\,. \tag{4.5.9}$$

and $\gamma_1, \tau_1$ defined in (1.2.28).

Fix $N_0 := N_0(\varepsilon)$ such that

$$1 \leq \varepsilon^2 N_0^{\tau+s_0+d} \leq 2 \tag{4.5.10}$$

and define the increasing sequence of scales

$$N_k = \big[N_0^{\bar{\chi}^k}\big]\,, \quad k \geq 0\,. \tag{4.5.11}$$

**Remark 4.5.2.** *Condition* (4.5.10) *is used in Lemma 4.7.2, see* (4.7.9)*, and in the proof of Proposition 4.6.1, see* (4.6.47)*. The first inequality* $\varepsilon^{-2} \leq N_0^{\tau+s_0+d}$ *in* (4.5.10) *is also used in Lemma 4.8.6.*

Finally we define, for $\eta \in [1/2, 1]$, the sets

$$\Lambda(\varepsilon; \eta, X_{r,\mu}) := \bigcap_{k\geq 1} \mathcal{G}^0_{N_k,\eta} \bigcap_{N\geq N_0^2} \mathtt{G}^0_{N,\eta} \bigcap \tilde{\mathcal{G}} \tag{4.5.12}$$

where $\mathcal{G}^0_{N,\eta}$ is defined in (4.5.5), the set $\mathtt{G}^0_{N,\eta}$ is defined in (4.5.6), and $\tilde{\mathcal{G}}$ in (4.5.8). These are the sets $\Lambda(\varepsilon; \eta, X_{r,\mu}) \subset \widetilde{\Lambda}$ appearing in the statement of Proposition 4.1.5. By (4.5.7) these sets clearly satisfy the property 1 listed in Proposition 4.1.5. We shall prove the measure properties 2 and 3 in section 4.8.

**Remark 4.5.3.** *The second intersection in* (4.5.12) *is restricted to the indices* $N \geq N_0^2$ *for definiteness: we could have set* $N \geq N_0^{\alpha(\tau)}$ *for some exponent* $\alpha(\tau)$ *which increases linearly with* $\tau$*. Indeed, the right invertibility properties of* $\Pi_N[\mathcal{L}_{r,\mu}]_{|\mathcal{H}_{2N}}$ *at the scales* $N \leq N_0^2$ *are deduced in section 4.6 by the unperturbed Melnikov non-resonance conditions* (1.2.7), (1.2.16)-(1.2.17), (4.1.5) *and a perturbative argument which holds for* $N \leq N_0^{\alpha(\tau)}$ *with* $\alpha(\tau)$ *linear in* $\tau$*, see* (4.6.47).

## 4.6 Right inverse of $[\mathcal{L}_{r,\mu}]_N^{2N}$ for $\bar{N} \leq N < N_0^2$

The goal of this section is to prove the following proposition, which implies item 1 of Proposition 4.1.5 with $N(\varepsilon) := N_0^2$ and $N_0 = N_0(\varepsilon)$ satisfying (4.5.10).

**Proposition 4.6.1.** *There are* $\bar{N}$ *and* $\varepsilon_0 > 0$ *such that, for all* $\varepsilon \in (0, \varepsilon_0)$*, for all* $\bar{N} \leq N < N_0^2(\varepsilon)$*, the operator*

$$[\mathcal{L}_{r,\mu}]_N^{2N} = \Pi_N[\mathcal{L}_{r,\mu}]_{|\mathcal{H}_{2N}} \, ,$$

*which is defined for all* $\lambda \in \widetilde{\Lambda}$*, has a right inverse* $\big([\mathcal{L}_{r,\mu}]_N^{2N}\big)^{-1} : \mathcal{H}_N \to \mathcal{H}_{2N}$ *satisfying, for all* $s \geq s_0$*,*

$$\Big|\Big(\frac{[\mathcal{L}_{r,\mu}]_N^{2N}}{1+\varepsilon^2\lambda}\Big)^{-1}\Big|_{\mathrm{Lip},s} \leq C(s) N^{\tau_0'+1}\big(N^{\varsigma s} + |r|_{\mathrm{Lip},+,s}\big) \tag{4.6.1}$$

*where* $\tau_0'$ *is a constant depending only on* $\tau_0$*. Moreover* (4.1.21) *holds.*

The proof of Proposition 4.6.1 is given in the rest of this section.

We decompose the operator $\mathcal{L}_{r,\mu}$ in (4.1.9), with $X_{r,\mu}$ defined in (4.1.6)-(4.1.7), as

$$\mathcal{L}_{r,\mu} = \mathcal{L}_{\mathtt{d}} + \rho(\varepsilon, \lambda, \varphi) \tag{4.6.2}$$

with

$$\mathcal{L}_{\mathtt{d}} := J\bar{\mu}\cdot\partial_\varphi + D_V + \mu_k \mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} + \mathfrak{c}\Pi_{\mathbb{S}} \tag{4.6.3}$$

$$\rho(\varepsilon,\lambda,\varphi) := J(\omega-\bar{\mu})\cdot\partial_\varphi + (\mu-\mu_k)\mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} + r(\varepsilon,\lambda,\varphi)\,. \tag{4.6.4}$$

Notice that the operator $\mathcal{L}_{\mathtt{d}}$ is independent of $(\varepsilon,\lambda)$ and $\varphi$, and $\rho$ is small in $\varepsilon$ since

$$\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2(\zeta+\lambda\bar{\mu}+\varepsilon^2\zeta\lambda) = \bar{\mu}+O(\varepsilon^2)\,, \quad |\omega|_{\mathrm{lip}} = O(\varepsilon^2)\,, \tag{4.6.5}$$

where $\bar{\mu}$ is the vector defined in (1.2.3) (see (1.2.25)),

$$\mu = \mu_k + O(\varepsilon^2) \quad \text{for some } k\in\mathbb{F}\,, \quad |\mu|_{\mathrm{lip}} = O(\varepsilon^2)\,, \tag{4.6.6}$$

by item 2 of Definition 4.1.2, and $|r|_{\mathrm{Lip},+,s_1} = O(\varepsilon^2)$ by item 1 of Definition 4.1.2.

In order to prove Proposition 4.6.1 we shall first find a right inverse of $[\mathcal{L}_{\mathtt{d}}]^{2N}_N$, thanks to the non-resonance conditions (1.2.7), (1.2.16)-(1.2.17), (4.1.5), and we shall prove that it has off-diagonal decay estimates, see (4.6.43). Then we shall deduce the existence of a right inverse for $[\mathcal{L}_{r,\mu}]^{2N}_N$ by a perturbative Neumann series argument.

Notice that the space $e^{\mathrm{i}\ell\cdot\varphi}\mathbf{H}$, where $\mathbf{H}$ is defined in (4.1.1), is invariant under $\mathcal{L}_{\mathtt{d}}$ and

$$D_m^{1/2}\mathcal{L}_{\mathtt{d}}D_m^{1/2}(e^{\mathrm{i}\ell\cdot\varphi}h) = e^{\mathrm{i}\ell\cdot\varphi}(M_\ell h)\,, \quad \forall h = h(x)\in\mathbf{H}\,, \tag{4.6.7}$$

where $M_\ell$ is the operator, acting on functions $h(x)\in\mathbf{H}$ of the space variable $x$ only, defined by

$$M_\ell := D_m^{1/2}\big(\mathrm{i}\bar{\mu}\cdot\ell\, J + D_V + \mu_k\mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} + \mathfrak{c}\Pi_{\mathbb{S}}\big)D_m^{1/2}\,. \tag{4.6.8}$$

**Lemma 4.6.2. (Invertibility of $M_\ell$)** *For all $\ell\in\mathbb{Z}^{|\mathbb{S}|}$, $|\ell|<N_0^2$, the self-adjoint operator $M_\ell$ of $\mathbf{H}$ is invertible and*

$$\|M_\ell^{-1}\|_0 \lesssim \gamma_0^{-1}\langle\ell\rangle^{\tau_0}\,. \tag{4.6.9}$$

Proof. We write $M_\ell = D_m^{1/2}M'_\ell D_m^{1/2}$ where

$$M'_\ell := \mathrm{i}\bar{\mu}\cdot\ell J + D_V + \mu_k\mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} + \mathfrak{c}\Pi_{\mathbb{S}}\,, \tag{4.6.10}$$

and we remind that, in case $(i)$ of (4.1.1), $M_\ell$ acts on $H$, and we have set $\mathcal{J}=0$ (see (4.1.6)), while in case $(ii)$, $M_\ell$ acts on $H\times H$ and $\mathcal{J}(h_1,h_2,h_3,h_4) = (-h_4,h_3,h_2,-h_1)$, by (4.1.8), (4.1.3), (4.1.4).

In case $(i)$ of (4.1.1), $M'_\ell$ is represented, in the Hilbert basis $((\Psi_j,0),(0,\Psi_j))_{j\in\mathbb{N}}$ of $\mathbf{H}=H$, by the block-diagonal matrix $\mathrm{Diag}_{j\in\mathbb{N}}M'_{\ell,j}$ where

$$M'_{\ell,j} := \begin{pmatrix} \mu_j+\mathfrak{c}\delta^j_{\mathbb{S}} & \mathrm{i}\bar{\mu}\cdot\ell \\ -\mathrm{i}\bar{\mu}\cdot\ell & \mu_j+\mathfrak{c}\delta^j_{\mathbb{S}} \end{pmatrix} \qquad \text{and} \qquad \delta^j_{\mathbb{S}} := \begin{cases} 1 & \text{if } j\in\mathbb{S} \\ 0 & \text{if } j\notin\mathbb{S}\,. \end{cases}$$

The eigenvalues of $M'_{\ell,j}$ are

$$\begin{cases} \pm\bar{\mu}\cdot\ell+\mu_j+\mathfrak{c} & \text{if } j\in\mathbb{S} \\ \pm\bar{\mu}\cdot\ell+\mu_j & \text{if } j\notin\mathbb{S}\,. \end{cases}$$

In case $(ii)$ of (4.1.1), $M'_\ell$ is represented, in the Hilbert basis

$$\Big(\Big(0,-\frac{\Psi_j}{\sqrt{2}}\,,\,\frac{\Psi_j}{\sqrt{2}},0\Big)\,,\,\Big(\frac{\Psi_j}{\sqrt{2}},0,0,\frac{\Psi_j}{\sqrt{2}}\Big)\,,\,\Big(\frac{\Psi_j}{\sqrt{2}},0,0,-\frac{\Psi_j}{\sqrt{2}}\Big)\,,\,\Big(0,\frac{\Psi_j}{\sqrt{2}},\frac{\Psi_j}{\sqrt{2}},0\Big)\Big)_{j\in\mathbb{N}}$$

of $\mathbf{H}=H\times H$, by the block-diagonal matrix $\mathrm{Diag}_{j\in\mathbb{N}}M'_{\ell,j}$, where

$$M'_{\ell,j}:=$$
$$\begin{pmatrix} \mu_j+\mathfrak{c}\delta^j_{\mathbb{S}}-\mu_k\delta^j_{(\mathbb{S}\cup\mathbb{F})^c} & \mathrm{i}\bar{\mu}\cdot\ell & 0 & 0 \\ -\mathrm{i}\bar{\mu}\cdot\ell & \mu_j+\mathfrak{c}\delta^j_{\mathbb{S}}-\mu_k\delta^j_{(\mathbb{S}\cup\mathbb{F})^c} & 0 & 0 \\ 0 & 0 & \mu_j+\mathfrak{c}\delta^j_{\mathbb{S}}+\mu_k\delta^j_{(\mathbb{S}\cup\mathbb{F})^c} & \mathrm{i}\bar{\mu}\cdot\ell \\ 0 & 0 & -\mathrm{i}\bar{\mu}\cdot\ell & \mu_j+\mathfrak{c}\delta^j_{\mathbb{S}}+\mu_k\delta^j_{(\mathbb{S}\cup\mathbb{F})^c} \end{pmatrix}$$

and, in this case, the eigenvalues of $M'_{\ell,j}$ are

$$\begin{cases} \pm\bar{\mu}\cdot\ell+\mu_j+\mathfrak{c} & \text{if } \ j\in\mathbb{S}\,, \\ \pm\bar{\mu}\cdot\ell+\mu_j & \text{if } \ j\in\mathbb{F}\,, \\ \pm\bar{\mu}\cdot\ell+\mu_j\pm\mu_k & \text{if } \ j\notin\mathbb{S}\cup\mathbb{F}\,. \end{cases} \tag{4.6.11}$$

By (4.1.5), (1.2.7), (1.2.16)-(1.2.17), in both cases we have $\|(M'_{\ell,j})^{-1}\|\le\gamma_0^{-1}\langle\ell\rangle^{\tau_0}$ for all $j\in\mathbb{N}$. Hence $M'_\ell$ defined in (4.6.10) is invertible and $\|(M'_\ell)^{-1}\|_0\le\gamma_0^{-1}\langle\ell\rangle^{\tau_0}$. In conclusion $M_\ell=D_m^{1/2}M'_\ell D_m^{1/2}$ is invertible and (4.6.9) holds, by the bound $\|D_m^{-1/2}\|_0\le C(m)$. ■

**Remark 4.6.3.** *The role of the term* $\mathfrak{c}\Pi_{\mathbb{S}}$ *in* (4.1.6), (4.1.7) *is precisely to prove Lemma 4.6.2. Otherwise, if* $\mathfrak{c}=0$, *then some of the eigenvalues* $\pm\bar{\mu}\cdot\ell+\mu_j$, $j\in\mathbb{S}$, *would vanish. We have some flexibility in defining the extended operators* $\mathcal{L}_r$, $\mathcal{L}_{r,\mu}$ *in the complementary subspace* $H_{\mathbb{S}\cup\mathbb{F}}$*: the important point is to define a positive definite operator leaving* $H_{\mathbb{S}\cup\mathbb{F}}$ *invariant.*

Let $\mathbf{H}_N$ be the finite dimensional subspace of $\mathbf{H}$ in (4.1.1),

$$\mathbf{H}_N:=\Big\{h(x)=\sum_{|j|\le N}u_je^{\mathrm{i}j\cdot x};\,u_j\in\mathbb{C}^r\Big\}\subset\mathbf{H}\,,\ r:=\begin{cases}2 \text{ in case (4.1.1)-}(i)\\ 4 \text{ in case (4.1.1)-}(ii)\end{cases} \tag{4.6.12}$$

and denote by $\Pi_N$ the corresponding $L^2_x$-projector.

Notice that, since the space $\mathbf{H}_N$ is not invariant under the operator $M_\ell$, the invertibility of the infinite dimensional operator $M_\ell$ does not imply the invertibility of the finite dimensional restriction $\Pi_N M_{\ell|\mathbf{H}_N}$. However, in the next lemma we prove that, for $N$ large enough, the operator

$$M_{\ell,N} := [M_\ell]_N^{2N} := \Pi_N (M_\ell)_{|\mathbf{H}_{2N}} : \mathbf{H}_{2N} \to \mathbf{H}_N \tag{4.6.13}$$

has a *right* inverse.

We shall use the following decomposition of the operator $M_\ell$ in (4.6.8):

$$M_\ell = D_m^{1/2}\big(\mathrm{i}\bar\mu\cdot\ell\, J + D_m + \mu_k \mathcal{J} + \bar R\big) D_m^{1/2}\,, \tag{4.6.14}$$

where the operator $\mathrm{i}\bar\mu\cdot\ell\, J + D_m + \mu_k\mathcal{J}$ is diagonal in the exponential basis $\{e^{\mathrm{i}j\cdot x},\ j\in\mathbb{Z}^d\}$, and the self-adjoint operator

$$\bar R := D_V - D_m - \mu_k \mathcal{J}\Pi_{\mathbb{S}\cup\mathbb{F}} + \mathfrak{c}\Pi_{\mathbb{S}} \tag{4.6.15}$$

satisfies, by (3.4.1) and Lemma 3.3.8, the off-diagonal estimate

$$|\bar R|_{+,s} < C(s) < +\infty\,, \quad \forall s\geq s_0\,. \tag{4.6.16}$$

**Lemma 4.6.4. (Right inverse of $M_{\ell,N}$)** *Let $M_{\ell,N}^* : \mathbf{H}_N \to \mathbf{H}_{2N}$ denote the adjoint operator of $M_{\ell,N}$. For all $N \geq \bar N$ large enough, $|\ell| \leq N < N_0^2$, the operator $M_{\ell,N}M_{\ell,N}^* : \mathbf{H}_N \to \mathbf{H}_N$ is invertible and*

$$\|\big(M_{\ell,N}M_{\ell,N}^*\big)^{-1}\|_0 \lesssim N^{2\tau_0}\,. \tag{4.6.17}$$

*As a consequence, the operator $B_{\ell,N}$ defined in* (4.6.13) *has the right inverse*

$$M_{\ell,N}^{-1} := M_{\ell,N}^*\big(M_{\ell,N}M_{\ell,N}^*\big)^{-1} : \mathbf{H}_N \to \mathbf{H}_{2N}\,, \tag{4.6.18}$$

*which satisfies*

$$\|M_{\ell,N}^{-1}\|_0 \lesssim N^{\tau_0}\,, \qquad |M_{\ell,N}^{-1}|_{s_0} \lesssim N^{\tau_0+s_0+\frac{d}{2}}\,. \tag{4.6.19}$$

PROOF. Since $M_\ell = M_\ell^*$, $M_{\ell,N}^* = \Pi_{2N}(M_\ell^*)_{|\mathbf{H}_N} = \Pi_{2N}(M_\ell)_{|\mathbf{H}_N}$. Using (4.6.14)-(4.6.15), we have

$$\begin{aligned}\|M_{\ell,N}^* - (M_\ell)_{|\mathbf{H}_N}\|_0 = \|\Pi_{2N}^\perp (M_\ell)_{|\mathbf{H}_N}\|_0 &\overset{(4.6.14)}{=} \|\Pi_{2N}^\perp (D_m^{1/2}\bar R D_m^{1/2})_{|\mathbf{H}_N}\|_0 \\ &\lesssim |\Pi_{2N}^\perp (D_m^{1/2}\bar R D_m^{1/2})_{|\mathbf{H}_N}|_{s_0} \\ &\overset{(3.3.14)}{\lesssim_s} N^{-(s-s_0)}|\bar R|_{+,s} \\ &\overset{(4.6.16)}{\leq} C(s)N^{-(s-s_0)}\,. \end{aligned} \tag{4.6.20}$$

For all $|\ell| \leq N < N_0^2$, we get, by (4.6.9), that, for any $h \in \mathbf{H}_N$,

$$\|M_\ell h\|_0 \gtrsim \gamma_0 \langle \ell \rangle^{-\tau_0} \|h\|_0 \gtrsim \gamma_0 N^{-\tau_0} \|h\|_0 \,. \tag{4.6.21}$$

Choosing $s > s_0 + \tau_0 + 2$ in (4.6.20), we deduce by (4.6.21) that, for $N \geq \bar{N}$ large enough, $\forall h \in \mathbf{H}_N$,

$$\begin{aligned} \|M_{\ell,N}^* h\|_0 &\geq \|M_\ell h\|_0 - \big\| \big( M_{\ell,N}^* - (M_\ell)_{|\mathbf{H}_N} \big) h \big\|_0 \\ &\geq \gamma_0 c N^{-\tau_0} \|h\|_0 - C(s) N^{-\tau_0 - 2} \|h\|_0 \\ &\gtrsim \gamma_0 N^{-\tau_0} \|h\|_0 \end{aligned}$$

and therefore

$$(M_{\ell,N} M_{\ell,N}^* h, h)_0 = \|M_{\ell,N}^* h\|_0^2 \gtrsim \gamma_0^2 N^{-2\tau_0} \|h\|_0^2 \,.$$

Since $\mathbf{H}_N$ is of finite dimension, we conclude that $M_{\ell,N} M_{\ell,N}^*$ is invertible and (4.6.17) holds (we do not track anymore the dependence with respect to the constant $\gamma_0$).

It is clear that $M_{\ell,N}^{-1}$ defined in (4.6.18) is a right inverse of $M_{\ell,N}$. Moreover, for all $h \in \mathbf{H}_N$,

$$\begin{aligned} \|M_{\ell,N}^{-1} h\|_0^2 &= \big( M_{\ell,N} M_{\ell,N}^* (M_{\ell,N} M_{\ell,N}^*)^{-1} h, (M_{\ell,N} M_{\ell,N}^*)^{-1} h \big)_0 \\ &= \big( h, (M_{\ell,N} M_{\ell,N}^*)^{-1} h \big)_0 \\ &\leq \|(M_{\ell,N} M_{\ell,N}^*)^{-1}\|_0 \|h\|_0^2 \end{aligned}$$

and (4.6.17) implies that $\|M_{\ell,N}^{-1}\|_0 \lesssim N^{\tau_0}$. The second inequality of (4.6.19) is an obvious consequence of the latter. ■

Our aim is now to obtain upper bounds of the $| \ |_s$-norms, for $s \geq s_0$, of the right inverse operator $M_{\ell,N}^{-1}$ defined in (4.6.18). We write

$$\begin{aligned} M_{\ell,N} = \Pi_N (M_\ell)_{|\mathbf{H}_{2N}} &= \big( \Pi_N M_\ell \Pi_N + \Pi_N M_\ell \Pi_N^\perp \big)_{|\mathbf{H}_{2N}} \\ &= \mathcal{D}_N + R_1 + R_2 \end{aligned} \tag{4.6.22}$$

where, recalling (4.6.14),

$$\begin{aligned} &\mathcal{D}_N := \Pi_N \mathcal{D}_{|\mathbf{H}_N} \,, \quad \mathcal{D} := D_m^{1/2} \big( \mathrm{i}\bar{\mu} \cdot \ell J + D_m + \mu_k \mathcal{J} \big) D_m^{1/2} \,, \\ &R_1 := \Pi_N (D_m^{1/2} \bar{R} D_m^{1/2} \Pi_N)_{|\mathbf{H}_{2N}} \\ &R_2 := \Pi_N (D_m^{1/2} \bar{R} D_m^{1/2})_{|\mathbf{H}_N^\perp \cap \mathbf{H}_{2N}} = \Pi_N (D_m^{1/2} \bar{R} D_m^{1/2} \Pi_N^\perp)_{|\mathbf{H}_{2N}} \,. \end{aligned} \tag{4.6.23}$$

By (4.6.16) we have

$$|R_1|_s \leq C(s) \,, \quad |R_2|_s \leq C(s) \,, \quad \forall s \geq s_0 \,. \tag{4.6.24}$$

We decompose accordingly the adjoint operator as

$$\begin{aligned} M_{\ell,N}^* = \Pi_{2N}(M_\ell)_{|\mathbf{H}_N} &= \Pi_N(M_\ell)_{|\mathbf{H}_N} + \Pi_N^\perp \Pi_{2N}(M_\ell)_{|\mathbf{H}_N} \\ &= \mathcal{D}_N + R_1^* + R_2^* \,. \end{aligned} \tag{4.6.25}$$

By (4.6.23), (3.3.13), (4.6.24), $M_{\ell,N}^*$ satisfies the estimate

$$|M_{\ell,N}^*|_s \le |\mathcal{D}_N|_s + |R_1^*|_s + |R_2^*|_s \le CN^2 + C(s) \lesssim_s N^2 \,. \tag{4.6.26}$$

Thus the right inverse $M_{\ell,N}^{-1}$ defined in (4.6.18), satisfies, by (3.3.6) and (4.6.26),

$$\begin{aligned} |M_{\ell,N}^{-1}|_s &\lesssim_s |M_{\ell,N}^*|_{s_0} |(M_{\ell,N} M_{\ell,N}^*)^{-1}|_s + |M_{\ell,N}^*|_s |(M_{\ell,N} M_{\ell,N}^*)^{-1}|_{s_0} \\ &\lesssim_s N^2 |(M_{\ell,N} M_{\ell,N}^*)^{-1}|_s \,. \end{aligned} \tag{4.6.27}$$

In Lemma 4.6.6 below we shall bound $|(M_{\ell,N} M_{\ell,N}^*)^{-1}|_s$ by a multi-scale argument. Without any loss of generality, we consider the case $(ii)$ of (4.1.1). We first give a general result which is a reformulation of the multiscale step Proposition 4.1 of [23], with stronger assumptions. Here $\mathfrak{I} := \{1, 2, 3, 4\}$.

**Lemma 4.6.5.** *Given $\varsigma \in (0, 1/2)$, $C_1 \ge 2$, $\tilde\tau > 0$, there are $\tilde\tau' > \tilde\tau$ (depending only on $\tilde\tau$ and $d$), $s^* > s_0$, $\eta > 0$, $\tilde N \ge 1$ with the following property. For any $N \ge \tilde N$, for any finite $\widetilde E \subset \mathbb{Z}^d \times \mathfrak{I}$ with $\mathrm{diam}(\widetilde E) \le N$, for any $A = D + R \in \mathcal{M}_{\widetilde E}^{\widetilde E}$ with $D$ diagonal, assume that*

*i)* **($L^2$-bound)** $\|A^{-1}\|_0 \le N^{\tilde\tau}$,

*ii)* **(Off-diagonal decay)** $|R|_{s^*} \le \eta$,

*iii)* **(Separation properties of the singular sites)** *There is a partition of the singular sites $\Omega := \{i \in \widetilde E \ : \ |D_i^i| < 1/4\} \subset \cup_\alpha \Omega_\alpha$ with*

$$\mathrm{diam}(\Omega_\alpha) \le N^{C_1\varsigma/(C_1+1)} \,, \quad \mathrm{d}(\Omega_\alpha, \Omega_\beta) \ge N^{2\varsigma/(C_1+1)} \,, \ \forall \alpha \ne \beta \,.$$

*Then*

$$|A^{-1}|_s \le C(s) N^{\tilde\tau'} (N^{\varsigma s} + |R|_s) \,, \quad \forall s \ge s_0 \,. \tag{4.6.28}$$

As usual, what is interesting in bound (4.6.28) is its "tamed" dependence with respect to $N$ ($\varsigma < 1$); the constant $C(s)$ may grow very strongly with $s$, but it does not depend on $N$.

**Lemma 4.6.6.** *There exist $t_0$, depending only on $\tau_0$, and $\bar N$ such that for all $\bar N \le N < N_0^2$, $|\ell| \le N$, we have that, $\forall s \ge s_0$,*

$$|(M_{\ell,N} M_{\ell,N}^*)^{-1}|_s \le C(s) N^{t_0 + \varsigma s} \,. \tag{4.6.29}$$

PROOF. We identify the operator $\mathcal{D}_N$ in (4.6.23) with the diagonal matrix

$$\mathrm{Diag}_{j\in[-N,N]^d}\mathcal{D}_j^j$$

where each $\mathcal{D}_j^j$ is in $\mathcal{L}(\mathbb{C}^4)$ (we are in case $(ii)$ of (4.1.1)). Using the unitary basis of $\mathbb{C}^4$ defined in (4.2.10), we identify each $\mathcal{D}_j^j$ with the $4\times 4$ diagonal matrix (see (4.2.11))

$$\mathcal{D}_j^j = \mathrm{Diag}(\mathcal{D}_{(j,\mathfrak{a})}^{(j,\mathfrak{a})})_{\mathfrak{a}\in\{1,2,3,4\}}\,, \quad \mathcal{D}_{(j,\mathfrak{a})}^{(j,\mathfrak{a})} := \langle j\rangle_m\big(\langle j\rangle_m + \sigma_1(\mathfrak{a})\mu_k + \sigma_2(\mathfrak{a})\bar{\mu}\cdot\ell\big) \tag{4.6.30}$$

with signs $\sigma_1(\mathfrak{a}), \sigma_2(\mathfrak{a})$ defined as in (4.4.7).

Notice that the singular sites of $M_\ell$ are those in (4.4.8) with $\omega = \bar{\mu}$, $\mu = \mu_k$, $\theta = 0$ and $\ell$ fixed.

By Lemma 4.4.8, which we can apply with $K = 4$ ($\ell$ being fixed), for $\Gamma \geq \bar{\Gamma}(\Theta)$, every $\Gamma$-chain of singular sites for $M_\ell$ is of length smaller than $(4\Gamma)^{C_2}$. Let us take

$$C_1 = 2C_2 + 3\,, \qquad \Gamma = N^{2/\chi} \quad \text{with} \quad \chi = \varsigma^{-1}(C_1+1)\,. \tag{4.6.31}$$

As a consequence, arguing as at the end of section 4.4, we deduce that, for $N$ large enough (depending on $\Theta$), the set $\Omega$ of the singular sites for $M_{\ell,N}$ can be partitioned as

$$\Omega = \cup_\alpha \Omega_\alpha\,, \quad \text{with} \quad \mathrm{diam}(\Omega_\alpha) \leq (4\Gamma)^{C_2}\times\Gamma \leq N^{C_1/\chi}\,,\ \mathrm{d}(\Omega_\alpha,\Omega_\beta) > N^{2/\chi}\,. \tag{4.6.32}$$

We now multiply

$$M_{\ell,N}M_{\ell,N}^* \overset{(4.6.22),(4.6.25)}{=} \mathcal{D}_N^2 + \mathcal{D}_N R_1^* + R_1\mathcal{D}_N + R_1R_1^* + R_2R_2^* \tag{4.6.33}$$

(notice that $\mathcal{D}_N R_2^*$, $R_2\mathcal{D}_N$, $R_1R_2^*$, $R_2R_1^*$ are zero) in both sides by the diagonal matrix

$$\begin{aligned} & d_{\Theta,N}^{-1} := (|\mathcal{D}_N| + \Theta\mathrm{Id}_N)^{-1} \qquad \text{where} \\ & |\mathcal{D}_N| := \mathrm{Diag}_{|j|\leq N,\mathfrak{a}\in\mathfrak{I}}(|\mathcal{D}_{j,\mathfrak{a}}^{j,\mathfrak{a}}|)_{|j|\leq N}\,, \quad \mathrm{Id}_N := \Pi_N\mathrm{Id}_{|\mathbf{H}_N}\,. \end{aligned} \tag{4.6.34}$$

We write

$$P_{\ell,N} := d_{\Theta,N}^{-1}M_{\ell,N}M_{\ell,N}^*d_{\Theta,N}^{-1} = \widetilde{\mathcal{D}}_N + \varrho_{\Theta,N} \tag{4.6.35}$$

where, by (4.6.33),

$$\widetilde{\mathcal{D}}_N := d_{\Theta,N}^{-1}\mathcal{D}_N^2 d_{\Theta,N}^{-1}\,, \qquad \varrho_{\Theta,N} := d_{\Theta,N}^{-1}\big(\mathcal{D}_N R_1^* + R_1\mathcal{D}_N + R_1R_1^* + R_2R_2^*\big)d_{\Theta,N}^{-1}\,. \tag{4.6.36}$$

We apply the multiscale Lemma 4.6.5 to $P_{\ell,N}$ with $C_1$ defined in (4.6.31) and $\tilde{\tau} := 2\tau_0 + 5$. Let us verify its assumptions. The operator $P_{\ell,N}$ defined in (4.6.35) is invertible as $M_{\ell,N}M_{\ell,N}^*$

(Lemma 4.6.4) and, by the definition of $d_{\Theta,N}$ in (4.6.34) and (4.6.30), for $N_0^2 > N \geq \bar{N}$ large enough (depending on $\Theta$), for all $|\ell| \leq N$, and using (4.6.17), we obtain

$$\begin{aligned} \|P_{\ell,N}^{-1}\|_0 &= \big\|d_{\Theta,N}\big(M_{\ell,N}M_{\ell,N}^*\big)^{-1}d_{\Theta,N}\big\|_0 \\ &\lesssim N^4\big\|\big(M_{\ell,N}M_{\ell,N}^*\big)^{-1}\big\|_0 \\ &\leq N^{2\tau_0+5} = N^{\tilde{\tau}} \end{aligned} \tag{4.6.37}$$

by the definition of $\tilde{\tau} := 2\tau_0 + 5$. Thus Assumption $i)$ of Lemma 4.6.5 holds.

Then we estimate the $|\ |_s$-decay norm of the operator $\varrho_{\Theta,N}$ in (4.6.36). By (3.3.20), (4.6.24) and $|d_{\Theta,N}^{-1}|_s \leq \Theta^{-1}$, $|d_{\Theta,N}^{-1}\mathcal{D}_N|_s \leq 1$ which directly follow by the definition (4.6.34), we get, for all $s \geq s_0$,

$$\begin{aligned} |\varrho_{\Theta,N}|_s &\lesssim_s |d_{\Theta,N}^{-1}\mathcal{D}_N|_s|R_1|_s|d_{\Theta,N}^{-1}|_s + |d_{\Theta,N}^{-1}|_s^2\big(|R_1|_{s_0}|R_1|_s + |R_2|_{s_0}|R_2|_s\big) \\ &\lesssim_s \Theta^{-1}\,. \end{aligned} \tag{4.6.38}$$

In particular, provided that $\Theta$ has been chosen large enough (depending on $\varsigma, C_1, \tau_0$), Assumption $ii)$ of Lemma 4.6.5 is satisfied.

To check Assumption $iii)$ it is enough to notice that, by the definition of $\widetilde{\mathcal{D}}_N$ in (4.6.36), and of $d_{\Theta,N}^{-1}$ in (4.6.34), for all $i \in [-N,N]^d \times \mathfrak{I}$, we have

$$\widetilde{\mathcal{D}}_i^i = \frac{|\mathcal{D}_i^i|^2}{(|\mathcal{D}_i^i| + \Theta)^2}\,, \qquad \text{and} \quad |\mathcal{D}_i^i| \geq \Theta \iff |\widetilde{\mathcal{D}}_i^i| \geq 1/4\,.$$

As a consequence, the separation properties for the singular sites of $M_{\ell,N}$ proved in (4.6.32) (with $\chi = \varsigma^{-1}(C_1+1)$), imply that Assumption $iii)$ of Lemma 4.6.5 is satisfied, provided that $N$ is large enough (depending only on $\Theta$). Lemma 4.6.5 implies that there is $t_0$, depending only on $\tau_0$, such that (see (4.6.28))

$$|P_{\ell,N}^{-1}|_s \lesssim_s N^{t_0}\big(N^{\varsigma s} + |\varrho_{\Theta,N}|_s\big)\,,$$

which gives, using (4.6.35), (3.3.20), $|d_{\Theta,N}^{-1}|_s \leq 1$,

$$\begin{aligned} |(M_{\ell,N}M_{\ell,N}^*)^{-1}|_s = |d_{\Theta,N}^{-1}P_{\ell,N}^{-1}d_{\Theta,N}^{-1}|_s &\lesssim_s |d_{\Theta,N}^{-1}|_s^2|P_{\ell,N}^{-1}|_s \\ &\lesssim_s N^{t_0}(N^{\varsigma s} + |\varrho_{\Theta,N}|_s)\,. \end{aligned} \tag{4.6.39}$$

Finally, estimate (4.6.29) follows by (4.6.39) and (4.6.38). ∎

Proof of Proposition 4.6.1 concluded. Recalling the decomposition (4.6.2), the first goal is to define a right inverse of the operator

$$\mathcal{L}_{\mathtt{d},N} : \mathcal{H}_{2N} \to \mathcal{H}_N\,, \quad \mathcal{L}_{\mathtt{d},N} := \Pi_N D_m^{1/2}\mathcal{L}_{\mathtt{d}}D_m^{1/2}|_{\mathcal{H}_{2N}}\,, \tag{4.6.40}$$

where $\mathcal{L}_{\mathtt{d}}$ is defined in (4.6.3). Recalling (4.6.7), (4.6.13) and Lemma 4.6.4, the linear operator $\mathcal{L}_{\mathtt{d},N}^{-1} : \mathcal{H}_N \to \mathcal{H}_{2N}$ defined by

$$\mathcal{L}_{\mathtt{d},N}^{-1}(e^{\mathrm{i}\ell\cdot\varphi}g) = e^{\mathrm{i}\ell\cdot\varphi}M_{\ell,N}^{-1}(g)\,, \quad \forall \ell \in [-N,N]^{|\mathbb{S}|}\,, \quad \forall g \in \mathbf{H}_N\,,$$

is a right inverse of $\mathcal{L}_{\mathtt{d},N}$. Using that $\mathcal{L}_{\mathtt{d},N}^{-1}$ is diagonal in time it results

$$|\mathcal{L}_{\mathtt{d},N}^{-1}|_s \le N^{|\mathbb{S}|/2} \max_{|\ell|\le N} |M_{\ell,N}^{-1}|_s\,. \tag{4.6.41}$$

By (4.6.27) and Lemma 4.6.6, we have the bound

$$|M_{\ell,N}^{-1}|_s \lesssim_s N^{t_0+\varsigma s+2}\,, \quad \forall s \ge s_0\,. \tag{4.6.42}$$

Therefore, by (4.6.41), (4.6.42) and the second estimate in (4.6.19), we get

$$|\mathcal{L}_{\mathtt{d},N}^{-1}|_s \lesssim_s N^{\frac{|\mathbb{S}|}{2}+t_0+\varsigma s+2}\,,\ \forall s \ge s_0\,, \qquad |\mathcal{L}_{\mathtt{d},N}^{-1}|_{s_0} \lesssim N^{\frac{b}{2}+\tau_0+s_0}\,, \quad b = d+|\mathbb{S}|\,. \tag{4.6.43}$$

Finally we use a perturbative Neumann series argument to prove the existence of a right inverse of

$$\begin{aligned}\mathcal{L}_{r,\mu,N}^{+} := D_m^{1/2}[\mathcal{L}_{r,\mu}]_N^{2N} D_m^{1/2} &\overset{(4.1.19)}{=} \varPi_N D_m^{1/2}\mathcal{L}_{r,\mu}(\varepsilon,\lambda) D_m^{1/2}{}_{|\mathcal{H}_{2N}} \\ &\overset{(4.6.2),(4.6.40)}{=} \mathcal{L}_{\mathtt{d},N} + \rho_N\end{aligned} \tag{4.6.44}$$

where

$$\rho_N := \varPi_N D_m^{1/2}\rho(\varepsilon,\lambda,\varphi) D_m^{1/2}{}_{|\mathcal{H}_{2N}} \tag{4.6.45}$$

satisfies, by (4.6.4), (4.6.5), (4.6.6), Lemma 3.3.8, and item 1 of Definition 4.1.2,

$$|\rho_N|_{\mathrm{Lip},s} \lesssim_s \varepsilon^2 N^2 + |r|_{\mathrm{Lip},+,s}\,, \quad |\rho_N|_{\mathrm{Lip},s_1} \lesssim_{s_1} \varepsilon^2 N^2\,. \tag{4.6.46}$$

Using the second estimates in (4.6.43) and (4.6.46), $N \le N_0^2$, (4.5.10) and (4.5.1), we get

$$|\mathcal{L}_{\mathtt{d},N}^{-1}|_{s_0}|\rho_N|_{\mathrm{Lip},s_0} \lesssim_{s_1} N^{\frac{b}{2}+\tau_0+s_0+2}\varepsilon^2 \le N_0^{b+2\tau_0+2s_0+4}\varepsilon^2 \ll 1\,. \tag{4.6.47}$$

Hence $\mathrm{Id}_{\mathcal{H}_N} + \rho_N\mathcal{L}_{\mathtt{d},N}^{-1}$ is invertible and the operator $\mathcal{L}_{r,\mu,N}^{+}$ in (4.6.44) has the right inverse

$$(\mathcal{L}_{r,\mu,N}^{+})^{-1} = D_m^{-1/2}([\mathcal{L}_{r,\mu}]_N^{2N})^{-1}D_m^{-1/2} := \mathcal{L}_{\mathtt{d},N}^{-1}\big(\mathrm{Id}_{\mathcal{H}_N} + \rho_N\mathcal{L}_{\mathtt{d},N}^{-1}\big)^{-1} \tag{4.6.48}$$

which satisfies the following tame estimates (see Lemma 3.3.12) for all $s \ge s_0$,

$$|(\mathcal{L}_{r,\mu,N}^{+})^{-1}|_{\mathrm{Lip},s} \lesssim_s |\mathcal{L}_{\mathtt{d},N}^{-1}|_{\mathrm{Lip},s} + |\mathcal{L}_{\mathtt{d},N}^{-1}|_{\mathrm{Lip},s_0}^2|\rho_N|_{\mathrm{Lip},s}\,. \tag{4.6.49}$$

Since $\mathcal{L}_{\mathtt{d},N}^{-1}$ does not depend on $\lambda$, $|\mathcal{L}_{\mathtt{d},N}^{-1}|_{\mathrm{Lip},s} = |\mathcal{L}_{\mathtt{d},N}^{-1}|_s$, and (4.6.48), (4.6.49), (4.6.43), (4.6.46) imply

$$|D_m^{-1/2}([\mathcal{L}_{r,\mu}]_N^{2N})^{-1}D_m^{-1/2}|_{\mathrm{Lip},s} \lesssim_s N^{\tau_0'}\big(N^{\varsigma s} + |r|_{\mathrm{Lip},+,s}\big)$$

where

$$\tau_0' := \max\left\{(|\mathbb{S}|/2) + t_0 + 2, b + 2s_0 + 2\tau_0 + 4\right\}$$

and, using (3.3.28), the estimate (4.6.1) follows.

Let us now prove (4.1.21). Calling $\rho'$ the operator defined in (4.6.4) associated to $(\mu', r')$ we have $\rho - \rho' = (\mu - \mu')\mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} + r - r'$ and, by (4.6.48),

$$\begin{aligned}
&D_m^{-1/2}\big(([\mathcal{L}_{r,\mu}]_N^{2N})^{-1} - ([\mathcal{L}_{r',\mu'}]_N^{2N})^{-1}\big)D_m^{-1/2} \\
&= \mathcal{L}^{-1}_{\mathtt{d},N}\Big[\big(\mathrm{Id}_{\mathcal{H}_N} + \rho_N\mathcal{L}^{-1}_{\mathtt{d},N}\big)^{-1} - \big(\mathrm{Id}_{\mathcal{H}_N} + \rho'_N\mathcal{L}^{-1}_{\mathtt{d},N}\big)^{-1}\Big] \\
&= \mathcal{L}^{-1}_{\mathtt{d},N}\big(\mathrm{Id}_{\mathcal{H}_N} + \rho_N\mathcal{L}^{-1}_{\mathtt{d},N}\big)^{-1}(\rho'_N - \rho_N)\mathcal{L}^{-1}_{\mathtt{d},N}\big(\mathrm{Id}_{\mathcal{H}_N} + \rho'_N\mathcal{L}^{-1}_{\mathtt{d},N}\big)^{-1} \\
&= (\mathcal{L}^{+}_{r,\mu,N})^{-1}(\rho'_N - \rho_N)(\mathcal{L}^{+}_{r',\mu',N})^{-1}\,.
\end{aligned} \tag{4.6.50}$$

In conclusion, by (4.6.50) (3.3.20), (4.6.49), and $|\mathcal{L}^{-1}_{\mathtt{d},N}|_{s_0}|\rho_N|_{\mathrm{Lip},s_1} \leq 1$, we get

$$\begin{aligned}
\big|D_m^{-1/2}\big(([\mathcal{L}_{r,\mu}]_N^{2N})^{-1} - ([\mathcal{L}_{r',\mu'}]_N^{2N})^{-1}\big)D_m^{-1/2}\big|_{s_1} &\lesssim_{s_1} |\mathcal{L}^{-1}_{\mathtt{d},N}|^2_{s_1}|\rho_N - \rho'_N|_{s_1} \\
&\lesssim_{s_1} N^{2(\tau_0' + \varsigma s_1)}\big(|\mu - \mu'|N^2 + |r - r'|_{+,s_1}\big)
\end{aligned}$$

which, using (3.3.28), implies (4.1.21). This completes the proof of Proposition 4.6.1.

**Remark 4.6.7.** *In this section we have proved directly the Lipschitz estimate of* $([\mathcal{L}_{r,\mu}]_N^{2N})^{-1}$ *instead of arguing as in Proposition 4.7.6, because for a right inverse we do not have the formula* (4.7.23).

## 4.7 Inverse of $\mathcal{L}_{r,\mu,N}$ for $N \geq N_0^2$

In Proposition 4.6.1 we proved item 1 of Proposition 4.1.5. The aim of this section is prove item 2 of Proposition 4.1.5, namely that for all $N \geq N_0^2$ and $\lambda \in \Lambda(\varepsilon; 1, X_{r,\mu})$ the operator $\mathcal{L}_{r,\mu,N}$ defined in (4.1.22) is invertible and its inverse $\mathcal{L}^{-1}_{r,\mu,N}$ has off-diagonal decay, see Proposition 4.7.6. The proof is based on inductive applications of the multiscale step Proposition 4.3.4, thanks to Proposition 4.4.5 about the separation properties of the bad sites.

#### The set $\Lambda(\varepsilon; 1, X_{r,\mu})$ is good at any scale

We first prove the following proposition.

**Proposition 4.7.1.** *The* $\Lambda(\varepsilon; 1, X_{r,\mu})$ *in* (4.5.12) *satisfies*

$$\Lambda(\varepsilon; 1, X_{r,\mu}) \subset \bigcap_{N \leq N_0} \mathcal{G}_N \bigcap_{k \geq 1} \mathcal{G}_{N_k} \tag{4.7.1}$$

*where* $\mathcal{G}_N$ *are defined in* (4.4.5).

We first consider the small scales $N \leq N_0$.

**Lemma 4.7.2.** *Let $\breve{A}(\varepsilon, \lambda, \theta)$ be the matrix in* (4.2.14) *corresponding to $r = 0$, see remark 4.2.1. There is $\bar{N}$ such that for all $\bar{N} \leq N \leq N_0 \leq (2\varepsilon^{-2})^{\frac{1}{\tau+s_0+d}}$ (see* (4.5.10)*), $\forall \lambda \in \widetilde{\Lambda}$, $\forall j_0 \in \mathbb{Z}^d$, $\theta \in \mathbb{R}$,*

$$\begin{aligned} &\|D_m^{-1/2} \breve{A}_{N,j_0}^{-1}(\varepsilon, \lambda, \theta) D_m^{-1/2}\|_0 \leq N^\tau \quad \Longrightarrow \\ &|D_m^{-1/2} A_{N,j_0}^{-1}(\varepsilon, \lambda, \theta) D_m^{-1/2}|_s \leq N^{\tau' + \varsigma s}, \ \forall s \in [s_0, s_1], \end{aligned} \tag{4.7.2}$$

*namely the matrix $A_{N,j_0}(\varepsilon, \lambda, \theta)$ is $N$-good according to Definition 4.3.1.*

Proof. For brevity, the dependence of the operators with respect to $(\varepsilon, \lambda)$ is kept implicit.
**Step 1.** We first prove that there is $\bar{N}$ such that $\forall N \geq \bar{N}$, $\forall j_0 \in \mathbb{Z}^d$, $\forall \theta \in \mathbb{R}$,

$$\begin{aligned} &\|D_m^{-1/2} \breve{A}_{N,j_0}^{-1}(\theta) D_m^{-1/2}\|_0 \leq N^\tau \implies \\ &|D_m^{-1/2} \breve{A}_{N,j_0}^{-1}(\theta) D_m^{-1/2}|_s \leq C(s) N^{\tau_1' + \varsigma s}, \ \forall s \geq s_0, \end{aligned} \tag{4.7.3}$$

where $\tau_1' := \tilde{\tau}' + (|\mathbb{S}|/2)$ and $\tilde{\tau}' > \tau$ is the constant provided by Lemma 4.6.5 associated to $\tilde{\tau} = \tau$.

Recall that the matrix $\breve{A}(\theta)$ represents the $L^2$-self-adjoint operator $\mathcal{L}_{0,\mu}(\theta)$ defined in (4.2.17), which is independent of $\varphi$. For all $\ell \in \mathbb{Z}^{|\mathbb{S}|}$, $h \in \mathbf{H}$, we have that

$$D_m^{1/2} \mathcal{L}_{0,\mu}(\theta) D_m^{1/2}(e^{\mathrm{i}\ell\cdot\varphi} h) = e^{\mathrm{i}\ell\cdot\varphi} M_\ell(\theta) h, \quad M_\ell(\theta) = \mathcal{D}_\ell(\theta) + T_\mu,$$

where

$$\begin{aligned} &\mathcal{D}_\ell(\theta) := \mathrm{i}(\omega \cdot \ell + \theta) J D_m + D_m^2 + \mu \mathcal{J} D_m, \\ &T_\mu := D_m^{1/2}\big(D_V - D_m - \mu \mathcal{J} \Pi_{\mathbb{S}\cup\mathbb{F}} + \mathfrak{c}\Pi_{\mathbb{S}}\big) D_m^{1/2}. \end{aligned}$$

Note that, by (3.4.1), Lemma 3.3.8, and since $\mu = O(1)$ (item 2 of Definition 4.1.2), it results

$$|T_\mu|_s \leq C(s), \quad \forall s \geq s_0. \tag{4.7.4}$$

In order to prove (4.7.3), since

$$|D_m^{-1/2} \breve{A}_{N,j_0}^{-1}(\theta) D_m^{-1/2}|_s \leq N^{|\mathbb{S}|/2} \max_{|\ell| \leq N} |(M_\ell(\theta))_{N,j_0}^{-1}|_s, \tag{4.7.5}$$

it is sufficient to bound the $| \ |_s$-norms of $(M_\ell(\theta))_{N,j_0}^{-1}$. We apply the multiscale Lemma 4.6.5.

We identify as usual the operator $\mathcal{D}_\ell(\theta)$ with the diagonal matrix $\mathrm{Diag}_j((\mathcal{D}_\ell(\theta))_j^j)$ where

$$\begin{aligned} &(\mathcal{D}_\ell(\theta))_j^j = \mathrm{Diag}([\mathcal{D}_\ell(\theta)]_{(j,\mathfrak{a})}^{(j,\mathfrak{a})})_{\mathfrak{a} \in \{1,2,3,4\}}, \\ &[\mathcal{D}_\ell(\theta)]_{(j,\mathfrak{a})}^{(j,\mathfrak{a})} := \langle j \rangle_m \big(\langle j \rangle_m + \sigma_1(\mathfrak{a})\mu + \sigma_2(\mathfrak{a})(\omega \cdot \ell + \theta)\big) \end{aligned}$$

with signs $\sigma_1(\mathfrak{a}), \sigma_2(\mathfrak{a})$ defined as in (4.4.7). Notice that the singular sites of $\mathcal{D}_\ell(\theta)$ are those in (4.4.8) with $\ell$ fixed. By Lemma 4.4.8, which we use with $K = 4$, $\ell$ being fixed, there is $C_2$ (independent of $\Theta$) such that, for $\Gamma > \bar{\Gamma}(\Theta)$, any $\Gamma$-chain of singular sites has length $L \leq (4\Gamma)^{C_2}$. As in (4.6.31), we can apply this result with $C_1 = 2C_2 + 3$, $\chi = \varsigma^{-1}(C_1 + 1)$ and $\Gamma = N^{2/\chi}$ (for any $N \geq \bar{N}$ large enough), and we find that the operator $\Theta^{-1}(M_\ell(\theta))_{N,j_0}$ satisfies Assumption *iii*) of Lemma 4.6.5, where we take $\tilde{\tau} = \tau$. Assumption *ii*) is also satisfied, provided that $\Theta$ has been chosen large enough, more precisely $\Theta \geq \eta^{-1}C(s^*)$, with the constant $C(s^*)$ of (4.7.4).

By Lemma 4.6.5, there is $\tilde{\tau}' > \tau$ such that $\forall N \geq \bar{N}$, $\forall j_0 \in \mathbb{Z}^d$, $\forall \theta \in \mathbb{R}$, $\forall \ell \in \mathbb{Z}^{|\mathbb{S}|}$,

$$\|(M_\ell(\theta))_{N,j_0}^{-1}\|_0 \leq N^\tau \quad \Longrightarrow \quad \forall s \geq s_0\,,\ |(M_\ell(\theta))_{N,j_0}^{-1}|_s \lesssim_s N^{\tilde{\tau}'}\big(N^{\varsigma s} + |T_\mu|_s\big)\,. \tag{4.7.6}$$

Since

$$\|D_m^{-1/2}\breve{A}_{N,j_0}^{-1}(\theta)D_m^{-1/2}\|_0 = \max_{|\ell|\leq N} \|(M_\ell(\theta))_{N,j_0}^{-1}\|_0\,,$$

the premise in (4.7.3) implies the premise in (4.7.6) and therefore (4.7.5), (4.7.6), (4.7.4) imply (4.7.3) since $\tau_1' := \tilde{\tau}' + |\mathbb{S}|/2$.

**Step 2.** We now apply a perturbative argument to the operator

$$D_m^{1/2}A_{N,j_0}(\theta)D_m^{1/2} = D_m^{1/2}\breve{A}_{N,j_0}(\theta)D_m^{1/2} + \rho_{N,j_0}\,, \tag{4.7.7}$$

where $\rho$ is the matrix which represents $D_m^{1/2} r D_m^{1/2}$, and, by item 1 of Definition 4.1.2,

$$|\rho_{N,j_0}|_{s_1} \leq |r|_{+,s_1} \leq C_1\varepsilon^2\,. \tag{4.7.8}$$

If $\|D_m^{-1/2}\breve{A}_{N,j_0}^{-1}(\theta)D_m^{-1/2}\|_0 \leq N^\tau$ then

$$|D_m^{-1/2}\breve{A}_{N,j_0}^{-1}(\theta)D_m^{-1/2}|_{s_0}|\rho_{N,j_0}|_{s_1} \lesssim N^{\tau+s_0+(d/2)}\varepsilon^2 \lesssim N_0^{\tau+s_0+(d/2)}\varepsilon^2 \ll 1 \tag{4.7.9}$$

since $N_0 \leq (2\varepsilon^{-2})^{\frac{1}{\tau+s_0+d}}$ (see (4.5.10)). Then Lemma 3.3.12 implies that $D_m^{1/2}A_{N,j_0}(\theta)D_m^{1/2}$ in (4.7.7) is invertible, and $\forall s \in [s_0, s_1]$

$$\begin{aligned} |D_m^{-1/2}A_{N,j_0}^{-1}(\theta)D_m^{-1/2}|_s &\lesssim_{s_1} |D_m^{-1/2}\breve{A}_{N,j_0}^{-1}(\theta)D_m^{-1/2}|_s + |D_m^{-1/2}\breve{A}_{N,j_0}^{-1}(\theta)D_m^{-1/2}|_{s_0}^2|\rho_{N,j_0}|_{s_1} \\ &\overset{(4.7.9)}{\lesssim_{s_1}} |D_m^{-1/2}\breve{A}_{N,j_0}^{-1}(\theta)D_m^{-1/2}|_s \overset{(4.7.3)}{\lesssim_{s_1}} N^{\tau_1'+\varsigma s} \leq N^{\tau'+\varsigma s}\,, \end{aligned}$$

because $\tau' > \tau_1' = \tilde{\tau}' + (|\mathbb{S}|/2)$ (see (4.5.3)) and for $\bar{N}$ large enough. ■

At small scales $N \leq N_0$, any $\lambda \in \widetilde{\Lambda}$ is $N$-good.

**Lemma 4.7.3. (Initialization)** *For all $N \leq N_0$ and $\varepsilon$ small, the set $\mathcal{G}_N$ defined in* (4.4.5) *is $\mathcal{G}_N = \widetilde{\Lambda}$.*

PROOF. Lemma 4.7.2 implies that $\forall \lambda \in \widetilde{\Lambda}$, $\forall j_0 \in \mathbb{Z}^d$, the set $B_N(j_0; \lambda)$ defined in (4.4.3) satisfies

$$B_N(j_0;\lambda) \subset \breve{B}_N^0 := \Big\{ \theta \in \mathbb{R} \,:\, \| D_m^{-1/2} \breve{A}_{N,j_0}^{-1}(\varepsilon,\lambda,\theta) D_m^{-1/2} \|_0 > N^\tau/2 \Big\} . \tag{4.7.10}$$

Thus, in order to prove that $\mathcal{G}_N = \widetilde{\Lambda}$, it is sufficient to show that the set $\breve{B}_N^0$ in (4.7.10) satisfies the complexity bound (4.4.4). Note that, since $\|D_m^{-1/2}\|_0 \leq m^{-1/4}$, we have

$$\begin{aligned} \breve{B}_N^0 &\subset \big\{ \theta \in \mathbb{R} \,:\, \| \breve{A}_{N,j_0}^{-1}(\varepsilon,\lambda,\theta) \|_0 > C N^\tau \,, \quad C := \sqrt{m}/2 \big\} \\ &= \big\{ \theta \in \mathbb{R} \,:\, \exists \text{ an eigenvalue of } \breve{A}_{N,j_0}(\varepsilon,\lambda,\theta) \text{ with modulus less than } N^{-\tau}/C \big\} . \end{aligned} \tag{4.7.11}$$

Let $\Pi_{N,j_0}$ denote the $L^2$-projector on the subspace

$$H_{N,j_0} := \Big\{ (q(x),p(x)) = \sum_{|j-j_0| \leq N} (q_j,p_j) e^{\mathrm{i} j \cdot x} \Big\} .$$

Since $\breve{A}_{N,j_0}(\varepsilon,\lambda,\theta)$ represents the operator $\mathcal{L}_{0,\mu}(\theta)$ in (4.2.17) which does not depend on $\varphi$ (see remark 4.2.1), the spectrum of $\breve{A}_{N,j_0}(\varepsilon,\lambda,\theta)$ is formed by

$$\begin{gathered} \pm(\omega \cdot \ell + \theta) - \beta_j \,,\ j = 1, \ldots, (2N+1)^d \,,\ \ell \in \mathbb{Z}^{|\mathbb{S}|} , \\ \beta_j \text{ eigenvalue of } \Pi_{N,j_0} \big( D_V + \mu \mathcal{J} \Pi_{\mathbb{S} \cup \mathbb{F}}^\perp + \mathfrak{c} \Pi_{\mathbb{S}} \big) \Pi_{N,j_0} \,, \end{gathered}$$

and, by (4.7.11), we have

$$\begin{gathered} \breve{B}_N^0 \subset \bigcup_{|\ell| \leq N, j = 1, \ldots, (2N+1)^d, \sigma = \pm} \mathcal{R}_{\ell,j}^\sigma \,, \\ \mathcal{R}_{\ell,j}^\sigma := \big\{ \theta \in \mathbb{R} \,:\, |\sigma(\theta + \omega \cdot \ell) - \beta_j| \leq N^{-\tau}/C \big\} . \end{gathered} \tag{4.7.12}$$

It follows that $\breve{B}_N^0$ is included in the union of $N^{|\mathbb{S}|+d+1}$ intervals $I_q$ of length $2N^{-\tau}/C$. By eventually dividing the intervals $I_q$ we deduce that $\breve{B}_N^0$ is included in the union of $N^{d+2+|\mathbb{S}|}$ intervals $I_q$ of length $N^{-\tau}$. ■

**Lemma 4.7.4.** *For all $k \geq 0$ we have*

$$\mathcal{G}_{N_k} \bigcap \mathcal{G}_{N_{k+1},1}^0 \bigcap \tilde{\mathcal{G}} \subset \mathcal{G}_{N_{k+1}} \tag{4.7.13}$$

*where the set $\mathcal{G}_N$ is defined in* (4.4.5)*, $\mathcal{G}_{N,\eta}^0$ in* (4.5.5) *and $\tilde{\mathcal{G}}$ in* (4.5.8)*.*

PROOF. Let $\lambda \in \mathcal{G}_{N_k} \cap \mathcal{G}_{N_{k+1},1}^0 \cap \tilde{\mathcal{G}}$. In order to prove that $\lambda \in \mathcal{G}_{N_{k+1}}$ (Definition 4.4.4), since $\lambda \in \mathcal{G}_{N_{k+1},1}^0$ (set defined in (4.5.5)), it is sufficient to prove that the sets $B_{N_{k+1}}(j_0;\lambda)$ in (4.4.4) and $B_{N_{k+1}}^0(j_0;\lambda,1)$ in (4.5.4) satisfy:

$$\forall j_0 \in \mathbb{Z}^d \,, \quad B_{N_{k+1}}(j_0;\lambda) \subset B_{N_{k+1}}^0(j_0;\lambda,1) \,,$$

or equivalently, that

$$
\begin{aligned}
&\|D_m^{-1/2} A^{-1}_{N_{k+1},j_0}(\varepsilon,\lambda,\theta) D_m^{-1/2}\|_0 \le N_{k+1}^{\tau} \quad \Longrightarrow \\
&|D_m^{-1/2} A^{-1}_{N_{k+1},j_0}(\varepsilon,\lambda,\theta) D_m^{-1/2}|_s \le N_{k+1}^{\tau'+\varsigma s}\,, \quad \forall s\in[s_0,s_1]\,.
\end{aligned}
\tag{4.7.14}
$$

We prove (4.7.14) applying the multiscale step Proposition 4.3.4 to the matrix $A_{N_{k+1},j_0}$. By (4.2.13) the assumption (H1) holds. The assumption (H2) is the premise in (4.7.14). Let us verify (H3). By remark 4.4.3, a site

$$
k \in E := \Big( (0,j_0) + [-N_{n+1}, N_{n+1}]^b \Big) \times \mathfrak{I}\,, \tag{4.7.15}
$$

which is $N_k$-good for $A(\varepsilon,\lambda,\theta) := \mathcal{L}_{r,\mu} + \theta Y$ (see Definition 4.4.2 with $A = A(\varepsilon,\lambda,\theta)$) is also

$$
(A_{N_{n+1},j_0}(\varepsilon,\lambda,\theta), N_k) - \text{good}
$$

(see Definition 4.3.3 with $A = A_{N_{n+1},j_0}(\varepsilon,\lambda,\theta)$). As a consequence we have the inclusion

$$
\begin{aligned}
&\Big\{ \ (A_{N_{n+1},j_0}(\varepsilon,\lambda,\theta), N_k)\text{−bad sites } \Big\} \subset \\
&\Big\{ N_k\text{−bad sites of } A(\varepsilon,\lambda,\theta) \text{ with } |\ell| \le N_{k+1} \Big\}
\end{aligned}
\tag{4.7.16}
$$

and (H3) is proved if the latter $N_k$-bad sites (in the right hand side of (4.7.16)) are contained in a disjoint union $\cup_\alpha \Omega_\alpha$ of clusters satisfying (4.3.7) (with $N = N_k$). This is a consequence of Proposition 4.4.5 applied to the infinite dimensional matrix $A(\varepsilon,\lambda,\theta)$. Since $\lambda \in \mathcal{G}_{N_k}$ then assumption (i) of Proposition 4.4.5 holds with $N = N_k$. Assumption (ii) holds by (4.5.1). Assumption (iii) of Proposition 4.4.5 holds because $\lambda \in \tilde{\mathcal{G}}$, see (4.5.8). Therefore the $N_k$-bad sites of $A(\varepsilon,\lambda,\theta)$ satisfy (4.4.6) with $N = N_k$, and therefore (H3) holds.

Then the multiscale step Proposition 4.3.4 applied to the matrix $A_{N_{k+1},j_0}(\varepsilon,\lambda,\theta)$ implies that if

$$
\|D_m^{-1/2} A^{-1}_{N_{k+1},j_0}(\varepsilon,\lambda,\theta) D_m^{-1/2}\|_0 \le N_{k+1}^{\tau}
$$

then

$$
\begin{aligned}
|D_m^{-1/2} A^{-1}_{N_{k+1},j_0}(\varepsilon,\lambda,\theta) D_m^{-1/2}|_s &\le \frac{1}{4} N_{k+1}^{\tau'} \big( N_{k+1}^{\varsigma s} + |T|_{+,s} \big) \\
&\overset{(4.2.13)}{\le} N_{k+1}^{\tau'+\varsigma s}\,,\ \forall s\in[s_0,s_1]\,,
\end{aligned}
\tag{4.7.17}
$$

proving (4.7.14). ■

**Corollary 4.7.5.** *For all $n \ge 1$ we have*

$$
\bigcap_{k=1}^{n} \mathcal{G}^0_{N_k,1} \bigcap \tilde{\mathcal{G}} \subset \mathcal{G}_{N_n}\,. \tag{4.7.18}
$$

Proof. For $n = 1$ the inclusion (4.7.18) follows by (4.7.13) at $k = 0$ and the fact that $\mathcal{G}_{N_0} = \widetilde{\Lambda}$ by Lemma 4.7.3. Then we argue by induction. Supposing that (4.7.18) holds at the step $n$ then

$$\bigcap_{k=1}^{n+1} \mathcal{G}^0_{N_k,1} \bigcap \tilde{\mathcal{G}} = \mathcal{G}^0_{N_{n+1},1} \bigcap \Big( \bigcap_{k=1}^{n} \mathcal{G}^0_{N_k,1} \bigcap \tilde{\mathcal{G}} \Big) \stackrel{(4.7.18)_n}{\subset} \mathcal{G}^0_{N_{n+1},1} \bigcap \mathcal{G}_{N_n} \bigcap \tilde{\mathcal{G}} \stackrel{(4.7.13)_n}{\subset} \mathcal{G}_{N_{n+1}}$$

proving (4.7.18) at the step $n + 1$. ■

Proof of proposition 4.7.1 concluded. Corollary 4.7.5 implies that

$$\bigcap_{k\geq 1} \mathcal{G}^0_{N_k,1} \bigcap \tilde{\mathcal{G}} \subset \bigcap_{n\geq 1} \mathcal{G}_{N_n} \,. \tag{4.7.19}$$

Then we conclude that the set $\Lambda(\varepsilon; 1, X_{r,\mu})$ defined in (4.5.12) satisfies

$$\Lambda(\varepsilon; 1, X_{r,\mu}) = \bigcap_{k\geq 1} \mathcal{G}^0_{N_k,1} \bigcap_{N\geq N_0^2} \mathtt{G}^0_{N,1} \bigcap \tilde{\mathcal{G}} \stackrel{(4.7.19)}{\subset} \bigcap_{n\geq 1} \mathcal{G}_{N_n}$$

proving (4.7.1), since $\mathcal{G}_N = \widetilde{\Lambda}$, for all $N \leq N_0$, by Lemma 4.7.3.

### Inverse of $\mathcal{L}_{r,\mu,N}$ for $N \geq N_0^2$

We can finally prove the following proposition.

**Proposition 4.7.6.** *For all $N \geq N_0^2$, $\lambda \in \Lambda(\varepsilon; 1, X_{r,\mu})$, the operator $\mathcal{L}_{r,\mu,N}$ defined in* (4.1.22) *is invertible and satisfies* (4.1.23). *Moreover* (4.1.24) *holds.*

Proof. Let $\lambda \in \Lambda(\varepsilon; 1, X_{r,\mu})$. For all $N \geq N_0^2$ there is $M \in \mathbb{N}$ such that $N = M^\chi$, for some $\chi \in [\bar{\chi}, \bar{\chi}^2]$ and $\lambda \in \mathcal{G}_M$. In fact

1. If $N \geq N_1$, then $N \in [N_{n+1}, N_{n+2}]$ for some $n \in \mathbb{N}$, and we have $N = N_n^\chi$ for some $\chi \in [\bar{\chi}, \bar{\chi}^2]$. Moreover if $\lambda \in \Lambda(\varepsilon; 1, X_{r,\mu})$ then $\lambda \in \mathcal{G}_{N_n}$ by (4.7.1).
2. If $N_0^2 \leq N < N_1$, it is enough to write $N = M^{\bar{\chi}}$ for some integer $M < N_0$. Moreover if $\lambda \in \Lambda(\varepsilon; 1, X_{r,\mu})$ then $\lambda \in \mathcal{G}_M$ by (4.7.1).

We now apply the multiscale step Proposition 4.3.4 to the matrix $A_N(\varepsilon, \lambda)$ (which represents $\mathcal{L}_{r,\mu,N}$ as stated in remark 4.2.2), for $N \geq N_0^2$, with $E = [-N, N]^b \times \mathfrak{I}$ and $N' \rightsquigarrow N$, $N \rightsquigarrow M$. The assumptions (4.3.3)-(4.3.6) hold, for all $\chi \in [\bar{\chi}, \bar{\chi}^2]$, by the choice of the constants $\bar{\chi}$, $\tau'$, $s_1$ at the beginning of section 4.5. Assumption (H1) holds by (4.2.13). Assumption (H2) holds because $\lambda \in \Lambda(\varepsilon; 1, X_{r,\mu}) \subset \mathtt{G}^0_{N,1}$ for $N \geq N_0^2$, see (4.5.12). Moreover, arguing as in Lemma 4.7.4 -for the matrix $A(\varepsilon, \theta, \lambda)$ with $\theta = 0$, $j_0 = 0$-, the hypothesis (H3) of Proposition 4.3.4 holds. Then the multiscale step Proposition 4.3.4 implies that, $\forall \lambda \in \Lambda(\varepsilon; 1, X_{r,\mu}) \subset \mathcal{G}_M \cap \mathtt{G}^0_{N,1}$, we have

$$|D_m^{-1/2} A_N^{-1} D_m^{-1/2}|_s \overset{(4.3.9)}{\leq} C(s) N^{\tau'} \big( N^{\varsigma s} + \varepsilon^2 |r|_{+,s} \big) \leq C(s) N^{\tau'} \big( N^{\varsigma s} + |r|_{+,s} \big) . \tag{4.7.20}$$

We claim the following direct consequence: on the set $\Lambda(\varepsilon; 1, X_{r,\mu})$ we have

$$\Big| D_m^{-1/2} \Big( \frac{\mathcal{L}_{r,\mu,N}}{1+\varepsilon^2 \lambda} \Big)^{-1} D_m^{-1/2} \Big|_{\mathrm{Lip},s} \leq C(s) N^{2(\tau' + \varsigma s_1 + 1)} \big( N^{\varsigma(s-s_1)} + |r|_{\mathrm{Lip},+,s} \big) . \tag{4.7.21}$$

For all $\lambda$ in the set $\Lambda(\varepsilon; 1, X_{r,\mu})$, the operator

$$U(\varepsilon, \lambda) := D_m^{-1/2} \Big( \frac{\mathcal{L}_{r,\mu,N}}{1+\varepsilon^2 \lambda} \Big)^{-1} D_m^{-1/2}$$

satisfies, by (4.7.20) and $|r|_{+,s_1} \leq C_1 \varepsilon^2$ (see item 1 of Definition 4.1.2), the estimates

$$|U(\varepsilon,\lambda)|_s \leq C(s) N^{\tau'} (N^{\varsigma s} + |r|_{+,s}) , \quad |U(\varepsilon,\lambda)|_{s_1} \leq C(s_1) N^{\tau' + \varsigma s_1} . \tag{4.7.22}$$

Moreover, for all $\lambda_1, \lambda_2 \in \Lambda(\varepsilon; 1, X_{r,\mu})$, we write (using that $\dfrac{\omega}{1+\varepsilon^2\lambda}$ is independent of $\lambda$)

$$\begin{aligned} \frac{U(\lambda_2) - U(\lambda_1)}{\lambda_2 - \lambda_1} &= -U(\lambda_2) \frac{U^{-1}(\lambda_2) - U^{-1}(\lambda_1)}{\lambda_2 - \lambda_1} U(\lambda_1) \\ &= -U(\lambda_2) D_m^{1/2} \frac{1}{\lambda_2 - \lambda_1} \Big( \frac{X_{r,\mu,N}(\lambda_2)}{1+\varepsilon^2\lambda_2} - \frac{X_{r,\mu,N}(\lambda_1)}{1+\varepsilon^2\lambda_1} \Big) D_m^{1/2} U(\lambda_1) \end{aligned} \tag{4.7.23}$$

where $X_{r,\mu,N} := \mathit{\Pi}_N (X_{r,\mu})_{|\mathcal{H}_N}$. Decomposing

$$\frac{1}{\lambda_2 - \lambda_1} \Big( \frac{X_{r,\mu,N}(\lambda_2)}{1+\varepsilon^2\lambda_2} - \frac{X_{r,\mu,N}(\lambda_1)}{1+\varepsilon^2\lambda_1} \Big) = \frac{X_{r,\mu,N}(\lambda_2) - X_{r,\mu,N}(\lambda_1)}{(\lambda_2 - \lambda_1)(1+\varepsilon^2\lambda_2)} - \frac{\varepsilon^2 X_{r,\mu,N}(\lambda_1)}{(1+\varepsilon^2\lambda_2)(1+\varepsilon^2\lambda_1)}$$

we deduce by (4.7.23), (4.7.22) and $\mu = O(1)$, $|\mu|_{\mathrm{lip}} = O(\varepsilon^2)$, $|r|_{+,s_1}, |r|_{\mathrm{lip},+,s_1} \leq C_1 \varepsilon^2$, the estimates

$$\begin{aligned} &|U|_{\mathrm{lip},s} \lesssim_s N^{2(\tau' + \varsigma s_1 + 2)} (N^{\varsigma(s-s_1)} + |r|_{+,s} + |r|_{\mathrm{lip},+,s}) , \\ &|U|_{\mathrm{lip},s_1} \lesssim_{s_1} N^{2(\tau' + \varsigma s_1 + 1)} . \end{aligned} \tag{4.7.24}$$

Finally (4.7.22) and (4.7.24) imply (4.7.21). The inequalities (3.3.28) and (4.7.21) imply (4.1.23).

We finally prove (4.1.24). Denoting $A'_N$ the matrix which represents $\mathcal{L}_{r',\mu',N}$ as in remark 4.2.2 we have that $A_N - A'_N$ represents $\mathit{\Pi}_N ((\mu - \mu') \mathcal{J} \Pi^\perp_{\mathbb{S} \cup \mathbb{F}} + r - r')_{|\mathcal{H}_N}$. Then it is enough to write

$$\begin{aligned} &\big| D_m^{-1/2} \big( A_N^{-1} - (A'_N)^{-1} \big) D_m^{-1/2} \big|_{s_1} \\ &\lesssim_{s_1} |D_m^{-1/2} A_N^{-1} D_m^{-1/2}|_{s_1} |A_N - A'_N|_{+,s_1} |D_m^{-1/2} (A'_N)^{-1} D_m^{-1/2}|_{s_1} \\ &\lesssim_{s_1} N^{2(\tau' + \varsigma s_1)} \big( |\mu - \mu'| N^2 + |r - r'|_{+,s_1} \big) \end{aligned}$$

using (4.7.20) at $s = s_1$ and the bounds $|r|_{+,s_1}, |r'|_{+,s_1} \lesssim \varepsilon^2$. This estimate and (3.3.28) imply (4.1.24). ■

## 4.8 Measure estimates

The aim of this section is to prove the measure estimates (4.1.17)-(4.1.18) in Proposition 4.1.5.

### Preliminaries

We first give several lemmas on basic properties of eigenvalues of self-adjoint matrices, which are a consequence of their variational characterization.

**Lemma 4.8.1.** *Let $A(\xi)$ be a family of self-adjoint matrices in $\mathcal{M}_E^E$, $E$ finite, defined for $\xi \in \widetilde{\Lambda} \subseteq \mathbb{R}$, satisfying, for some $\beta > 0$,*

$$\mathfrak{d}_\xi A(\xi) \geq \beta \mathrm{Id}$$

*(recall the notation (1.6.4)). We list the eigenvalues of $A(\xi)$ in non decreasing order*

$$\mu_1(\xi) \leq \ldots \leq \mu_q(\xi) \leq \ldots \leq \mu_{|E|}(\xi)$$

*according to their variational characterization*

$$\mu_q(\xi) := \inf_{F \in \mathcal{F}_q} \max_{y \in F, \|y\|_0 = 1} \langle A(\xi) y, y \rangle_0 \tag{4.8.1}$$

*where $\mathcal{F}_q$ is the set of all subspaces $F$ of $\mathbb{C}^{|E|}$ of dimension $q$. Then*

$$\mathfrak{d}_\xi \mu_q(\xi) \geq \beta > 0 \,, \quad \forall q = 1, \ldots, |E| \,. \tag{4.8.2}$$

Proof. By the assumption $\mathfrak{d}_\xi A(\xi) \geq \beta \mathrm{Id}$, we have, for all $\xi_2 > \xi_1$, $\xi_1, \xi_2 \in \widetilde{\Lambda}$, $y \in F$, $\|y\|_0 = 1$, that

$$\langle A(\xi_2) y, y \rangle_0 > \langle A(\xi_1) y, y \rangle_0 + \beta(\xi_2 - \xi_1) \,.$$

Therefore

$$\max_{y \in F, \|y\|_0 = 1} \langle A(\xi_2) y, y \rangle_0 \geq \max_{y \in F, \|y\|_0 = 1} \langle A(\xi_1) y, y \rangle_0 + \beta(\xi_2 - \xi_1)$$

and, by (4.8.1), for all $q = 1, \ldots, |E|$,

$$\mu_q(\xi_2) \geq \mu_q(\xi_1) + \beta(\xi_2 - \xi_1) \,.$$

Hence $\mathfrak{d}_\xi \mu_q(\xi) \geq \beta$. ∎

**Lemma 4.8.2.** *i) Let $A(\xi)$ be a family of self-adjoint matrices in $\mathcal{M}_E^E$, $E$ finite, Lipschitz with respect to $\xi \in \widetilde{\Lambda} \subseteq \mathbb{R}$, satisfying, for some $\beta > 0$,*

$$\mathfrak{d}_\xi A(\xi) \geq \beta \mathrm{Id} \,.$$

*Then there are intervals* $(I_q)_{1\le q\le |E|}$ *in* $\mathbb{R}$ *such that*

$$\Big\{\xi\in\widetilde\Lambda\,:\,\|A^{-1}(\xi)\|_0\ge\alpha^{-1}\Big\}\subseteq\bigcup_{1\le q\le|E|}I_q\,,\qquad |I_q|\le 2\alpha\beta^{-1}\,,\tag{4.8.3}$$

*and in particular the Lebesgue measure*

$$\Big|\Big\{\xi\in\widetilde\Lambda\,:\,\|A^{-1}(\xi)\|_0\ge\alpha^{-1}\Big\}\Big|\le 2|E|\alpha\beta^{-1}\,.\tag{4.8.4}$$

*ii) Let* $A(\xi):=Z+\xi W$ *be a family of self-adjoint matrices in* $\mathcal{M}_E^E$*, Lipschitz with respect to* $\xi\in\widetilde\Lambda\subseteq\mathbb{R}$*, with* $W$ *invertible and*

$$\beta_1\mathrm{Id}\le Z\le\beta_2\mathrm{Id}$$

*with* $\beta_1>0$*. Then there are intervals* $(I_q)_{1\le q\le|E|}$ *such that*

$$\Big\{\xi\in\widetilde\Lambda\,:\,\|A^{-1}(\xi)\|_0\ge\alpha^{-1}\Big\}\subseteq\bigcup_{1\le q\le|E|}I_q\,,\qquad |I_q|\le 2\alpha\beta_2\beta_1^{-1}\|W^{-1}\|_0\,.\tag{4.8.5}$$

Proof. Proof of $i$). Let $(\mu_q(\xi))_{1\le q\le|E|}$ be the eigenvalues of $A(\xi)$ listed as in Lemma 4.8.1. We have

$$\Big\{\xi\in\widetilde\Lambda\,:\,\|A^{-1}(\xi)\|_0\ge\alpha^{-1}\Big\}=\bigcup_{1\le q\le|E|}\Big\{\xi\in\widetilde\Lambda\,:\,\mu_q(\xi)\in[-\alpha,\alpha]\Big\}.$$

By (4.8.2) each

$$I_q:=\big\{\xi\in\widetilde\Lambda\,:\,\mu_q(\xi)\in[-\alpha,\alpha]\big\}$$

is included in an interval of length less than $2\alpha\beta^{-1}$.
Proof of $ii$). Let $U:=W^{-1}Z$ and consider the family of self-adjoint matrices

$$\tilde A(\xi):=A(\xi)U=(Z+\xi W)U=ZW^{-1}Z+\xi Z\,.$$

We have the inclusion

$$\Big\{\xi\in\widetilde\Lambda\,:\,\|A^{-1}(\xi)\|_0\ge\alpha^{-1}\Big\}\subset\Big\{\xi\in\widetilde\Lambda\,:\,\|\tilde A^{-1}(\xi)\|_0\ge(\alpha\|U\|_0)^{-1}\Big\}\,.\tag{4.8.6}$$

Since $\mathfrak{d}_\xi\tilde A(\xi)\ge\beta_1\mathrm{Id}$ we derive by item $i$) that

$$\Big\{\xi\in\widetilde\Lambda\,:\,\|\tilde A^{-1}(\xi)\|_0\ge(\alpha\|U\|_0)^{-1}\Big\}\subset\bigcup_{1\le q\le|E|}I_q\tag{4.8.7}$$

where $I_q$ are intervals with measure

$$|I_q|\le 2\alpha\|U\|_0\beta_1^{-1}\le 2\alpha\|W^{-1}\|_0\|Z\|_0\beta_1^{-1}\le 2\alpha\|W^{-1}\|_0\beta_2\beta_1^{-1}\,.\tag{4.8.8}$$

Then (4.8.6), (4.8.7), (4.8.8) imply (4.8.5). ■

The variational characterization of the eigenvalues also implies the following lemma.

**Lemma 4.8.3.** *Let $A$, $A_1$ be self adjoint matrices $\mathcal{M}_E^E$, $E$ finite. Then their eigenvalues, ranked in nondecreasing order, satisfy the Lipschitz property*

$$|\mu_q(A) - \mu_q(A_1)| \leq \|A - A_1\|_0\,, \quad \forall q = 1, \ldots, |E|\,.$$

We finish this section stating a simple perturbative lemma, proved by a Neumann series argument.

**Lemma 4.8.4.** *Let $B, B' \in \mathcal{M}_E^E$ with $E := [-N, N]^b \times \mathfrak{I}$, and assume that*

$$\|D_m^{-1/2} B^{-1} D_m^{-1/2}\|_0 \leq K_1\,, \quad \|B' - B\|_0 \leq \alpha\,. \tag{4.8.9}$$

*If $4(N^2 + m)^{1/2}\alpha K_1 \leq 1$, then*

$$\|D_m^{-1/2} (B')^{-1} D_m^{-1/2}\|_0 \leq K_1 + 4(N^2 + m)^{1/2}\alpha K_1^2\,. \tag{4.8.10}$$

**Measure estimate of $\Lambda \setminus \tilde{\mathcal{G}}$**

We estimate the complementary set $\Lambda \setminus \tilde{\mathcal{G}}$ where $\tilde{\mathcal{G}}$ is the set defined in (4.5.8).

**Lemma 4.8.5.** $|\Lambda \setminus \tilde{\mathcal{G}}| \leq \varepsilon^2$.

Proof. Since $\bar{\omega}_\varepsilon := \bar{\omega}$ satisfies (1.2.30) we have

$$\begin{aligned}\big|n + (1 + \varepsilon^2\lambda)^2 \sum_{1 \leq i \leq j \leq |\mathbb{S}|} p_{ij}\bar{\omega}_i\bar{\omega}_j\big| &\geq \big|n + \sum_{1 \leq i \leq j \leq |\mathbb{S}|} p_{ij}\bar{\omega}_i\bar{\omega}_j\big| - C|\lambda|\varepsilon^2|p| \\ &\geq \frac{\gamma_1}{\langle p\rangle^{\tau_1}} - C'\varepsilon^2|p| \geq \frac{\gamma_1/2}{\langle p\rangle^{\tau_1}}\end{aligned}$$

for all $|p| \leq \big(\frac{\gamma_1}{2C'\varepsilon^2}\big)^{\frac{1}{\tau_1+1}}$. Moreover, if $n = 0$ then, for all $\lambda \in \Lambda$, we have

$$\big|(1 + \varepsilon^2\lambda)^2 \sum_{1 \leq i \leq j \leq |\mathbb{S}|} p_{ij}\bar{\omega}_i\bar{\omega}_j\big| \geq (1 - c\varepsilon^2)^2 \frac{\gamma_1}{\langle p\rangle^{\tau_1}}\,.$$

As a consequence, recalling the definition of $\tilde{\mathcal{G}}$ in (4.5.8), Definition 4.1.4 and $\gamma_2 := \gamma_1/2$, we have

$$\Lambda \setminus \tilde{\mathcal{G}} \subset \bigcup_{(n,p) \in \mathcal{N}} \mathcal{R}_{n,p}\,, \quad \mathcal{N} := \Big\{n \neq 0\,,\ |p| > \Big(\frac{\gamma_1}{2C'\varepsilon^2}\Big)^{\frac{1}{\tau_1+1}}\,,\ |n| \leq C|p|\Big\}\,, \tag{4.8.11}$$

where

$$\mathcal{R}_{n,p} := \Big\{\lambda \in \Lambda\,:\, |f_{n,p}(\lambda)| < \frac{2\gamma_2}{\langle p\rangle^{\tau_2}}\Big\}, \quad f_{n,p}(\lambda) := \frac{n}{(1 + \varepsilon^2\lambda)^2} + \sum_{1 \leq i \leq j \leq |\mathbb{S}|} p_{ij}\bar{\omega}_i\bar{\omega}_j\,.$$

Since $\partial_\lambda f_{n,p}(\lambda) := -\dfrac{2n\varepsilon^2}{(1+\varepsilon^2\lambda)^3}$, if $|n| \neq 0$, we have $|\mathcal{R}_{n,p}| \lesssim \varepsilon^{-2}\gamma_2 \langle p\rangle^{-\tau_2}$ and by (4.8.11)

$$|\Lambda \setminus \tilde{\mathcal{G}}| \lesssim_{\gamma_0} \sum_{|p| > \left(\frac{\gamma_1}{2C'\varepsilon^2}\right)^{\frac{1}{\tau_1+1}}} |p| \frac{\varepsilon^{-2}}{|p|^{\tau_2}} \lesssim_{\gamma_0} \varepsilon^{-2} \int_{\left(\frac{\gamma_1}{2C'\varepsilon^2}\right)^{\frac{1}{\tau_1+1}}}^{+\infty} \rho^{-\tau_2+\frac{1}{2}|\mathbb{S}|(|\mathbb{S}|-1)} d\rho$$

$$\lesssim_{\gamma_0} \varepsilon^{\frac{2(\tau_2-(|\mathbb{S}|(|\mathbb{S}|-1)/2)-2-\tau_1)}{\tau_1+1}} \leq \varepsilon^2$$

for $\tau_2$ large with respect to $\tau_1$, i.e. $\tau_2$ defined as in (4.5.9), and $\varepsilon$ small. ■

**Measure estimate of $\widetilde{\Lambda} \setminus \mathtt{G}^0_{N,\frac{1}{2}}$ for $N \geq N_0^2$**

We estimate the complementary set $\widetilde{\Lambda} \setminus \mathtt{G}^0_{N,\frac{1}{2}}$ where $\mathtt{G}^0_{N,\eta}$ is defined in (4.5.6).

**Lemma 4.8.6.** *If $N \geq N_0^2$ then*

$$|\widetilde{\Lambda} \setminus \mathtt{G}^0_{N,\frac{1}{2}}| \leq \varepsilon^{-2} N^{-\tau+d+|\mathbb{S}|+1} \leq N^{-\frac{\tau}{2}+2d+|\mathbb{S}|+s_0} . \tag{4.8.12}$$

PROOF. By (4.5.6) we have the inclusion

$$\widetilde{\Lambda} \setminus \mathtt{G}^0_{N,\frac{1}{2}} \subset \left\{\lambda \in \widetilde{\Lambda} \,:\, \|P_N^{-1}(\lambda)\|_0 > N^\tau/4\right\}$$

where (see remark 4.2.2)

$$P_N(\lambda) := D_m^{1/2} \frac{A_N(\varepsilon,\lambda)}{1+\varepsilon^2\lambda} D_m^{1/2} = D_m^{1/2} \mathit{\Pi}_N \Big[ J\bar{\omega}_\varepsilon \cdot \partial_\varphi + \frac{X_{r,\mu}(\varepsilon,\lambda)}{1+\varepsilon^2\lambda} \Big]_{|\mathcal{H}_N} D_m^{1/2} .$$

By assumption 3 of Definition 4.1.2, the matrix $P_N(\lambda)$ satisfies $\mathfrak{d}_\lambda P_N(\lambda) \leq -c\varepsilon^2$ with $c := c_2\sqrt{m}$. The first inequality in (4.8.12) follows by Lemma 4.8.2-*i*) (in particular (4.8.4)) applied to $-P_N(\lambda)$ with $E = [-N,N]^{d+|\mathbb{S}|} \times \mathfrak{I}$, $\alpha = 4N^{-\tau}$, $\beta = c\varepsilon^2$, taking $N$ large. The second inequality in (4.8.12) follows because, by (4.5.10),

$$\varepsilon^{-2} \leq N_0^{\tau+d+s_0} \leq N^{\frac{\tau+d+s_0}{2}}$$

for $N \geq N_0^2$. ■

**Measure estimate of $\widetilde{\Lambda} \setminus \mathcal{G}^0_{N_k,\frac{1}{2}}$ for $k \geq 1$**

We estimate the complementary set $\widetilde{\Lambda} \setminus \mathcal{G}^0_{N_k,\frac{1}{2}}$ where $\mathcal{G}^0_{N,\eta}$ is defined in (4.5.5).

**Proposition 4.8.7.** *For all* $N \geq N_1 := N_0^{\bar{\chi}}$ *the set* $\mathcal{B}^0_{N,\frac{1}{2}} := \widetilde{\Lambda} \setminus \mathcal{G}^0_{N,\frac{1}{2}}$ *has measure*

$$|\mathcal{B}^0_{N,\frac{1}{2}}| \leq N^{-1} . \tag{4.8.13}$$

We first obtain complexity estimates for the set $B^0_N(j_0; \lambda, 1/2)$ defined in (4.5.4). We argue differently for $|j_0| \geq 3(1+|\bar{\mu}|)N$ and $|j_0| < 3(1+|\bar{\mu}|)N$.

**Lemma 4.8.8.** *For all* $\lambda \in \widetilde{\Lambda}$*, for all* $|j_0| \geq 3(1+|\bar{\mu}|)N$*, we have*

$$B^0_N(j_0; \lambda, 1/2) \subset \bigcup_{q=1}^{N^{d+|\mathbb{S}|+1}} I_q \tag{4.8.14}$$

*where* $|I_q|$ *are intervals with length* $|I_q| \leq N^{-\tau}$*.*

Proof. Recalling (4.2.22), (4.2.14) and (4.2.15), we have

$$D_m^{1/2} A_{N,j_0}(\varepsilon, \lambda, \theta) D_m^{1/2} = D_m^{1/2} A_{N,j_0}(\varepsilon, \lambda) D_m^{1/2} + \theta\, D_m^{1/2} Y_{N,j_0} D_m^{1/2} . \tag{4.8.15}$$

We claim that, if $|j_0| \geq 3(1+|\bar{\mu}|)N$ and $N \geq \bar{N}(V, d, |\mathbb{S}|)$ is large, then

$$\frac{|j_0|^2}{10} \mathrm{Id} \leq D_m^{1/2} A_{N,j_0}(\varepsilon, \lambda) D_m^{1/2} \leq 4|j_0|^2 \mathrm{Id} . \tag{4.8.16}$$

Indeed by (4.2.12), (4.2.18)-(4.2.19) the eigenvalues $\nu_{\ell,j}$ of $A_{N,j_0}(\varepsilon, \lambda)$ satisfy

$$\nu_{\ell,j} = \delta^{\pm}_{\ell,j} + O(\|D_m^{1/2} T_{N,j_0}(\varepsilon, \lambda) D_m^{1/2}\|_0) \quad \text{where} \quad \delta^{\pm}_{\ell,j} := \langle j \rangle_m \big( \langle j \rangle_m \pm \omega \cdot \ell \pm \mu \big) . \tag{4.8.17}$$

Since $|\ell| \leq N$ and $|j| \geq |j_0| - N$ we see that, for $|j_0| \geq 3(1+|\bar{\mu}|)N$, for $N \geq \bar{N}(V, d, |\mathbb{S}|)$ large enough,

$$\frac{2|j_0|^2}{9} \leq \delta^{\pm}_{\ell,j} \leq 3|j_0|^2 . \tag{4.8.18}$$

Hence (4.8.17), (4.8.18) and (4.2.13) imply (4.8.16). As a consequence, Lemma 4.8.2-*ii*) applied to the matrix in (4.8.15) with $Z = D_m^{1/2} A_{N,j_0}(\varepsilon, \lambda) D_m^{1/2}$, $\alpha = 2N^{-\tau}$, $\beta_1 = |j_0|^2/10$, $\beta_2 \leq 4|j_0|^2$, $W = D_m^{1/2} Y_{N,j_0} D_m^{1/2}$, $\|W^{-1}\|_0 \leq C$, imply that

$$B^0_N(j_0; \lambda, 1/2) \subset \bigcup_{q=1}^{4(2N+1)^{d+|\mathbb{S}|}} I_q \qquad \text{with} \qquad |I_q| \leq C N^{-\tau} .$$

Dividing further these intervals we obtain (4.8.14).■

We now consider the case $|j_0| \leq 3(1+|\bar{\mu}|)N$. We can no longer argue directly as in Lemma 4.8.8. In this case the aim is to bound the measure of

$$B^0_{2,N}(j_0; \lambda) := \Big\{ \theta \in \mathbb{R} \, : \, \|D_m^{-1/2} A^{-1}_{N,j_0}(\varepsilon, \lambda, \theta) D_m^{-1/2}\|_0 > N^{\tau}/4 \Big\} . \tag{4.8.19}$$

The continuity property of the eigenvalues (Lemma 4.8.3) allows then to derive a complexity estimate for $B^0_N(j_0;\lambda,1/2)$ in terms of the measure $|B^0_{2,N}(j_0;\lambda)|$ (Lemma 4.8.10). Lemma 4.8.11 is devoted to the estimate of the bi-dimensional Lebesgue measure

$$\Big|\Big\{(\lambda,\theta)\in\widetilde{\Lambda}\times\mathbb{R}\,:\,\theta\in B^0_{2,N}(j_0;\lambda)\Big\}\Big|\,.$$

Such an estimate is then used in Lemma 4.8.12 to justify that the measure of the section $|B^0_{2,N}(j_0;\lambda)|$ has an appropriate bound for "most" $\lambda$ (by a Fubini type argument).

We first show that, for $|j_0|\le 3(1+|\bar\mu|)N$, the set $B^0_{2,N}(j_0;\lambda)$ is contained in an interval of size $O(N)$ centered at the origin.

**Lemma 4.8.9.** *$\forall|j_0|<3(1+|\bar\mu|)N$, $\forall\lambda\in\widetilde{\Lambda}$, we have*

$$B^0_{2,N}(j_0;\lambda)\subset I_N:=\big(-5(1+|\bar\mu|)N,5(1+|\bar\mu|)N\big)\,.$$

PROOF. The eigenvalues $\nu_{\ell,j}(\theta)$ of $D_m^{1/2}A_{N,j_0}(\varepsilon,\lambda,\theta)D_m^{1/2}$ satisfy

$$\begin{aligned}&\nu_{\ell,j}(\theta)=\delta^{\pm}_{\ell,j}(\theta)+O(|T|_{+,s_1})\\ &\text{where}\quad \delta^{\pm}_{\ell,j}:=\langle j\rangle_m\big(\langle j\rangle_m\pm(\omega\cdot\ell+\theta)\pm\mu\big)\,.\end{aligned}\tag{4.8.20}$$

If $|\theta|\ge 5(1+|\bar\mu|)N$ then, using also (4.2.13), each eigenvalue satisfies $|\nu_{\ell,j}(\theta)|\ge 1$, and therefore $\theta$ belongs to the complementary of the set $B^0_{2,N}(j_0;\lambda)$ defined in (4.8.19). ■

**Lemma 4.8.10.** *$\forall|j_0|\le 3(1+|\bar\mu|)N$, $\forall\lambda\in\widetilde{\Lambda}$, we have*

$$B^0_N(j_0;\lambda,1/2)\subset\bigcup_{q=1,\ldots,[\hat C\mathtt{M}N^{\tau+1}]}I_q$$

*where $I_q$ are intervals with $|I_q|\le N^{-\tau}$ and $\mathtt{M}:=|B^0_{2,N}(j_0;\lambda)|$.*

PROOF. Suppose that $\theta\in B^0_N(j_0;\lambda,1/2)$ where $B^0_N(j_0;\lambda,\eta)$ is defined in (4.5.4). Then there exists an eigenvalue of $D_m^{1/2}A_{N,j_0}(\varepsilon,\lambda,\theta)D_m^{1/2}$ with modulus less than $2N^{-\tau}$. Now, by (4.8.15), and since $|j_0|\le 3(1+|\bar\mu|)N$, we have

$$\begin{aligned}\|D_m^{1/2}\big(A_{N,j_0}(\varepsilon,\lambda,\theta+\Delta\theta)-A_{N,j_0}(\varepsilon,\lambda,\theta)\big)D_m^{1/2}\|_0&=|\Delta\theta|\|D_m^{1/2}Y_{N,j_0}D_m^{1/2}\|_0\\ &\le|\Delta\theta|\,5(1+|\bar\mu|)N\,.\end{aligned}$$

Hence, by Lemma 4.8.3, if $5(1+|\bar\mu|)N|\Delta\theta|\le N^{-\tau}$ then $\theta+\Delta\theta\in B^0_{2,N}(j_0;\lambda)$ because $A_{N,j_0}(\varepsilon,\lambda,\theta+\Delta\theta)$ has an eigenvalue with modulus less than $4N^{-\tau}$. Hence

$$[\theta-cN^{-(\tau+1)},\theta+cN^{-(\tau+1)}]\subset B^0_{2,N}(j_0;\lambda)\,.$$

Therefore $B_N^0(j_0;\lambda,1/2)$ is included in an union of intervals $J_m$ with disjoint interiors,

$$B_N^0(j_0;\lambda,1/2) \subset \bigcup_m J_m \subset B_{2,N}^0(j_0;\lambda), \quad \text{with length} \quad |J_m| \geq 2cN^{-(\tau+1)} \tag{4.8.21}$$

(if some of the intervals $[\theta - cN^{-(\tau+1)}, \theta + cN^{-(\tau+1)}]$ overlap, then we glue them together). We decompose each $J_m$ as an union of (non overlapping) intervals $I_q$ of length between $cN^{-(\tau+1)}/2$ and $cN^{-(\tau+1)}$. Then, by (4.8.21), we get a new covering

$$\begin{aligned} & B_N^0(j_0;\lambda,1/2) \subset \bigcup_{q=1,\ldots,Q} I_q \subset B_{2,N}^0(j_0;\lambda) \\ & \text{with} \;\; cN^{-(\tau+1)}/2 \leq |I_q| \leq cN^{-(\tau+1)} \leq N^{-\tau} \end{aligned}$$

and, since the intervals $I_q$ do not overlap,

$$QcN^{-(\tau+1)}/2 \leq \sum_{q=1}^{Q} |I_q| \leq |B_{2,N}^0(j_0;\lambda)| =: \mathtt{M}\,.$$

As a consequence $Q \leq \hat{C}\,\mathtt{M}\,N^{\tau+1}$, proving the lemma. ■

In the next lemma we use the crucial sign condition assumption 3 of Definition 4.1.2.

**Lemma 4.8.11.** $\forall |j_0| < 3(1+|\bar{\mu}|)N$, *the set*

$$\begin{aligned} & \mathbf{B}_{2,N}^0(j_0) := \mathbf{B}_{2,N}^0(j_0;\varepsilon) := \\ & \Big\{(\lambda,\theta) \in \widetilde{\Lambda} \times \mathbb{R} \,:\, \Big\| D_m^{-1/2} A_{N,j_0}^{-1}(\varepsilon,\lambda,\theta) D_m^{-1/2} \Big\|_0 > N^\tau/4 \Big\} \end{aligned} \tag{4.8.22}$$

*has measure*

$$|\mathbf{B}_{2,N}^0(j_0)| \leq C\varepsilon^{-2} N^{-\tau+d+|\mathbb{S}|+1}\,. \tag{4.8.23}$$

Proof. By Lemma 4.8.9, the set $\mathbf{B}_{2,N}^0(j_0) \subset \widetilde{\Lambda} \times I_N$. In order to estimate the "bad" $(\lambda,\theta)$ where at least one eigenvalue of $D_m^{1/2} A_{N,j_0}(\varepsilon,\lambda,\theta) D_m^{1/2}$ has modulus less than $4N^{-\tau}$, we introduce the variable

$$\vartheta := \frac{\theta}{1+\varepsilon^2\lambda} \quad \text{where} \quad \vartheta \in 2I_N\,, \tag{4.8.24}$$

and we consider the self adjoint matrix (recall that $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$)

$$\begin{aligned} P_{N,j_0}(\lambda) &:= D_m^{1/2} \frac{A_{N,j_0}(\varepsilon,\lambda,\theta)}{1+\varepsilon^2\lambda} D_m^{1/2} \\ &= D_m^{1/2} \Big( J\bar{\omega}_\varepsilon \cdot \partial_\varphi + \frac{[X_{r,\mu}(\varepsilon,\lambda)]_{N,j_0}}{1+\varepsilon^2\lambda} + \vartheta Y_{N,j_0} \Big) D_m^{1/2}\,. \end{aligned} \tag{4.8.25}$$

By the assumption 3 of Definition 4.1.2 we get

$$\mathfrak{d}_\lambda P_{N,j_0}(\lambda) \leq -c\varepsilon^2\,, \quad c := c_2\sqrt{m}\,.$$

By Lemma 4.8.2-*i*), for each fixed $\vartheta$, the set of $\lambda \in \widetilde{\Lambda}$ such that at least one eigenvalue is $\leq 4N^{-\tau}$ has measure at most $O(\varepsilon^{-2}N^{-\tau+d+|\mathbb{S}|})$. Then, integrating on $\vartheta \in I_N$, whose length is $|I_N| = O(N)$, we deduce (4.8.23). ∎

As a consequence of Lemma 4.8.11 for "most" $\lambda$ the measure of $B^0_{2,N}(j_0;\lambda)$ is "small".

**Lemma 4.8.12.** $\forall |j_0| < 3(1+|\bar{\mu}|)N$, *the set*

$$\mathcal{F}_N(j_0) := \Big\{\lambda \in \widetilde{\Lambda}\ :\ |B^0_{2,N}(j_0;\lambda)| \geq \varepsilon^{-2}\hat{C}^{-1}N^{-\tau+2d+|\mathbb{S}|+3}\Big\} \tag{4.8.26}$$

*where* $\hat{C}$ *is the positive constant of Lemma 4.8.10, has measure*

$$|\mathcal{F}_N(j_0)| \leq CN^{-d-2}\,. \tag{4.8.27}$$

Proof. By Fubini theorem, recalling (4.8.22) and (4.8.19), we have

$$|\mathbf{B}^0_{2,N}(j_0)| = \int_{\widetilde{\Lambda}} |B^0_{2,N}(j_0;\lambda)|\,d\lambda\,. \tag{4.8.28}$$

Let $\mu := \tau - 2d - |\mathbb{S}| - 3$. By (4.8.28) and (4.8.23),

$$\begin{aligned} C\varepsilon^{-2}N^{-\tau+d+|\mathbb{S}|+1} &\geq \int_{\widetilde{\Lambda}} |B^0_{2,N}(j_0;\lambda)|\,d\lambda \\ &\geq \varepsilon^{-2}\hat{C}^{-1}N^{-\mu}\Big|\Big\{\lambda \in \widetilde{\Lambda}\ :\ |B^0_{2,N}(j_0;\lambda)| \geq \varepsilon^{-2}\hat{C}^{-1}N^{-\mu}\Big\}\Big| \\ &:= \varepsilon^{-2}\hat{C}^{-1}N^{-\mu}|\mathcal{F}_N(j_0)| \end{aligned}$$

whence (4.8.27). ∎

As a corollary we get

**Lemma 4.8.13.** *Let* $N \geq N_1 := [N_0^{\bar{\chi}}]$, *see* (4.5.11). *Then* $\forall |j_0| < 3(1+|\bar{\mu}|)N$, $\forall \lambda \notin \mathcal{F}_N(j_0)$, *we have*

$$B^0_N(j_0;\lambda,1/2) \subset \bigcup_{q=1,\ldots,N^{2d+|\mathbb{S}|+5}} I_q \tag{4.8.29}$$

*with* $I_q$ *intervals satisfying* $|I_q| \leq N^{-\tau}$.

Proof. By the definition of $\mathcal{F}_N(j_0)$ in (4.8.26), for all $\lambda \notin \mathcal{F}_N(j_0)$, we have

$$|B^0_{2,N}(j_0;\lambda)| < \varepsilon^{-2}\hat{C}^{-1}N^{-\tau+2d+|\mathbb{S}|+3}\,.$$

Then Lemma 4.8.10 implies that $\forall |j_0| < 3(1+|\bar\mu|)N$,

$$B^0_N(j_0;\lambda,1/2) \subset \bigcup_{q=1,\dots,\varepsilon^{-2}N^{2d+|\mathbb{S}|+4}} I_q\,.$$

For all $N \geq N_1 = [N_0^{\bar\chi}]$ we have

$$\varepsilon^{-2}N^{2d+|\mathbb{S}|+4} \stackrel{(4.5.10)}{\leq} N_0^{\tau+s_0+d}N^{2d+|\mathbb{S}|+4} \leq CN^{\frac{\tau+s_0+d}{\bar\chi}+2d+|\mathbb{S}|+4} \leq N^{2d+|\mathbb{S}|+5}$$

by (4.5.2). This proves (4.8.29). ■

Proof of Proposition 4.8.7 concluded. By Lemmata 4.8.8 and 4.8.13, for all $N \geq N_1$, $\lambda \in \widetilde\Lambda$,

$$\lambda \notin \bigcup_{|j_0|<3(1+|\bar\mu|)N} \mathcal{F}_N(j_0) \quad \Longrightarrow \quad \lambda \in \mathcal{G}^0_{N,\frac12}$$

(see the definition of $\mathcal{G}^0_{N,\frac12}$ in (4.5.5)) and therefore

$$\mathcal{B}^0_{N,\frac12} \subseteq \bigcup_{|j_0|<3(1+|\bar\mu|)N} \mathcal{F}_N(j_0)\,. \tag{4.8.30}$$

In conclusion, (4.8.30) and (4.8.27) imply that, for $N \geq N_1$,

$$|\mathcal{B}^0_{N,\frac12}| \leq \sum_{|j_0|<3(1+|\bar\mu|)N} |\mathcal{F}_N(j_0)| \lesssim N^d N^{-d-2} \leq N^{-1}\,.$$

### Stability of the $L^2$-good parameters under variation of $X_{r,\mu}$

In order to prove (4.1.18) we prove the "stability" of the sets $\mathtt{G}^0_{N,\eta} := \mathtt{G}^0_{N,\eta}(X_{r,\mu})$ and $\mathcal{G}^0_{N,\eta} := \mathcal{G}^0_{N,\eta}(X_{r,\mu})$ defined respectively in (4.5.6) and (4.5.5) with respect to small variations of the operator $X_{r,\mu}$.

**Lemma 4.8.14.** *Assume* $|r'-r|_{+,s_1} + |\mu'-\mu| \leq \delta$.
*i) If* $N^{\tau+1}\sqrt\delta$ *is small enough, then, for* $(1/2)+\sqrt\delta \leq \eta \leq 1$,

$$\mathtt{G}^0_{N,\eta-\sqrt\delta}(X_{r,\mu}) \cap \widetilde\Lambda' \subset \mathtt{G}^0_{N,\eta}(X_{r',\mu'})\,. \tag{4.8.31}$$

*ii) If* $N_k^{\tau+1}\sqrt\delta$, $k \geq 1$, *is small enough, then, for* $(1/2)+\sqrt\delta \leq \eta \leq 1$,

$$\mathcal{G}^0_{N_k,\eta-\sqrt\delta}(X_{r,\mu}) \cap \widetilde\Lambda' \subset \mathcal{G}^0_{N_k,\eta}(X_{r',\mu'})\,. \tag{4.8.32}$$

PROOF. Call $A_N(\varepsilon,\lambda)$, resp. $A'_N(\varepsilon,\lambda)$, the truncated operator associated to $\mathcal{L}_{r,\mu}$, resp. $\mathcal{L}_{r',\mu'}$, see remark 4.2.2, defined for $\lambda \in \widetilde{\Lambda}$, resp. $\lambda \in \widetilde{\Lambda}'$. By assumption, for any $\lambda \in \widetilde{\Lambda} \cap \widetilde{\Lambda}'$ we have

$$\|A_N(\varepsilon,\lambda) - A'_N(\varepsilon,\lambda)\|_0 \lesssim \|r - r'\|_0 + |\mu - \mu'| \le C\delta \,. \tag{4.8.33}$$

Proof of $i$). Assume that $\lambda \in \mathtt{G}^0_{N,\eta-\sqrt{\delta}}(X_{r,\mu}) \cap \widetilde{\Lambda}'$ where $\mathtt{G}^0_{N,\eta} := \mathtt{G}^0_{N,\eta}(X_{r,\mu})$ is defined in (4.5.6). Then

$$\|D_m^{-1/2} A_N^{-1}(\varepsilon,\lambda) D_m^{-1/2}\|_0 \le (\eta - \sqrt{\delta}) N^\tau \,. \tag{4.8.34}$$

Now we apply Lemma 4.8.4 to $B = A_N(\varepsilon,\lambda)$, $B' = A'_N(\varepsilon,\lambda)$. By (4.8.34), (4.8.33), the assumption (4.8.9) holds with $K_1 = (\eta - \sqrt{\delta})N^\tau$ and $\alpha = C\delta$. If $\delta N^{\tau+1}$ is small enough then (4.8.10) applies, and we deduce

$$\|D_m^{-1/2} (A'_N)^{-1}(\varepsilon,\lambda) D_m^{-1/2}\|_0 \le (\eta - \sqrt{\delta}) N^\tau + 4C\delta N^{2\tau+1} \le \eta N^\tau$$

provided that $\sqrt{\delta} N^{\tau+1} \le 1/(4C)$. Hence $\lambda \in \mathtt{G}^0_{N,\eta}(X_{r',\mu'})$, proving (4.8.31).

Proof of $ii$). Assume that $\lambda \in \mathcal{G}^0_{N_k,\eta-\sqrt{\delta}}(X_{r,\mu}) \cap \widetilde{\Lambda}'$ where $\mathcal{G}^0_{N,\eta} := \mathcal{G}^0_{N,\eta}(X_{r,\mu})$ is defined in (4.5.5). Let $B^0_{N_k}(j_0;\lambda,\eta)$, resp. $(B'_{N_k})^0(j_0;\lambda,\eta)$, be the the set defined in (4.5.4) corresponding to $X_{r,\mu}$, resp. $X_{r',\mu'}$, at $N = N_k$. Applying the same perturbative argument of item $i$) to the matrices $D_m^{-1/2} A_{N_k,j_0}(\varepsilon,\lambda,\theta) D_m^{-1/2}$ and $D_m^{-1/2} A'_{N_k,j_0}(\varepsilon,\lambda,\theta) D_m^{-1/2}$, we prove that, if $N_k^{\tau+1}\sqrt{\delta}$ is small enough, then, for all $|j_0| \le 3(1 + |\bar{\mu}|)N_k$, we have the inclusion

$$(B'_{N_k})^0(j_0;\lambda,\eta) \subset B^0_{N_k}(j_0;\lambda,\eta - \sqrt{\delta}) \,.$$

Hence, by Lemma 4.8.8, we have $\lambda \in \mathcal{G}^0_{N_k,\eta}(X_{r',\mu'})$, proving (4.8.32). ■

**Conclusion: proof of** (4.1.17)-(4.1.18)

We finally prove that the sets $\Lambda(\varepsilon;\eta,X_{r,\mu})$, $\eta \in [1/2,1]$, defined in (4.5.12) satisfy the measure estimates (4.1.17)-(4.1.18).

PROOF OF (4.1.17). We have to estimate the measure of the complementary set

$$\Lambda(\varepsilon;1/2,X_{r,\mu})^c \cap \widetilde{\Lambda} = \bigcup_{k\ge1} \mathcal{B}^0_{N_k,\frac12} \bigcup_{N \ge N_0^2} \big(\mathtt{G}^0_{N,\frac12}\big)^c \bigcup \tilde{\mathcal{G}}^c \cap \widetilde{\Lambda} \tag{4.8.35}$$

where $\mathcal{B}^0_{N_k,\frac12} = \Lambda \setminus \mathcal{G}^0_{N_k,\frac12}$ with $\mathcal{G}^0_{N,\eta}$ defined in (4.5.5), the set $\mathtt{G}^0_{N,\eta}$ is defined in (4.5.6), and $\tilde{\mathcal{G}}$ in (4.5.8).

**Lemma 4.8.15.** $\Big|\bigcup_{k\ge1} \mathcal{B}^0_{N_k,\frac12}\Big| \le \varepsilon^2 \,.$

Proof. By Proposition 4.8.7 we have

$$\Big|\bigcup_{k\geq 1}\mathcal{B}^0_{N_k,\frac{1}{2}}\Big| \leq \sum_{k\geq 1} N_k^{-1} \lesssim N_1^{-1} \overset{(4.5.11)}{\lesssim} N_0^{-\bar{\chi}} \overset{(4.5.10)}{\lesssim} \varepsilon^{\frac{2\bar{\chi}}{\tau+s_0+d}} \leq \varepsilon^2$$

since $\bar{\chi}$ is large according to the second inequality in (4.5.2). ■

**Lemma 4.8.16.** $\Big|\bigcup_{N\geq N_0^2}(\mathtt{G}^0_{N,\frac{1}{2}})^c\cap\widetilde{\Lambda}\Big| \leq \dfrac{\varepsilon}{3}$.

Proof. By Lemma 4.8.6 we have

$$\begin{aligned}\Big|\bigcup_{N\geq N_0^2}\big(\mathtt{G}^0_{N,\frac{1}{2}}\big)^c\cap\widetilde{\Lambda}\Big| \leq \sum_{N\geq N_0^2} N^{-\frac{\tau}{2}+2d+2|\mathbb{S}|+s_0} &\lesssim N_0^{-\tau+4d+4|\mathbb{S}|+2s_0+2}\\ &\overset{(4.5.10)}{\lesssim} \varepsilon^{2\frac{(\tau-4d-4|\mathbb{S}|-2s_0-2)}{\tau+s_0+d}}\\ &\leq \frac{\varepsilon}{3}\end{aligned}$$

since $\tau$ is large according to (4.5.1). ■

By Lemmata 4.8.15, 4.8.16, 4.8.5 we deduce that the complementary set (4.8.35) has measure

$$\begin{aligned}|\Lambda(\varepsilon;1/2,X_{r,\mu})^c\cap\widetilde{\Lambda}| &\leq \Big|\bigcup_{k\geq 1}\mathcal{B}^0_{N_k,\frac{1}{2}}\Big| + \Big|\bigcup_{N\geq N_0^2}\big(\mathtt{G}^0_{N,\frac{1}{2}}\big)^c\cap\widetilde{\Lambda}\Big| + |\tilde{\mathcal{G}}^c\cap\widetilde{\Lambda}|\\ &\leq 2\varepsilon^2+\frac{\varepsilon}{3}\leq\varepsilon\end{aligned}$$

proving (4.1.17).

**Remark 4.8.17.** *We could prove that the measure of $\Lambda(\varepsilon;1/2,X_{r,\mu})^c$ is smaller than $\varepsilon^p$, for any $p$, optimizing the choice of the constants. Indeed the measure of the set of Lemma 4.8.5, respectively 4.8.15, decreases taking the constant $\tau_2$ in (4.5.9), respectively $\bar{\chi}$ in (4.5.2), larger. The set in Lemma 4.8.16 has measure $\varepsilon^{2-\alpha}$ as the constant $\tau$ defined in (4.5.1) increases. If we had intersected in (4.5.12) for $N\geq N_0^{\alpha(\tau)}$, as explained in remark 4.5.3, the new set in Lemma 4.8.16 would have arbitrarily small measure as well.*

Proof of (4.1.18). Recalling the definition of the sets $\Lambda(\varepsilon;\eta,X_{r,\mu})$, $\eta\in[1/2,1]$, in (4.5.12) and Lemma 4.8.14, we have that, for $N\geq\bar{N}$, $(1/2)+\sqrt{\delta}\leq\eta\leq 1$,

$$\begin{aligned}N^{\tau+2}\sqrt{\delta}<1 &\quad\overset{(4.8.31)}{\Longrightarrow}\quad \Lambda(\varepsilon;\eta-\sqrt{\delta},X_{r,\mu})\cap\mathtt{G}^0_{N,\eta}(X_{r',\mu'})^c\cap\widetilde{\Lambda}'=\emptyset\\ N_k^{\tau+2}\sqrt{\delta}<1 &\quad\overset{(4.8.32)}{\Longrightarrow}\quad \Lambda(\varepsilon;\eta-\sqrt{\delta},X_{r,\mu})\cap\mathcal{G}^0_{N_k,\eta}(X_{r',\mu'})^c\cap\widetilde{\Lambda}'=\emptyset\,.\end{aligned}$$

Hence

$$\begin{aligned}
&\Lambda(\varepsilon;\eta-\sqrt{\delta},X_{r,\mu})\cap\Lambda(\varepsilon;\eta,X_{r',\mu'})^c\cap\widetilde{\Lambda}' \\
&\subset\bigcup_{k\geq 1}\Big(\Lambda(\varepsilon;\eta-\sqrt{\delta},X_{r,\mu})\cap\mathcal{G}^0_{N_k,\eta}(X_{r',\mu'})^c\Big)\cap\widetilde{\Lambda}' \\
&\quad\bigcup_{N\geq N_0^2}\Big(\Lambda(\varepsilon;\eta-\sqrt{\delta},X_{r,\mu})\cap\mathtt{G}^0_{N,\eta}(X_{r',\mu'})^c\Big)\cap\widetilde{\Lambda}' \\
&\subset\bigcup_{N_k^{\tau+2}\sqrt{\delta}\geq 1}\mathcal{G}^0_{N_k,\eta}(X_{r',\mu'})^c\bigcup_{N^{\tau+2}\sqrt{\delta}\geq 1}\mathtt{G}^0_{N,\eta}(X_{r',\mu'})^c\cap\widetilde{\Lambda}' \\
&\subset\bigcup_{N_k^{\tau+2}\sqrt{\delta}\geq 1}\mathcal{G}^0_{N_k,\frac{1}{2}}(X_{r',\mu'})^c\bigcup_{N^{\tau+2}\sqrt{\delta}\geq 1}\mathtt{G}^0_{N,\frac{1}{2}}(X_{r',\mu'})^c\cap\widetilde{\Lambda}'
\end{aligned}\tag{4.8.36}$$

by (4.5.7). Finally, by (4.8.36), Proposition 4.8.7 and Lemma 4.8.6 we deduce the measure estimate

$$\begin{aligned}
|\Lambda(\varepsilon;\eta-\sqrt{\delta},X_{r,\mu})\cap\Lambda(\varepsilon;\eta,X_{r',\mu'})^c\cap\widetilde{\Lambda}'| &\leq \delta^{\frac{1}{2(\tau+2)}}+\delta^{\frac{(\tau/2)-2d+2|\mathbb{S}|+4}{2(\tau+2)}} \\
&\leq 2\delta^{\frac{1}{2(\tau+2)}}
\end{aligned}$$

by (4.5.1). This proves (4.1.18) for $\alpha < 1/(2(\tau+2))$.

# Chapter 5

# Nash-Moser theorem

The goal of this Chapter is to state the Nash-Moser implicit function Theorem 5.1.2, which proves the existence of a torus embedding $\varphi \mapsto i(\varphi)$ of the form (5.1.1) which is a zero of the nonlinear operator $\mathcal{F}$ defined in (5.1.2). Theorem 5.1.2 implies, going back to the original coordinates, Theorem 1.2.1.

## 5.1 Statement

In this section we state a Nash-Moser implicit function theorem (Theorem 5.1.2) which proves the existence of a solution

$$\varphi \mapsto i(\varphi) = (\theta(\varphi), y(\varphi), z(\varphi)) = (\varphi + \vartheta(\varphi), y(\varphi), z(\varphi)) , \tag{5.1.1}$$

with $z(\varphi) = (Q(\varphi), P(\varphi)) \in H_{\mathbb{S}}^{\perp}$, $\forall \varphi \in \mathbb{T}^{|\mathbb{S}|}$, of the nonlinear operator

$$\begin{aligned} \mathcal{F}(i) := \mathcal{F}(\lambda; i) &:= \omega \cdot \partial_\varphi i(\varphi) - X_K(i(\varphi)) \\ &= \begin{pmatrix} \omega \cdot \partial_\varphi \vartheta(\varphi) + \omega - \bar{\mu} - \varepsilon^2 (\partial_y R)(i(\varphi), \xi) \\ \omega \cdot \partial_\varphi y(\varphi) + \varepsilon^2 (\partial_\theta R)(i(\varphi), \xi) \\ \omega \cdot \partial_\varphi z(\varphi) - J D_V z(\varphi) - \varepsilon^2 \big(0, (\nabla_Q R)(i(\varphi), \xi)\big) \end{pmatrix} \end{aligned} \tag{5.1.2}$$

which depends on the one dimensional parameter

$$\lambda \in \Lambda := [-\lambda_0, \lambda_0]$$

(the set $\Lambda$ is fixed in (1.2.26)) through the frequency vector $\omega = (1 + \varepsilon^2 \lambda) \bar{\omega}_\varepsilon$ where $\bar{\omega}_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ is introduced in (1.2.25), and the amplitudes $\xi := \xi(\lambda)$ are defined in (1.2.27). A solution $i(\varphi)$ of (5.1.2) is an embedded invariant torus for the Hamiltonian system (2.2.11)-(2.2.12), filled by quasi-periodic solutions with frequency $\omega$.

We look for reversible solutions of $\mathcal{F}(\lambda; i) = 0$, namely satisfying $\tilde{S}i(\varphi) = i(-\varphi)$ (the involution $\tilde{S}$ is defined in (2.2.16)), i.e.

$$\theta(-\varphi) = -\theta(\varphi)\,, \quad y(-\varphi) = y(\varphi)\,, \quad z(-\varphi) = (Sz)(\varphi)\,. \tag{5.1.3}$$

**Remark 5.1.1.** *The reversibility property slightly simplifies the argument in Proposition 11.2.4 because the right hand side in* (11.2.41) *has zero average, and therefore the equation* (11.2.41) *is directly solvable. Otherwise we would have to add a counterterm in the second component of the operator $\mathcal{F}$ as in [24], [8].*

The Sobolev norm of the periodic component of the embedded torus

$$\mathfrak{I}(\varphi) := i(\varphi) - (\varphi, 0, 0) := (\vartheta(\varphi), y(\varphi), z(\varphi))\,, \quad \vartheta(\varphi) := \theta(\varphi) - \varphi\,, \tag{5.1.4}$$

is

$$\|\mathfrak{I}\|_{\mathrm{Lip},s} := \|\vartheta\|_{\mathrm{Lip},H^s_\varphi} + \|y\|_{\mathrm{Lip},H^s_\varphi} + \|z\|_{\mathrm{Lip},s}\,. \tag{5.1.5}$$

The solutions of $\mathcal{F}(\lambda; i) = 0$ will be found by a Nash-Moser iterative scheme. Evaluating $\mathcal{F}$ at the trivial embedding

$$i_0(\varphi) := (\varphi, 0, 0)$$

we have

$$\mathcal{F}(i_0) = \begin{pmatrix} \omega - \bar{\mu} - \varepsilon^2(\partial_y R)(\varphi, 0, 0, \xi) \\ \varepsilon^2(\partial_\theta R)(\varphi, 0, 0, \xi) \\ -\varepsilon^2\big(0, \nabla_Q R(\varphi, 0, 0, \xi)\big) \end{pmatrix} \tag{5.1.6}$$

which satisfies, since $\omega$ is $O(\varepsilon^2)$-close to $\bar{\mu}$,

$$\|\mathcal{F}(i_0)\|_{\mathrm{Lip},s} \le C(s)\varepsilon^2, \quad \forall s \ge s_0\,. \tag{5.1.7}$$

In order to construct a better approximate solution we first compute in section 5.2 the shifted tangential frequency vector induced by the nonlinearity, up to $O(\varepsilon^4)$. Then in section 5.3 we construct the first approximate solution $i_1(\varphi)$, defined for all $\lambda \in \Lambda$, by using the unperturbed Melnikov non-resonance conditions (1.2.6)-(1.2.7) on the linear unperturbed frequencies, in such a way that (see (5.3.5))

$$\|\mathcal{F}(i_1)\|_{\mathrm{Lip},s} \le C(s)\varepsilon^4\,, \quad \forall s \ge s_0\,.$$

Subsequently, given an approximate solution $i_n(\varphi)$, the main point is to construct a much better approximate solution $i_{n+1}(\varphi)$. We use an inductive Nash-Moser iterative scheme. The key step concerns the approximate right invertibility properties of the linearized operators $d_i\mathcal{F}(i_n)$ obtained along the iteration, that we obtain restricting the values of $\lambda$ to subsets $\mathbf{\Lambda}_n \subset \Lambda$ with large measure, see Theorem 11.2.1. The following theorem will be proved in Chapter 11, relying on the results of Chapters 6-10 concerning the invertibility properties of the linearized operator $d_i\mathcal{F}(i_n)$.

**Theorem 5.1.2. (Nash-Moser)** *Assume* (1.1.3) *and the non-resonance conditions* (1.2.6)-(1.2.8), (1.2.16)-(1.2.19). *Assume also the twist condition* (1.2.12) *and the non-degeneracy conditions* (1.2.21)-(1.2.22). *Fix a direction* $\bar{\omega}_\varepsilon := \bar{\mu} + \varepsilon^2\zeta$, $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})$, *as in* (1.2.25) *such that the Diophantine conditions* (1.2.29)-(1.2.30) *hold. Define* $\Lambda = [-\lambda_0, \lambda_0]$ *as in* (1.2.26). *Then there are Sobolev indices* $s_2 > s_1 > s_0$, *a constant* $\varepsilon_0 > 0$, *and, for all* $\varepsilon \in (0, \varepsilon_0)$ *there exist*

1. *a Cantor-like set* $\mathcal{C}_\infty \subset \Lambda$ *of asymptotically full measure as* $\varepsilon \to 0$, *i.e.*

$$\lim_{\varepsilon\to 0} \frac{|\mathcal{C}_\infty|}{|\Lambda|} = 1 \,, \tag{5.1.8}$$

*more precisely, there is a map* $\varepsilon \mapsto b(\varepsilon)$, *independent of* $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})$ *such that the Diophantine conditions* (1.2.29)-(1.2.30) *hold, and satisfying* $|\Lambda \setminus \mathcal{C}_\infty| \leq b(\varepsilon)$, $\lim_{\varepsilon\to 0} b(\varepsilon) = 0$;

2. *a Lipschitz function*

$$i_\infty(\varphi;\lambda) - (\varphi,0,0) = (\vartheta_\infty, y_\infty, z_\infty) : \mathcal{C}_\infty \to H^{s_2}_\varphi \times H^{s_2}_\varphi \times (\mathcal{H}^{s_2} \cap H^\perp_{\mathbb{S}})$$

*satisfying*

$$\|i_\infty - (\varphi,0,0)\|_{\mathrm{Lip},s_1} \leq C(s_1)\varepsilon^2 \,, \quad \|i_\infty - (\varphi,0,0)\|_{\mathrm{Lip},s_2} \leq \varepsilon \,, \tag{5.1.9}$$

*such that the torus* $i_\infty(\varphi;\lambda)$, $\lambda \in \mathcal{C}_\infty$, *is a solution of* $\mathcal{F}(\lambda; i_\infty(\lambda)) = 0$.

*Moreover, for any* $\lambda \in \mathcal{C}_\infty$, *the function* $i_\infty - (\varphi,0,0)$ *is of class* $C^\infty$ *in* $(\varphi,x)$, *and Lipschitz in* $\lambda$ *as a map valued in* $H^s_\varphi \times H^s_\varphi \times (\mathcal{H}^s \cap H^\perp_{\mathbb{S}})$, $\forall s \geq s_2$.

*As a consequence the embedded torus* $\varphi \mapsto i_\infty(\varphi;\lambda)$ *is invariant for the Hamiltonian system* (2.2.11)-(2.2.12), *and it is filled by quasi-periodic solutions with frequency* $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$.

Going back to the original coordinates via (2.1.10), (2.2.2), (2.2.3), (2.2.6), Theorem 5.1.2 implies the existence, for all $\lambda \in \mathcal{C}_\infty$, of a quasi-periodic solution of the wave equation (1.2.1) of the form

$$u(t,x) = \sum_{j\in\mathbb{S}} \mu_j^{-\frac{1}{2}} \sqrt{2(\xi_j + (y_\infty)_j(\omega t))} \cos\big(\omega_j t + (\vartheta_\infty)_j(\omega t)\big)\Psi_j(x) + D_V^{-\frac{1}{2}} Q_\infty(\omega t)$$

with frequency $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$. This proves Theorem 1.2.1 with $\mathcal{G}_{\varepsilon,\zeta} := \mathcal{C}_\infty$ and $\bar{s} = s_2$.

The proof of Theorem 5.1.2 occupies the rest of the Monograph from section 5.2 until Chapter 11.

We first prove, as a corollary of Theorem 5.1.2, the result (1.2.35) about the density, close to $\bar{\mu}$, of the frequency vectors $\omega$ of the quasi-periodic solutions of (1.2.1) obtained in Theorem 5.1.2.

## Proof of (1.2.35)

Let $\Omega$ be the set of the frequency vectors $\omega$ of the quasi-periodic solutions of (1.1.1) provided by Theorem 1.2.1. Such frequency vectors have the form

$$\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2(\zeta + \lambda\bar{\mu} + \varepsilon^2\lambda\zeta)\,, \quad \zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|})\,,\ \lambda \in \Lambda\,, \tag{5.1.10}$$

where

- $\zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|}) \setminus B_\varepsilon$ (the set $B_\varepsilon$ is defined in Lemma 2.3.1), so that $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2\zeta$ satisfies the Diophantine conditions (1.2.29)-(1.2.30);
- $\lambda \in \Lambda \setminus \mathcal{G}_{\varepsilon,\zeta}$ where $\mathcal{G}_{\varepsilon,\zeta} := \mathcal{C}_\infty$ is the set defined in Theorem 5.1.2.

We define

$$\mathfrak{B}_\varepsilon := \Big\{(\zeta,\lambda) \in \mathscr{A}([1,2]^{|\mathbb{S}|}) \times \Lambda \,:\, \zeta \in B_\varepsilon \text{ or } \lambda \notin \mathcal{G}_{\varepsilon,\zeta}\Big\}$$
$$\mathfrak{G}_\varepsilon := (\mathscr{A}([1,2]^{|\mathbb{S}|}) \times \Lambda) \setminus \mathfrak{B}_\varepsilon\,.$$

By Lemma 2.3.1 the Lebesgue measure of $B_\varepsilon$ satisfies $|B_\varepsilon| \le \varepsilon$, and, using also the measure estimate provided in item 1 of Theorem 5.1.2, we deduce that

$$|\mathfrak{B}_\varepsilon| \le \varepsilon|\Lambda| + b(\varepsilon)|\mathscr{A}([1,2]^{|\mathbb{S}|})| =: b_1(\varepsilon) \tag{5.1.11}$$

where $\lim_{\varepsilon\to 0} b_1(\varepsilon) = 0$.

In view of (5.1.10), in order to prove (1.2.35), we have to estimate the measure of the set

$$B'_\varepsilon := \Big\{\beta \in \mathcal{C}_1 := \mathscr{A}([1,2]^{|\mathbb{S}|}) + \Lambda\bar{\mu} \,:\, \nexists(\zeta,\lambda) \in \mathfrak{G}_\varepsilon \text{ such that } \beta = \zeta + \lambda\bar{\mu} + \varepsilon^2\lambda\zeta\Big\}\,. \tag{5.1.12}$$

**Lemma 5.1.3.** *$|B'_\varepsilon| \to 0$ as $\varepsilon \to 0$.*

Proof. Define the map

$$\Psi_\varepsilon : \mathscr{A}([1,2]^{|\mathbb{S}|}) \times \Lambda \to \mathbb{R}^{|\mathbb{S}|} \times \Lambda\,, \quad \Psi_\varepsilon(\zeta,\lambda) := \big(\zeta + \lambda\bar{\mu} + \varepsilon^2\lambda\zeta, \lambda\big)\,,$$

which is a diffeomorphism onto its image. Thus, recalling (5.1.11),

$$|\Psi_\varepsilon(\mathfrak{B}_\varepsilon)| \lesssim |\mathfrak{B}_\varepsilon| \lesssim b_1(\varepsilon)\,. \tag{5.1.13}$$

For any $\beta \in \mathcal{C}_1$, let

$$U_{\beta,\varepsilon} := \Big\{\lambda \in \Lambda \,:\, \frac{\beta - \lambda\bar{\mu}}{1+\varepsilon^2\lambda} \in \mathscr{A}([1,2]^{|\mathbb{S}|})\Big\}\,, \tag{5.1.14}$$

i.e. $\lambda \in U_{\beta,\varepsilon}$ if and only if $(\beta,\lambda)$ is in the image $\Psi_\varepsilon(\mathscr{A}([1,2]^{|\mathbb{S}|}) \times \Lambda)$. Thus, recalling (5.1.12), we deduce that

$$\beta \in B'_\varepsilon\,,\ \lambda \in U_{\beta,\varepsilon} \quad \Longrightarrow \quad (\beta,\lambda) \in \Psi_\varepsilon(\mathfrak{B}_\varepsilon)\,.$$

Therefore

$$\int_{B'_\varepsilon} |U_{\beta,\varepsilon}| d\beta \le |\Psi_\varepsilon(\mathfrak{B}_\varepsilon)| \overset{(5.1.13)}{\lesssim} b_1(\varepsilon) \,. \tag{5.1.15}$$

Our aim is now to justify that the measure of $U_{\beta,\varepsilon}$ satisfies $|U_{\beta,\varepsilon}| \ge \sqrt{b_1(\varepsilon)}$ for all $\beta \in \mathscr{A}([1,2]^{|\mathbb{S}|}) + \Lambda\bar{\mu}$ but a subset the measure of which vanishes as $\varepsilon \to 0$.

First note that $\mathscr{A}([1,2]^{|\mathbb{S}|})$ is a convex subset of $\mathbb{R}^{|\mathbb{S}|}$, with interior $\mathscr{A}((1,2)^{|\mathbb{S}|})$. For $\varepsilon > 0$, define

$$\begin{aligned} V_\varepsilon &:= \Big\{ x \in \mathbb{R}^{|\mathbb{S}|} \ : \ \forall \lambda \in \Lambda \,, \ \frac{x}{1+\varepsilon^2\lambda} \in \mathscr{A}((1,2)^{|\mathbb{S}|}) \Big\} \\ &= \Big\{ x \in \mathbb{R}^{|\mathbb{S}|} \ : \ \Big[ \frac{1}{1+\varepsilon^2\lambda_0}, \frac{1}{1-\varepsilon^2\lambda_0} \Big] \cdot x \subset \mathscr{A}((1,2)^{|\mathbb{S}|}) \Big\} \subset \mathscr{A}((1,2)^{|\mathbb{S}|}) \,. \end{aligned} \tag{5.1.16}$$

Each set $V_\varepsilon$ is convex and open,

$$V_{\varepsilon'} \subset V_\varepsilon \,, \ \forall \, 0 < \varepsilon < \varepsilon' \qquad \text{and} \qquad \bigcup_{\varepsilon > 0} V_\varepsilon = \mathscr{A}((1,2)^{|\mathbb{S}|}) \,. \tag{5.1.17}$$

Then, for any $\beta \in \mathcal{C}_1$ define

$$U'_{\beta,\varepsilon} := \Big\{ \lambda \in \Lambda \ : \ \beta - \lambda\bar{\mu} \in V_\varepsilon \Big\} \overset{(5.1.16),(5.1.14)}{\subset} U_{\beta,\varepsilon} \,.$$

Recalling (5.1.17) we have $U'_{\beta,\varepsilon'} \subset U'_{\beta,\varepsilon}$, $\forall 0 < \varepsilon < \varepsilon'$. At last define, for $\delta > 0$,

$$D_{\varepsilon,\delta} := \big\{ \beta \in \mathcal{C}_1 \ : \ |U'_{\beta,\varepsilon}| \ge \delta \big\} \,. \tag{5.1.18}$$

The following properties holds:

(i) For $0 < \varepsilon < \varepsilon'$ and $0 < \delta < \delta'$, we have $D_{\varepsilon',\delta'} \subset D_{\varepsilon,\delta}$.

(ii) Since the sets $U'_{\beta,\varepsilon}$ are open (hence of strictly positive measure if nonempty),

$$\bigcup_{\delta > 0} D_{\varepsilon,\delta} = \big\{ \beta \in \mathcal{C}_1 \ : \ U'_{\beta,\varepsilon} \neq \emptyset \big\} =: D_\varepsilon \,.$$

Moreover $D_\varepsilon = V_\varepsilon + \Lambda\bar{\mu}$.

(iii) By items (i)-(ii) and (5.1.17) we deduce

$$\bigcup_{\varepsilon > 0} \Big( \bigcup_{\delta > 0} D_{\varepsilon,\delta} \Big) = \bigcup_{\varepsilon > 0} D_\varepsilon = \mathscr{A}((1,2)^{|\mathbb{S}|}) + \Lambda\bar{\mu} = \mathcal{C}_1 \,.$$

**Claim:**

$$\lim_{\varepsilon\to 0_+,\delta\to 0_+} \left|\mathcal{C}_1\backslash D_{\varepsilon,\delta}\right| = \left|\mathcal{C}_1\backslash(\mathscr{A}((1,2)^{|\mathbb{S}|}) + \Lambda\bar{\mu})\right| = 0\,. \tag{5.1.19}$$

The first equality follows by items (i)-(iii) above. To justify this last equality, let us introduce the $(|\mathbb{S}|-1)$-dimensional linear subspace of $\mathbb{R}^{|\mathbb{S}|}$, $E := \bar{\mu}^{\perp}$. Let $K$ be the orthogonal projection of $\mathscr{A}([1,2]^{|\mathbb{S}|})$ onto $E$; note that $K$ is a convex compact subset of $E$ of nonempty interior in $E$. Moreover, since $\mathscr{A}([1,2]^{|\mathbb{S}|})$ is convex, it can be decomposed as

$$\mathscr{A}([1,2]^{|\mathbb{S}|}) = \bigcup_{x\in K} \{x + [\alpha_-(x), \alpha_+(x)]\bar{\mu}\}$$

where the functions $\alpha_+, \alpha_- : K \to \mathbb{R}$ are respectively concave and convex, with $\alpha_-(x) \le \alpha_+(x)$, for all $x \in K$, and $\alpha_-(x) < \alpha_+(x)$ for $x \in \mathrm{int}(K)$. Hence, since $\alpha_\pm$ are continuous on $\mathrm{int}(K)$,

$$\mathscr{A}((1,2)^{|\mathbb{S}|}) = \mathrm{int}(\mathscr{A}([1,2]^{|\mathbb{S}|})) = \left\{x + (\alpha_-(x), \alpha_+(x))\bar{\mu}\,;\ x \in \mathrm{int}(K)\right\},$$

and

$$\begin{aligned}\mathcal{C}_1\backslash(\mathscr{A}((1,2)^{|\mathbb{S}|}) + \Lambda\bar{\mu}) = \Big\{x + \lambda\bar{\mu}\ :\ &\big(x \in \partial K\,,\ \lambda \in [\alpha_-(x) - \lambda_0, \alpha_+(x) + \lambda_0]\big)\\ &\text{or } \big(x \in K\,,\ \lambda = \alpha_\pm(x) \pm \lambda_0\big)\Big\}\,,\end{aligned}$$

which gives $\left|\mathcal{C}_1\backslash(\mathscr{A}((1,2)^{|\mathbb{S}|}) + \Lambda\bar{\mu})\right| = 0$.

Setting $D'_\varepsilon := D_{\varepsilon,\sqrt{b_1(\varepsilon)}}$, where $D_{\varepsilon,\delta}$ is defined in (5.1.18), the estimate (5.1.19) implies

$$\lim_{\varepsilon\to 0_+} \left|\mathcal{C}_1\backslash D'_\varepsilon\right| = 0\,. \tag{5.1.20}$$

Moreover, by the definition of $D'_\varepsilon$ and the inclusion $U'_{\beta,\varepsilon} \subset U_{\beta,\varepsilon}$, we deduce

$$\forall \beta \in D'_\varepsilon\,,\quad |U_{\beta,\varepsilon}| \ge \sqrt{b_1(\varepsilon)}\,,$$

and therefore

$$|B'_\varepsilon \cap D'_\varepsilon|\sqrt{b_1(\varepsilon)} \le \int_{B'_\varepsilon\cap D'_\varepsilon} |U_\beta|d\beta \overset{(5.1.15)}{\lesssim} b_1(\varepsilon)\,. \tag{5.1.21}$$

Finally, since $B'_\varepsilon \subset (B'_\varepsilon \cap D'_\varepsilon) \cup (\mathcal{C}_1\backslash D'_\varepsilon)$, we deduce, by (5.1.21) and (5.1.20), that $\lim_{\varepsilon\to 0} |B'_\varepsilon| = 0$. Lemma 5.1.3 is proved. ■

Now, recalling (5.1.12) and (5.1.10), we have

$$\bar{\mu} + \bigcup_{\varepsilon>0} \varepsilon^2(\mathcal{C}_1\backslash B'_\varepsilon) \subset \Omega \tag{5.1.22}$$

where $\Omega$ is the set of the frequency vectors $\omega$ of the quasi-periodic solutions of (1.1.1) provided by Theorem 1.2.1. Notice that, by (1.2.26), (1.2.25), (1.2.23), we get that

$$\zeta + \lambda\bar{\mu} + \lambda\varepsilon^2\zeta \in \mathscr{A}\big([\frac{1}{2}, 4]^{|\mathbb{S}|}\big) , \quad \forall \zeta \in \mathscr{A}([1,2]^{|\mathbb{S}|}) , \ \forall \lambda \in \Lambda ,$$

and therefore $\mathcal{C}_1 = \mathscr{A}([1,2]^{|\mathbb{S}|}) + \Lambda\bar{\mu}$ does not contain 0. Moreover $\mathcal{C}_1$ is a compact convex subset of $\mathbb{R}^{|\mathbb{S}|}$, with nonempty interior, which implies that

$$\big\{ x \in \mathbb{R}_+\mathcal{C}_1 \ : \ \mathbb{R}_+ x \cap \mathcal{C}_1 \text{ is a singleton} \big\} \subset \partial(\mathbb{R}_+\mathcal{C}_1) . \tag{5.1.23}$$

Thus, given $y \in \mathbb{R}_+\mathcal{C}_1$, the measure $\big|\{ r > 0 \ : \ y/r \in \mathcal{C}_1 \}\big| > 0$ except for $y \in \partial(\mathbb{R}_+\mathcal{C}_1)$, which is of zero measure. Using (5.1.23) and Lemma 5.1.3, we can obtain

$$\lim_{r \to 0_+} \frac{\Big| \Big( \bigcup_{\varepsilon > 0} \varepsilon^2 (\mathcal{C}_1 \backslash B'_\varepsilon) \Big) \cap B(0,r) \Big|}{\Big| \mathbb{R}_+\mathcal{C}_1 \cap B(0,r) \Big|} = 1 . \tag{5.1.24}$$

We omit the details. Recalling (5.1.22), (5.1.24) implies (1.2.35).

## 5.2 Shifted tangential frequencies up to $O(\varepsilon^4)$

In this section we evaluate the average of the first component in (5.1.6):

$$\omega - \bar{\mu} - \varepsilon^2 \langle \partial_y R(\varphi, 0, 0, \xi) \rangle \tag{5.2.1}$$

where

$$\langle f \rangle := \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} f(\varphi)\, d\varphi . \tag{5.2.2}$$

Evaluating (2.2.13) at $(\theta, y, Q) = (\varphi, 0, 0)$, we get, inserting the expression of $g(\varepsilon, x, u)$ in (2.1.6), that for each $m = 1, \dots, |\mathbb{S}|$,

$$\begin{aligned} \langle \partial_{y_m} R(\varphi, 0, 0, \xi) \rangle &= \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} \int_{\mathbb{T}^d} g(\varepsilon, x, v(\varphi, 0, \xi)) \frac{\mu_m^{-1/2}}{\sqrt{2\xi_m}} \cos\varphi_m \, \Psi_m(x) \, dx \, d\varphi \\ &= \mathtt{r}_{m,3} + \varepsilon \mathtt{r}_{m,4} + \varepsilon^2 \mathtt{r}_{m,5} \end{aligned} \tag{5.2.3}$$

where

$$\mathtt{r}_{m,3} := \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} \int_{\mathbb{T}^d} a(x) \big( v(\varphi, 0, \xi) \big)^3 \frac{\mu_m^{-1/2}}{\sqrt{2\xi_m}} \cos\varphi_m \, \Psi_m(x) \, dx \tag{5.2.4}$$

$$\mathtt{r}_{m,4} := \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} \int_{\mathbb{T}^d} a_4(x) \big( v(\varphi, 0, \xi) \big)^4 \frac{\mu_m^{-1/2}}{\sqrt{2\xi_m}} \cos\varphi_m \, \Psi_m(x) \, dx \tag{5.2.5}$$

$$\mathtt{r}_{m,5} := \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} \int_{\mathbb{T}^d} \mathfrak{r}(\varepsilon, x, v(\varphi, 0, \xi)) \frac{\mu_m^{-1/2}}{\sqrt{2\xi_m}} \cos\varphi_m \, \Psi_m(x) \, dx . \tag{5.2.6}$$

We now compute the terms in (5.2.3).

**Lemma 5.2.1.** $\mathtt{r}_{m,3}$ *in* (5.2.4) *is*

$$\mathtt{r}_{m,3} = [\mathscr{A}\xi]_m \tag{5.2.7}$$

*where* $\mathscr{A} := (\mathscr{A}_m^j)_{j,m\in\mathbb{S}}$ *is the symmetric twist matrix defined in* (1.2.9).

Proof. Using (2.2.9), we expand the integral (5.2.4) as

$$\begin{aligned}\mathtt{r}_{m,3} = &\sum_{j_1,j_2,j_3\in\mathbb{S}} \mu_{j_1}^{-1/2}\mu_{j_2}^{-1/2}\mu_{j_3}^{-1/2}\mu_m^{-1/2}\sqrt{\xi_{j_1}\xi_{j_2}\xi_{j_3}}\frac{2}{\sqrt{\xi_m}} \\ &\times \frac{1}{(2\pi)^{|\mathbb{S}|}}\int_{\mathbb{T}^{|\mathbb{S}|}} \cos\varphi_{j_1}\cos\varphi_{j_2}\cos\varphi_{j_3}\cos\varphi_m d\varphi \\ &\times \int_{\mathbb{T}^d} a(x)\Psi_{j_1}(x)\Psi_{j_2}(x)\Psi_{j_3}(x)\Psi_m(x)\,dx\,.\end{aligned} \tag{5.2.8}$$

The integral $\int_{\mathbb{T}^{|\mathbb{S}|}} \cos\varphi_{j_1}\cos\varphi_{j_2}\cos\varphi_{j_3}\cos\varphi_m d\varphi$ does not vanish only if

1. $j_1 = j_2 = j_3 = m$,
2. $j_1 = j_2 \neq j_3 = m$ and permutation of the indices (3 times).

Hence, by (5.2.8),

$$\begin{aligned}\mathtt{r}_{3,m} &= 2\mu_m^{-2}\xi_m\frac{1}{(2\pi)^{|\mathbb{S}|}}\int_{\mathbb{T}^{|\mathbb{S}|}}\cos^4\varphi_m d\varphi\int_{\mathbb{T}^d} a(x)\Psi_m^4(x)\,dx \\ &+ 3\sum_{j_1\neq m}\mu_{j_1}^{-1}\mu_m^{-1}\xi_{j_1}2\frac{1}{(2\pi)^{|\mathbb{S}|}}\int_{\mathbb{T}^{|\mathbb{S}|}}\cos^2\varphi_{j_1}\cos^2\varphi_m d\varphi\int_{\mathbb{T}^d} a(x)\Psi_{j_1}^2(x)\Psi_m^2(x)\,dx \\ &= \frac{3}{4}\mu_m^{-2}T_m^m\xi_m + \frac{3}{2}\sum_{j\neq m}\mu_j^{-1}\mu_m^{-1}T_m^j\xi_j\end{aligned}$$

having set

$$T_m^j := \int_{\mathbb{T}^d} a(x)\Psi_j^2(x)\Psi_m^2(x)\,dx\,,\ j,m\in\mathbb{S}\,,$$

and noting that

$$\begin{aligned}&\frac{2}{(2\pi)^{|\mathbb{S}|}}\int_{\mathbb{T}^{|\mathbb{S}|}}\cos^4\varphi_m d\varphi = \frac{1}{\pi}\int_{\mathbb{T}}\cos^4\varphi_m d\varphi_m = \frac{3}{4} \\ &\frac{6}{(2\pi)^{|\mathbb{S}|}}\int_{\mathbb{T}^{|\mathbb{S}|}}\cos^2\varphi_{j_1}\cos^2\varphi_m d\varphi = \frac{6}{(2\pi)^2}\Big(\int_{\mathbb{T}}\cos^2\theta\,d\theta\Big)^2 = \frac{3}{2}\,.\end{aligned}$$

Recalling the definition of the twist matrix $\mathscr{A} := (\mathscr{A}_m^j)_{j,m\in\mathbb{S}}$ in (1.2.9)-(1.2.10) we deduce (5.2.7). ■

**Lemma 5.2.2.** *For all $m \in \mathbb{S}$, each $\mathtt{r}_{m,4}$ in* (5.2.5) *is* $\mathtt{r}_{m,4} = 0$.

PROOF. Since the function $v$ defined in (2.2.9) satisfies the symmetry

$$v(\varphi + \vec{\pi}, 0, \xi) = -v(\varphi, 0, \xi)\,, \quad \vec{\pi} := (\pi, \dots, \pi) \in \mathbb{R}^{|\mathbb{S}|}\,, \tag{5.2.9}$$

the function $g(\varphi) := (v(\varphi, 0, \xi))^4 \cos(\varphi_m)$ satisfies $g(\varphi + \vec{\pi}) = -g(\varphi)$ and therefore its integral

$$\int_{\mathbb{T}^{|\mathbb{S}|}} g(\varphi) d\varphi = \int_{\mathbb{T}^{|\mathbb{S}|}} g(\varphi + \vec{\pi}) d\varphi = -\int_{\mathbb{T}^{|\mathbb{S}|}} g(\varphi) d\varphi$$

is equal to zero. Hence $\mathtt{r}_{m,4} = 0$. ∎

By (5.2.3), Lemmata 5.2.1 and 5.2.2 we deduce that

$$\langle \partial_y R(\varphi, 0, 0, \xi) \rangle = \mathscr{A}\xi + \varepsilon^2 \mathtt{r}_5(\varepsilon, \xi) \tag{5.2.10}$$

where $\mathscr{A}$ is the symmetric *twist* matrix defined in (1.2.9), and $\mathtt{r}_5(\varepsilon, \xi) := (\mathtt{r}_{5,m})_{m=1,\dots,|\mathbb{S}|} \in \mathbb{R}^{|\mathbb{S}|}$. As a consequence, the term (5.2.1) is $O(\varepsilon^4)$, more precisely, since $\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon$, it results

$$\begin{aligned} \omega - \bar{\mu} - \varepsilon^2 \langle \partial_y R(\varphi, 0, 0, \xi) \rangle &= (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon - \bar{\mu} - \varepsilon^2 \mathscr{A}\xi - \varepsilon^4 \mathtt{r}_5(\varepsilon, \xi) \\ &\overset{(1.2.27)}{=} -\varepsilon^4 \mathtt{r}_5(\varepsilon, \xi)\,. \end{aligned} \tag{5.2.11}$$

## 5.3 First approximate solution

We now define the first approximate torus embedding solution

$$i_1(\varphi) = \big(\theta_1(\varphi), y_1(\varphi), Q_1(\varphi), P_1(\varphi)\big)\,, \qquad \theta_1(\varphi) = \varphi + \vartheta_1(\varphi)\,, \tag{5.3.1}$$

in such a way that $\mathcal{F}(i_1) = O(\varepsilon^4)$. Given a function $f : \mathbb{T}^{|\mathbb{S}|} \to \mathbb{R}^{|\mathbb{S}|}$, we denote by $\langle f \rangle \in \mathbb{R}^{|\mathbb{S}|}$ its average with respect to $\varphi$ (as in (5.2.2)) and by $[f](\varphi)$ its zero mean part, so that

$$f(\varphi) = \langle f \rangle + [f](\varphi)\,. \tag{5.3.2}$$

**Lemma 5.3.1. (First approximate solution)** *Let $\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon$ with $\bar{\omega}_\varepsilon = \bar{\mu} + \varepsilon^2\zeta$ as in* (1.2.25) *and define $\xi := \xi(\lambda)$ as in* (1.2.27)*. Then there exists a unique solution $i_1$, with the form in* (5.3.1)*, with average $\langle y_1 \rangle = 0$, independent of $\lambda \in \Lambda$, of the system*

$$\begin{cases} \bar{\mu} \cdot \partial_\varphi \vartheta_1 - \varepsilon^2 \big[(\partial_y R)(\varphi, 0, 0, \xi)\big] = 0 \\ \bar{\mu} \cdot \partial_\varphi y_1 + \varepsilon^2 (\partial_\theta R)(\varphi, 0, 0, \xi) = 0 \\ \big(\bar{\mu} \cdot \partial_\varphi - JD_V\big)(Q_1, P_1) - \varepsilon^2\big(0, (\nabla_Q R)(\varphi, 0, 0, \xi)\big) = 0 \end{cases} \tag{5.3.3}$$

*satisfying, for all* $s \geq s_0$,

$$\|i_1 - (\varphi, 0, 0)\|_{\mathrm{Lip},s} = \|i_1 - (\varphi, 0, 0)\|_s \leq C(s)\varepsilon^2 . \tag{5.3.4}$$

*It results*

$$\|\mathcal{F}(i_1)\|_{\mathrm{Lip},s} \leq C(s)\varepsilon^4 , \quad \forall s \geq s_0 . \tag{5.3.5}$$

PROOF. SOLUTION OF THE FIRST EQUATION IN (5.3.3). Since $\bar{\mu}$ is a Diophantine vector by (1.2.6), we solve the first equation in (5.3.3), finding

$$\vartheta_1 = \varepsilon^2 (\bar{\mu} \cdot \partial_\varphi)^{-1} \big[ (\partial_y R)(\varphi, 0, 0, \xi) \big] \tag{5.3.6}$$

where $(\bar{\mu} \cdot \partial_\varphi)^{-1}$ is defined as in (1.6.1) (with $\bar{\mu}$ instead of $\omega$).
SOLUTION OF THE SECOND EQUATION IN (5.3.3). Since

$$(\partial_\theta R)(\varphi, 0, 0, \xi) = \partial_\varphi \big( R(\varphi, 0, 0, \xi) \big) ,$$

it has zero average in $\varphi$. Then, since $\bar{\mu}$ is Diophantine by (1.2.6), the second equation in (5.3.3) admits the unique solution with zero average

$$y_1 = -\varepsilon^2 (\bar{\mu} \cdot \partial_\varphi)^{-1} \big[ (\partial_\theta R)(\varphi, 0, 0, \xi) \big] . \tag{5.3.7}$$

SOLUTION OF THE THIRD EQUATION IN (5.3.3). The operator $\bar{\mu} \cdot \partial_\varphi - JD_V$ is represented, in the basis $\{ e^{\mathrm{i}\ell \cdot \varphi} (\Psi_j(x), 0), e^{\mathrm{i}\ell \cdot \varphi} (0, \Psi_j(x)) \}_{j \in \mathbb{N}}$ (see (3.2.26)) by the diagonal matrix

$$\mathrm{Diag}_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{N}} \begin{pmatrix} \mathrm{i}\bar{\mu} \cdot \ell & -\mu_j \\ \mu_j & \mathrm{i}\bar{\mu} \cdot \ell \end{pmatrix} ,$$

and therefore, by the unperturbed first Melnikov condition (1.2.7), it is invertible. Moreover, arguing as in the end of Lemma 7.2.2, it satisfies the estimate

$$\| \big( \bar{\mu} \cdot \partial_\varphi - JD_V \big)^{-1} h \|_s \leq C(s) \|h\|_{s+\tau_0} . \tag{5.3.8}$$

Then the third equation in (5.3.3) admits the unique solution

$$(Q_1, P_1) = \varepsilon^2 \big( \bar{\mu} \cdot \partial_\varphi - JD_V \big)^{-1} \big( 0, (\nabla_Q R)(\varphi, 0, 0, \xi) \big) . \tag{5.3.9}$$

The estimate (5.3.4) follows by the definition of $i_1(\varphi) = (\theta_1(\varphi), y_1(\varphi), z_1(\varphi))$ where $z_1(\varphi) := (Q_1(\varphi), P_1(\varphi))$, in (5.3.6), (5.3.7), (5.3.9), using (5.3.8) and the Diophantine condition (1.2.6). Finally, comparing (5.1.2) with system (5.3.3), we have

$$\mathcal{F}(i_1) =$$

$$\begin{pmatrix} \omega - \bar{\mu} - \varepsilon^2 \langle (\partial_y R)(i_1(\varphi), \xi) \rangle + (\omega - \bar{\mu}) \cdot \partial_\varphi \vartheta_1(\varphi) - \varepsilon^2 \big( [\partial_y R(i_1(\varphi), \xi)] - [\partial_y R(\varphi, 0, 0, \xi)] \big) \\ (\omega - \bar{\mu}) \cdot \partial_\varphi y_1(\varphi) + \varepsilon^2 \big( (\partial_\theta R)(i_1(\varphi), \xi) - (\partial_\theta R)(\varphi, 0, 0, \xi) \big) \\ (\omega - \bar{\mu}) \cdot \partial_\varphi z_1(\varphi) - \varepsilon^2 \Big( \big( 0, (\nabla_Q R)(i_1(\varphi), \xi) \big) - \big( 0, (\nabla_Q R)(\varphi, 0, 0, \xi) \big) \Big) \end{pmatrix}$$

and (5.3.5) follows using the estimate (5.3.4), the fact that $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$ with $\bar{\omega}_\varepsilon = \bar{\mu}+\varepsilon^2\zeta$, and (5.2.11). ■

The successive approximate solutions $i_n$, $n \geq 2$, of the functional equation $\mathcal{F}(i) = 0$, are defined through a Nash-Moser iterative scheme. The main point to define $i_{n+1}$ is the construction of an approximate right inverse of the linearized operators $d_i\mathcal{F}(i_n)$ at the approximate torus $i_n$, that we obtain in the next Chapters 6-10.

# Chapter 6

# Linearized operator at an approximate solution

In order to implement a convergent Nash-Moser scheme (Chapter 11) that leads to a solution of $\mathcal{F}(\lambda, i) = 0$ where $\mathcal{F}(\lambda, i)$ is the nonlinear operator defined in (5.1.2), the key step is to prove the existence of an approximate right inverse of the linearized operator $d_i\mathcal{F}(\lambda; \underline{i})$ in (6.1.2). The first step is Proposition 6.1.1 where we introduce suitable symplectic coordinates which reduce the problem to the search of an approximate inverse of the operator $\mathcal{L}_\omega$ in (6.1.23) acting in the normal components only. This will be studied in Chapters 7-10.

## 6.1 Symplectic approximate decoupling

We linearize $\mathcal{F}(\lambda, i)$ at an arbitrary torus

$$\underline{i}(\varphi) = (\underline{\theta}(\varphi), \underline{y}(\varphi), \underline{z}(\varphi)) , \tag{6.1.1}$$

obtaining

$$d_i\mathcal{F}(\lambda; \underline{i})[\widehat{\imath}] = \omega\cdot\partial_\varphi\widehat{\imath} - d_i X_K(\underline{i}(\varphi))[\widehat{\imath}] . \tag{6.1.2}$$

We denote by

$$\underline{\mathfrak{I}}(\varphi) := \underline{i}(\varphi) - (\varphi, 0, 0) := (\underline{\vartheta}(\varphi), \underline{y}(\varphi), \underline{z}(\varphi)) , \quad \underline{\vartheta}(\varphi) := \underline{\theta}(\varphi) - \varphi , \tag{6.1.3}$$

the periodic component of the torus $\varphi \mapsto \underline{i}(\varphi)$ with norm as in (5.1.5). We assume the following condition for $\underline{i}$ which is satisfied by any approximate solution obtained along the Nash-Moser iteration performed in Chapter 11 (see precisely (11.2.11)):

- The map $\lambda \mapsto \underline{\mathfrak{I}}(\lambda)$ is Lipschitz with respect to $\lambda \in \Lambda_{\underline{\mathfrak{I}}} \subset \Lambda$, and

$$\|\underline{\mathfrak{I}}\|_{\mathrm{Lip}, s_1+2} \leq C(s_1)\varepsilon^2 , \quad \|\underline{\mathfrak{I}}\|_{\mathrm{Lip}, s_2} \leq \varepsilon . \tag{6.1.4}$$

We implement the general strategy proposed in [24], used also in [8], [30], where, instead of inverting $d_i\mathcal{F}(\lambda;\underline{i})$ (where all the $(\theta,y,z)$ components are coupled, see (5.1.2)) we invert the linear operator $\mathbb{D}(\underline{i})$ in (6.1.22), which has a triangular form. The operator $\mathbb{D}(\underline{i})$ is found by a natural geometrical construction. We define the "error function"

$$Z(\varphi) := (Z_1, Z_2, Z_3)(\varphi) := \mathcal{F}(\lambda;\underline{i})(\varphi) = \omega\cdot\partial_\varphi \underline{i}(\varphi) - X_K(\underline{i}(\varphi))\,. \tag{6.1.5}$$

Notice that, if $Z = 0$ then the torus $\underline{i}$ is invariant for $X_K$; in general, we say that $\underline{i}$ is "*approximately invariant*", up to order $O(Z)$. Given $\underline{i}(\varphi)$ satisfying (6.1.4) we first construct an isotropic torus $i_\delta(\varphi)$ which is close to $\underline{i}$, see (6.1.6) and (6.2.6). By (6.1.7), $\mathcal{F}(i_\delta)$ is also $O(Z)$. Since the torus $i_\delta$ is isotropic, the diffeomorphism $(\phi,\zeta,w)\mapsto G_\delta(\phi,\zeta,w)$ defined in (6.1.9) is symplectic. In these coordinates, the torus $i_\delta$ reads $(\phi,0,0)$, and the transformed Hamiltonian system becomes (6.2.12), where, by (6.1.15) the terms $\partial_\phi \mathtt{K}_{00}, \mathtt{K}_{10}-\omega, \mathtt{K}_{01}$ are $O(Z)$. Neglecting such terms in the linearized operator (6.1.21) at $(\phi,0,0)$, we obtain the linear operator $\mathbb{D}(\underline{i})$ in (6.1.22).

The main result of this section is the following Proposition.

**Proposition 6.1.1.** *Let $\underline{i}(\varphi)$ be a torus of the form* (6.1.1)*, defined for all $\lambda\in\Lambda_{\underline{\mathfrak{I}}}$, satisfying* (6.1.4)*. Then*

- **(isotropic torus)** *there is an isotropic torus $i_\delta(\varphi) = (\underline{\theta}(\varphi), y_\delta(\varphi), \underline{z}(\varphi))$ satisfying, for some $\underline{\tau} := \underline{\tau}(|\mathbb{S}|,\tau_1) > 0$,*

$$\|y_\delta - \underline{y}\|_{\mathrm{Lip},s} \lesssim_s \|Z\|_{\mathrm{Lip},s+\underline{\tau}} + \|Z\|_{\mathrm{Lip},s_0+\underline{\tau}}\|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+\underline{\tau}} \tag{6.1.6}$$

$$\|\mathcal{F}(i_\delta)\|_{\mathrm{Lip},s} \lesssim_s \|Z\|_{\mathrm{Lip},s+\underline{\tau}} + \|Z\|_{\mathrm{Lip},s_0+\underline{\tau}}\|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+\underline{\tau}}\,. \tag{6.1.7}$$

  *Given another $\underline{i}'$ satisfying* (6.1.4) *we have*

$$\|i_\delta(\underline{i}) - i_\delta(\underline{i}')\|_{s_1} \lesssim_{s_1} \|\underline{\mathfrak{I}} - \underline{\mathfrak{I}}'\|_{s_1+1}\,. \tag{6.1.8}$$

- **(symplectic diffeomorphism)** *the change of variable $G_\delta : (\phi,\zeta,w)\to(\theta,y,z)$ of the phase space $\mathbb{T}^{|\mathbb{S}|}\times\mathbb{R}^{|\mathbb{S}|}\times H_{\mathbb{S}}^\perp$ defined by*

$$\begin{pmatrix}\theta\\ y\\ z\end{pmatrix} := G_\delta\begin{pmatrix}\phi\\ \zeta\\ w\end{pmatrix} := \begin{pmatrix}\underline{\theta}(\phi)\\ y_\delta(\phi) + [\partial_\phi\underline{\theta}(\phi)]^{-\top}\zeta - \big[(\partial_\theta\underline{\tilde{z}})(\underline{\theta}(\phi))\big]^\top Jw\\ \underline{z}(\phi)+w\end{pmatrix} \tag{6.1.9}$$

  *where $\underline{\tilde{z}}(\theta) := \underline{z}(\underline{\theta}^{-1}(\theta))$, is symplectic.*

  *In the new coordinates $(\phi,\zeta,w)$, the isotropic torus $i_\delta$ is the trivial embedded torus $(\varphi,0,0)$, i.e.*

$$i_\delta(\varphi) = G_\delta(\varphi,0,0)\,. \tag{6.1.10}$$

*The linearized diffeomorphism* $DG_\delta(\varphi,0,0)$ *satisfies, for all* $s \geq s_0$,

$$\|DG_\delta(\varphi,0,0)[\widehat{\imath}]\|_{\mathrm{Lip},s} + \big\|\big(DG_\delta(\varphi,0,0)\big)^{-1}[\widehat{\imath}]\big\|_{\mathrm{Lip},s} \lesssim_s \|\widehat{\imath}\|_{\mathrm{Lip},s} + \|\mathfrak{I}\|_{\mathrm{Lip},s+2}\|\widehat{\imath}\|_{\mathrm{Lip},s_0} \tag{6.1.11}$$

*and*

$$\|D^2G_\delta(\varphi,0,0)[\widehat{\imath}_1,\widehat{\imath}_2]\|_{\mathrm{Lip},s} \lesssim_s \|\widehat{\imath}_1\|_{\mathrm{Lip},s}\|\widehat{\imath}_2\|_{\mathrm{Lip},s_0} + \|\widehat{\imath}_1\|_{\mathrm{Lip},s_0}\|\widehat{\imath}_2\|_{\mathrm{Lip},s} + \|\mathfrak{I}\|_{\mathrm{Lip},s+3}\|\widehat{\imath}_1\|_{\mathrm{Lip},s_0}\|\widehat{\imath}_2\|_{\mathrm{Lip},s_0}\,. \tag{6.1.12}$$

- **(Transformed Hamiltonian)** *Under the symplectic change of variables* $G_\delta$*, the Hamiltonian vector field* $X_K$ *(the Hamiltonian* $K$ *is defined in* (2.2.7)*) transforms into*

$$X_{\mathtt{K}} = (DG_\delta)^{-1}X_K \circ G_\delta \qquad \text{where} \qquad \mathtt{K} := K \circ G_\delta\,. \tag{6.1.13}$$

*The Hamiltonian* $\mathtt{K}$ *is reversible, i.e.* $\mathtt{K}\circ\tilde{S} = \mathtt{K}$*. The 2-jets of the Taylor expansion of the Hamiltonian* $\mathtt{K}$ *at the trivial torus* $(\phi,0,0)$,

$$\begin{aligned}\mathtt{K}(\phi,\zeta,w) &= \mathtt{K}_{00}(\phi) + \mathtt{K}_{10}(\phi)\cdot\zeta + (\mathtt{K}_{01}(\phi),w)_{L^2(\mathbb{T}_x)} + \frac{1}{2}\mathtt{K}_{20}(\phi)\zeta\cdot\zeta \\ &\quad + \big(\mathtt{K}_{11}(\phi)\zeta,w\big)_{L^2(\mathbb{T}_x)} + \frac{1}{2}\big(\mathtt{K}_{02}(\phi)w,w\big)_{L^2(\mathbb{T}_x)} + \mathtt{K}_{\geq 3}(\phi,\zeta,w)\end{aligned} \tag{6.1.14}$$

*where* $\mathtt{K}_{\geq 3}$ *collects the terms at least cubic in the variables* $(\zeta,w)$*, satisfy the following properties:*

*i) The vector field*

$$X_{\mathtt{K}}(\phi,0,0) = \begin{pmatrix}\mathtt{K}_{10}(\phi)\\ -\partial_\phi\mathtt{K}_{00}(\phi)\\ J\mathtt{K}_{01}(\phi)\end{pmatrix} = \begin{pmatrix}\omega\\0\\0\end{pmatrix} - \big(DG_\delta(\phi,0,0)\big)^{-1}Z_\delta(\phi) \tag{6.1.15}$$

*where* $Z_\delta(\phi) := \mathcal{F}(i_\delta)(\phi)$*. The functions* $\mathtt{K}_{00} : \mathbb{T}^{|\mathbb{S}|} \to \mathbb{R}$, $\mathtt{K}_{01} : \mathbb{T}^{|\mathbb{S}|} \to \mathbb{R}^{|\mathbb{S}|}$ *and* $\mathtt{K}_{01} : \mathbb{T}^{|\mathbb{S}|} \to H_{\mathbb{S}}^{\perp}$*, that we regard as an element of* $\mathcal{H}^s(\mathbb{T}^{|\mathbb{S}|}\times\mathbb{T}^d,\mathbb{R}^2)$*, satisfy the estimate*

$$\|\partial_\phi\mathtt{K}_{00}\|_{\mathrm{Lip},s} + \|\mathtt{K}_{10}-\omega\|_{\mathrm{Lip},s} + \|\mathtt{K}_{01}\|_{\mathrm{Lip},s} \lesssim_s \|Z\|_{\mathrm{Lip},s+\underline{\tau}} + \|Z\|_{\mathrm{Lip},s_0+\underline{\tau}}\|\mathfrak{I}\|_{\mathrm{Lip},s+\underline{\tau}}\,. \tag{6.1.16}$$

*ii) The average* $\langle\mathtt{K}_{20}\rangle := (2\pi)^{-|\mathbb{S}|}\int_{\mathbb{T}^{|\mathbb{S}|}}\mathtt{K}_{20}(\phi)d\phi$ *satisfies*

$$\|\langle\mathtt{K}_{20}\rangle - \varepsilon^2\mathscr{A}\|_{\mathrm{Lip}} \lesssim \varepsilon^4 \tag{6.1.17}$$

*where $\mathscr{A}$ is the twist matrix in* (1.2.9), *and*

$$\|\mathtt{K}_{20}\zeta\|_{\mathrm{Lip},s} \lesssim_s \varepsilon^2\big(\|\zeta\|_{\mathrm{Lip},s} + \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+\tau}\|\zeta\|_{\mathrm{Lip},s_0}\big) \tag{6.1.18}$$
$$\|\mathtt{K}_{11}\zeta\|_{\mathrm{Lip},s} \lesssim_s \varepsilon^2\big(\|\zeta\|_{\mathrm{Lip},s} + \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+\tau}\|\zeta\|_{\mathrm{Lip},s_0}\big) \tag{6.1.19}$$
$$\|\mathtt{K}_{11}^\top w\|_{\mathrm{Lip},s} \lesssim_s \varepsilon^2\big(\|w\|_{\mathrm{Lip},s} + \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+\tau}\|w\|_{\mathrm{Lip},s_0}\big)\,. \tag{6.1.20}$$

- **(Linearized operator in the new coordinates)** *The linearized operator*

$$\omega\cdot\partial_\varphi - d_{(\phi,\zeta,w)}X_{\mathtt{K}}(\phi,0,0)$$

*is*

$$\begin{pmatrix}\widehat\phi\\ \widehat\zeta\\ \widehat w\end{pmatrix} \mapsto \begin{pmatrix}\omega\cdot\partial_\varphi\widehat\phi - \partial_\phi\mathtt{K}_{10}(\phi)[\widehat\phi] - \mathtt{K}_{20}(\phi)\widehat\zeta - \mathtt{K}_{11}^\top(\phi)\widehat w\\ \omega\cdot\partial_\varphi\widehat\zeta + \partial_{\phi\phi}\mathtt{K}_{00}(\phi)[\widehat\phi] + [\partial_\phi\mathtt{K}_{10}(\phi)]^\top\widehat\zeta + [\partial_\phi\mathtt{K}_{01}(\phi)]^\top\widehat w\\ \omega\cdot\partial_\varphi\widehat w - J\{\partial_\phi\mathtt{K}_{01}(\phi)[\widehat\phi] + \mathtt{K}_{11}(\phi)\widehat\zeta + \mathtt{K}_{02}(\phi)\widehat w\}\end{pmatrix}. \tag{6.1.21}$$

In order to find an approximate inverse of the linear operator in (6.1.21) it is sufficient to invert the operator

$$\mathbb{D}\begin{pmatrix}\widehat\phi\\ \widehat\zeta\\ \widehat w\end{pmatrix} := \mathbb{D}(\underline{i})\begin{pmatrix}\widehat\phi\\ \widehat\zeta\\ \widehat w\end{pmatrix} := \begin{pmatrix}\omega\cdot\partial_\varphi\widehat\phi - \mathtt{K}_{20}(\phi)\widehat\zeta - \mathtt{K}_{11}^\top(\phi)\widehat w\\ \omega\cdot\partial_\varphi\widehat\zeta\\ \omega\cdot\partial_\varphi\widehat w - J\mathtt{K}_{02}(\phi)\widehat w - J\mathtt{K}_{11}(\phi)\widehat\zeta\end{pmatrix} \tag{6.1.22}$$

which is obtained by neglecting in (6.1.21) the terms $\partial_\phi\mathtt{K}_{10}$, $\partial_{\phi\phi}\mathtt{K}_{00}$, $\partial_\phi\mathtt{K}_{00}$, $\partial_\phi\mathtt{K}_{01}$, which vanish if $Z = 0$ by (6.1.16). The linear operator $\mathbb{D}(\underline{i})$ can be inverted in a "triangular" way. Indeed the second component in (6.1.22) for the action variable is decoupled from the others. Then one inverts the operator in the third component, i.e. the operator

$$\mathcal{L}_\omega := \mathcal{L}_\omega(\underline{i}) := \Pi_{\mathbb{S}}^\perp\big(\omega\cdot\partial_\varphi - J\mathtt{K}_{02}(\phi)\big)_{|H_{\mathbb{S}}^\perp}\,, \tag{6.1.23}$$

and finally the first one. The invertibility properties of $\mathcal{L}_\omega$ will be obtained in Proposition 11.1.1 using the results of Chapters 7-10. We now provide the explicit expression of $\mathcal{L}_\omega(\underline{i})$.

**Lemma 6.1.2. (Linearized operator in the normal directions)** *The linear operator* $\mathtt{K}_{02}(\phi) := \mathtt{K}_{02}(\underline{i};\phi)$ *has the form*

$$\mathtt{K}_{02}(\phi) = D_V + \varepsilon^2\mathtt{B}(\phi) + \mathtt{r}_\varepsilon(\phi)\,, \tag{6.1.24}$$

*where* $\mathtt{B} := \mathtt{B}(\varepsilon,\lambda)$ *is the self-adjoint operator*

$$\mathtt{B}\begin{pmatrix}Q\\ P\end{pmatrix} := \begin{pmatrix}\Pi_{\mathbb{S}}^\perp D_V^{-1/2}\big(3a(x)(v(\phi,0,\xi))^2 + \varepsilon 4a_4(x)(v(\phi,0,\xi))^3\big)D_V^{-1/2}Q\big)\\ 0\end{pmatrix} \tag{6.1.25}$$

*with the function $v$ defined in* (2.2.9)*, the functions $a(x)$ and $a_4(x)$ are in* (2.1.5)*, the vector $\xi = \xi(\lambda) \in [1,2]^{|\mathbb{S}|}$ in* (1.2.27)*, and $\mathtt{r}_\varepsilon := \mathtt{r}_\varepsilon(\underline{\mathfrak{I}})$ is a self-adjoint remainder satisfying*

$$|\mathtt{r}_\varepsilon|_{\mathrm{Lip},+,s_1} \lesssim_{s_1} \varepsilon^4 , \tag{6.1.26}$$

$$|\mathtt{r}_\varepsilon|_{\mathrm{Lip},+,s} \lesssim_s \varepsilon^2(\varepsilon^2 + \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+2}) . \tag{6.1.27}$$

*Moreover, given another torus $\underline{i}' = (\varphi, 0, 0) + \underline{\mathfrak{I}}'$ satisfying* (6.1.4)*, we have*

$$|\mathtt{r}_\varepsilon - \mathtt{r}'_\varepsilon|_{+,s_1} \lesssim_{s_1} \varepsilon^2 \|\underline{\mathfrak{I}} - \underline{\mathfrak{I}}'\|_{s_1+2} . \tag{6.1.28}$$

The next Chapters 7-10 will be devoted to obtain an approximate right inverse of $\mathcal{L}_\omega(\underline{i})$, as stated in Proposition 11.1.1. In Chapter 7 we shall conjugate $\mathcal{L}_\omega(\underline{i})$ to an operator (see (7.3.4)-(7.3.5) and (11.1.8)) which is in a suitable form to apply Proposition 8.2.1, and so proving Proposition 11.1.1. Proposition 8.2.1 is proved in Chapters 9 and 10.

The rest of this Chapter is devoted to the proof of Proposition 6.1.1 and Lemma 6.1.2.

## 6.2 Proof of Proposition 6.1.1

By (1.2.29), for all $\lambda \in \Lambda$ the frequency vector $\omega = (1 + \varepsilon^2\lambda)\bar{\omega}_\varepsilon$ satisfies the Diophantine condition

$$|\omega \cdot \ell| \geq \frac{\gamma_2}{\langle \ell \rangle^{\tau_1}} , \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|} \setminus \{0\} , \quad \text{where} \quad \gamma_2 = \gamma_1/2 = \gamma_0/4 . \tag{6.2.1}$$

We recall that the constant $\gamma_0$ in (1.2.6) depends only on the potential $V(x)$, and it is considered as a fixed $O(1)$ quantity, and thus we shall not track its dependence in the estimates.

An invariant torus $\underline{i}$ for the Hamiltonian vector field $X_K$, supporting a Diophantine flow, is isotropic (see e.g. Lemma 1 in [24] and Lemma C.1.2), namely the pull-back 1-form $\underline{i}^*\varkappa$ is closed, where $\varkappa$ is the Liouville 1-form defined in (2.2.15). This is equivalent to say that the 2-form

$$\underline{i}^*\mathcal{W} = \underline{i}^* d\varkappa = d(\underline{i}^*\varkappa) = 0$$

vanishes, where $\mathcal{W} = d\varkappa$ is defined in (2.2.10). Given an "approximately invariant" torus $\underline{i}$, the 1-form $\underline{i}^*\varkappa$ is only "approximately closed". In order to make this statement quantitative we consider

$$\begin{aligned} \underline{i}^*\varkappa &= \sum_{k=1,\ldots,|\mathbb{S}|} a_k(\varphi) d\varphi_k , \\ a_k(\varphi) &:= \big([\partial_\varphi \underline{\theta}(\varphi)]^\top \underline{y}(\varphi)\big)_k + \frac{1}{2}(\partial_{\varphi_k}\underline{z}(\varphi), J\underline{z}(\varphi))_{L^2(\mathbb{T}_x)} \end{aligned} \tag{6.2.2}$$

and we quantify how small is the pull-back 2-form

$$\underline{i}^*\mathcal{W} = d\,\underline{i}^*\varkappa = \sum_{k,j=1,\dots,|\mathbb{S}|,k<j} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j\,, \qquad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi)\,, \tag{6.2.3}$$

in terms of the error function $Z(\varphi)$ defined in (6.1.5).

**Lemma 6.2.1.** *The coefficients $A_{kj}$ in* (6.2.3) *satisfy*

$$\|A_{kj}\|_{\mathrm{Lip},s} \lesssim_s \|Z\|_{\mathrm{Lip},s+\tau_1+1} + \|Z\|_{\mathrm{Lip},s_0+1}\|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+\tau_1+1}\,. \tag{6.2.4}$$

Proof. The coefficients $A_{kj}$ satisfy the identity (see [24], Lemma 5, and (C.2.7))

$$\omega\cdot\partial_\varphi A_{kj} = \mathcal{W}\big(\partial_\varphi Z(\varphi)\underline{e}_k, \partial_\varphi \underline{i}(\varphi)\underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi \underline{i}(\varphi)\underline{e}_k, \partial_\varphi Z(\varphi)\underline{e}_j\big)$$

where $\underline{e}_k$ denote the $k$-th vector of the canonical basis of $\mathbb{R}^{|\mathbb{S}|}$. Then by (6.1.4) we get

$$\|\omega\cdot\partial_\varphi A_{kj}\|_{\mathrm{Lip},s} \lesssim_s \|Z\|_{\mathrm{Lip},s+1} + \|Z\|_{\mathrm{Lip},s_0+1}\|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1}\,.$$

Notice that the functions $A_{kj}(\varphi)$ defined in (6.2.3) have zero mean value in $\varphi$, so that

$$A_{kj}(\varphi) = (\omega\cdot\partial_\varphi)^{-1}\Big(\mathcal{W}\big(\partial_\varphi Z(\varphi)\underline{e}_k, \partial_\varphi \underline{i}(\varphi)\underline{e}_j\big) + \mathcal{W}\big(\partial_\varphi \underline{i}(\varphi)\underline{e}_k, \partial_\varphi Z(\varphi)\underline{e}_j\big)\Big)\,.$$

Now, since $\omega$ is Diophantine according to (6.2.1), by (2.3.7) we have

$$\|(\omega\cdot\partial_\varphi)^{-1} g\|_{\mathrm{Lip},s} \le C\|g\|_{\mathrm{Lip},s+\tau_1}$$

where the constant $\gamma_2 = \gamma_0/4$ is included in $C$ because it is considered a fixed constant $O(1)$. By the expression of $\mathcal{W}$ in (2.2.10), the tame estimate (3.5.1) and (6.1.4) we deduce (6.2.4). ∎

We now modify the approximate torus $\underline{i}$ to obtain an isotropic torus $i_\delta = i_\delta(\underline{i})$ nearby, which is still approximately invariant. We denote the Laplacian

$$\Delta_\varphi := \sum_{k=1,\dots,|\mathbb{S}|} \partial^2_{\varphi_k}\,.$$

**Lemma 6.2.2. (Isotropic torus)** *The torus $i_\delta(\varphi) := (\underline{\theta}(\varphi), y_\delta(\varphi), \underline{z}(\varphi))$ defined by*

$$y_\delta(\varphi) := \underline{y}(\varphi) - [\partial_\varphi\underline{\theta}(\varphi)]^{-T}\rho(\varphi)\,,\ \ \rho = (\rho_j)_{j\in\mathbb{S}}\,,\quad \rho_j(\varphi) := \Delta_\varphi^{-1}\sum_{k=1,\dots,|\mathbb{S}|}\partial_{\varphi_k}A_{kj}(\varphi) \tag{6.2.5}$$

*is isotropic. There is $\underline{\tau} := \underline{\tau}(|\mathbb{S}|,\tau_1)$ such that* (6.1.6)-(6.1.7) *hold. Moreover*

$$\|y_\delta - \underline{y}\|_{\mathrm{Lip},s} \lesssim_s \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1} \tag{6.2.6}$$

*and* (6.1.8) *holds.*

Along the section we denote by $\underline{\tau} := \underline{\tau}(|\mathbb{S}|, \tau_1)$ possibly different (larger) "loss of derivatives" constants.

Proof. The proof of the isotropy of the torus $i_\delta$ is in Lemma 6 of [24], see Lemma C.2.5. Let us prove the bounds (6.1.6)-(6.1.8) and (6.2.6). First notice that, since the map $A \mapsto A^{-1}$ is $C^\infty$ on the open set of invertible matrices $\{A \in M_{|\mathbb{S}|}(\mathbb{R}) \ : \ \det A \neq 0\}$, we derive from (6.1.4) and Lemma 3.5.5 that, for $\varepsilon$ small, the map

$$D\underline{\theta}^{-1} : \varphi \mapsto [D\underline{\theta}(\varphi)]^{-1}$$

satisfies the tame estimates

$$\|D\underline{\theta}^{-1}\|_{\mathrm{Lip},s} \lesssim_s 1 + \|\mathfrak{I}\|_{s+1}\,, \quad \forall s \geq s_0\,. \tag{6.2.7}$$

Then (6.1.6) and (6.2.6) follow by (6.2.5), (6.2.7), (6.2.2), (6.2.3), (6.2.4) and (6.1.4). Moreover, we have that the difference

$$\mathcal{F}(i_\delta) - \mathcal{F}(\underline{i}) = \begin{pmatrix} 0 \\ \omega\cdot\partial_\varphi(y_\delta - \underline{y}) \\ 0 \end{pmatrix} + \varepsilon^2\big(X_R(i_\delta) - X_R(\underline{i})\big)$$

and (6.1.7) follows by (6.1.6), Lemma 3.5.5 and (6.1.4). Finally the bound (6.1.8) follows by (6.2.5), (6.2.3), (6.2.2), (6.1.4). ∎

It is proved in [24] (see Lemma C.1.3) that the diffeomorphism $G_\delta : (\phi, \zeta, w) \to (\theta, y, z)$ defined in (6.1.9) is symplectic because the torus $i_\delta$ is isotropic (Lemma 6.2.2). By construction, (6.1.10) holds. Since $G_\delta$ is symplectic the Hamiltonian system generated by $K$ transforms as in (6.1.13) into the Hamiltonian system with Hamiltonian $\mathtt{K} = K \circ G_\delta$. By (5.1.3) the transformation $G_\delta$ in (6.1.9) is also reversibility preserving and so the Hamiltonian $\mathtt{K}$ is reversible, i.e. $\mathtt{K} \circ \tilde{S} = \mathtt{K}$.

**Lemma 6.2.3.** *The tame estimates* (6.1.11) *and* (6.1.12) *hold.*

Proof. We write (6.1.9) as

$$G_\delta \begin{pmatrix} \phi \\ \zeta \\ w \end{pmatrix} := \begin{pmatrix} \underline{\theta}(\phi) \\ y_\delta(\phi) + M(\phi)\zeta - \sum_{j=1}^{|\mathbb{S}|} (m_j(\phi), Jw)_{L^2_x} \underline{e}_j \\ \underline{z}(\phi) + w \end{pmatrix}$$

where $(\underline{e}_j)$ denotes the canonical basis of $\mathbb{R}^{|\mathbb{S}|}$ and we set

$$M(\phi) := [\partial_\phi \underline{\theta}(\phi)]^{-\top}\,, \quad m_j(\phi) := (\partial_{\theta_j} \tilde{\underline{z}})(\underline{\theta}(\phi)) = [(\partial_\varphi \underline{\theta})^{-\top}(\phi) \nabla \underline{z}(\phi)]_j\,.$$

The tame estimate (6.2.7) implies

$$\|M\|_{\mathrm{Lip},s} \lesssim_s 1 + \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1} \tag{6.2.8}$$

and using (3.5.1) and (6.1.4) we have

$$\begin{aligned}\|m_j\|_{\mathrm{Lip},s} = \|(\partial_{\theta_j}\underline{\tilde{z}})(\underline{\theta}(\phi))\|_{\mathrm{Lip},s} &\lesssim_s \|\underline{z}\|_{\mathrm{Lip},s+1} + \|\underline{z}\|_{\mathrm{Lip},s_0+1}\|\underline{\vartheta}\|_{\mathrm{Lip},s+1} \\ &\lesssim_s \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1}\,. \end{aligned}\tag{6.2.9}$$

Now

$$DG_\delta(\varphi,0,0)[\widehat{\imath}(\varphi)] := DG_\delta(\varphi,0,0)\begin{pmatrix}\widehat{\phi}(\varphi)\\ \widehat{\zeta}(\varphi)\\ \widehat{w}(\varphi)\end{pmatrix} = \begin{pmatrix}\widehat{a}(\varphi)\\ \widehat{b}(\varphi)\\ \widehat{c}(\varphi)\end{pmatrix}$$

where

$$\begin{aligned}\widehat{a} &:= \partial_\varphi\underline{\theta}(\varphi)[\widehat{\phi}]\ ,\\ \widehat{b} &:= \partial_\varphi y_\delta(\varphi)[\widehat{\phi}] + M(\varphi)[\widehat{\zeta}] - \sum_{j=1}^{|\mathbb{S}|}(m_j(\varphi), J\widehat{w})_{L^2_x}\underline{e}_j\ ,\\ \widehat{c} &:= \partial_\varphi\underline{z}(\varphi)[\widehat{\phi}] + \widehat{w}\end{aligned}\tag{6.2.10}$$

and

$$D^2G_\delta(\varphi,0,0)[\widehat{\imath}_1(\varphi),\widehat{\imath}_2(\varphi)] = \begin{pmatrix}\widehat{\alpha}(\varphi)\\ \widehat{\beta}(\varphi)\\ \widehat{\gamma}(\varphi)\end{pmatrix}$$

where

$$\begin{aligned}\widehat{\alpha} &:= \partial^2_\varphi\underline{\theta}[\widehat{\phi}_1,\widehat{\phi}_2]\,,\\ \widehat{\beta} &:= \partial^2_\varphi y_\delta[\widehat{\phi}_1,\widehat{\phi}_2] + \partial_\varphi M[\widehat{\phi}_1]\widehat{\zeta}_2 + \partial_\varphi M[\widehat{\phi}_2]\widehat{\zeta}_1\\ &\quad - \sum_{j=1}^{|\mathbb{S}|}\Big(\big(\partial_\varphi m_j[\widehat{\phi}_1], J\widehat{w}_2\big)_{L^2_x} + \big(\partial_\varphi m_j[\widehat{\phi}_2], J\widehat{w}_1\big)_{L^2_x}\Big)\underline{e}_j\,,\\ \widehat{\gamma} &:= \partial^2_\varphi\underline{z}[\widehat{\phi}_1,\widehat{\phi}_2]\,.\end{aligned}\tag{6.2.11}$$

The tame estimates (6.1.11) and (6.1.12) are a consequence of (6.2.10), (6.2.11), (3.5.1), (6.1.4), (6.2.6) and (6.2.8), (6.2.9). ■

Then we consider the Taylor expansion (6.1.14) of the Hamiltonian $\mathtt{K}$ at the trivial torus $(\phi,0,0)$. Notice that the Taylor coefficient $\mathtt{K}_{00}(\phi)\in\mathbb{R}$, $\mathtt{K}_{10}(\phi)\in\mathbb{R}^{|\mathbb{S}|}$, $\mathtt{K}_{01}(\phi)\in H^\perp_{\mathbb{S}}$, $\mathtt{K}_{20}(\phi)$ is a $|\mathbb{S}|\times|\mathbb{S}|$ real matrix, $\mathtt{K}_{02}(\phi)$ is a linear self-adjoint operator of $H^\perp_{\mathbb{S}}$ and $\mathtt{K}_{11}(\phi)\in\mathcal{L}(\mathbb{R}^{|\mathbb{S}|},H^\perp_{\mathbb{S}})$.

The Hamilton equations associated to (6.1.14) are

$$\begin{cases} \dot{\phi} = \mathtt{K}_{10}(\phi) + \mathtt{K}_{20}(\phi)\zeta + \mathtt{K}_{11}^{\top}(\phi)w + \partial_\zeta \mathtt{K}_{\geq 3}(\phi,\zeta,w) \\ \dot{\zeta} = -\partial_\phi \mathtt{K}_{00}(\phi) - [\partial_\phi \mathtt{K}_{10}(\phi)]^{\top}\zeta - [\partial_\phi \mathtt{K}_{01}(\phi)]^{\top} w \\ \qquad -\partial_\phi\big(\frac{1}{2}\mathtt{K}_{20}(\phi)\zeta\cdot\zeta + (\mathtt{K}_{11}(\phi)\zeta, w)_{L^2(\mathbb{T}_x)} + \frac{1}{2}(\mathtt{K}_{02}(\phi)w,w)_{L^2(\mathbb{T}_x)} + \mathtt{K}_{\geq 3}(\phi,\zeta,w)\big) \\ \dot{w} = J\big(\mathtt{K}_{01}(\phi) + \mathtt{K}_{11}(\phi)\zeta + \mathtt{K}_{02}(\phi)w + \nabla_w \mathtt{K}_{\geq 3}(\phi,\zeta,w)\big) \end{cases} \tag{6.2.12}$$

where $\partial_\phi \mathtt{K}_{10}^{\top}$ is the $|\mathbb{S}| \times |\mathbb{S}|$ transposed matrix and

$$\partial_\phi \mathtt{K}_{01}^{\top}(\phi)\,,\ \mathtt{K}_{11}^{\top}(\phi) : H_{\mathbb{S}}^{\perp} \to \mathbb{R}^{|\mathbb{S}|}\,, \quad \forall \phi \in \mathbb{T}^{|\mathbb{S}|}\,,$$

are defined by the duality relation

$$(\partial_\phi \mathtt{K}_{01}[\widehat{\phi}], w)_{L^2_x} = \widehat{\phi}\cdot[\partial_\phi \mathtt{K}_{01}]^{\top} w\,, \quad \forall \widehat{\phi} \in \mathbb{R}^{|\mathbb{S}|},\ w \in H_{\mathbb{S}}^{\perp}\,, \tag{6.2.13}$$

where $\cdot$ denotes the scalar product in $\mathbb{R}^{|\mathbb{S}|}$. The transposed operator $\mathtt{K}_{11}^{\top}$ is similarly defined and it turns out to be the following operator: for all $w \in H_{\mathbb{S}}^{\perp}$, and denoting $\underline{e}_k$ the $k$-th vector of the canonical basis of $\mathbb{R}^{|\mathbb{S}|}$,

$$\mathtt{K}_{11}^{\top}(\phi)w = \sum_{k=1,\dots,|\mathbb{S}|} \big(\mathtt{K}_{11}^{\top}(\phi)w\cdot\underline{e}_k\big)\underline{e}_k = \sum_{k=1,\dots,|\mathbb{S}|} \big(w, \mathtt{K}_{11}(\phi)\underline{e}_k\big)_{L^2(\mathbb{T}_x)}\underline{e}_k \in \mathbb{R}^{|\mathbb{S}|}\,. \tag{6.2.14}$$

The terms $\partial_\phi \mathtt{K}_{00}$, $\mathtt{K}_{10} - \omega$, $\mathtt{K}_{01}$ in the Taylor expansion (6.1.14) vanish if $Z = 0$.

**Lemma 6.2.4.** (6.1.15) *and* (6.1.16) *hold.*

Proof. Formula (6.1.15) is proved in Lemma 8 of [24] (see Lemma C.2.6). Then (6.1.4), (6.1.6), (6.1.7), (6.1.11) imply (6.1.16). ■

Notice that, if $\mathcal{F}(\underline{i}) = 0$, namely $\underline{i}(\varphi)$ is an invariant torus for the Hamiltonian vector field $X_K$, supporting quasi-periodic solutions with frequency $\omega$, then, by (6.1.15)-(6.1.16), the Hamiltonian in (6.1.14) simplifies to the KAM (variable coefficients) normal form

$$\mathtt{K} = const + \omega\cdot\zeta + \frac{1}{2}\mathtt{K}_{20}(\phi)\zeta\cdot\zeta + \big(\mathtt{K}_{11}(\phi)\zeta, w\big)_{L^2(\mathbb{T})} + \frac{1}{2}\big(\mathtt{K}_{02}(\phi)w, w\big)_{L^2(\mathbb{T})} + \mathtt{K}_{\geq 3}\,. \tag{6.2.15}$$

We now estimate $\mathtt{K}_{20}, \mathtt{K}_{11}$ in (6.1.14).

**Lemma 6.2.5.** (6.1.17)-(6.1.20) *hold.*

Proof.

Proof of (6.1.17)-(6.1.18). By Lemma 9 of [24] (see Lemma C.2.7) and the form of $K$ in (2.2.7), we have

$$\begin{aligned} \mathtt{K}_{20}(\phi) &= [\partial_\phi \underline{\theta}(\phi)]^{-1} \partial_{yy} K(i_\delta(\phi))[\partial_\varphi \underline{\theta}(\phi)]^{-\top} \\ &= \varepsilon^2 [\partial_\phi \underline{\theta}(\phi)]^{-1} \partial_{yy} R(i_\delta(\phi))[\partial_\phi \underline{\theta}(\phi)]^{-\top} \qquad (6.2.16) \\ &= \varepsilon^2 \partial_{yy} R(i_0(\phi)) + r_{20} \qquad (6.2.17) \end{aligned}$$

where $i_0(\phi) = (\phi, 0, 0)$ and

$$\begin{aligned} r_{20} := \varepsilon^2 &\Big( [\partial_\phi \underline{\theta}(\phi)]^{-1} \partial_{yy} R(i_\delta(\phi))[\partial_\phi \underline{\theta}(\phi)]^{-\top} - \partial_{yy} R(i_\delta(\phi)) \Big) \\ &+ \varepsilon^2 \Big( \partial_{yy} R(i_\delta(\phi)) - \partial_{yy} R(i_0(\phi)) \Big) . \end{aligned}$$

By Lemma 3.5.5, (6.1.4), (6.2.6) we have

$$\| \partial_{yy} R(i_\delta(\phi)) - \partial_{yy} R(i_0(\phi)) \|_{\mathrm{Lip}, s_0} \leq C(s_1) \varepsilon^2$$

and, using also $\| (\partial_\phi \underline{\theta}(\phi))^{-1} - \mathrm{Id}_{|\mathbb{S}|} \|_{\mathrm{Lip}, s_0} \leq \| \underline{\mathfrak{I}} \|_{\mathrm{Lip}, s_0 + 1} \lesssim_{s_1} \varepsilon^2$, that

$$\| [\partial_\phi \underline{\theta}(\phi)]^{-1} \partial_{yy} R(i_\delta(\phi))[\partial_\phi \underline{\theta}(\phi)]^{-\top} - \partial_{yy} R(i_\delta(\phi)) \|_{\mathrm{Lip}, s_0} \leq C(s_1) \varepsilon^2 .$$

Therefore

$$\| r_{20} \|_{\mathrm{Lip}, s_0} \leq C(s_1) \varepsilon^4 . \qquad (6.2.18)$$

Moreover, by (6.2.16) and Lemma 3.5.5 the norm of $\mathtt{K}_{20}$ (which is the sum of the norms of its $|\mathbb{S}| \times |\mathbb{S}|$ matrix entries) satisfies

$$\| \mathtt{K}_{20} \|_{\mathrm{Lip}, s} \lesssim_s \varepsilon^2 (1 + \| \underline{\mathfrak{I}} \|_{\mathrm{Lip}, s + \underline{\tau}})$$

and (6.1.18) follows by the tame estimates (3.5.1) for the product of functions.

Next, recalling the expression of $R$ in (2.2.8), with $G$ as in (2.1.7), by computations similar to those in section 5.2, it turns out that the average with respect to $\phi$ of $\partial_{yy} R(i_0(\phi))$ is

$$\langle \partial_{yy} R(i_0(\phi)) \rangle = \mathscr{A} + r \quad \text{with} \quad \| r \|_{\mathrm{Lip}} \leq C \varepsilon^2 \qquad (6.2.19)$$

where $\mathscr{A}$ is the twist matrix defined in (1.2.9) (in particular there is no contribution from $a_4(x) u^4$).

The estimate (6.1.17) follows by (6.2.17), (6.2.18) and (6.2.19).

Proof of (6.1.19)-(6.1.20). By Lemma 9 of [24] (see Lemma C.2.7) and the form of $K$ in (2.2.7) we have

$$\begin{aligned} \mathtt{K}_{11}(\phi) &= \partial_y \nabla_z K(i_\delta(\phi))[\partial_\varphi \underline{\theta}(\phi)]^{-\top} + J(\partial_\theta \underline{\tilde{z}})(\underline{\theta}(\phi))(\partial_{yy} K)(i_\delta(\phi))[\partial_\varphi \underline{\theta}(\phi)]^{-\top} \\ &= \varepsilon^2 \partial_y \nabla_z R(i_\delta(\phi))[\partial_\varphi \underline{\theta}(\phi)]^{-\top} + \varepsilon^2 J(\partial_\theta \underline{\tilde{z}})(\underline{\theta}(\phi))(\partial_{yy} R)(i_\delta(\phi))[\partial_\varphi \underline{\theta}(\phi)]^{-\top} , \qquad (6.2.20) \end{aligned}$$

and using (6.1.4), (6.2.6), we deduce (6.1.19). The bound (6.1.20) for $\mathtt{K}_{11}^\top$ follows by (6.2.14) and (6.1.19). ■

Finally formula (6.1.21) is obtained linearizing (6.2.12).

## 6.3 Proof of Lemma 6.1.2

We have to compute the quadratic term $\frac{1}{2}(\mathtt{K}_{02}(\phi)w, w)_{L^2(\mathbb{T}_x)}$ in the Taylor expansion (6.1.14) of the Hamiltonian $\mathtt{K}(\phi, 0, w)$. The operator $\mathtt{K}_{02}(\phi)$ is

$$\begin{aligned} \mathtt{K}_{02}(\phi) &= \partial_w \nabla_w \mathtt{K}(\phi, 0, 0) \\ &= \partial_w \nabla_w (K \circ G_\delta)(\phi, 0, 0) = D_V + \varepsilon^2 \partial_w \nabla_w (R \circ G_\delta)(\phi, 0, 0) \end{aligned} \tag{6.3.1}$$

where the Hamiltonian $K$ is defined in (2.2.7) and $G_\delta$ is the symplectic diffeomorphism (6.1.9). Differentiating with respect to $w$ the Hamiltonian

$$(R \circ G_\delta)(\phi, \zeta, w) = R(\underline{\theta}(\phi), y_\delta(\phi) + L_1(\phi)\zeta + L_2(\phi)w, \underline{z}(\phi) + w)$$

where, for brevity, we set

$$L_1(\phi) := [\partial_\phi \underline{\theta}(\phi)]^{-T}, \quad L_2(\phi) := -[\partial_\theta \underline{\tilde{z}}(\underline{\theta}(\phi))]^\top J, \quad \underline{\tilde{z}}(\theta) = \underline{z}(\underline{\theta}^{-1}(\theta)), \tag{6.3.2}$$

(see (6.1.9)), we get

$$\nabla_w (R \circ G_\delta)(\phi, \zeta, w) = L_2(\phi)^\top \partial_y R(G_\delta(\phi, \zeta, w)) + \nabla_z R(G_\delta(\phi, \zeta, w)).$$

Differentiating such identity with respect to $w$, and recalling (6.1.10), we get

$$\partial_w \nabla_w (R \circ G_\delta)(\phi, 0, 0) = \partial_z \nabla_z R(i_\delta(\phi)) + r(\phi) \tag{6.3.3}$$

with a self adjoint remainder $r(\phi) := r_1(\phi) + r_2(\phi) + r_3(\phi)$ given by

$$\begin{aligned} r_1(\phi) &:= L_2(\phi)^\top \partial_{yy} R(i_\delta(\phi)) L_2(\phi), \\ r_2(\phi) &:= L_2(\phi)^\top \nabla_z \partial_y R(i_\delta(\phi)), \\ r_3(\phi) &:= \partial_y \nabla_z R(i_\delta(\phi)) L_2(\phi). \end{aligned} \tag{6.3.4}$$

Each operator $r_1, r_2, r_3$ is the composition of at least one operator with "finite rank $\mathbb{R}^{|\mathbb{S}|}$ in the space variable", and therefore it has the "finite dimensional" form

$$r_l(\phi)[h] = \sum_{j=1,\ldots,|\mathbb{S}|} \big(h\,,\, g_j^{(l)}(\phi,\cdot)\big)_{L^2_x} \chi_j^{(l)}(\phi,\cdot), \quad \forall h \in H_{\mathbb{S}}^\perp,\ l = 1,2,3, \tag{6.3.5}$$

for functions $g_j^{(l)}(\phi,\cdot), \chi_j^{(l)}(\phi,\cdot) \in H_{\mathbb{S}}^\perp$. Indeed, writing the operator $L_2(\phi) : H_{\mathbb{S}}^\perp \to \mathbb{R}^{|\mathbb{S}|}$ as

$$L_2(\phi)[h] = \sum_{j=1,\ldots,|\mathbb{S}|} \big(h\,,\, L_2(\phi)^\top [\underline{e}_j]\big)_{L^2_x} \underline{e}_j, \quad \forall h \in H_{\mathbb{S}}^\perp,$$

we get, by (6.3.4),

$$r_1(\phi)[h] = \sum_{j=1,\dots,|\mathbb{S}|} \big(h\,,\, L_2(\phi)^\top[\underline{e}_j]\big)_{L^2_x} A_1[\underline{e}_j]\,, \quad A_1 := L_2(\phi)^\top \partial_{yy} R(i_\delta(\phi))\,, \tag{6.3.6}$$

$$r_2(\phi)[h] = \sum_{j=1,\dots,|\mathbb{S}|} \big(h, A_2^\top[\underline{e}_j]\big)_{L^2_x} L_2(\phi)^\top[\underline{e}_j]\,, \quad A_2 := \partial_z \partial_y R(i_\delta(\phi))\,, \tag{6.3.7}$$

$$r_3(\phi)[h] = \sum_{j=1,\dots,|\mathbb{S}|} \big(h, L_2(\phi)^\top[\underline{e}_j]\big)_{L^2_x} A_3[\underline{e}_j]\,, \quad A_3 := \partial_y \nabla_z R(i_\delta(\phi)) = A_2^\top\,. \tag{6.3.8}$$

**Lemma 6.3.1.** *For all $l = 1, 2, 3$, $j = 1, \dots, |\mathbb{S}|$, we have, for all $s \geq s_0$,*

$$\begin{aligned} \|g_j^{(l)}\|_{\mathrm{Lip},s} + \|\chi_j^{(l)}\|_{\mathrm{Lip},s} &\lesssim_s 1 + \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1} \\ \min\{\|g_j^{(l)}\|_{\mathrm{Lip},s}, \|\chi_j^{(l)}\|_{\mathrm{Lip},s}\} &\lesssim_s \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1}\,. \end{aligned} \tag{6.3.9}$$

PROOF. Recalling the expression of $L_2$ in (6.3.2) and (6.3.6)-(6.3.8) we have:
**i)** $g_j^{(1)}(\phi) = g_j^{(3)}(\phi) = \chi_j^{(2)}(\phi) = L_2(\phi)^\top[\underline{e}_j] = J(\partial_{\theta_j}\underline{\tilde{z}})(\underline{\theta}(\phi))$. Therefore by (6.2.9),

$$\|g_j^{(1)}\|_{\mathrm{Lip},s} = \|g_j^{(3)}\|_{\mathrm{Lip},s} = \|\chi_j^{(2)}\|_{\mathrm{Lip},s} \lesssim_s \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1}\,.$$

**ii)** $\chi_j^{(1)} = L_2(\phi)^\top \partial_{yy} R(i_\delta(\phi))[\underline{e}_j]$. Then, recalling (6.3.2) and using Lemma 3.5.5, we have

$$\begin{aligned} \|\chi_j^{(1)}\|_{\mathrm{Lip},s} &\lesssim_s \|\partial_\theta \underline{\tilde{z}}(\underline{\theta}(\phi))\|_{\mathrm{Lip},s}(1 + \|\mathfrak{I}_\delta\|_{\mathrm{Lip},s_0}) + \|\partial_\theta \underline{\tilde{z}}(\underline{\theta}(\phi))\|_{\mathrm{Lip},s_0}\|\mathfrak{I}_\delta\|_{\mathrm{Lip},s} \\ &\overset{(6.1.4)(6.2.9)(6.2.6)}{\lesssim_s} \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1}\,. \end{aligned}$$

**iii)** $\chi_j^{(3)} = g_j^{(2)} = \partial_{y_j} \nabla_z R(i_\delta((\phi))$. Then, using Lemma 3.5.5 and (6.2.6), we get

$$\|\chi_j^{(3)}\|_{\mathrm{Lip},s} = \|g_j^{(2)}\|_{\mathrm{Lip},s} \lesssim_s 1 + \|\mathfrak{I}_\delta\|_{\mathrm{Lip},s} \lesssim_s 1 + \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1}\,.$$

Items **i**) − **iii**) imply the estimates (6.3.9). ∎

We now use Lemmata 3.3.7 and 6.3.1 to derive bounds on the decay norms of the remainders $r_l$ defined in (6.3.5). Recalling Definition 3.3.4, we have to estimate the norm $|r_l|_{\mathrm{Lip},+,s} = |D_m^{1/2} r_l \Pi_{\mathbb{S}}^\perp D_m^{1/2}|_{\mathrm{Lip},s}$ of the extended operators, acting on the whole $\mathcal{H}^0 = L^2(\mathbb{T}^d \times \mathbb{T}^{|\mathbb{S}|}; \mathbb{C}^2)$, defined by

$$\begin{aligned} D_m^{1/2} r_l \Pi_{\mathbb{S}}^\perp D_m^{1/2} h &:= \sum_{j=1,\dots,|\mathbb{S}|} \big(\Pi_{\mathbb{S}}^\perp D_m^{1/2} h\,,\, g_j^{(l)}\big)_{L^2_x} (D_m^{1/2}\chi_j^{(l)}) \\ &= \sum_{j=1,\dots,|\mathbb{S}|} \big(h\,,\, D_m^{1/2} g_j^{(l)}\big)_{L^2_x} (D_m^{1/2}\chi_j^{(l)})\,, \quad h \in \mathcal{H}^0\,, \end{aligned} \tag{6.3.10}$$

where we used that $g_j^{(l)} \in H_{\mathbb{S}}^{\perp}$. Then the decay norm of $r = r_1 + r_2 + r_3$ satisfies

$$
\begin{aligned}
|r|_{\mathrm{Lip},+,s} &\lesssim \max_{l=1,2,3} |D_m^{1/2} r_l \Pi_{\mathbb{S}}^{\perp} D_m^{1/2}|_{\mathrm{Lip},s} \\
&\overset{(6.3.10),(3.3.29)}{\lesssim_s} \max_{l=1,2,3,j=1,\ldots,|\mathbb{S}|} \|D_m^{1/2} g_j^{(l)}\|_{\mathrm{Lip},s} \|D_m^{1/2}\chi_j^{(l)}\|_{\mathrm{Lip},s_0} + \|D_m^{1/2} g_j^{(l)}\|_{\mathrm{Lip},s_0} \|D_m^{1/2}\chi_j^{(l)}\|_{\mathrm{Lip},s} \\
&\overset{(6.3.9),(6.1.4)}{\lesssim_s} \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+2} \,. 
\end{aligned}
\tag{6.3.11}
$$

Finally, by (6.3.1), (6.3.3) we have

$$
\begin{aligned}
\mathtt{K}_{02}(\phi) &= D_V + \varepsilon^2 \partial_z \nabla_z R(i_\delta(\phi)) + \varepsilon^2 r(\phi) \\
&= D_V + \varepsilon^2 \partial_z \nabla_z R(\phi,0,0) + \varepsilon^2 \big(\partial_z \nabla_z R(i_\delta(\phi)) - \partial_z \nabla_z R(\phi,0,0)\big) + \varepsilon^2 r(\phi) \\
&= D_V + \varepsilon^2 \mathtt{B} + \mathtt{r}_\varepsilon
\end{aligned}
\tag{6.3.12}
$$

where $\mathtt{B}$ is defined in (6.1.25) and

$$
\mathtt{r}_\varepsilon := \varepsilon^2 \partial_z \nabla_z R(\phi,0,0) - \varepsilon^2 \mathtt{B} + \varepsilon^2 \big(\partial_z \nabla_z R(i_\delta(\phi)) - \partial_z \nabla_z R(\phi,0,0)\big) + \varepsilon^2 r(\phi) \,. \tag{6.3.13}
$$

This is formula (6.1.24). We now prove that $\mathtt{r}_\varepsilon$ satisfies (6.1.27). In (6.3.11) we have yet estimated $|r|_{\mathrm{Lip},+,s}$. Recalling Definition 3.3.4, and the expression of $R$ in (2.2.8) (see also (2.2.14)) we have to estimate the decay norm of the extended operator, acting on the whole $\mathcal{H}^0 = L^2(\mathbb{T}^d \times \mathbb{T}^{|\mathbb{S}|}; \mathbb{C}^2)$, defined as

$$
\begin{aligned}
&\partial_z \nabla_z R(i_\delta(\phi)) \Pi_{\mathbb{S}}^{\perp} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} := \\
&\begin{pmatrix} \Pi_{\mathbb{S}}^{\perp} D_V^{-\frac12} (\partial_u g)(\varepsilon, x, v(\underline{\theta}(\phi), y_\delta(\phi), \xi) + D_V^{-\frac12} \underline{Q}(\phi)) D_V^{-1/2} \Pi_{\mathbb{S}}^{\perp} h_1 \\ 0 \end{pmatrix}
\end{aligned}
\tag{6.3.14}
$$

where $g(\varepsilon, x, u) = \partial_u G(\varepsilon, x, u)$ is the nonlinearity in (1.2.2). Hence, by Proposition 3.4.6 and Lemma 3.3.8, we have

$$
\begin{aligned}
&|\partial_z \nabla_z R(i_\delta(\phi)) - \partial_z \nabla_z R(\phi,0,0)|_{\mathrm{Lip},+,s} \lesssim_s \\
&\big\|(\partial_u g)(\varepsilon, x, v(\underline{\theta}(\phi), y_\delta(\phi), \xi) + D_V^{-1/2} \underline{Q}(\phi)) - (\partial_u g)(\varepsilon, x, v(\phi, 0, \xi))\big\|_{\mathrm{Lip},s} \\
&\overset{\textit{Lemma } 3.5.5,(6.1.4)}{\lesssim_s} \|\mathfrak{I}_\delta\|_{\mathrm{Lip},s} \overset{(6.2.6)}{\lesssim_s} \|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+1} \,.
\end{aligned}
\tag{6.3.15}
$$

Moreover, recalling (2.1.6) and the definition of $\mathtt{B}$ in (6.1.25), we have

$$
|\partial_z \nabla_z R(\phi,0,0) - \mathtt{B}|_{\mathrm{Lip},+,s} \lesssim_s \varepsilon^2 \,. \tag{6.3.16}
$$

In conclusion, the operator $\mathtt{r}_\varepsilon$ defined in (6.3.13) satisfies, by (6.3.16), (6.3.15), (6.3.11), the estimate (6.1.27). In particular, (6.1.26) holds by (6.1.4).

There remains to prove (6.1.28). Let $\underline{i}' = (\varphi, 0, 0) + \underline{\mathfrak{I}}'$ be another torus embedding satisfying (6.1.4) and let $\mathtt{r}'_\varepsilon$ be the associated remainder in (6.3.13). We have

$$\mathtt{r}'_\varepsilon(\phi) - \mathtt{r}_\varepsilon(\phi) = \varepsilon^2 \big(\partial_z \nabla_z R(i'_\delta(\phi)) - \partial_z \nabla_z R(i_\delta(\phi))\big) + \varepsilon^2 (r'(\phi) - r(\phi)) \, .$$

By Lemma 3.5.5 and the expression (6.3.14) of $\partial_z \nabla_z R$,

$$|\partial_z \nabla_z R(i'_\delta(\phi)) - \partial_z \nabla_z R(i_\delta(\phi))|_{+,s_1} \lesssim_{s_1} \|\mathfrak{I}'_\delta - \mathfrak{I}_\delta\|_{s_1} \overset{(6.1.8)}{\lesssim_{s_1}} \|\underline{\mathfrak{I}}' - \underline{\mathfrak{I}}\|_{s_1+1} \, . \tag{6.3.17}$$

Moreover let ${g'}_j^{(l)}$ and ${\chi'}_j^{(l)}$ be the functions obtained in (6.3.5)-(6.3.8) from $i'$. Using again Lemma 3.5.5 as well as (6.1.8) and (6.2.6), we obtain, for any $j = 1, 2, 3$, $j = 1, \ldots, |\mathbb{S}|$,

$$\|{g'}_j^{(l)} - g_j^{(l)}\|_s + \|{\chi'}_j^{(l)} - \chi_j^{(l)}\|_s \lesssim_s \|\underline{\mathfrak{I}}' - \underline{\mathfrak{I}}\|_{s+1} + \|\underline{\mathfrak{I}}\|_{s+1} \|\underline{\mathfrak{I}}' - \underline{\mathfrak{I}}\|_{s_0+1} \, .$$

Hence, by Lemma 3.3.7, $|r' - r|_{+,s_1} \lesssim_{s_1} \|\underline{\mathfrak{I}}' - \underline{\mathfrak{I}}\|_{s_1+2}$. With (6.3.17), this gives (6.1.28).

# Chapter 7

# Splitting of low-high normal subspaces up to $O(\varepsilon^4)$

The main result of this Chapter is Proposition 7.3.1. Its goal is to transform the linear operator $\mathcal{L}_\omega$ in (6.1.23), into a form (see (7.3.5)) suitable to apply Proposition 8.2.1 in the next Chapter, which will enable to prove the existence of an approximate right inverse of $\mathcal{L}_\omega$ for most values of the parameter $\lambda$.

In the next section we fix the set $\mathbb{M}$ in the splitting $H_{\mathbb{S}}^\perp = H_{\mathbb{M}} \oplus H_{\mathbb{M}}^\perp$.

## 7.1 Choice of $\mathbb{M}$

We first remind that, by Lemma 6.1.2, the linear operator $\mathcal{L}_\omega$ defined in (6.1.23), acting in the normal subspace $H_{\mathbb{S}}^\perp$, has the form

$$\mathcal{L}_\omega = \omega \cdot \partial_\varphi - J(D_V + \varepsilon^2 \mathtt{B} + \mathtt{r}_\varepsilon) \, , \quad \omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon \, , \tag{7.1.1}$$

where $\mathtt{B} = \mathtt{B}(\varepsilon, \lambda)$ is the self-adjoint operator in (6.1.25) and the self-adjoint remainder $\mathtt{r}_\varepsilon$ satisfies (6.1.26). Recalling (6.1.25), (2.2.9), (1.2.27), we derive, for $\|a_4\|_{L^\infty} \leq 1$, the following bounds for the $L^2$-operatorial norm of $\mathtt{B}$:

$$\|\mathtt{B}\|_0 \leq C(\|a\|_{L^\infty} + \varepsilon) \, , \quad \|\mathtt{B}\|_{\mathrm{lip},0} \leq C(\|a\|_{L^\infty} + \varepsilon)\|\mathscr{A}^{-1}\| \tag{7.1.2}$$

where $\|\mathscr{A}^{-1}\|$ is some norm of the inverse twist matrix $\mathscr{A}^{-1}$.

Dividing (7.1.1) by $1+\varepsilon^2\lambda$ we now consider the operator

$$\begin{aligned} \frac{\mathcal{L}_\omega}{1+\varepsilon^2\lambda} &= \bar{\omega}_\varepsilon \cdot \partial_\varphi - J\Big(\mathtt{A} + \frac{\mathtt{r}_\varepsilon}{1+\varepsilon^2\lambda}\Big) \, , \\ \mathtt{A} := \mathtt{A}(\lambda) &= \frac{D_V}{1+\varepsilon^2\lambda} + \frac{\varepsilon^2 \mathtt{B}}{1+\varepsilon^2\lambda} \, . \end{aligned} \tag{7.1.3}$$

We denote by $\varrho$ the self-adjoint operator

$$\varrho := \frac{\varepsilon^2 \mathtt{B}}{1+\varepsilon^2\lambda} \tag{7.1.4}$$

and, according to the splitting $H_{\mathbb{S}}^{\perp} = H_{\mathbb{M}} \oplus H_{\mathbb{M}}^{\perp}$, and taking in $H_{\mathbb{M}}$ the basis

$$\{(\Psi_j(x), 0), (0, \Psi_j(x))\}_{j\in\mathbb{M}},$$

we represent $\mathtt{A}$ as (recall that $D_V\Psi_j = \mu_j\Psi_j$ and the notation (3.2.1))

$$\mathtt{A} = \begin{pmatrix} \mathrm{Diag}_{j\in\mathbb{M}} \frac{\mu_j}{1+\varepsilon^2\lambda}\mathrm{Id}_2 & 0 \\ 0 & \frac{D_V}{1+\varepsilon^2\lambda} \end{pmatrix} + \begin{pmatrix} \varrho_{\mathbb{M}}^{\mathbb{M}} & \varrho_{\mathbb{M}}^{\mathbb{M}^c} \\ \varrho_{\mathbb{M}^c}^{\mathbb{M}} & \varrho_{\mathbb{M}^c}^{\mathbb{M}^c} \end{pmatrix}. \tag{7.1.5}$$

Recalling (6.1.25), since the functions $a, a_4 \in C^\infty(\mathbb{T}^d)$, using (2.2.9), (1.2.27), and Proposition 3.4.6, we derive that the operator $\varrho$ defined in (7.1.4) satisfies

$$|\varrho|_{\mathrm{Lip},+,s} \le C(s)\varepsilon^2. \tag{7.1.6}$$

In the next lemma we fix the subset of indices $\mathbb{M}$. We recall the notation $\mathtt{A}_{\mathbb{M}^c}^{\mathbb{M}^c} = \Pi_{H_{\mathbb{M}^c}}\mathtt{A}_{|H_{\mathbb{M}^c}} = \Pi_{H_{\mathbb{M}}^{\perp}}\mathtt{A}_{|H_{\mathbb{M}}^{\perp}}$.

**Lemma 7.1.1. (Choice of $\mathbb{M}$)** *Let*

$$\Omega_j(\varepsilon,\lambda) := [\mathscr{B}\mathscr{A}^{-1}\bar\omega_\varepsilon]_j + \frac{\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_j}{1+\varepsilon^2\lambda}, \quad j\in\mathbb{M}, \tag{7.1.7}$$

*where $\mathscr{A}$, $\mathscr{B}$ are the "Birkhoff" matrices defined in* (1.2.9), (1.2.10)*, the $\mu_j$ are the unperturbed frequencies defined in* (1.1.5)*, and the vectors $\bar\mu$, $\bar\omega_\varepsilon$ are defined respectively in* (1.2.3), (1.2.25)*. There is a constant $C > 0$ such that, if $\mathbb{M} \subset \mathbb{N}$ contains the subset $\mathbb{F}$ defined in* (1.2.15)*, i.e. $\mathbb{F}\subset\mathbb{M}$, and*

$$\min_{j\in\mathbb{M}^c} \mu_j \ge C\left(1+\varepsilon^2\|a\|_{L^\infty} + \|a\|_{L^\infty}\|\mathscr{A}^{-1}\| + \varepsilon\|\mathscr{A}^{-1}\|\right), \tag{7.1.8}$$

*then*

$$\partial_\lambda \mathtt{A}_{\mathbb{M}^c}^{\mathbb{M}^c} \le -\Big(\max_{j\in\mathbb{F}}|\partial_\lambda\Omega_j| + \varepsilon^2\Big)\mathrm{Id}. \tag{7.1.9}$$

Proof. Differentiating the expression of $\mathtt{A}$ in (7.1.3) we get

$$\partial_\lambda\mathtt{A} := -\frac{\varepsilon^2 D_V}{(1+\varepsilon^2\lambda)^2} + \frac{\varepsilon^2\partial_\lambda\mathtt{B}}{1+\varepsilon^2\lambda} - \frac{\varepsilon^4\mathtt{B}}{(1+\varepsilon^2\lambda)^2}.$$

Then by (7.1.2) and the fact that $[D_V]_{\mathbb{M}^c}^{\mathbb{M}^c} \geq \min_{j\in\mathbb{M}^c} \mu_j$, we get

$$\partial_\lambda \mathtt{A}_{\mathbb{M}^c}^{\mathbb{M}^c} \leq -\frac{\varepsilon^2}{(1+\varepsilon^2\lambda)^2} \min_{j\in\mathbb{M}^c} \mu_j + C\varepsilon^4(\|a\|_{L^\infty}+\varepsilon) + C\varepsilon^2(\|a\|_{L^\infty}+\varepsilon)\|\mathscr{A}^{-1}\| . \tag{7.1.10}$$

Now, recalling (7.1.7), we have

$$\partial_\lambda \Omega_j(\varepsilon,\lambda) = \frac{-\varepsilon^2}{(1+\varepsilon^2\lambda)^2}\big(\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_j\big) \tag{7.1.11}$$

and so (7.1.10) and (7.1.11) imply

$$\begin{aligned}\partial_\lambda \mathtt{A}_{\mathbb{M}^c}^{\mathbb{M}^c} + \max_{j\in\mathbb{F}} |\partial_\lambda\Omega_j| \leq & -\frac{\varepsilon^2}{(1+\varepsilon^2\lambda)^2}\Big(\min_{j\in\mathbb{M}^c}\mu_j - \max_{j\in\mathbb{F}}\big|\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_j\big|\Big) \\ & + C\varepsilon^4(\|a\|_{L^\infty}+\varepsilon) + C\varepsilon^2(\|a\|_{L^\infty}+\varepsilon)\|\mathscr{A}^{-1}\| .\end{aligned}$$

Thus (7.1.9) holds taking $\mathbb{M}$ large enough such that (recall that $\mu_j \to +\infty$)

$$\min_{j\in\mathbb{M}^c}\mu_j \geq \max_{j\in\mathbb{F}}\big|\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_j\big| + C(1+\varepsilon^2\|a\|_{L^\infty} + \|a\|_{L^\infty}\|\mathscr{A}^{-1}\| + \varepsilon\|\mathscr{A}^{-1}\|) . \tag{7.1.12}$$

By (1.2.15), (1.2.13), the fact that $\mu_j \geq \sqrt{\beta}$ (see (1.1.5)), we get

$$\begin{aligned}\max_{j\in\mathbb{F}}\big|\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar\mu]_j\big| \leq \mathfrak{g} &= \max_{j\in\mathbb{S}^c}\{(\mathscr{B}\mathscr{A}^{-1}\bar\mu)_j - \mu_j\} \\ &\leq C(\|a\|_{L^\infty}\|\mathscr{A}^{-1}\|+1)\end{aligned} \tag{7.1.13}$$

for some $C := C(\beta,\mathbb{S})$, having used that (1.2.10) we have $\sup_{k\in\mathbb{S}}|G_k^j| \leq C\|a\|_{L^\infty}\beta$. By (7.1.12)-(7.1.13) and (7.1.8) we deduce (7.1.9). ■

In the sequel of the Monograph the subset $\mathbb{M}$ is kept fixed. Note that the condition (7.1.8) can be fulfilled taking $\mathbb{M}$ large enough because $\mu_j \to +\infty$ as $j \to +\infty$.

In the next part of the Chapter we perform one step of averaging to eliminate, as much as possible, the terms of order $O(\varepsilon^2)$ of $\varrho_{\mathbb{M}}^{\mathbb{M}}$, $\varrho_{\mathbb{M}}^{\mathbb{M}^c}$, $\varrho_{\mathbb{M}^c}^{\mathbb{M}}$ in (7.1.5).

## 7.2 Homological equations

According to the splitting $H = H_{\mathbb{M}} \oplus H_{\mathbb{M}}^\perp$ we consider the linear map

$$\mathtt{S} \mapsto J\bar\mu\cdot\partial_\varphi \mathtt{S} + [J\mathtt{S}, JD_V]$$

where $\bar\mu \in \mathbb{R}^{|\mathbb{S}|}$ is the unperturbed tangential frequency vector defined in (1.2.3), and

$$\begin{aligned}&\mathtt{S}(\varphi) = \begin{pmatrix}\mathtt{d}(\varphi) & \mathtt{a}(\varphi)^* \\ \mathtt{a}(\varphi) & 0\end{pmatrix} \in \mathcal{L}(H_{\mathbb{S}}^\perp) , \ \forall \varphi \in \mathbb{T}^{|\mathbb{S}|} , \\ &\mathtt{d}(\varphi) = \mathtt{d}^*(\varphi) \in \mathcal{L}(H_{\mathbb{M}}) , \ \mathtt{a}(\varphi) \in \mathcal{L}(H_{\mathbb{M}}, H_{\mathbb{M}}^\perp)\end{aligned} \tag{7.2.1}$$

is self-adjoint.

Since $D_V$ and $J$ commute, we have

$$\begin{aligned} & J\bar\mu \cdot \partial_\varphi \mathtt{S} + [J\mathtt{S}, JD_V] \\ & = \begin{pmatrix} J\bar\mu \cdot \partial_\varphi \mathtt{d} + D_V \mathtt{d} + J\mathtt{d} J D_V & J\bar\mu \cdot \partial_\varphi \mathtt{a}^* + D_V \mathtt{a}^* + J\mathtt{a}^* J D_V \\ J\bar\mu \cdot \partial_\varphi \mathtt{a} + D_V \mathtt{a} + J\mathtt{a} J D_V & 0 \end{pmatrix} . \end{aligned} \tag{7.2.2}$$

Recalling the definition of $\Pi_{\mathtt{D}}$ in (3.2.28) (with $\mathbb{F} \rightsquigarrow \mathbb{M}$) and of $\Pi_{\mathtt{0}}$ in (3.2.30) we decompose the self-adjoint operator $\varrho = \varepsilon^2 \mathtt{B}(1+\varepsilon^2\lambda)^{-1}$ defined in (7.1.4) as

$$\varrho = \Pi_{\mathtt{D}} \varrho + \Pi_{\mathtt{0}} \varrho \, . \tag{7.2.3}$$

The term $\Pi_{\mathtt{0}}\varrho$ has the form

$$\begin{aligned} & \Pi_{\mathtt{0}} \varrho(\varphi) = \begin{pmatrix} \varrho_1(\varphi) & \varrho_2(\varphi)^* \\ \varrho_2(\varphi) & 0 \end{pmatrix} \in \mathcal{L}(H_{\mathbb{S}}^\perp) \, , \\ & \varrho_1(\varphi) \in \mathcal{L}(H_{\mathbb{M}}) \, , \ \varrho_2(\varphi) \in \mathcal{L}(H_{\mathbb{M}}, H_{\mathbb{M}}^\perp) \, , \end{aligned} \tag{7.2.4}$$

where $\varrho_1(\varphi) = \varrho_1^*(\varphi)$, and, recalling (3.2.29), (3.2.18), the $\varphi$-average

$$[\widehat{\varrho}_1]_j^j(0) = \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} (\widehat{\varrho}_1)_j^j(\varphi) \, d\varphi \in M_- \, , \qquad \forall j \in \mathbb{M} \, . \tag{7.2.5}$$

The aim is to solve the "homological" equation

$$J\bar\mu \cdot \partial_\varphi \mathtt{S} + [J\mathtt{S}, JD_V] = J\Pi_{\mathtt{0}} \varrho \tag{7.2.6}$$

which, recalling (7.2.2), (7.2.4), amounts to solve the decoupled pair of equations

$$J\bar\mu \cdot \partial_\varphi \mathtt{d} + D_V \mathtt{d} + J\mathtt{d} J D_V = J\varrho_1 \tag{7.2.7}$$

$$J\bar\mu \cdot \partial_\varphi \mathtt{a} + D_V \mathtt{a} + J\mathtt{a} J D_V = J\varrho_2 \, . \tag{7.2.8}$$

Note that, taking the adjoint equation of (7.2.8), multiplying by $J$ on the left and the right, and since $J$ and $D_V$ commute, we obtain also the equation in top right in (7.2.6), see (7.2.2) and (7.2.4).

The arguments of this section are similar to those developed in section 9.3, actually simpler because the equation (7.2.8) has constant coefficients in $\varphi$, unlike the corresponding equation (9.2.6) where the operator $V_0$ depends on $\varphi \in \mathbb{T}^{|\mathbb{S}|}$. Thus in the sequel we shall often refer to section 9.3.

We first find a solution $\mathtt{d}$ of the equation (7.2.7). We recall that a linear operator $\mathtt{d}(\varphi) \in \mathcal{L}(H_{\mathbb{M}})$ is represented by a finite dimensional square matrix $(\mathtt{d}_i^j(\varphi))_{i,j\in\mathbb{M}}$ with entries $\mathtt{d}_i^j(\varphi) \in \mathcal{L}(H_j, H_i) \simeq \mathrm{Mat}_2(\mathbb{R})$. To solve (7.2.7) we use the second order Melnikov non-resonance conditions (1.2.16)-(1.2.19), that depend just on the unperturbed linear frequencies defined by (1.1.5).

**Lemma 7.2.1. (Homological equation** (7.2.7)**)** *Assume the second order Melnikov non resonance conditions* (1.2.16)-(1.2.19). *Then the equation* (7.2.7) *has a solution* $\mathtt{d}(\varphi) = (\mathtt{d}_i^j(\varphi))_{i,j\in\mathbb{M}}$, $\mathtt{d}(\varphi) = \mathtt{d}^*(\varphi)$, *satisfying*

$$\|\mathtt{d}_i^j\|_{\mathrm{Lip},H^s(\mathbb{T}^{|\mathbb{S}|})} \le C\|(\varrho_1)_i^j\|_{\mathrm{Lip},H^{s+2\tau_0}(\mathbb{T}^{|\mathbb{S}|})}\,, \quad \forall i,j\in\mathbb{M}\,. \tag{7.2.9}$$

PROOF. Since the symplectic operator $J$ leaves invariant each subspace $H_j$ and recalling (2.1.11), the equation (7.2.7) is equivalent to

$$J\bar\mu\cdot\partial_\varphi \mathtt{d}_i^j(\varphi) + \mu_i\mathtt{d}_i^j(\varphi) + \mu_j J\mathtt{d}_i^j(\varphi)J = J(\varrho_1(\varphi))_i^j\,,\ \forall i,j\in\mathbb{M}\,,\quad J = \begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}\,,$$

and, by a Fourier series expansion with respect to the variable $\varphi\in\mathbb{T}^{|\mathbb{S}|}$ writing

$$d_i^j(\varphi) = \sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|}} \widehat{d}_i^j(\ell)e^{\mathrm{i}\ell\cdot\varphi}\,,\quad \widehat{d}_i^j(\ell)\in\mathrm{Mat}_2(\mathbb{C})\,,\quad \overline{\widehat{d}_i^j(\ell)} = \widehat{d}_i^j(-\ell)\,,$$

to

$$\mathrm{i}(\bar\mu\cdot\ell)J\widehat{\mathtt{d}}_i^j(\ell) + \mu_i\widehat{\mathtt{d}}_i^j(\ell) + \mu_j J\widehat{\mathtt{d}}_i^j(\ell)J = J[\widehat{\varrho}_1]_i^j(\ell)\,,\ \forall i,j\in\mathbb{M}\,,\ \ell\in\mathbb{Z}^{|\mathbb{S}|}\,. \tag{7.2.10}$$

Using the second order Melnikov non resonance conditions (1.2.16)-(1.2.19), and since, by (7.2.5), the Fourier coefficients $[\widehat{\varrho}_1]_j^j(0)\in M_-$, the equations (7.2.10) can be solved arguing as in Lemma 9.3.2 and the estimate (7.2.9) follows by standard arguments. ∎

We now solve the equation (7.2.8) in the unknown $\mathtt{a}\in\mathcal{L}(H_{\mathbb{M}},H_{\mathbb{M}}^\perp)$. We use again the second order Melnikov non-resonance conditions (1.2.16)-(1.2.19).

**Lemma 7.2.2. (Homological equation** (7.2.8)**)** *Assume the second order Melnikov non resonance conditions* (1.2.16)-(1.2.19). *Then the homological equation* (7.2.8) *has a solution* $\mathtt{a}\in L^2(\mathbb{T}^{|\mathbb{S}|},\mathcal{L}(H_{\mathbb{M}},H_{\mathbb{M}}^\perp))$ *satisfying*

$$|\mathtt{a}|_{\mathrm{Lip},s} \le C(s)\|\varrho_2\|_{\mathrm{Lip},s+2\tau_0}\,. \tag{7.2.11}$$

PROOF. Writing $\mathtt{a}^j(\varphi) := \mathtt{a}(\varphi)_{|H_j}$, $\varrho_2^j(\varphi) := (\varrho_2(\varphi))_{|H_j}\in\mathcal{L}(H_j,H_{\mathbb{M}}^\perp)$ and recalling (2.1.11), the equation (7.2.8) amounts to

$$J\bar\mu\cdot\partial_\varphi\mathtt{a}^j(\varphi) + D_V\mathtt{a}^j(\varphi) + \mu_j J\mathtt{a}^j(\varphi)J = J\varrho_2^j(\varphi)\,,\quad \forall j\in\mathbb{M}\,. \tag{7.2.12}$$

Writing, by a Fourier series expansion with respect to $\varphi\in\mathbb{T}^{|\mathbb{S}|}$,

$$\begin{aligned}\mathtt{a}^j(\varphi) &= \sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|}}\widehat{\mathtt{a}}^j(\ell)e^{\mathrm{i}\ell\cdot\varphi}\,,\ \widehat{\mathtt{a}}^j(\ell) = \frac{1}{(2\pi)^{|\mathbb{S}|}}\int_{\mathbb{T}^{|\mathbb{S}|}}\mathtt{a}^j(\varphi)e^{-\mathrm{i}\ell\cdot\varphi}\,d\varphi\,,\\ \varrho_2^j(\varphi) &= \sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|}} J\widehat{\varrho}_2^j(\ell)e^{\mathrm{i}\ell\cdot\varphi}\,,\end{aligned}$$

the equation (7.2.12) amounts to

$$\mathrm{i}\bar{\mu}\cdot\ell J\widehat{\mathsf{a}}^j(\ell)+D_V\widehat{\mathsf{a}}^j(\ell)+\mu_j J\widehat{\mathsf{a}}^j(\ell)J=J\widehat{\varrho}_2^j(\ell)\,,\quad \forall j\in\mathbb{M}\,,\ \ell\in\mathbb{Z}^{|\mathbb{S}|}\,. \tag{7.2.13}$$

According to the $L^2$-orthogonal splitting $H_{\mathbb{M}}^{\perp}=\oplus_{j\in\mathbb{M}^c}H_j$ the linear operator $\widehat{\mathsf{a}}^j(\ell)$, which maps $H_j$ into the complexification of $H_{\mathbb{M}}^{\perp}$ and satisfies $\overline{\widehat{\mathsf{a}}^j(\ell)}=\widehat{\mathsf{a}}^j(-\ell)$, is identified (as in (3.2.11)-(3.2.12) with index $k\in\mathbb{M}^c$) with a sequence of $2\times 2$ matrices

$$(\widehat{\mathsf{a}}_k^j(\ell))_{k\in\mathbb{M}^c}\,,\quad \widehat{\mathsf{a}}_k^j(\ell)\in\mathrm{Mat}_2(\mathbb{C})\,,\quad \overline{\widehat{\mathsf{a}}_k^j(\ell)}=\widehat{\mathsf{a}}_k^j(-\ell)\,. \tag{7.2.14}$$

Similarly $\widehat{\varrho}_2^j(\ell)\equiv([\widehat{\varrho}_2]_k^j(\ell))_{k\in\mathbb{M}^c}$. Thus (7.2.13) amounts to the following sequence of equations

$$\begin{gathered}\mathrm{i}\bar{\mu}\cdot\ell J\widehat{\mathsf{a}}_k^j(\ell)+\mu_k\widehat{\mathsf{a}}_k^j(\ell)+\mu_j J\widehat{\mathsf{a}}_k^j(\ell)J=J[\widehat{\varrho}_2]_k^j(\ell)\,,\\ j\in\mathbb{M}\,,\ k\in\mathbb{M}^c\,,\ \ell\in\mathbb{Z}^{|\mathbb{S}|}\,,\ J=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\,.\end{gathered} \tag{7.2.15}$$

Note that the equation (7.2.15) is like (7.2.10). Since $j\neq k$ (indeed $j\in\mathbb{M}$, $k\in\mathbb{M}^c$), by the second order Melnikov non resonance conditions (1.2.16)-(1.2.19), each equation (7.2.15) has a unique solution for any $\varrho_2$, the reality condition (7.2.14) holds, and

$$\|\widehat{\mathsf{a}}_k^j(\ell)\|\le C\langle\ell\rangle^{\tau_0}\|[\widehat{\varrho}_2]_k^j(\ell)\|\,. \tag{7.2.16}$$

We now estimate $|\mathsf{a}|_s$. By Lemma 3.3.10 we have $|\mathsf{a}|_s\simeq_s\|\mathsf{a}\|_s$ (we identify each $\mathsf{a}^j\in\mathcal{L}(H_j,H_{\mathbb{M}}^{\perp})$ with a function of $H_{\mathbb{M}}^{\perp}\times H_{\mathbb{M}}^{\perp}$ as in (3.2.6), (3.2.7)). Given a function

$$u(\varphi,x)=\sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|},k\in\mathbb{M}^c}u_{\ell,k}e^{\mathrm{i}\ell\cdot\varphi}\Psi_k(x)\in H_{\mathbb{M}}^{\perp}$$

the Sobolev norm $\|u\|_s$ defined in (3.3.2) is equivalent, using (3.1.4), (3.1.3), to

$$\begin{aligned}\|u\|_s^2&\simeq_s\|u\|^2_{L^2_\varphi(H^s_x\cap H_{\mathbb{M}}^{\perp})}+\|u\|^2_{H^s_\varphi(L^2_x\cap H_{\mathbb{M}}^{\perp})}\\&\simeq_s\sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|},k\in\mathbb{M}^c}\big(\mu_k^{2s}+\langle\ell\rangle^{2s}\big)|u_{\ell,k}|^2\,.\end{aligned} \tag{7.2.17}$$

In conclusion (7.2.17) and (7.2.16) and the fact that by Young inequality

$$\begin{gathered}\langle\ell\rangle^{2\tau_0}\mu_k^{2s}\le\frac{\langle\ell\rangle^{2\tau_0 p}}{p}+\frac{\mu_k^{2sq}}{q}\lesssim_{s,\tau_0}\langle\ell\rangle^{2(\tau_0+s)}+\mu_k^{2(\tau_0+s)}\,,\\ p:=\frac{2(\tau_0+s)}{2\tau_0}\,,\ q:=\frac{2(\tau_0+s)}{2s}\,,\end{gathered}$$

imply $|\mathsf{a}|_s\le C(s)\|\varrho_2\|_{s+\tau_0}$. The estimate (7.2.11) for the Lipschitz norm follows as usual. ■

## 7.3 Averaging step

We consider the family of invertible symplectic transformations

$$\mathtt{P}(\varphi) := e^{J\mathtt{S}(\varphi)}\,, \quad \mathtt{P}^{-1}(\varphi) := e^{-J\mathtt{S}(\varphi)}\,, \quad \varphi \in \mathbb{T}^{|\mathbb{S}|}\,, \tag{7.3.1}$$

where $\mathtt{S} := \mathtt{S}(\varphi)$ is the self-adjoint operator in $\mathcal{L}(H_{\mathbb{S}}^{\perp})$ of the form (7.2.1) with $\mathtt{d}(\varphi)$, $\mathtt{a}(\varphi)$ defined in Lemmata 7.2.1, 7.2.2. By Lemma 3.3.10, the estimates (7.2.9), (7.2.11) and (7.1.6) imply

$$|\mathtt{S}|_{\mathrm{Lip},+,s} \lesssim_s |\mathtt{S}|_{\mathrm{Lip},s+\frac{1}{2}} \lesssim_s \|\varrho\|_{\mathrm{Lip},s+2\tau_0+1} \leq C(s)\varepsilon^2 \tag{7.3.2}$$

and the transformation $\mathtt{P}(\varphi)$ in (7.3.1) satisfies, for $\varepsilon$ small, the estimates

$$|\mathtt{P}|_{\mathrm{Lip},s_1} \leq 2\,, \quad |\mathtt{P}|_{\mathrm{Lip},s}\,,\ |\mathtt{P}^{-1}|_{\mathrm{Lip},s} \leq C(s)\,, \quad \forall s \geq s_1\,. \tag{7.3.3}$$

In the next proposition, which is the main result of this Chapter, we conjugate the whole operator $\bar\omega_\varepsilon \cdot \partial_\varphi - J(\mathtt{A} + \dfrac{\mathtt{r}_\varepsilon}{1+\varepsilon^2\lambda})$ defined in (7.1.3) by $\mathtt{P}(\varphi)$.

**Proposition 7.3.1. (Averaging)** *Assume the second order Melnikov non resonance conditions* (1.2.16)-(1.2.19) *where the set* $\mathbb{M}$ *is fixed in Lemma 7.1.1. Let* $\mathtt{P}(\varphi)$ *be the symplectic transformation* (7.3.1) *of* $H_{\mathbb{S}}^{\perp}$*, where* $\mathtt{S}(\varphi)$ *is the self-adjoint operator of the form* (7.2.1) *with* $\mathtt{d}(\varphi)$, $\mathtt{a}(\varphi)$ *defined in Lemmata 7.2.1, 7.2.2. Then the conjugated operator*

$$\mathtt{P}^{-1}(\varphi)\Big[\bar\omega_\varepsilon \cdot \partial_\varphi - J\Big(\mathtt{A} + \frac{\mathtt{r}_\varepsilon}{1+\varepsilon^2\lambda}\Big)\Big]\mathtt{P}(\varphi) = \bar\omega_\varepsilon \cdot \partial_\varphi - J\mathtt{A}^+ \tag{7.3.4}$$

*acting in the normal subspace* $H_{\mathbb{S}}^{\perp}$*, has the following form, with respect to the splitting* $H_{\mathbb{S}}^{\perp} = H_{\mathbb{M}} \oplus H_{\mathbb{M}}^{\perp}$*,*

$$\mathtt{A}^+ = \mathtt{A}_0 + \varrho^+\,, \quad \mathtt{A}_0 := \frac{D_V}{1+\varepsilon^2\lambda} + \Pi_{\mathtt{D}}\varrho = \begin{pmatrix} D(\varepsilon,\lambda) & 0 \\ 0 & \mathtt{A}_{\mathbb{M}^c}^{\mathbb{M}^c} \end{pmatrix}\,, \tag{7.3.5}$$

*where, in the basis* $\{(\Psi_j,0),(0,\Psi_j)\}_{j\in\mathbb{M}}$ *of* $H_{\mathbb{M}}$*, the operator* $D(\varepsilon,\lambda)$ *is represented by the diagonal matrix*

$$D(\varepsilon,\lambda) = \mathrm{Diag}_{j\in\mathbb{M}}\Omega_j(\varepsilon,\lambda)\mathrm{Id}_2\,, \quad \mathrm{Id}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{7.3.6}$$

*with* $\Omega_j(\varepsilon,\lambda)$ *defined in* (7.1.7)*,* $\mathtt{A}$ *is defined in* (7.1.3)*, and*

$$|\varrho^+|_{\mathrm{Lip},+,s} \lesssim_s \varepsilon^4 + |\mathtt{r}_\varepsilon|_{\mathrm{Lip},+,s}\,. \tag{7.3.7}$$

*Moreover, given another self-adjoint operator* $\mathtt{r}'_\varepsilon$ *satisfying* (6.1.26)*, we have that*

$$|\varrho^+ - (\varrho^+)'|_{+,s_1} \lesssim_{s_1} |\mathtt{r}_\varepsilon - \mathtt{r}'_\varepsilon|_{+,s_1}\,. \tag{7.3.8}$$

The rest of this Chapter is dedicated to prove Proposition 7.3.1. We first study the conjugated operator

$$\mathtt{P}^{-1}(\varphi)(\bar{\omega}_\varepsilon \cdot \partial_\varphi - J\mathtt{A})\mathtt{P}(\varphi)$$

where $\mathtt{A}$ is defined in (7.1.3). We have the Lie series expansion

$$\mathtt{P}^{-1}(\varphi)(\bar{\omega}_\varepsilon \cdot \partial_\varphi - J\mathtt{A})\mathtt{P}(\varphi) = (\bar{\omega}_\varepsilon \cdot \partial_\varphi - J\mathtt{A}) + \mathrm{Ad}_{(-J\mathtt{S})}(\mathtt{X}_0) + \sum_{k\geq 2} \frac{1}{k!}\mathrm{Ad}^k_{(-J\mathtt{S})}(\mathtt{X}_0) \tag{7.3.9}$$

where $\mathtt{X}_0 := \bar{\omega}_\varepsilon \cdot \partial_\varphi - J\mathtt{A}$. Recalling that $\mathtt{A} = (1+\varepsilon^2\lambda)^{-1}D_V + \varrho$ by (7.1.3)-(7.1.4), we expand the commutator as

$$\begin{aligned} \mathrm{Ad}_{(-J\mathtt{S})}(X_0) &= \bar{\omega}_\varepsilon \cdot \partial_\varphi(J\mathtt{S}) + [J\mathtt{S}, J\mathtt{A}] \\ &= J(\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathtt{S}) + (1+\varepsilon^2\lambda)^{-1}[J\mathtt{S}, JD_V] + [J\mathtt{S}, J\varrho] \\ &= J(\bar{\mu} \cdot \partial_\varphi \mathtt{S}) + [J\mathtt{S}, JD_V] - \varepsilon^2\lambda(1+\varepsilon^2\lambda)^{-1}[J\mathtt{S}, JD_V] + [J\mathtt{S}, J\varrho] \\ &\quad + J(\bar{\omega}_\varepsilon - \bar{\mu})\cdot\partial_\varphi \mathtt{S} \\ &\overset{(7.2.6),(1.2.25)}{=} J\Pi_{\mathtt{D}}\varrho - \varepsilon^2\lambda(1+\varepsilon^2\lambda)^{-1}[J\mathtt{S}, JD_V] + [J\mathtt{S}, J\varrho] + \varepsilon^2(J\zeta\cdot\partial_\varphi)\mathtt{S}\,. \end{aligned} \tag{7.3.10}$$

As a consequence of (7.3.9), (7.3.10), (7.2.3), (7.1.3)-(7.1.4) we deduce (7.3.4) with

$$\mathtt{A}^+ = \mathtt{A}_0 + \varrho^+\,, \quad \mathtt{A}_0 := (1+\varepsilon^2\lambda)^{-1}D_V + \Pi_{\mathtt{D}}\varrho \tag{7.3.11}$$

and

$$\begin{aligned} -J\varrho^+ &:= -\varepsilon^2\lambda(1+\varepsilon^2\lambda)^{-1}[J\mathtt{S}, JD_V] + [J\mathtt{S}, J\varrho] + \varepsilon^2(J\zeta\cdot\partial_\varphi)\mathtt{S} \\ &\quad - (1+\varepsilon^2\lambda)^{-1}\mathtt{P}^{-1}J\mathtt{r}_\varepsilon\mathtt{P} + \sum_{k\geq 2}\frac{1}{k!}\mathrm{Ad}^k_{(-J\mathtt{S})}(\mathtt{X}_0)\,. \end{aligned} \tag{7.3.12}$$

The estimate (7.3.7) for $\varrho^+$ follows by (7.3.2), (7.3.3), (7.1.6) and the same arguments used in Lemmata 9.4.2-9.4.3. Similarly we deduce (7.3.8).

Now, recalling (3.2.28)-(3.2.29) (with $\mathbb{F} \rightsquigarrow \mathbb{M}$, $\mathbb{G} \rightsquigarrow \mathbb{M}^c$), and (7.1.3)-(7.1.4), we have that the operator $\mathtt{A}_0$ in (7.3.11) can be decomposed, with respect to the splitting $H_{\mathbb{M}} \oplus H_{\mathbb{M}^c}$, as

$$\mathtt{A}_0 = \frac{D_V}{1+\varepsilon^2\lambda} + \Pi_{\mathtt{D}}\varrho = \begin{pmatrix} D(\varepsilon,\lambda) & 0 \\ 0 & \mathtt{A}^{\mathbb{M}^c}_{\mathbb{M}^c} \end{pmatrix}\,,$$

where, taking in $H_{\mathbb{M}}$ the basis $\{(\Psi_j(x),0)\,,(0,\Psi_j(x))\}_{j\in\mathbb{M}}$, we have

$$\begin{aligned} D(\varepsilon,\lambda) &= \frac{[D_V]^{\mathbb{M}}_{\mathbb{M}}}{1+\varepsilon^2\lambda} + \mathrm{Diag}_{j\in\mathbb{M}}\big(\pi_+[\widehat{\varrho}^{\,j}_j(0)]\big) \\ &\overset{(7.1.4)}{=} \frac{1}{1+\varepsilon^2\lambda}\mathrm{Diag}_{j\in\mathbb{M}}\begin{pmatrix} \mu_j & 0 \\ 0 & \mu_j \end{pmatrix} + \varepsilon^2\mathrm{Diag}_{j\in\mathbb{M}}\Big(\frac{\pi_+[\widehat{\mathtt{B}}^{\,j}_j(0)]}{1+\varepsilon^2\lambda}\Big)\,. \end{aligned} \tag{7.3.13}$$

We now prove that $D(\varepsilon,\lambda)$ has the form (7.3.6).

**Lemma 7.3.2. (Shifted normal frequencies)** *The operator $D(\varepsilon,\lambda)$ in* (7.3.13) *has the form* (7.3.6) *with $\Omega_j(\varepsilon,\lambda)$ defined in* (7.1.7).

Proof. By (7.3.13) and recalling the definition of $\pi_+$ in (3.2.21), the operator $D(\varepsilon,\lambda)$ is

$$D(\varepsilon,\lambda) = \mathrm{Diag}_{j\in\mathbb{M}} \frac{\mu_j + \varepsilon^2 \mathtt{b}_j}{1+\varepsilon^2\lambda} \mathrm{Id}_2 \,, \quad \mathtt{b}_j := \frac{1}{2}\mathrm{Tr}(\widehat{\mathtt{B}}_j^j(0))\,, \tag{7.3.14}$$

and, by the definition of $\mathtt{B}$ in (6.1.25), we have

$$\mathrm{Tr}(\widehat{\mathtt{B}}_j^j(0)) = \frac{3}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} \big(\Psi_j, D_V^{-1/2}\big(a(x)(v(\varphi,0,\xi))^2 D_V^{-1/2}\Psi_j\big)\big)_{L^2(\mathbb{T}^d)}\, d\varphi\,, \tag{7.3.15}$$

$$+ \frac{4\varepsilon}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} \big(\Psi_j, D_V^{-1/2}\big(a_4(x)(v(\varphi,0,\xi))^3 D_V^{-1/2}\Psi_j\big)\big)_{L^2(\mathbb{T}^d)}\, d\varphi\,. \tag{7.3.16}$$

We first compute the term (7.3.15). Expanding the expression of $v$ in (2.2.9) we get

$$\begin{aligned}
(7.3.15) &= \frac{3}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|+d}} (D_V^{-1/2}\Psi_j(x)) a(x) (v(\varphi,0,\xi))^2 D_V^{-1/2}\Psi_j(x)\, d\varphi dx \\
&= \frac{3}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|+d}} \frac{\Psi_j(x)}{\sqrt{\mu_j}} a(x) \Big(\sum_{k\in\mathbb{S}} \mu_k^{-1/2}\sqrt{2\xi_k}\cos\varphi_k \Psi_k(x)\Big)^2 \frac{\Psi_j(x)}{\sqrt{\mu_j}}\, d\varphi dx \\
&= \frac{3}{(2\pi)^{|\mathbb{S}|}} \sum_{k_1,k_2\in\mathbb{S}} \int_{\mathbb{T}^{|\mathbb{S}|+d}} a(x)\ \mu_{k_1}^{-1/2}\mu_{k_2}^{-1/2}\sqrt{2\xi_{k_1}}\sqrt{2\xi_{k_2}}\cos\varphi_{k_1}\cos\varphi_{k_2} \\
&\qquad\qquad \times \Psi_{k_1}(x)\Psi_{k_2}(x)\frac{\Psi_j^2(x)}{\mu_j}\, d\varphi dx \\
&= \frac{3}{(2\pi)^{|\mathbb{S}|}} \sum_{k\in\mathbb{S}} \int_{\mathbb{T}^d} a(x)\ \mu_k^{-1} 2\xi_k \Psi_k^2(x)\Psi_j^2(x)\mu_j^{-1} dx \int_{\mathbb{T}^{|\mathbb{S}|}} \cos^2\varphi_k\, d\varphi \\
&= \frac{6}{(2\pi)^{|\mathbb{S}|}} \mu_j^{-1} \sum_{k\in\mathbb{S}} \mu_k^{-1}\xi_k \int_{\mathbb{T}^d} a(x)\ \Psi_k^2(x)\Psi_j^2(x) dx\, (2\pi)^{|\mathbb{S}|-1} \int_{\mathbb{T}} \cos^2\theta\, d\theta \\
&= 3\mu_j^{-1} \sum_{k\in\mathbb{S}} \mu_k^{-1} (\Psi_j^2, a(x)\Psi_k^2)_{L^2} \xi_k = 2(\mathscr{B}\xi)_j
\end{aligned} \tag{7.3.17}$$

using (1.2.9)-(1.2.10). On the other hand, the term in (7.3.16) is equal to zero, because

$$(7.3.16) = \frac{4\varepsilon}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^d}\int_{\mathbb{T}^{|\mathbb{S}|}} \frac{\Psi_j(x)}{\sqrt{\mu_j}} a_4(x)\big(v(\varphi,0,\xi)\big)^3 \frac{\Psi_j(x)}{\sqrt{\mu_j}}\, d\varphi dx \tag{7.3.18}$$

and, by (5.2.9), the integral

$$\int_{\mathbb{T}^{|\mathbb{S}|}} \big(v(\varphi,0,\xi)\big)^3 d\varphi = \int_{\mathbb{T}^{|\mathbb{S}|}} \big(v(\varphi+\vec{\pi},0,\xi)\big)^3 d\varphi = -\int_{\mathbb{T}^{|\mathbb{S}|}} \big(v(\varphi,0,\xi)\big)^3 d\varphi\,,$$

is equal to zero.

In conclusion, we deduce by (7.3.14), (7.3.15), (7.3.16), (7.3.17), (7.3.18) that $\mathtt{b}_j = (\mathscr{B}\xi)_j$, and, inserting the value $\xi := \xi(\lambda) = \varepsilon^{-2}\mathscr{A}^{-1}((1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon - \bar{\mu})$ defined in (1.2.27), we get that the eigenvalues of $D(\varepsilon, \lambda)$ in (7.3.14), are equal to

$$\frac{\mu_j + \varepsilon^2 \mathtt{b}_j}{1+\varepsilon^2\lambda} = \Omega_j(\varepsilon, \lambda)$$

with $\Omega_j(\varepsilon, \lambda)$ defined in (7.1.7). ∎

The new operator $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J\mathtt{A}^+$ obtained in Proposition 7.3.1 is in a suitable form to admit, for most values of the parameter $\lambda$, an approximate right inverse according to Proposition 8.2.1 in the next Chapter.

# Chapter 8

# Approximate right inverse in normal directions

The goal of this Chapter is to state the crucial Proposition 8.2.1 for the existence of an approximate right inverse for a class of quasi-periodic linear Hamiltonian operators, acting in the normal subspace $H_{\mathbb{S}}^{\perp}$, of the form $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho)$ where $A_0$ and $\rho$ are self-adjoint operators, $A_0$ is an admissible split operator according to Definition 8.1.1, and $\rho$ is "small", see (8.2.2). We shall use Proposition 8.2.1 to construct the sequence of approximate solutions along the iterative nonlinear Nash-Moser scheme of Chapter 11, more precisely to prove Proposition 11.1.1.

## 8.1 Split admissible operators

We first define the following class of admissible split operators $A_0$.

**Definition 8.1.1. (Admissible split operators)** *Let $C_1, c_1, c_2 > 0$ be constants. We denote by $\mathcal{C}(C_1, c_1, c_2)$ the class of self-adjoint operators*

$$A_0(\varepsilon, \lambda, \varphi) = \frac{D_V}{1+\varepsilon^2\lambda} + R_0(\varepsilon, \lambda, \varphi), \qquad D_V = (-\Delta + V(x))^{1/2}, \tag{8.1.1}$$

*acting on $H_{\mathbb{S}}^{\perp}$, defined for all $\lambda \in \widetilde{\Lambda} \subset \Lambda$, that satisfy*

1. $|R_0|_{\mathrm{Lip},+,s_1} \le C_1 \varepsilon^2$,
2. *$A_0$ is block diagonal with respect to the splitting $H_{\mathbb{S}}^{\perp} = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, i.e. $A_0$ has the form (see (3.2.1))*

$$A_0 := A_0(\varepsilon, \lambda, \varphi) = \begin{pmatrix} D_0(\varepsilon, \lambda) & 0 \\ 0 & V_0(\varepsilon, \lambda, \varphi) \end{pmatrix}, \tag{8.1.2}$$

*and, moreover, in the basis of the eigenfunctions* $\{(\Psi_j,0),\ (0,\Psi_j)\}_{j\in\mathbb{F}}$ *(see* (3.1.9)*), the operator* $D_0$ *is represented by the diagonal matrix*

$$D_0 := D_0(\varepsilon,\lambda) = \mathrm{Diag}_{j\in\mathbb{F}}\,\mu_j(\varepsilon,\lambda)\mathrm{Id}_2\,, \quad \mu_j(\varepsilon,\lambda)\in\mathbb{R}\,. \tag{8.1.3}$$

*The eigenvalues* $\mu_j(\varepsilon,\lambda)$ *satisfy*

$$|\mu_j(\varepsilon,\lambda)-\mu_j| \le C_1\varepsilon^2\,, \tag{8.1.4}$$

*where* $\mu_j$ *are defined in* (1.1.5)*, and*

$$\mathfrak{d}_\lambda(\mu_i-\mu_j)(\varepsilon,\lambda)\ge c_2\varepsilon^2 \quad \textit{or} \quad \mathfrak{d}_\lambda(\mu_i-\mu_j)(\varepsilon,\lambda)\le -c_2\varepsilon^2\,, \quad i\neq j\,, \tag{8.1.5}$$

$$\mathfrak{d}_\lambda(\mu_i+\mu_j)(\varepsilon,\lambda)\ge c_2\varepsilon^2 \quad \textit{or} \quad \mathfrak{d}_\lambda(\mu_i+\mu_j)(\varepsilon,\lambda)\le -c_2\varepsilon^2\,, \tag{8.1.6}$$

$$\forall j\in\mathbb{F}\,,\ \Big(c_2\varepsilon^2\le \mathfrak{d}_\lambda\mu_j(\varepsilon,\lambda)\le c_2^{-1}\varepsilon^2 \text{ or } -c_2^{-1}\varepsilon^2\le \mathfrak{d}_\lambda\mu_j(\varepsilon,\lambda)\le -c_2\varepsilon^2\Big)\,, \tag{8.1.7}$$

*and, in addition, for all* $j\in\mathbb{F}$,

$$\begin{cases}\mathfrak{d}_\lambda\big(V_0(\varepsilon,\lambda)+\mu_j(\varepsilon,\lambda)\mathrm{Id}\big)\le -c_1\varepsilon^2\\ \mathfrak{d}_\lambda\big(V_0(\varepsilon,\lambda)-\mu_j(\varepsilon,\lambda)\mathrm{Id}\big)\le -c_1\varepsilon^2\,.\end{cases} \tag{8.1.8}$$

We now verify that the operator $\mathtt{A}_0$ defined in (7.3.5) is an admissible split operator according to Definition 8.1.1.

**Lemma 8.1.2. ($\mathtt{A}_0$ is split admissible)** *There exist positive constants* $C_1,c_1,c_2$ *such that the operator* $\mathtt{A}_0$ *defined in* (7.3.5)-(7.3.6) *belongs to the class* $\mathcal{C}(C_1,c_1,c_2)$ *of split admissible operators introduced in Definition 8.1.1. Notice that* $\mathtt{A}_0$ *is defined for all* $\lambda\in\Lambda$.

Proof. Remind that the decomposition in (7.3.5) refers to the splitting $H_{\mathbb{S}}^{\perp}=H_{\mathbb{M}}\oplus H_{\mathbb{M}^c}$, where the finite set $\mathbb{M}\supset\mathbb{F}$ has been fixed in Lemma 7.1.1, whereas the decomposition in (8.1.2) concerns the splitting $H_{\mathbb{S}}^{\perp}=H_{\mathbb{F}}\oplus H_{\mathbb{G}}$.

By (7.3.5) the operator $\mathtt{A}_0$ has the form (8.1.1) with $R_0=\Pi_{\mathtt{D}}\varrho$ and (3.3.35), (7.1.6) imply

$$|R_0|_{\mathrm{Lip},+,s_1}=|\Pi_{\mathtt{D}}\varrho|_{\mathrm{Lip},+,s_1}\le C_1\varepsilon^2\,.$$

In addition, with respect to the splitting $H_{\mathbb{S}}^{\perp}=H_{\mathbb{F}}\oplus H_{\mathbb{G}}$, and taking in $H_{\mathbb{F}}$ the basis of the eigenfunctions $\{(\Psi_j,0),\ (0,\Psi_j)\}_{j\in\mathbb{F}}$, the operator $\mathtt{A}_0$ in (7.3.5)-(7.3.6) has the form (8.1.2)-(8.1.3) (recall that $\mathbb{F}\subset\mathbb{M}$) with

$$D_0(\varepsilon,\lambda)=\mathrm{Diag}_{j\in\mathbb{F}}\Omega_j(\varepsilon,\lambda)\mathrm{Id}_2\,, \quad \mu_j(\varepsilon,\lambda)=\Omega_j(\varepsilon,\lambda)\,,$$

and the operator $V_0$, which acts in $H_{\mathbb{G}}$, admits the decomposition, with respect to the splitting $H_{\mathbb{G}}=H_{\mathbb{M}\setminus\mathbb{F}}\oplus H_{\mathbb{M}^c}$, and taking in $H_{\mathbb{M}\setminus\mathbb{F}}$ the basis of the eigenfunctions $\{(\Psi_j,0),\ (0,\Psi_j)\}_{j\in\mathbb{M}\setminus\mathbb{F}}$,

$$V_0=\begin{pmatrix}\mathrm{Diag}_{j\in\mathbb{M}\setminus\mathbb{F}}\Omega_j(\varepsilon,\lambda)\mathrm{Id}_2 & 0\\ 0 & \mathtt{A}_{\mathbb{M}^c}^{\mathbb{M}^c}\end{pmatrix}\,. \tag{8.1.9}$$

By (7.1.7) and (1.2.25) we have, for all $j \in \mathbb{M}$, $\Omega_j(\varepsilon,\lambda) = \Omega_j(0,\lambda)+O(\varepsilon^2)$, where $\Omega_j(0,\lambda) = \mu_j$ are the unperturbed linear frequencies defined in (1.1.5), and the derivative of the functions $\Omega_j(\varepsilon,\lambda)$ (which are defined for all $\lambda \in \Lambda$) is

$$\partial_\lambda \Omega_j(\varepsilon,\lambda) = -\frac{\varepsilon^2}{(1+\varepsilon^2\lambda)^2}(\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar{\mu}]_j), \quad \forall j \in \mathbb{M}\,. \tag{8.1.10}$$

Then, by (8.1.10) and the definition of $\mathbb{F}$ and $\mathbb{G}$ in (1.2.15), for any $j \in \mathbb{M}\setminus\mathbb{F} \subset \mathbb{G}$, we have

$$\begin{aligned}\partial_\lambda \Omega_j + \max_{j\in\mathbb{F}} |\partial_\lambda \Omega_j| &= -\frac{\varepsilon^2}{(1+\varepsilon^2\lambda)^2}\Big((\mu_j - [\mathscr{B}\mathscr{A}^{-1}\bar{\mu}]_j) - \max_{j\in\mathbb{F}}|\mu_j - (\mathscr{B}\mathscr{A}^{-1}\bar{\mu})_j|\Big)\\ &< -\frac{\varepsilon^2}{(1+\varepsilon^2\lambda)^2}\mathfrak{c}\end{aligned} \tag{8.1.11}$$

for some $\mathfrak{c} > 0$. By (8.1.11) (recall that $\mu_j(\varepsilon,\lambda) = \Omega_j(\varepsilon,\lambda)$) and (7.1.9) we conclude that, for $\varepsilon$ small, property (8.1.8) holds for some $c_1 > 0$. The other properties (8.1.5)-(8.1.7) follow, for $\varepsilon$ small, by (8.1.10) and the assumptions (1.2.21)-(1.2.22). ■

## 8.2 Approximate right inverse

The main result of this Chapter is the following proposition which provides an approximate solution of the linear equation

$$\mathfrak{L}h = g \qquad \text{where} \qquad \mathfrak{L} := \bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho)\,.$$

**Proposition 8.2.1. (Approximate right inverse in normal directions)** *Let $\bar{\omega}_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ be $(\gamma_1,\tau_1)$-Diophantine and satisfy property $(\mathbf{NR})_{\gamma_1,\tau_1}$ in Definition 4.1.4 with $\gamma_1,\tau_1$ fixed in (1.2.28). Fix $s_1 > s_0$ according to Proposition 4.1.5 (more precisely Proposition 4.3.4), and $s_1 < s_2 < s_3$ such that*

$$(i)\ s_2 - s_1 > 300(\tau' + 3s_1 + 3)\,, \qquad (ii)\ s_3 - s_1 \geq 3(s_2 - s_1) \tag{8.2.1}$$

*where the constant $\tau'$ appears in the multiscale Proposition 4.1.5 and we assume that $\tau' > 2\tau + 3$.*

*Then there is $\varepsilon_0 > 0$ such that, $\forall \varepsilon \in (0,\varepsilon_0)$, for each self-adjoint operator*

$$A_0 = \frac{D_V}{1+\varepsilon^2\lambda} + R_0$$

*acting in $H_{\mathbb{S}}^{\perp}$, belonging to the class $\mathcal{C}(C_1,c_1,c_2)$ of admissible split operators (see Definition 8.1.1), for any self-adjoint operator $\rho \in L^2(\mathbb{T}^{|\mathbb{S}|},\mathcal{L}(H_{\mathbb{S}}^{\perp}))$, defined in $\widetilde{\Lambda} \subset \Lambda$, satisfying*

$$|\rho|_{\mathrm{Lip},+,s_1} \leq \varepsilon^3\,, \quad |R_0|_{\mathrm{Lip},+,s_2} + |\rho|_{\mathrm{Lip},+,s_2} \leq \varepsilon^{-1}\,, \tag{8.2.2}$$

*there are closed subsets $\mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho) \subset \widetilde{\Lambda}$, $1/2 \leq \eta \leq 5/6$, satisfying*

1. $\mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho)\subseteq\mathbf{\Lambda}(\varepsilon;\eta',A_0,\rho)$, *for all* $1/2\le\eta\le\eta'\le 5/6$,

2. $\big|[\mathbf{\Lambda}(\varepsilon;1/2,A_0,\rho)]^c\cap\widetilde{\Lambda}\big|\le b(\varepsilon)$ *where* $\lim_{\varepsilon\to 0} b(\varepsilon)=0$,

3. *if* $A_0'=(1+\varepsilon^2\lambda)^{-1}D_V+R_0'\in\mathcal{C}(C_1,c_1,c_2)$ *and* $\rho'$ *satisfy*
$$|R_0'-R_0|_{+,s_1}+|\rho'-\rho|_{+,s_1}\le\delta\le\varepsilon^3\,, \tag{8.2.3}$$
*for all* $\lambda\in\widetilde{\Lambda}\cap\widetilde{\Lambda}'\subset\Lambda$, *then, for all* $(1/2)+\delta^{2/5}\le\eta\le 5/6$,
$$\big|\widetilde{\Lambda}'\cap[\mathbf{\Lambda}(\varepsilon;\eta,A_0',\rho')]^c\cap\mathbf{\Lambda}(\varepsilon;\eta-\delta^{2/5},A_0,\rho)\big|\le\delta^{\alpha/3}; \tag{8.2.4}$$

*and, for any* $\nu\in(0,\varepsilon)$, *there exists a linear operator*
$$\mathfrak{L}^{-1}_{approx}:=\mathfrak{L}^{-1}_{approx,\nu}\in\mathcal{L}(\mathcal{H}^{s_3}\cap H^{\perp}_{\mathbb{S}})$$
*such that, for any function* $g:\widetilde{\Lambda}\to\mathcal{H}^{s_3}\cap H^{\perp}_{\mathbb{S}}$ *satisfying*
$$\|g\|_{\mathrm{Lip},s_1}\le\varepsilon^2\nu\,,\quad |R_0|_{\mathrm{Lip},+,s_3}+|\rho|_{\mathrm{Lip},+,s_3}+\|g\|_{\mathrm{Lip},s_3}\le\varepsilon^2\nu^{-1}\,, \tag{8.2.5}$$
*the function* $h:=\mathfrak{L}^{-1}_{approx}g$, $h:\mathbf{\Lambda}(\varepsilon;5/6,A_0,\rho)\to\mathcal{H}^{s_3}\cap H^{\perp}_{\mathbb{S}}$ *satisfies*
$$\|h\|_{\mathrm{Lip},s_1}\le\varepsilon^2\nu^{\frac{4}{5}}\,,\quad \|h\|_{\mathrm{Lip},s_3}\le\varepsilon^2\nu^{-\frac{11}{10}}\,, \tag{8.2.6}$$
*and*
$$(\bar{\omega}_\varepsilon\cdot\partial_\varphi-J(A_0+\rho))h=g+r \tag{8.2.7}$$
*with*
$$\|r\|_{\mathrm{Lip},s_1}\le\varepsilon^2\nu^{3/2}\,. \tag{8.2.8}$$
*Furthermore, setting* $Q':=2(\tau'+\varsigma s_1)+3$ *(where* $\varsigma=1/10$ *and* $\tau'$ *is given by Proposition 4.1.5), for all* $g\in\mathcal{H}^{s_0+Q'}\cap H^{\perp}_{\mathbb{S}}$,
$$\|\mathfrak{L}^{-1}_{approx}\,g\|_{\mathrm{Lip},s_0}\lesssim_{s_1}\|g\|_{\mathrm{Lip},s_0+Q'}\,. \tag{8.2.9}$$

The required bounds in (8.2.2)-(8.2.5) will be verified along the nonlinear Nash-Moser scheme of Chapter 11. Proposition 8.2.1 will be applied to the operator $\bar{\omega}_\varepsilon\cdot\partial_\varphi-J\mathtt{A}^+$ where $\mathtt{A}^+=\mathtt{A}^0+\varrho^+$ is defined in Proposition 7.3.1, to prove Proposition 11.1.1. Notice that $\mathtt{A}^0$ is split admissible by Lemma 8.1.2 and the coupling operator $\varrho^+$ satisfies $|\varrho^+|_{\mathrm{Lip},+,s_1}\lesssim_{s_1}\varepsilon^4\le\varepsilon^3$ by (7.3.7) and (6.1.26).

Proposition 8.2.1 is proved in Chapter 10 using the results of Chapters 4 and 9.

We remark that the value of $\varepsilon_0$ given in Proposition 8.2.1 may depend on $s_3$, because, in the estimates of the proof, there are quantities of the form $C(s_3)\varepsilon^a$, $a>0$, and we choose $\varepsilon$ small so that $C(s_3)\varepsilon^a<1$. However, such terms appear by quantities $C(s_3)\nu^{a_1}$, $a_1>0$, and so we may require $C(s_3)\nu^{a_1}<1$ assuming only $\nu$ small enough, i.e. $\nu\le\widetilde{\nu}(s_3)$ (see Remark 10.4.1). As a result, the following more specific statement holds.

**Proposition 8.2.2.** *The conclusion of Proposition 8.2.1 can be modified from* (8.2.4) *in the following way.*

*For any* $s \geq s_3$*, there is* $\widetilde{\nu}(s) > 0$ *such that: for any* $0 < \nu < \min(\varepsilon, \widetilde{\nu}(s))$*, there exists a linear operator*

$$\mathfrak{L}^{-1}_{approx} := \mathfrak{L}^{-1}_{approx,\nu,s} \in \mathcal{L}(\mathcal{H}^s \cap H^{\perp}_{\mathbb{S}})$$

*such that, for any function* $g : \widetilde{\Lambda} \to \mathcal{H}^s \cap H^{\perp}_{\mathbb{S}}$ *satisfying*

$$\|g\|_{\mathrm{Lip},s_1} \leq \varepsilon^2 \nu\,, \quad |R_0|_{\mathrm{Lip},+,s} + |\rho|_{\mathrm{Lip},+,s} + \|g\|_{\mathrm{Lip},s} \leq \varepsilon^2 \nu^{-1}\,, \tag{8.2.10}$$

*the function* $h := \mathfrak{L}^{-1}_{approx} g$*,* $h : \mathbf{\Lambda}(\varepsilon; 5/6, A_0, \rho) \to \mathcal{H}^s \cap H^{\perp}_{\mathbb{S}}$ *satisfies*

$$\|h\|_{\mathrm{Lip},s_1} \leq \varepsilon^2 \nu^{\frac{4}{5}}\,, \quad \|h\|_{\mathrm{Lip},s} \leq \varepsilon^2 \nu^{-\frac{11}{10}}\,, \tag{8.2.11}$$

*and*

$$(\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho))h = g + r \tag{8.2.12}$$

*with*

$$\|r\|_{\mathrm{Lip},s_1} \leq \varepsilon^2 \nu^{3/2}\,. \tag{8.2.13}$$

*Furthermore, setting* $Q' := 2(\tau' + \varsigma s_1) + 3$ *(where* $\varsigma = 1/10$ *and* $\tau'$ *is given by Proposition 4.1.5), for all* $g \in \mathcal{H}^{s_0+Q'} \cap H^{\perp}_{\mathbb{S}}$*,*

$$\|\mathfrak{L}^{-1}_{approx}\, g\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \|g\|_{\mathrm{Lip},s_0+Q'}\,. \tag{8.2.14}$$

# Chapter 9

# Splitting between low-high normal subspaces

The main result of this Chapter is Corollary 9.1.2 below. Its goal is to block-diagonalize a quasi-periodic Hamiltonian operator of the form $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho)$ according to the splitting $H_{\mathbb{S}}^\perp = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, up to a very small coupling term, see the conjugation (9.1.34) where $\rho_{\mathfrak{n}}$ is small according to (9.1.32).

## 9.1 Splitting step and corollary

The proof of Corollary 9.1.2 is based on an iterative application of the following Proposition.

**Proposition 9.1.1. (Splitting step)** *Let $\bar{\omega}_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ be $(\gamma_1, \tau_1)$-Diophantine and satisfy property $(\mathbf{NR})_{\gamma_1,\tau_1}$ in Definition 4.1.4 with $\gamma_1, \tau_1$ fixed in* (1.2.28)*. Assume $s_2 - s_1 > 120(\tau' + 3s_1 + 3)$, i.e. condition* (8.2.1)*-(i). Then, given $C_1 > 0$, $c_1 > 0$, $c_2 > 0$, there is $\varepsilon_0 > 0$ such that, $\forall \varepsilon \in (0, \varepsilon_0)$, for each self-adjoint operator*

$$A_0 = \frac{D_V}{1+\varepsilon^2\lambda} + R_0 = \begin{pmatrix} D_0(\varepsilon,\lambda) & 0 \\ 0 & V_0(\varepsilon,\lambda,\varphi) \end{pmatrix}$$

*belonging to the class $\mathcal{C}(C_1, c_1, c_2)$ of admissible split operators of Definition 8.1.1 (see* (8.1.1)*-*(8.1.2)*), defined in $\widetilde{\Lambda}$, with*

$$D_0(\varepsilon,\lambda) = \mathrm{Diag}_{j\in\mathbb{F}}\, \mu_j(\varepsilon,\lambda)\mathrm{Id}_2\,, \quad \mu_j(\varepsilon,\lambda) \in \mathbb{R}\,,$$

*as in* (8.1.3)*, there are*

- *closed subsets $\Lambda(\varepsilon;\eta, A_0) \subset \widetilde{\Lambda}$, $1/2 \le \eta \le 1$, satisfying the properties*

  1. *$\Lambda(\varepsilon;\eta, A_0) \subseteq \Lambda(\varepsilon;\eta', A_0)$ for all $1/2 \le \eta \le \eta' \le 1$,*

2. $\big|[\Lambda(\varepsilon;1/2,A_0)]^c\cap\widetilde{\Lambda}\big|\leq b(\varepsilon)$ *where* $\lim_{\varepsilon\to 0} b(\varepsilon)=0$,

3. *if* $A_0'=(1+\varepsilon^2\lambda)^{-1}D_V+R_0'\in\mathcal{C}(C_1,c_1,c_2)$ *with* $|R_0'-R_0|_{+,s_1}\leq\delta\leq\varepsilon^{5/2}$ *for all* $\lambda\in\widetilde{\Lambda}\cap\widetilde{\Lambda}'$, *then, for all* $(1/2)+\sqrt{\delta}\leq\eta\leq 1$,

$$\big|\widetilde{\Lambda}'\cap[\Lambda(\varepsilon;\eta,A_0')]^c\cap\Lambda(\varepsilon;\eta-\sqrt{\delta},A_0)\big|\leq\delta^{\alpha}\,,\quad \alpha>0\,, \tag{9.1.1}$$

*and, for each self-adjoint operator* $\rho$ *acting in* $H_{\mathbb{S}}^{\perp}$, *defined for* $\lambda\in\widetilde{\Lambda}$, *satisfying*

$$|\rho|_{\mathrm{Lip},+,s_1}\leq\delta_1\leq\varepsilon^3\,,\qquad \delta_1(|R_0|_{\mathrm{Lip},+,s_2}+|\rho|_{\mathrm{Lip},+,s_2})\leq\varepsilon\,, \tag{9.1.2}$$

*there exist*

- *a symplectic linear invertible transformation*

$$e^{J\mathcal{S}(\varphi)}\in\mathcal{L}(H_{\mathbb{S}}^{\perp})\,,\quad \varphi\in\mathbb{T}^{|\mathbb{S}|}\,, \tag{9.1.3}$$

  *defined for all* $\lambda\in\Lambda(\varepsilon;1,A_0)$, *where* $\mathcal{S}(\varphi):=\mathcal{S}(\varepsilon,\lambda)(\varphi)$ *is a self-adjoint operator in* $\mathcal{L}(H_{\mathbb{S}}^{\perp})$, *satisfying the estimates* (9.1.7)-(9.1.8) *below;*

- *a self-adjoint operator* $A^+$ *of the form*

$$A^+=\frac{D_V}{1+\varepsilon^2\lambda}+R_0^++\rho^+=\begin{pmatrix}D_0^+(\varepsilon,\lambda) & 0\\ 0 & V_0^+(\varepsilon,\lambda,\varphi)\end{pmatrix}+\rho^+\,, \tag{9.1.4}$$

  *defined for all* $\lambda\in\Lambda(\varepsilon;1,A_0)$, *with*

$$R_0^+,\rho^+\ L^2\text{-}self-adjoint\,,\qquad D_0^+(\varepsilon,\lambda)=\mathrm{Diag}_{j\in\mathbb{F}}\,\mu_j^+(\varepsilon,\lambda)\mathrm{Id}_2\,, \tag{9.1.5}$$

  *and* $\mu_j^+(\varepsilon,\lambda)\in\mathbb{R}$;

*such that, for all* $\lambda\in\Lambda(\varepsilon;1,A_0)$,

$$\big(\bar{\omega}_\varepsilon\cdot\partial_\varphi-JA_0-J\rho\big)e^{J\mathcal{S}(\varphi)}=e^{J\mathcal{S}(\varphi)}\big(\bar{\omega}_\varepsilon\cdot\partial_\varphi-JA^+\big)\,, \tag{9.1.6}$$

*and the following estimates hold:*

- *The self-adjoint operator* $\mathcal{S}$ *satisfies*

$$|\mathcal{S}|_{\mathrm{Lip},s_1+1}\leq\delta_1^{\frac{7}{8}}\,,\ |\mathcal{S}|_{\mathrm{Lip},s_2+1}\leq\delta_1^{-\frac{1}{4}}\big(|R_0|_{\mathrm{Lip},+,s_2}+|\rho|_{\mathrm{Lip},+,s_2}\big)+\delta_1^{-\frac{3}{4}}\,, \tag{9.1.7}$$

  *and more generally, for all* $s\geq s_2$,

$$|\mathcal{S}|_{\mathrm{Lip},s+1}\leq C(s)\Big[\delta_1^{-\frac{1}{4}}\big(|R_0|_{\mathrm{Lip},+,s}+|\rho|_{\mathrm{Lip},+,s}\big)+\delta_1^{-\frac{3}{4}}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\Big]\,; \tag{9.1.8}$$

- *The operators in* (9.1.4)-(9.1.5) *satisfy*

$$|\mu_j^+(\varepsilon,\lambda) - \mu_j(\varepsilon,\lambda)|_{\rm Lip} \leq \delta_1^{3/4} = o(\varepsilon^2)\,, \tag{9.1.9}$$
$$\|V_0^+ - V_0\|_{{\rm Lip},0} \leq \delta_1^{3/4} = o(\varepsilon^2)\,, \tag{9.1.10}$$

*and*

$$|R_0^+ - R_0|_{{\rm Lip},+,s_1} \leq \delta_1^{3/4}\,, \quad |\rho^+|_{{\rm Lip},+,s_1} \leq \delta_1^{3/2}/2\,, \tag{9.1.11}$$
$$|R_0^+|_{{\rm Lip},+,s_2} + |\rho^+|_{{\rm Lip},+,s_2} \leq \delta_1^{-1/4}\big(|R_0|_{{\rm Lip},+,s_2} + |\rho|_{{\rm Lip},+,s_2}\big) + \delta_1^{-3/4} \tag{9.1.12}$$
$$\stackrel{(9.1.2)}{\leq} \varepsilon\delta_1^{-3/2}\,, \tag{9.1.13}$$

*and, more generally, for all* $s \geq s_2$*,*

$$|R_0^+|_{{\rm Lip},+,s} + |\rho^+|_{{\rm Lip},+,s} \lesssim_s \delta_1^{-\frac{1}{4}}\big(|R_0|_{{\rm Lip},+,s} + |\rho|_{{\rm Lip},+,s}\big) + \delta_1^{-\frac{3}{4}}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}} \tag{9.1.14}$$

*where* $\varsigma = 1/10$ *is fixed in* (4.1.16)*.*

*At last, given* $A_0' = (1+\varepsilon^2\lambda)^{-1}D_V + R_0' \in \mathcal{C}(C_1,c_1,c_2)$ *and* $\rho'$ *satisfying* (9.1.2)*, then, for all* $\lambda \in \Lambda(\varepsilon;1,A_0) \cap \Lambda(\varepsilon;1,A_0')$*, we have the estimates*

$$|{R_0'}^+ - R_0^+|_{+,s_1} \leq |R_0' - R_0|_{+,s_1} + C|\rho' - \rho|_{+,s_1} \tag{9.1.15}$$
$$|{\rho'}^+ - \rho^+|_{+,s_1} \leq \delta_1^{1/2}|R_0' - R_0|_{+,s_1} + \delta_1^{-1/20}|\rho' - \rho|_{+,s_1}\,. \tag{9.1.16}$$

Note that, according to (9.1.6), (9.1.4), (9.1.11), the new coupling term $\rho^+$ is much smaller than $\rho$, in low norm $|\ |_{{\rm Lip},+,s_1}$. Moreover, by (9.1.13), the new $\rho^+$ satisfies also the second assumption (9.1.2) with $\delta_1^{3/2}$ instead of $\delta_1$. Notice also that the Cantor sets $\Lambda(\varepsilon;\eta,A_0)$ depend only on $A_0$ and not on $\rho$. Finally we point out that (9.1.15)-(9.1.16) will be used for the measure estimate of Cantor sets of "good" parameters $\lambda$, in relation with property 3 (see (9.1.1)): for this application, an estimates of low norms $|\ |_{+,s_1}$, without the control of the Lipschitz dependence, is enough.

Applying iteratively Proposition 9.1.1 we deduce the following corollary.

**Corollary 9.1.2. (Splitting)** *Let* $\bar\omega_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ *be* $(\gamma_1,\tau_1)$*-Diophantine and satisfy property* $(\mathbf{NR})_{\gamma_1,\tau_1}$ *in Definition 4.1.4 with* $\gamma_1,\tau_1$ *fixed in* (1.2.28)*. Let*

$$A_0 = \frac{D_V}{1+\varepsilon^2\lambda} + R_0$$

*be a self-adjoint operator in the class of admissible split operators* $\mathcal{C}(C_1,c_1,c_2)$ *(see Definition 8.1.1) and* $\rho$ *be a self-adjoint operator in* $\mathcal{L}(H_{\mathbb{S}}^{\perp})$*, defined for* $\lambda \in \widetilde\Lambda$*, satisfying* (8.2.2)*. Then there exist*

- *closed sets* $\Lambda_\infty(\varepsilon;\eta,A_0,\rho)\subset\widetilde{\Lambda}$, $1/2\le\eta\le 5/6$, *satisfying the properties*
  1. $\Lambda_\infty(\varepsilon;\eta,A_0,\rho)\subseteq\Lambda_\infty(\varepsilon;\eta',A_0,\rho)$ *for all* $1/2\le\eta\le\eta'\le 5/6$,
  2. $\big|\big[\Lambda_\infty(\varepsilon;1/2,A_0,\rho)\big]^c\cap\widetilde{\Lambda}\big|\le b_1(\varepsilon)$ *where* $\lim_{\varepsilon\to 0}b_1(\varepsilon)=0$,
  3. *if* $A_0'=(1+\varepsilon^2\lambda)^{-1}D_V+R_0'\in\mathcal{C}(C_1,c_1,c_2)$ *and* $\rho'$ *satisfy*

$$|R_0-R_0'|_{+,s_1}+|\rho-\rho'|_{+,s_1}\le\delta\le\varepsilon^2 \tag{9.1.17}$$

  *for all* $\lambda\in\widetilde{\Lambda}\cap\widetilde{\Lambda}'$, *then, for all* $(1/2)+\delta^{2/5}\le\eta\le 5/6$,

$$\big|\widetilde{\Lambda}'\cap[\Lambda_\infty(\varepsilon;\eta,A_0',\rho')]^c\cap\Lambda_\infty(\varepsilon;\eta-\delta^{2/5},A_0,\rho)\big|\le\delta^{\alpha/2}; \tag{9.1.18}$$

- *a sequence of symplectic linear invertible transformations*

$$\mathcal{P}_0:=\mathrm{Id}\,,\quad \mathcal{P}_{\mathfrak{n}}:=\mathcal{P}_{\mathfrak{n}}(\varepsilon,\lambda)(\varphi)=e^{J\mathcal{S}_1(\varepsilon,\lambda)(\varphi)}\dots e^{J\mathcal{S}_{\mathfrak{n}}(\varepsilon,\lambda)(\varphi)}\,,\quad \mathfrak{n}\ge 1\,, \tag{9.1.19}$$

*defined for all* $\lambda\in\Lambda_\infty(\varepsilon;5/6,A_0,\rho)$, *acting in* $H_{\mathbb{S}}^\perp$, *satisfying, for*

$$\delta_1=\varepsilon^3\,, \tag{9.1.20}$$

*the estimates*

$$\text{for } \mathfrak{n}\ge 1,\ |\mathcal{P}_{\mathfrak{n}}^{\pm1}-\mathcal{P}_{\mathfrak{n}-1}^{\pm1}|_{\mathrm{Lip},+,s_1}\le\delta_1^{(\frac{3}{2})^{\mathfrak{n}-1}\frac{3}{4}},\ |\mathcal{P}_{\mathfrak{n}}^{\pm1}-\mathrm{Id}|_{\mathrm{Lip},+,s_1}\le 2\delta_1^{\frac{3}{4}}\,, \tag{9.1.21}$$

$$|\mathcal{P}_{\mathfrak{n}}^{\pm1}|_{\mathrm{Lip},+,s_2}\le(C(s_2))^{\mathfrak{n}}\delta_1^{-(3/2)^{\mathfrak{n}-1}\frac{3}{4}}\big[\varepsilon\delta_1^{-\frac{1}{2}}+1\big]\,, \tag{9.1.22}$$

*and, more generally, for all* $s\ge s_2$,

$$|\mathcal{P}_{\mathfrak{n}}^{\pm1}|_{\mathrm{Lip},+,s}\le(C(s))^{\mathfrak{n}}\delta_1^{-(\frac{3}{2})^{\mathfrak{n}-1}(\frac{3}{4}+\alpha(s))}\big[(|R_0|_{\mathrm{Lip},+,s}+|\rho|_{\mathrm{Lip},+,s})\delta_1^{\frac{1}{2}+\frac{2\alpha(s)}{3}}+1\big] \tag{9.1.23}$$

*where*

$$\alpha(s):=3\varsigma\frac{s-s_2}{s_2-s_1}\,; \tag{9.1.24}$$

- *a sequence of self-adjoint block diagonal operators of the form*

$$A_{\mathfrak{n}}=\frac{D_V}{1+\varepsilon^2\lambda}+R_{\mathfrak{n}}:=\begin{pmatrix}D_{\mathfrak{n}}(\varepsilon,\lambda) & 0\\ 0 & V_{\mathfrak{n}}(\varepsilon,\lambda,\varphi)\end{pmatrix}\,,\quad \mathfrak{n}\ge 1\,, \tag{9.1.25}$$

*defined for* $\lambda\in\Lambda_\infty(\varepsilon;5/6,A_0,\rho)$, *belonging to the class* $\mathcal{C}(2C_1,c_1/2,c_2/2)$ *of admissible split operators, with*

$$D_{\mathfrak{n}}(\varepsilon,\lambda)=\mathrm{Diag}_{j\in\mathbb{F}}\,\mu_j^{(\mathfrak{n})}(\varepsilon,\lambda)\mathrm{Id}_2\,, \tag{9.1.26}$$

*satisfying*

$$|\mu_j^{(\mathfrak{n})}(\varepsilon,\lambda) - \mu_j(\varepsilon,\lambda)|_{\mathrm{Lip}} = O(\delta_1^{3/4}) = o(\varepsilon^2)\,, \tag{9.1.27}$$

$$\|V_{\mathfrak{n}} - V_0\|_{\mathrm{Lip},0} = O(\delta_1^{3/4}) = o(\varepsilon^2)\,, \tag{9.1.28}$$

$$|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} \le C_1\varepsilon^2 + 2\delta_1^{3/4} \le 2C_1\varepsilon^2\,, \tag{9.1.29}$$

$$|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} \ll (C(s_2))^{\mathfrak{n}}\delta_1^{-(\frac{3}{2})^{\mathfrak{n}-1}\frac{3}{4}}\big[\varepsilon\delta_1^{-\frac{1}{2}} + 1\big]\,, \tag{9.1.30}$$

*and, more generally, for all* $s \ge s_2$,

$$|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s} \le (C(s))^{\mathfrak{n}}\delta_1^{-(\frac{3}{2})^{\mathfrak{n}-1}(\frac{3}{4}+\alpha(s))}\big[\big(|R_0|_{\mathrm{Lip},+,s} + |\rho|_{\mathrm{Lip},+,s}\big)\delta_1^{\frac{1}{2}+\frac{2\alpha(s)}{3}} + 1\big] \tag{9.1.31}$$

*where* $\alpha(s)$ *is defined in* (9.1.24)*;*

- *a sequence of* $L^2$ *self-adjoint operators* $\rho_{\mathfrak{n}} \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$, $\mathfrak{n} \ge 1$, *defined for* $\lambda \in \Lambda_\infty(\varepsilon; 5/6, A_0, \rho)$, *satisfying*

$$|\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} \le \delta_1^{(\frac{3}{2})^{\mathfrak{n}}}\,, \quad |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} \ll (C(s_2))^{\mathfrak{n}}\delta_1^{-(\frac{3}{2})^{\mathfrak{n}-1}\frac{3}{4}}[\varepsilon\delta_1^{-\frac{1}{2}} + 1]\,, \tag{9.1.32}$$

*and, more generally, for all* $s \ge s_2$,

$$|\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s} \le (C(s))^{\mathfrak{n}}\delta_1^{-(\frac{3}{2})^{\mathfrak{n}-1}(\frac{3}{4}+\alpha(s))}\big[(|R_0|_{\mathrm{Lip},+,s} + |\rho|_{\mathrm{Lip},+,s})\delta_1^{\frac{1}{2}+\frac{2\alpha(s)}{3}} + 1\big] \tag{9.1.33}$$

*where* $\alpha(s)$ *is defined in* (9.1.24)*;*

*such that, for all* $\lambda \in \Lambda_\infty(\varepsilon; 5/6, A_0, \rho)$,

$$\big(\bar\omega_\varepsilon \cdot \partial_\varphi - JA_0 - J\rho\big)\mathcal{P}_{\mathfrak{n}}(\varphi) = \mathcal{P}_{\mathfrak{n}}(\varphi)\big(\bar\omega_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}} - J\rho_{\mathfrak{n}}\big)\,. \tag{9.1.34}$$

*At last, given*

$$A_0' = \frac{D_V}{1+\varepsilon^2\lambda} + R_0' \in \mathcal{C}(C_1, c_1, c_2)$$

*and a self-adjoint operator* $\rho' \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$ *satisfying* (8.2.2)*, if*

$$|A_0' - A_0|_{+,s_1} + |\rho' - \rho|_{+,s_1} \le \delta \le \varepsilon^3\,, \quad \forall \lambda \in \widetilde\Lambda \cap \widetilde\Lambda'\,, \tag{9.1.35}$$

*then, for all* $\lambda \in \Lambda_\infty(\varepsilon; 5/6, A_0, \rho) \cap \Lambda_\infty(\varepsilon; 5/6, A_0', \rho')$, *for all* $\mathfrak{n} \in \mathbb{N}$,

$$|A_{\mathfrak{n}}' - A_{\mathfrak{n}}|_{+,s_1} \le \delta^{4/5}\,. \tag{9.1.36}$$

Note that each operator $A_{\mathfrak{n}}$ in (9.1.25) is block-diagonal according to the splitting $H_{\mathbb{S}}^{\perp} = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$ in (3.1.5), i.e. it has the same form as $A_0$ in (8.1.2) but the coupling term $\rho_{\mathfrak{n}}$ in (9.1.34) is much smaller than $\rho$ (in the low norm $| \ |_{\mathrm{Lip},+,s_1}$), compare the first inequality in (9.1.32) (where $\delta_1 = \varepsilon^3$ by (9.1.20)) and the first inequality in (8.2.2). The tame estimates (9.1.23) and (9.1.33) for all $s \geq s_2$, will be used in the proof of Proposition 8.2.1, which provides an approximate right inverse required in the Nash-Moser nonlinear iteration in Chapter 11.

PROOF. Let us define the sequences of real numbers $(\delta_{\mathfrak{n}})_{\mathfrak{n}\geq 1}$ and $(\eta_{\mathfrak{n}})_{\mathfrak{n}\geq 0}$ by

$$\delta_{\mathfrak{n}} := \delta_1^{(\frac{3}{2})^{\mathfrak{n}-1}}, \ \delta_1 = \varepsilon^3, \quad \text{and} \quad \eta_0 := 0, \ \eta_{\mathfrak{n}+1} := \eta_{\mathfrak{n}} + \delta_{\mathfrak{n}+1}^{\frac{3}{8}} = \eta_{\mathfrak{n}} + \delta_1^{\frac{3}{8}(\frac{3}{2})^{\mathfrak{n}}}. \tag{9.1.37}$$

We shall prove by induction the following statements: for any $\mathfrak{n} \in \mathbb{N}$

- $(\mathbf{P})_{\mathfrak{n}}$ there exist:
  - (i) symplectic linear invertible transformations $\mathcal{P}_0, \ldots, \mathcal{P}_{\mathfrak{n}}$ of the form (9.1.19), defined respectively for $\lambda$ in decreasing subsets

    $$\Lambda_{\mathfrak{n}}(\varepsilon; 5/6, A_0, \rho) \subset \ldots \subset \Lambda_1(\varepsilon; 5/6, A_0, \rho) \subset \Lambda_0 := \widetilde{\Lambda},$$

    satisfying (9.1.21)-(9.1.23) at any order $k = 0, \ldots, \mathfrak{n}$. The sets $\Lambda_{\mathfrak{n}}(\varepsilon; \eta, A_0, \rho)$, $1/2 \leq \eta \leq 5/6$, are defined inductively by

    $$\Lambda_0 := \widetilde{\Lambda} \quad \text{and} \quad \Lambda_{\mathfrak{n}}(\varepsilon; \eta, A_0, \rho) := \bigcap_{k=0}^{\mathfrak{n}-1} \Lambda(\varepsilon; \eta + \eta_k, A_k), \ \mathfrak{n} \geq 1, \tag{9.1.38}$$

    where the sets $\Lambda(\varepsilon; \eta + \eta_k, A_k)$ are those defined by Proposition 9.1.1. Notice that, for $\varepsilon$ small enough, $\eta_k \leq 1/6$ for any $k \geq 0$ (and so $\eta + \eta_k \leq 1$).
  - (ii) Self-adjoint admissible split operators $A_0, \ldots, A_{\mathfrak{n}}$ as in (9.1.25)-(9.1.26), in the class $\mathcal{C}(2C_1, c_1/2, c_2/2)$ (see Definition 8.1.1), satisfying (9.1.27)-(9.1.31) at any order $k = 0, \ldots, \mathfrak{n}$, and such that the conjugation identity (9.1.34) holds for all $\lambda$ in $\Lambda_{\mathfrak{n}}(\varepsilon; 5/6, A_0, \rho)$, with $\rho_k$ satisfying (9.1.32)-(9.1.33) for $k = 0, \ldots, \mathfrak{n}$.
  - (iii) Moreover we have, $\forall \mathfrak{n} \geq 1$,

    $$|A_{\mathfrak{n}} - A_{\mathfrak{n}-1}|_{\mathrm{Lip},+,s_1} \leq \delta_{\mathfrak{n}}^{\frac{3}{4}} = \delta_1^{\frac{3}{4}(\frac{3}{2})^{\mathfrak{n}-1}} \text{ on } \Lambda_{\mathfrak{n}}(\varepsilon; 5/6, A_0, \rho). \tag{9.1.39}$$

**Initialization.** The statement $(\mathbf{P})_0$-(i) holds with $\mathcal{P}_0 :=$ Id, and (9.1.21)-(9.1.23) trivially hold. The conjugation identity (9.1.34) at $\mathfrak{n} = 0$ trivially holds with $\rho_0 := \rho$. Then, in order to prove $(\mathbf{P})_0$-(ii), it is sufficient to notice that the self-adjoint operator $A_0 \in \mathcal{C}(C_1, c_1, c_2)$ has the form (9.1.25), (9.1.26) with $\mu_j^{(0)}(\varepsilon, \lambda) = \mu_j(\varepsilon, \lambda)$, and (9.1.27)-(9.1.29) hold. The estimate (9.1.30) (which for $\mathfrak{n} = 0$ is $|R_0|_{\mathrm{Lip},+,s_2} \ll \delta_1^{-1/2}[\varepsilon\delta_1^{-1/2} + 1]$) and (9.1.32) are consequences of Assumption (8.2.2), since $\delta_1 = \varepsilon^3$. Finally notice that (9.1.31) and (9.1.33) are trivially satisfied.

**Induction.** Next assume that $(\mathbf{P})_{\mathfrak{n}}$ holds. In order to define $\mathcal{P}_{\mathfrak{n}+1}$ and $A_{\mathfrak{n}+1}$ we apply the "splitting step" Proposition 9.1.1 with $A_0, \rho$ replaced by $A_{\mathfrak{n}}, \rho_{\mathfrak{n}}$. In fact, by the inductive assumption $(\mathbf{P})_{\mathfrak{n}}$, the operator $A_{\mathfrak{n}}$ in (9.1.25) belongs to the class $\mathcal{C}(2C_1, c_1/2, c_2/2)$ of admissible split operators, according to Definition 8.1.1. Moreover, by (9.1.32), (9.1.30), we have

$$|\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} \leq \delta_1^{(\frac{3}{2})^{\mathfrak{n}}} \overset{(9.1.37)}{=} \delta_{\mathfrak{n}+1}\,, \qquad \delta_{\mathfrak{n}+1}\big(|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2}\big) \leq \varepsilon\,, \tag{9.1.40}$$

which is (9.1.2) with $\rho_{\mathfrak{n}}$, $R_{\mathfrak{n}}$, $\delta_{\mathfrak{n}+1}$ instead of $\rho$, $R_0$, $\delta_1$.

We define, for $1/2 \leq \eta \leq 5/6$, the set

$$\Lambda_{\mathfrak{n}+1}(\varepsilon;\eta, A_0, \rho) := \Lambda_{\mathfrak{n}}(\varepsilon;\eta, A_0,\rho) \cap \Lambda\big(\varepsilon;\eta + \eta_{\mathfrak{n}}, A_{\mathfrak{n}}\big)\,. \tag{9.1.41}$$

in agreement with (9.1.38) at $\mathfrak{n}+1$.

By (9.1.40), Proposition 9.1.1 implies the existence of a self-adjoint operator $\mathcal{S}_{\mathfrak{n}+1} \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$, defined for $\lambda \in \Lambda_{\mathfrak{n}+1}(\varepsilon; 5/6, A_0, \rho) \subset \Lambda\big(\varepsilon; 1, A_{\mathfrak{n}}\big)$, satisfying (see (9.1.7)-(9.1.8))

$$\begin{aligned}
&|\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s_1+1} \leq \delta_{\mathfrak{n}+1}^{\frac{7}{8}} = \delta_1^{\frac{7}{8}(\frac{3}{2})^{\mathfrak{n}}}\,,\\
&|\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s_2+1} \leq \delta_{\mathfrak{n}+1}^{-\frac{1}{4}}\big(|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2}\big) + \delta_{\mathfrak{n}+1}^{-\frac{3}{4}}\,,\\
&|\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s+1} \leq C(s)\Big[\delta_{\mathfrak{n}+1}^{-\frac{1}{4}}\big(|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s}\big) + \delta_{\mathfrak{n}+1}^{-\frac{3}{4}}\delta_{\mathfrak{n}+1}^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\Big]\,,
\end{aligned} \tag{9.1.42}$$

such that, for any $\lambda \in \Lambda_{\mathfrak{n}+1}(\varepsilon; 5/6, A_0, \rho)$, we have

$$\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}} - J\rho_{\mathfrak{n}}\big)e^{J\mathcal{S}_{\mathfrak{n}+1}} = e^{J\mathcal{S}_{\mathfrak{n}+1}}\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}+1} - J\rho_{\mathfrak{n}+1}\big) \tag{9.1.43}$$

where the operator

$$A_{\mathfrak{n}+1} = \frac{D_V}{1+\varepsilon^2\lambda} + R_{\mathfrak{n}+1} \tag{9.1.44}$$

is block-diagonal as in (9.1.25)-(9.1.26). By (9.1.11)-(9.1.13), (9.1.14), we have

$$\begin{aligned}
|R_{\mathfrak{n}+1} - R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} &\leq \delta_{\mathfrak{n}+1}^{3/4} = \delta_1^{\frac{3}{4}(\frac{3}{2})^{\mathfrak{n}}}\,,\\
|\rho_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_1} &\leq \frac{1}{2}\,\delta_1^{(\frac{3}{2})^{\mathfrak{n}+1}}\,,\\
|R_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_2} + |\rho_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_2} &\leq \delta_{\mathfrak{n}+1}^{-\frac{1}{4}}\big(|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2}\big) + \delta_{\mathfrak{n}+1}^{-\frac{3}{4}}\,,\\
|R_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s} + |\rho_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s} &\lesssim_s \delta_{\mathfrak{n}+1}^{-\frac{1}{4}}\big(|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s}\big) + \delta_{\mathfrak{n}+1}^{-\frac{3}{4}}\delta_{\mathfrak{n}+1}^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\,.
\end{aligned} \tag{9.1.45}$$

In particular, we have the first bound in (9.1.32) at the step $\mathfrak{n}+1$.

THE SYMPLECTIC TRANSFORMATION $\mathcal{P}_{\mathfrak{n}+1}$. We define, for $\lambda \in \Lambda_{\mathfrak{n}+1}(\varepsilon; 5/6, A_0, \rho)$, the symplectic linear invertible transformation

$$\mathcal{P}_{\mathfrak{n}+1} := \mathcal{P}_{\mathfrak{n}} e^{J\mathcal{S}_{\mathfrak{n}+1}} \stackrel{(9.1.19)}{=} e^{J\mathcal{S}_1} \dots e^{J\mathcal{S}_{\mathfrak{n}}} e^{J\mathcal{S}_{\mathfrak{n}+1}} \tag{9.1.46}$$

which has the form (9.1.19) at order $\mathfrak{n}+1$. By the inductive assumption $(\mathbf{P})_{\mathfrak{n}}$, the conjugation identity (9.1.34) holds for all $\lambda \in \Lambda_{\mathfrak{n}}(\varepsilon; \eta, A_0, \rho)$, and we deduce, by (9.1.43), that, for all $\lambda$ in the set $\Lambda_{\mathfrak{n}+1}(\varepsilon; \eta, A_0, \rho)$ defined in (9.1.41), we have

$$\begin{aligned}
\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_0 - J\rho\big)\mathcal{P}_{\mathfrak{n}+1} &\stackrel{(9.1.46)}{=} \big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_0 - J\rho\big)\mathcal{P}_{\mathfrak{n}} e^{J\mathcal{S}_{\mathfrak{n}+1}} \\
&\stackrel{(9.1.34)}{=} \mathcal{P}_{\mathfrak{n}}\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}} - J\rho_{\mathfrak{n}}\big) e^{J\mathcal{S}_{\mathfrak{n}+1}} \\
&\stackrel{(9.1.43)}{=} \mathcal{P}_{\mathfrak{n}} e^{J\mathcal{S}_{\mathfrak{n}+1}}\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}+1} - J\rho_{\mathfrak{n}+1}\big) \\
&\stackrel{(9.1.46)}{=} \mathcal{P}_{\mathfrak{n}+1}\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}+1} - J\rho_{\mathfrak{n}+1}\big)
\end{aligned}$$

which is (9.1.34) at the step $\mathfrak{n}+1$.

We have $\mathcal{P}_{\mathfrak{n}+1} - \mathcal{P}_{\mathfrak{n}} = \mathcal{P}_{\mathfrak{n}}(e^{J\mathcal{S}_{\mathfrak{n}+1}} - \mathrm{Id})$. By (9.1.42), using (3.3.44)-(3.3.45) and the definition of $\delta_{\mathfrak{n}+1}$ in (9.1.37),

$$|e^{J\mathcal{S}_{\mathfrak{n}+1}} - \mathrm{Id}|_{\mathrm{Lip},+,s_1} \lesssim |J\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_1} \lesssim |J\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s_1+1} \lesssim \delta_{\mathfrak{n}+1}^{7/8}\,. \tag{9.1.47}$$

Moreover, by the second inequality of (9.1.21), $|\mathcal{P}_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} \leq 2$. Hence

$$|\mathcal{P}_{\mathfrak{n}+1} - \mathcal{P}_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} \leq \delta_{\mathfrak{n}+1}^{3/4}$$

which is the first inequality in (9.1.21) at the step $\mathfrak{n}+1$ (we obtain in the same way the estimate for $\mathcal{P}_{\mathfrak{n}+1}^{-1}$). As a consequence,

$$|\mathcal{P}_{\mathfrak{n}+1} - \mathrm{Id}|_{\mathrm{Lip},+,s_1} \leq \sum_{k=1}^{\mathfrak{n}+1} \delta_k^{3/4} \leq 2\delta_1^{3/4}$$

which is the second inequality in (9.1.21) at the step $\mathfrak{n}+1$ (we obtain in the same way the estimate for $\mathcal{P}_{\mathfrak{n}+1}^{-1}$). Estimates (9.1.22)-(9.1.23) at the step $\mathfrak{n}+1$ are proved below.

$A_{\mathfrak{n}+1}$ IN (9.1.44) IS A SPLIT ADMISSIBLE OPERATOR IN $\mathcal{C}(2C_1, c_1/2, c_2/2)$, see Definition 8.1.1. By (9.1.45) we have

$$|A_{\mathfrak{n}+1} - A_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} = |R_{\mathfrak{n}+1} - R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} \leq \delta_{\mathfrak{n}+1}^{\frac{3}{4}} = \delta_1^{\frac{3}{4}(\frac{3}{2})^{\mathfrak{n}}} \quad \text{on } \Lambda_{\mathfrak{n}+1}(\varepsilon; 5/6, A_0, \rho)\,,$$

which is (9.1.39) at the step $\mathfrak{n}+1$. As a consequence

$$|R_{\mathfrak{n}+1} - R_0|_{\mathrm{Lip},+,s_1} = |A_{\mathfrak{n}+1} - A_0|_{\mathrm{Lip},+,s_1} \leq \sum_{k=1}^{\mathfrak{n}+1} \delta_1^{\frac{3}{4}(\frac{3}{2})^{k-1}} \leq 2\delta_1^{3/4} \tag{9.1.48}$$

and (9.1.29) at order $\mathfrak{n}+1$ follows: in fact, $|R_0|_{\mathrm{Lip},+,s_1} \leq C_1\varepsilon^2$ by item 1 of Definition 8.1.1 and

$$|R_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_1} \leq |R_0|_{\mathrm{Lip},+,s_1} + 2\delta_1^{3/4} \leq C_1\varepsilon^2 + 2\varepsilon^{9/4} \leq 2C_1\varepsilon^2$$

for $\varepsilon$ small enough, since $\delta_1 = \varepsilon^3$. Recalling (9.1.25), the estimate (9.1.28) at order $\mathfrak{n}+1$ is also a direct consequence of (9.1.48), as well as

$$\|D_{\mathfrak{n}+1} - D_0\|_{\mathrm{Lip},0} = O(\delta_1^{3/4}) = o(\varepsilon^2)\,,$$

which implies (9.1.27) at order $\mathfrak{n}+1$.

Now, since $A_0 \in \mathcal{C}(C_1, c_1, c_2)$ (Definition 8.1.1), either $\mathfrak{d}_\lambda(\mu_i - \mu_j)(\varepsilon,\lambda) \geq c_2\varepsilon^2$ or $\mathfrak{d}_\lambda(\mu_i - \mu_j)(\varepsilon,\lambda) \leq -c_2\varepsilon^2$, see (8.1.5). Thus, by (9.1.27) at order $\mathfrak{n}+1$, we deduce that $\mathfrak{d}_\lambda(\mu_i^{(\mathfrak{n}+1)} - \mu_j^{(\mathfrak{n}+1)})(\varepsilon,\lambda) \geq c_2\varepsilon^2/2$ in the first case and $\mathfrak{d}_\lambda(\mu_i^{(\mathfrak{n}+1)} - \mu_j^{(\mathfrak{n}+1)})(\varepsilon,\lambda) \leq -c_2\varepsilon^2/2$ in the second case, for $\varepsilon$ small enough.

In a similar way (9.1.27) provides properties (8.1.6)-(8.1.7) with constant $c_2/2$ instead of $c_2$ at order $\mathfrak{n}+1$, and (9.1.27)-(9.1.28) provide (8.1.8) with constant $c_1/2$ instead of $c_1$ at order $\mathfrak{n}+1$.

Estimates of $|R_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s}$ and $|\rho_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s}$. We first consider the case $s = s_2$. By (9.1.45) and (9.1.30), (9.1.32), and recalling the definition of $\delta_{\mathfrak{n}+1}$ in (9.1.37),

$$\begin{aligned}|R_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_2} + |\rho_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_2} &\leq \delta_{\mathfrak{n}+1}^{-\frac14}\big(|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2}\big) + \delta_{\mathfrak{n}+1}^{-\frac34}\\ &\ll (C(s_2))^{\mathfrak{n}+1}\delta_{\mathfrak{n}+1}^{-\frac34}(\varepsilon\delta_1^{-\frac12}+1)\,,\end{aligned}$$

which gives (9.1.30) and the second bound in (9.1.32) at the step $\mathfrak{n}+1$.

To estimate the $s$-norms for any $s \geq s_2$, let us introduce the notation

$$u_{\mathfrak{n}}(s) := |R_{\mathfrak{n}}|_{\mathrm{Lip},+,s} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s} \quad \text{and} \quad \alpha(s) := 3\varsigma\frac{s-s_2}{s_2-s_1} \quad \text{as in (9.1.24)}\,. \tag{9.1.49}$$

By (9.1.45), we have the inductive bound

$$u_{\mathfrak{n}+1}(s) = |R_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s} + |\rho_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s} \leq C'(s)\big[\delta_{\mathfrak{n}+1}^{-\frac14}u_{\mathfrak{n}}(s) + \delta_{\mathfrak{n}+1}^{-\frac34}\,\delta_{\mathfrak{n}+1}^{-\alpha(s)}\big]\,.$$

Then, since (9.1.31) and (9.1.33) hold at order $\mathfrak{n}$, we obtain

$$\begin{aligned}u_{\mathfrak{n}+1}(s) &\leq C'(s)\big[\delta_{\mathfrak{n}+1}^{-\frac14}(C(s))^{\mathfrak{n}}\delta_{\mathfrak{n}}^{-\frac34-\alpha(s)}\big(u_0(s)\delta_1^{\frac12+\frac{2\alpha(s)}{3}}+1\big) + \delta_{\mathfrak{n}+1}^{-\frac34}\delta_{\mathfrak{n}+1}^{-\alpha(s)}\big]\\ &\leq (C(s))^{\mathfrak{n}+1}\delta_{\mathfrak{n}+1}^{-\frac34-\alpha(s)}\big(u_0(s)\delta_1^{\frac12+\frac{2\alpha(s)}{3}}+1\big)\end{aligned}$$

for $C(s)$ large enough, and using that $\delta_{\mathfrak{n}+1} = \delta_{\mathfrak{n}}^{3/2}$. Hence the estimates (9.1.31) and (9.1.33) are proved also at order $\mathfrak{n}+1$.

ESTIMATES OF $|\mathcal{P}_{\mathfrak{n}+1}^{\pm 1}|_{\mathrm{Lip},+,s}$. We first consider the case $s = s_2$. By (3.3.26) and the first estimate in (9.1.42) we obtain

$$|e^{J\mathcal{S}_{\mathfrak{n}+1}}|_{\mathrm{Lip},s_2+1} \le 1 + C(s_2) \sum_{k\ge 1} \frac{(C(s_2))^{k-1}}{k!} |\mathcal{S}_{\mathfrak{n}+1}|^{k-1}_{\mathrm{Lip},s_1} |\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s_2+1}$$
$$\le 1 + C'(s_2)|\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s_2+1} \,. \tag{9.1.50}$$

Hence by (9.1.46) and (3.3.24), we get

$$|\mathcal{P}_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_2} \le C(s_2)\big(|\mathcal{P}_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} |e^{J\mathcal{S}_{\mathfrak{n}+1}}|_{\mathrm{Lip},s_1+1} + |\mathcal{P}_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} |e^{J\mathcal{S}_{\mathfrak{n}+1}}|_{\mathrm{Lip},s_2+1}\big)$$
$$\overset{(9.1.47),(9.1.50)}{\le} C(s_2)\big(|\mathcal{P}_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} + |\mathcal{P}_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} |\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s_2+1}\big) \tag{9.1.51}$$

for some new constant $C(s_2)$. Therefore, by (9.1.51), properties (9.1.22), (9.1.21) at order $\mathfrak{n}$, the second inequality in (9.1.42), and recalling the definition of $\delta_{\mathfrak{n}}$ in (9.1.37), we have

$$|\mathcal{P}_{\mathfrak{n}+1}|_{\mathrm{Lip},+,s_2} \le C(s_2)\Big((C(s_2))^{\mathfrak{n}} \delta_{\mathfrak{n}}^{-\frac{3}{4}} (\varepsilon \delta_1^{-\frac{1}{2}} + 1) + 2\delta_{\mathfrak{n}+1}^{-\frac{1}{4}} \big(|R_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_2}\big) + \delta_{\mathfrak{n}+1}^{-\frac{3}{4}}\Big)$$
$$\overset{(9.1.30),(9.1.32)}{\le} (C(s_2))^{\mathfrak{n}+1} \delta_{\mathfrak{n}+1}^{-\frac{3}{4}} (\varepsilon \delta_1^{-\frac{1}{2}} + 1)$$

provided the constant $C(s_2)$ is large enough, proving (9.1.22) at the step $\mathfrak{n}+1$. We obtain in the same way the estimate for $\mathcal{P}_{\mathfrak{n}+1}^{-1}$.

Now, for any $s \ge s_2$, we derive by the last inequality in (9.1.42), (9.1.49), (9.1.31), (9.1.33),

$$|\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s+1} \le C(s)\big[\delta_{\mathfrak{n}+1}^{-\frac{1}{4}} u_{\mathfrak{n}}(s) + \delta_{\mathfrak{n}+1}^{-\frac{3}{4}-\alpha(s)}\big]$$
$$\le (C(s))^{\mathfrak{n}+1} \delta_{\mathfrak{n}+1}^{-\frac{3}{4}-\alpha(s)} \big(u_0(s) \delta_1^{\frac{1}{2}+\frac{2\alpha(s)}{3}} + 1\big) \,. \tag{9.1.52}$$

As in the case $s = s_2$ (see (9.1.50)) we obtain

$$|e^{J\mathcal{S}_{\mathfrak{n}+1}}|_{\mathrm{Lip},s+1} \le 1 + C(s)|\mathcal{S}_{\mathfrak{n}+1}|_{\mathrm{Lip},s+1} \,.$$

Since $\mathcal{P}_{\mathfrak{n}+1} = \mathcal{P}_{\mathfrak{n}} e^{J\mathcal{S}_{\mathfrak{n}+1}}$ we derive the bound (9.1.23) on $|\mathcal{P}_{\mathfrak{n}+1}^{\pm 1}|_{\mathrm{Lip},+,s}$, exactly as in the case $s = s_2$, using the interpolation inequality (3.3.24), the inductive assumptions (9.1.23) and (9.1.21), (9.1.52), and taking the constant $C(s)$ of (9.1.23) large enough.

This completes the iterative proof of $(\mathbf{P}_{\mathfrak{n}})_{\mathfrak{n}\ge 0}$.

THE CANTOR-LIKE SETS $\Lambda_\infty(\varepsilon;\eta,A_0,\rho)$. We define, for $1/2 \le \eta \le 5/6$, the set

$$\Lambda_\infty(\varepsilon;\eta,A_0,\rho) := \bigcap_{\mathfrak{n}=0}^{\infty} \Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho) = \bigcap_{k=0}^{\infty} \Lambda(\varepsilon;\eta+\eta_k,A_k) \tag{9.1.53}$$

where $\Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)$ are defined in (9.1.38) and $\Lambda(\varepsilon;\eta+\eta_k,A_k)$ are defined by Proposition 9.1.1. We recall that the sequence $(\eta_k)$ is defined in (9.1.37).

The sets $\Lambda_\infty(\varepsilon;\eta,A_0,\rho)$ satisfy Property 1 of Corollary 9.1.2 as the sets $\Lambda(\varepsilon;\eta+\eta_k,A_k)$ satisfy Property 1 of Proposition 9.1.1.

PROOF OF PROPERTY 2 FOR THE SETS $\Lambda_\infty(\varepsilon;\eta,A_0,\rho)$. The complementary set of $\Lambda_\infty(\varepsilon;\eta,A_0,\rho)$ may be decomposed as (the sets $\Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)$ in (9.1.38) are decreasing in $\mathfrak{n}$)

$$\begin{aligned}
\Lambda_\infty(\varepsilon;\eta,A_0,\rho)^c &\overset{(9.1.53)}{=} \bigcup_{\mathfrak{n}=0}^{\infty}\Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)^c \\
&\overset{\Lambda_0=\widetilde\Lambda}{=} \widetilde\Lambda^c\cup\bigcup_{k=0}^{\infty}\Big(\Lambda_k(\varepsilon;\eta,A_0,\rho)\cap\Lambda_{k+1}(\varepsilon;\eta,A_0,\rho)^c\Big) \\
&\overset{(9.1.38)}{=} \widetilde\Lambda^c\cup\bigcup_{k=0}^{\infty}\Big(\Lambda_k(\varepsilon;\eta,A_0,\rho)\cap\big(\Lambda_k(\varepsilon;\eta,A_0,\rho)\cap\Lambda(\varepsilon;\eta+\eta_k,A_k)\big)^c\Big) \\
&= \widetilde\Lambda^c\cup\bigcup_{k=0}^{\infty}\Big(\Lambda_k(\varepsilon;\eta,A_0,\rho)\cap\Lambda(\varepsilon;\eta+\eta_k,A_k)^c\Big) \\
&\overset{\Lambda_0=\widetilde\Lambda}{=} \widetilde\Lambda^c\cup\Lambda(\varepsilon;\eta,A_0)^c\bigcup_{k=1}^{\infty}\Big(\Lambda_k(\varepsilon;\eta,A_0,\rho)\cap\Lambda(\varepsilon;\eta+\eta_k,A_k)^c\Big)\,. \qquad (9.1.54)
\end{aligned}$$

By Property 2 of Proposition 9.1.1 we have

$$|\Lambda(\varepsilon;1/2,A_0)^c\cap\widetilde\Lambda|\le b(\varepsilon)\quad\text{with }\lim_{\varepsilon\to0}b(\varepsilon)=0\,. \qquad (9.1.55)$$

Moreover, for $k\ge1$, we have, by the definition of $\Lambda_k(\varepsilon;\eta,A_0,\rho)$ in (9.1.38),

$$\begin{aligned}
&\Lambda_k\big(\varepsilon;\eta,A_0,\rho\big)\bigcap\Lambda\Big(\varepsilon;\eta+\eta_k,A_k\Big)^c = \\
&\Lambda_k\big(\varepsilon;\eta,A_0,\rho\big)\bigcap\Lambda\Big(\varepsilon;\eta+\eta_k,A_k\Big)^c\bigcap\Lambda\big(\varepsilon;\eta+\eta_{k-1},A_{k-1}\big)\,. \qquad (9.1.56)
\end{aligned}$$

Now, since, for all $k\ge1$ we have $|A_k-A_{k-1}|_{+,s_1}\le\delta_k^{3/4}$ on $\Lambda_k$ (see (9.1.39)) and $\eta_k=\eta_{k-1}+\delta_k^{3/8}$ (see (9.1.37)), we deduce by (9.1.1) and (9.1.56) the Lebesgue measure estimate,

$$\Big|\Lambda_k\big(\varepsilon;\frac12,A_0,\rho\big)\bigcap\Lambda\Big(\varepsilon;\frac12+\eta_k,A_k\Big)^c\Big|\le\delta_k^{3\alpha/4}\,,\quad\forall k\ge1\,. \qquad (9.1.57)$$

In conclusion, by (9.1.54), (9.1.55), (9.1.57) we obtain, recalling (9.1.37),

$$\begin{aligned}
|\Lambda_\infty(\varepsilon;1/2,A_0,\rho)^c\cap\widetilde\Lambda| &\le b(\varepsilon)+\sum_{k=1}^{\infty}\delta_1^{\frac{3\alpha}{4}(\frac32)^{k-1}} \\
&\le b(\varepsilon)+2\delta_1^{\frac{3\alpha}{4}}=b(\varepsilon)+2\varepsilon^{\frac{9\alpha}{4}}=:b_1(\varepsilon)
\end{aligned}$$

since $\delta_1 = \varepsilon^3$. Property 2 of Corollary 9.1.2 is proved.

PROOF OF (9.1.36). Actually we prove that, if $A_0' \in \mathcal{C}(C_1, c_1, c_2)$, $(R_0', \rho')$ satisfy (8.2.2) as $(R_0, \rho)$ and $(A_0', \rho')$ satisfy (9.1.35), then

$$|A_{\mathfrak{n}}' - A_{\mathfrak{n}}|_{+,s_1} \leq \delta^{4/5}, \quad \forall \lambda \in \Lambda_{\mathfrak{n}}(\varepsilon; 5/6, A_0, \rho) \cap \Lambda_{\mathfrak{n}}(\varepsilon; 5/6, A_0', \rho'), \tag{9.1.58}$$

where the sets $\Lambda_{\mathfrak{n}}$ are defined in (9.1.38). Let

$$\Delta_{\mathfrak{n}}^{(1)} := |R_{\mathfrak{n}}' - R_{\mathfrak{n}}|_{+,s_1}, \quad \Delta_{\mathfrak{n}}^{(2)} := |\rho_{\mathfrak{n}}' - \rho_{\mathfrak{n}}|_{+,s_1}, \quad \Delta_{\mathfrak{n}} := \Delta_{\mathfrak{n}}^{(1)} + \Delta_{\mathfrak{n}}^{(2)}. \tag{9.1.59}$$

The assumption (9.1.35) means that $\Delta_0 \leq \delta$ on $\widetilde{\Lambda} \cap \widetilde{\Lambda}'$. We recall that $(A_{\mathfrak{n}+1}, \rho_{\mathfrak{n}+1})$ (resp. $(A_{\mathfrak{n}+1}', \rho_{\mathfrak{n}+1}')$) is built applying Proposition 9.1.1 to $(A_{\mathfrak{n}}, \rho_{\mathfrak{n}})$ (resp. $(A_{\mathfrak{n}}', \rho_{\mathfrak{n}}')$) instead of $(A_0, \rho_0)$ (resp. $(A_0', \rho_0')$). Hence, by (9.1.15)-(9.1.16) (with $\delta_1$ replaced by $\delta_{\mathfrak{n}+1}$), for all $\mathfrak{n} \in \mathbb{N}$, for all $\lambda \in \Lambda_{\mathfrak{n}+1}(\varepsilon; 5/6, A_0, \rho) \cap \Lambda_{\mathfrak{n}+1}(\varepsilon; 5/6, A_0', \rho')$, we have the iterative inequalities

$$\Delta_{\mathfrak{n}+1}^{(1)} \leq \Delta_{\mathfrak{n}}^{(1)} + C\Delta_{\mathfrak{n}}^{(2)} \qquad \text{and} \qquad \Delta_{\mathfrak{n}+1}^{(2)} \leq \delta_{\mathfrak{n}+1}^{1/2} \Delta_{\mathfrak{n}}^{(1)} + \delta_{\mathfrak{n}+1}^{-1/20} \Delta_{\mathfrak{n}}^{(2)}. \tag{9.1.60}$$

As a consequence $\Delta_{\mathfrak{n}+1} \leq \delta_{\mathfrak{n}+1}^{-\frac{1}{19}} \Delta_{\mathfrak{n}}$ and, recalling the definition of $\delta_{\mathfrak{n}}$ in (9.1.37), we deduce that, for any $\mathfrak{n} \geq 1$, for all $\lambda \in \Lambda_{\mathfrak{n}}(\varepsilon; \eta, A_0, \rho) \cap \Lambda_{\mathfrak{n}}(\varepsilon; \eta, A_0', \rho')$, we have

$$\Delta_{\mathfrak{n}} \leq (\delta_{\mathfrak{n}} \ldots \delta_1)^{-\frac{1}{19}} \Delta_0 \leq \delta_1^{-\frac{1}{19}(1+\ldots+(\frac{3}{2})^{\mathfrak{n}-1})} \delta \leq \delta_1^{-\frac{2}{19}(\frac{3}{2})^{\mathfrak{n}}} \delta. \tag{9.1.61}$$

Let $\mathfrak{n}_0 \geq 1$ be the integer such that

$$\delta_1^{(\frac{3}{2})^{\mathfrak{n}_0}} < \delta < \delta_1^{(\frac{3}{2})^{\mathfrak{n}_0 - 1}}. \tag{9.1.62}$$

Thus, by (9.1.61) and the second inequality in (9.1.62), we get

$$\forall \mathfrak{n} \leq \mathfrak{n}_0, \quad \Delta_{\mathfrak{n}} \leq \delta_1^{-\frac{2}{19}(\frac{3}{2})^{\mathfrak{n}_0}} \delta \leq \delta^{-\frac{3}{19}} \delta = \delta^{\frac{16}{19}}. \tag{9.1.63}$$

On the other hand, recalling (9.1.59) we bound, using the first estimate in (9.1.32),

$$\forall \mathfrak{n} \geq \mathfrak{n}_0 + 1, \quad \Delta_{\mathfrak{n}}^{(2)} \leq |\rho_{\mathfrak{n}}|_{+,s_1} + |\rho_{\mathfrak{n}}'|_{+,s_1} \leq 2\delta_1^{(\frac{3}{2})^{\mathfrak{n}}}. \tag{9.1.64}$$

Applying iteratively the first inequality in (9.1.60) we get, $\forall \mathfrak{n} \geq \mathfrak{n}_0 + 1$,

$$\begin{aligned} |A_{\mathfrak{n}}' - A_{\mathfrak{n}}|_{+,s_1} = \Delta_{\mathfrak{n}}^{(1)} & \leq \Delta_{\mathfrak{n}_0}^{(1)} + C\big(\Delta_{\mathfrak{n}_0}^{(2)} + \ldots + \Delta_{\mathfrak{n}-1}^{(2)}\big) \\ & \overset{(9.1.63),(9.1.64)}{\leq} \delta^{\frac{16}{19}} + 2C(\delta_1^{(\frac{3}{2})^{\mathfrak{n}_0}} + \ldots + \delta_1^{(\frac{3}{2})^{\mathfrak{n}-1}}) \\ & \leq \delta^{\frac{16}{19}} + 3C\delta_1^{(\frac{3}{2})^{\mathfrak{n}_0}} \\ & \overset{(9.1.62)}{\leq} \delta^{\frac{16}{19}} + 3C\delta \leq \delta^{\frac{4}{5}} \end{aligned} \tag{9.1.65}$$

for $\delta$ small enough. In conclusion (9.1.63) and (9.1.65) imply (9.1.58).

PROOF OF PROPERTY 3 FOR THE SETS $\Lambda_\infty(\varepsilon;\eta,A_0,\rho)$. By (9.1.54) (with $A_0',\rho'$ instead of $A_0,\rho$) and (9.1.53) we deduce, for all $(1/2)+\delta^{2/5}\le\eta\le 5/6$, the inclusion

$$\mathcal{M} := \widetilde{\Lambda}'\cap\Lambda_\infty(\varepsilon;\eta,A_0',\rho')^c\cap\Lambda_\infty(\varepsilon;\eta-\delta^{2/5},A_0,\rho)\subset\mathcal{M}_0\bigcup\Big(\bigcup_{\mathfrak{n}\ge1}\mathcal{M}_{\mathfrak{n}}\Big) \tag{9.1.66}$$

where

$$\mathcal{M}_0 := \widetilde{\Lambda}'\cap\Lambda(\varepsilon;\eta,A_0')^c\cap\Lambda(\varepsilon;\eta-\delta^{2/5},A_0)$$

and

$$\mathcal{M}_{\mathfrak{n}} := \Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0',\rho')\bigcap\Lambda(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}}')^c\bigcap\Lambda_{\mathfrak{n}}(\varepsilon;\eta-\delta^{2/5},A_0,\rho)\bigcap\Lambda(\varepsilon;\eta+\eta_{\mathfrak{n}}-\delta^{2/5},A_{\mathfrak{n}})\,.$$

By the assumption (9.1.17) we have $|R_0'-R_0|_{+,s_1}\le\delta$ on $\widetilde{\Lambda}\cap\widetilde{\Lambda}'$, and therefore property 3 of Proposition 9.1.1 and the fact that $\delta^{1/2}\le\delta^{2/5}$ imply that, for all $(1/2)+\delta^{2/5}\le\eta\le5/6$,

$$|\mathcal{M}_0|\le|\widetilde{\Lambda}'\cap\Lambda(\varepsilon;\eta,A_0')^c\cap\Lambda(\varepsilon;\eta-\delta^{1/2},A_0)|\le\delta^\alpha\,. \tag{9.1.67}$$

For $\mathfrak{n}\ge1$, we have, by (9.1.38), the inclusion

$$\mathcal{M}_{\mathfrak{n}}\subset\Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0',\rho')\cap\Lambda(\varepsilon;\eta+\eta_{\mathfrak{n}-1},A_{\mathfrak{n}-1}')\cap\Lambda(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}}')^c\,.$$

and therefore, since $|R_{\mathfrak{n}}'-R_{\mathfrak{n}-1}'|_{+,s_1}=|A_{\mathfrak{n}}'-A_{\mathfrak{n}-1}'|_{+,s_1}\le\delta_{\mathfrak{n}}^{3/4}$ on $\Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0',\rho')$ by (9.1.39), the estimate (9.1.1) and $\eta_{\mathfrak{n}}=\eta_{\mathfrak{n}-1}+\delta_{\mathfrak{n}}^{3/8}$, imply

$$|\mathcal{M}_{\mathfrak{n}}|\le\delta_{\mathfrak{n}}^{3\alpha/4}\,. \tag{9.1.68}$$

On the other hand, by (9.1.58), $|A_{\mathfrak{n}}-A_{\mathfrak{n}}'|_{+,s_1}\le\delta^{4/5}$ for any

$$\lambda\in\mathcal{M}_n\subset\Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0',\rho')\bigcap\Lambda_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)$$

and we deduce, by (9.1.1), the measure estimate

$$|\mathcal{M}_{\mathfrak{n}}|\le\big|\Lambda(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}}')^c\bigcap\Lambda(\varepsilon;\eta+\eta_{\mathfrak{n}}-\delta^{2/5},A_{\mathfrak{n}})\big|\le\delta^{4\alpha/5}\,. \tag{9.1.69}$$

Finally (9.1.66), (9.1.67), (9.1.68), (9.1.69) imply the measure estimate

$$|\mathcal{M}|\le\delta^\alpha+\sum_{\mathfrak{n}\ge1}\min(\delta_{\mathfrak{n}}^{3\alpha/4},\delta^{4\alpha/5})\le\delta^{\alpha/2}$$

for $\delta$ small, proving (9.1.18). The proof of Corollary 9.1.2 is complete. ■

The rest of the Chapter is devoted to the Proof of Proposition 9.1.1.

## 9.2 The linearized homological equation

We consider the linear map

$$\mathcal{S} \mapsto J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0]$$

where $\mathcal{S} := \mathcal{S}(\varphi)$, $\varphi \in \mathbb{T}^{|\mathbb{S}|}$, has the form

$$\begin{gathered} \mathcal{S}(\varphi) = \begin{pmatrix} d(\varphi) & a(\varphi)^* \\ a(\varphi) & 0 \end{pmatrix} \in \mathcal{L}(H_{\mathbb{S}}^{\perp}), \\ d(\varphi) = d^*(\varphi) \in \mathcal{L}(H_{\mathbb{F}}), \quad a(\varphi) \in \mathcal{L}(H_{\mathbb{F}}, H_{\mathbb{G}}), \end{gathered} \tag{9.2.1}$$

and it is self-adjoint.

Recalling (8.1.2) and using that $D_0$ and $J$ commute, $J^2 = -\mathrm{Id}$, we have

$$\begin{aligned} & J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] = \\ & \begin{pmatrix} J\bar{\omega}_\varepsilon \cdot \partial_\varphi d + D_0 d + JdJD_0 & J\bar{\omega}_\varepsilon \cdot \partial_\varphi a^* + Ja^*JV_0 + D_0 a^* \\ J\bar{\omega}_\varepsilon \cdot \partial_\varphi a - JV_0 Ja + JaJD_0 & 0 \end{pmatrix}. \end{aligned} \tag{9.2.2}$$

The key step in the proof of the splitting Proposition 9.1.1 (see section 9.4) is, given $\rho(\varphi)$ of the form

$$\begin{gathered} \rho(\varphi) = \begin{pmatrix} \rho_1(\varphi) & \rho_2(\varphi)^* \\ \rho_2(\varphi) & 0 \end{pmatrix} \in \mathcal{L}(H_{\mathbb{S}}^{\perp}), \\ \rho_1(\varphi) = \rho_1^*(\varphi) \in \mathcal{L}(H_{\mathbb{F}}), \ \rho_2(\varphi) \in \mathcal{L}(H_{\mathbb{F}}, H_{\mathbb{G}}), \end{gathered} \tag{9.2.3}$$

to solve (approximately) the "homological" equation

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] = J\rho. \tag{9.2.4}$$

The equation (9.2.4) amounts to solve the pair of decoupled equations

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi d + D_0 d + JdJD_0 = J\rho_1, \tag{9.2.5}$$

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi a - JV_0 Ja + JaJD_0 = J\rho_2. \tag{9.2.6}$$

Note that, taking the adjoint equation of (9.2.6), multiplying by $J$ on the left and the right, since $V_0$ and $D_0$ are self-adjoint and $JD_0J = -D_0$, $J^* = -J$, we obtain

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi a^* + Ja^*JV_0 + D_0 a^* = J\rho_2^*,$$

which is the equation in the top right in (9.2.4), (9.2.2), (9.2.3).

**Remark 9.2.1.** *We shall solve only approximately the homological equation* (9.2.4) *up to terms which are Fourier supported on high frequencies. The main reason is that the multiscale Proposition 4.1.5 provides tame estimates of the inverses of finite dimensional restrictions of infinite dimensional operators. This is sufficient for proving Proposition 9.1.1.*

We shall decompose an operator $\rho$ of the form (9.2.3), as well as $\mathcal{S}$ in (9.2.1), in the following way. The operator $\rho_1 \in \mathcal{L}(H_{\mathbb{F}})$ can be represented as a finite dimensional self-adjoint square matrix $((\rho_1)_i^j)_{i,j\in\mathbb{F}}$ with entries $(\rho_1)_i^j \in \mathcal{L}(H_j, H_i)$. Using in each subspace $H_j$, $j \in \mathbb{F}$, the basis $((\Psi_j, 0), (0, \Psi_j))$, see (3.1.9), we identify each operator $(\rho_1)_i^j(\varphi) \in \mathcal{L}(H_j, H_i)$ with a $2\times 2$-real matrix that we still denote by $(\rho_1)_i^j(\varphi) \in \mathrm{Mat}_2(\mathbb{R})$. We shall also Fourier expand

$$(\rho_1)_i^j(\varphi) = \sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|}} [\widehat{\rho}_1]_i^j(\ell) e^{\mathrm{i}\ell\cdot\varphi}, \quad [\widehat{\rho}_1]_i^j(\ell) \in \mathrm{Mat}_2(\mathbb{C}), \quad \overline{[\widehat{\rho}_1]_i^j(\ell)} = [\widehat{\rho}_1]_i^j(-\ell), \tag{9.2.7}$$

where $[\widehat{\rho}_1]_i^j(0)$ is the average

$$[\widehat{\rho}_1]_i^j(0) = \frac{1}{(2\pi)^{|\mathbb{S}|}} \int_{\mathbb{T}^{|\mathbb{S}|}} (\rho_1)_i^j(\varphi)\, d\varphi \in \mathrm{Mat}_2(\mathbb{R}). \tag{9.2.8}$$

The operator $\rho_2 \in \mathcal{L}(H_{\mathbb{F}}, H_{\mathbb{G}})$ is identified, as in (3.2.3), with $(\rho_2^j)_{j\in\mathbb{F}}$ where $\rho_2^j \in \mathcal{L}(H_j, H_{\mathbb{G}})$, which, using in $H_j$ the basis $((\Psi_j, 0), (0, \Psi_j))$, can be identified with a vector of $H_{\mathbb{G}} \times H_{\mathbb{G}}$.

We also recall that, for a $\varphi$-dependent family of operators $\rho(\varphi)$, $\varphi \in \mathbb{T}^{|\mathbb{S}|}$, of the form (9.2.3), we have the estimates (3.3.44)-(3.3.45).

The next lemma provides an approximate solution of the homological equation (9.2.4).

**Lemma 9.2.2. (Homological equations)** *Given $C_1 > 0$, $c_1 > 0$, $c_2 > 0$, there is $\varepsilon_1 > 0$ such that $\forall \varepsilon \in (0, \varepsilon_1)$, for each admissible split operator $A_0 \in \mathcal{C}(C_1, c_1, c_2)$ (see Definition 8.1.1), defined for $\lambda \in \widetilde{\Lambda}$, there are closed subsets $\Lambda(\varepsilon; \eta, A_0) \subset \widetilde{\Lambda}$, $1/2 \le \eta \le 1$, satisfying the properties 1-3 of Proposition 9.1.1, such that, if $\rho \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$ has the form (9.2.3), $\rho$ is $L^2$-self-adjoint, defined for $\lambda \in \widetilde{\Lambda}$, and satisfies*

$$|\rho|_{\mathrm{Lip},+,s_1} \le \delta_1^{\frac{9}{10}}, \ \ \delta_1 \le \varepsilon^3, \qquad \delta_1^{\frac{11}{10}}(|R_0|_{\mathrm{Lip},+,s_2} + |\rho|_{\mathrm{Lip},+,s_2}) \le \varepsilon, \tag{9.2.9}$$

$$[\widehat{\rho}_1]_j^j(0) \in M_- \ (\textit{recall}\ (9.2.8), (3.2.18)), \quad \forall j \in \mathbb{F}, \ \forall \lambda \in \widetilde{\Lambda}, \tag{9.2.10}$$

*then there is a linear self-adjoint operator $\mathcal{S} := \mathcal{S}(\varepsilon, \lambda)(\varphi) \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$ of the form (9.2.1), defined for all $\lambda \in \Lambda(\varepsilon; 1, A_0)$, satisfying (9.1.7)-(9.1.8), such that*

$$|J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] - J\rho|_{\mathrm{Lip},+,s_1} \le \delta_1^{\frac{7}{4}}, \tag{9.2.11}$$

$$|J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0]|_{\mathrm{Lip},+,s_2} \le \delta_1^{-\frac{1}{4}}\big(|R_0|_{\mathrm{Lip},+,s_2} + |\rho|_{\mathrm{Lip},+,s_2}\big) + \delta_1^{-\frac{3}{4}}, \tag{9.2.12}$$

*and, more generally, for all $s \ge s_2$,*

$$|J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0]|_{\mathrm{Lip},+,s} \le C(s)\Big[\delta_1^{-\frac{1}{4}}\big(|R_0|_{\mathrm{Lip},+,s} + |\rho|_{\mathrm{Lip},+,s}\big) + \delta_1^{-\frac{3}{4}}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\Big]. \tag{9.2.13}$$

*At last, denoting by $\mathcal{S}_{A_0,\rho}$ and $\mathcal{S}_{A_0',\rho'}$ the operators defined as above associated, respectively, to $(A_0,\rho)$ and $(A_0',\rho')$, we have, for all $\lambda \in \Lambda(\varepsilon;1,A_0) \cap \Lambda(\varepsilon;1,A_0')$,*

$$|\mathcal{S}_{A_0,\rho} - \mathcal{S}_{A_0',\rho'}|_{+,s_1} \le \delta_1^{\frac{3}{4}} |A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}} |\rho - \rho'|_{+,s_1} \,, \tag{9.2.14}$$

*and*

$$\begin{aligned} &\big|(J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S}_{A_0,\rho} + [J\mathcal{S}_{A_0,\rho}, JA_0] - J\rho) - (J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S}_{A_0',\rho'} + [J\mathcal{S}_{A_0',\rho'}, JA_0] - J\rho')\big|_{+,s_1} \\ &\le \delta_1^{\frac{3}{4}} |A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}} |\rho - \rho'|_{+,s_1} \,. \end{aligned} \tag{9.2.15}$$

The proof of Lemma 9.2.2 is given in the next section.

## 9.3 Solution of homological equations: proof of Lemma 9.2.2

**Step 1: approximate solution of the homological equation** (9.2.5)**.** We represent a linear operator $d(\varphi) \in \mathcal{L}(H_{\mathbb{F}})$ by a finite dimensional square matrix $(d_i^j(\varphi))_{i,j\in\mathbb{F}}$ with entries $d_i^j(\varphi) \in \mathcal{L}(H_j,H_i) \simeq \mathrm{Mat}_2(\mathbb{R})$. Since the symplectic operator $J$ leaves invariant each subspace $H_j$ and

$$D_0 = \mathrm{Diag}_{j\in\mathbb{F}}\, \mu_j(\varepsilon,\lambda)\mathrm{Id}_2$$

(see (8.1.3)), the equation (9.2.5) is equivalent to

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi d_i^j(\varphi) + \mu_i(\varepsilon,\lambda) d_i^j(\varphi) + \mu_j(\varepsilon,\lambda) J d_i^j(\varphi) J = J(\rho_1)_i^j(\varphi) \,, \qquad \forall i,j\in\mathbb{F} \quad \text{where} \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} ,$$

and, by a Fourier series expansion with respect to the variable $\varphi \in \mathbb{T}^{|\mathbb{S}|}$, writing

$$d_i^j(\varphi) = \sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|}} \widehat{d}_i^j(\ell) e^{\mathrm{i}\ell\cdot\varphi} \,, \quad \widehat{d}_i^j(\ell) \in \mathrm{Mat}_2(\mathbb{C}) \,, \quad \overline{\widehat{d}_i^j(\ell)} = \widehat{d}_i^j(-\ell) \,, \tag{9.3.1}$$

to

$$\mathrm{i}(\bar{\omega}_\varepsilon \cdot \ell) J \widehat{d}_i^j(\ell) + \mu_i(\varepsilon,\lambda)\widehat{d}_i^j(\ell) + \mu_j(\varepsilon,\lambda) J \widehat{d}_i^j(\ell) J = J[\widehat{\rho}_1]_i^j(\ell) \,, \quad \forall i,j\in\mathbb{F}\,,\ \ell\in\mathbb{Z}^{|\mathbb{S}|} \,. \tag{9.3.2}$$

In order to solve (9.3.2) we have to study the linear operator

$$T_{ij\ell} : \mathrm{Mat}_2(\mathbb{C}) \to \mathrm{Mat}_2(\mathbb{C}) \,, \quad \mathtt{d} \mapsto \mathrm{i}\bar{\omega}_\varepsilon \cdot \ell\, J\mathtt{d} + \mu_i(\varepsilon,\lambda)\mathtt{d} + \mu_j(\varepsilon,\lambda) J\mathtt{d}J \,. \tag{9.3.3}$$

In the basis $(M_1, M_2, M_3, M_4)$ of $\mathrm{Mat}_2(\mathbb{C})$ defined in (3.2.19)-(3.2.20) the linear operator $T_{ij\ell}$ is represented by the following self-adjoint matrix (recall that $JM_l = M_l J$ for $l = 1, 2$ and $JM_l = -M_l J$ for $l = 3, 4$, and $JM_1 = -M_2$, $JM_3 = -M_4$)

$$\begin{pmatrix} (\mu_i - \mu_j)(\varepsilon, \lambda) & \mathrm{i}\bar{\omega}_\varepsilon \cdot \ell & 0 & 0 \\ -\mathrm{i}\bar{\omega}_\varepsilon \cdot \ell & (\mu_i - \mu_j)(\varepsilon, \lambda) & 0 & 0 \\ 0 & 0 & (\mu_i + \mu_j)(\varepsilon, \lambda) & \mathrm{i}\bar{\omega}_\varepsilon \cdot \ell \\ 0 & 0 & -\mathrm{i}\bar{\omega}_\varepsilon \cdot \ell & (\mu_i + \mu_j)(\varepsilon, \lambda) \end{pmatrix}. \tag{9.3.4}$$

As a consequence the eigenvalues of $T_{ij\ell}$ are

$$\pm\bar{\omega}_\varepsilon \cdot \ell + \mu_i(\varepsilon, \lambda) - \mu_j(\varepsilon, \lambda), \quad \pm\bar{\omega}_\varepsilon \cdot \ell + \mu_i(\varepsilon, \lambda) + \mu_j(\varepsilon, \lambda). \tag{9.3.5}$$

To impose non-resonance conditions we define the sets, for $1/2 \le \eta \le 1$,

$$\begin{aligned} \Lambda^1(\varepsilon; \eta, A_0) := \Big\{ \lambda \in \widetilde{\Lambda} \, : \, & |\bar{\omega}_\varepsilon \cdot \ell \pm \mu_j(\varepsilon, \lambda) \pm \mu_i(\varepsilon, \lambda)| \ge \frac{\gamma_1}{2\eta \langle \ell \rangle^\tau}, \\ & \forall (\ell, i, j) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{F} \times \mathbb{F}, \ (\ell, i, j) \neq (0, j, j) \Big\} \end{aligned} \tag{9.3.6}$$

where the constant $\gamma_1 = \gamma_0/2$ (recall that $\gamma_0$ is fixed in (1.2.6)) and

$$\tau \ge (3/2)\tau_1 + 3 + |\mathbb{S}| \tag{9.3.7}$$

where $\tau_1$ is defined in (1.2.28). The inequalities in (9.3.6) are second-order Melnikov non-resonance concerning only a finite number of normal frequencies.

**Remark 9.3.1.** *Since $\mu_j(\varepsilon, \lambda) > c_0 > 0$, for any $j \in \mathbb{N}$, if $\gamma_1 \le 2c_0$, then the inequality $|\bar{\omega}_\varepsilon \cdot \ell + \mu_j(\varepsilon, \lambda) + \mu_i(\varepsilon, \lambda)| \ge \frac{\gamma_1}{2\eta\langle \ell \rangle^\tau}$ in (9.3.6) holds for $\ell = 0$ and $j = i$, for all $\lambda \in \Lambda$, $\eta \in [1/2, 1]$.*

In the next lemma we find a solution $d_N$ of the projected homological equation

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi d_N + D_0 d_N + J d_N J D_0 = \Pi_N J \rho_1 \tag{9.3.8}$$

where here the projector $\Pi_N$ applies to functions depending only on the variable $\varphi$, namely

$$\Pi_N : \quad h(\varphi) = \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} h_\ell e^{\mathrm{i}\ell\cdot\varphi} \quad \mapsto \quad (\Pi_N h)(\varphi) := \sum_{|\ell| \le N} h_\ell e^{\mathrm{i}\ell\cdot\varphi}. \tag{9.3.9}$$

**Lemma 9.3.2. (Homological equation** (9.2.5)**)** *Let $\rho$ be a self-adjoint operator of the form* (9.2.3)*, defined for $\lambda \in \widetilde{\Lambda}$, satisfying* (9.2.9)*, and assume that the average $[\widehat{\rho}_1]_j^j(0) \in M_-$, $\forall j \in \mathbb{F}$, $\forall \lambda$, i.e.* (9.2.10) *holds. Let*

$$N \in \Big[\delta_1^{-\frac{3}{s_2 - s_1}} - 1, \delta_1^{-\frac{3}{s_2 - s_1}} + 1\Big]. \tag{9.3.10}$$

*Then, for all* $\lambda \in \Lambda^1(\varepsilon; 1, A_0)$ *(set defined in* (9.3.6)*) there exists* $d_N := d_N(\lambda; \varphi) \in \mathcal{L}(H_{\mathbb{F}})$, $d_N = d_N^*$, *satisfying*

$$|d_N|_{\mathrm{Lip}, s_1+1} \ll \delta_1^{\frac{7}{8}}, \quad |d_N|_{\mathrm{Lip}, s+1} \ll \delta_1^{-\frac{1}{40}} |\rho|_{\mathrm{Lip}, s}, \ \forall s \geq s_1, \tag{9.3.11}$$

*which solves the projected homological equation* (9.3.8)*, and which is an approximate solution of the homological equation* (9.2.5) *in the sense that*

$$|J\bar{\omega}_\varepsilon \cdot \partial_\varphi d_N + D_0 d_N + J d_N J D_0 - J\rho_1|_{\mathrm{Lip}, s_1} \ll \delta_1^{7/4}. \tag{9.3.12}$$

*At last, denoting by* $d'_N$ *the operator associated to* $D'_0, \rho'$ *as above, we have, for all* $\lambda \in \Lambda^1(\varepsilon; 1, A'_0) \cap \Lambda^1(\varepsilon; 1, A_0)$,

$$|d'_N - d_N|_{s_1} \leq \delta_1^{\frac{7}{8}} |D'_0 - D_0|_{s_1} + \delta_1^{-\frac{1}{40}} |\rho'_1 - \rho_1|_{s_1} \tag{9.3.13}$$

*and*

$$\begin{aligned} &\big|\big(J\bar{\omega}_\varepsilon \cdot \partial_\varphi d'_N + D'_0 d'_N + J d'_N J D'_0 - J\rho'_1\big) - \big(J\bar{\omega}_\varepsilon \cdot \partial_\varphi d_N + D_0 d_N + J d_N J D_0 - J\rho_1\big)\big|_{s_1} \\ &\leq |\rho'_1 - \rho_1|_{s_1}. \end{aligned} \tag{9.3.14}$$

PROOF. We split the proof in some steps.

STEP 1. *For any* $(\ell, i, j) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{F} \times \mathbb{F}$, *for all* $\lambda \in \Lambda^1(\varepsilon; 1, A_0)$, *there exists a solution* $\widehat{d}_i^j(\ell)$ *of the equation* (9.3.2)*, satisfying the reality condition* (9.3.1) *and*

$$\|\widehat{d}_i^j(\ell)\| \leq C\langle \ell \rangle^\tau \|[\widehat{\rho}_1]_i^j(\ell)\| \tag{9.3.15}$$

*where* $\| \ \|$ *denotes some norm in* $\mathrm{Mat}_2(\mathbb{C})$.

Recalling (9.3.5) and the definition of $\Lambda^1(\varepsilon; 1, A_0)$ in (9.3.6), we have that, for all $\lambda \in \Lambda^1(\varepsilon; 1, A_0)$, the operators $T_{ij\ell}$ defined in (9.3.3) are isomorphisms of $\mathrm{Mat}_2(\mathbb{C})$ for $\ell \neq 0$ or $i \neq j$, and

$$\|T_{ij\ell}^{-1}\| \leq C\langle \ell \rangle^\tau \tag{9.3.16}$$

where $\| \ \|$ denotes some norm in $\mathcal{L}(\mathrm{Mat}_2(\mathbb{C}))$ and $C := C(\gamma_1)$. Hence, in this case, there exists a unique solution $\widehat{d}_i^j(\ell) = T_{ij\ell}^{-1}\big(J[\widehat{\rho}_1]_i^j(\ell)$ of the equation (9.3.2).

Let us consider the case $\ell = 0$ and $i = j$. By (9.3.4), the linear operator $T_{jj0}$ is represented, in the basis $(M_1, M_2, M_3, M_4)$ of $\mathrm{Mat}_2(\mathbb{C})$, by the matrix

$$2\mu_j(\varepsilon, \lambda) \begin{pmatrix} 0_2 & 0 \\ 0 & \mathrm{Id}_2 \end{pmatrix}.$$

Moreover there is $c_0 > 0$ such that $\mu_j(\varepsilon, \lambda) > c_0$ for any $j \in \mathbb{N}$, $\varepsilon \in (0, \varepsilon_0)$, $\lambda \in \Lambda$. Hence the range of $T_{jj0}$, is the subspace $M_- = \mathrm{Span}(M_3, M_4)$ of $\mathrm{Mat}_2(\mathbb{C})$ defined in (3.2.18), (3.2.20), and

$$\forall \widetilde{\rho} \in M_-, \quad \exists!\, \widetilde{d} \in M_- \quad \text{solving} \quad T_{jj0}\widetilde{d} = \widetilde{\rho} \quad \text{with} \quad \|\widetilde{d}\| \lesssim \|\widetilde{\rho}\|;$$

with a little abuse of notation, we denote $\widetilde{d}$ by $T^{-1}_{jj0}(\widehat{\rho})$. By the assumption $[\widehat{\rho}_1]^j_j(0) \in M_-$ and the fact that $M_-$ is stable under the action of $J$, we have that $J[\widehat{\rho}_1]^j_j(0)$ is in $M_-$ and there is a unique solution $\widehat{d}^j_j(0) \in M_-$ of the equation (9.3.2).

To conclude the proof of Step 1, it remains to show that the matrices $\widehat{d}^j_i(\ell) \in \mathrm{Mat}_2(\mathbb{C})$ satisfy the reality condition (9.3.1). It is consequence of $T_{ij(-\ell)} = \overline{T_{ij\ell}}$, the fact that $[\widehat{\rho}_1]^j_i(\ell)$ satisfies the reality condition (9.2.7), and the uniqueness of the solutions of $T_{ij\ell}(\widehat{d}^j_i(\ell)) = J[\widehat{\rho}_1]^j_i(\ell)$ (with the condition $d^j_i(\ell) \in M_-$ if $\ell = 0$ and $i = j$).

Step 2. Definition of $d_N$ and proof that $d_N = d_N^*$. We define $d_N := \Pi_N d(\varphi)$ where $\Pi_N$ is the projector defined in (9.3.9) and

$$d(\varphi) = (d^j_i(\varphi))_{i,j\in\mathbb{F}}\,, \quad d^j_i(\varphi) = \sum_{\ell\in\mathbb{Z}^{|\mathbb{S}|}} \widehat{d}^j_i(\ell) e^{\mathrm{i}\ell\cdot\varphi}\,, \quad \widehat{d}^j_i(\ell) = T^{-1}_{ij\ell}\big(J[\widehat{\rho}_1]^j_i(\ell)\,, \tag{9.3.17}$$

is the unique solution of (9.2.5). Note that, taking the adjoint equation of (9.2.5), multiplying by $J$ on the left and the right, since $V_0$ and $D_0$ are self-adjoint, the operators $D_0$ and $J$ commute, $J^* = -J$, and using that $\rho_1$ is self-adjoint, we obtain

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi d^* + D_0 d^* + J d^* J D_0 = J\rho_1\,.$$

By uniqueness $d^* = d$ is the unique solution of (9.2.5). Then also $d_N^* = d_N$.

Step 3. Proof of (9.3.11). By (9.3.15) (and (3.3.44), (4.1.13)) we have

$$|d_N|_{s+1} \lesssim N^{\tau+1}|\rho_1|_s\,. \tag{9.3.18}$$

We claim also the estimate

$$|d_N|_{\mathrm{lip},s+1} \lesssim N^{2\tau+1}|\rho_1|_{\mathrm{Lip},s} \tag{9.3.19}$$

which implies, together with (9.3.18),

$$|d_N|_{\mathrm{Lip},s+1} \lesssim N^{2\tau+1}|\rho_1|_{\mathrm{Lip},s}\,. \tag{9.3.20}$$

To prove (9.3.19) notice that, by (9.3.16),

$$\begin{aligned}\|T^{-1}_{ij\ell}(\lambda_1) - T^{-1}_{ij\ell}(\lambda_2)\| &= \|T^{-1}_{ij\ell}(\lambda_1)\big(T_{ij\ell}(\lambda_2) - T_{i,j,\ell}(\lambda_1)\big)T^{-1}_{ij\ell}(\lambda_2)\big)\| \\ &\lesssim \langle\ell\rangle^{2\tau}\|T_{ij\ell}(\lambda_2) - T_{ij\ell}(\lambda_1)\| \lesssim \langle\ell\rangle^{2\tau}\varepsilon^2|\lambda_2 - \lambda_1|\end{aligned} \tag{9.3.21}$$

recalling the definition of $T_{ij\ell}$ in (9.3.3), and (8.1.7). By (9.3.16) and (9.3.21) we estimate the lip seminorm of $d^j_i(\ell) = T^{-1}_{ij\ell}\big([\widehat{\rho}_1]^j_j(\ell)\big)$ as

$$\|d^j_i(\ell)\|_{\mathrm{lip}} \lesssim \langle\ell\rangle^{2\tau}\varepsilon^2\|[\widehat{\rho}_1]^j_j(\ell)\| + \langle\ell\rangle^{\tau}\|[\widehat{\rho}_1]^j_j(\ell)\|_{\mathrm{lip}}$$

and (9.3.19) follows. The estimates (9.3.20), (9.3.10) and (8.2.1)-(i), finally imply

$$|d_N|_{\mathrm{Lip},s+1} \le C N^{2\tau+1} |\rho_1|_{\mathrm{Lip},s} \ll \delta_1^{-\frac{1}{40}} |\rho_1|_{\mathrm{Lip},s}$$

which gives the second inequality in (9.3.11). By (9.2.9) we deduce

$$|d_N|_{\mathrm{Lip},s_1+1} \ll \delta_1^{-\frac{1}{40}} \delta_1^{\frac{9}{10}} = \delta_1^{\frac{7}{8}}$$

which is the first inequality in (9.3.11).

Step 4. Proof of (9.3.12). By the definition of $d_N$ we have

$$J\bar{\omega}_\varepsilon \cdot \partial_\varphi d_N + D_0 d_N + J d_N J D_0 - J\rho_1 = -\Pi_N^\perp J \rho_1 \,. \tag{9.3.22}$$

Then, recalling (3.3.44),

$$\begin{aligned} |J\bar{\omega}_\varepsilon \cdot \partial_\varphi d_N + D_0 d_N + J d_N J D_0 - J\rho_1|_{\mathrm{Lip},s_1} &= |\Pi_N^\perp J \rho_1|_{\mathrm{Lip},s_1} \\ &\overset{(4.1.13),(9.2.9)}{\lesssim_{s_2}} N^{-(s_2-s_1)} \delta_1^{-\frac{11}{10}} \\ &\overset{(9.3.10)}{\lesssim_{s_2}} \delta_1^3 \delta_1^{-\frac{11}{10}} \ll \delta_1^{7/4} \,. \end{aligned}$$

Step 5. Proof of (9.3.13)-(9.3.14). We denote by $T'_{ij\ell}$ the linear operator as in (9.3.3) associated to $A'_0$, i.e. with $\mu'_j(\varepsilon,\lambda)$ instead of $\mu_j(\varepsilon,\lambda)$. Arguing as for (9.3.21) we get that, if $\lambda \in \Lambda^1(\varepsilon;1,A_0) \cap \Lambda^1(\varepsilon;1,A'_0)$, then

$$\left\| T_{ij\ell}^{-1} - (T'_{ij\ell})^{-1} \right\| \le C \langle \ell \rangle^{2\tau} \max_{j\in\mathbb{F}} \left| \mu_j(\varepsilon,\lambda) - \mu'_j(\varepsilon,\lambda) \right| . \tag{9.3.23}$$

Let us denote $d_N := d_{N,D_0,\rho}$ and $d'_N := d_{N,D'_0,\rho'}$ where we highlight the dependence of $d_N$ with respect to $D_0$ and $\rho$. Using that $d_N$ depends linearly on $\rho$, actually only on $\rho_1$, see (9.3.17), we get

$$\begin{aligned} |d_{N,D_0,\rho} - d_{N,D'_0,\rho'}|_{s_1} &\le |d_{N,D_0,\rho} - d_{N,D'_0,\rho}|_{s_1} + |d_{N,D'_0,\rho-\rho'}|_{s_1} \\ &\overset{(9.3.23),(9.3.18)}{\lesssim} N^{2\tau} |\rho_1|_{s_1} \max_{j\in\mathbb{F}} |\mu_j(\varepsilon,\lambda) - \mu'_j(\varepsilon,\lambda)| + N^{\tau+1} |\rho_1 - \rho'_1|_{s_1} \end{aligned}$$

and therefore, by (9.2.9), (9.3.10), (8.2.1)-(i), $|\rho_1|_{+,s} \simeq |\rho_1|_s$ we conclude that

$$\begin{aligned} |d_{N,D_0,\rho} - d_{N,D'_0,\rho'}|_{s_1} &\le \delta_1^{-\frac{1}{40}} \delta_1^{\frac{9}{10}} |D_0 - D'_0|_{s_1} + \delta_1^{-\frac{1}{40}} |\rho_1 - \rho'_1|_{s_1} \\ &\le \delta_1^{\frac{4}{5}} |D_0 - D'_0|_{s_1} + \delta_1^{-\frac{1}{40}} |\rho_1 - \rho'_1|_{s_1} \end{aligned}$$

proving (9.3.13). Finally (9.3.14) is an immediate consequence of (9.3.22). ■

We now prove *measure estimates* for the sets $\Lambda^1(\varepsilon;\eta,A_0)$ defined in (9.3.6).

**Lemma 9.3.3.** *$|[\Lambda^1(\varepsilon;\eta,A_0)]^c \cap \widetilde{\Lambda}| \le \varepsilon$ for all $1/2 \le \eta \le 1$.*

PROOF. For $\ell \in \mathbb{Z}^{|\mathbb{S}|}$, $\eta \in [1/2,1]$, we define the set

$$\Lambda^1_\ell(\varepsilon;\eta,A_0) := \Big\{\lambda \in \widetilde{\Lambda} \, : \, |\bar\omega_\varepsilon \cdot \ell \pm \mu_j(\varepsilon,\lambda) \pm \mu_i(\varepsilon,\lambda)| \ge \frac{\gamma_1}{2\eta\langle \ell\rangle^\tau}\,, \quad \forall (i,j) \in \mathbb{F}\times\mathbb{F} \text{ with } |\ell| + |i-j| \neq 0\Big\} \tag{9.3.24}$$

where $\gamma_1 = \gamma_0/2$ and $\tau > \tau_1$, see (9.3.7). By the unperturbed second order Melnikov non-resonance conditions (1.2.16)-(1.2.17), the Diophantine property (1.2.29), since $\bar\omega_\varepsilon = \bar\mu + \varepsilon^2\zeta$ (see (1.2.25)) and $\mu_j(\varepsilon,\lambda) = \mu_j + O(\varepsilon^2)$ (see (8.1.4)), we have

$$\text{if } \varepsilon^2 \langle \ell\rangle^{\tau_0+1}\gamma_0^{-1} \le c \text{ small enough} \quad \Rightarrow \quad \Lambda^1_\ell(\varepsilon;1/2,A_0) = \widetilde{\Lambda}\,.$$

Hence, for $\varepsilon$ small enough,

$$\Lambda^1(\varepsilon;\eta,A_0) = \bigcap_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} \Lambda^1_\ell(\varepsilon;\eta,A_0) = \bigcap_{|\ell| > c_1\varepsilon^{-2/\tau_0}} \Lambda^1_\ell(\varepsilon;\eta,A_0) \tag{9.3.25}$$

where $c_1 := (c\gamma_0)^{1/\tau_0}$. Now, using (8.1.5)-(8.1.6), we deduce that the complementary set $[\Lambda^1_\ell(\varepsilon;\eta,A_0)]^c$ is included in the union of $4\mathtt{f}^2$ intervals of length $\dfrac{\gamma_1}{\eta\langle\ell\rangle^\tau c_2\varepsilon^2}$ where $\mathtt{f}$ denotes the cardinality of $\mathbb{F}$. Hence its measure satisfies

$$|[\Lambda^1_\ell(\varepsilon;\eta,A_0)]^c \cap \widetilde{\Lambda}| \le \frac{4\gamma_1\mathtt{f}^2}{\eta\langle\ell\rangle^\tau c_2\varepsilon^2}\,,$$

and, for any $L > 0$, $1/2 \le \eta \le 1$, $\tau > |\mathbb{S}|$, we have

$$\sum_{|\ell| \ge L} \big|[\Lambda^1_\ell(\varepsilon;\eta,A_0)]^c\big| \le \sum_{|\ell|\ge L} \frac{4\gamma_1\mathtt{f}^2}{\eta\langle\ell\rangle^\tau c_2\varepsilon^2} \le \frac{C}{L^{\tau-|\mathbb{S}|}\varepsilon^2}\,. \tag{9.3.26}$$

The lemma follows by (9.3.25), (9.3.26) with $L = c_1\varepsilon^{-2/\tau_0}$ and (9.3.7). ∎

**Remark 9.3.4.** *The measure of the set $[\Lambda^1(\varepsilon;\eta,A_0)]^c$ can be made smaller than $\varepsilon^p$ as $\varepsilon$ tends to 0, for any $p$, if we take the exponent $\tau$ large enough. This is analogous to the situation described in remark 4.8.17.*

**Lemma 9.3.5.** *Assume that $|A_0' - A_0|_{+,s_1} \le \delta \le \varepsilon^3$ on $\widetilde{\Lambda} \cap \widetilde{\Lambda}'$. Then, for $\eta \in [(1/2)+\sqrt{\delta},1]$, we have*

$$|\widetilde{\Lambda}' \cap [\Lambda^1(\varepsilon;\eta,A_0')]^c \cap \Lambda^1(\varepsilon;\eta-\sqrt{\delta},A_0)| \le \delta^{\frac{1}{12}}\,. \tag{9.3.27}$$

PROOF. We first prove the estimate

$$\big|[\Lambda^1(\varepsilon;\eta,A_0')]^c \bigcap \Lambda^1(\varepsilon;\eta-\sqrt{\delta},A_0)\big| \lesssim \min(\varepsilon^{-2}\delta^{\frac{1}{2}-\frac{|\mathbb{S}|}{2\tau}},\varepsilon)\,. \tag{9.3.28}$$

Denoting $\Lambda^1_\ell(\varepsilon;\eta,A_0')$ the set (9.3.24) associated to $A_0'$, we claim that there exists $c(s_0)>0$ such that

$$[\Lambda^1_\ell(\varepsilon;\eta,A_0')]^c \cap \Lambda^1(\varepsilon;\eta-\sqrt{\delta},A_0)=\emptyset\,, \quad \text{if } \langle\ell\rangle^\tau \le c(s_0)\gamma_1/(\sqrt{\delta})\,. \tag{9.3.29}$$

Indeed, denoting by $\mu_j'(\varepsilon,\lambda)\in\mathbb{R}$ the eigenvalues of $D_0'=\Pi_{\mathbb{F}}A_0'|_{H_{\mathbb{F}}}$, since $\|A_0'-A_0\|_0 \le C(s_0)|A_0'-A_0|_{+,s_0}\le C(s_0)\delta$ (see (3.3.27)), we have

$$|\mu_j'(\varepsilon,\lambda)-\mu_j(\varepsilon,\lambda)|\le C(s_0)\delta\,, \quad \forall j\in\mathbb{F}\,. \tag{9.3.30}$$

If $\lambda\notin\Lambda^1_\ell(\varepsilon;\eta,A_0')$ then, recalling (9.3.24), there exist $i,j\in\mathbb{F}$ (with $i\neq j$ if $\ell=0$), signs $\epsilon_i=\pm1$ and $\epsilon_j=\pm1$ such that

$$|\bar\omega_\varepsilon\cdot\ell+\epsilon_i\mu_i'(\varepsilon,\lambda)+\epsilon_j\mu_j'(\varepsilon,\lambda)|<\frac{\gamma_1}{2\eta\langle\ell\rangle^\tau}$$

and so, by (9.3.30),

$$|\bar\omega_\varepsilon\cdot\ell+\epsilon_i\mu_i(\varepsilon,\lambda)+\epsilon_j\mu_j(\varepsilon,\lambda)|<\frac{\gamma_1}{2\eta\langle\ell\rangle^\tau}+2C(s_0)\delta<\frac{\gamma_1}{2(\eta-\sqrt{\delta})\langle\ell\rangle^\tau}\,,$$

for all $(1/2)+\sqrt{\delta}\le\eta\le1$, $\langle\ell\rangle^\tau\le c(s_0)\gamma_1/\sqrt{\delta}$, for some $c(s_0)$ small enough, proving (9.3.29). Hence

$$\begin{aligned}[\Lambda^1(\varepsilon;\eta,A_0')]^c\bigcap\Lambda^1(\varepsilon;\eta-\sqrt{\delta},A_0) &\subset \bigcup_{\ell\in\mathbb{Z}^{|\mathbb{S}|}}[\Lambda^1_\ell(\varepsilon;\eta,A_0')]^c\cap\Lambda^1(\varepsilon;\eta-\sqrt{\delta},A_0)\\ &\overset{(9.3.29)}{\subset} \bigcup_{|\ell|\ge(\frac{c(s_0)\gamma_1}{\sqrt{\delta}})^{1/\tau}}[\Lambda^1_\ell(\varepsilon;\eta,A_0')]^c\end{aligned}$$

which implies (9.3.28) by (9.3.26) with $L=(\gamma_1/\sqrt{\delta})^{1/\tau}$ (applied to $A_0'$) and Lemma 9.3.3.

Finally, by (9.3.7) and since $\tau_1\ge|\mathbb{S}|$, we have $|\mathbb{S}|/2\tau<1/4$ and so

$$\min(\varepsilon^{-2}\delta^{\frac{1}{2}-\frac{|\mathbb{S}|}{2\tau}},\varepsilon)\le\min(\varepsilon^{-2}\delta^{\frac14},\varepsilon)\le\delta^{\frac{1}{12}} \tag{9.3.31}$$

where the last inequality follows distinguishing the cases $\varepsilon\le\delta^{\frac{1}{12}}$ and $\varepsilon>\delta^{\frac{1}{12}}$. By (9.3.28) and (9.3.31), the estimate (9.3.27) is proved. ∎

**Step 2: approximate solution of the homological equation** (9.2.6)**.** We decompose $a\in\mathcal{L}(H_{\mathbb{F}},H_{\mathbb{G}})$ as $a=(a^j)_{j\in\mathbb{F}}$ and $\rho_2\in\mathcal{L}(H_{\mathbb{F}},H_{\mathbb{G}})$ as $\rho_2=(\rho_2^j)_{j\in\mathbb{F}}$ where $a^j:=a_{|H_j}\in$

$\mathcal{L}(H_j, H_{\mathbb{G}})$ and $\rho_2^j := (\rho_2)_{|H_j} \in \mathcal{L}(H_j, H_{\mathbb{G}})$. Recalling the form of $D_0$ in (8.1.3), the equation (9.2.6) is equivalent to

$$T_j(a^j) = J\rho_2^j \,, \quad \forall j \in \mathbb{F} \,, \tag{9.3.32}$$

where we define $T_j$ as the linear operator which maps $a^j : \mathbb{T}^{|\mathbb{S}|} \to \mathcal{L}(H_j, H_{\mathbb{G}})$ to

$$T_j(a^j) := J\bar{\omega}_\varepsilon \cdot \partial_\varphi a^j - JV_0 J a^j + \mu_j(\varepsilon, \lambda) J a^j J \,. \tag{9.3.33}$$

We shall consider $T_j$ as an unbounded linear operator from the space $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H_{\mathbb{G}}))$ to itself.

In the sequel $i_{\mathbb{G}}$ denotes the injection

$$i_{\mathbb{G}} : H_{\mathbb{G}} \hookrightarrow H \,. \tag{9.3.34}$$

Before applying the multiscale Proposition 4.1.5 we need to extend the linear operator $T_j$ defined in (9.3.33), which acts on $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H_{\mathbb{G}}))$, to a linear operator $T_j^\sharp$ acting on the *whole* space $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H)) \simeq L^2(\mathbb{T}^{|\mathbb{S}|}, H \times H)$ (by (3.2.6)) and satisfying the properties of Definition 4.1.2.

We define, for any $a^j \in L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H))$, the operator

$$\begin{aligned} &T_j^\sharp(a^j) := \\ &J\bar{\omega}_\varepsilon \cdot \partial_\varphi a^j + \frac{D_V}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}\cup\mathbb{F}} a^j + \frac{\mathfrak{c}}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}} a^j - J i_{\mathbb{G}} V_0 \Pi_{\mathbb{G}} J a^j + \mu_j(\varepsilon,\lambda) J \Pi_{\mathbb{G}} a^j J \end{aligned} \tag{9.3.35}$$

where $\Pi_{\mathbb{S}\cup\mathbb{F}}, \Pi_{\mathbb{G}}, \Pi_{\mathbb{S}}$ are the $L^2$-orthogonal projectors on the subspaces $H_{\mathbb{S}\cup\mathbb{F}}, H_{\mathbb{G}}, H_{\mathbb{S}}$ and $\mathfrak{c} > 0$ is a positive constant that we fix according to (4.1.5). Clearly $T_j^\sharp$ is an extension of $T_j$, i.e.

$$T_j^\sharp(a^j) = T_j(a^j) \,, \quad \forall a^j \in L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H_{\mathbb{G}})) \subset L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H)) \,.$$

Recalling the decomposition (3.2.15)-(3.2.16), i.e.

$$\begin{aligned} \mathcal{L}(H_j, H) &= \mathcal{L}(H_j, H_{\mathbb{S}\cup\mathbb{F}}) \oplus \mathcal{L}(H_j, H_{\mathbb{G}}) \,, \\ a^j(\varphi) &= \Pi_{\mathbb{S}\cup\mathbb{F}} a^j(\varphi) + \Pi_{\mathbb{G}} a^j(\varphi) \,, \end{aligned} \tag{9.3.36}$$

we may write the operator $T_j^\sharp$ in (9.3.35) as

$$T_j^\sharp(a^j) = J\bar{\omega}_\varepsilon \cdot \partial_\varphi a^j + \frac{D_V}{1+\varepsilon^2\lambda} a^j + \mu_j(\varepsilon,\lambda) J\Pi_{\mathbb{G}} a^j J + \frac{\mathfrak{c}}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}} a^j + \mathcal{R}_j a^j \tag{9.3.37}$$

where, using that $J$ commutes with $D_V$ and $\Pi_{\mathbb{G}}$, $J^2 = -\mathrm{Id}$,

$$\mathcal{R}_j a^j := J\Big(\frac{D_V}{1+\varepsilon^2\lambda} - i_{\mathbb{G}} V_0\Big)\Pi_{\mathbb{G}} J a^j = -JR_0 \Pi_{\mathbb{G}} J a^j \tag{9.3.38}$$

and we used (8.1.1)-(8.1.2) to obtain the last equality. Notice that, according to the decomposition (9.3.36), the operator $T_j^\sharp$ is represented by the block-diagonal matrix of operators

$$T_j^\sharp = \begin{pmatrix} J\bar{\omega}_\varepsilon \cdot \partial_\varphi + \dfrac{D_V}{1+\varepsilon^2\lambda} + \dfrac{\mathfrak{c}}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}} & 0 \\ 0 & T_j \end{pmatrix}, \tag{9.3.39}$$

and, according to the further splitting

$$\mathcal{L}(H_j, H) = \mathcal{L}(H_j, H_{\mathbb{S}}) \oplus \mathcal{L}(H_j, H_{\mathbb{F}}) \oplus \mathcal{L}(H_j, H_{\mathbb{G}}),$$
$$a^j(\varphi) = \Pi_{\mathbb{S}} a^j(\varphi) + \Pi_{\mathbb{F}} a^j(\varphi) + \Pi_{\mathbb{G}} a^j(\varphi),$$

recalling (9.3.35), by the matrix of operators

$$T_j^\sharp = J\bar{\omega}_\varepsilon \cdot \partial_\varphi + \begin{pmatrix} \dfrac{D_V}{1+\varepsilon^2\lambda} + \dfrac{\mathfrak{c}\,\mathrm{Id}}{1+\varepsilon^2\lambda} & 0 & 0 \\ 0 & \dfrac{D_V}{1+\varepsilon^2\lambda} & 0 \\ 0 & 0 & -JV_0J + \mu_j(\varepsilon,\lambda)\mathcal{J} \end{pmatrix} \tag{9.3.40}$$

where $\mathcal{J}$ is defined in (4.1.8). Note also that

$$T_j \Pi_{\mathbb{G}} = \Pi_{\mathbb{G}} T_j^\sharp. \tag{9.3.41}$$

Now, given $g^j := J\rho_2^j : \mathbb{T}^{|\mathbb{S}|} \to \mathcal{L}(H_j, H_{\mathbb{G}})$ we define

$$g_j^\sharp : \mathbb{T}^{|\mathbb{S}|} \to \mathcal{L}(H_j, H) \simeq H \times H, \quad g_j^\sharp(\varphi) := i_{\mathbb{G}} g^j(\varphi), \tag{9.3.42}$$

(recall (3.2.6), (9.3.34)) and we look for an (approximate) solution $a_j^\sharp : \mathbb{T}^{|\mathbb{S}|} \to \mathcal{L}(H_j, H) \simeq H \times H$ of the equation

$$T_j^\sharp(a_j^\sharp) = g_j^\sharp. \tag{9.3.43}$$

As already said, $T_j^\sharp$ is an (unbounded) operator on $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j,, H))$ and an extension of $T_j$. It is important to notice that, since the subspaces $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H_{\mathbb{S}\cup\mathbb{F}}))$ and $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H_{\mathbb{G}}))$ are invariant under the operator $T_j^\sharp$ (see e.g. (9.3.39)), a solution of (9.3.43) satisfies $T_j(\Pi_{\mathbb{G}} a_j^\sharp) = g_j^\sharp$, by (9.3.41), and therefore $a^j := \Pi_{\mathbb{G}} a_j^\sharp$ solves equation (9.3.32) for the non extended operator $T_j$. Actually in Lemma 9.3.9 we find an approximate solution of the equation (9.3.32).

In the sequel $T_j^\sharp$ is regarded as an operator acting on $L^2(\mathbb{T}^{|\mathbb{S}|}; H \times H)$: in fact, using the basis $((\Psi_j, 0), (0, \Psi_j))$ of $H_j$, see (3.1.9), we have the identification (3.2.6) and $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H)) \simeq L^2(\mathbb{T}^{|\mathbb{S}|}; H \times H)$. We will apply the multiscale Proposition 4.1.5 (in case-($ii$) in (4.1.1)) to the operator

$$\mathcal{L}_{r,\mu} = (1 + \varepsilon^2\lambda) T_j^\sharp.$$

Actually, recalling (9.3.37), $\Pi_{\mathbb{G}} = \Pi^{\perp}_{\mathbb{F}\cup\mathbb{S}}$, and the definition of $\mathcal{J}$ in (4.1.8), we have that

$$\begin{aligned} \mathcal{L}_{r,\mu} &= (1+\varepsilon^2\lambda)T_j^\sharp = J\omega\cdot\partial_\varphi + X_{r,\mu}\,, \\ X_{r,\mu} &= D_V + \mathfrak{c}\Pi_{\mathbb{S}} + \mu\mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} + r\,, \end{aligned} \tag{9.3.44}$$

has the form (4.1.9) with $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$ and $X_{r,\mu}$ as in (4.1.7) with

$$\mu = (1+\varepsilon^2\lambda)\mu_j(\varepsilon,\lambda)\,, \qquad r = (1+\varepsilon^2\lambda)\mathcal{R}_j \overset{(9.3.38)}{=} -(1+\varepsilon^2\lambda)JR_0\Pi_{\mathbb{G}}J\,. \tag{9.3.45}$$

We now prove that $X_{r,\mu}$ in (9.3.44) satisfies the properties stated in Definition 4.1.2, beginning with its self-adjointness.

**Lemma 9.3.6. (Self-adjointness)** *The operator $X_{r,\mu}$ defined in* (9.3.44)-(9.3.45) *is self-adjoint with respect to the scalar product $\langle\ ,\ \rangle_0$ in* (3.2.22).

Proof. The self-adjointness of $D_V$ and $\Pi_{\mathbb{S}}$ with respect to $\langle\ ,\ \rangle_0$ directly follows by the fact that $D_V$ and $\Pi_{\mathbb{S}}$ are $L^2$- self-adjoint. Now let $a := a(\varphi)$, $b := b(\varphi)$ belong to $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j,, H))$. We obtain, using that $\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}}$ is $L^2$-self-adjoint, that

$$\begin{aligned} \langle a, \mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} b\rangle_0 &\overset{(4.1.8)}{=} \int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}((J\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}}bJ)^*a)d\varphi \\ &= \int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}(Jb^*\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}}Ja)d\varphi = \int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}(Jb^*J\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}}a)d\varphi \end{aligned}$$

because $J$ and $\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}}$ commute. Thus, using that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, we deduce that

$$\langle a, \mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} b\rangle_0 = \int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}(b^*J\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}}aJ)d\varphi \overset{(4.1.8),(3.2.22)}{=} \langle \mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}} a, b\rangle_0\,.$$

Hence the operator $\mathcal{J}\Pi^{\perp}_{\mathbb{S}\cup\mathbb{F}}$ is self-adjoint with respect to $\langle\ ,\ \rangle_0$.

It remains to prove that the operator $\mathcal{R}_j$ in (9.3.38) is self-adjoint with respect to $\langle\ ,\ \rangle_0$. Since $R_0$ is self-adjoint and $R_0(H_{\mathbb{G}}) \subset H_{\mathbb{G}}$, we deduce that $B := R_0\Pi_{\mathbb{G}}$ is $L^2$-self-adjoint, i.e. it satisfies $B^* = B$, and, recalling (9.3.38), we obtain

$$\langle a, \mathcal{R}_j b\rangle_0 = -\int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}((JBJb)^*a)d\varphi = -\int_{\mathbb{T}^{|\mathbb{S}|}} \mathrm{Tr}(b^*JBJa)d\varphi = \langle \mathcal{R}_j a, b\rangle_0\,.$$

This completes the proof of the lemma. ■

**Lemma 9.3.7. (Off-diagonal decay)** *The operator $\mathcal{R}_j$ defined in* (9.3.38) *satisfies*

$$|\mathcal{R}_j|_{\mathrm{Lip},s} \le C(s)|R_0|_{\mathrm{Lip},s}\,, \qquad |\mathcal{R}_j|_{\mathrm{Lip},+,s} \le C(s)|R_0|_{\mathrm{Lip},+,s}\,, \tag{9.3.46}$$

*in particular $|\mathcal{R}_j|_{\mathrm{Lip},+,s_1} \le C_1'\varepsilon^2$ for some constant $C_1'$ depending only on $C_1$.*

Proof. We identify $a(\varphi) \in \mathcal{L}(H_j, H)$ with the vector $(a^{(1)}(\varphi), a^{(2)}(\varphi), a^{(3)}(\varphi), a^{(4)}(\varphi))$ in $H \times H$ as in (3.2.7). Then, using (3.2.13), we have

$$\mathcal{R}_j a = \big( - JR_0\Pi_{\mathbb{G}}(a_2, -a_1), -JR_0\Pi_{\mathbb{G}}(a_3, -a_4)\big) .$$

Hence

$$|\mathcal{R}_j|_{\mathrm{Lip},s} \sim |R_0\Pi_{\mathbb{G}}|_{\mathrm{Lip},s} \quad \text{and} \quad |\mathcal{R}_j|_{\mathrm{Lip},+,s} \sim |R_0\Pi_{\mathbb{G}}|_{\mathrm{Lip},+,s} .$$

Lemmata 3.3.5 and 3.3.8 imply that

$$|R_0\Pi_{\mathbb{G}}|_{\mathrm{Lip},s} \overset{(3.3.21)}{\lesssim_s} |\Pi_{\mathbb{G}}|_{\mathrm{Lip},s}|R_0|_{\mathrm{Lip},s} \overset{(3.3.33)}{\lesssim_s} |R_0|_{\mathrm{Lip},s}$$
$$|R_0\Pi_{\mathbb{G}}|_{\mathrm{Lip},+,s} \overset{(3.3.24)}{\lesssim_s} |\Pi_{\mathbb{G}}|_{\mathrm{Lip},s+\frac{1}{2}}|R_0|_{\mathrm{Lip},+,s} \overset{(3.3.33)}{\lesssim_s} |R_0|_{\mathrm{Lip},+,s}$$

proving (9.3.46). Since $|R_0|_{\mathrm{Lip},+,s_1} \leq C_1\varepsilon^2$ by Definition 8.1.1, the second estimate in (9.3.46) implies $|\mathcal{R}_j|_{\mathrm{Lip},+,s_1} \leq C_1'\varepsilon^2$. ■

**Lemma 9.3.8. (Sign condition)** *The operator $X_{r,\mu}$ defined in* (9.3.44) *satisfies, for some $c > 0$ depending on the constant $c_1$ in* (8.1.8),

$$\mathfrak{d}_\lambda\Big(\frac{X_{r,\mu}}{1+\varepsilon^2\lambda}\Big) \leq -c\,\varepsilon^2 \mathrm{Id} .$$

Proof. According to (9.3.40) the operator $\dfrac{X_{r,\mu}}{1+\varepsilon^2\lambda}$ is represented by the matrix of operators

$$\begin{pmatrix} \dfrac{D_V + \mathfrak{c}\,\mathrm{Id}}{1+\varepsilon^2\lambda} & 0 & 0 \\ 0 & \dfrac{D_V}{1+\varepsilon^2\lambda} & 0 \\ 0 & 0 & -JV_0J + \mu_j(\varepsilon,\lambda)\,\mathcal{J} \end{pmatrix}$$

where $\mathcal{J}a = JaJ$. We clearly have

$$\partial_\lambda \frac{D_V + \mathfrak{c}\,\mathrm{Id}}{1+\varepsilon^2\lambda} = -\frac{\varepsilon^2(D_V + \mathfrak{c}\,\mathrm{Id})}{(1+\varepsilon^2\lambda)^2} \leq -\frac{\varepsilon^2 D_V}{(1+\varepsilon^2\lambda)^2} \leq -c\varepsilon^2\mathrm{Id} \tag{9.3.47}$$

for some $c > 0$. Then it is sufficient to prove that, for all $a \in L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H_{\mathbb{G}}))$, and for all $\lambda_1, \lambda_2 \in \widetilde{\Lambda}$ with $\lambda_1 \neq \lambda_2$, using the notation (1.6.3),

$$\Big\langle -J\frac{\Delta V_0}{\Delta\lambda}Ja + \frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\,JaJ, a\Big\rangle_0 \leq -c\varepsilon^2\|a\|_0^2 ,$$

for the scalar product $\langle\ ,\ \rangle_0$ defined in (3.2.22). We have

$$\begin{aligned}\Big\langle -J\frac{\Delta V_0}{\Delta\lambda}Ja + \frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\,JaJ, a\Big\rangle_0 &= \Big\langle \frac{\Delta V_0}{\Delta\lambda}Ja, Ja\Big\rangle_0 - \frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\,\langle aJ, Ja\rangle_0 \\ &\leq \Big\langle \frac{\Delta V_0}{\Delta\lambda}Ja, Ja\Big\rangle_0 + \Big|\frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\Big|\|aJ\|_0\|Ja\|_0 \\ &= \Big\langle \frac{\Delta V_0}{\Delta\lambda}Ja, Ja\Big\rangle_0 + \Big|\frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\Big|\|Ja\|_0^2 \qquad (9.3.48)\end{aligned}$$

noting that $\|aJ\|_0 = \|Ja\|_0 = \|a\|_0$ from the definition (3.2.22) of the scalar product $\langle\ ,\ \rangle_0$. Let

$$(Ja)^{(1)}(\varphi) = Ja(\varphi)(\Psi_j, 0)\,, \quad (Ja)^{(2)}(\varphi) = Ja(\varphi)(0, \Psi_j)\,, \quad (Ja)^{(i)} \in L^2(\mathbb{T}^{|\mathbb{S}|}, H_{\mathbb{G}})\,.$$

Using the notations $\langle\ ,\ \rangle_0$ and $\|\ \|_0$ for the scalar product and its associated norm both in $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H_{\mathbb{G}}))$ and in $L^2(\mathbb{T}^{|\mathbb{S}|}; H_{\mathbb{G}})$, we obtain by (9.3.48)

$$\begin{aligned}&\Big\langle -J\frac{\Delta V_0}{\Delta\lambda}Ja + \frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\,JaJ, a\Big\rangle_0 \\ &\leq \sum_{i=1}^{2}\Big\langle \frac{\Delta V_0}{\Delta\lambda}(Ja)^{(i)}, (Ja)^{(i)}\Big\rangle_0 + \Big|\frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\Big|\|(Ja)^{(i)}\|_0^2\,.\end{aligned} \qquad (9.3.49)$$

Now the assumption (8.1.8) implies that for all $h \in L^2(\mathbb{T}^{|\mathbb{S}|}, H_{\mathbb{G}})$

$$\Big\langle \frac{\Delta V_0}{\Delta\lambda}h, h\Big\rangle_0 + \Big|\frac{\Delta\mu_j(\varepsilon,\lambda)}{\Delta\lambda}\Big|\|h\|_0^2 \leq -c_1\varepsilon^2\|h\|_0^2\,. \qquad (9.3.50)$$

The estimates (9.3.47), (9.3.49) and (9.3.50) imply the lemma. ■

By Lemmata 9.3.6, 9.3.7, 9.3.8 we apply the multiscale Proposition 4.1.5 to the operator $\mathcal{L}_{r,\mu} = (1+\varepsilon^2\lambda)T_j^\sharp$ in (9.3.44) where $r$ is given in (9.3.45). As a consequence there exist, for any $j \in \mathbb{F}$, closed subsets

$$\Lambda_j^2(\varepsilon;\eta,A_0)\,,\ 1/2 \leq \eta \leq 1\,,\ \text{satisfying Properties 1–3 of Proposition 9.1.1}\,, \qquad (9.3.51)$$

and $\bar N \in \mathbb{N}$, such that, for all $\forall N \geq \bar N$, for all $\lambda \in \Lambda_j^2(\varepsilon;1,A_0)$, there are operators $(T_j^\sharp)_N^{-1}$ defined in Proposition 4.1.5 as

$$\begin{aligned}&i)\ \text{the right inverse of } \Pi_N(T_j^\sharp)_{|\mathcal{H}_{2N}}\ \text{ if } \bar N \leq N < N(\varepsilon)\,, \\ &ii)\ \text{the inverse of } \Pi_N(T_j^\sharp)_{|\mathcal{H}_N}\ \text{ if } N \geq N(\varepsilon)\,,\end{aligned} \qquad (9.3.52)$$

where $\mathcal{H}_N$ are the finite dimensional subspaces defined in (4.1.10). By (4.1.20), (4.1.23), (9.3.45)-(9.3.46) and $|R_0|_{\mathrm{Lip},+,s_1} \leq C_1\varepsilon^2$ by Definition 8.1.1, the operators $(T_j^\sharp)_N^{-1}$ satisfy the following tame estimates: $\forall s \geq s_0$,

$$|(T_j^\sharp)_N^{-1}|_{\mathrm{Lip},s_1} \leq C(s_1) N^{2(\tau'+\varsigma s_1)+3} \tag{9.3.53}$$

$$|(T_j^\sharp)_N^{-1}|_{\mathrm{Lip},s} \leq C(s) N^{2(\tau'+\varsigma s_1)+3}\big(N^{\varsigma(s-s_1)} + |R_0|_{\mathrm{Lip},+,s}\big)\,. \tag{9.3.54}$$

To justify that the sets $\Lambda_j^2(\varepsilon;\eta,A_0)$ satisfy Property 3 of Proposition 9.1.1, we remark that if the operators $A_0$ and $A_0'$ satisfy $|R_0'-R_0|_{+,s_1} \leq \delta \leq \varepsilon^{5/2}$, then $\mu,\mu',r,r'$ defined in (9.3.45) satisfy

$$|r'-r|_{+,s_1} + |\mu'-\mu| \leq C\delta \leq \varepsilon^2\,,$$

and refer to (4.1.18) in Proposition 4.1.5 (note that the exponent $\alpha$ of Property 3 of Proposition 9.1.1 may be taken slightly smaller than in (4.1.18)).

We now obtain an approximate solution of the homological equation (9.3.32), finding an approximate solution of (9.3.43).

**Lemma 9.3.9. (Homological equation** (9.3.32)**)** *Let*

$$N \in \Big[\delta_1^{-\frac{3}{s_2-s_1}} - 1, \delta_1^{-\frac{3}{s_2-s_1}} + 1\Big]\,. \tag{9.3.55}$$

*Then, for all $j \in \mathbb{F}$, for all $\lambda \in \Lambda_j^2(\varepsilon;1,A_0)$ (set introduced in* (9.3.51)*), the function $a^j$ defined as*

$$a^j := \Pi_{\mathbb{G}} b_j^\sharp\,,\quad b_j^\sharp := (T_j^\sharp)_N^{-1} g_{j,N}^\sharp\,,\quad g_{j,N}^\sharp := \mathit{\Pi}_N g_j^\sharp\,,\quad g_j^\sharp := i_{\mathbb{G}} g^j\,,\ g^j = J(\rho_2)_{|H_j} \tag{9.3.56}$$

*(recall the notation* (9.3.42), (9.2.3)*), satisfies*

$$|a^j|_{\mathrm{Lip},s_1+1} \ll \delta_1^{\frac{7}{8}}\,, \tag{9.3.57}$$

$$|a^j|_{\mathrm{Lip},s+1} \lesssim_s \delta_1^{-\frac{1}{40}}\big(|\rho|_{\mathrm{Lip},s} + |R_0|_{\mathrm{Lip},+,s}\big) + \delta_1^{\frac12}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\,,\quad \forall s \geq s_2\,, \tag{9.3.58}$$

*and it is an approximate solution of the homological equation* (9.3.32) *in the sense that*

$$|T_j a^j - g^j|_{\mathrm{Lip},s_1+1} \ll \delta_1^{7/4}\,. \tag{9.3.59}$$

*At last, denoting by $a^j := a^j_{A_0,\rho}$ and $(a^j)' := a^j_{A_0',\rho'}$ the operators defined as above associated to $(A_0,\rho)$ and $(A_0',\rho')$ respectively, for all $\lambda \in \Lambda^2(\varepsilon;1,A_0) \cap \Lambda^2(\varepsilon;1,A_0')$, we have*

$$\big|a^j_{A_0,\rho} - a^j_{A_0',\rho'}\big|_{s_1+1} \ll \delta_1^{\frac45}|A_0-A_0'|_{+,s_1} + \delta_1^{-\frac{1}{40}}|\rho-\rho'|_{+,s_1} \tag{9.3.60}$$

*and*

$$\big|\big(T_{j,A_0}a^j_{A_0,\rho} - g^j_\rho\big) - \big(T_{j,A_0'}a^j_{A_0',\rho'} - g^j_{\rho'}\big)\big|_{s_1+\frac12} \ll \delta_1^{\frac45}|A_0-A_0'|_{+,s_1} + \delta_1^{-\frac{1}{40}}|\rho-\rho'|_{+,s_1}\,. \tag{9.3.61}$$

PROOF. The function $b_j^\sharp$ defined in (9.3.56) satisfies, by (9.3.52), the property

$$\begin{aligned} &i)\ \Pi_{2N} b_j^\sharp = b_j^\sharp \quad \text{if} \quad \bar N \le N < N(\varepsilon) \\ &ii)\ \Pi_N b_j^\sharp = b_j^\sharp \quad \text{if} \quad N \ge N(\varepsilon)\,, \end{aligned} \tag{9.3.62}$$

and it is a solution of $\Pi_N T_j^\sharp b_j^\sharp = \Pi_N g_j^\sharp$. Then, up to terms which are Fourier supported on high harmonics, it is an approximate solution of the equation (9.3.43), more precisely,

$$T_j^\sharp b_j^\sharp = \Pi_N T_j^\sharp b_j^\sharp + \Pi_N^\perp T_j^\sharp b_j^\sharp = \Pi_N g_j^\sharp + \Pi_N^\perp T_j^\sharp b_j^\sharp = g_j^\sharp + z_N \tag{9.3.63}$$

where

$$z_N := -\Pi_N^\perp g_j^\sharp + \Pi_N^\perp T_j^\sharp b_j^\sharp\,. \tag{9.3.64}$$

Applying $\Pi_{\mathbb{G}}$ to both sizes in (9.3.63), we deduce, by (9.3.41) and (9.3.56), that $a^j := \Pi_{\mathbb{G}} b_j^\sharp$ is an approximate solution of the homological equation (9.3.32), in the sense that

$$T_j a^j - J(\rho_2)_{|H_j} = \Pi_{\mathbb{G}} z_N\,, \qquad J(\rho_2)_{|H_j} = J\rho_2^j = g^j\,. \tag{9.3.65}$$

We now prove (9.3.57)-(9.3.59).

ESTIMATES OF $b_j^\sharp$. By (3.2.6) we identify the $\varphi$-dependent family of operators $g_j^\sharp = i_{\mathbb{G}} J(\rho_2)_{|H_j}$ in $L^2(\mathbb{T}^{|\mathbb{S}|}, \mathcal{L}(H_j, H))$ defined in (9.3.56) with a function of $L^2(\mathbb{T}^{|\mathbb{S}|}, H \times H)$. We have the equivalence of the norms

$$|g_j^\sharp|_s \sim_s \|g_j^\sharp\|_s\,. \tag{9.3.66}$$

In fact, let us define, for all $\varphi \in \mathbb{T}^{|\mathbb{S}|}$, the functions

$$g_j^{\sharp,1}(\varphi) := g_j^\sharp(\varphi)\big[(\Psi_j, 0)\big] \in H \quad \text{and} \quad g_j^{\sharp,2}(\varphi) := g_j^\sharp(\varphi)\big[(0, \Psi_j)\big] \in H\,.$$

We have to prove that $|g_j^\sharp|_s \sim_s \|g_j^{\sharp,1}\|_s + \|g_j^{\sharp,2}\|_s$. By Definition 3.3.4, $|g_j^\sharp|_s = |\widetilde{g_j^\sharp}|_s$, where the operator $\widetilde{g_j^\sharp}$ is defined on the whole $L^2(\mathbb{T}^{|\mathbb{S}|}, H)$ by

$$\widetilde{g_j^\sharp}\big[(h^{(1)}, h^{(2)})\big] := \sum_{i=1}^{2} g_j^{\sharp,i}\big(h^{(i)}, \Psi_j\big)_{L^2_x}\,.$$

Now, by Lemma 3.3.7,

$$|\widetilde{g_j^\sharp}|_s \le C(s)\big(\|g_j^{\sharp,1}\|_s + \|g_j^{\sharp,2}\|_s\big)\,.$$

Conversely

$$\|g_j^{\sharp,1}\|_s = \|\widetilde{g_j^\sharp}\big[(\Psi_j, 0)\big]\|_s \lesssim_s |\widetilde{g_j^\sharp}|_s \|\Psi_j\|_s \lesssim_s |\widetilde{g_j^\sharp}|_s$$

and similarly $\|g_j^{\sharp,2}\|_s \lesssim_s |\widetilde{g_j^{\sharp}}|_s$. This proves the norm equivalence (9.3.66).

We have $g^{\sharp}_{j,N} = \Pi_N g^{\sharp}_{j,N}$, and (4.1.13), (9.3.56), (9.3.53), (9.2.9), imply

$$\begin{aligned} |b_j^{\sharp}|_{\mathrm{Lip},s_1+1} &\leq N |b_j^{\sharp}|_{\mathrm{Lip},s_1} \\ &\lesssim_{s_1} N |(T_j^{\sharp})_N^{-1}|_{\mathrm{Lip},s_1} |g^{\sharp}_{j,N}|_{\mathrm{Lip},s_1} \\ &\lesssim_{s_1} N^{2(\tau'+\varsigma s_1+2)} |\rho|_{\mathrm{Lip},s_1} \lesssim_{s_1} N^Q \delta_1^{\frac{9}{10}} \end{aligned} \tag{9.3.67}$$

having set

$$Q := 2(\tau' + \varsigma s_1 + 2) \quad \text{with} \quad \varsigma = 1/10 \quad (\text{as in } (4.1.16))\,. \tag{9.3.68}$$

Similarly, for $s \geq s_1$, using (4.1.13), (9.3.56), the tame estimate (3.3.21), (9.3.53)-(9.3.54), (9.2.9), we get

$$\begin{aligned} |b_j^{\sharp}|_{\mathrm{Lip},s+1} &\leq N |b_j^{\sharp}|_{\mathrm{Lip},s} \\ &\lesssim_s N |(T_j^{\sharp})_N^{-1}|_{\mathrm{Lip},s} |g^{\sharp}_{j,N}|_{\mathrm{Lip},s_1} + N |(T_j^{\sharp})_N^{-1}|_{\mathrm{Lip},s_1} |g^{\sharp}_{j,N}|_{\mathrm{Lip},s} \\ &\lesssim_s N^Q \big( N^{\varsigma(s-s_1)} + |R_0|_{\mathrm{Lip},+,s} \big) \delta_1^{\frac{9}{10}} + N^Q |g^{\sharp}_{j,N}|_{\mathrm{Lip},s} \\ &\lesssim_s N^Q \big( N^{\varsigma(s-s_1)} \delta_1^{\frac{9}{10}} + |R_0|_{\mathrm{Lip},+,s} \delta_1^{\frac{9}{10}} + |\rho|_{\mathrm{Lip},s} \big)\,. \end{aligned} \tag{9.3.69}$$

Estimates of $z_N$ defined in (9.3.64). We first claim that

$$|\Pi_N^{\perp} T_j^{\sharp} \Pi_N|_{\mathrm{Lip},s} \leq |\Pi_N^{\perp} T_j^{\sharp} \Pi_{2N}|_{\mathrm{Lip},s} \lesssim_s N + |R_0|_{\mathrm{Lip},s}\,. \tag{9.3.70}$$

Indeed, recalling the expression (9.3.37) of $T_j^{\sharp}$, setting $\mathcal{J}(a^j) := J a^j J$, and writing $D_V = D_m + (D_V - D_m)$, we get

$$\begin{aligned} \Pi_N^{\perp} T_j^{\sharp} \Pi_{2N} = \Pi_N^{\perp} J \bar{\omega}_{\varepsilon} \cdot \partial_{\varphi} \Pi_{2N} + \frac{1}{1+\varepsilon^2\lambda} \big( \Pi_N^{\perp} D_m \Pi_{2N} + \Pi_N^{\perp} (D_V - D_m) \Pi_{2N} \big) \\ + \mu_j(\varepsilon,\lambda) \Pi_N^{\perp} \mathcal{J} \Pi_{\mathbb{G}} \Pi_{2N} + \Pi_N^{\perp} \mathcal{R}_j \Pi_{2N} + \Pi_N^{\perp} \frac{\mathfrak{c}}{1+\varepsilon^2\lambda} \Pi_{\mathbb{S}} \Pi_{2N}\,. \end{aligned} \tag{9.3.71}$$

In view of (9.3.71), Lemmas 3.3.8, 9.3.7 and Proposition 3.4.1 imply (9.3.70).

Then, by (9.3.64), (9.3.62), (9.3.56), (9.3.70), (3.3.20), $|R_0|_{\mathrm{Lip},s_1} \leq 1$, we obtain

$$\begin{aligned} |z_N|_{\mathrm{Lip},s} &\leq |g_j^{\sharp}|_{\mathrm{Lip},s} + |\Pi_N^{\perp} T_j^{\sharp} b_j^{\sharp}|_{\mathrm{Lip},s} \\ &\lesssim_s |\rho|_{\mathrm{Lip},s} + N |b_j^{\sharp}|_{\mathrm{Lip},s} + (N + |R_0|_{\mathrm{Lip},s}) |b_j^{\sharp}|_{\mathrm{Lip},s_1} \\ &\overset{(9.3.69),(9.3.67)}{\lesssim_s} N^Q \big( |\rho|_{\mathrm{Lip},s} + |R_0|_{\mathrm{Lip},+,s} \delta_1^{\frac{9}{10}} + N^{\varsigma(s-s_1)} \delta_1^{\frac{9}{10}} \big)\,. \end{aligned} \tag{9.3.72}$$

By (9.3.64) we have $\Pi_N z_N = 0$, and, using (9.3.72) and the assumption (9.2.9), we derive that

$$|z_N|_{\mathrm{Lip},s_1+1} \le N^{-(s_2-s_1-1)}|z_N|_{\mathrm{Lip},s_2} \lesssim_{s_2} N^{Q+1-(s_2-s_1)}\delta_1^{-\frac{11}{10}} + N^{Q+1-(1-\varsigma)(s_2-s_1)}\delta_1^{\frac{9}{10}} . \tag{9.3.73}$$

Proof of (9.3.57)-(9.3.59). By the choice of $N$ in (9.3.55), the condition (8.2.1)-($i$), and (9.3.68) we get

$$N^{Q+1} = o(\delta_1^{-\frac{1}{40}}) , \tag{9.3.74}$$

$$N^{Q+1-(s_2-s_1)}\delta_1^{-\frac{11}{10}} = o(\delta_1^{\frac{7}{4}}) , \ N^{Q+1-(1-\varsigma)(s_2-s_1)}\delta_1^{\frac{9}{10}} = o(\delta_1^{\frac{7}{4}}) , \tag{9.3.75}$$

$$\forall s \ge s_2 , \quad N^{Q+1+\varsigma(s-s_1)}\delta_1^{\frac{9}{10}} = N^{Q+1+\varsigma(s_2-s_1)}\delta_1^{\frac{9}{10}} N^{\varsigma(s-s_2)} \le \delta_1^{\frac{1}{2}}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}} . \tag{9.3.76}$$

Then, by (9.3.67), (9.3.74) and (9.3.73), (9.3.75), for $\delta_1$ small enough, we deduce the bounds

$$|b_j^\sharp|_{\mathrm{Lip},s_1+1} \ll \delta_1^{7/8}, \qquad |z_N|_{\mathrm{Lip},s_1+1} \ll \delta_1^{7/4} . \tag{9.3.77}$$

In addition, (9.3.69), (9.3.74), (9.3.76) imply the estimate in high norm

$$|b_j^\sharp|_{\mathrm{Lip},s+1} \le \delta_1^{-\frac{1}{40}}\big(|\rho|_{\mathrm{Lip},s} + |R_0|_{\mathrm{Lip},+,s}\big) + \delta_1^{\frac{1}{2}}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}} . \tag{9.3.78}$$

By Lemma 3.3.8, the function $a^j := \Pi_{\mathbb{G}} b_j^\sharp$ satisfies the same estimates as $b_j^\sharp$ in (9.3.77), (9.3.78). In particular (9.3.57)-(9.3.58) hold and, by (9.3.65), Lemma 3.3.8, (9.3.77), we get

$$|T_j a^j - g^j|_{\mathrm{Lip},s_1+1} = |\Pi_{\mathbb{G}} z_N|_{\mathrm{Lip},s_1+1} \ll \delta_1^{7/4}$$

proving (9.3.59).

**Proof of** (9.3.60)**-**(9.3.61)**.** We denote by $(a^j)'$, $b_j^{\sharp\prime}$ the functions obtained in (9.3.56) from $(A_0', \rho')$ instead of $A_0, \rho$, and, similarly, by $T_j^{\sharp\prime}$ the linear operator defined in (9.3.37) from $(A_0', \rho')$. Recall that to define $T_j^\sharp$ we applied Proposition 4.1.5 with $\mu$ and $r$ defined in (9.3.45). In particular, we have

$$|\mu - \mu'| \lesssim |A_0 - A_0'|_{+,s_1} , \quad |r - r'|_{+,s_1} \lesssim |R_0 - R_0'|_{+,s_1} = |A_0 - A_0'|_{+,s_1} .$$

Hence, by (4.1.21), (4.1.24) in Proposition 4.1.5, for any $\lambda \in \Lambda_j^2(\varepsilon; 1, A_0) \cap \Lambda_j^2(\varepsilon; 1, A_0')$, we have

$$\big|(T_j^\sharp)_N^{-1} - (T_j^{\sharp\prime})_N^{-1}\big|_{s_1} \le N^{2(\tau'+\varsigma s_1)+3}|A_0 - A_0'|_{+,s_1} . \tag{9.3.79}$$

By (9.3.56) and Lemma 3.3.8 we have

$$
\begin{aligned}
|(a^j)' - a^j|_{s_1+1} &\lesssim_{s_1} |b_j^{\sharp\prime} - b_j^\sharp|_{s_1+1} \\
&\overset{(9.3.62),(3.3.20)}{\lesssim_{s_1}} N\big|(T_j^{\sharp\prime})_N^{-1} - (T_j^\sharp)_N^{-1}\big|_{s_1}|g_{j,N}^{\sharp\prime}|_{s_1} + N\big|(T_j^\sharp)_N^{-1}\big|_{s_1}|g_{j,N}^{\sharp\prime} - g_{j,N}^\sharp|_{s_1} \\
&\overset{(9.3.79),(9.3.53),(9.2.9)}{\lesssim_{s_1}} N^{2(\tau'+\varsigma s_1)+4}|A_0' - A_0|_{+,s_1}\delta_1^{\frac{9}{10}} + N^{2(\tau'+\varsigma s_1)+4}|\rho - \rho'|_{s_1} \\
&\overset{(9.3.68)}{\lesssim_{s_1}} N^Q\delta_1^{\frac{9}{10}}|A_0' - A_0|_{+,s_1} + N^Q|\rho - \rho'|_{+,s_1} \qquad (9.3.80) \\
&\overset{(9.3.74)}{\ll} \delta_1^{\frac{7}{8}}|A_0' - A_0|_{+,s_1} + \delta_1^{-\frac{1}{40}}|\rho - \rho'|_{+,s_1}
\end{aligned}
$$

proving (9.3.60). In addition

$$
T_j^\sharp b_j^\sharp = g_j^\sharp + z_N\,, \quad T_j^{\sharp\prime} b_j^{\sharp\prime} = g_j^{\sharp\prime} + z_N'\,,
$$

where, by (9.3.64),

$$
\begin{aligned}
|z_N - z_N'|_{s_1+\frac12} &= \big|\big(-\mathit{\Pi}_N^\perp g_j^\sharp + \mathit{\Pi}_N^\perp T_j^\sharp b_j^\sharp\big) - \big(-\mathit{\Pi}_N^\perp g_j^{\sharp\prime} + \mathit{\Pi}_N^\perp T_j^{\sharp\prime} b_j^{\sharp\prime}\big)\big|_{s_1+\frac12} \\
&\le |g_j^\sharp - g_j^{\sharp\prime}|_{s_1+\frac12} + \big|\mathit{\Pi}_N^\perp(T_j^\sharp - T_j^{\sharp\prime})b_j^\sharp\big|_{s_1+\frac12} + \big|\mathit{\Pi}_N^\perp T_j^{\sharp\prime}\mathit{\Pi}_{2N}(b_j^\sharp - b_j^{\sharp\prime})\big|_{s_1+\frac12} \\
&\overset{(9.3.37),(9.3.38),(9.3.70)}{\lesssim_{s_1}} |g_j^\sharp - g_j^{\sharp\prime}|_{s_1+\frac12} + \big|A_0 - A_0'\big|_{s_1+\frac12}\big|b_j^\sharp\big|_{s_1+\frac12} \\
&\qquad + \big(N + \big|R_0\big|_{\mathrm{Lip},s_1+\frac12}\big)\big|b_j^\sharp - b_j^{\sharp\prime}\big|_{s_1+\frac12} \\
&\lesssim_{s_1} |g_j^\sharp - g_j^{\sharp\prime}|_{s_1+\frac12} + \big|A_0 - A_0'\big|_{+,s_1}\big|b_j^\sharp\big|_{s_1+\frac12} + (N + \big|R_0\big|_{\mathrm{Lip},+,s_1})\big|b_j^\sharp - b_j^{\sharp\prime}\big|_{s_1+\frac12} \\
&\overset{(9.3.56),(9.3.77),(9.3.80)}{\lesssim_{s_1}} |\rho - \rho'|_{+,s_1} + \delta_1^{\frac78}|A_0 - A_0'|_{+,s_1} \\
&\qquad + N^{Q+1}\delta_1^{\frac{9}{10}}|A_0' - A_0|_{+,s_1} + N^{Q+1}|\rho - \rho'|_{+,s_1} \\
&\overset{(9.3.74)}{\ll} \delta_1^{\frac45}|A_0' - A_0|_{+,s_1} + \delta_1^{-\frac{1}{40}}|\rho - \rho'|_{+,s_1}\,. \qquad (9.3.81)
\end{aligned}
$$

Finally, recalling (9.3.65) and Lemma 3.3.8, we obtain

$$
\big|(T_j a^j - g^j) - (T_j'(a^j)' - (g^j)')\big|_{s_1+\frac12} = |\Pi_{\mathbb{G}}(z_N - z_N')|_{s_1+\frac12} \lesssim_{s_1} |z_N - z_N'|_{s_1+\frac12}
$$

and (9.3.81) implies the estimate (9.3.61). ■

**Step 3: Conclusion of the proof of Lemma 9.2.2.** We consider the sets

$$
\Lambda(\varepsilon;\eta,A_0) := \bigcap_{j\in\mathbb{F}}\Lambda_j^2(\varepsilon;\eta,A_0)\bigcap\Lambda^1(\varepsilon;\eta,A_0)\,, \quad \frac12 \le \eta \le 1\,, \qquad (9.3.82)
$$

where the sets $\Lambda^1(\varepsilon;\eta,A_0)$ are defined in (9.3.6) and the sets $\Lambda^2_j(\varepsilon;\eta,A_0)$ in (9.3.51). By Lemmata 9.3.3 and 9.3.5 and (9.3.51), the sets $\Lambda(\varepsilon;\eta,A_0)$, $1/2\le\eta\le 1$, satisfy the properties 1-3 listed in Proposition 9.1.1 for some $0<\alpha<1/12$.

For all $\lambda\in\Lambda(\varepsilon;1,A_0)$ we define the self-adjoint operator

$$\mathcal{S}(\varphi):=\begin{pmatrix} d_N(\varphi) & a^*(\varphi)\\ a(\varphi) & 0\end{pmatrix} \tag{9.3.83}$$

where $d_N(\varphi)=d_N^*(\varphi)$ is defined in Lemma 9.3.2 and $a:=(a^j)_{j\in\mathbb{F}}$ by Lemma 9.3.9. The operator $\mathcal{S}$ defined in (9.3.83) satisfies the estimates (9.1.7)-(9.1.8) by (9.3.11) and (9.3.57)-(9.3.58).

Proof of (9.2.11). The estimate (9.2.11) follows by (9.3.12), (3.3.44)-(3.3.45), and (9.3.59), recalling the definition of $T_j$ in (9.3.33) and of $g^j$ in (9.3.56).

Proof of (9.2.12)-(9.2.13). By (9.2.2), (9.3.8), Lemma 9.3.9, (3.3.44)-(3.3.45) and the fact that

$$T_j(a_j)=T_j(\Pi_{\mathbb{G}}b_j^\sharp)=\Pi_{\mathbb{G}}T_j^\sharp(b_j^\sharp)\,,$$

(see (9.3.56), (9.3.41)), we derive

$$\begin{aligned}
&|J\bar\omega_\varepsilon\cdot\partial_\varphi\mathcal{S}+[J\mathcal{S},JA_0]|_{\mathrm{Lip},+,s}\\
&\le |\mathit{\Pi}_N\rho_1|_{\mathrm{Lip},+,s}+\sum_{j\in\mathbb{F}}\big|T_j(a_j)\big|_{\mathrm{Lip},s+\frac12}\\
&\lesssim_s |\rho_1|_{\mathrm{Lip},+,s}+\sum_{j\in\mathbb{F}}\big|\mathit{\Pi}_N T_j^\sharp(b_j^\sharp)\big|_{\mathrm{Lip},s+\frac12}+\sum_{j\in\mathbb{F}}\big|\mathit{\Pi}_N^\perp T_j^\sharp(b_j^\sharp)\big|_{\mathrm{Lip},s+\frac12}\\
&\lesssim_s |\rho_1|_{\mathrm{Lip},+,s}+\big|\mathit{\Pi}_N\rho_2\big|_{\mathrm{Lip},s+\frac12}+\sum_{j\in\mathbb{F}}\big|\mathit{\Pi}_N^\perp T_j^\sharp(b_j^\sharp)\big|_{\mathrm{Lip},s+\frac12}\,.
\end{aligned} \tag{9.3.84}$$

Now

$$\begin{aligned}
\big|\mathit{\Pi}_N^\perp T_j^\sharp(b_j^\sharp)\big|_{\mathrm{Lip},s+\frac12} &\overset{(9.3.70)}{\lesssim_s} \big(N+|R_0|_{\mathrm{Lip},s+\frac12}\big)\|b_j^\sharp\|_{\mathrm{Lip},s_1}+\big(N+|R_0|_{\mathrm{Lip},s_1}\big)\|b_j^\sharp\|_{\mathrm{Lip},s+\frac12}\\
&\lesssim_s \big(N+\big|R_0\big|_{\mathrm{Lip},+,s}\big)\|b_j^\sharp\|_{\mathrm{Lip},s_1}+(N+\varepsilon^2)\|b_j^\sharp\|_{\mathrm{Lip},s+\frac12}\\
&\overset{(9.3.67),(9.3.69)}{\lesssim_s} \big(N+|R_0|_{\mathrm{Lip},+,s}\big)N^Q\delta_1^{9/10}\\
&\qquad + N^{Q+1}\big((N^{\varsigma(s-s_1)}+|R_0|_{\mathrm{Lip},+,s})\delta_1^{9/10}+|\rho|_{\mathrm{Lip},+,s}\big)\\
&\lesssim_s N^{Q+1}\big(|R_0|_{\mathrm{Lip},+,s}+|\rho|_{\mathrm{Lip},+,s}\big)+N^{Q+1}\delta_1^{9/10}N^{\varsigma(s-s_1)}\\
&\lesssim_s \delta_1^{-1/40}\big(|R_0|_{\mathrm{Lip},+,s}+|\rho|_{\mathrm{Lip},+,s}\big)+\delta_1^{7/8}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\,.
\end{aligned} \tag{9.3.85}$$

Estimates (9.2.12) (for $\varepsilon$ small enough) and (9.2.13) are immediate consequences of (9.3.84) and (9.3.85).

Proof of (9.2.14). By (9.3.83), Lemma 9.3.2, Lemma 3.3.10, and Lemma 9.3.9, for any $\lambda \in \Lambda(\varepsilon; 1, A_0) \cap \Lambda(\varepsilon; 1, A_0')$, we have

$$\begin{aligned} |\mathcal{S}_{A_0,\rho} - \mathcal{S}_{A_0',\rho'}|_{+,s_1} &\lesssim |d_N - d_N'|_{s_1+\frac{1}{2}} + \sum_{j\in\mathbb{F}} |a_j - a_j'|_{s_1+\frac{1}{2}} \\ &\leq \delta_1^{\frac{4}{5}} |A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}} |\rho - \rho'|_{+,s_1} \end{aligned}$$

proving (9.2.14).

Proof of (9.2.15). Also (9.2.15) is a consequence of Lemma 9.3.2 and Lemma 9.3.9, more precisely of (9.3.14), (3.3.44)-(3.3.45) and (9.3.61). This concludes the proof of Lemma 9.2.2.

## 9.4 Splitting step: Proof of Proposition 9.1.1

Recalling (3.2.28), (3.2.29), (3.2.30) we decompose the coupling term $\rho \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$ as

$$\rho = \Pi_{\mathtt{D}}\rho + \Pi_{\mathtt{O}}\rho \, . \tag{9.4.1}$$

The operator $\Pi_{\mathtt{O}}\rho$ has the form (9.2.3), satisfies (9.2.10) and, by (3.3.35) and (9.1.2), it satisfies also (9.2.9). We then apply Lemma 9.2.2 with $A_0$ and $\rho \rightsquigarrow \Pi_{\mathtt{O}}\rho$. It provides the existence of closed subsets $\Lambda(\varepsilon; \eta, A_0) \subset \widetilde{\Lambda}$, $1/2 \leq \eta \leq 1$, satisfying the properties 1-3 of Proposition 9.1.1, and a self-adjoint operator $\mathcal{S}(\varphi) := \mathcal{S}(\varepsilon, \lambda)(\varphi) \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$ of the form (9.2.1), defined for all $\lambda \in \Lambda(\varepsilon; 1, A_0)$, such that

$$|J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] - J\Pi_{\mathtt{O}}\rho|_{\mathrm{Lip},+,s_1} \leq \delta_1^{\frac{7}{4}} \tag{9.4.2}$$

(see (9.2.11)) and (9.1.7)-(9.1.8), (9.2.12), (9.2.13) hold.

We now conjugate the Hamiltonian operator

$$\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA \, , \quad A = A_0 + \rho \overset{(8.1.1)}{=} \frac{D_V}{1+\varepsilon^2\lambda} + R_0 + \rho \, , \tag{9.4.3}$$

by the symplectic linear invertible transformations $e^{J\mathcal{S}(\varphi)}$. We first notice that the conjugated operator

$$e^{-J\mathcal{S}(\varphi)}\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA(\varphi)\big)e^{J\mathcal{S}(\varphi)} = \bar{\omega}_\varepsilon \cdot \partial_\varphi - JA^+(\varphi) \tag{9.4.4}$$

is Hamiltonian, because $e^{J\mathcal{S}(\varphi)}$ is symplectic and Lemma 3.2.3 implies that the (unbounded) operator $A^+(\varphi)$ is self-adjoint. We are going to prove that (see (9.4.13), (9.1.4))

$$A^+(\varphi) = \frac{D_V}{1+\varepsilon^2\lambda} + R_0^+(\varphi) + \rho^+(\varphi) \, ,$$

where

$$R_0^+ := R_0 + \Pi_{\mathtt{D}}\rho \,, \tag{9.4.5}$$

and the estimates (9.1.11)-(9.1.13), (9.1.14) of Proposition 9.1.1 hold. First notice that the bounded operator $\rho^+(\varphi)$ is self-adjoint because the operators $A^+(\varphi)$, $D_V$ and $R_0^+(\varphi)$ are self-adjoint.

**Lie series expansion.** Consider the 1-parameter family of operators

$$X_t := e^{-tJ\mathcal{S}(\varphi)}\big(\bar\omega_\varepsilon \cdot \partial_\varphi - JA\big)e^{tJ\mathcal{S}(\varphi)} \,, \quad t \in [0,1] \,, \tag{9.4.6}$$

connecting $X_0 := \bar\omega_\varepsilon \cdot \partial_\varphi - JA$ to

$$X_1 := e^{-J\mathcal{S}(\varphi)}(\bar\omega_\varepsilon \cdot \partial_\varphi - JA)e^{J\mathcal{S}(\varphi)} \stackrel{(9.4.4)}{=} \bar\omega_\varepsilon \cdot \partial_\varphi - JA^+ \,. \tag{9.4.7}$$

Since the path $t \mapsto X_t$ solves the problem

$$\begin{cases} \dfrac{dX_t}{dt} = -J\mathcal{S}X_t + X_t J\mathcal{S} = [-J\mathcal{S}, X_t] = \mathrm{Ad}_{(-J\mathcal{S})}(X_t) \\ X_0 = \bar\omega_\varepsilon \cdot \partial_\varphi - JA \,, \end{cases} \tag{9.4.8}$$

we deduce from (9.4.7), (9.4.6), (9.4.8) the Lie series expansion

$$\begin{aligned} \bar\omega_\varepsilon \cdot \partial_\varphi - JA^+ = X_1 &= \sum_{k \geq 0} \frac{1}{k!}\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0) \\ &= X_0 + \mathrm{Ad}_{(-J\mathcal{S})}(X_0) + \sum_{k \geq 2} \frac{1}{k!}\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0) \,. \end{aligned} \tag{9.4.9}$$

Now

$$\begin{aligned} \mathrm{Ad}_{(-J\mathcal{S})}(X_0) &= X_0 J\mathcal{S} - J\mathcal{S}X_0 \\ &= J\mathcal{S}\,\bar\omega_\varepsilon \cdot \partial_\varphi + \bar\omega_\varepsilon \cdot \partial_\varphi(J\mathcal{S}) - JAJ\mathcal{S} - (J\mathcal{S}\,\bar\omega_\varepsilon \cdot \partial_\varphi - J\mathcal{S}JA) \\ &= \bar\omega_\varepsilon \cdot \partial_\varphi(J\mathcal{S}) + [J\mathcal{S}, JA] \\ &= J\,\bar\omega_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] + [J\mathcal{S}, J\rho] \end{aligned} \tag{9.4.10}$$

recalling that $A = A_0 + \rho$, see (9.4.3). Comparing (9.4.9) and (9.4.10) we obtain

$$\begin{aligned} -JA^+ &= -JA + J\bar\omega_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] + [J\mathcal{S}, J\rho] + \sum_{k \geq 2} \frac{1}{k!}\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0) \\ &= -JA_0 - J\Pi_{\mathtt{D}}\rho - J\rho^+ \end{aligned} \tag{9.4.11}$$

where, recalling the decomposition (9.4.1),

$$-J\rho^+ := \big(J\bar\omega_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] - J\Pi_{\mathtt{0}}\rho\big) + [J\mathcal{S}, J\rho] + \sum_{k \geq 2} \frac{1}{k!}\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0) \,. \tag{9.4.12}$$

Note that, by (9.4.2), the first addendum in the expression of $-J\rho^+$ in (9.4.12) is very small (in low norm) and the others terms in (9.4.12) are "quadratic" in $\rho$ (note that the term $\rho^+$ satisfies the estimate (9.4.17)). By (9.4.11) we have that

$$\begin{aligned} A^+ = A_0 + \Pi_{\mathtt{D}}\rho + \rho^+ &\overset{(8.1.1)}{=} \frac{D_V}{1+\varepsilon^2\lambda} + R_0 + \Pi_{\mathtt{D}}\rho + \rho^+ \\ &\overset{(9.4.5)}{=} \frac{D_V}{1+\varepsilon^2\lambda} + R_0^+ + \rho^+ \,. \end{aligned} \tag{9.4.13}$$

**Lemma 9.4.1.** *$A^+$ has the form* (9.1.4)-(9.1.5) *and* (9.1.9)-(9.1.10) *holds.*

Proof. By (9.4.13), (8.1.2), (3.2.28), (3.2.29), (8.1.3), the operator $A^+$ has the form (9.1.4) with

$$\begin{aligned} D_0^+ &= D_0 + D_+(\rho_{\mathbb{F}}^{\mathbb{F}}) = \mathrm{Diag}_{j\in\mathbb{F}}\big(\mu_j(\varepsilon,\lambda)\mathrm{Id}_2 + \pi_+[\widehat{\rho}_j^{\,j}(0)]\big)\,, \\ V_0^+ &= V_0 + \rho_{\mathbb{G}}^{\mathbb{G}}\,. \end{aligned}$$

Since $\rho_j^j(\varphi)$ is a $2\times 2$ symmetric matrix, (3.2.21) implies (9.1.5) with

$$\mu_j^+(\varepsilon,\lambda) = \mu_j(\varepsilon,\lambda) + \frac{1}{2}\mathrm{Tr}[\widehat{\rho}_j^{\,j}(0)]\,.$$

Hence

$$|\mu_j^+(\varepsilon,\lambda) - \mu_j(\varepsilon,\lambda)|_{\mathrm{Lip}} + \|V_0^+ - V_0\|_{\mathrm{Lip},0} \lesssim \|\rho\|_{\mathrm{Lip},0}$$

and by (9.1.2) the bounds (9.1.9)-(9.1.10) follow. ■

The first estimate in (9.1.11) follows by

$$|R_0^+ - R_0|_{\mathrm{Lip},+,s_1} \overset{(9.4.5)}{=} |\Pi_{\mathtt{D}}\rho|_{\mathrm{Lip},+,s_1} \overset{(3.3.35),(9.1.2)}{\lesssim_{s_1}} \delta_1 \le \delta_1^{\frac{3}{4}} \tag{9.4.14}$$

for $\delta_1 \le \varepsilon^3$ small. For the estimate of $|\rho^+|_{\mathrm{Lip},+,s_1}$ we use the following lemma.

**Lemma 9.4.2.** *We have*

$$|\mathrm{Ad}_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s_1} \lesssim_{s_1} \delta_1 \tag{9.4.15}$$

$$|\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s_1} \le \delta_1^{1+\frac{3}{4}(k-1)}\,,\ \forall k \ge 2\,. \tag{9.4.16}$$

Proof. By (9.4.10) and (3.3.24), we have

$$\begin{aligned} |\mathrm{Ad}_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s_1} &= \big|J\bar\omega_\varepsilon\cdot\partial_\varphi\mathcal{S} + [J\mathcal{S},JA_0] + [J\mathcal{S},J\rho]\big|_{\mathrm{Lip},+,s_1} \\ &\le \big|J\bar\omega_\varepsilon\cdot\partial_\varphi\mathcal{S} + [J\mathcal{S},JA_0] - \Pi_{\mathtt{0}}\rho\big|_{\mathrm{Lip},+,s_1} \\ &\quad + \big|\Pi_{\mathtt{0}}\rho\big|_{\mathrm{Lip},+,s_1} + C(s_1)|J\mathcal{S}|_{\mathrm{Lip},s_1+\frac{1}{2}}|\rho|_{\mathrm{Lip},+,s_1} \\ &\overset{(9.4.2),(3.3.35),(9.1.2),(9.1.7)}{\le} \delta_1^{\frac{7}{4}} + C(s_1)\delta_1 + C(s_1)\delta_1^{\frac{7}{8}+1} \\ &\lesssim_{s_1} \delta_1 \end{aligned}$$

for $\delta_1$ small, proving (9.4.15). In order to prove (9.4.16), let $m_k := |\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s_1}$. For $k \geq 1$, we get, for $\delta_1$ small,

$$m_{k+1} = \big|[-J\mathcal{S}, \mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)]\big|_{\mathrm{Lip},+,s_1} \overset{(3.3.24)}{\leq} C(s_1)|\mathcal{S}|_{\mathrm{Lip},s_1+\frac{1}{2}}\, m_k \overset{(9.1.7)}{\leq} C(s_1)\delta_1^{\frac{7}{8}} m_k \ll \delta_1^{\frac{3}{4}} m_k\,.$$

This iterative inequality and (9.4.15) imply (9.4.16). ∎

We derive by (9.4.12) the bound

$$\begin{aligned} |\rho^+|_{\mathrm{Lip},+,s_1} &\leq |J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0] - J\Pi_{\mathtt{0}}\rho|_{\mathrm{Lip},+,s_1} + |[J\mathcal{S}, J\rho]|_{\mathrm{Lip},+,s_1} \\ &\quad + \sum_{k\geq 2} \frac{1}{k!}\big|\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)\big|_{\mathrm{Lip},+,s_1} \\ &\overset{(9.4.2),(3.3.24),(9.1.2),(9.1.7),(9.4.16)}{\leq} \delta_1^{\frac{7}{4}} + C(s_1)\delta_1^{\frac{7}{8}+1} + \delta_1^{\frac{7}{4}} \ll \delta_1^{\frac{3}{2}} \end{aligned} \tag{9.4.17}$$

for $\delta_1 \leq \varepsilon^3$ small. This proves the second estimate in (9.1.11).

There remains to estimate the high norms $|R_0^+|_{\mathrm{Lip},+,s} + |\rho^+|_{\mathrm{Lip},+,s}$ for $s \geq s_2$. First, by (9.4.5) and (3.3.35),

$$|R_0^+|_{\mathrm{Lip},+,s} = |R_0 + \Pi_{\mathtt{D}}\rho|_{\mathrm{Lip},+,s} \leq |R_0|_{\mathrm{Lip},+,s} + C(s)|\rho|_{\mathrm{Lip},+,s} \tag{9.4.18}$$

which implies the bound for $|R_0^+|_{\mathrm{Lip},+,s}$ in (9.1.14) (actually (9.4.18) is much better than the estimate (9.1.14) for $|R_0^+|_{\mathrm{Lip},+,s}$). For $|\rho^+|_{\mathrm{Lip},+,s}$ we use the following lemma.

**Lemma 9.4.3.** *For $k \geq 1$,*

$$|\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s} \leq (C(s))^k \mathcal{M}_s \delta_1^{3\frac{k-1}{4}} \tag{9.4.19}$$

*where*

$$\mathcal{M}_s := \delta_1^{-\frac{1}{4}}\big(|R_0|_{\mathrm{Lip},+,s} + |\rho|_{\mathrm{Lip},+,s}\big) + \delta_1^{-\frac{3}{4}}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\,. \tag{9.4.20}$$

Proof. By (9.4.10) and (3.3.24) we have

$$\begin{aligned} |\mathrm{Ad}_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s} &\leq \big|J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0]\big|_{\mathrm{Lip},+,s} + \big|[J\mathcal{S}, J\rho]\big|_{\mathrm{Lip},+,s} \\ &\leq \big|J\bar{\omega}_\varepsilon \cdot \partial_\varphi \mathcal{S} + [J\mathcal{S}, JA_0]\big|_{\mathrm{Lip},+,s} + C(s)\big|\mathcal{S}\big|_{\mathrm{Lip},s_1+\frac{1}{2}}|\rho|_{\mathrm{Lip},+,s} \\ &\quad + C(s)|\mathcal{S}|_{\mathrm{Lip},s+\frac{1}{2}}|\rho|_{\mathrm{Lip},+,s_1} \\ &\overset{(9.2.13),(9.1.7),(9.1.8)}{\leq} C(s)\Big[\delta_1^{-\frac{1}{4}}\big(|R_0|_{\mathrm{Lip},+,s} + |\rho|_{\mathrm{Lip},+,s}\big) + \delta_1^{-\frac{3}{4}}\delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\Big]\,. \end{aligned}$$

Hence, the estimate (9.4.19) is proved for $k = 1$, recall (9.4.20). In order to prove (9.4.19) for $k \geq 2$, let $M_k := |\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s}$. For $k \geq 1$, we have, by (3.3.24),

$$
\begin{aligned}
M_{k+1} &= \big|[-J\mathcal{S}, \mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)]\big|_{\mathrm{Lip},+,s} \\
&\leq C(s)|\mathcal{S}|_{\mathrm{Lip},s_1+\frac{1}{2}} M_k + C(s)|\mathcal{S}|_{\mathrm{Lip},s+\frac{1}{2}} |\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)|_{\mathrm{Lip},+,s_1} \\
&\overset{(9.1.7),(9.1.8),(9.4.20),(9.4.16)}{\leq} C(s)\big(\delta_1^{\frac{7}{8}} M_k + \mathcal{M}_s \delta_1^{1+\frac{3}{4}(k-1)}\big)\,. \qquad (9.4.21)
\end{aligned}
$$

Then (9.4.19) follows by iteration from (9.4.21), provided $\varepsilon$ is small enough (independently of $s$). ■

Finally, by (9.4.12) and (9.4.10) we get

$$
-J\rho^+ = -J\Pi_{\mathsf{0}}\rho + \sum_{k\geq 1} \frac{1}{k!} \mathrm{Ad}^k_{(-J\mathcal{S})}(X_0) \qquad (9.4.22)
$$

so that

$$
\begin{aligned}
|\rho^+|_{\mathrm{Lip},+,s} &\leq |\Pi_{\mathsf{0}}\rho|_{\mathrm{Lip},+,s} + \sum_{k\geq 1} \frac{1}{k!} \big|\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0)\big|_{\mathrm{Lip},+,s} \\
&\overset{(3.3.35),(9.4.19)}{\leq} C(s)|\rho|_{\mathrm{Lip},+,s} + \mathcal{M}_s \delta_1^{-3/4} \sum_{k\geq 1} \frac{\big(C(s)\delta_1^{3/4}\big)^k}{k!} \\
&\leq C(s)|\rho|_{\mathrm{Lip},+,s} + \mathcal{M}_s C(s) \frac{e^{C(s)\delta_1^{3/4}} - 1}{C(s)\delta_1^{3/4}} \\
&\overset{(9.4.20)}{\lesssim_s} \delta_1^{-1/4}\big(|R_0|_{\mathrm{Lip},+,s} + |\rho|_{\mathrm{Lip},+,s}\big) + \delta_1^{-3/4} \delta_1^{-3\varsigma\frac{s-s_2}{s_2-s_1}}\,. \qquad (9.4.23)
\end{aligned}
$$

The estimate (9.1.14) is a consequence of (9.4.18) and (9.4.23). We can prove in the same way (9.1.13) taking $\delta_1$ small enough (depending on $s_2$).

At last, by (9.4.5) and (3.3.35) we get

$$
|{R'_0}^+ - R_0^+|_{+,s_1} \leq |R'_0 - R_0|_{+,s_1} + C|\rho' - \rho|_{+,s_1}
$$

for some positive $C := C(s_1)$, proving (9.1.15).

**Lemma 9.4.4.** (9.1.16) *holds.*

PROOF. By (9.4.22) we have

$$
\begin{aligned}
-J(\rho^+ - (\rho^+)') &= -J\Pi_{\mathsf{0}}(\rho - \rho') + \mathrm{Ad}_{(-J\mathcal{S})}(X_0) - \mathrm{Ad}_{(-J\mathcal{S}')}(X'_0) \\
&\quad + \sum_{k\geq 2} \frac{1}{k!}\big(\mathrm{Ad}^k_{(-J\mathcal{S})}(X_0) - \mathrm{Ad}^k_{(-J\mathcal{S}')}(X'_0)\big)\,. \qquad (9.4.24)
\end{aligned}
$$

Set

$$V_k := \mathrm{Ad}^k_{(-J\mathcal{S})}(X_0) - \mathrm{Ad}^k_{(-J\mathcal{S}')}(X_0'), \qquad v_k := |V_k|_{+,s_1}, \quad k \geq 1. \tag{9.4.25}$$

By (9.2.15) in Lemma 9.2.2 (that we applied with $\Pi_0\rho$ instead of $\rho$), and recalling (9.4.10), we have

$$\begin{aligned} |V_1 - J\Pi_0(\rho-\rho')|_{+,s_1} &\leq \delta_1^{\frac{3}{4}}|A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}}|\Pi_0(\rho-\rho')|_{+,s_1} \\ &\quad + |[J\mathcal{S}, J\Pi_0\rho] - [J\mathcal{S}', J\Pi_0\rho']|_{+,s_1} \qquad (9.4.26) \\ &\lesssim_{s_1} \delta_1^{\frac{3}{4}}|A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}}|\rho-\rho'|_{+,s_1} \qquad (9.4.27) \end{aligned}$$

because the term in (9.4.26) is bounded as

$$\begin{aligned} |[J\mathcal{S}, J\Pi_0\rho] - [J\mathcal{S}', J\Pi_0\rho']|_{+,s_1} &\overset{(3.3.23)}{\lesssim_{s_1}} |\mathcal{S}-\mathcal{S}'|_{s_1+\frac{1}{2}}|\Pi_0\rho|_{+,s_1} + |\mathcal{S}'|_{s_1+\frac{1}{2}}|\Pi_0(\rho-\rho')|_{+,s_1} \\ &\overset{(3.3.35)}{\lesssim_{s_1}} |\mathcal{S}-\mathcal{S}'|_{+,s_1}|\rho|_{+,s_1} + |\mathcal{S}'|_{s_1+\frac{1}{2}}|\rho-\rho'|_{+,s_1} \\ &\overset{(9.2.14),(9.1.2),(9.1.7)}{\lesssim_{s_1}} \big(\delta_1^{\frac{3}{4}}|A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}}|\rho-\rho'|_{+,s_1}\big)\delta_1 \\ &\qquad + \delta_1^{\frac{7}{8}}|\rho-\rho'|_{+,s_1} \end{aligned}$$

(the estimate (9.1.7) is applied to $\mathcal{S}'$). By (9.4.27) and (3.3.35) we deduce that

$$v_1 := |V_1|_{+,s_1} \leq C(s_1)\big(\delta_1^{\frac{3}{4}}|A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}}|\rho-\rho'|_{+,s_1}\big). \tag{9.4.28}$$

Now, recalling (9.4.25),

$$V_{k+1} = \mathrm{Ad}_{(-J\mathcal{S})}(V_k) + \big[J\mathcal{S}' - J\mathcal{S}, \mathrm{Ad}^k_{(-J\mathcal{S}')}(X_0')\big].$$

Hence, for $k \geq 1$, we get, setting $m_k' := |\mathrm{Ad}^k_{(-J\mathcal{S}')}(X_0')|_{+,s_1}$,

$$\begin{aligned} v_{k+1} := |V_{k+1}|_{+,s_1} &\lesssim_{s_1} |\mathcal{S}|_{+,s_1}v_k + |\mathcal{S}-\mathcal{S}'|_{+,s_1}m_k' \\ &\leq C(s_1)\delta_1^{\frac{7}{8}}v_k + C(s_1)\delta_1^{1+\frac{3}{4}(k-1)}\big(\delta_1^{\frac{3}{4}}|A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{30}}|\rho-\rho'|_{+,s_1}\big) \end{aligned}$$

by (9.1.7), (9.2.14) and Lemma 9.4.2. From the previous iterative estimate and (9.4.28), we may derive

$$\sum_{k\geq 2} v_k \leq \delta_1^{\frac{3}{4}}\big(|A_0 - A_0'|_{+,s_1} + |\rho-\rho'|_{+,s_1}\big). \tag{9.4.29}$$

Finally, recalling (9.4.24) and (9.4.25), we get

$$\begin{aligned} |\rho^+ - \rho'^+|_{+,s_1} &\leq \big|V_1 - J\Pi_0(\rho-\rho')\big|_{+,s_1} + \sum_{k\geq 2} v_k \\ &\overset{(9.4.27),(9.4.29)}{\leq} \delta_1^{\frac{1}{2}}|A_0 - A_0'|_{+,s_1} + \delta_1^{-\frac{1}{20}}|\rho-\rho'|_{+,s_1} \end{aligned}$$

which is (9.1.16). ∎

# Chapter 10

# Construction of approximate right inverse

The goal of this Chapter is to prove Proposition 8.2.1.

## 10.1 Splitting of low-high normal subspaces

The first step in order to find an approximate solution of the equation

$$\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho)\big)h = g\,, \tag{10.1.1}$$

see (8.2.7), is to apply the splitting Corollary 9.1.2, whose assumptions are verified by the hypothesis of Proposition 8.2.1. As a consequence there are sets $\Lambda_\infty(\varepsilon; \eta, A_0, \rho) \subset \tilde{\Lambda}$, $1/2 \le \eta \le 5/6$, and a sequence of symplectic linear transformations $\mathcal{P}_{\mathfrak{n}}(\varphi) \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$, $\mathfrak{n} \ge 1$, satisfying (9.1.21)-(9.1.23), such that, for all $\lambda \in \Lambda_\infty(\varepsilon; 5/6, A_0, \rho)$, the conjugation (9.1.34) holds for any $\mathfrak{n} \ge 1$, namely

$$\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_0 - J\rho\big)\mathcal{P}_{\mathfrak{n}}(\varphi) = \mathcal{P}_{\mathfrak{n}}(\varphi)\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}} - J\rho_{\mathfrak{n}}\big)\,, \tag{10.1.2}$$

where the operators $A_{\mathfrak{n}}$ have the form described in (9.1.25)-(9.1.26) and satisfy (9.1.27)-(9.1.31), in particular $A_{\mathfrak{n}} \in \mathcal{C}(2C_1, c_1/2, c_2/2)$ is an admissible split operator according to Definition 8.1.1. Moreover the coupling operator $\rho_{\mathfrak{n}} \in \mathcal{L}(H_{\mathbb{S}}^{\perp})$ in (10.1.2) satisfies the estimates (9.1.32)-(9.1.33), in particular it is small in the low norm $|\ |_{\mathrm{Lip},+,s_1}$.

Thus, given a function $g$ satisfying (8.2.5), in order to find a function $h$ such that (8.2.7) holds with a remainder $r$ satisfying (8.2.8), we make the change of variables

$$g'(\varphi) := \mathcal{P}_{\mathfrak{n}}^{-1}(\varphi)g(\varphi) \in H_{\mathbb{S}}^{\perp}\,, \qquad h'(\varphi) := \mathcal{P}_{\mathfrak{n}}^{-1}(\varphi)h(\varphi) \in H_{\mathbb{S}}^{\perp}\,, \tag{10.1.3}$$

and we look for an approximate solution of the equation

$$\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - JA_{\mathfrak{n}} - J\rho_{\mathfrak{n}}\big)h' = g'\,. \tag{10.1.4}$$

In sections 10.2 and 10.3 we shall solve approximately the more general equation

$$\big(\bar{\omega}_\varepsilon \cdot \partial_\varphi - J\mathfrak{A}(\varepsilon,\lambda,\varphi) - J\varrho\big)h' = g' \tag{10.1.5}$$

with an admissible split operator $\mathfrak{A} \in \mathcal{C}(2C_1, c_1/2, c_2/2)$ as in Definition 8.1.1, i.e. of the form

$$\mathfrak{A}(\varepsilon,\lambda,\varphi) = \frac{D_V}{1+\varepsilon^2\lambda} + \mathfrak{R}(\varepsilon,\lambda,\varphi)\,, \quad |\mathfrak{R}|_{\mathrm{Lip},+,s_1} \le 2C_1\varepsilon^2\,, \tag{10.1.6}$$

and a self-adjoint operator $\varrho \in \mathcal{L}(H_{\mathbb{S}}^\perp)$ satisfying suitable smallness conditions (see (10.3.2)).

Multiplying by $J$ both sizes of the equation (10.1.5), we look for an approximate solution $h'$ of

$$\mathcal{L}h' = Jg' \qquad \text{with} \qquad \mathcal{L} := \mathcal{L}_D + \varrho\,, \quad \mathcal{L}_D := J\bar{\omega}_\varepsilon \cdot \partial_\varphi + \mathfrak{A}\,. \tag{10.1.7}$$

Notice that the operator $\mathcal{L}_D$ is block-diagonal according to the splitting $H_{\mathbb{S}}^\perp = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, as the admissible split operator $\mathfrak{A} \in \mathcal{C}(2C_1, c_1/2, c_2/2)$.

In section 10.2 we shall first find an approximate right inverse of $\mathcal{L}_D$. Then in section 10.3 we shall obtain, by a Neumann series argument, an approximate right inverse of the whole $\mathcal{L} = \mathcal{L}_D + \varrho$ under a suitable smallness condition on $\varrho$, see (10.3.2).

Finally in section 10.4 we shall apply the results of sections 10.2 and 10.3 to obtain an approximate solution of the equation (10.1.5) with $\mathfrak{A} = A_{\mathfrak{n}}$ and $\varrho = \rho_{\mathfrak{n}}$, namely of equation (10.1.4), and, ultimately, of (10.1.1). The number $\mathfrak{n}$ of splitting steps will be chosen in section 10.4 large enough (see (10.4.3)-(10.4.4)), so that the coupling term $\rho_{\mathfrak{n}}$ is a small perturbation of $\mathcal{L}_D$, satisfying the smallness condition (10.3.2).

## 10.2 Approximate right inverse of $\mathcal{L}_D$

We first obtain an approximate right inverse of the block diagonal operator $\mathcal{L}_D$ introduced in (10.1.7).

**Proposition 10.2.1. (Approximate right inverse of $\mathcal{L}_D$).** *Let $\mathfrak{A}$ be a split admissible operator in $\mathcal{C}(2C_1, c_1/2, c_2/2)$ according to Definition 8.1.1. Then there are closed subsets $\Lambda(\varepsilon;\eta,\mathfrak{A})$, $1/2 \le \eta \le 1$, satisfying Properties 1-3 of Proposition 4.1.5, and $\bar{N} \in \mathbb{N}$ such that, for all $N \ge \bar{N}$, there exists a linear operator $\mathcal{I}_D := \mathcal{I}_{D,N}$, defined for $\lambda \in \Lambda(\varepsilon;\eta,\mathfrak{A})$, with the following properties:*

- *setting*

$$Q' := 2(\tau' + \varsigma s_1) + 3\,, \tag{10.2.1}$$

*where $\tau'$ is given by Proposition 4.1.5, we have*

$$\|\mathcal{I}_D g'\|_{\mathrm{Lip},s_0} \lesssim_{s_0} N^{Q'} \|g'\|_{\mathrm{Lip},s_0}\,, \quad \|\mathcal{I}_D g'\|_{\mathrm{Lip},s_1} \lesssim_{s_1} N^{Q'} \|g'\|_{\mathrm{Lip},s_1}\,, \tag{10.2.2}$$

*and,* $\forall s \geq s_1$,

$$\|\mathcal{I}_D g'\|_{\mathrm{Lip},s} \leq C N^{Q'} \|g'\|_{\mathrm{Lip},s} + C(s) N^{Q'} (N^{\varsigma(s-s_1)} + |\mathfrak{R}|_{\mathrm{Lip},+,s}) \|g'\|_{\mathrm{Lip},s_1} \tag{10.2.3}$$

*where* $C$ *is a constant independent of* $s$ *(it depends on* $s_1$*) and* $\mathfrak{R}$ *is introduced in* (10.1.6)*. Furthermore*

$$\|\mathcal{I}_D g'\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \|g'\|_{\mathrm{Lip},s_0+Q'} \,. \tag{10.2.4}$$

- $\mathcal{I}_D$ *is an approximate right inverse of* $\mathcal{L}_D$*, in the sense that*

$$\begin{aligned} &\|(\mathcal{L}_D \mathcal{I}_D - \mathrm{Id}) g'\|_{\mathrm{Lip},s_1} \\ &\quad \lesssim_{s_3} N^{Q'+1-(s_3-s_1)} \big( \|g'\|_{\mathrm{Lip},s_3} + \big( N^{\varsigma(s_3-s_1)} + |\mathfrak{R}|_{\mathrm{Lip},+,s_3} \big) \|g'\|_{\mathrm{Lip},s_1} \big) \,. \end{aligned} \tag{10.2.5}$$

The rest of this section is devoted to the proof of Proposition 10.2.1.

Recalling Definition 8.1.1, the split admissible operator $\mathfrak{A} \in \mathcal{C}(2C_1, c_1/2, c_2/2)$ in (10.1.6) is block-diagonal according to the splitting $H_{\mathbb{S}}^{\perp} = H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, i.e.

$$\mathfrak{A}(\varepsilon, \lambda, \varphi) = \begin{pmatrix} \mathfrak{D}(\varepsilon, \lambda) & 0 \\ 0 & \mathfrak{W}(\varepsilon, \lambda, \varphi) \end{pmatrix} , \tag{10.2.6}$$

and, moreover, in the basis of the eigenfunctions $\{(\Psi_j, 0),\ (0, \Psi_j)\}_{j \in \mathbb{F}}$ (see (3.1.9)), the operator $\mathfrak{D}$ is represented by the diagonal matrix

$$\mathfrak{D} := \mathrm{Diag}_{j \in \mathbb{F}}\, \mathfrak{m}_j(\varepsilon, \lambda) \mathrm{Id}_2 \,, \tag{10.2.7}$$

where each $\mathfrak{m}_j(\varepsilon, \lambda) \in \mathbb{R}$ satisfies (see (8.1.4))

$$|\mathfrak{m}_j(\varepsilon, \lambda) - \mu_j| \leq C_1 \varepsilon^2 \,,$$

and the estimates (see (8.1.7))

$$\forall j \in \mathbb{F} \quad \Big( \frac{c_2}{2} \varepsilon^2 \leq \mathfrak{d}_\lambda \mathfrak{m}_j(\varepsilon, \lambda) \leq 2 c_2^{-1} \varepsilon^2 \quad \text{or} \quad -2 c_2^{-1} \varepsilon^2 \leq \mathfrak{d}_\lambda \mathfrak{m}_j(\varepsilon, \lambda) \leq -\frac{c_2}{2} \varepsilon^2 \Big) \tag{10.2.8}$$

and, by (8.1.8),

$$\mathfrak{d}_\lambda \mathfrak{W}(\varepsilon, \lambda) \leq -\frac{c_1}{2} \varepsilon^2 \mathrm{Id} \,. \tag{10.2.9}$$

By (10.2.6), in order to find an approximate solution of $(J \bar{\omega}_\varepsilon \cdot \partial_\varphi + \mathfrak{A}) h' = g'$, where $g' \in L^2(\mathbb{T}^{|\mathbb{S}|}, H_{\mathbb{S}}^{\perp})$, we have to solve approximately the pair of *decoupled* equations

$$\big( J \bar{\omega}_\varepsilon \cdot \partial_\varphi + \mathfrak{D}(\varepsilon, \lambda) \big) h'_{\mathbb{F}} = g'_{\mathbb{F}} \tag{10.2.10}$$

$$\big( J \bar{\omega}_\varepsilon \cdot \partial_\varphi + \mathfrak{W}(\varepsilon, \lambda, \varphi) \big) h'_{\mathbb{G}} = g'_{\mathbb{G}} \tag{10.2.11}$$

having used the notation

$$
\begin{gathered}
h' = \begin{pmatrix} h'_{\mathbb{F}} \\ h'_{\mathbb{G}} \end{pmatrix}, \ g' = \begin{pmatrix} g'_{\mathbb{F}} \\ g'_{\mathbb{G}} \end{pmatrix}, \\
h'_{\mathbb{F}} := \Pi_{\mathbb{F}} h', \ h'_{\mathbb{G}} := \Pi_{\mathbb{G}} h', \ g'_{\mathbb{F}} := \Pi_{\mathbb{F}} g', \ g'_{\mathbb{G}} := \Pi_{\mathbb{G}} g'.
\end{gathered} \tag{10.2.12}
$$

Solution of (10.2.10). Recalling (10.2.6) and (10.2.7), and decomposing

$$
\begin{aligned}
h'_{\mathbb{F}} &= \sum_{j \in \mathbb{F}} h'_j(\varphi) \Psi_j(x), \quad h'_j(\varphi) \in \mathbb{R}^2, \\
g'_{\mathbb{F}} &= \sum_{j \in \mathbb{F}} g'_j(\varphi) \Psi_j(x), \quad g'_j(\varphi) \in \mathbb{R}^2,
\end{aligned} \tag{10.2.13}
$$

the equation (10.2.10) reduces to the decoupled system of scalar equations

$$
\big( J \bar{\omega}_\varepsilon \cdot \partial_\varphi + \mathfrak{m}_j(\varepsilon, \lambda) \big) h'_j(\varphi) = g'_j(\varphi), \quad j \in \mathbb{F}. \tag{10.2.14}
$$

Setting, for $1/2 \le \eta \le 1$,

$$
\Lambda^1(\varepsilon; \eta, \mathfrak{A}) := \Big\{ \lambda \in \widetilde{\Lambda} \,:\, |\bar{\omega}_\varepsilon \cdot \ell + \mathfrak{m}_j(\varepsilon, \lambda)| \ge \frac{\gamma_1}{2\eta \langle \ell \rangle^\tau}, \ \forall (\ell, j) \in \mathbb{Z}^{|\mathbb{S}|} \times \mathbb{F} \Big\} \tag{10.2.15}
$$

with $\gamma_1 = \gamma_0/2$ and $\tau$ satisfying (9.3.7), we have the following usual lemma.

**Lemma 10.2.2. (Solution of** (10.2.10)**)** *For all* $\lambda \in \Lambda^1(\varepsilon; 1, \mathfrak{A})$ *the equation* (10.2.10) *has a solution* $h'_{\mathbb{F}} = \sum_{j \in \mathbb{F}} h'_j(\varphi) \Psi_j(x)$*, written as in* (10.2.13)*, satisfying*

$$
\| h'_j \|_{\mathrm{Lip}, H^s(\mathbb{T}^{|\mathbb{S}|})} \le C \| g'_j \|_{\mathrm{Lip}, H^{s+2\tau}(\mathbb{T}^{|\mathbb{S}|})}, \quad \forall j \in \mathbb{F}. \tag{10.2.16}
$$

*Moreover, for all* $N \in \mathbb{N} \backslash \{0\}$,

$$
\begin{aligned}
\| \varPi_N h'_j \|_{\mathrm{Lip}, H^s(\mathbb{T}^{|\mathbb{S}|})} &\le C \| \varPi_N g'_j \|_{\mathrm{Lip}, H^{s+2\tau}(\mathbb{T}^{|\mathbb{S}|})} \\
&\le C N^{2\tau} \| \varPi_N g'_j \|_{\mathrm{Lip}, H^s(\mathbb{T}^{|\mathbb{S}|})}, \quad \forall j \in \mathbb{F}.
\end{aligned} \tag{10.2.17}
$$

Proof. By the Fourier expansion

$$
g'_j(\varphi) = \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} \widehat{g'_j}(\ell) e^{\mathrm{i} \ell \cdot \varphi}, \quad \widehat{g'_j}(\ell) \in \mathbb{C}^2, \quad \overline{\widehat{g'_j}(\ell)} = \widehat{g'_j}(-\ell), \tag{10.2.18}
$$

$$
h'_j(\varphi) = \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} \widehat{h'_j}(\ell) e^{\mathrm{i} \ell \cdot \varphi}, \quad \widehat{h'_j}(\ell) \in \mathbb{C}^2, \quad \overline{\widehat{h'_j}(\ell)} = \widehat{h'_j}(-\ell), \tag{10.2.19}
$$

each equation (10.2.14) amounts to

$$M_{j,\ell}\widehat{h'_j}(\ell) = \widehat{g'_j}(\ell)\,,\ \ \ell \in \mathbb{Z}^{|\mathbb{S}|}\,, \quad \text{where} \quad M_{j,\ell} := \mathrm{i}\bar{\omega}_\varepsilon \cdot \ell J + \mathfrak{m}_j \mathrm{Id}_2\,, \quad j \in \mathbb{F}\,, \tag{10.2.20}$$

are $2 \times 2$ self-adjoint matrices with eigenvalues $\mathfrak{m}_j \pm \bar{\omega}_\varepsilon \cdot \ell$. As a consequence, for any $\lambda$ in the set $\Lambda^1(\varepsilon; 1, \mathfrak{A})$ defined in (10.2.15), the matrices $M_{j,\ell}$ are invertible and $\|M_{j,\ell}^{-1}\| \leq C\langle \ell \rangle^\tau$, for some positive constant $C := C(\gamma_0)$, for any $\ell \in \mathbb{Z}^{|\mathbb{S}|}$. Hence the equation (10.2.14) has the unique solution

$$h'_j(\varphi) = \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} M_{j,\ell}^{-1} \widehat{g'_j}(\ell) e^{\mathrm{i}\ell\cdot\varphi} \qquad \text{satisfying} \qquad \|h'_j\|_{H^s(\mathbb{T}^{|\mathbb{S}|})} \leq C \|g'_j\|_{H^{s+\tau}(\mathbb{T}^{|\mathbb{S}|})}\,. \tag{10.2.21}$$

In addition $\widehat{h'_j}(\ell) = M_{j,\ell}^{-1}\widehat{g'_j}(\ell)$ satisfies the reality condition (10.2.19) since, taking the complex conjugated equation in (10.2.20), we obtain, by (10.2.18),

$$M_{j,-\ell}\overline{\widehat{h'_j}(\ell)} = \widehat{g'_j}(-\ell)$$

and, therefore, by uniqueness $\overline{\widehat{h'_j}(\ell)} = \widehat{h'_j}(-\ell)$.

Moreover, since $\|M_{j,\ell}^{-1}\| \leq C\langle \ell \rangle^\tau$ and $\|M_{j,\ell}\|_{\mathrm{lip}} \simeq |\mathfrak{m}_j|_{\mathrm{lip}} \leq 2c_2^{-1}\varepsilon^2$ (see (10.2.8)), we get

$$\begin{aligned}
\|h'_j\|_{\mathrm{lip}, H^s(\mathbb{T}^{|\mathbb{S}|})} &\lesssim \Big( \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} \langle \ell \rangle^{2s} \big( \|M_{j,\ell}^{-1}\|_{\mathrm{lip}} |\widehat{g'_j}(\ell)| + \|M_{j,\ell}^{-1}\| |\widehat{g'_j}(\ell)|_{\mathrm{lip}} \big)^2 \Big)^{1/2} \\
&\lesssim \Big( \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} \langle \ell \rangle^{2s} \big( \|M_{j,\ell}^{-1}\|^2 \|M_{j,\ell}\|_{\mathrm{lip}} |\widehat{g'_j}(\ell)| + \|M_{j,\ell}^{-1}\| |\widehat{g'_j}(\ell)|_{\mathrm{lip}} \big)^2 \Big)^{1/2} \\
&\lesssim \Big( \sum_{\ell \in \mathbb{Z}^{|\mathbb{S}|}} \langle \ell \rangle^{2s} \big( \langle \ell \rangle^{2\tau} \varepsilon^2 |\widehat{g'_j}(\ell)| + \langle \ell \rangle^{\tau} |\widehat{g'_j}(\ell)|_{\mathrm{lip}} \big)^2 \Big)^{1/2} \\
&\lesssim \varepsilon^2 \|g'_j\|_{H^{s+2\tau}(\mathbb{T}^{|\mathbb{S}|})} + \|g'_j\|_{\mathrm{lip}, H^{s+\tau}(\mathbb{T}^{|\mathbb{S}|})}\,.
\end{aligned} \tag{10.2.22}$$

The bounds (10.2.21) and (10.2.22) imply (10.2.16). The estimate (10.2.17) is obtained in the same way, just considering sums over $|\ell| \leq N$. ∎

Moreover, with arguments similar to those used in Lemmas 9.3.3 and 9.3.5, using (10.2.8), the fact that $\mathfrak{m}_j(\varepsilon, \lambda) = \mu_j + O(\varepsilon^2)$, and the unperturbed first order Melnikov non-resonance conditions (1.2.7), we deduce the following measure estimate.

**Lemma 10.2.3. (Measure estimate)** *Let $\tau \geq (3/2)\tau_0 + 3 + |\mathbb{S}|$ ($\tau_0$ is the Diophantine exponent in (1.2.7)). Then the sets $\Lambda^1(\varepsilon; \eta, \mathfrak{A})$ defined in (10.2.15) satisfy, for $\varepsilon$ small enough (depending on $C_1, c_1$),*

$$|[\Lambda^1(\varepsilon; \eta, \mathfrak{A})]^c \cap \widetilde{\Lambda}| \leq \varepsilon\,, \quad \forall 1/2 \leq \eta \leq 1\,. \tag{10.2.23}$$

*Moreover, if $|\mathfrak{A}' - \mathfrak{A}|_{+,s_1} \leq \delta \leq \varepsilon^3$ on $\widetilde{\Lambda} \cap \widetilde{\Lambda}'$, then, for $\eta \in [(1/2) + \sqrt{\delta}, 1]$,*

$$|\widetilde{\Lambda}' \cap [\Lambda^1(\varepsilon; \eta, \mathfrak{A}')]^c \cap \Lambda^1(\varepsilon; \eta - \sqrt{\delta}, \mathfrak{A})| \leq \delta^{\frac{1}{12}}\,. \tag{10.2.24}$$

**Remark 10.2.4.** *The measure of the set* $[\Lambda^1(\varepsilon;\eta,\mathfrak{A})]^c$ *is smaller than* $\varepsilon^p$, *for any* $p$, *taking the exponent* $\tau$ *in* (10.2.15) *large enough. This is analogous to the situation described in remark 9.3.4.*

Approximate solution of (10.2.11). We now solve approximately the equation (10.2.11), that we write as

$$T(h'_{\mathbb{G}}) = g'_{\mathbb{G}}\,, \qquad T := J\bar{\omega}_\varepsilon\cdot\partial_\varphi + \mathfrak{W}(\varepsilon,\lambda,\varphi)\,, \tag{10.2.25}$$

where we regard $T$ as an unbounded operator of $L^2(\mathbb{T}^{|\mathbb{S}|},H_{\mathbb{G}})$. We first extend the operator $T$ to an unbounded linear operator $T^\sharp$ acting on the whole space

$$L^2(\mathbb{T}^{|\mathbb{S}|},H)\,,\quad H = H_{\mathbb{S}\cup\mathbb{F}}\oplus H_{\mathbb{G}} = H_{\mathbb{S}}\oplus H_{\mathbb{F}}\oplus H_{\mathbb{G}}\,,$$

(see (3.1.7)), by defining

$$\begin{aligned} T^\sharp &:= J\bar{\omega}_\varepsilon\cdot\partial_\varphi + \frac{D_V}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}\cup\mathbb{F}} + \frac{\mathfrak{c}}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}} + i_{\mathbb{G}}\mathfrak{W}(\varepsilon,\lambda,\varphi)\Pi_{\mathbb{G}} \\ &\overset{(10.2.6),(10.1.6)}{=} J\bar{\omega}_\varepsilon\cdot\partial_\varphi + \frac{D_V}{1+\varepsilon^2\lambda} + \frac{\mathfrak{c}}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}} + \mathfrak{R}(\varepsilon,\lambda,\varphi)\Pi_{\mathbb{G}} \end{aligned} \tag{10.2.26}$$

where $\mathfrak{c}>0$ is a positive constant that we fix according to (4.1.5), $i_{\mathbb{G}}$ is the canonical injection defined in (9.3.34) and $\mathfrak{R}(\varepsilon,\lambda,\varphi)$ is given in (10.1.6).

According to the decomposition $H = H_{\mathbb{S}\cup\mathbb{F}}\oplus H_{\mathbb{G}}$ the operator $T^\sharp$ is represented by the matrix of operators

$$T^\sharp = \begin{pmatrix} J\bar{\omega}_\varepsilon\cdot\partial_\varphi + \frac{D_V}{1+\varepsilon^2\lambda} + \frac{\mathfrak{c}}{1+\varepsilon^2\lambda}\Pi_{\mathbb{S}} & 0 \\ 0 & T \end{pmatrix}\,,\qquad T\Pi_{\mathbb{G}} = \Pi_{\mathbb{G}}T^\sharp\,, \tag{10.2.27}$$

and, according to the decomposition $H = H_{\mathbb{S}}\oplus H_{\mathbb{F}}\oplus H_{\mathbb{G}}$, and recalling (10.2.25), by the matrix of operators

$$T^\sharp = J\bar{\omega}_\varepsilon\cdot\partial_\varphi + \begin{pmatrix} \frac{D_V+\mathfrak{c}\,\mathrm{Id}}{1+\varepsilon^2\lambda} & 0 & 0 \\ 0 & \frac{D_V}{1+\varepsilon^2\lambda} & 0 \\ 0 & 0 & \mathfrak{W}(\varepsilon,\lambda,\varphi) \end{pmatrix}. \tag{10.2.28}$$

We look for an approximate solution of

$$T^\sharp h^{\sharp\prime} = g^{\sharp\,\prime}_{\mathbb{G}} \qquad \text{where} \qquad g^{\sharp\,\prime}_{\mathbb{G}} := i_{\mathbb{G}}g'_{\mathbb{G}}\,. \tag{10.2.29}$$

With this aim we apply the multiscale Proposition 4.1.5 (in case (4.1.1)-($i$) and (4.1.6)) to the operator

$$\mathcal{L}_r = (1+\varepsilon^2\lambda)T^\sharp \overset{(10.2.26)}{=} J\omega\cdot\partial_\varphi + D_V + \mathfrak{c}\Pi_{\mathbb{S}} + r = J\omega\cdot\partial_\varphi + X_r \tag{10.2.30}$$

where

$$\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon\,, \quad r = (1+\varepsilon^2\lambda)\mathfrak{R}\Pi_{\mathbb{G}}\,, \quad X_r = D_V + \mathfrak{c}\Pi_{\mathbb{S}} + r\,.$$

The operator $X_r$ belongs to the class $\mathfrak{C}(C_1, c_1)$ (see Definition 4.1.2-($i$)) for some positive constant $C_1, c_1 > 0$. Indeed $X_r$ is self-adjoint with respect to the scalar product $L^2(\mathbb{T}^{|\mathbb{S}|}, H)$ (argue as in Lemma 9.3.6) and $r = (1+\varepsilon^2\lambda)\mathfrak{R}\Pi_{\mathbb{G}}$ has off-diagonal decay by (10.1.6). More precisely, arguing as in Lemma 9.3.7, we obtain

$$|r|_{\mathrm{Lip},+,s} \overset{(3.3.24)}{\lesssim_s} (1+\varepsilon^2\lambda)|\mathfrak{R}|_{\mathrm{Lip},+,s}|\Pi_{\mathbb{G}}|_{\mathrm{Lip},s+\frac{1}{2}} \overset{(3.3.33)}{\lesssim_s} |\mathfrak{R}|_{\mathrm{Lip},+,s}\,, \tag{10.2.31}$$

and therefore, by (10.1.6), $|r|_{\mathrm{Lip},+,s_1} \leq C(s_1)\varepsilon^2$. Finally, recalling (10.2.30) and (10.2.28), according to the decomposition $H = H_{\mathbb{S}} \oplus H_{\mathbb{F}} \oplus H_{\mathbb{G}}$, we represent

$$\frac{X_r}{1+\varepsilon^2\lambda} = \begin{pmatrix} \dfrac{D_V + \mathfrak{c}\,\mathrm{Id}}{1+\varepsilon^2\lambda} & 0 & 0 \\ 0 & \dfrac{D_V}{1+\varepsilon^2\lambda} & 0 \\ 0 & 0 & \mathfrak{W}(\varepsilon,\lambda,\varphi) \end{pmatrix}$$

and we deduce by the assumption (10.2.9), and arguing as in Lemma 9.3.8, that

$$\mathfrak{d}_\lambda\Big(\frac{X_r}{1+\varepsilon^2\lambda}\Big) \leq -c\,\varepsilon^2\,\mathrm{Id}$$

for some constant $c > 0$. We have proved that $X_r$ is in $\mathfrak{C}(C_1, c_1)$ according to Definition 4.1.2-($i$), for suitable positive constants $C_1, c_1$.

Thus Proposition 4.1.5 implies the existence of $\varepsilon_0 > 0$ (depending only on fixed positive constants as $C_1, c_1, \gamma_1, \tau_1$), of closed subsets

$$\Lambda^2(\varepsilon;\eta,\mathfrak{A})\,,\ 1/2 \leq \eta \leq 1\,,\ \text{satisfying Properties}\, 1-3 \text{ of Proposition 4.1.5}\,, \tag{10.2.32}$$

and $\bar{N} \in \mathbb{N}$, such that, for all $\lambda \in \Lambda^2(\varepsilon;1,\mathfrak{A})$, $\forall N \geq \bar{N}$,

$$\forall s \geq s_0\,,\ |(T_N^\sharp)^{-1}|_{\mathrm{Lip},s} \leq C(s)N^{Q'}\big(N^{\varsigma(s-s_1)} + |\mathfrak{R}|_{\mathrm{Lip},+,s}\big) \tag{10.2.33}$$

$$|(T_N^\sharp)^{-1}|_{\mathrm{Lip},s_0} \leq C(s_0)N^{Q'}\,, \qquad |(T_N^\sharp)^{-1}|_{\mathrm{Lip},s_1} \leq C(s_1)N^{Q'} \tag{10.2.34}$$

(see (4.1.20), (4.1.23), (10.2.31)) where $Q'$ is defined in (10.2.1) and $(T_N^\sharp)^{-1}$ denotes

$i$) the right inverse of $\mathit{\Pi}_N(T^\sharp)_{|\mathcal{H}_{2N}}$ defined in Proposition 4.1.5 if $\bar{N} \leq N < N(\varepsilon)$,

$ii$) the inverse of $\mathit{\Pi}_N(T^\sharp)_{|\mathcal{H}_N}$ if $N \geq N(\varepsilon)$,

(10.2.35)

and $\mathcal{H}_N$ are the finite dimensional subspaces defined in (4.1.10). We lay the stress on the fact that $\bar{N}$ can be regarded as a fixed constant, being independent of $\varepsilon$. Note also that (10.2.34) is a straightforward consequence of (10.2.33), by (10.1.6).

Now, given $g'_{\mathbb{G}} \in L^2(\mathbb{T}^{|\mathbb{S}|}, H_{\mathbb{G}})$, we define the following approximate solution of the equation (10.2.29),

$$h'_{\mathbb{G}} := \Pi_{\mathbb{G}} h^{\sharp'} , \qquad h^{\sharp'} := (T_N^\sharp)^{-1} \mathtt{g}_N^\sharp , \qquad \mathtt{g}_N^\sharp := \mathit{\Pi}_N {g_{\mathbb{G}}^\sharp}' , \quad {g_{\mathbb{G}}^\sharp}' := i_{\mathbb{G}} g'_{\mathbb{G}} , \tag{10.2.36}$$

and notice that, by (10.2.35),

$$h^{\sharp'} \in \mathcal{H}_{2N} \text{ if } N < N(\varepsilon) , \qquad \text{and} \qquad h^{\sharp'} \in \mathcal{H}_N \text{ if } N \geq N(\varepsilon) . \tag{10.2.37}$$

Finally, for all $\lambda$ in

$$\Lambda(\varepsilon; \eta, \mathfrak{A}) := \Lambda^1(\varepsilon; \eta, \mathfrak{A}) \cap \Lambda^2(\varepsilon; \eta, \mathfrak{A}) , \quad 1/2 \leq \eta \leq 1 , \tag{10.2.38}$$

where $\Lambda^{(1)}$, $\Lambda^{(2)}$ are the sets introduced in (10.2.15), (10.2.32), we define the following approximate right inverse $\mathcal{I}_D$ of the operator $\mathcal{L}_D$: given $g' \in L^2(\mathbb{T}^{|\mathbb{S}|}, H_{\mathbb{S}}^\perp)$, let

$$h' := \mathcal{I}_D g' := \Pi_N h'_{\mathbb{F}} + h'_{\mathbb{G}} , \qquad \Pi_N h'_{\mathbb{F}} := \sum_{j \in \mathbb{F}} (\mathit{\Pi}_N h'_j(\varphi)) \Psi_j(x) , \tag{10.2.39}$$

where $h'_{\mathbb{F}}$ is the solution of equation (10.2.10) given in Lemma 10.2.2, $h'_{\mathbb{G}}$ is defined in (10.2.36) and the projector $\mathit{\Pi}_N$ applies to functions depending only on the variable $\varphi$ as in (9.3.9).

**Lemma 10.2.5.** *The operator* $\mathcal{I}_D$ *defined in* (10.2.39) *satisfies* (10.2.2)-(10.2.3).

Proof. We first estimate the function $h'_{\mathbb{G}} = \Pi_{\mathbb{G}} h^{\sharp'} = \Pi_{\mathbb{G}} (T_N^\sharp)^{-1} \mathtt{g}_N^\sharp$ in (10.2.36). Since $|\Pi_{\mathbb{G}}|_{\mathrm{Lip},s} \leq C(s)$ by (3.3.33), the estimate (3.3.12) implies that, $\forall s \geq s_0$,

$$\begin{aligned} \|h'_{\mathbb{G}}\|_{\mathrm{Lip},s} &= \|\Pi_{\mathbb{G}} h^{\sharp'}\|_{\mathrm{Lip},s} = \|\Pi_{\mathbb{G}} (T_N^\sharp)^{-1} \mathtt{g}_N^\sharp\|_{\mathrm{Lip},s} && (10.2.40) \\ &\leq C \|(T_N^\sharp)^{-1} \mathtt{g}_N^\sharp\|_{\mathrm{Lip},s} + C(s) \|(T_N^\sharp)^{-1} \mathtt{g}_N^\sharp\|_{\mathrm{Lip},s_1} \\ &\overset{(3.3.12)}{\leq} C |(T_N^\sharp)^{-1}|_{\mathrm{Lip},s_1} \|\mathtt{g}_N^\sharp\|_{\mathrm{Lip},s} + C(s) |(T_N^\sharp)^{-1}|_{\mathrm{Lip},s} \|\mathtt{g}_N^\sharp\|_{\mathrm{Lip},s_1} \\ &\overset{(10.2.33),(10.2.34)}{\leq} C N^{Q'} \|\mathtt{g}_N^\sharp\|_{\mathrm{Lip},s} + C(s) N^{Q'} (N^{\varsigma(s-s_1)} + |\mathfrak{R}|_{\mathrm{Lip},+,s}) \|\mathtt{g}_N^\sharp\|_{\mathrm{Lip},s_1} \\ &\overset{(10.2.36),(4.1.13)}{\leq} C N^{Q'} \|g'_{\mathbb{G}}\|_{\mathrm{Lip},s} + C(s) N^{Q'} (N^{\varsigma(s-s_1)} + |\mathfrak{R}|_{\mathrm{Lip},+,s}) \|g'_{\mathbb{G}}\|_{\mathrm{Lip},s_1} && (10.2.41) \end{aligned}$$

where $C$ is a positive constant which depends on $s_1$. Moreover (10.2.40), (3.3.33), (3.3.8), (10.2.34), (10.2.36),(4.1.13) imply

$$\|h'_{\mathbb{G}}\|_{\mathrm{Lip},s_1} \lesssim_{s_1} \|h^{\sharp'}\|_{\mathrm{Lip},s_1} \lesssim_{s_1} N^{Q'} \|g'_{\mathbb{G}}\|_{\mathrm{Lip},s_1} , \tag{10.2.42}$$

$$\|h'_{\mathbb{G}}\|_{\mathrm{Lip},s_0} \lesssim_{s_0} \|h^{\sharp'}\|_{\mathrm{Lip},s_0} \lesssim_{s_0} N^{Q'} \|g'_{\mathbb{G}}\|_{\mathrm{Lip},s_0} . \tag{10.2.43}$$

In addition, Lemma 10.2.2 and the fact that $\Psi_j(x) \in C^\infty(\mathbb{T}^d)$ for all $j \in \mathbb{F}$ imply that $\Pi_N h'_{\mathbb{F}}$ defined in (10.2.39) satisfies, for $s \geq s_0$,

$$\begin{aligned}\|\Pi_N h'_{\mathbb{F}}\|_{\mathrm{Lip},s} &\overset{(3.5.2)}{\leq} \max_{j\in\mathbb{F}} \big( C\|\Psi_j\|_{s_1} \|\mathit{\Pi}_N h'_j\|_{\mathrm{Lip},H^s(\mathbb{T}^{|\mathbb{S}|})} + C(s)\|\Psi_j\|_s \|\mathit{\Pi}_N h'_j\|_{\mathrm{Lip},H^{s_1}(\mathbb{T}^{|\mathbb{S}|})} \big) \\ &\overset{(10.2.17)}{\leq} N^{2\tau} \big( C \max_{j\in\mathbb{F}} \|g'_j\|_{\mathrm{Lip},H^s(\mathbb{T}^{|\mathbb{S}|})} + C(s) \max_{j\in\mathbb{F}} \|g'_j\|_{\mathrm{Lip},H^{s_1}(\mathbb{T}^{|\mathbb{S}|})} \big) \\ &\leq N^{2\tau} \big( C\|g'_{\mathbb{F}}\|_{\mathrm{Lip},s} + C(s)\|g'_{\mathbb{F}}\|_{\mathrm{Lip},s_1} \big) \end{aligned} \tag{10.2.44}$$

using that $g'_j(\varphi) = \langle g'_{\mathbb{F}}(\varphi,\cdot), \Psi_j \rangle_{L^2_x}$, see (10.2.13).

In conclusion, by (10.2.41), (10.2.44), the fact that $2\tau < Q'$, (10.2.42), (10.2.43), (10.1.6), the function $h' := \mathcal{I}_D g'$ defined in (10.2.39) satisfies (10.2.3) and (10.2.2). ■

**Lemma 10.2.6.** *The operator $\mathcal{I}_D$ defined in* (10.2.39) *satisfies* (10.2.4).

PROOF. Recalling (10.2.39), that $\Psi_j(x) \in C^\infty(\mathbb{T}^d)$ for all $j \in \mathbb{F}$, and arguing as in (10.2.44), we obtain

$$\begin{aligned}\|\Pi_N h'_{\mathbb{F}}\|_{\mathrm{Lip},s_0} \lesssim_{s_0} \max_{j\in\mathbb{F}} \|h'_j\|_{\mathrm{Lip},H^{s_0}(\mathbb{T}^{|\mathbb{S}|})} &\overset{(10.2.16)}{\lesssim_{s_0}} \max_{j\in\mathbb{F}} \|g'_j\|_{\mathrm{Lip},H^{s_0+2\tau}(\mathbb{T}^{|\mathbb{S}|})} \\ &\lesssim_{s_1} \|g'_{\mathbb{F}}\|_{\mathrm{Lip},s_0+Q'}\end{aligned}$$

where $Q' = 2(\tau' + \varsigma s_1) + 3$ is defined in (10.2.1). Hence, in order to prove (10.2.4), it is sufficient to show that

$$\|h'_{\mathbb{G}}\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \|g'_{\mathbb{G}}\|_{\mathrm{Lip},s_0+Q'} \,. \tag{10.2.45}$$

We use a dyadic decomposition argument. First, given an integer $N > \bar{N}$, we define a sequence $(M_p)_{0\leq p\leq q}$ of positive integers, $q \geq 1$, by

$$M_0 := \bar{N}\,, \quad M_p := 2M_{p-1}\,, \forall p \in [\![1, q-1]\!] \quad \text{and} \quad 2M_{q-1} \leq M_q := N < 4M_{q-1}\,,$$

so that

$$[\![0, N]\!] = [\![0, M_0]\!] \cup \ldots \cup [\![M_{q-1}, M_q]\!]\,.$$

For $0 \leq p \leq q$, we set $\mathit{\Pi}_p := \mathit{\Pi}_{M_p}$ and we define the dyadic projectors

$$\Delta_0 := \mathit{\Pi}_0\,, \quad \Delta_p := \mathit{\Pi}_p - \mathit{\Pi}_{p-1} = \mathit{\Pi}_p \mathit{\Pi}^\perp_{p-1}\,, \ \forall p \in [\![1, q]\!]\,, \quad \mathit{\Pi}^\perp_{p-1} := \mathrm{Id} - \mathit{\Pi}_{p-1}\,.$$

For any function $h \in \mathcal{H}_N$ (recall that $N = M_q$) we consider its dyadic decomposition

$$h = \sum_{p=0}^{q} h_p \qquad \text{where} \qquad h_p := \Delta_p h\,, \tag{10.2.46}$$

and, for $0 \le p \le q$, we denote

$$H_p := \sum_{\ell=0}^{p} h_\ell = \Pi_p h \,. \tag{10.2.47}$$

In order to estimate $h'_{\mathbb{G}} := \Pi_{\mathbb{G}} h^{\sharp'}$ we recall that, by (10.2.36) and (10.2.35), we have

$$\Pi_N T^\sharp h^{\sharp'} = \mathsf{g}_N^\sharp \,, \quad \forall N \ge \bar N \,, \tag{10.2.48}$$

and the function $h^{\sharp'}$ satisfies (10.2.37). The key estimate is the following:

- Let $g_p^\sharp := \Delta_p \mathsf{g}_N^\sharp$. Then, each function $h_p^{\sharp'} = \Delta_p h^{\sharp'}$, $p \in [\![0, q]\!]$, satisfies

$$\begin{aligned} \| h_p^{\sharp'} \|_{\mathrm{Lip}, s_0} \lesssim_{s_1} & \ \bar N^{Q'} \big\| g_{p-1}^\sharp + g_p^\sharp + g_{p+1}^\sharp + g_{p+2}^\sharp \big\|_{\mathrm{Lip}, s_0 + Q'} \\ & + M_p^{-1} \big( \| h^{\sharp'} \|_{\mathrm{Lip}, s_0} + \| g'_{\mathbb{G}} \|_{\mathrm{Lip}, s_0} \big) \end{aligned} \tag{10.2.49}$$

  with the convention that $g_l^\sharp := 0$ for $l < 0$ or $l > q$.

We split the proof of (10.2.49) in different cases.

**Case I:** $0 \le p \le q-3$ and $M_{p+1} \le N(\varepsilon)$. Applying in (10.2.48) the projectors $\Pi_{p+1}$ and $\Pi_{p+2}$ for $0 \le p \le q-3$, and using the splitting $\mathrm{Id} = \Pi_{p+1} + \Pi_{p+1}^\perp$, we get

$$\Pi_{p+2} T^\sharp \Pi_{p+1} h^{\sharp'} + \Pi_{p+2} T^\sharp \Pi_{p+1}^\perp h^{\sharp'} = G_{p+2}^\sharp \qquad \text{where} \qquad G_{p+2}^\sharp := \Pi_{p+2} \mathsf{g}_N^\sharp \,, \tag{10.2.50}$$

and therefore

$$\mathcal{T}_{p+1}^\sharp H_{p+1}^{\sharp'} = G_{p+2}^\sharp - \Pi_{p+2} T^\sharp \Pi_{p+1}^\perp h^{\sharp'} \tag{10.2.51}$$

where $\mathcal{T}_{p+1}^\sharp := \Pi_{p+2} (T^\sharp)_{|\mathcal{H}_{M_{p+1}}}$ and $H_{p+1}^{\sharp'} := \Pi_{p+1} h^{\sharp'}$.

We claim that $\mathcal{T}_{p+1}^\sharp$ has a left inverse. Indeed, by (10.2.32)-(10.2.35) (applied with $M_{p+1}$ instead of $N$) each operator $\Pi_{p+1} (T^\sharp)_{|\mathcal{H}_{M_{p+2}}}$ has a right inverse

$$R_{p+1} := (T_{M_{p+1}}^\sharp)^{-1} : \mathcal{H}_{M_{p+1}} \mapsto \mathcal{H}_{M_{p+2}} \,.$$

Taking the adjoints in the identity $\Pi_{p+1} T^\sharp R_{p+1} = \mathrm{Id}_{\mathcal{H}_{M_{p+1}}}$, we obtain

$$R_{p+1}^* \mathcal{T}_{p+1}^\sharp = \mathrm{Id}_{\mathcal{H}_{M_{p+1}}} \qquad \text{where} \qquad \mathcal{T}_{p+1}^\sharp = \Pi_{p+2} (T^\sharp)_{|\mathcal{H}_{M_{p+1}}} \,,$$

i.e. $R_{p+1}^* : \mathcal{H}_{M_{p+2}} \to \mathcal{H}_{M_{p+1}}$ is a left inverse of $\mathcal{T}_{p+1}^\sharp$, that we denote by $(\mathcal{T}_{p+1}^\sharp)^{-1} := R_{p+1}^*$. By (3.3.13) and since $R_{p+1} = (T_{M_{p+1}}^\sharp)^{-1}$ satisfies (10.2.34) we deduce that

$$\big| (\mathcal{T}_{p+1}^\sharp)^{-1} \big|_{\mathrm{Lip}, s_1} \le C(s_1) M_{p+1}^{Q'} \,. \tag{10.2.52}$$

Applying the left inverse $(\mathcal{T}^\sharp_{p+1})^{-1}$ in (10.2.51) we deduce that $H^{\sharp'}_{p+1} = \Pi_{p+1} h^{\sharp'}$ may be expressed as

$$H^{\sharp'}_{p+1} = (\mathcal{T}^\sharp_{p+1})^{-1} G^\sharp_{p+2} - (\mathcal{T}^\sharp_{p+1})^{-1} \Pi_{p+2} T^\sharp \Pi^\perp_{p+1} h^{\sharp'} .$$

Finally, applying the projector $\Delta_p$ we get, setting $g^\sharp_p := \Delta_p \mathtt{g}^\sharp_N$,

$$\begin{aligned} h^{\sharp'}_p := \Delta_p H^{\sharp'}_{p+1} = \Delta_p (\mathcal{T}^\sharp_{p+1})^{-1} (g^\sharp_{p-1} + g^\sharp_p + g^\sharp_{p+1} + g^\sharp_{p+2}) & \quad (10.2.53) \\ + \Delta_p (\mathcal{T}^\sharp_{p+1})^{-1} G^\sharp_{p-2} & \quad (10.2.54) \\ - \Delta_p (\mathcal{T}^\sharp_{p+1})^{-1} \Pi_{p+2} T^\sharp \Pi^\perp_{p+1} h^{\sharp'} . & \quad (10.2.55) \end{aligned}$$

For $p = 0, 1$ the previous formula holds with $g^\sharp_{-1} := 0$ and $G^\sharp_{-2} := G^\sharp_{-1} := 0$. In particular for $p = 0, 1$ the term (10.2.54) is not present.

We now estimate separately the terms in (10.2.53)-(10.2.55).

Estimate of (10.2.53). By (3.3.8), (10.2.52) and since $M_{p+1} = 2M_p$ we have, for $p = 0, \ldots, q-3$,

$$\begin{aligned} & \|\Delta_p (\mathcal{T}^\sharp_{p+1})^{-1} (g^\sharp_{p-1} + g^\sharp_p + g^\sharp_{p+1} + g^\sharp_{p+2})\|_{\mathrm{Lip}, s_0} \\ & \lesssim_{s_1} M_p^{Q'} \|g^\sharp_{p-1} + g^\sharp_p + g^\sharp_{p+1} + g^\sharp_{p+2}\|_{\mathrm{Lip}, s_0} \\ & \lesssim_{s_1} \bar N^{Q'} \|g^\sharp_{p-1} + g^\sharp_p + g^\sharp_{p+1} + g^\sharp_{p+2}\|_{\mathrm{Lip}, s_0 + Q'} . \qquad (10.2.56) \end{aligned}$$

Note that we make appear the multiplicative constant $\bar N^{Q'}$ to deal with the cases $p = 0, 1$, where $g^\sharp_0$ is in the last sum (we use $M_0, M_1 \le 4\bar N$). Notice that the constant $C(s_1)$ in (10.2.56) does not depend on $\bar N$.

Estimate of (10.2.54). Denoting

$$B_p := \left\{ i \in \mathbb{Z}^{|\mathbb{S}|+d} \ : \ |i| \le M_p \right\} , \qquad (10.2.57)$$

we have, if $2 \le p \le q-3$ (for $p = 0, 1$ this term is not present)

$$\begin{aligned} \|\Delta_p (\mathcal{T}^\sharp_{p+1})^{-1} G^\sharp_{p-2}\|_{\mathrm{Lip}, s_0} & \overset{(3.3.8)}{\lesssim_{s_0}} \big|\Delta_p (\mathcal{T}^\sharp_{p+1})^{-1}_{|\mathcal{H}_{M_{p-2}}}\big|_{\mathrm{Lip}, s_0} \|G^\sharp_{p-2}\|_{\mathrm{Lip}, s_0} \\ & \overset{(3.3.14),(10.2.50)}{\lesssim_{s_0}} \mathtt{d}(B_{p-2}, B^c_{p-1})^{-(s_1 - s_0)} \big|\Delta_p (\mathcal{T}^\sharp_{p+1})^{-1}_{|\mathcal{H}_{M_{p-2}}}\big|_{\mathrm{Lip}, s_1} \|\mathtt{g}^\sharp_N\|_{\mathrm{Lip}, s_0} \\ & \overset{(10.2.57),(10.2.52),(10.2.36)}{\lesssim_{s_1}} M_p^{-(s_1 - s_0)} M_p^{Q'} \|g'_{\mathbb{G}}\|_{\mathrm{Lip}, s_0} , \qquad (10.2.58) \end{aligned}$$

keeping in mind that $M_{l+1} = 2M_l$ for $0 \le l \le p$.

Estimate of (10.2.55). By (3.3.8) we have

$$\begin{aligned}
&\|\Delta_p(\mathcal{T}_{p+1}^\sharp)^{-1}\Pi_{p+2}T^\sharp\Pi_{p+1}^\perp h^{\sharp\prime}\|_{\mathrm{Lip},s_0} \\
&\lesssim_{s_0} \big|\Delta_p(\mathcal{T}_{p+1}^\sharp)^{-1}\Pi_{p+2}T^\sharp\Pi_{p+1}^\perp\big|_{\mathrm{Lip},s_0}\|h^{\sharp\prime}\|_{\mathrm{Lip},s_0} \\
&\overset{(3.3.14)}{\lesssim_{s_0}} \mathtt{d}(B_p,B_{p+1}^c)^{-(s_1-s_0)}\big|(\mathcal{T}_{p+1}^\sharp)^{-1}\Pi_{p+2}T^\sharp\Pi_{p+1}^\perp\big|_{\mathrm{Lip},s_1}\|h^{\sharp\prime}\|_{\mathrm{Lip},s_0} \\
&\overset{(10.2.57),(10.2.52)}{\lesssim_{s_1}} M_p^{-(s_1-s_0)}M_p^{Q'+1}\|h^{\sharp\prime}\|_{\mathrm{Lip},s_0}\,, \qquad (10.2.59)
\end{aligned}$$

using that $|\Pi_{p+2}T^\sharp|_{\mathrm{Lip},s_1}\lesssim_{s_1} M_{p+2}\lesssim_{s_1} M_p$. Now, since $s_1$ is large according to (4.3.6), and $Q'=2(\tau'+\varsigma s_1)+3$, we have that $Q'-(s_1-s_0)<-1$ and by (10.2.53)-(10.2.55) and (10.2.56), (10.2.58), (10.2.59) we obtain the estimate (10.2.49), for any $0\le p\le q-3$ and $M_{p+1}\le N(\varepsilon)$.

**Case II:** $0\le p\le q-3$ and $M_{p+1}>N(\varepsilon)$. We have just to replace (10.2.50) with

$$\Pi_{p+1}T^\sharp\Pi_{p+1}h^{\sharp\prime}+\Pi_{p+1}T^\sharp\Pi_{p+1}^\perp h^{\sharp\prime}=G_{p+1}^\sharp$$

and apply the inverse $(\mathcal{T}_{p+1}^\sharp)^{-1}$ of $\mathcal{T}_{p+1}^\sharp:=T_{M_{p+1}}^\sharp$ (recall that by (10.2.35) this operator admits an inverse), which satisfies (10.2.34). Then $H_{p+1}^{\sharp\prime}:=\Pi_{p+1}h^{\sharp\prime}$ satisfies

$$H_{p+1}^{\sharp\prime}=(\mathcal{T}_{p+1}^\sharp)^{-1}G_{p+1}^\sharp-(\mathcal{T}_{p+1}^\sharp)^{-1}\Pi_{p+1}T^\sharp\Pi_{p+1}^\perp h^{\sharp\prime}$$

and, applying $\Delta_p$, we derive the estimate (10.2.49) in the same way (notice that since the functions $g_{p-1}^\sharp,\dots,g_{p+2}^\sharp$ are orthogonal for the scalar product associated to the $s_0$-norm, we have $\|g_{p-1}^\sharp+\ldots+g_{p+1}^\sharp\|_{\mathrm{Lip},s_0}\le\|g_{p-1}^\sharp+\ldots+g_{p+1}^\sharp+g_{p+2}^\sharp\|_{\mathrm{Lip},s_0}$).

**Case III:** $q-2\le p\le q$. By (10.2.36) we write

$$\begin{aligned}
h_p^{\sharp\prime} &= \Delta_p h^{\sharp\prime}=\Delta_p(T_N^\sharp)^{-1}\mathtt{g}_N^\sharp \\
&= \Delta_p(T_N^\sharp)^{-1}(g_{p-1}^\sharp+\ldots+g_q^\sharp)+\Delta_p(T_N^\sharp)^{-1}\Pi_{p-2}\mathtt{g}_N^\sharp\,.
\end{aligned}$$

Now recalling (10.2.57), and since $p-2\ge q-4$, $M_q=N$, we have that

$$\mathtt{d}(B_{p-1}^c,B_{p-2})\ge M_{p-2}\ge N/32\,.$$

Hence, arguing as above, we get by (10.2.34), (3.3.14), (3.3.8),

$$\begin{aligned}
\|h_p^{\sharp\prime}\|_{\mathrm{Lip},s_0} &\lesssim_{s_1} N^{Q'}(\|g_{p-1}^\sharp\|_{\mathrm{Lip},s_0}+\ldots+\|g_q^\sharp\|_{\mathrm{Lip},s_0})+N^{-(s_1-s_0)}|(T_N^\sharp)^{-1}|_{\mathrm{Lip},s_1}\|\mathtt{g}_N^\sharp\|_{\mathrm{Lip},s_0} \\
&\lesssim_{s_1} \big(\|g_{p-1}^\sharp\|_{\mathrm{Lip},s_0+Q'}+\ldots+\|g_q^\sharp\|_{\mathrm{Lip},s_0+Q'}\big)+N^{-1}\|g_{\mathbb{G}}'\|_{\mathrm{Lip},s_0}\,.
\end{aligned}$$

In conclusion, since $N^{-1}\le M_p^{-1}$, the estimate (10.2.49) is proved for any $p\in[\![0,q]\!]$.

By (10.2.49) we have

$$\|h_p^{\sharp'}\|^2_{\mathrm{Lip},s_0} \lesssim_{s_1} \bar N^{2Q'}\big(\|g^\sharp_{p-1}\|^2_{\mathrm{Lip},s_0+Q'} + \|g^\sharp_p\|^2_{\mathrm{Lip},s_0+Q'} + \|g^\sharp_{p+1}\|^2_{\mathrm{Lip},s_0+Q'} + \|g^\sharp_{p+2}\|^2_{\mathrm{Lip},s_0+Q'}\big) \\ + M_p^{-2}\|h^{\sharp'}\|^2_{\mathrm{Lip},s_0} + M_p^{-2}\|g'_{\mathbb G}\|^2_{\mathrm{Lip},s_0}\,. \tag{10.2.60}$$

Taking the sum for $0 \le p \le q$ of (10.2.60), recalling that $g_p^\sharp := \Delta_p \mathsf{g}_N^\sharp$, the definition of $\mathsf{g}_N^\sharp$ in (10.2.36), and that $M_0 = \bar N$, we obtain

$$\|h^{\sharp'}\|^2_{\mathrm{Lip},s_0} \lesssim_{s_1} \bar N^{2Q'}\|g'_{\mathbb G}\|^2_{\mathrm{Lip},s_0+Q'} + \bar N^{-2}\|g'_{\mathbb G}\|^2_{\mathrm{Lip},s_0} + \bar N^{-2}\|h^{\sharp'}\|^2_{\mathrm{Lip},s_0}\,.$$

For $\bar N^{-1} \le \delta(s_1)$ small enough, this implies

$$\|h^{\sharp'}\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \bar N^{Q'}\|g'_{\mathbb G}\|_{\mathrm{Lip},s_0+Q'}\,.$$

Then the function $h'_{\mathbb G} := \Pi_{\mathbb G} h^{\sharp'}$ satisfies, by Lemma 3.3.8, the estimate

$$\|h'_{\mathbb G}\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \bar N^{Q'}\|g'_{\mathbb G}\|_{\mathrm{Lip},s_0+Q'}\,.$$

Since $\bar N$ is a fixed constant, depending on $s_1$, this inequality implies (10.2.45). The proof of the lemma is complete. ∎

We now prove that $\mathcal I_D$ is an approximate right inverse of $\mathcal L_D$ satisfying (10.2.5).

**Lemma 10.2.7.** (10.2.5) *holds.*

PROOF. We have to estimate

$$(\mathcal L_D \mathcal I_D - \mathrm{Id})g' \overset{(10.2.39)}{=} \mathcal L_D h' - g' \overset{(10.1.7)}{=} (J\bar\omega_\varepsilon\cdot\partial_\varphi + \mathfrak A)h' - g'\,. \tag{10.2.61}$$

Recalling the definition of $h' := \mathcal I_D g'$ in (10.2.39), (10.2.10)-(10.2.11), Lemma 10.2.2, (10.2.25) (10.2.27), (10.2.35), (10.2.36), we have

$$(J\bar\omega_\varepsilon\cdot\partial_\varphi + \mathfrak A)h' - g' = \begin{pmatrix} -\Pi_N^\perp g'_{\mathbb F} \\ \Pi_{\mathbb G} \mathit{\Pi}_N^\perp\big(T^\sharp h^{\sharp'} - g_{\mathbb G}^{\sharp\,'}\big)\end{pmatrix}\,. \tag{10.2.62}$$

Hence by (10.2.61) and (10.2.62) we have

$$\begin{aligned}\big\|\big(J\bar\omega_\varepsilon\cdot\partial_\varphi + \mathfrak A\big)h' - g'\big\|_{\mathrm{Lip},s_1} &\lesssim_{s_1} \|\Pi_N^\perp g'_{\mathbb F}\|_{\mathrm{Lip},s_1} + \|\mathit{\Pi}_N^\perp(T^\sharp h^{\sharp'} - g_{\mathbb G}^{\sharp\,'})\|_{\mathrm{Lip},s_1} \\ &\lesssim_{s_1} \|\Pi_N^\perp g'_{\mathbb F}\|_{\mathrm{Lip},s_1} + \|\mathit{\Pi}_N^\perp T^\sharp h^{\sharp'}\|_{\mathrm{Lip},s_1} + \|\mathit{\Pi}_N^\perp g_{\mathbb G}^{\sharp\,'}\|_{\mathrm{Lip},s_1} \\ &\lesssim_{s_1} N^{-(s_3-s_1)}\|g'\|_{\mathrm{Lip},s_3} + N^{-(s_3-s_1)}\|\mathit{\Pi}_N^\perp T^\sharp h^{\sharp'}\|_{\mathrm{Lip},s_3}\end{aligned} \tag{10.2.63}$$

by the smoothing property (4.1.13).

Recalling the form of $T^\sharp$ in (10.2.26), and writing $D_V = D_m + (D_V - D_m)$, we obtain

$$\begin{aligned} \Pi_N^\perp T^\sharp h^{\sharp\prime} &= \Pi_N^\perp T^\sharp \Pi_{2N} h^{\sharp\prime} \\ &= \Pi_N^\perp J\bar\omega_\varepsilon \cdot \partial_\varphi \Pi_{2N} h^{\sharp\prime} + \frac{1}{1+\varepsilon^2\lambda}\big(\Pi_N^\perp D_m \Pi_{2N} h^{\sharp\prime} + \Pi_N^\perp (D_V - D_m) h^{\sharp\prime}\big) \\ &\quad + \frac{1}{1+\varepsilon^2\lambda}\Pi_N^\perp \mathfrak{c}\Pi_{\mathbb{S}} \Pi_{2N} h^{\sharp\prime} + \Pi_N^\perp \mathfrak{R}\Pi_{\mathbb{G}} h^{\sharp\prime} . \end{aligned} \tag{10.2.64}$$

Notice that, by (10.2.37), if $N \ge N(\varepsilon)$ then $h^{\sharp\prime} \in \mathcal{H}_N$ and the first and second terms in (10.2.64) are zero.

In conclusion, by (10.2.63), (10.2.64), using (3.3.8), (3.3.33), we get

$$\begin{aligned} & \|(J\bar\omega_\varepsilon \cdot \partial_\varphi + \mathfrak{A})h' - g'\|_{\mathrm{Lip},s_1} \\ & \lesssim_{s_3} N^{-(s_3-s_1)}\big(\|g'\|_{\mathrm{Lip},s_3} + \|(D_V - D_m)h^{\sharp\prime}\|_{\mathrm{Lip},s_3} + \|\mathfrak{R}\Pi_{\mathbb{G}} h^{\sharp\prime}\|_{\mathrm{Lip},s_3} + N\|h^\sharp\|_{\mathrm{Lip},s_3}\big) \\ & \overset{(3.4.1),(10.1.6)}{\lesssim_{s_3}} N^{-(s_3-s_1)}\big(\|g'\|_{\mathrm{Lip},s_3} + N\|h^{\sharp\prime}\|_{\mathrm{Lip},s_3} + |\mathfrak{R}|_{\mathrm{Lip},s_3}\|h^{\sharp\prime}\|_{\mathrm{Lip},s_1}\big) \\ & \overset{(10.2.40),(10.2.41),(10.2.42)}{\lesssim_{s_3}} N^{-(s_3-s_1)} N^{Q'+1}\big(\|g'\|_{\mathrm{Lip},s_3} + (N^{\varsigma(s_3-s_1)} + |\mathfrak{R}|_{\mathrm{Lip},+,s_3})\|g'\|_{\mathrm{Lip},s_1}\big)\big) \end{aligned} \tag{10.2.65}$$

(and recall that $g'_{\mathbb{G}} = \Pi_{\mathbb{G}} g'$ by (10.2.12)). Recalling (10.2.61), the estimate (10.2.65) proves (10.2.5). ∎

Now by Lemma 10.2.3 and (10.2.32), $\Lambda^1(\varepsilon;\eta,\mathfrak{A})$ and $\Lambda^2(\varepsilon;\eta,\mathfrak{A})$ satisfy properties 1-3 of Proposition 4.1.5. Since these properties are preserved under finite intersection, this implies the following lemma.

**Lemma 10.2.8.** *The sets $\Lambda(\varepsilon;\eta,\mathfrak{A})$ defined in (10.2.38) satisfy properties 1-3 of Proposition 4.1.5.*

## 10.3 Approximate right inverse of $\mathcal{L} = \mathcal{L}_D + \varrho$

In this section we construct an approximate right inverse of $\mathcal{L} = \mathcal{L}_D + \varrho$ by a perturbative Neumann series argument, for $\varrho \in \mathcal{L}(H_{\mathbb{S}}^\perp)$ small, using the approximate right inverse $\mathcal{I}_D$ of $\mathcal{L}_D$ found in Proposition 10.2.1.

We denote

$$U(s) := U_{\mathfrak{A},\varrho}(s) := |\mathfrak{R}|_{\mathrm{Lip},+,s} + |\varrho|_{\mathrm{Lip},+,s} . \tag{10.3.1}$$

**Proposition 10.3.1. (Approximate right inverse of $\mathcal{L}$).** *There is $c_0 > 0$ (depending on $s_1$) such that, for $N \ge \bar N$, if $\varrho$ satisfies*

$$|\varrho|_{\mathrm{Lip},s_1} N^{Q'} = |\varrho|_{\mathrm{Lip},s_1} N^{2(\tau'+\varsigma s_1)+3} \le c_0 , \tag{10.3.2}$$

*then* $\forall \lambda \in \Lambda(\varepsilon;\eta,\mathfrak{A})$, $1/2 \le \eta \le 1$ *(where the set* $\Lambda(\varepsilon;\eta,\mathfrak{A})$ *is defined in Proposition 10.2.1), the operator* $\mathrm{Id} + \mathcal{I}_D \varrho$ *is invertible and*

$$\mathcal{I} := \mathcal{I}_N := (\mathrm{Id} + \mathcal{I}_D \varrho)^{-1} \mathcal{I}_D \tag{10.3.3}$$

*is an approximate right inverse of* $\mathcal{L}$*, in the sense that*

$$\begin{aligned} &\|(\mathcal{L}\mathcal{I} - \mathrm{Id})g'\|_{\mathrm{Lip},s_1} \lesssim_{s_3} \\ &N^{Q'+1-(s_3-s_1)}\big(\|g'\|_{\mathrm{Lip},s_3} + (N^{\varsigma(s_3-s_1)} + N^{Q'}U(s_3))\|g'\|_{\mathrm{Lip},s_1}\big)\,. \end{aligned} \tag{10.3.4}$$

*Moreover the operator* $\mathcal{I}$ *satisfies*

$$\|\mathcal{I}g'\|_{\mathrm{Lip},s_1} \lesssim_{s_1} N^{Q'}\|g'\|_{\mathrm{Lip},s_1}\,, \tag{10.3.5}$$

$$\|\mathcal{I}g'\|_{\mathrm{Lip},s} \le CN^{Q'}\|g'\|_{\mathrm{Lip},s} + C(s)N^{Q'}\big(N^{\varsigma(s-s_1)} + N^{Q'}U(s)\big)\|g'\|_{\mathrm{Lip},s_1}\,, \quad \forall s \ge s_1\,, \tag{10.3.6}$$

*where the constant* $C$ *is independent of* $s$ *(it depends on* $s_1$*). Furthermore*

$$\|\mathcal{I}g'\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \|g'\|_{\mathrm{Lip},s_0+Q'}\,. \tag{10.3.7}$$

The proof of Proposition 10.3.1 is given in the rest of this section.

We first justify that the operator $\mathrm{Id}+\mathcal{I}_D\varrho$ is invertible and provide appropriate estimates for its inverse, as an application of Lemma 3.3.13. By (10.2.2) and (3.3.8) we have

$$\begin{aligned} \|\mathcal{I}_D\varrho h'\|_{\mathrm{Lip},s_0} &\lesssim_{s_0} N^{Q'}\|\varrho h'\|_{\mathrm{Lip},s_0} \lesssim_{s_0} N^{Q'}|\varrho|_{\mathrm{Lip},s_0}\|h'\|_{\mathrm{Lip},s_0}\,, \\ \|\mathcal{I}_D\varrho h'\|_{\mathrm{Lip},s_1} &\lesssim_{s_1} N^{Q'}\|\varrho h'\|_{\mathrm{Lip},s_1} \lesssim_{s_1} N^{Q'}|\varrho|_{\mathrm{Lip},s_1}\|h'\|_{\mathrm{Lip},s_1}\,, \end{aligned} \tag{10.3.8}$$

and, by (10.2.3), (10.3.1),

$$\begin{aligned} \|\mathcal{I}_D\varrho h'\|_{\mathrm{Lip},s} &\le CN^{Q'}\|\varrho h'\|_{\mathrm{Lip},s} + C(s)N^{Q'}\big(N^{\varsigma(s-s_1)} + U(s)\big)\|\varrho h'\|_{\mathrm{Lip},s_1} \\ &\overset{(3.3.12)}{\le} CN^{Q'}|\varrho|_{\mathrm{Lip},s_1}\|h'\|_{\mathrm{Lip},s} + C(s)N^{Q'}|\varrho|_{\mathrm{Lip},s}\|h'\|_{\mathrm{Lip},s_1} \\ &\quad + C(s)N^{Q'}|\varrho|_{\mathrm{Lip},s_1}\big(N^{\varsigma(s-s_1)} + U(s)\big)\|h'\|_{\mathrm{Lip},s_1} \end{aligned} \tag{10.3.9}$$

where $C = C(s_1)$. Hence, there is a positive constant $c_0$ (depending on $s_1$) such that, if

$$|\varrho|_{\mathrm{Lip},s_1}N^{Q'} \le c_0\,, \tag{10.3.10}$$

then the operator $\mathcal{I}_D\varrho$ satisfies, by (10.3.8), (10.3.9) and recalling also (10.3.1),

$$\|\mathcal{I}_D\varrho h'\|_{\mathrm{Lip},s_0} \le \frac{1}{2}\|h'\|_{\mathrm{Lip},s_0} \tag{10.3.11}$$

$$\|\mathcal{I}_D\varrho h'\|_{\mathrm{Lip},s_1} \le \frac{1}{2}\|h'\|_{\mathrm{Lip},s_1} \tag{10.3.12}$$

$$\|\mathcal{I}_D\varrho h'\|_{\mathrm{Lip},s} \le \frac{1}{2}\|h'\|_{\mathrm{Lip},s} + C(s)\big(N^{\varsigma(s-s_1)} + N^{Q'}U(s)\big)\|h'\|_{\mathrm{Lip},s_1}\,. \tag{10.3.13}$$

By (10.3.11)-(10.3.13) and Lemma 3.3.13 (applied with $R = \mathcal{I}_D \varrho$ and $E = L^2(\mathbb{T}^{|\mathbb{S}|}, H_{\mathbb{S}}^{\perp})$), the operator $\mathrm{Id} + \mathcal{I}_D \varrho$ is invertible and its inverse satisfies the tame estimates

$$\|(\mathrm{Id} + \mathcal{I}_D\varrho)^{-1} g'\|_{\mathrm{Lip},s_0} \leq 2\|g'\|_{\mathrm{Lip},s_0} \tag{10.3.14}$$

$$\|(\mathrm{Id} + \mathcal{I}_D\varrho)^{-1} g'\|_{\mathrm{Lip},s_1} \leq 2\|g'\|_{\mathrm{Lip},s_1} \tag{10.3.15}$$

$$\|(\mathrm{Id} + \mathcal{I}_D\varrho)^{-1} g'\|_{\mathrm{Lip},s} \leq 2\|g'\|_{\mathrm{Lip},s} + C(s)\big(N^{\varsigma(s-s_1)} + N^{Q'} U(s)\big)\|g'\|_{\mathrm{Lip},s_1} \tag{10.3.16}$$

for all $s \geq s_1$. We now estimate the operator $\mathcal{I}$ defined in (10.3.3).

**Lemma 10.3.2.** *The operator* $\mathcal{I} = (\mathrm{Id} + \mathcal{I}_D\varrho)^{-1}\mathcal{I}_D$ *defined in* (10.3.3) *satisfies* (10.3.5)-(10.3.6) *and* (10.3.7).

Proof. By (10.3.15) and (10.2.2) we have

$$\|\mathcal{I}g'\|_{\mathrm{Lip},s_1} \leq 2\|\mathcal{I}_D g'\|_{\mathrm{Lip},s_1} \lesssim_{s_1} N^{Q'} \|g'\|_{\mathrm{Lip},s_1}$$

which is (10.3.5). In addition (10.3.16), (10.2.2), (10.2.3), (10.3.1), imply

$$\begin{aligned} \|\mathcal{I}g'\|_{\mathrm{Lip},s} &\leq 2\|\mathcal{I}_D g'\|_{\mathrm{Lip},s} + C(s)\big(N^{\varsigma(s-s_1)} + N^{Q'}U(s)\big)\|\mathcal{I}_D g'\|_{\mathrm{Lip},s_1} \\ &\leq C N^{Q'} \|g'\|_{\mathrm{Lip},s} + C(s) N^{Q'}\big(N^{\varsigma(s-s_1)} + N^{Q'}U(s)\big)\|g'\|_{\mathrm{Lip},s_1} \end{aligned}$$

proving (10.3.6). Finally, by (10.3.14),

$$\begin{aligned} \|\mathcal{I}g'\|_{\mathrm{Lip},s_0} = \|(\mathrm{Id} + \mathcal{I}_D\varrho)^{-1}\mathcal{I}_D g'\|_{\mathrm{Lip},s_0} &\leq 2\|\mathcal{I}_D g'\|_{\mathrm{Lip},s_0} \\ &\overset{(10.2.4)}{\lesssim_{s_1}} \|g'\|_{\mathrm{Lip},s_0+Q'} \end{aligned}$$

which is (10.3.7). ■

We now prove that $\mathcal{I}$ is an approximate right inverse of $\mathcal{L}$ satisfying (10.3.4).

**Lemma 10.3.3.** (10.3.4) *holds.*

Proof. Recalling that $\mathcal{L} = \mathcal{L}_D + \varrho$ by (10.1.7), and setting

$$h' := \mathcal{I}g' \overset{(10.3.3)}{=} (\mathrm{Id} + \mathcal{I}_D\varrho)^{-1}(\mathcal{I}_D g')\,, \tag{10.3.17}$$

we have

$$\begin{aligned} (\mathcal{L}\mathcal{I} - \mathrm{Id})g' &= \mathcal{L}_D h' + \varrho h' - g' \\ &= \mathcal{L}_D(\mathrm{Id} + \mathcal{I}_D\varrho)h' - \mathcal{L}_D\mathcal{I}_D\varrho h' + \varrho h' - g' \\ &= \mathcal{L}_D\mathcal{I}_D g' - \mathcal{L}_D\mathcal{I}_D\varrho h' + \varrho h' - g' \\ &= (\mathcal{L}_D\mathcal{I}_D - \mathrm{Id})(g' - \varrho h')\,. \end{aligned} \tag{10.3.18}$$

Then we estimate (10.3.18) as

$$
\begin{aligned}
\|(\mathcal{L}\mathcal{I}-\mathrm{Id})g'\|_{\mathrm{Lip},s_1} &= \|(\mathcal{L}_D\mathcal{I}_D-\mathrm{Id})(g'-\varrho h')\|_{\mathrm{Lip},s_1} \\
&\overset{(10.2.5),(10.3.1)}{\lesssim_{s_3}} N^{Q'+1-(s_3-s_1)}\Big(\|g'\|_{\mathrm{Lip},s_3}+\|\varrho h'\|_{\mathrm{Lip},s_3} \\
&\qquad + (N^{\varsigma(s_3-s_1)}+U(s_3))(\|g'\|_{\mathrm{Lip},s_1}+\|\varrho h'\|_{\mathrm{Lip},s_1})\Big) \\
&\overset{(3.3.8),(10.3.17),(10.3.5)}{\lesssim_{s_3}} N^{Q'+1-(s_3-s_1)}\Big(\|g'\|_{\mathrm{Lip},s_3}+|\varrho|_{\mathrm{Lip},s_1}\|h'\|_{\mathrm{Lip},s_3}+|\varrho|_{\mathrm{Lip},s_3}N^{Q'}\|g'\|_{\mathrm{Lip},s_1} \\
&\qquad + (N^{\varsigma(s_3-s_1)}+U(s_3))\big(\|g'\|_{\mathrm{Lip},s_1}+|\varrho|_{\mathrm{Lip},s_1}N^{Q'}\|g'\|_{\mathrm{Lip},s_1}\big)\Big) \\
&\overset{(10.3.17),(10.3.6),(10.3.2),(10.3.1)}{\lesssim_{s_3}} N^{Q'+1-(s_3-s_1)}\Big(\|g'\|_{\mathrm{Lip},s_3}+\big(N^{\varsigma(s_3-s_1)}+N^{Q'}U(s_3)\big)\|g'\|_{\mathrm{Lip},s_1}\Big)
\end{aligned}
$$

which proves (10.3.4). ■

## 10.4 Approximate right inverse of $\bar{\omega}_\varepsilon\cdot\partial_\varphi - J(A_0+\rho)$

In this section we complete the proof of Proposition 8.2.1. As said in section 10.1, by the hypotheses of Proposition 8.2.1, the assumptions of Corollary 9.1.2 are satisfied by $(A_0,\rho)$. Corollary 9.1.2 then provides, for all $\lambda$ in $\Lambda_\infty(\varepsilon;5/6,A_0,\rho)$, the conjugation (see (10.1.2))

$$\big(\bar{\omega}_\varepsilon\cdot\partial_\varphi - JA_0 - J\rho\big)\mathcal{P}_{\mathfrak{n}}(\varphi) = \mathcal{P}_{\mathfrak{n}}(\varphi)\big(\bar{\omega}_\varepsilon\cdot\partial_\varphi - JA_{\mathfrak{n}} - J\rho_{\mathfrak{n}}\big)\,,$$

where $A_{\mathfrak{n}}$ is a self-adjoint block diagonal operator of the form

$$A_{\mathfrak{n}} = \frac{D_V}{1+\varepsilon^2\lambda} + R_{\mathfrak{n}}\,,$$

see (9.1.25)-(9.1.26), satisfying (9.1.27)-(9.1.31), in particular $A_{\mathfrak{n}}\in\mathcal{C}(2C_1,c_1/2,c_2/2)$ is an admissible split operator according to Definition 8.1.1. The sequence $(\mathcal{P}_{\mathfrak{n}}(\varphi))$ of symplectic transformations satisfies (9.1.21)-(9.1.23) and $R_{\mathfrak{n}}$, $\rho_{\mathfrak{n}}$ satisfy (9.1.29)-(9.1.33); in particular

$$|\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s_1} \le \delta_1^{(\frac{3}{2})^{\mathfrak{n}}}\,. \tag{10.4.1}$$

We define, for $1/2\le\eta\le 5/6$, the sets

$$\boldsymbol{\Lambda}(\varepsilon;\eta,A_0,\rho) := \Lambda_\infty(\varepsilon;\eta,A_0,\rho)\bigcap\Big(\bigcap_{\mathfrak{n}\ge 0}\mathtt{\Lambda}(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}})\Big) \tag{10.4.2}$$

where $\Lambda_\infty$ is defined in Corollary 9.1.2, $\mathtt{\Lambda}$ in Proposition 10.2.1 (i.e. (10.2.38)) and the sequence $(\eta_{\mathfrak{n}})$ is defined in (9.1.37). Note that we can apply Proposition 10.2.1 with $\mathfrak{A}=A_{\mathfrak{n}}$

for any $\mathfrak{n} \geq 0$, because $A_{\mathfrak{n}} \in \mathcal{C}(2C_1, c_1/2, c_2/2)$. In Lemma 10.4.4 below we shall prove that the sets $\mathbf{\Lambda}(\varepsilon; \eta, A_0, \rho)$ satisfy properties 1-3 of Proposition 8.2.1.

Choice of the cut-off $N$ and of the number $\mathfrak{n}$ of splitting steps.
For any $\nu \in (0, \varepsilon)$, we choose $N \in \mathbb{N}$ such that

$$N \in \Big[\nu^{-\frac{3}{s_3 - s_1}} - \frac{1}{2}, \nu^{-\frac{3}{s_3 - s_1}} + \frac{1}{2}\Big) \tag{10.4.3}$$

and the number $\mathfrak{n} \in \mathbb{N}$ of splitting steps in Corollary 9.1.2 as

$$\mathfrak{n} := \min \Big\{ k \in \mathbb{N} \, : \, \delta_1^{(\frac{3}{2})^k} N^{Q'} = \delta_1^{(\frac{3}{2})^k} N^{2(\tau' + \varsigma s_1) + 3} \leq c_0 \Big\}$$

where $\delta_1 = \varepsilon^3$ (see (9.1.20) and (9.1.2)) and $c_0$ is the strictly positive constant of Proposition 10.3.1. Hence

$$\delta_1^{(\frac{3}{2})^{\mathfrak{n}}} N^{Q'} \leq c_0 \, , \qquad \text{and, if } \ \mathfrak{n} \geq 1 \, , \quad \delta_1^{(\frac{3}{2})^{\mathfrak{n}-1}} N^{Q'} > c_0 \, . \tag{10.4.4}$$

**Remark 10.4.1.** *In the sequel we suppose that $\nu$ is small enough, possibly depending on $s_3$ (the aim is to obtain estimates without any multiplicative constant depending on $s_3$). Since $\nu \in (0, \varepsilon)$, any smallness condition for $\nu$ is satisfied if $\varepsilon$ is small enough, possibly depending on $s_3$, which is a large, but fixed, constant. For this reason, often we will not explicitly indicate such dependence. However it is useful to point out that, for $\varepsilon$ fixed, our estimates still hold for $\nu$ small enough. In fact, in Proposition 8.2.2, we extend some estimates on the $s_3$-norm to estimates on the $s$-norm for any $s \geq s_3$, with a smallness condition on $\varepsilon$ which does not depend on $s$ but only on $s_3$, while the smallness condition on $\nu$ depends on $s$.*

In the sequel of this section the notation $a \ll b$ means that $a/b \to 0$ as $\nu \to 0$.
By (10.4.3) that, for $\nu$ small enough, $N \geq \bar{N}$.
By (10.4.3), (8.2.1) and recalling that $Q' = 2(\tau' + \varsigma s_1) + 3$ and $\varsigma = 1/10$, we have

$$N^{Q'+1} \ll \nu^{-\frac{1}{20}} \, , \quad N^{\varsigma(s_3 - s_1)} \ll \nu^{-\frac{2}{5}} \, , \quad N^{-(1-\varsigma)(s_3 - s_1)} \ll \nu^{\frac{5}{2}} \tag{10.4.5}$$

$$N^{Q'+1-(s_3 - s_1)} \leq \nu^{-\frac{1}{20}} N^{-(s_3 - s_1)} \lesssim_{s_3} \nu^{-\frac{1}{20}} \nu^3 \ll \nu^{\frac{11}{4}} \, . \tag{10.4.6}$$

Moreover, by (10.4.3), (10.4.4), (8.2.1), we have

$$\begin{aligned} &\delta_1^{(\frac{3}{2})^{\mathfrak{n}}} \lesssim \nu^{\frac{3Q'}{s_3 - s_1}} \ll \nu^{-\frac{1}{20}} \, , \quad \mathfrak{n} \leq \frac{1}{\ln(3/2)} \ln \ln \nu^{-1} \, , \\ &[C(s)]^{\mathfrak{n}} = e^{\mathfrak{n} \ln C(s)} \leq e^{\frac{\ln C(s)}{\ln(3/2)} \ln \ln \nu^{-1}} \leq \big( \ln \nu^{-1} \big)^{\frac{\ln C(s)}{\ln(3/2)}} \, , \end{aligned} \tag{10.4.7}$$

and, for $\mathfrak{n} \geq 1$,

$$\delta_1^{-(\frac{3}{2})^{\mathfrak{n}-1}} \lesssim \nu^{\frac{-3Q'}{s_3 - s_1}} \ll \nu^{-\frac{1}{20}} \, . \tag{10.4.8}$$

Recalling (10.3.1), we introduce the notation

$$U_{\mathfrak{n}}(s) = U_{A_{\mathfrak{n}},\rho_{\mathfrak{n}}}(s) = |R_{\mathfrak{n}}|_{\mathrm{Lip},+,s} + |\rho_{\mathfrak{n}}|_{\mathrm{Lip},+,s}\,. \tag{10.4.9}$$

The bounds (9.1.31), (9.1.33) where $\alpha(s) = 3\varsigma\dfrac{s-s_2}{s_2-s_1}$ and (10.4.7), (10.4.8) provide the estimate

$$\begin{aligned} U_{\mathfrak{n}}(s_3) &\lesssim_{s_3} 2(\ln(\nu^{-1}))^{\frac{\ln C(s_3)}{\ln(3/2)}}\nu^{-\frac{3Q'}{s_3-s_1}(\frac{3}{4}+3\varsigma\frac{s_3-s_2}{s_2-s_1})}\big[\big(|R_0|_{\mathrm{Lip},+,s_3}+|\rho|_{\mathrm{Lip},+,s_3}\big)\delta_1^{\frac{1}{2}+\frac{2\alpha(s_3)}{3}}+1\big] \\ &\overset{(8.2.5)}{\lesssim_{s_3}} 2(\ln(\nu^{-1}))^{\frac{\ln C(s_3)}{\ln(3/2)}}\nu^{-\frac{3Q'}{s_3-s_1}(\frac{3}{4}+3\varsigma\frac{s_3-s_2}{s_2-s_1})}\big[\nu^{-1}\varepsilon^2+1\big] \\ &\lesssim_{s_3} 2(\ln(\nu^{-1}))^{\frac{\ln C(s_3)}{\ln(3/2)}}\nu^{-\frac{9Q'}{4(s_3-s_1)}}\nu^{-\frac{9Q'\varsigma}{s_2-s_1}}\big[\nu^{-1}\varepsilon^2+1\big] \\ &\overset{(8.2.1)}{\leq} C(s_3)\nu^{-\frac{11}{10}}\,. \end{aligned} \tag{10.4.10}$$

The estimate (9.1.23) provides similarly

$$|\mathcal{P}_{\mathfrak{n}}^{\pm 1}|_{\mathrm{Lip},+,s_3} \le C(s_3)\nu^{-\frac{11}{10}}\,. \tag{10.4.11}$$

Solution of (8.2.7).

By (10.4.1) and (10.4.4), the smallness condition (10.3.2) is satisfied by $\varrho = \rho_{\mathfrak{n}}$. Then Proposition 10.3.1 applies to the operator $\mathcal{L} := J\bar{\omega}_\varepsilon \cdot \partial_\varphi + A_{\mathfrak{n}} + \rho_{\mathfrak{n}}$, implying the existence of an approximate right inverse $\mathcal{I} := \mathcal{I}_{N,\mathfrak{n}}$, defined for $\lambda \in \mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho)$ (see (10.4.2)), satisfying (10.3.4)-(10.3.7). The operator $\mathcal{I}$ satisfies

$$\|\mathcal{I}\widetilde{g}\|_{\mathrm{Lip},s_1} \overset{(10.3.5)}{\lesssim_{s_1}} N^{Q'}\|\widetilde{g}\|_{\mathrm{Lip},s_1} \overset{(10.4.5)}{\le} \nu^{-\frac{1}{20}}\|\widetilde{g}\|_{\mathrm{Lip},s_1} \tag{10.4.12}$$

and

$$\begin{aligned} \|\mathcal{I}\widetilde{g}\|_{\mathrm{Lip},s_3} &\overset{(10.3.6)}{\le} CN^{Q'}\|\widetilde{g}\|_{\mathrm{Lip},s_3} + C(s_3)\big(N^{\varsigma(s_3-s_1)} + N^{Q'}U_{\mathfrak{n}}(s_3)\big)\|\widetilde{g}\|_{\mathrm{Lip},s_1} \\ &\overset{(10.4.5),(10.4.10)}{\le} \nu^{-\frac{1}{20}}\|\widetilde{g}\|_{\mathrm{Lip},s_3} + \big(\nu^{-\frac{2}{5}} + \nu^{-\frac{1}{20}}\nu^{-\frac{11}{10}}\big)\|\widetilde{g}\|_{\mathrm{Lip},s_1} \\ &\le \nu^{-\frac{1}{20}}\|\widetilde{g}\|_{\mathrm{Lip},s_3} + \nu^{-\frac{6}{5}}\|\widetilde{g}\|_{\mathrm{Lip},s_1}\,. \end{aligned} \tag{10.4.13}$$

In addition estimates (10.3.4), (10.4.5)-(10.4.6), (10.4.10) give

$$\begin{aligned} \|(\mathcal{L}\mathcal{I} - \mathrm{Id})\widetilde{g}\|_{\mathrm{Lip},s_1} &\le \nu^{\frac{11}{4}}\Big(\|\widetilde{g}\|_{\mathrm{Lip},s_3} + \big(\nu^{-\frac{2}{5}} + \nu^{-\frac{1}{20}}\nu^{-\frac{11}{10}}\big)\|\widetilde{g}\|_{\mathrm{Lip},s_1}\Big) \\ &\le \nu^{\frac{11}{4}}\|\widetilde{g}\|_{\mathrm{Lip},s_3} + \nu^{\frac{3}{2}}\|\widetilde{g}\|_{\mathrm{Lip},s_1}\,. \end{aligned} \tag{10.4.14}$$

Now let $g \in H_{\mathbb{S}}^{\perp}$ satisfy the assumption (8.2.5) and consider the function $g' = \mathcal{P}_{\mathfrak{n}}^{-1}(\varphi)g$ introduced in (10.1.3). We define, for any $\lambda$ in $\mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho)$, the approximate solution

$h' = \mathcal{I}(Jg')$ of the equation (10.1.7) where $\mathcal{I}$ is the approximate right inverse, obtained in Proposition 10.3.1, of the operator $\mathcal{L} = J\bar{\omega}_\varepsilon \cdot \partial_\varphi + A_{\mathfrak{n}} + \rho_{\mathfrak{n}}$, and we consider the "approximate solution" $h$ of the equation (10.1.1) as

$$h := \mathcal{P}_{\mathfrak{n}}(\varphi)h' \qquad \text{where} \qquad h' := \mathcal{I}(Jg') \,, \quad g' = \mathcal{P}_{\mathfrak{n}}^{-1}(\varphi)g \,. \tag{10.4.15}$$

It means that we define the approximate right inverse of $\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho)$ as

$$\mathfrak{L}^{-1}_{approx} := \mathcal{P}_{\mathfrak{n}}(\varphi)\mathcal{I}J\mathcal{P}_{\mathfrak{n}}(\varphi)^{-1} \,. \tag{10.4.16}$$

We claim that the function $h$ defined in (10.4.15) is the required solution of (8.2.7) with a remainder $r$ satisfying (8.2.8).

**Lemma 10.4.2.** *The approximate solution $h$ defined in* (10.4.15) *satisfies* (8.2.6) *and* (8.2.9).

PROOF. We have

$$\begin{aligned} \|h\|_{\mathrm{Lip},s_1} &\overset{(10.4.15)}{=} \|\mathcal{P}_{\mathfrak{n}}\mathcal{I}(J\mathcal{P}_{\mathfrak{n}}^{-1}g)\|_{\mathrm{Lip},s_1} \overset{(9.1.21),(3.3.8)}{\lesssim_{s_1}} \|\mathcal{I}(J\mathcal{P}_{\mathfrak{n}}^{-1}g)\|_{\mathrm{Lip},s_1} \\ &\overset{(10.4.12)}{\lesssim_{s_1}} \nu^{-\frac{1}{20}}\|\mathcal{P}_{\mathfrak{n}}^{-1}g\|_{\mathrm{Lip},s_1} \overset{(9.1.21)}{\lesssim_{s_1}} \nu^{-\frac{1}{20}}\|g\|_{\mathrm{Lip},s_1} \lesssim_{s_1} \varepsilon^2\nu^{\frac{19}{20}} \end{aligned} \tag{10.4.17}$$

by the assumption $\|g\|_{\mathrm{Lip},s_1} \le \varepsilon^2\nu$, see (8.2.5). Therefore $\|h\|_{\mathrm{Lip},s_1} \le \varepsilon^2\nu^{4/5}$, for $\nu$ small enough (depending on $s_1$), proving the first inequality in (8.2.6).

Let us now estimate $\|h\|_{\mathrm{Lip},s_3}$. We have

$$\begin{aligned} \|g'\|_{\mathrm{Lip},s_3} &\overset{(10.4.15)}{=} \|\mathcal{P}_{\mathfrak{n}}^{-1}g\|_{\mathrm{Lip},s_3} \\ &\overset{(3.3.8)}{\lesssim_{s_3}} |\mathcal{P}_{\mathfrak{n}}^{-1}|_{\mathrm{Lip},s_3}\|g\|_{\mathrm{Lip},s_1} + |\mathcal{P}_{\mathfrak{n}}^{-1}|_{\mathrm{Lip},s_1}\|g\|_{\mathrm{Lip},s_3} \\ &\overset{(10.4.11),(9.1.21)}{\lesssim_{s_3}} \nu^{-\frac{11}{10}}\|g\|_{\mathrm{Lip},s_1} + \|g\|_{\mathrm{Lip},s_3} \\ &\overset{(8.2.5)}{\lesssim_{s_3}} \varepsilon^2(\nu^{-\frac{1}{10}} + \nu^{-1}) \lesssim_{s_3} \varepsilon^2\nu^{-1} \,. \end{aligned} \tag{10.4.18}$$

Then (10.4.13), (10.4.18) and $\|g'\|_{\mathrm{Lip},s_1} \lesssim_{s_1} \|g\|_{\mathrm{Lip},s_1} \lesssim_{s_1} \varepsilon^2\nu$, imply

$$\|\mathcal{I}(Jg')\|_{\mathrm{Lip},s_3} \lesssim_{s_3} \varepsilon^2(\nu^{-\frac{1}{20}}\nu^{-1} + \nu^{-\frac{6}{5}}\nu) \lesssim_{s_3} \varepsilon^2\nu^{-\frac{21}{20}} \tag{10.4.19}$$

and finally, using (3.3.8), (9.1.21), (10.4.11),(10.4.19) and $\|\mathcal{I}(Jg')\|_{\mathrm{Lip},s_1} \lesssim_{s_1} \varepsilon^2\nu^{\frac{19}{20}}$ (see (10.4.17)), we conclude that

$$\|h\|_{\mathrm{Lip},s_3} = \|\mathcal{P}_{\mathfrak{n}}\mathcal{I}(Jg')\|_{\mathrm{Lip},s_3} \lesssim_{s_3} \varepsilon^2\big(\nu^{-\frac{3}{20}} + \nu^{-\frac{21}{20}}\big) \le \varepsilon^2\nu^{-\frac{11}{10}} \tag{10.4.20}$$

for $\nu$ small enough (depending on $s_3$). This proves the second inequality in (8.2.6).

At last, by (9.1.21) and (10.3.7), for any $g \in \mathcal{H}^{s_0+Q'} \cap H_{\mathbb{S}}^{\perp}$,

$$\|\mathfrak{L}_{approx}^{-1} g\|_{\mathrm{Lip},s_0} = \|\mathcal{P}_{\mathfrak{n}} \mathcal{I} J \mathcal{P}_{\mathfrak{n}}^{-1} g\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \|J\mathcal{P}_{\mathfrak{n}}^{-1} g\|_{\mathrm{Lip},s_0+Q'} \lesssim_{s_1} \|g\|_{\mathrm{Lip},s_0+Q'}$$

using that $s_0 + Q' \le s_1$. This proves (8.2.9). ∎

We finally estimate the error $r$ in the equation (8.2.7).

**Lemma 10.4.3.** $\|r\|_{\mathrm{Lip},s_1} \le \varepsilon^2 \nu^{3/2}$ *, where* $r = \big(\bar\omega_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho)\big)h - g$.

Proof. By (10.1.2), (10.4.15) and recalling that $\mathcal{L} = J\bar\omega_\varepsilon \cdot \partial_\varphi + A_{\mathfrak{n}} + \rho_{\mathfrak{n}}$ we have

$$\begin{aligned} r &= \big(\bar\omega_\varepsilon \cdot \partial_\varphi - J(A_0 + \rho)\big)h - g \\ &= \mathcal{P}_{\mathfrak{n}}\big[\big(\bar\omega_\varepsilon \cdot \partial_\varphi - J(A_{\mathfrak{n}} + \rho_{\mathfrak{n}})\big)h' - g'\big] \\ &= -\mathcal{P}_{\mathfrak{n}} J\big[\mathcal{L}h' - Jg'\big] \\ &= -\mathcal{P}_{\mathfrak{n}} J\big[(\mathcal{L}\mathcal{I} - \mathrm{Id})Jg'\big] \,. \end{aligned} \tag{10.4.21}$$

Now, by the estimates in the proof of Lemma 10.4.2, see (10.4.18), we know that $g' = \mathcal{P}_{\mathfrak{n}}^{-1} g$ satisfies

$$\|g'\|_{\mathrm{Lip},s_3} \lesssim_{s_3} \varepsilon^2(\nu^{-1} + \nu^{-\frac{1}{10}}) \lesssim_{s_3} \varepsilon^2 \nu^{-1}\,, \qquad \|g'\|_{\mathrm{Lip},s_1} \lesssim \varepsilon^2 \nu\,. \tag{10.4.22}$$

Hence by (10.4.21), (3.3.8), (9.1.21), we get

$$\begin{aligned} \|r\|_{\mathrm{Lip},s_1} &\lesssim \|(\mathcal{L}\mathcal{I} - \mathrm{Id})Jg'\|_{\mathrm{Lip},s_1} \\ &\overset{(10.4.14)}{\lesssim} \nu^{\frac{11}{4}} \|g'\|_{\mathrm{Lip},s_3} + \nu^{\frac{3}{2}} \|g'\|_{\mathrm{Lip},s_1} \\ &\overset{(10.4.22)}{\lesssim_{s_3}} \varepsilon^2(\nu^{\frac{7}{4}} + \nu^{\frac{5}{2}}) \le \varepsilon^2 \nu^{\frac{3}{2}} \end{aligned}$$

proving the lemma. ∎

The next lemma completes the proof of Proposition 8.2.1.

**Lemma 10.4.4. (Measure estimates)** *The sets* $\mathbf{\Lambda}(\varepsilon;\eta, A_0, \rho)$, $1/2 \le \eta \le 5/6$, *defined in* (10.4.2) *satisfy properties 1-3 of Proposition 8.2.1.*

Proof. Property 1 follows immediately because the sets $\Lambda_\infty$ defined in Corollary 9.1.2, and $\mathsf{\Lambda}$ in Proposition 10.2.1 are increasing in $\eta$.

For the proof of properties 2 and 3 , we observe that whereas $A_0$ is defined for any $\lambda \in \widetilde\Lambda$, by Corollary 9.1.2 the operators $A_{\mathfrak{n}}$, $\mathfrak{n} \ge 1$, are defined for $\lambda \in \Lambda_\infty(\varepsilon; 5/6, A_0, \rho) \subset \widetilde\Lambda$. Hence for $\mathfrak{n} \ge 1$, the set $\mathsf{\Lambda}(\varepsilon, \eta, A_{\mathfrak{n}})$ $(1/2 \le \eta \le 1)$ is considered as a subset of $\Lambda_\infty(\varepsilon; 5/6, A_0, \rho)$, and we shall apply Proposition 10.2.1 (more precisely Lemma 10.2.8) in this setting.

PROOF OF PROPERTY 2 OF PROPOSITION 8.2.1. Definining

$$\mathtt{\Lambda}_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho) := \Lambda_\infty(\varepsilon;\eta,A_0,\rho)\bigcap\Big(\bigcap_{k=0}^{\mathfrak{n}-1}\mathtt{\Lambda}(\varepsilon;\eta+\eta_k,A_k)\Big)\,,\quad \forall\mathfrak{n}\geq 1\,, \tag{10.4.23}$$

we write the set in (10.4.2) (recall that $\eta_0 = 0$) as

$$\mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho) = \Lambda_\infty(\varepsilon;\eta,A_0,\rho)\bigcap\mathtt{\Lambda}(\varepsilon;\eta,A_0)\bigcap\Big(\bigcap_{\mathfrak{n}\geq 1}\mathtt{\Lambda}_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)\Big)\,. \tag{10.4.24}$$

Then its complementary set may be decomposed as (arguing as in (9.1.54))

$$\begin{aligned}\mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho)^c = \Lambda_\infty(\varepsilon;\eta,A_0,\rho)^c\bigcup\mathtt{\Lambda}(\varepsilon;\eta,A_0)^c \\ \bigcup\Big(\bigcup_{\mathfrak{n}\geq 1}\big(\mathtt{\Lambda}_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)\bigcap\mathtt{\Lambda}(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}})^c\big)\Big)\,.\end{aligned} \tag{10.4.25}$$

By property 2 of Corollary 9.1.2 we have

$$\big|\Lambda_\infty(\varepsilon;1/2,A_0,\rho)^c\cap\widetilde{\Lambda}\big| \leq b_1(\varepsilon) \quad\text{with}\quad \lim_{\varepsilon\to 0} b_1(\varepsilon) = 0\,. \tag{10.4.26}$$

By Lemma 10.2.8,

$$\big|\mathtt{\Lambda}(\varepsilon;1/2,A_0)^c\cap\widetilde{\Lambda}\big| \lesssim \varepsilon\,. \tag{10.4.27}$$

Moreover, for $\mathfrak{n}\geq 1$, by (10.4.23), we have

$$\begin{aligned}&\mathtt{\Lambda}_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)\cap\mathtt{\Lambda}(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}})^c \\ &\subset \Lambda_\infty(\varepsilon;5/6,A_0,\rho)\cap\mathtt{\Lambda}(\varepsilon;\eta+\eta_{\mathfrak{n}-1},A_{\mathfrak{n}-1})\cap\mathtt{\Lambda}(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}})^c\,.\end{aligned}$$

Then, since, for all $\mathfrak{n}\geq 1$, we have $|A_{\mathfrak{n}} - A_{\mathfrak{n}-1}|_{+,s_1} \leq \delta_{\mathfrak{n}}^{3/4}$ on $\Lambda_\infty(\varepsilon;5/6,A_0,\rho)$ (see (9.1.39)) and $\eta+\eta_{\mathfrak{n}-1} = \eta+\eta_{\mathfrak{n}} - \delta_{\mathfrak{n}}^{3/8}$ (see (9.1.37)), we deduce by Lemma 10.2.8 that, for any $1/2\leq\eta\leq 5/6$,

$$\big|\mathtt{\Lambda}_{\mathfrak{n}}(\varepsilon;\eta,A_0,\rho)\bigcap\mathtt{\Lambda}(\varepsilon;\eta+\eta_{\mathfrak{n}},A_{\mathfrak{n}})^c\big| \leq \delta_{\mathfrak{n}}^{3\alpha/4}\,. \tag{10.4.28}$$

In conclusion, by (10.4.25) and (10.4.26), (10.4.27), (10.4.28) at $\eta = 1/2$, we deduce

$$\begin{aligned}\big|\mathbf{\Lambda}(\varepsilon;1/2,A_0,\rho)^c\cap\widetilde{\Lambda}\big| &\leq b_1(\varepsilon) + C\varepsilon + \sum_{\mathfrak{n}\geq 1}\delta_{\mathfrak{n}}^{3\alpha/4} \\ &\leq b_1(\varepsilon) + C\varepsilon + c\delta_1^{3\alpha/4} \leq b(\varepsilon)\end{aligned}$$

with $\lim_{\varepsilon\to 0} b(\varepsilon) = 0$, since $\delta_1 = \varepsilon^3$. This proves property 2 of Proposition 8.2.1 for the sets $\mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho)$.

Proof of property 3 of Proposition 8.2.1. By (10.4.25) (with $A_0', \rho'$ instead of $A_0, \rho$) and (10.4.24) we deduce, for all $(1/2) + \delta^{2/5} \leq \eta \leq 5/6$, the inclusion

$$\begin{aligned} \mathcal{M} &:= \widetilde{\Lambda}' \bigcap \mathbf{\Lambda}(\varepsilon; \eta, A_0', \rho')^c \bigcap \mathbf{\Lambda}(\varepsilon; \eta - \delta^{2/5}, A_0, \rho) \\ &\subset \mathcal{M}_\infty \bigcup \mathcal{M}_0 \bigcup \Big( \bigcup_{\mathfrak{n} \geq 1} \mathcal{M}_{\mathfrak{n}} \Big) , \end{aligned} \tag{10.4.29}$$

where

$$\begin{aligned} \mathcal{M}_\infty &:= \widetilde{\Lambda}' \cap [\Lambda_\infty(\varepsilon; \eta, A_0', \rho')]^c \cap \Lambda_\infty(\varepsilon; \eta - \delta^{2/5}, A_0, \rho) \\ \mathcal{M}_0 &:= \widetilde{\Lambda}' \cap \mathtt{\Lambda}(\varepsilon; \eta, A_0')^c \cap \mathtt{\Lambda}(\varepsilon; \eta - \delta^{2/5}, A_0) \\ \mathcal{M}_{\mathfrak{n}} &:= \mathtt{\Lambda}_{\mathfrak{n}}(\varepsilon; \eta, A_0', \rho') \cap \mathtt{\Lambda}(\varepsilon; \eta + \eta_{\mathfrak{n}}, A_{\mathfrak{n}}')^c \cap \mathtt{\Lambda}(\varepsilon; \eta + \eta_{\mathfrak{n}} - \delta^{2/5}, A_{\mathfrak{n}}) , \quad \text{for } \mathfrak{n} \geq 1 . \end{aligned}$$

By (8.2.3) and noting that $\delta^{1/2} \leq \delta^{2/5}$, we deduce by (9.1.18) that, for all $(1/2) + \delta^{2/5} \leq \eta \leq 5/6$,

$$|\mathcal{M}_\infty| \leq \delta^{\alpha/2} . \tag{10.4.30}$$

Moreover Lemma 10.2.8 implies that

$$|\mathcal{M}_0| \leq \big| \widetilde{\Lambda}' \cap \mathtt{\Lambda}(\varepsilon; \eta, A_0')^c \cap \mathtt{\Lambda}(\varepsilon; \eta - \delta^{1/2}, A_0) \big| \leq \delta^\alpha . \tag{10.4.31}$$

For $\mathfrak{n} \geq 1$, we have, by (10.4.28) (applied to $(A_0', \rho')$)

$$|\mathcal{M}_{\mathfrak{n}}| \leq \big| \mathtt{\Lambda}_{\mathfrak{n}}(\varepsilon; \eta, A_0', \rho') \cap \mathtt{\Lambda}(\varepsilon; \eta + \eta_{\mathfrak{n}}, A_{\mathfrak{n}}')^c \big| \leq \delta_{\mathfrak{n}}^{3\alpha/4} . \tag{10.4.32}$$

On the other hand, assumption (8.2.3) implies (9.1.35), and so (9.1.36) holds, i.e. $|A_{\mathfrak{n}} - A_{\mathfrak{n}}'|_{+,s_1} \leq \delta^{4/5}$ for any $\lambda$ in $\mathcal{M}_{\mathfrak{n}} \subset \Lambda_\infty(\varepsilon; 5/6, A_0', \rho') \bigcap \Lambda_\infty(\varepsilon; 5/6, A_0, \rho)$, and we deduce, by Lemma 10.2.8, the measure estimate

$$\begin{aligned} |\mathcal{M}_{\mathfrak{n}}| &\leq \big| \Lambda_\infty(\varepsilon; 5/6, A_0', \rho') \cap \mathtt{\Lambda}(\varepsilon; \eta + \eta_{\mathfrak{n}}, A_{\mathfrak{n}}')^c \cap \mathtt{\Lambda}(\varepsilon; \eta + \eta_{\mathfrak{n}} - \delta^{2/5}, A_{\mathfrak{n}}) \big| \\ &\lesssim \delta^{4\alpha/5} . \end{aligned} \tag{10.4.33}$$

Finally (10.4.29), (10.4.30), (10.4.31), (10.4.33), (10.4.32) imply the measure estimate

$$|\mathcal{M}| \leq \delta^{\alpha/2} + \delta^\alpha + C \sum_{\mathfrak{n} \geq 1} \min(\delta^{4\alpha/5}, \delta_{\mathfrak{n}}^{3\alpha/4}) . \tag{10.4.34}$$

Now, using that $\delta_{\mathfrak{n}} = \delta_1^{(\frac{3}{2})^{\mathfrak{n}-1}}$ with $\delta_1 = \varepsilon^3$, there is a constant $C > 0$ such that

$$\delta^{\frac{4\alpha}{5}} \leq \delta_{\mathfrak{n}}^{\frac{3\alpha}{4}} \quad \Longrightarrow \quad \mathfrak{n} \leq C \ln(\ln(\delta^{-1})) .$$

Hence

$$\sum_{\mathfrak{n}\geq 1}\min(\delta^{\frac{4\alpha}{5}},\delta_{\mathfrak{n}}^{\frac{3\alpha}{4}})\leq C\ln(\ln(\delta^{-1}))\delta^{\frac{4\alpha}{5}}+\sum_{\mathfrak{n}\geq 1,\delta_{\mathfrak{n}}^{\frac{3\alpha}{4}}<\delta^{\frac{4\alpha}{5}}}\delta_{\mathfrak{n}}^{\frac{3\alpha}{4}}$$
$$\leq C\ln(\ln(\delta^{-1}))\delta^{\frac{4\alpha}{5}}+C\delta^{\frac{4\alpha}{5}}\,. \tag{10.4.35}$$

The estimates (10.4.34) and (10.4.35) imply, for $\delta$ small, that

$$|\mathcal{M}|\leq\delta^{\alpha/3}\,,$$

which is (8.2.4). ■

Proof of Proposition 8.2.2. The result of Proposition 8.2.1 holds for any fixed $s_3$ satisfying (8.2.1). The sets $\mathbf{\Lambda}(\varepsilon;\eta,A_0,\rho)$ do not depend on the choice of $s_3$. On the other hand, the smallness condition for $\varepsilon$ may depend on $s_3$ but, as explained in Remark 10.4.1, what is really used in the proof of Proposition 8.2.1 is that $\nu$ is small enough. This naturally leads to Proposition 8.2.2, the proof of which is exactly the same as for Proposition 8.2.1, replacing $s_3$ by $s$, and the smallness condition on $\varepsilon$ by a smallness condition on $\nu$ (depending on $s$). Notice that the approximate inverse $\mathfrak{L}^{-1}_{approx}$ depends not only on $\nu$, but also on $s$, in particular through the choice of $N$, that, replacing $s_3$ with $s$ in (10.4.3), results in

$$N\in\Big[\nu^{-\frac{3}{s-s_1}}-\frac{1}{2},\nu^{-\frac{3}{s-s_1}}+\frac{1}{2}\Big)\,.$$

# Chapter 11

# Proof of the Nash-Moser Theorem

In this Chapter we finally prove the Nash-Moser Theorem 5.1.2, finding, by an iterative scheme (see section 11.2), a solution of the nonlinear equation $\mathcal{F}(\lambda; i) = 0$ where $\mathcal{F}$ is the operator defined in (5.1.2).

By the procedure described in Chapter 6 (see Proposition 6.1.1) in order to find an approximate inverse for the linearized operator at an approximate solution $\underline{i}$, and thus implement a convergent Nash-Moser scheme, it is sufficient to prove the existence of an approximate right inverse of the operator $\mathbb{D}(\underline{i})$ defined in (6.1.22). This is achieved in Proposition 11.2.4.

## 11.1 Approximate right inverse of $\mathcal{L}_\omega$

We first give the key result about the existence of an approximate right inverse of the operator $\mathcal{L}_\omega := \mathcal{L}_\omega(\underline{i})$ defined in (6.1.23), acting on the normal subspace $H_{\mathbb{S}}^{\perp}$. We recall that $\underline{i}(\varphi) = (\varphi, 0, 0) + \underline{\mathfrak{I}}(\varphi)$ is defined for all $\lambda \in \Lambda_{\underline{\mathfrak{I}}}$.

**Proposition 11.1.1. (Approximate right inverse of $\mathcal{L}_\omega(\underline{i})$)** *Let $\bar{\omega}_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ be $(\gamma_1, \tau_1)$-Diophantine and satisfy property $(\mathbf{NR})_{\gamma_1,\tau_1}$ in Definition 4.1.4 with $\gamma_1, \tau_1$ fixed in (1.2.28). Assume (8.2.1).*

*Then there is $\varepsilon_0 > 0$ such that, $\forall \varepsilon \in (0, \varepsilon_0)$, for all $\underline{\mathfrak{I}}$, defined for all $\lambda \in \Lambda_{\underline{\mathfrak{I}}}$, satisfying $\|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s_2+2} \leq \varepsilon$, there are closed subsets $\mathbf{\Lambda}(\varepsilon; \eta, \underline{\mathfrak{I}}) \subset \Lambda_{\underline{\mathfrak{I}}}$, $1/2 \leq \eta \leq 5/6$, satisfying*

1. *$\mathbf{\Lambda}(\varepsilon; \eta, \underline{\mathfrak{I}}) \subseteq \mathbf{\Lambda}(\varepsilon; \eta', \underline{\mathfrak{I}})$ for all $1/2 \leq \eta \leq \eta' \leq 5/6$,*
2. *$\left|[\mathbf{\Lambda}(\varepsilon; 1/2, \underline{\mathfrak{I}})]^c \cap \Lambda_{\underline{\mathfrak{I}}}\right| \leq b(\varepsilon)$ where $\lim_{\varepsilon \to 0} b(\varepsilon) = 0$,*
3. *if $\|\underline{\mathfrak{I}}' - \underline{\mathfrak{I}}\|_{s_1+2} \leq \delta \leq \varepsilon^{3/2}$ for all $\lambda \in \Lambda_{\underline{\mathfrak{I}}} \cap \Lambda_{\underline{\mathfrak{I}}'}$, then, for all $(1/2) + \delta^{2/5} \leq \eta \leq 5/6$,*

$$\left|\Lambda_{\underline{\mathfrak{I}}'} \cap [\mathbf{\Lambda}(\varepsilon; \eta, \underline{\mathfrak{I}}')]^c \cap \mathbf{\Lambda}(\varepsilon; \eta - \delta^{2/5}, \underline{\mathfrak{I}})\right| \leq \delta^{\alpha/3}; \tag{11.1.1}$$

*and, for any* $\nu \in (0, \varepsilon^{\frac{3}{2}})$ *such that* $\|\underline{\mathfrak{I}}\|_{\mathrm{Lip}, s_3+4} \leq \varepsilon \nu^{-\frac{9}{10}}$, *there exists a linear operator*

$$\mathcal{L}_{appr}^{-1} := \mathcal{L}_{appr,\nu}^{-1}$$

*such that, for any function* $g : \Lambda_{\underline{\mathfrak{I}}} \to \mathcal{H}^{s_3+2} \cap H_{\mathbb{S}}^{\perp}$ *satisfying*

$$\|g\|_{\mathrm{Lip}, s_1} \leq \varepsilon^2 \nu \,, \quad \|g\|_{\mathrm{Lip}, s_3+2} \leq \varepsilon^2 \nu^{-\frac{9}{10}} \,, \tag{11.1.2}$$

*the function* $h := \mathcal{L}_{appr}^{-1} g$, $h : \mathbf{\Lambda}(\varepsilon; 5/6, \underline{\mathfrak{I}}) \to \mathcal{H}^{s_3+2} \cap H_{\mathbb{S}}^{\perp}$ *satisfies*

$$\|h\|_{\mathrm{Lip}, s_1} \leq C(s_1) \varepsilon^2 \nu^{\frac{4}{5}} \,, \quad \|h\|_{\mathrm{Lip}, s_3+2} \leq C(s_3) \varepsilon^2 \nu^{-\frac{11}{10}} \,, \tag{11.1.3}$$

*and, setting* $\mathcal{L}_\omega := \mathcal{L}_\omega(\underline{i})$ *defined in* (6.1.23), *we have*

$$\|\mathcal{L}_\omega h - g\|_{\mathrm{Lip}, s_1} \leq C(s_1) \varepsilon^2 \nu^{\frac{3}{2}} \,. \tag{11.1.4}$$

*Furthermore, setting* $Q' := 2(\tau' + \varsigma s_1) + 3$ *(where* $\varsigma = 1/10$, *see* (4.1.16)*) and* $\tau'$ *are given by Proposition 4.1.5), for all* $g \in \mathcal{H}^{s_0+Q'} \cap H_{\mathbb{S}}^{\perp}$,

$$\|\mathcal{L}_{appr}^{-1} g\|_{\mathrm{Lip}, s_0} \lesssim_{s_1} \|g\|_{\mathrm{Lip}, s_0+Q'} \,. \tag{11.1.5}$$

*We underline that the estimate* (11.1.5) *is independent of* $\nu$ *which defines* $\mathcal{L}_{appr} = \mathcal{L}_{appr,\nu}$.

**Remark 11.1.2.** *We lay the stress on the fact that along the proof, any smallness condition depending on* $s_3$ *will concern* $\nu$ *rather than* $\varepsilon$ *(while* $\varepsilon$ *is small depending on* $s_1$ *and* $s_2$*). Since* $\nu \in (0, \varepsilon^{\frac{3}{2}})$, *these conditions obviously will be satisfied for* $\varepsilon$ *small enough (depending on* $s_3$*). However to obtain* $C^\infty$ *solutions in section 11.3, we shall need to replace* $s_3$ *with some* $s$ *that is slowly increasing along the Nash Moser scheme, and we shall use that, for* $\varepsilon$ *fixed, the estimates of Proposition 11.1.1 hold for any* $\nu$ *small enough (depending on* $s$*).*

Proof. The proposition is a consequence of Lemma 6.1.2, Proposition 7.3.1, Lemma 8.1.2 and Proposition 8.2.1. By (6.1.23), (6.1.24) and recalling that $\omega = (1+\varepsilon^2\lambda)\bar{\omega}_\varepsilon$ we write the operator $\mathcal{L}_\omega$ acting on $H_{\mathbb{S}}^{\perp}$,

$$\mathcal{L}_\omega = (1+\varepsilon^2\lambda)\Big[\bar{\omega}_\varepsilon \cdot \partial_\varphi - J\Big(\mathtt{A} + \frac{\mathtt{r}_\varepsilon}{1+\varepsilon^2\lambda}\Big)\Big] \quad \text{where} \quad \mathtt{A} = \frac{D_V}{1+\varepsilon^2\lambda} + \frac{\varepsilon^2 \mathtt{B}}{1+\varepsilon^2\lambda} \,,$$

as in (7.1.3). Notice that the term $\mathtt{r}_\varepsilon = \mathtt{r}_\varepsilon(\underline{\mathfrak{I}})$ depends on the torus $\underline{i}$ at which we linearize, see (6.1.1), unlike $\mathtt{B}$ in (6.1.25), and thus $\mathtt{A}$. Moreover, by (6.1.27) and the assumptions of the proposition

$$\|\underline{\mathfrak{I}}\|_{\mathrm{Lip}, s_2+2} \leq \varepsilon \,, \quad \|\underline{\mathfrak{I}}\|_{\mathrm{Lip}, s_3+4} \leq \varepsilon \nu^{-\frac{9}{10}} \,,$$

we get

$$|\mathtt{r}_\varepsilon|_{\mathrm{Lip}, +, s_2} \leq C(s_2) \varepsilon^2 (\varepsilon^2 + \varepsilon) \ll \varepsilon^{\frac{5}{2}} \tag{11.1.6}$$

$$|\mathtt{r}_\varepsilon|_{\mathrm{Lip}, +, s_3+2} \leq C(s_3) \varepsilon^2 (\varepsilon^2 + \varepsilon \nu^{-\frac{9}{10}}) \ll \varepsilon^{\frac{5}{2}} \nu^{-1} \,. \tag{11.1.7}$$

Applying Proposition 7.3.1 we get

$$\mathcal{L}_\omega = (1+\varepsilon^2\lambda)\mathtt{P}(\varphi)\big[\bar\omega_\varepsilon\cdot\partial_\varphi - J(\mathtt{A}_0+\varrho^+)\big]\mathtt{P}^{-1}(\varphi)\,, \tag{11.1.8}$$

where, as in (7.3.5), (7.1.4),

$$\mathtt{A}_0 := \frac{D_V}{1+\varepsilon^2\lambda} + R_0\,, \quad R_0 = \Pi_{\mathtt{D}}\varrho\,, \quad \varrho = \frac{\varepsilon^2}{1+\varepsilon^2\lambda}\mathtt{B}\,, \tag{11.1.9}$$

and the coupling term $\varrho^+$ satisfies (7.3.7), i.e.

$$|\varrho^+|_{\mathrm{Lip},+,s} \le C(s)(\varepsilon^4 + |\mathtt{r}_\varepsilon|_{\mathrm{Lip},+,s})\,. \tag{11.1.10}$$

By (3.3.35) and (7.1.6) we have

$$|R_0|_{\mathrm{Lip},+,s} = |\Pi_{\mathtt{D}}\varrho|_{\mathrm{Lip},+,s} \le C(s)\varepsilon^2\,. \tag{11.1.11}$$

We want to apply Proposition 8.2.1 to the operator

$$\bar\omega_\varepsilon\cdot\partial_\varphi - J(\mathtt{A}_0+\varrho^+) \qquad \text{with} \qquad \breve g = \frac{\mathtt{P}^{-1}(\varphi)g}{1+\varepsilon^2\lambda}\,, \tag{11.1.12}$$

namely with the substitutions $A_0 \rightsquigarrow \mathtt{A}_0$, $R_0 = \Pi_{\mathtt{D}}\varrho$, $\rho \rightsquigarrow \varrho^+$ and $g \rightsquigarrow \breve g$. First of all Lemma 8.1.2 proves that $\mathtt{A}_0$ is an admissible split operator in the class $\mathcal{C}(C_1, c_1, c_2)$. Then, by (11.1.10), (6.1.26), (11.1.11), (11.1.6)-(11.1.7), we verify that

$$\begin{aligned} &|\varrho^+|_{\mathrm{Lip},+,s_1} \le \varepsilon^3\,,\\ &|R_0|_{\mathrm{Lip},+,s_2} + |\varrho^+|_{\mathrm{Lip},+,s_2} \le \varepsilon^{-1}\,,\\ &\nu\big(|R_0|_{\mathrm{Lip},+,s_3+2} + |\varrho^+|_{\mathrm{Lip},+,s_3+2}\big) \ll \varepsilon^2\,, \end{aligned} \tag{11.1.13}$$

for $\nu$ small depending on $s_3$ (for the last estimate we use the first inequality in (11.1.7)), which implies the conditions required in (8.2.2) and (8.2.5) with $s_3 \rightsquigarrow s_3+2$. Notice also that the function $\breve g$ defined in (11.1.12) satisfies, by (7.3.3) and (3.3.8), the estimate

$$\|\breve g\|_{\mathrm{Lip},s} \le C(s)\|g\|_{\mathrm{Lip},s}\,, \quad \forall s \ge s_1\,,$$

and therefore, by (11.1.2), the function $\breve g$ satisfies the assumption

$$\|\breve g\|_{\mathrm{Lip},s_1} \le \varepsilon^2 C(s_1)\nu\,, \quad \|\breve g\|_{\mathrm{Lip},s_3+2} \le C(s_3)\varepsilon^2\nu^{-\frac{9}{10}} \le \varepsilon^2(C(s_1)\nu)^{-1}$$

required in (8.2.5) with $\nu \rightsquigarrow C(s_1)\nu$ and $s_3 \rightsquigarrow s_3+2$, for $\nu$ small depending on $s_3$. Note also that, since $\nu \in (0, \varepsilon^{3/2})$, we have $C(s_1)\nu \in (0,\varepsilon)$, as required in Proposition 8.2.1. Therefore Proposition 8.2.1 applies and there are closed subsets (independent of $g$)

$$\mathbf{\Lambda}(\varepsilon;\eta,\underline{\mathfrak{I}}) := \mathbf{\Lambda}(\varepsilon;\eta,\mathtt{A}_0,\varrho^+) \subset \Lambda_{\underline{\mathfrak{I}}}\,, \quad 1/2 \le \eta \le 5/6\,, \tag{11.1.14}$$

satisfying the properties 1-3 listed in Proposition 8.2.1 with $\widetilde{\Lambda} = \Lambda_{\underline{\mathfrak{I}}}$, and an operator $\mathfrak{L}_{approx}^{-1} = \mathfrak{L}_{approx,\nu}^{-1}$ such that the function

$$\breve{h} := \mathfrak{L}_{approx}^{-1}\breve{g}\,, \ \breve{h} : \mathbf{\Lambda}(\varepsilon; 5/6, \underline{\mathfrak{I}}) \to \mathcal{H}^{s_3+2} \cap H_{\mathbb{S}}^{\perp}\,,$$

satisfies (8.2.6) with $s_3 \rightsquigarrow s_3 + 2$, i.e.

$$\|\breve{h}\|_{\mathrm{Lip},s_1} \lesssim_{s_1} \varepsilon^2 \nu^{\frac{4}{5}}\,, \quad \|\breve{h}\|_{\mathrm{Lip},s_3+2} \lesssim_{s_1} \varepsilon^2 \nu^{-\frac{11}{10}}\,, \tag{11.1.15}$$

and, see (8.2.7), (8.2.8),

$$(\bar{\omega}_\varepsilon \cdot \partial_\varphi - J(\mathtt{A}_0 + \varrho^+))\breve{h} = \breve{g} + \breve{r} \quad \text{with} \quad \|\breve{r}\|_{\mathrm{Lip},s_1} \lesssim_{s_1} \varepsilon^2 \nu^{\frac{3}{2}}\,. \tag{11.1.16}$$

Set

$$\mathcal{L}_{appr}^{-1} := \frac{1}{1+\varepsilon^2\lambda} \mathtt{P}(\varphi) \mathfrak{L}_{approx}^{-1} \mathtt{P}^{-1}(\varphi)\,, \quad h := \mathtt{P}(\varphi)\breve{h} = \mathcal{L}_{appr}^{-1} g\,. \tag{11.1.17}$$

By (11.1.16)-(11.1.17) and (11.1.8), (11.1.12) we get

$$\mathcal{L}_\omega h - g = (1+\varepsilon^2\lambda)\mathtt{P}(\varphi)\breve{r}\,. \tag{11.1.18}$$

By (11.1.17), (7.3.3), (3.3.8), (11.1.18) we get

$$\begin{aligned} \|h\|_{\mathrm{Lip},s} &= \|\mathtt{P}(\varphi)\breve{h}\|_{\mathrm{Lip},s} \le C(s)\|\breve{h}\|_{\mathrm{Lip},s} \\ \|\mathcal{L}_\omega h - g\|_{\mathrm{Lip},s} &= \|(1+\varepsilon^2\lambda)\mathtt{P}(\varphi)\breve{r}\|_{\mathrm{Lip},s} \le C(s)\|\breve{r}\|_{\mathrm{Lip},s}\,, \end{aligned}$$

and, from the estimates of $\breve{h}$ and $\breve{r}$ in (11.1.15), (11.1.16), we deduce (11.1.3), (11.1.4). Furthermore the estimate (11.1.5) follows by (11.1.17), (7.3.3), (3.3.8) and (8.2.9).

Let us finally prove that the sets $\mathbf{\Lambda}(\varepsilon; \eta, \underline{\mathfrak{I}})$ defined in (11.1.14) satisfy the properties 1-3 listed in the statement of the Proposition 11.1.1. Items 1-2 are immediate consequences the corresponding ones in Proposition 8.2.1. For proving item 3, first notice that, $\varrho$ (defined in (7.1.4)), $R_0 = \Pi_{\mathtt{D}}\varrho$ and hence $\mathtt{A}_0$ (defined in (11.1.9)) do not depend on the torus $\underline{\mathfrak{I}}$ at which we linearize. Then, if $\|\underline{\mathfrak{I}} - \underline{\mathfrak{I}}'\|_{s_1+2} \le \delta \le \varepsilon^{3/2}$, we have

$$\begin{aligned} |\mathtt{A}_0' - \mathtt{A}_0|_{+,s_1} + |\varrho^+ - (\varrho^+)'|_{+,s_1} &= |\varrho^+ - (\varrho^+)'|_{+,s_1} \\ &\overset{(7.3.8)}{\lesssim_{s_1}} |\mathtt{r}_\varepsilon - \mathtt{r}_\varepsilon'|_{+,s_1} \\ &\overset{(6.1.28)}{\lesssim_{s_1}} \varepsilon^2 \|\underline{\mathfrak{I}} - \underline{\mathfrak{I}}'\|_{s_1+2} \le \varepsilon^{\frac{3}{2}}\delta \le \varepsilon^3\,, \end{aligned}$$

so that condition (8.2.3) of Proposition 8.2.1 is satisfied. Hence, by the property 3 (see (8.2.4)) in Proposition 8.2.1, and the inclusion $\mathbf{\Lambda}(\varepsilon; \eta - \delta^{\frac{2}{5}}, \underline{\mathfrak{I}}) \subset \mathbf{\Lambda}(\varepsilon; \eta - (\varepsilon^{\frac{3}{2}}\delta)^{\frac{2}{5}}, \underline{\mathfrak{I}})$, we have

$$\big|\Lambda_{\underline{\mathfrak{I}}'} \cap [\mathbf{\Lambda}(\varepsilon; \eta, \underline{\mathfrak{I}}')]^c \cap \mathbf{\Lambda}(\varepsilon; \eta - \delta^{\frac{2}{5}}, \underline{\mathfrak{I}})\big| \le (\varepsilon^{\frac{3}{2}}\delta)^{\frac{\alpha}{3}}\,.$$

This proves property 3 for the sets $\mathbf{\Lambda}(\varepsilon; \eta, \underline{\mathfrak{I}})$ in Proposition 11.1.1. ■

## 11.2 Nash-Moser iteration

In Chapter 5 we observed that the nonlinear operator $\mathcal{F}$ defined in (5.1.2) evaluated at the trivial torus $i_0(\varphi) := (\varphi, 0, 0)$ satisfies (see (5.1.7))

$$\|\mathcal{F}(\lambda; i_0)\|_{\mathrm{Lip},s} \leq C(s)\varepsilon^2 \quad \text{on } \Lambda\,, \quad \forall s \geq s_0\,, \tag{11.2.1}$$

and, in Lemma 5.3.1 we have defined, for all $\lambda \in \Lambda$, a torus embedding $i_1(\varphi)$ such that (see (5.3.5))

$$\|\mathcal{F}(\lambda; i_1)\|_{\mathrm{Lip},s} \leq C(s)\varepsilon^4 \quad \text{on } \Lambda\,, \quad \forall s \geq s_0\,. \tag{11.2.2}$$

In this section we define a sequence of torus embeddings $(i_n(\varphi))_{n\geq 2}$, $n \geq 2$, defined for $\lambda$ belonging to a decreasing sequence of subsets $\mathbf{\Lambda}_n \subset \Lambda$ which converges, for all $\lambda$ belonging to the intersection $\cap_{n\geq 2}\mathbf{\Lambda}_n$ to a solution $i_\infty(\varphi)$ of $\mathcal{F}(\lambda; i_\infty) = 0$.

Fixing $\nu_1 = C(s_1)\varepsilon^2$ such that $\|\mathcal{F}(i_1)\|_{\mathrm{Lip},s_1} \leq \varepsilon^2\nu_1$ (see (11.2.2)), we define the decreasing sequence $(\nu_n)_{n\geq 1}$ by

$$\nu_n := \nu_1^{\mathfrak{q}^{n-1}}\,, \qquad \nu_{n+1} = \nu_n^{\mathfrak{q}}\,, \qquad \mathfrak{q} := \frac{3}{2} - \sigma_\star\,, \tag{11.2.3}$$

for some $\sigma_\star \in (0, 1/4)$. Theorem 5.1.2 will be a consequence of the following result.

**Theorem 11.2.1. (Nash-Moser)** *Let $\bar{\omega}_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ be $(\gamma_1, \tau_1)$-Diophantine and satisfy property $(\mathbf{NR})_{\gamma_1,\tau_1}$ in Definition 4.1.4 with $\gamma_1, \tau_1$ fixed in* (1.2.28)*. Assume* (8.2.1) *and $s_2 - s_1 \geq \underline{\tau} + 2$, $s_1 \geq s_0 + 2 + \underline{\tau} + Q'$, where $\underline{\tau}$ is the loss of derivatives defined in Proposition 6.1.1 and $Q' := 2(\tau' + \varsigma s_1) + 3$ is defined in Proposition 10.2.1.*

*Then, there exists $s_3$ large enough, a constant $\sigma_\star := \sigma_\star(s_3) > 0$, satisfying $\sigma_\star(s_3) \to 0$ as $s_3 \to +\infty$, and $\varepsilon_0 > 0$, such that, defining the sequence $(\nu_n)$ by* (11.2.3)*, for all $0 < \varepsilon \leq \varepsilon_0$, for all $n \geq 1$, there exist*

1. *a subset $\mathbf{\Lambda}_n \subseteq \mathbf{\Lambda}_{n-1}$, $\mathbf{\Lambda}_1 := \mathbf{\Lambda}_0 := \Lambda$, satisfying*

$$\begin{aligned} &|\mathbf{\Lambda}_1 \setminus \mathbf{\Lambda}_2| \leq b(\varepsilon) \quad \text{with} \quad \lim_{\varepsilon\to 0} b(\varepsilon) = 0\,, \\ &|\mathbf{\Lambda}_{n-1} \setminus \mathbf{\Lambda}_n| \leq \nu_{n-2}^{\alpha_*}\,,\ \forall n \geq 3\,, \end{aligned} \tag{11.2.4}$$

   *where $\alpha_* = \alpha/4$ and $\alpha > 0$ is the exponent in* (11.1.1)*,*

2. *a torus $i_n(\varphi) = (\varphi, 0, 0) + \mathfrak{I}_n(\varphi)$, defined for all $\lambda \in \mathbf{\Lambda}_n$, satisfying*

$$\|\mathfrak{I}_n\|_{\mathrm{Lip},s_1+2} \leq C(s_1)\varepsilon^2\,, \quad \|\mathfrak{I}_n\|_{\mathrm{Lip},s_2+2} \leq \varepsilon\,, \quad \|\mathfrak{I}_n\|_{\mathrm{Lip},s_3+2} \leq \varepsilon^2\nu_n^{-\frac{4}{5}}\,, \tag{11.2.5}$$

$$\|i_n - i_{n-1}\|_{\mathrm{Lip},s_1+2} \leq C(s_1)\nu_{n-1} + \varepsilon^2\nu_{n-1}^{\frac{4}{5}-\sigma_\star}\,,\ n \geq 2\,, \tag{11.2.6}$$

$$\|i_n - i_{n-1}\|_{\mathrm{Lip},s_2+2} \leq \varepsilon\nu_{n-1}^{\frac{1}{5}}\,, \quad n \geq 2\,, \tag{11.2.7}$$

   *such that*

$$\|\mathcal{F}(i_n)\|_{\mathrm{Lip},s_1} \leq \varepsilon^2\nu_n\,. \tag{11.2.8}$$

The rest of this section is devoted to the proof of Theorem 11.2.1.

**First step.** For $n = 1$ the torus $i_1(\varphi)$ defined in Lemma 5.3.1, for all $\lambda \in \Lambda = \mathbf{\Lambda}_1$, satisfies (5.3.4) and (5.3.5), which imply (11.2.5) and (11.2.8) at $n = 1$, for $\varepsilon$ small enough (recall that $\nu_1 = C(s_1)\varepsilon^2$).

**Iteration.** We now proceed by induction. Assume that we have already defined subsets $\mathbf{\Lambda}_n \subseteq \mathbf{\Lambda}_{n-1} \subseteq \ldots \subseteq \Lambda$ satisfying (11.2.4), and $n$ tori $i_1, \ldots, i_n$, of the form $i_n(\varphi) = (\varphi, 0, 0) + \mathfrak{I}_n(\varphi)$, satisfying (11.2.5)-(11.2.7) and such that (11.2.8) holds. We are going to define a subset $\mathbf{\Lambda}_{n+1} \subseteq \mathbf{\Lambda}_n$ and, for all $\lambda \in \mathbf{\Lambda}_{n+1}$, the subsequent better approximate torus embedding $i_{n+1}$ which satisfies (11.2.5)-(11.2.7) and (11.2.8) at order $n + 1$.

We shall define $i_{n+1}$ by a Nash-Moser type iterative scheme, using Propositions 6.1.1 and 11.1.1.

- **Notation.** From now, $\sigma$ denotes an arbitrarily small, strictly positive, constant. In the sequel, when it appears in some inequality (at the exponent), it means that for any $\sigma > 0$, if $s_3$ has been chosen large enough and $\nu_n$ is small enough (depending on $s_3$ and $\sigma$), the inequality holds. In particular, since $\nu_n \leq \nu_1 \leq C(s_1)\varepsilon^2$, the inequality holds for $\varepsilon$ small enough (depending on $s_3$).

**Step 1. Regularization.** We first consider the regularized approximate torus

$$\breve{i}_n(\varphi) := (\varphi, 0, 0) + \breve{\mathfrak{I}}_n(\varphi), \quad \breve{\mathfrak{I}}_n := \Pi_{N_n}\mathfrak{I}_n, \tag{11.2.9}$$

where $\Pi_N$ is the Fourier projector defined in (4.1.11) and

$$N_n \in \Big[\nu_n^{-\frac{3}{s_3 - s_1}} - 1, \nu_n^{-\frac{3}{s_3 - s_1}} + 1\Big]. \tag{11.2.10}$$

**Lemma 11.2.2.** *The torus $\breve{i}_n(\varphi) := (\varphi, 0, 0) + \breve{\mathfrak{I}}_n(\varphi)$ satisfies*

$$\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip}, s_1+2} \leq C(s_1)\varepsilon^2, \quad \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip}, s_2+2} \leq \varepsilon, \quad \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip}, s_3+\underline{\tau}+6} \leq \varepsilon^2 \nu_n^{-\frac{4}{5}-\sigma}, \tag{11.2.11}$$

$$\|i_n - \breve{i}_n\|_{\mathrm{Lip}, s_1+2} \ll \varepsilon^2 \nu_n^2, \quad \|i_n - \breve{i}_n\|_{\mathrm{Lip}, s_3+2} \leq \varepsilon^2 \nu_n^{-\frac{4}{5}}, \tag{11.2.12}$$

$$\|\mathcal{F}(\breve{i}_n)\|_{\mathrm{Lip}, s_1} \leq 2\varepsilon^2 \nu_n, \quad \|\mathcal{F}(\breve{i}_n)\|_{\mathrm{Lip}, s_3+2+\underline{\tau}} \leq \varepsilon^2 \nu_n^{-\frac{4}{5}-2\sigma}, \tag{11.2.13}$$

*where $\sigma > 0$ can be taken arbitrarily small taking $s_3 - s_1$ large enough.*

Proof. The first two estimates in (11.2.11) follow by (11.2.5) and (4.1.13). The third estimate in (11.2.11) follows by

$$\begin{aligned} \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip}, s_3+\underline{\tau}+6} &\overset{(4.1.13)}{\leq} N_n^{\underline{\tau}+4} \|\mathfrak{I}_n\|_{\mathrm{Lip}, s_3+2} \\ &\overset{(11.2.10),(11.2.5)}{\leq} C\nu_n^{-\frac{3(\underline{\tau}+4)}{s_3-s_1}} \varepsilon^2 \nu_n^{-\frac{4}{5}} \leq \varepsilon^2 \nu_n^{-\frac{4}{5}-\sigma} \end{aligned}$$

for $3(\underline{\tau}+4)/(s_3-s_1)<\sigma$.

Proof of (11.2.12). We have

$$\|i_n-\breve{\imath}_n\|_{\mathrm{Lip},s_1+2} \stackrel{(11.2.9)}{=} \|\Pi_{N_n}^{\perp}\mathfrak{I}_n\|_{\mathrm{Lip},s_1+2} \stackrel{(4.1.13)}{\leq} N_n^{-(s_3-s_1)}\|\mathfrak{I}_n\|_{\mathrm{Lip},s_3+2} \stackrel{(11.2.10),(11.2.5)}{\lesssim_{s_3}} \nu_n^3\varepsilon^2\nu_n^{-\frac{4}{5}}$$

which implies the first bound in (11.2.12). Similarly (11.2.9), (4.1.13) and (11.2.5) imply the second estimate in (11.2.12).

Proof of (11.2.13). Recalling the definition of the operator $\mathcal{F}$ in (5.1.2), and using Lemma 3.5.5, (11.2.5), (11.2.11), we have

$$\|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s_1} \leq \|\mathcal{F}(i_n)\|_{\mathrm{Lip},s_1} + C(s_1)\|i_n-\breve{\imath}_n\|_{\mathrm{Lip},s_1+2} \stackrel{(11.2.8),(11.2.12)}{\leq} 2\varepsilon^2\nu_n$$

proving the first inequality in (11.2.13). Finally Lemma 3.5.5 and (11.2.11) imply

$$\|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s_3+2+\underline{\tau}} \leq \|\mathcal{F}(i_0)\|_{\mathrm{Lip},s_3+2+\underline{\tau}} + C(s_3)\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_3+4+\underline{\tau}} \stackrel{(11.2.1),(11.2.11)}{\lesssim_{s_3}} \varepsilon^2+\varepsilon^2\nu_n^{-\frac{4}{5}-\sigma} \leq \varepsilon^2\nu_n^{-\frac{4}{5}-2\sigma}$$

for $\varepsilon$ small, proving the second inequality in (11.2.13). ■

**Step 2. Isotropic torus.** We associate to the regularized torus $\breve{\imath}_n$ defined in (11.2.9) the isotropic torus $i_{n,\delta}$ defined by Proposition 6.1.1 with $\underline{i}=\breve{\imath}_n$. Notice that, by (11.2.11), the torus $\breve{\imath}_n$ satisfies the assumption (6.1.4) required in Proposition 6.1.1. The torus $i_{n,\delta}$ is an approximate solution "essentially as good" as $i_n$ (compare (11.2.8) and (11.2.16)).

**Lemma 11.2.3.** *The isotropic torus $i_{n,\delta}(\varphi)=(\varphi,0,0)+\mathfrak{I}_{n,\delta}(\varphi)$ defined by Proposition 6.1.1 with $\underline{i}=\breve{\imath}_n$, satisfies*

$$\|i_{n,\delta}-\breve{\imath}_n\|_{\mathrm{Lip},s_0+2} \lesssim_{s_0} \varepsilon^2\nu_n\,, \qquad \|i_{n,\delta}-\breve{\imath}_n\|_{\mathrm{Lip},s_1+4} \ll \varepsilon^2\nu_n^{1-\sigma}\,, \tag{11.2.14}$$

$$\|i_{n,\delta}-\breve{\imath}_n\|_{\mathrm{Lip},s_3+4} \ll \varepsilon^2\nu_n^{-\frac{4}{5}-2\sigma}\,, \quad \|\mathfrak{I}_{n,\delta}\|_{\mathrm{Lip},s_3+4} \ll \varepsilon^2\nu_n^{-\frac{4}{5}-2\sigma}\,, \tag{11.2.15}$$

*and*

$$\begin{aligned} &\|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s_0} \lesssim_{s_0} \varepsilon^2\nu_n\,,\\ &\|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s_1+2} \leq \varepsilon^2\nu_n^{1-\sigma}\\ &\|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s_3+2} \leq \varepsilon^2\nu_n^{-\frac{4}{5}-2\sigma}\,. \end{aligned} \tag{11.2.16}$$

Proof.
Proof of (11.2.14). Since $s_0 + 2 + \underline{\tau} \leq s_1$, we have by (11.2.11),

$$\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_0+2+\underline{\tau}} \leq C(s_1)\varepsilon^2 .$$

Hence by (6.1.6) (with $s = s_0 + 2$), we have

$$\|i_{n,\delta} - \breve{\imath}_n\|_{\mathrm{Lip},s_0+2} \lesssim_{s_0} \|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s_0+2+\underline{\tau}} \lesssim_{s_0} \|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s_1} \overset{(11.2.13)}{\lesssim_{s_0}} \varepsilon^2\nu_n \tag{11.2.17}$$

which is the first estimate in (11.2.14). Similarly, since $s_1 + 2 + \underline{\tau} \leq s_2$, we have by the second estimate in (11.2.11),

$$\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_1+4+\underline{\tau}} \leq \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_2+2} \leq \varepsilon .$$

Hence by (6.1.6) (with $s = s_1 + 4$),

$$\|i_{n,\delta} - \breve{\imath}_n\|_{\mathrm{Lip},s_1+4} \lesssim_{s_1} \|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s_1+4+\underline{\tau}} . \tag{11.2.18}$$

Now, by the interpolation inequality (3.5.10) we have

$$\|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s_1+4+\underline{\tau}} \lesssim_{s_3} \|\mathcal{F}(\breve{\imath}_n)\|^{\theta}_{\mathrm{Lip},s_1}\|\mathcal{F}(\breve{\imath}_n)\|^{1-\theta}_{\mathrm{Lip},s_3} \tag{11.2.19}$$

where

$$\theta := 1 - \frac{4+\underline{\tau}}{s_3 - s_1}, \quad 1-\theta := \frac{4+\underline{\tau}}{s_3 - s_1} . \tag{11.2.20}$$

Therefore (11.2.19), (11.2.13), (11.2.20) imply

$$\|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s_1+4+\underline{\tau}} \lesssim_{s_3} (\varepsilon^2\nu_n)^{\theta}(\varepsilon^2\nu_n^{-\frac{4}{5}-2\sigma})^{1-\theta} \ll \varepsilon^2\nu_n^{1-\sigma} \tag{11.2.21}$$

for $s_3 - s_1$ large enough and $\nu_n$ small. Inequalities (11.2.18) and (11.2.21) prove the second estimate in (11.2.14).
Proof of (11.2.15). By (6.2.6) we have

$$\|\mathfrak{I}_{n,\delta} - \breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_3+4} \lesssim_{s_3} \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_3+5} \overset{(11.2.11)}{\ll} \varepsilon^2\nu_n^{-\frac{4}{5}-2\sigma}$$

for $\nu_n$ small enough, proving the first inequality in (11.2.15). The second inequality in (11.2.15) follows as a consequence of the first one and the third estimate in (11.2.11).
Proof of (11.2.16). Recalling the definition of the operator $\mathcal{F}$ in (5.1.2) and using Lemma 3.5.5, (11.2.11), we have

$$\|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s} \leq \|\mathcal{F}(\breve{\imath}_n)\|_{\mathrm{Lip},s} + C(s)\|i_{n,\delta} - \breve{\imath}_n\|_{\mathrm{Lip},s+2}$$

both for $s = s_0$ and $s = s_1 + 2$, and therefore

$$\|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s_0} \overset{(11.2.13)(11.2.14)}{\lesssim_{s_0}} \varepsilon^2 \nu_n \,,$$
$$\|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s_1+2} \overset{(11.2.21)(11.2.14)}{\leq} \varepsilon^2 \nu_n^{1-\sigma} \,,$$

proving the first two estimates in (11.2.16). Finally

$$\begin{aligned} \|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s_3+2} &\leq \|\mathcal{F}(i_0)\|_{\mathrm{Lip},s_3+2} + \|\mathcal{F}(i_{n,\delta}) - \mathcal{F}(i_0)\|_{\mathrm{Lip},s_3+2} \\ &\overset{(11.2.1), Lemma\, 3.5.5}{\lesssim_{s_3}} \varepsilon^2 + \|\mathfrak{I}_{n,\delta}\|_{\mathrm{Lip},s_3+4} \\ &\overset{(11.2.15)}{\leq} \varepsilon^2 \nu_n^{-\frac{4}{5}-2\sigma} \end{aligned} \tag{11.2.22}$$

for $\nu_n$ small, which is the third inequality in (11.2.16). ■

**Step 3. Symplectic diffeomorphism.** We apply the symplectic change of variables $G_{n,\delta}$ defined in (6.1.9) with $(\underline{\theta}(\phi), y_\delta(\phi), \underline{z}(\phi)) = i_{n,\delta}(\phi)$, which transforms the isotropic torus $i_{n,\delta}$ into (see (6.1.10))

$$G_{n,\delta}^{-1}(i_{n,\delta}(\varphi)) = (\varphi, 0, 0)\,. \tag{11.2.23}$$

It conjugates the Hamiltonian vector field $X_K$ associated to $K$ defined in (2.2.7), with the Hamiltonian vector field (see (6.1.13))

$$X_{\mathtt{K}_n} = (DG_{n,\delta})^{-1} X_K \circ G_{n,\delta} \qquad \text{where} \qquad \mathtt{K}_n := K \circ G_{n,\delta}\,. \tag{11.2.24}$$

Denote by $\mathtt{u} := (\phi, \zeta, w)$ the symplectic coordinates induced by the diffeomorphism $G_{n,\delta}$ in (6.1.9). Under the symplectic map $G_{n,\delta}$, the nonlinear operator $\mathcal{F}$ in (5.1.2) is transformed into

$$\mathtt{F}_n(\mathtt{u}(\varphi)) := \omega\cdot\partial_\varphi \mathtt{u}(\varphi) - X_{\mathtt{K}_n}(\mathtt{u}(\varphi)) = (DG_{n,\delta}(\mathtt{u}(\varphi)))^{-1}\mathcal{F}(G_{n,\delta}(\mathtt{u}(\varphi)))\,. \tag{11.2.25}$$

By (11.2.25) and (11.2.23) (see also (6.1.15)) we have that

$$\mathtt{F}_n(\varphi, 0, 0) = (\omega, 0, 0) - X_{\mathtt{K}_n}(\varphi, 0, 0) = (DG_{n,\delta}(\varphi, 0, 0))^{-1}\mathcal{F}\big(i_{n,\delta}(\varphi)\big)\,. \tag{11.2.26}$$

Hence by (6.1.11), for $s \geq s_0$,

$$\begin{aligned} \|\mathtt{F}_n(\varphi, 0, 0)\|_{\mathrm{Lip},s} &\lesssim_s \|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s} + \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s+2}\|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s_0} \\ &\overset{(11.2.16)}{\lesssim_s} \|\mathcal{F}(i_{n,\delta})\|_{\mathrm{Lip},s} + \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s+2}\, \varepsilon^2 \nu_n\,. \end{aligned} \tag{11.2.27}$$

By (11.2.27), (11.2.16) and (11.2.11), we have

$$\begin{aligned} &\|\mathtt{F}_n(\varphi, 0, 0)\|_{\mathrm{Lip},s_0} \lesssim_{s_0} \varepsilon^2 \nu_n \\ &\|\mathtt{F}_n(\varphi, 0, 0)\|_{\mathrm{Lip},s_1+2} \leq \varepsilon^2 \nu_n^{1-2\sigma} \\ &\|\mathtt{F}_n(\varphi, 0, 0)\|_{\mathrm{Lip},s_3+2} \leq \varepsilon^2 \nu_n^{-\frac{4}{5}-3\sigma}\,. \end{aligned} \tag{11.2.28}$$

**Step 4. Approximate solution of the linear equation associated to a Nash-Moser step in new coordinates.** In order to look for a better approximate zero

$$\mathtt{u}_{n+1}(\varphi) = (\varphi, 0, 0) + \mathtt{h}_{n+1}(\varphi) \tag{11.2.29}$$

of the operator $\mathtt{F}_n(\mathtt{u})$ defined in (11.2.25), we expand

$$\mathtt{F}_n(\mathtt{u}_{n+1}) = \mathtt{F}_n(\varphi, 0, 0) + \big(\omega\cdot\partial_\varphi - d_{\mathtt{u}} X_{\mathtt{K}_n}(\varphi, 0, 0)\big)\mathtt{h}_{n+1} + Q_n(\mathtt{h}_{n+1}) \tag{11.2.30}$$

where $Q_n(\mathtt{h}_{n+1})$ is a quadratic remainder, and we want to solve (approximately) the linear equation

$$\mathtt{F}_n(\varphi, 0, 0) + \big(\omega\cdot\partial_\varphi - d_{\mathtt{u}} X_{\mathtt{K}_n}(\varphi, 0, 0)\big)\mathtt{h}_{n+1} = 0\,. \tag{11.2.31}$$

This is the goal of the present step. Notice that the operator $\omega\cdot\partial_\varphi - d_{\mathtt{u}} X_{\mathtt{K}_n}(\varphi, 0, 0)$ is provided by (6.1.21), and we decompose it as

$$\omega\cdot\partial_\varphi - d_{\mathtt{u}} X_{\mathtt{K}_n}(\varphi, 0, 0) = \mathbb{D}_n + \mathtt{R}_{Z_n} \tag{11.2.32}$$

where $\mathbb{D}_n$ is the operator in (6.1.22) with $\underline{i}$ replaced by $\breve{\imath}_n$, i.e.

$$\mathbb{D}_n \begin{pmatrix} \widehat\phi \\ \widehat\zeta \\ \widehat w \end{pmatrix} := \begin{pmatrix} \omega\cdot\partial_\varphi \widehat\phi - \mathtt{K}_{20}^{(n)}(\varphi)\widehat\zeta - [\mathtt{K}_{11}^{(n)}]^\top(\varphi)\widehat w \\ \omega\cdot\partial_\varphi\widehat\zeta \\ \mathcal{L}_\omega^{(n)}\widehat w - J\mathtt{K}_{11}^{(n)}(\varphi)\widehat\zeta \end{pmatrix} \tag{11.2.33}$$

with $\mathcal{L}_\omega^{(n)} := \mathcal{L}_\omega(\breve{\imath}_n)$ (see (6.1.23)), the functions $\mathtt{K}_{20}^{(n)}(\varphi)$, $\mathtt{K}_{11}^{(n)}(\varphi)$ are the Taylor coefficients in (6.1.14) of the Hamiltonian $\mathtt{K}_n$, and

$$\mathtt{R}_{Z_n} \begin{pmatrix} \widehat\phi \\ \widehat\zeta \\ \widehat w \end{pmatrix} := \begin{pmatrix} -\partial_\phi \mathtt{K}_{10}^{(n)}(\varphi)[\widehat\phi] \\ \partial_{\phi\phi}\mathtt{K}_{00}^{(n)}(\varphi)[\widehat\phi] + [\partial_\phi\mathtt{K}_{10}^{(n)}(\varphi)]^\top\widehat\zeta + [\partial_\phi\mathtt{K}_{01}^{(n)}(\varphi)]^\top\widehat w \\ -J\{\partial_\phi\mathtt{K}_{01}^{(n)}(\varphi)[\widehat\phi]\} \end{pmatrix}\,. \tag{11.2.34}$$

In the next proposition we define $\mathtt{h}_{n+1}$ as an approximate solution of the linear equation $\mathtt{F}_n(\varphi, 0, 0) + \mathbb{D}_n\mathtt{h}_{n+1} = 0$.

**Proposition 11.2.4. (Approximate right inverse of $\mathbb{D}_n$)** *For all $\lambda$ in the set*

$$\mathbf{\Lambda}_{n+1} := \mathbf{\Lambda}_n \cap \mathbf{\Lambda}(\varepsilon; \eta_n, \breve{\mathfrak{I}}_n)\,, \qquad \eta_1 := 1/2\,, \qquad \eta_n := \eta_{n-1} + \nu_{n-1}^{\frac14}\,,\ n \geq 2\,, \tag{11.2.35}$$

*where $\mathbf{\Lambda}(\varepsilon;\eta,\underline{\mathfrak{I}})$ is defined in Proposition 11.1.1, there exists $\mathtt{h}_{n+1}$ satisfying*

$$\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1+2} \lesssim_{s_1} \nu_n + \varepsilon^2\nu_n^{\frac45-5\sigma}\,, \quad \|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_3+2} \leq \varepsilon^2\nu_n^{-\frac{11}{10}-2\sigma}\,, \tag{11.2.36}$$

*such that*

$$\mathtt{r}_{n+1} := \mathtt{F}_n(\varphi, 0, 0) + \mathbb{D}_n\,\mathtt{h}_{n+1} \tag{11.2.37}$$

*satisfies*

$$\|\mathtt{r}_{n+1}\|_{\mathrm{Lip},s_1} \leq \varepsilon^2\nu_n^{\frac32-6\sigma}\,. \tag{11.2.38}$$

PROOF. Recalling (11.2.33) we look for an approximate solution $\mathtt{h}_{n+1} = (\widehat{\phi}, \widehat{\zeta}, \widehat{w})$ of

$$\mathbb{D}_n \begin{pmatrix} \widehat{\phi} \\ \widehat{\zeta} \\ \widehat{w} \end{pmatrix} = \begin{pmatrix} \omega \cdot \partial_\varphi \widehat{\phi} - \mathtt{K}_{20}^{(n)}(\varphi)\widehat{\zeta} - [\mathtt{K}_{11}^{(n)}]^\top(\varphi)\widehat{w} \\ \omega \cdot \partial_\varphi \widehat{\zeta} \\ \mathcal{L}_\omega^{(n)} \widehat{w} - J\mathtt{K}_{11}^{(n)}(\varphi)\widehat{\zeta} \end{pmatrix} = \begin{pmatrix} \mathtt{f}_1^{(n)} \\ \mathtt{f}_2^{(n)} \\ \mathtt{f}_3^{(n)} \end{pmatrix} \tag{11.2.39}$$

where $(\mathtt{f}_1^{(n)}, \mathtt{f}_2^{(n)}, \mathtt{f}_3^{(n)}) := -\mathtt{F}_n(\varphi, 0, 0)$. Since $X_{\mathtt{K}_n}$ is a reversible vector field, its components $\mathtt{f}_1^{(n)}$, $\mathtt{f}_2^{(n)}$, $\mathtt{f}_3^{(n)}$ satisfy the reversibility property

$$\mathtt{f}_1^{(n)}(\varphi) = \mathtt{f}_1^{(n)}(-\varphi)\,, \quad \mathtt{f}_2^{(n)}(\varphi) = -\mathtt{f}_2^{(n)}(-\varphi)\,, \quad \mathtt{f}_3^{(n)}(\varphi) = -(S\mathtt{f}_3^{(n)})(-\varphi)\,. \tag{11.2.40}$$

We solve approximately (11.2.39) in a triangular way.

STEP 1. APPROXIMATE SOLUTION OF THE SECOND EQUATION IN (11.2.39), i.e. $\omega \cdot \partial_\varphi \widehat{\zeta} = \mathtt{f}_2^{(n)}$. We solve the equation

$$\omega \cdot \partial_\varphi \widehat{\zeta} = \Pi_{N_n} \mathtt{f}_2^{(n)}(\varphi) \tag{11.2.41}$$

where $N_n$ is defined in (11.2.10) and the projector $\Pi_{N_n}$ applies to functions depending only on the variable $\varphi$ as in (9.3.9). Notice that $\omega$ is a Diophantine vector by (6.2.1) for all $\lambda \in \Lambda$, and that, by (11.2.40), the $\varphi$-average of $\Pi_{N_n} \mathtt{f}_2^{(n)}$ is zero. The solutions of (11.2.41) are

$$\widehat{\zeta} = \widehat{\zeta}_0 + [\widehat{\zeta}]\,, \quad [\widehat{\zeta}] := (\omega \cdot \partial_\varphi)^{-1} \Pi_{N_n} \mathtt{f}_2^{(n)}\,, \quad \widehat{\zeta}_0 \in \mathbb{R}^{|\mathbb{S}|}\,, \tag{11.2.42}$$

where $(\omega \cdot \partial_\varphi)^{-1}$ is defined in (1.6.1) and $\widehat{\zeta}_0$ is a free parameter that we fix below in (11.2.54). By (11.2.42), (6.2.1), (2.3.7), we have

$$\|[\widehat{\zeta}]\|_{\mathrm{Lip},s} \lesssim \|\Pi_{N_n} \mathtt{f}_2^{(n)}\|_{\mathrm{Lip},s+\tau_1} \overset{(4.1.13)}{\lesssim} N_n^{\tau_1} \|\mathtt{f}_2^{(n)}\|_{\mathrm{Lip},s} \tag{11.2.43}$$

($\gamma_2 = \gamma_0/4$ is considered as a fixed constant). Therefore, using (11.2.43), (11.2.28), (11.2.10), and taking $s_3 - s_1$ large enough, we have for $\nu_n$ small enough (depending on $s_3$)

$$\|[\widehat{\zeta}]\|_{\mathrm{Lip},s_1+2} \le \varepsilon^2 \nu_n^{1-3\sigma}\,, \quad \|[\widehat{\zeta}]\|_{\mathrm{Lip},s_3+2} \le \varepsilon^2 \nu_n^{-\frac{4}{5}-4\sigma}\,. \tag{11.2.44}$$

STEP 2. APPROXIMATE SOLUTION OF THE THIRD EQUATION IN (11.2.39), i.e.

$$\mathcal{L}_\omega^{(n)} \widehat{w} = \mathtt{f}_3^{(n)} + J\mathtt{K}_{11}^{(n)} \widehat{\zeta}_0 + J\mathtt{K}_{11}^{(n)} [\widehat{\zeta}]\,. \tag{11.2.45}$$

For some $\nu \in (0, \varepsilon^{\frac{3}{2}})$ such that

$$\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_3+4} \le \varepsilon \nu^{-\frac{9}{10}} \tag{11.2.46}$$

and which will be chosen later in (11.2.58), we consider the operator $\mathcal{L}_{appr,\nu}^{-1}$ defined in Proposition 11.1.1 (with $\underline{\mathfrak{I}} \rightsquigarrow \breve{\mathfrak{I}}_n$). Notice that, by (11.2.11), also the assumption $\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_2+2} \le \varepsilon$

required in Proposition 11.1.1 with $\underline{\mathfrak{I}} \rightsquigarrow \breve{\mathfrak{I}}_n$ holds. Moreover $\mathcal{L}^{-1}_{appr,\nu}$ satisfies (11.1.5) (independently of $\nu$). We define the approximate solution $\widehat{w}$ of (11.2.45),

$$\widehat{w} := \mathcal{L}^{-1}_{appr,\nu}\big(\mathtt{f}_3^{(n)} + J\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]\big) + \mathcal{L}^{-1}_{appr,\nu} J\mathtt{K}_{11}^{(n)}\widehat{\zeta}_0\,. \tag{11.2.47}$$

Step 3. Approximate solution of the first equation in (11.2.39). With (11.2.47), the first equation in (11.2.39) can be written as

$$\omega\cdot\partial_\varphi\widehat{\phi} = T_n\widehat{\zeta}_0 + g_n \tag{11.2.48}$$

where

$$\begin{aligned} T_n &:= \mathtt{K}_{20}^{(n)} + [\mathtt{K}_{11}^{(n)}]^\top \mathcal{L}^{-1}_{appr} J\mathtt{K}_{11}^{(n)}\,, \\ g_n &:= \mathtt{f}_1^{(n)} + \mathtt{K}_{20}^{(n)}[\widehat{\zeta}] + [\mathtt{K}_{11}^{(n)}]^\top \mathcal{L}^{-1}_{appr}\big(\mathtt{f}_3^{(n)} + J\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]\big)\,. \end{aligned} \tag{11.2.49}$$

We have to choose the constant $\widehat{\zeta}_0 \in \mathbb{R}^{|\mathbb{S}|}$ such that the right hand side in (11.2.48) has zero average. By (6.1.17), the matrix

$$\langle\mathtt{K}_{20}^{(n)}\rangle := (2\pi)^{-|\mathbb{S}|}\int_{\mathbb{T}^{|\mathbb{S}|}} \mathtt{K}_{20}^{(n)}(\varphi)d\varphi$$

is close to the invertible twist matrix $\varepsilon^2\mathscr{A}$ (see (1.2.12)) and therefore by (1.2.12) there is a constant $C$ such that

$$\|\langle\mathtt{K}_{20}^{(n)}\rangle^{-1}\| \le C\varepsilon^{-2}\,. \tag{11.2.50}$$

By (11.2.49) the operator $\langle T_n\rangle : \zeta_0 \in \mathbb{R}^{|\mathbb{S}|} \mapsto \langle T_n\zeta_0\rangle \in \mathbb{R}^{|\mathbb{S}|}$ is decomposed as

$$\langle T_n\rangle := \langle\mathtt{K}_{20}^{(n)}\rangle + \langle[\mathtt{K}_{11}^{(n)}]^\top \mathcal{L}^{-1}_{appr} J\mathtt{K}_{11}^{(n)}\rangle\,. \tag{11.2.51}$$

Now, for $\zeta_0 \in \mathbb{R}^{|\mathbb{S}|}$,

$$\begin{aligned} |\langle[\mathtt{K}_{11}^{(n)}]^\top \mathcal{L}^{-1}_{appr} J\mathtt{K}_{11}^{(n)}\zeta_0\rangle|_{\mathrm{Lip}} &\le \|[\mathtt{K}_{11}^{(n)}]^\top \mathcal{L}^{-1}_{appr} J\mathtt{K}_{11}^{(n)}\zeta_0\|_{\mathrm{Lip},s_0} \\ &\overset{(6.1.20),(11.2.11)}{\le} \varepsilon^2\|\mathcal{L}^{-1}_{appr} J\mathtt{K}_{11}^{(n)}\zeta_0\|_{\mathrm{Lip},s_0} \\ &\overset{(11.1.5)}{\le} C(s_1)\varepsilon^2\|\mathtt{K}_{11}^{(n)}\zeta_0\|_{\mathrm{Lip},s_0+Q'} \\ &\overset{(6.1.19)}{\le} C(s_1)\varepsilon^4\big(|\zeta_0|_{\mathrm{Lip}} + \|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_0+\underline{\tau}+Q'}|\zeta_0|_{\mathrm{Lip}}\big) \\ &\overset{(11.2.11)}{\le} C(s_1)\varepsilon^4|\zeta_0|_{\mathrm{Lip}} \end{aligned} \tag{11.2.52}$$

using the fact that $s_0 + \underline{\tau} + Q' \le s_1$. By (11.2.51), (11.2.50), (11.2.52), we deduce that, for $\varepsilon$ small enough, $\langle T_n\rangle$ is invertible and

$$\|\langle T_n\rangle^{-1}\| \le 2C\varepsilon^{-2}\,. \tag{11.2.53}$$

Thus, in view of (11.2.48), we define

$$\widehat{\zeta}_0 := -\langle T_n\rangle^{-1}\langle g_n\rangle \in \mathbb{R}^{|\mathbb{S}|}\,. \tag{11.2.54}$$

We now estimate the function $g_n$ defined in (11.2.49). By (6.1.18), (6.1.20), (11.2.11), (11.1.5), and using that $s_1 > s_0 + Q' + \underline{\tau}$,

$$\begin{aligned}\|g_n\|_{\mathrm{Lip},s_0} &\leq \|\mathtt{f}_1^{(n)}\|_{\mathrm{Lip},s_0} + \|\mathtt{K}_{20}^{(n)}[\widehat{\zeta}]\|_{\mathrm{Lip},s_0} + \|[\mathtt{K}_{11}^{(n)}]^\top \mathcal{L}_{appr}^{-1}\big(\mathtt{f}_3^{(n)} + J\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]\big)\|_{\mathrm{Lip},s_0}\\ &\lesssim_{s_0} \|\mathtt{f}_1^{(n)}\|_{\mathrm{Lip},s_0} + \varepsilon^2\|[\widehat{\zeta}]\|_{\mathrm{Lip},s_0} + \varepsilon^2\big(\|\mathtt{f}_3^{(n)}\|_{\mathrm{Lip},s_0+Q'} + \|\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]\|_{\mathrm{Lip},s_0+Q'}\big)\,.\end{aligned}$$

Hence by (11.2.44), (6.1.19), (11.2.11),

$$\begin{aligned}\|g_n\|_{\mathrm{Lip},s_0} &\lesssim_{s_0} \|\mathtt{f}_1^{(n)}\|_{\mathrm{Lip},s_0} + \varepsilon^4\nu_n^{1-3\sigma} + \varepsilon^2\big(\|\mathtt{f}_3^{(n)}\|_{\mathrm{Lip},s_0+Q'} + \varepsilon^4\nu_n^{1-3\sigma}\big)\\ &\overset{(11.2.28)}{\lesssim_{s_1}} \varepsilon^2\nu_n + \varepsilon^4\nu_n^{1-3\sigma} + \varepsilon^2\big(\varepsilon^2\nu_n^{1-2\sigma} + \varepsilon^4\nu_n^{1-3\sigma}\big)\\ &\lesssim_{s_1} \varepsilon^2\nu_n + \varepsilon^4\nu_n^{1-3\sigma}\,.\end{aligned} \tag{11.2.55}$$

Then the constant $\widehat{\zeta}_0$ defined in (11.2.54) satisfies, by (11.2.53), (11.2.55),

$$|\widehat{\zeta}_0|_{\mathrm{Lip}} \lesssim_{s_1} \nu_n + \varepsilon^2\nu_n^{1-3\sigma}\,. \tag{11.2.56}$$

Using (11.2.28), (6.1.19), (11.2.56), (11.2.44) and (11.2.11), we have

$$\begin{aligned}&\|\mathtt{f}_3^{(n)}\|_{\mathrm{Lip},s_1} + \|J\mathtt{K}_{11}^{(n)}\widehat{\zeta}_0\|_{\mathrm{Lip},s_1} + \|J\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]\|_{\mathrm{Lip},s_1} \leq \varepsilon^2\nu_n^{1-3\sigma}\,,\\ &\|\mathtt{f}_3^{(n)}\|_{\mathrm{Lip},s_3+2} + \|J\mathtt{K}_{11}^{(n)}\widehat{\zeta}_0\|_{\mathrm{Lip},s_3+2} + \|J\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]\|_{\mathrm{Lip},s_3+2} \leq \varepsilon^2\nu_n^{-\frac{4}{5}-5\sigma}\end{aligned} \tag{11.2.57}$$

for $\nu_n$ small enough.

We now choose the constant $\nu$ in (11.2.46) as

$$\nu := \nu_n^{1-3\sigma} \tag{11.2.58}$$

so that $\nu \in (0, \varepsilon^{3/2})$ and, by the third inequality in (11.2.11), condition (11.2.46) is satisfied, provided $\sigma$ is chosen small enough (and hence $s_3$ is large enough). We apply Proposition 11.1.1 with $g = \mathtt{f}_3^{(n)} + J\mathtt{K}_{11}^{(n)}\widehat{\zeta}_0 + J\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]$ (and $\underline{\mathfrak{I}} \rightsquigarrow \breve{\mathfrak{I}}_n$) noting that (11.2.57) implies (11.1.2) with $\nu$ defined in (11.2.58), again provided $\sigma$ is chosen small enough. As a consequence, by (11.1.3) the function $\widehat{w}$ defined in (11.2.47) satisfies (for $\nu_n$ small enough)

$$\begin{aligned}\|\widehat{w}\|_{\mathrm{Lip},s_1} &\leq C(s_1)\varepsilon^2\nu_n^{\frac{4}{5}(1-3\sigma)} \leq \varepsilon^2\nu_n^{\frac{4}{5}-3\sigma}\,,\\ \|\widehat{w}\|_{\mathrm{Lip},s_3+2} &\leq C(s_3)\varepsilon^2\nu_n^{-\frac{11}{10}(1-3\sigma)} \leq \varepsilon^2\nu_n^{-\frac{11}{10}+3\sigma}\,.\end{aligned} \tag{11.2.59}$$

By (11.2.59) and interpolation inequality (3.5.10), arguing as above (11.2.21), we obtain

$$\|\widehat{w}\|_{\mathrm{Lip},s_1+2} \leq \varepsilon^2 \nu_n^{\frac{4}{5}-4\sigma}, \tag{11.2.60}$$

for $s_3$ large enough and $\nu_n$ small enough (depending on $s_3$). Moreover, by (11.1.4),

$$\mathcal{L}_\omega^{(n)} \widehat{w} - \big(\mathtt{f}_3^{(n)} + J\mathtt{K}_{11}^{(n)} \widehat{\zeta}_0 + J\mathtt{K}_{11}^{(n)}[\widehat{\zeta}]\big) = \mathfrak{r}_{n+1}, \tag{11.2.61}$$

where $\mathfrak{r}_{n+1}$ satisfies

$$\|\mathfrak{r}_{n+1}\|_{\mathrm{Lip},s_1} \leq C(s_1)\varepsilon^2 \nu_n^{\frac{3}{2}(1-3\sigma)} \leq \varepsilon^2 \nu_n^{\frac{3}{2}-5\sigma}. \tag{11.2.62}$$

Now, since $\widehat{\zeta}_0$ has been chosen in (11.2.54) so that $T_n\widehat{\zeta}_0 + g_n$ has zero mean value, the equation (11.2.48) has the approximate solution

$$\widehat{\phi} := (\omega \cdot \partial_\varphi)^{-1} \Pi_{N_n}(T_n\widehat{\zeta}_0 + g_n). \tag{11.2.63}$$

Recalling the definition of $g_n$ and $T_n$ in (11.2.49) we have

$$g_n + T_n\widehat{\zeta}_0 = \mathtt{f}_1^{(n)} + \mathtt{K}_{20}^{(n)}\widehat{\zeta} + [\mathtt{K}_{11}^{(n)}]^\top \widehat{w} \tag{11.2.64}$$

where $\widehat{w}$ is defined in (11.2.47). By (11.2.64) we estimate

$$\begin{aligned}
\|g_n + T_n\widehat{\zeta}_0\|_{\mathrm{Lip},s_1} &\overset{(6.1.18),(6.1.20),(11.2.11)}{\leq} \|\mathtt{f}_1^{(n)}\|_{\mathrm{Lip},s_1} + C(s_1)\varepsilon^2\big(\|\widehat{\zeta}\|_{\mathrm{Lip},s_1} + \|\widehat{w}\|_{\mathrm{Lip},s_1}\big) \\
&\overset{(11.2.28),(11.2.44),(11.2.56),(11.2.59)}{\leq} \varepsilon^2\nu_n^{1-2\sigma} + C(s_1)\varepsilon^2\big(\nu_n + \varepsilon^2\nu_n^{1-3\sigma} + \varepsilon^2\nu_n^{\frac{4}{5}-3\sigma}\big) \\
&\leq \varepsilon^2\nu_n^{\frac{4}{5}-3\sigma}
\end{aligned} \tag{11.2.65}$$

and, similarly, by (6.1.18), (6.1.20),(11.2.11), (11.2.28), (11.2.44), (11.2.56), (11.2.59), we get

$$\begin{aligned}
\|g_n + T_n\widehat{\zeta}_0\|_{\mathrm{Lip},s_3} &\leq \|\mathtt{f}_1^{(n)}\|_{\mathrm{Lip},s_3} + C(s_3)\varepsilon^2\big(\|\widehat{\zeta}\|_{\mathrm{Lip},s_3} + \|\widehat{w}\|_{\mathrm{Lip},s_3}\big) \\
&\qquad + C(s_3)\varepsilon^2\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_3+\underline{\tau}}\big(\|\widehat{\zeta}\|_{\mathrm{Lip},s_0} + \|\widehat{w}\|_{\mathrm{Lip},s_0}\big) \\
&\leq \varepsilon^2\nu_n^{-\frac{4}{5}-3\sigma} + C(s_3)\varepsilon^2\nu_n^{-\frac{11}{10}+3\sigma} + C(s_3)\varepsilon^2\nu_n^{-\frac{4}{5}-\sigma}\nu_n^{\frac{4}{5}-3\sigma} \\
&\leq \varepsilon^2\nu_n^{-\frac{11}{10}}.
\end{aligned} \tag{11.2.66}$$

The function $\widehat{\phi}$ defined in (11.2.63) satisfies by (6.2.1), (2.3.7), (4.1.13), (11.2.10), for $s_3 - s_1$ large enough, and (11.2.65), (11.2.66),

$$\|\widehat{\phi}\|_{\mathrm{Lip},s_1} \leq \varepsilon^2\nu_n^{\frac{4}{5}-4\sigma}, \quad \|\widehat{\phi}\|_{\mathrm{Lip},s_3+2} \leq \varepsilon^2\nu_n^{-\frac{11}{10}-\sigma}. \tag{11.2.67}$$

Using again the interpolation inequality (3.5.10), we deduce by (11.2.67), the estimate

$$\|\widehat{\phi}\|_{\mathrm{Lip},s_1+2} \le \varepsilon^2 \nu_n^{\frac{4}{5}-5\sigma}, \tag{11.2.68}$$

for $s_3$ large enough and $\nu_n$ small enough. Moreover the remainder

$$\mathfrak{r}_3^{(n+1)} := \Pi_{N_n}^{\perp}(T_n \widehat{\zeta}_0 + g_n) \tag{11.2.69}$$

satisfies

$$\|\mathfrak{r}_3^{(n+1)}\|_{\mathrm{Lip},s_1} \overset{(4.1.13)}{\le} N_n^{-(s_3-s_1)} \|T_n \widehat{\zeta}_0 + g_n\|_{\mathrm{Lip},s_3} \overset{(11.2.10),(11.2.66)}{\le} \varepsilon^2 \nu_n^{\frac{3}{2}}. \tag{11.2.70}$$

Step 4. Conclusion. We set $\mathtt{h}_{n+1} := (\widehat{\phi}, \widehat{\zeta}, \widehat{w})$ defined in (11.2.63), (11.2.42), (11.2.54) and (11.2.47). The estimates in (11.2.36) follow by (11.2.67)-(11.2.68), (11.2.44), (11.2.56), (11.2.59)-(11.2.60) and recalling that $\nu_n \le \nu_1 = C(s_1)\varepsilon^2$. Finally, by (11.2.39), (11.2.42), (11.2.54), (11.2.61), we have that

$$\mathbb{D}_n \begin{pmatrix} \widehat{\phi} \\ \widehat{\zeta} \\ \widehat{w} \end{pmatrix} - \begin{pmatrix} \mathtt{f}_1^{(n)} \\ \mathtt{f}_2^{(n)} \\ \mathtt{f}_3^{(n)} \end{pmatrix} = \begin{pmatrix} -\mathfrak{r}_3^{(n+1)} \\ \Pi_{N_n}^{\perp} \mathtt{f}_2^{(n)} \\ \mathfrak{r}_{n+1} \end{pmatrix} \tag{11.2.71}$$

where $\mathfrak{r}_3^{(n+1)}$ is defined in (11.2.69) and $\mathfrak{r}_{n+1}$ in (11.2.61). The estimate (11.2.38) follows by (11.2.71), (11.2.70), the bound

$$\|\Pi_{N_n}^{\perp} \mathtt{f}_2^{(n)}\|_{\mathrm{Lip},s_1} \overset{(4.1.13)}{\le} N_n^{-(s_3-s_1)} \|\mathtt{f}_2^{(n)}\|_{\mathrm{Lip},s_3} \overset{(11.2.10),(11.2.28)}{\ll} \varepsilon^2 \nu_n^{\frac{3}{2}}$$

and (11.2.62). ■

The function $\mathtt{h}_{n+1}$ defined in Proposition 11.2.4 is an approximate solution of equation (11.2.31), according to the following corollary.

**Corollary 11.2.5.** *The function*

$$\mathtt{r}'_{n+1} := \mathtt{F}_n(\varphi, 0, 0) + \big(\omega \cdot \partial_\varphi - d_{\mathtt{u}} X_{\mathtt{K}_n}(\varphi, 0, 0)\big)\mathtt{h}_{n+1} \tag{11.2.72}$$

*satisfies*

$$\|\mathtt{r}'_{n+1}\|_{\mathrm{Lip},s_1} \le \varepsilon^2 \nu_n^{\frac{3}{2}-8\sigma}. \tag{11.2.73}$$

Proof. The term $\mathtt{r}'_{n+1}$ in (11.2.72) is, by (11.2.32) and (11.2.37),

$$\mathtt{r}'_{n+1} = \mathtt{r}_{n+1} + \mathtt{R}_{Z_n} \mathtt{h}_{n+1}. \tag{11.2.74}$$

By the expression of $\mathtt{R}_{Z_n}$ in (11.2.34), using the tame estimate (3.5.1) for the product of functions and (6.2.13), we get

$$\begin{aligned} \|\mathtt{R}_{Z_n}\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1} &\lesssim_{s_1} \big(\|\mathtt{K}_{10}^{(n)}-\omega\|_{\mathrm{Lip},s_1+1}+\|\mathtt{K}_{01}^{(n)}\|_{\mathrm{Lip},s_1+1}+\|\partial_\phi\mathtt{K}_{00}^{(n)}\|_{\mathrm{Lip},s_1+1}\big)\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1} \\ &\overset{(6.1.16),(11.2.11)}{\lesssim_{s_1}} \|\mathtt{F}_n(\varphi,0,0)\|_{\mathrm{Lip},s_1+1+\underline{\tau}}\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1} \end{aligned} \tag{11.2.75}$$

having used that $s_1+1+\underline{\tau}\le s_2$ and the estimate $\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_2}\le\varepsilon$ in (11.2.11). Now, by (11.2.28) and the interpolation inequality (3.5.10), we get, for $s_3-s_1$ large enough,

$$\|\mathtt{F}_n(\varphi,0,0)\|_{\mathrm{Lip},s_1+1+\underline{\tau}}\le\varepsilon^2\nu_n^{1-3\sigma}\,. \tag{11.2.76}$$

So we derive, by (11.2.75), (11.2.76) and (11.2.36), the estimate

$$\|\mathtt{R}_{Z_n}\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1}\le\varepsilon^2\nu_n^{\frac{9}{5}-8\sigma}\,. \tag{11.2.77}$$

In conclusion (11.2.74), (11.2.38), (11.2.77) imply (11.2.73). ■

**Step 5. Approximate solution $i_{n+1}$.** Finally, for all $\lambda$ in the set $\mathbf{\Lambda}_{n+1}$ introduced in (11.2.35), we define the new approximate solution of the Nash-Moser iteration in the original coordinates as

$$i_{n+1}:=i_{n,\delta}+h_{n+1}\,,\quad h_{n+1}:=DG_{n,\delta}(\varphi,0,0)\mathtt{h}_{n+1}\,, \tag{11.2.78}$$

where $i_{n,\delta}$ is the isotropic torus defined in Lemma 11.2.3 and $\mathtt{h}_{n+1}$ is the function defined in Proposition 11.2.4. By (6.1.11), (11.2.11) coupled with the inequality $s_1+4\le s_2+2$ and (11.2.36), we have

$$\|h_{n+1}\|_{\mathrm{Lip},s_1+2}\lesssim_{s_1}\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1+2}\lesssim_{s_1}\nu_n+\varepsilon^2\nu_n^{\frac{4}{5}-5\sigma} \tag{11.2.79}$$

$$\|h_{n+1}\|_{\mathrm{Lip},s_3+2}\lesssim_{s_3}\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_3+2}+\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_3+4}\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_0}\le\varepsilon^2\nu_n^{-\frac{11}{10}-3\sigma} \tag{11.2.80}$$

for $\nu_n$ small enough.

**Lemma 11.2.6.** *The term*

$$\varrho_{n+1}:=\mathcal{F}(i_{n,\delta})+d_i\mathcal{F}(i_{n,\delta})h_{n+1} \tag{11.2.81}$$

*satisfies*

$$\|\varrho_{n+1}\|_{\mathrm{Lip},s_1}\le\varepsilon^2\nu_n^{\frac{3}{2}-9\sigma}\,. \tag{11.2.82}$$

Proof. Differentiating the identity (see (11.2.25))

$$DG_{n,\delta}(\mathtt{u}(\varphi))\mathtt{F}_n(\mathtt{u}(\varphi))=\mathcal{F}(G_{n,\delta}(\mathtt{u}(\varphi))) \tag{11.2.83}$$

we obtain

$$\begin{aligned}&D^2G_{n,\delta}(\mathtt{u}(\varphi))[\mathtt{h},\mathtt{F}_n(\mathtt{u}(\varphi))]+DG_{n,\delta}(\mathtt{u}(\varphi))d_{\mathtt{u}}\mathtt{F}_n(\mathtt{u}(\varphi))[\mathtt{h}]\\&=d_i\mathcal{F}(G_{n,\delta}(\mathtt{u}(\varphi)))DG_{n,\delta}(\mathtt{u}(\varphi))[\mathtt{h}]\,.\end{aligned}$$

For $\mathtt{u}(\varphi)=(\varphi,0,0)$ and $\mathtt{h}=\mathtt{h}_{n+1}$ this gives, recalling (11.2.23), (11.2.78),

$$\begin{aligned}&D^2G_{n,\delta}(\varphi,0,0)[\mathtt{h}_{n+1},\mathtt{F}_n(\varphi,0,0)]+DG_{n,\delta}(\varphi,0,0)d_{\mathtt{u}}\mathtt{F}_n(\varphi,0,0)[\mathtt{h}_{n+1}]\\&=d_i\mathcal{F}(i_{n,\delta}(\varphi))h_{n+1}\,.\end{aligned}\tag{11.2.84}$$

By (11.2.83) we have $\mathcal{F}(i_{n,\delta})=DG_{n,\delta}(\varphi,0,0)\mathtt{F}_n(\varphi,0,0)$ and therefore, by (11.2.84),

$$\begin{aligned}&\mathcal{F}(i_{n,\delta})+d_i\mathcal{F}(i_{n,\delta})h_{n+1}\\&=DG_{n,\delta}(\varphi,0,0)\big(\mathtt{F}_n(\varphi,0,0)+d_{\mathtt{u}}\mathtt{F}_n(\varphi,0,0)[\mathtt{h}_{n+1}]\big)+D^2G_{n,\delta}(\varphi,0,0)[\mathtt{h}_{n+1},\mathtt{F}_n(\varphi,0,0)]\\&\overset{(11.2.72),(11.2.30)}{=}DG_{n,\delta}(\varphi,0,0)\mathtt{r}'_{n+1}+D^2G_{n,\delta}(\varphi,0,0)[\mathtt{h}_{n+1},\mathtt{F}_n(\varphi,0,0)]\,.\end{aligned}$$

In conclusion the term $\varrho_{n+1}$ in (11.2.81) satisfies

$$\begin{aligned}\|\varrho_{n+1}\|_{\mathrm{Lip},s_1}&\le\|DG_{n,\delta}(\varphi,0,0)\mathtt{r}'_{n+1}\|_{\mathrm{Lip},s_1}+\|D^2G_{n,\delta}(\varphi,0,0)[\mathtt{h}_{n+1},\mathtt{F}_n(\varphi,0,0)]\|_{\mathrm{Lip},s_1}\\&\overset{(6.1.11),(6.1.12)}{\lesssim_{s_1}}\big(1+\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_1+2}\big)\|\mathtt{r}'_{n+1}\|_{\mathrm{Lip},s_1}\\&\qquad+\big(1+\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_1+3}\big)\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1}\|\mathtt{F}_n(\varphi,0,0)\|_{\mathrm{Lip},s_1}\\&\overset{(11.2.11)}{\lesssim_{s_1}}\|\mathtt{r}'_{n+1}\|_{\mathrm{Lip},s_1}+\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s_1}\|\mathtt{F}_n(\varphi,0,0)\|_{\mathrm{Lip},s_1}\end{aligned}\tag{11.2.85}$$

where we use that $s_1+3\le s_2+2$. Hence by (11.2.85), (11.2.73), (11.2.36), (11.2.28), we get

$$\|\varrho_{n+1}\|_{\mathrm{Lip},s_1}\lesssim_{s_1}\varepsilon^2\nu_n^{\frac32-8\sigma}+(\nu_n+\varepsilon^2\nu_n^{\frac45-5\sigma})\varepsilon^2\nu_n^{1-2\sigma}$$

which gives (11.2.82). ■

**Lemma 11.2.7.** $\|\mathcal{F}(i_{n+1})\|_{\mathrm{Lip},s_1}\le\varepsilon^2\nu_n^{\frac32-10\sigma}$.

Proof. By (11.2.78) we have

$$\mathcal{F}(i_{n+1})=\mathcal{F}(i_{n,\delta}+h_{n+1})=\mathcal{F}(i_{n,\delta})+d_i\mathcal{F}(i_{n,\delta})h_{n+1}+Q(i_{n,\delta},h_{n+1})\tag{11.2.86}$$

and, by the form of $\mathcal{F}$ in (5.1.2),

$$\begin{aligned}Q(i_{n,\delta},h_{n+1})&=\widetilde{\mathcal{F}}(i_{n,\delta}+h_{n+1})-\widetilde{\mathcal{F}}(i_{n,\delta})-d_i\widetilde{\mathcal{F}}(i_{n,\delta})h_{n+1}\\&=\int_0^1(1-\tau)(D_i^2\widetilde{\mathcal{F}})(i_{n,\delta}+\tau h_{n+1})[h_{n+1},h_{n+1}]\,d\tau\end{aligned}$$

where $\widetilde{\mathcal{F}}$ is the "nonlinear part" of $\mathcal{F}$ defined as

$$\widetilde{\mathcal{F}}(i) := \begin{pmatrix} -\varepsilon^2(\partial_y R)(i(\varphi), \xi) \\ \varepsilon^2(\partial_\theta R)(i(\varphi), \xi) \\ -\varepsilon^2\big(0, (\nabla_Q R)(i(\varphi), \xi)\big) \end{pmatrix}$$

and $R$ is the Hamiltonian defined in (2.2.8) and (2.2.9). We have

$$\begin{aligned} \|Q(i_{n,\delta}, h_{n+1})\|_{\mathrm{Lip},s_1} &\lesssim_{s_1} \varepsilon^2\big(1 + \|\mathfrak{I}_{n,\delta}\|_{\mathrm{Lip},s_1} + \|h_{n+1}\|_{\mathrm{Lip},s_1}\big)\|h_{n+1}\|^2_{\mathrm{Lip},s_1} \\ &\overset{(11.2.11),(11.2.14)}{\lesssim_{s_1}} \varepsilon^2 \|h_{n+1}\|^2_{\mathrm{Lip},s_1} \\ &\overset{(11.2.79)}{\lesssim_{s_1}} \varepsilon^2 \nu_n^{\frac{8}{5}-10\sigma} . \end{aligned} \tag{11.2.87}$$

The lemma follows by (11.2.86), Lemma 11.2.6 and (11.2.87). ■

**Step 6. End of the Induction.** We are now ready to prove that the approximate solution $i_{n+1}$ defined in (11.2.78) for all $\lambda$ in the set $\mathbf{\Lambda}_{n+1} \subset \mathbf{\Lambda}_n$ introduced in (11.2.35) satisfies the estimates (11.2.5)-(11.2.7) and (11.2.8) at step $n+1$. By Lemma 11.2.7 and (11.2.3) we obtain

$$\|\mathcal{F}(i_{n+1})\|_{\mathrm{Lip},s_1} \le \varepsilon^2 \nu_n^{\frac{3}{2}-10\sigma} = \varepsilon^2 \nu_{n+1}$$

with

$$\sigma_\star := 10\sigma , \tag{11.2.88}$$

proving $(11.2.8)_{n+1}$. Moreover we have

$$\begin{aligned} \|i_{n+1} - i_n\|_{\mathrm{Lip},s_1+2} &\le \|i_{n+1} - i_{n,\delta}\|_{\mathrm{Lip},s_1+2} + \|i_{n,\delta} - \breve{\imath}_n\|_{\mathrm{Lip},s_1+2} + \|\breve{\imath}_n - i_n\|_{\mathrm{Lip},s_1+2} \\ &\overset{(11.2.78),(11.2.79),(11.2.14),(11.2.12)}{\lesssim_{s_1}} \nu_n + \varepsilon^2 \nu_n^{\frac{4}{5}-5\sigma} + \varepsilon^2 \nu_n^{1-\sigma} + \varepsilon^2 \nu_n^2 \\ &\le C(s_1)\nu_n + \varepsilon^2 \nu_n^{\frac{4}{5}-6\sigma} \end{aligned} \tag{11.2.89}$$

proving $(11.2.6)_{n+1}$ since $\sigma_\star := 10\sigma$. Similarly we get

$$\begin{aligned} \|i_{n+1} - \breve{\imath}_n\|_{\mathrm{Lip},s_3+2} &\le \|i_{n+1} - i_{n,\delta}\|_{\mathrm{Lip},s_3+2} + \|i_{n,\delta} - \breve{\imath}_n\|_{\mathrm{Lip},s_3+2} \\ &\overset{(11.2.78),(11.2.80),(11.2.15)}{\le} \varepsilon^2 \nu_n^{-\frac{11}{10}-4\sigma} \qquad (11.2.90) \\ &\overset{(11.2.3),(11.2.88)}{\ll} \varepsilon^2 \nu_{n+1}^{-\frac{4}{5}} \qquad (11.2.91) \end{aligned}$$

and, by (11.2.91) and (11.2.12),

$$\|i_{n+1} - i_n\|_{\mathrm{Lip},s_3+2} \le \|i_{n+1} - \breve{\imath}_n\|_{\mathrm{Lip},s_3+2} + \|\breve{\imath}_n - i_n\|_{\mathrm{Lip},s_3+2} \ll \varepsilon^2 \nu_{n+1}^{-\frac{4}{5}} \tag{11.2.92}$$

provided $\sigma$ is chosen small enough.

The bound $(11.2.7)_{n+1}$ for $\|i_{n+1}-i_n\|_{\mathrm{Lip},s_2}$ follows by interpolation: setting

$$s_2+2=\theta s_1+(1-\theta)s_3\,,\quad \theta=\frac{s_3-s_2-2}{s_3-s_1}\,,\quad 1-\theta=\frac{s_2+2-s_1}{s_3-s_1}\,,$$

we have, using $\nu_n=O(\varepsilon^2)$,

$$\begin{aligned}\|i_{n+1}-i_n\|_{\mathrm{Lip},s_2+2} &\overset{(3.5.10)}{\lesssim_{s_1,s_3}} \|i_{n+1}-i_n\|^{\theta}_{\mathrm{Lip},s_1}\|i_{n+1}-i_n\|^{1-\theta}_{\mathrm{Lip},s_3}\\ &\overset{(11.2.89),(11.2.92)}{\lesssim_{s_1,s_3}} \big(\max\{\nu_n,\varepsilon^2\nu_n^{\frac45-6\sigma}\}\big)^{\theta}\big(\varepsilon^2\nu_n^{-\frac{11}{10}-4\sigma}\big)^{1-\theta}\\ &\lesssim_{s_1,s_3} \big(\varepsilon^{\frac32}\nu_n^{\frac14}\big)^{\theta}\big(\varepsilon^{\frac32}\nu_n^{-\frac{11}{10}-3\sigma}\big)^{1-\theta}\leq\varepsilon^{\frac32}\nu_n^{\frac14-\sigma}\end{aligned}\tag{11.2.93}$$

for $s_3$ large enough. The bounds for $\mathfrak{I}_{n+1}$ follow by a telescoping argument. Using the first estimate in $(11.2.5)_1$ and $(11.2.6)_k$ for all $1\leq k\leq n+1$ we get

$$\begin{aligned}\|\mathfrak{I}_{n+1}\|_{\mathrm{Lip},s_1+2}\leq\|\mathfrak{I}_1\|_{\mathrm{Lip},s_1+2}+\sum_{k=1}^{n+1}\|\mathfrak{I}_{k+1}-\mathfrak{I}_k\|_{\mathrm{Lip},s_1+2}&\lesssim_{s_1}\varepsilon^2+\varepsilon^2\nu_1^{\frac45-\sigma_\star}\\ &\leq C(s_1)\varepsilon^2\end{aligned}$$

since $\nu_1=C(s_1)\varepsilon^2$. This proves the first inequality in $(11.2.5)_{n+1}$. Moreover

$$\begin{aligned}\|\mathfrak{I}_{n+1}\|_{\mathrm{Lip},s_3+2} &\leq\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s_3+2}+\|i_{n+1}-\breve{i}_n\|_{\mathrm{Lip},s_3+2}\\ &\overset{(11.2.11)(11.2.90)}{\leq}\varepsilon^2\nu_n^{-\frac45-\sigma}+\varepsilon^2\nu_n^{-\frac{11}{10}-4\sigma}\\ &\leq 2\varepsilon^2\nu_n^{-\frac{11}{10}-4\sigma}\overset{(11.2.3)}{\leq}\varepsilon^2\nu_{n+1}^{-\frac45}\end{aligned}\tag{11.2.94}$$

which is the third estimate in $(11.2.5)_{n+1}$. The second estimate in $(11.2.5)_{n+1}$ follows similarly by (5.3.4) and $(11.2.7)_k$, for all $k\leq n+1$.

In order to complete the proof of Theorem 11.2.1 it remains to prove the measure estimates (11.2.4).

**Lemma 11.2.8.** *The sets* $\mathbf{\Lambda}_n$ *defined iteratively in* (11.2.35) *with* $\mathbf{\Lambda}_1:=\Lambda$*, satisfy* (11.2.4).

Proof. The estimate $|\mathbf{\Lambda}_1\setminus\mathbf{\Lambda}_2|\leq b(\varepsilon)$ with $\lim_{\varepsilon\to0}b(\varepsilon)=0$ follows by item 2 of Proposition 11.1.1 with $\Lambda_{\underline{\mathfrak{I}}}=\Lambda_{\breve{\mathfrak{I}}_1}=\Lambda$. In order to prove (11.2.4) for $n\geq3$ notice that, by the definition of $\mathbf{\Lambda}_{n+1}$ (and $\mathbf{\Lambda}_n$) in (11.2.35), we have

$$\mathbf{\Lambda}_n\setminus\mathbf{\Lambda}_{n+1}=\mathbf{\Lambda}_{n-1}\bigcap\mathbf{\Lambda}(\varepsilon;\eta_{n-1},\breve{\mathfrak{I}}_{n-1})\bigcap[\mathbf{\Lambda}(\varepsilon;\eta_n,\breve{\mathfrak{I}}_n)]^c\,,\quad n\geq2\,.\tag{11.2.95}$$

In addition, for all $\lambda \in \mathbf{\Lambda}_n$, we have, for $\beta := \frac{3}{4}$,

$$\begin{aligned}
\|\breve{\mathfrak{I}}_n - \breve{\mathfrak{I}}_{n-1}\|_{s_1+2} &\stackrel{(11.2.9)}{=} \|\Pi_{N_n}\mathfrak{I}_n - \Pi_{N_n}\breve{\mathfrak{I}}_{n-1}\|_{s_1+2} \\
&\leq \|\Pi_{N_n}(\mathfrak{I}_n - \mathfrak{I}_{n-1})\|_{s_1+2} + \|\Pi_{N_n}(\mathfrak{I}_{n-1} - \breve{\mathfrak{I}}_{n-1})\|_{s_1+2} \\
&\stackrel{(4.1.12)}{\leq} N_n^2 \|i_n - i_{n-1}\|_{s_1} + \|i_{n-1} - \breve{\imath}_{n-1}\|_{s_1+2} \\
&\stackrel{(11.2.6),(11.2.12)}{\leq} \nu_n^{-\frac{6}{s_3-s_1}}\big(C(s_1)\nu_{n-1} + \varepsilon^2 \nu_{n-1}^{\frac{4}{5}-\sigma_\star}\big) + \varepsilon^2 \nu_{n-1}^2 \\
&\ll \nu_{n-1}^{\beta} \leq \varepsilon^{\frac{3}{2}}, \quad \forall n \geq 2, \qquad (11.2.96)
\end{aligned}$$

since $\nu_{n-1} \leq C(s_1)\varepsilon^2$. Now by (11.2.35), we have $\eta_n - \eta_{n-1} = \nu_{n-1}^{\frac{1}{4}} \geq \nu_{n-1}^{\frac{2}{5}\beta}$. Hence the estimates (11.1.1) and (11.2.96) imply that the set of (11.2.95) satisfies the measure estimate

$$|\mathbf{\Lambda}_n \setminus \mathbf{\Lambda}_{n+1}| \leq \nu_{n-1}^{\beta\alpha/3} \leq \nu_{n-1}^{\alpha/4},$$

which proves (11.2.4). ∎

**Proof of Theorem 5.1.2.** The torus embedding

$$i_\infty := i_0 + (i_1 - i_0) + \sum_{n\geq 2}(i_n - i_{n-1})$$

is defined for all $\lambda$ in $\mathcal{C}_\infty := \cap_{n\geq 1}\mathbf{\Lambda}_n$, and, by (5.3.4) and (11.2.6)-(11.2.7), it is convergent in $\| \, \|_{\mathrm{Lip},s_1}$ and $\| \, \|_{\mathrm{Lip},s_2}$-norms with

$$\|i_\infty - i_0\|_{\mathrm{Lip},s_1} \leq C(s_1)\varepsilon^2, \quad \|i_\infty - i_0\|_{\mathrm{Lip},s_2} \leq \varepsilon,$$

proving (5.1.9). By (11.2.8) we deduce that

$$\forall \lambda \in \mathcal{C}_\infty = \cap_{n\geq 1}\mathbf{\Lambda}_n, \quad \mathcal{F}(\lambda; i_\infty(\lambda)) = 0.$$

Finally, by (11.2.4) we deduce (5.1.8), i.e. that $\mathcal{C}_\infty$ is a set of asymptotically full measure.

**Remark 11.2.9.** *The previous result holds if the nonlinearity $g(x,u)$ and the potential $V(x)$ in* (1.1.1) *are of class $C^q$ for some $q$ large enough, depending on $s_3$.*

## 11.3 $C^\infty$ solutions

In this section we prove the last statement of Theorem 5.1.2 about $C^\infty$ solutions.

By Proposition 8.2.2 and a simple modification of the proof of Proposition 11.1.1 which substitutes any $s \geq s_3$ to $s_3$ (see Remark 11.1.2), we obtain the following result.

**Proposition 11.3.1.** *Same assumptions as in Proposition 11.1.1. From* (11.1.1)*, the conclusion can be modified in the following way: for any* $s \geq s_3$*, there is* $\widetilde{\nu}(s)$ *such that the following holds.*

*For any* $\nu \in (0, \varepsilon^{\frac{3}{2}}) \cap (0, \widetilde{\nu}(s))$ *such that* $\|\underline{\mathfrak{I}}\|_{\mathrm{Lip},s+4} \leq \varepsilon \nu^{-\frac{9}{10}}$*, there exists a linear operator* $\mathcal{L}_{appr}^{-1} := \mathcal{L}_{appr,\nu,s}^{-1}$ *such that, for any function* $g : \Lambda_{\underline{\mathfrak{I}}} \to \mathcal{H}^s \cap H_{\mathbb{S}}^{\perp}$ *satisfying*

$$\|g\|_{\mathrm{Lip},s_1} \leq \varepsilon^2 \nu \,, \quad \|g\|_{\mathrm{Lip},s} \leq \varepsilon^2 \nu^{-\frac{9}{10}} \,, \tag{11.3.1}$$

*the function* $h := \mathcal{L}_{appr}^{-1} g$*,* $h : \mathbf{\Lambda}(\varepsilon; 5/6, \underline{\mathfrak{I}}) \to \mathcal{H}^s \cap H_{\mathbb{S}}^{\perp}$ *satisfies*

$$\|h\|_{\mathrm{Lip},s_1} \leq C(s_1) \varepsilon^2 \nu^{\frac{4}{5}} \,, \quad \|h\|_{\mathrm{Lip},s+2} \leq C(s) \varepsilon^2 \nu^{-\frac{11}{10}} \,, \tag{11.3.2}$$

*and, setting* $\mathcal{L}_\omega := \mathcal{L}_\omega(\underline{i})$*, we have*

$$\|\mathcal{L}_\omega h - g\|_{\mathrm{Lip},s_1} \leq C(s_1) \varepsilon^2 \nu^{\frac{3}{2}} \,. \tag{11.3.3}$$

*Furthermore, setting* $Q' := 2(\tau' + \varsigma s_1) + 3$ *(where* $\varsigma = 1/10$ *and* $\tau'$ *are given by Proposition 4.1.5), for all* $g \in \mathcal{H}^{s_0 + Q'} \cap H_{\mathbb{S}}^{\perp}$*,*

$$\|\mathcal{L}_{appr}^{-1} g\|_{\mathrm{Lip},s_0} \lesssim_{s_1} \|g\|_{\mathrm{Lip},s_0+Q'} \,. \tag{11.3.4}$$

Note that in Proposition 11.3.1 the sets $\mathbf{\Lambda}(\varepsilon; 5/6, \underline{\mathfrak{I}})$ do not depend on $s$ and are the same as in Proposition 11.1.1.

Thanks to Proposition 11.3.1 we can modify the Nash-Moser scheme in order to keep along the iteration the control of higher and higher Sobolev norms. This new scheme relies on the following result, which is used in the iteration. We recall that the sequence $(\nu_n)$ is defined in (11.2.3), and the sequence $(\eta_n)$ in (11.2.35).

**Proposition 11.3.2.** *For any* $s \geq s_3$*, there is* $\widetilde{\nu}'(s) > 0$ *with the following property. Assume that for some* $n$ *such that* $\nu_n \leq \widetilde{\nu}'(s)$*, there is a map* $\mathfrak{I}_n : \lambda \mapsto \mathfrak{I}_n(\lambda)$*, defined for* $\lambda$ *in some set* $\mathbf{\Lambda}_n$*, such that,*

$$\|\mathfrak{I}_n\|_{\mathrm{Lip},s_1+2} \leq C(s_1) \varepsilon^2 \,, \quad \|\mathfrak{I}_n\|_{\mathrm{Lip},s_2+2} \leq \varepsilon \,, \quad \|\mathfrak{I}_n\|_{\mathrm{Lip},s+2} \leq \varepsilon^2 \nu_n^{-\frac{4}{5}} \,, \tag{11.3.5}$$

*and*

$$\|\mathcal{F}(i_n)\|_{\mathrm{Lip},s_1} \leq \varepsilon^2 \nu_n \quad \text{where} \quad i_n(\varphi) = (\varphi, 0, 0) + \mathfrak{I}_n(\varphi) \,.$$

*Let*

$$N_{n,s} \in \Big[\nu_n^{-\frac{3}{s-s_1}} - 1, \nu_n^{-\frac{3}{s-s_1}} + 1\Big] \quad \text{and} \quad \breve{\mathfrak{I}}_{n,s} := \Pi_{N_{n,s}} \mathfrak{I}_n \,. \tag{11.3.6}$$

*Then there exists a map* $\lambda \mapsto \mathfrak{I}_{n+1}(\lambda)$*, defined for* $\lambda \in \mathbf{\Lambda}_n \cap \mathbf{\Lambda}(\varepsilon; \eta_n, \breve{\mathfrak{I}}_{n,s})$*, which satisfies*

$$\|i_{n+1} - i_n\|_{\mathrm{Lip},s_1+2} \leq C(s_1) \nu_n + \varepsilon^2 \nu_n^{\frac{4}{5} - \sigma_\star}, \quad \|i_{n+1} - i_n\|_{\mathrm{Lip},s_2} \leq \varepsilon^2 \nu_n^{\frac{1}{5}} \,, \tag{11.3.7}$$

$$\|\mathfrak{I}_{n+1}\|_{\mathrm{Lip},s+3} \leq \varepsilon^2 \nu_{n+1}^{-\frac{4}{5}} \tag{11.3.8}$$

*and*

$$\|\mathcal{F}(i_{n+1})\|_{\mathrm{Lip},s_1} \le \varepsilon^2 \nu_{n+1}\,,$$

*where $\sigma_\star$ is chosen as in Theorem 11.2.1.*

Proof. The proof follows exactly the steps of the inductive part in the proof of Theorem 11.2.1, with $N_{n,s} \sim \nu_n^{-\frac{3}{s-s_1}}$ and with the same small exponent $\sigma$. We have just to observe that in Step 1 (regularization)

$$\begin{aligned}\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s+\underline{\tau}+7} &\overset{(4.1.13)}{\le} N_{n,s}^{\underline{\tau}+5}\|\mathfrak{I}_n\|_{\mathrm{Lip},s+2}\\ &\overset{(11.3.6),(11.3.5)}{\le} C(s)\nu_n^{-\frac{3(\underline{\tau}+5)}{s-s_1}}\varepsilon^2\nu_n^{-\frac{4}{5}} \le \varepsilon^2\nu_n^{-\frac{4}{5}-\sigma}\,,\end{aligned}$$

for $3(\underline{\tau}+5)/(s_3-s_1) < \sigma$ and $\nu_n$ small enough (depending on $s$). As a result

$$\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s+1+\underline{\tau}+6} \le \varepsilon^2\nu_n^{-\frac{4}{5}-\sigma}\,,$$

i.e. in the third estimate of (11.2.11) in Lemma 11.2.2, $s+1$ can be substituted to $s_3$. From this point, we use the substitution $s_3 \rightsquigarrow s+1$ in all the estimates of the induction where $s_3$ appears, except the second estimate of (11.2.12) and (11.2.92), where we keep $s_3$. Note that these last two estimates are useful only to obtain the bound (11.2.93) on the norm $\|i_{n+1}-i_n\|_{\mathrm{Lip},s_2+2}$ by an interpolation argument. All the estimates where we apply the substitution $s_3 \rightsquigarrow s+1$ hold provided that $\nu_n$ is smaller than some possibly very small, but positive constant, depending on $s$. In particular, the analogous of Proposition 11.2.4 uses Proposition 11.3.1, and the second estimate of (11.2.36) is replaced by

$$\|\mathtt{h}_{n+1}\|_{\mathrm{Lip},s+3} \le \varepsilon^2\nu_n^{-\frac{11}{10}-2\sigma}\,.$$

To prove (11.3.8), we use (11.2.94) with the substitution $s_3 \rightsquigarrow s+1$. Note that in this estimate we need only a bound for $\|\breve{\mathfrak{I}}_n\|_{\mathrm{Lip},s+3}$, not for $\|\mathfrak{I}_n\|_{\mathrm{Lip},s+3}$. ■

We now consider a non-decreasing sequence $(p_n)$ of integers with the following properties:

(i) $p_1 = 0$ and $\lim_{n\to\infty} p_n = \infty$ ;

(ii) $\forall n \ge 1\,,\ \nu_n \le \widetilde{\nu}'(s_3+p_n)$ ;

(iii) $\forall n \ge 1\,,\ p_{n+1} = p_n$ or $p_{n+1} = p_n + 1$ .

The sequence $(p_n)$ can be defined iteratively in the following way: $p_1 := 0$, so that property (ii) is satisfied for $n = 1$, provided $\varepsilon$ is small enough (the smallness condition depending on $s_3$). Once $p_n$ is defined (satisfying (ii)), we choose

$$p_{n+1} := \begin{cases} p_n + 1 & \text{if } \nu_{n+1} \le \widetilde{\nu}'(s_3+p_n+1) \\ p_n & \text{otherwise}\,. \end{cases}$$

Then property (ii) is satisfied at step $n+1$ in both cases, because $(\nu_n)$ is decreasing. The sequence $(p_n)$ satisfies (i) because $\lim_{n\to\infty} \nu_n = 0$ (so that the sequence cannot be stationary) and (iii) by definition.

Starting, as in section 11.2, with the torus $i_1(\varphi)$ defined in Lemma 5.3.1 for all $\lambda \in \Lambda = \mathbf{\Lambda}_1$ and using repeatedly Proposition 11.3.2, we obtain the following theorem.

**Theorem 11.3.3. (Nash-Moser $C^\infty$)** *Let $\bar{\omega}_\varepsilon \in \mathbb{R}^{|\mathbb{S}|}$ be $(\gamma_1, \tau_1)$-Diophantine and satisfy property $(\mathbf{NR})_{\gamma_1,\tau_1}$ in Definition 4.1.4 with $\gamma_1, \tau_1$ fixed in (1.2.28). Assume (8.2.1) and $s_2 - s_1 \geq \underline{\tau} + 2$, $s_1 \geq s_0 + 2 + \underline{\tau} + Q'$, where $\underline{\tau}$ is the loss of derivatives defined in Proposition 6.1.1 and $Q' := 2(\tau' + \varsigma s_1) + 3$ is defined in Proposition 10.2.1.*

*Then, for all $0 < \varepsilon \leq \varepsilon_0$ small enough, for all $n \geq 1$, there exist*

1. *a subset $\mathbf{\Lambda}_n \subseteq \mathbf{\Lambda}_{n-1}$, $\mathbf{\Lambda}_1 := \mathbf{\Lambda}_0 := \Lambda$, satisfying*

$$\begin{aligned} &|\mathbf{\Lambda}_1 \setminus \mathbf{\Lambda}_2| \leq b(\varepsilon) \quad \text{with} \quad \lim_{\varepsilon\to 0} b(\varepsilon) = 0\,, \\ &|\mathbf{\Lambda}_{n-1} \setminus \mathbf{\Lambda}_n| \leq \nu_{n-2}^{\alpha_*}\,, \ \forall n \geq 3\,, \end{aligned} \tag{11.3.9}$$

   *where $\alpha_* = \alpha/4$ and $\alpha > 0$ is the exponent in (11.1.1),*

2. *a torus $i_n(\varphi) = (\varphi, 0, 0) + \mathfrak{I}_n(\varphi)$, defined for all $\lambda \in \mathbf{\Lambda}_n$, satisfying*

$$\|\mathfrak{I}_n\|_{\mathrm{Lip},s_1} \leq C(s_1)\varepsilon^2\,, \quad \|\mathfrak{I}_n\|_{\mathrm{Lip},s_2+2} \leq \varepsilon\,, \quad \|\mathfrak{I}_n\|_{\mathrm{Lip},s_3+p_n+2} \leq \varepsilon^2 \nu_n^{-\frac{4}{5}}\,, \tag{11.3.10}$$

$$\|i_n - i_{n-1}\|_{\mathrm{Lip},s_1+2} \leq C(s_1)\nu_{n-1} + \varepsilon^2 \nu_{n-1}^{\frac{4}{5}-\sigma_\star}\,, \ n \geq 2\,, \tag{11.3.11}$$

$$\|i_n - i_{n-1}\|_{\mathrm{Lip},s_2} \leq \varepsilon \nu_{n-1}^{\frac{1}{5}}\,, \quad n \geq 2\,, \tag{11.3.12}$$

   *and*

$$\|\mathcal{F}(i_n)\|_{\mathrm{Lip},s_1} \leq \varepsilon^2 \nu_n\,. \tag{11.3.13}$$

As said at the end of section 11.2, (11.3.11) and (11.3.12) imply that the sequence $(i_n)$ converges in $\| \ \|_{\mathrm{Lip},s_1}$ and $\| \ \|_{\mathrm{Lip},s_2}$ norms to a map $i_\infty : \lambda \mapsto i_\infty(\lambda)$, defined on

$$\mathcal{C}_\infty = \cap_{n\geq 1} \mathbf{\Lambda}_n\,,$$

which satisfies

$$\forall \lambda \in \mathcal{C}_\infty\,, \quad \mathcal{F}(\lambda; i_\infty(\lambda)) = 0\,.$$

There remains to justify that

$$\|\mathfrak{I}_\infty\|_{\mathrm{Lip},s} = \|i_\infty - i_0\|_{\mathrm{Lip},s} < \infty\,, \quad \forall s > s_2\,.$$

For a given $s > s_2$, let us fix $\overline{n}$ large enough, so that

$$s \leq \frac{3}{4}s_1 + \frac{1}{4}\overline{s} \qquad \text{with} \ \ \overline{s} := s_3 + p_{\overline{n}} + 2\,. \tag{11.3.14}$$

Then, by (11.3.10), for all $k \geq \overline{n}+1$,

$$\|\mathfrak{I}_k - \mathfrak{I}_{k-1}\|_{\mathrm{Lip},\overline{s}} \leq \|\mathfrak{I}_k\|_{\mathrm{Lip},\overline{s}} + \|\mathfrak{I}_{k-1}\|_{\mathrm{Lip},\overline{s}} \leq 2\varepsilon^2 \nu_k^{-\frac{4}{5}} . \tag{11.3.15}$$

Hence, by interpolation,

$$\begin{aligned} \|\mathfrak{I}_k - \mathfrak{I}_{k-1}\|_{\mathrm{Lip},s} &\overset{(11.3.14)}{\leq} \|\mathfrak{I}_k - \mathfrak{I}_{k-1}\|_{\mathrm{Lip},\frac{3s_1+\overline{s}}{4}} \\ &\overset{(3.5.10)}{\leq} C(\overline{s}) \|\mathfrak{I}_k - \mathfrak{I}_{k-1}\|_{\mathrm{Lip},s_1}^{\frac{3}{4}} \|\mathfrak{I}_k - \mathfrak{I}_{k-1}\|_{\mathrm{Lip},\overline{s}}^{\frac{1}{4}} \\ &\overset{(11.3.11),(11.3.15)}{\leq} C(\overline{s}) \nu_{k-1}^{\frac{3}{5}-\sigma_\star} \nu_k^{-\frac{1}{5}} \leq \nu_{k-1}^{\frac{3}{10}-\sigma_\star} \end{aligned} \tag{11.3.16}$$

using that $\nu_k = \nu_{k-1}^{\mathfrak{q}}$, with $\mathfrak{q} < 3/2$, see (11.2.3). Moreover, recalling that $\overline{s} := s_3 + p_{\overline{n}} + 2$ (see (11.3.14)), we deduce by (11.3.10) that

$$\|\mathfrak{I}_{\overline{n}}\|_{\mathrm{Lip},\overline{s}} < \infty . \tag{11.3.17}$$

In conclusion, (11.3.17) and (11.3.16), imply, since $(3/10) - \sigma_\star > 0$, that

$$\|\mathfrak{I}_{\overline{n}}\|_{\mathrm{Lip},s} + \sum_{k \geq \overline{n}+1} \|\mathfrak{I}_k - \mathfrak{I}_{k-1}\|_{\mathrm{Lip},s} < \infty ,$$

and the sequence $(\mathfrak{I}_n)_{n \geq \overline{n}}$ converges in $H^s_\varphi \times H^s_\varphi \times (\mathcal{H}^s \cap H^\perp_{\mathbb{S}})$ to $\mathfrak{I}_\infty = i_\infty - i_0$. Since $s$ is arbitrary, we conclude that $i_\infty(\lambda)$ is $C^\infty$ for any $\lambda \in \mathcal{C}_\infty$.

# Chapter 12

# Genericity of the assumptions

The aim of this Chapter is to prove the genericity result stated in Theorem 1.2.3.

## 12.1 Genericity of non-resonance and non-degeneracy conditions

We fix $s > d/2$, so that we have the compact embedding $H^s(\mathbb{T}^d) \hookrightarrow C^0(\mathbb{T}^d)$. We denote by $B(w, r)$ the open ball of center $w$ and radius $r$ in $H^s(\mathbb{T}^d)$.

Recalling Definition 1.2.2 of $C^\infty$-dense open sets, it is straightforward to check the following lemma.

**Lemma 12.1.1.** *The following properties hold:*

1. *a finite intersection of $C^\infty$-dense open subsets of $\mathcal{U}$ is $C^\infty$-dense open in $\mathcal{U}$;*
2. *a countable intersection of $C^\infty$-dense open subsets of $\mathcal{U}$ is $C^\infty$-dense in $\mathcal{U}$;*
3. *given three open subsets $\mathcal{W} \subset \mathcal{V} \subset \mathcal{U}$ of $H^s(\mathbb{T}^d)$ (resp. $H^s(\mathbb{T}^d) \times H^s(\mathbb{T}^d)$), if $\mathcal{W}$ is $C^\infty$-dense in $\mathcal{V}$ and $\mathcal{V}$ is $C^\infty$-dense in $\mathcal{U}$, then $\mathcal{W}$ is $C^\infty$-dense in $\mathcal{U}$.*

Moreover we have the following useful result.

**Lemma 12.1.2.** *Let $U$ be a connected open subset of $H^s(\mathbb{T}^d)$ and let $f : U \to \mathbb{R}$ be a real analytic function. If $f \not\equiv 0$, then*

$$Z(f)^c := \{w \in U \ : \ f(w) \neq 0\}$$

*is a $C^\infty$-dense open subset of $U$.*

Proof. Since $f : U \to \mathbb{R}$ is continuous, $Z(f)^c$ is an open subset of $U$. Arguing by contradiction, we assume that $Z(f)^c$ is not $C^\infty$-dense in $U$. Then there are $w_0 \in U$, $s' \geq s$ and $\epsilon > 0$ such that: for all $h \in C^\infty(\mathbb{T}^d)$ satisfying $\|h\|_{H^{s'}} < \epsilon$, we have that $f(w_0 + h) = 0$. Let $\rho > 0$ be such that the ball $B(w_0, \rho) \subset U$. We claim that $f$ vanishes on $B(w_0, \rho) \cap C^\infty(\mathbb{T}^d)$. Indeed, if $h \in C^\infty(\mathbb{T}^d)$ satisfies $\|h\|_{H^s} < \rho$ we have the segment $[w_0, w_0 + h] \subset U$, and the map $\varphi : t \mapsto f(w_0 + th)$ is real analytic on an open interval of $\mathbb{R}$ which contains $[0, 1]$. Moreover $\varphi(t)$ vanishes on the whole interval $|t| < \epsilon\|h\|^{-1}_{H^{s'}}$. Hence $\varphi$ vanishes everywhere, and $f(w_0 + h) = \varphi(1) = 0$.

Now, by the fact that $C^\infty(\mathbb{T}^d)$ is a dense subset of $H^s(\mathbb{T}^d)$, and $f$ is continuous, we conclude the $f$ vanishes on the whole ball $B(w_0, \rho)$.

Let

$$V := \big\{w \in U \ : \ f \text{ vanishes on some open neighborhood of } w\big\}\,.$$

From the previous argument, $V$ is not empty, and it is by definition an open subset of $U$. Let us prove that it is closed in $U$ too. Assume that the sequence $(w_n)$ $H^s$-converges to some $w$ and satisfies: $w_n \in V$ for all $n$. Then there is $r > 0$ such that, for $n$ large enough, the ball $B(w_n, r)$ contains $w$ and is included in $U$. Using the same argument as before (with $w_n$ instead of $w_0$), we can conclude that $f$ vanishes on the whole ball $B(w_n, r)$, hence $w \in V$.

Since $U$ is connected, we finally obtain $V = U$, which contradicts the hypothesis $f \not\equiv 0$. ■

The proof of Theorem 1.2.3 uses results and arguments provided by Kappeler-Kuksin [89] that we recall below. We first introduce some preliminary information. For a real valued potential $V \in H^s(\mathbb{T}^d)$, we denote by $(\lambda_j(V))_{j\in\mathbb{N}}$ the sequence of the eigenvalues of the Sturm-Liouville operator $-\Delta + V(x)$, written in increasing order and counted with multiplicity

$$\lambda_0(V) < \lambda_1(V) \leq \lambda_2(V) \leq \ldots\,.$$

These eigenvalues are Lipschitz-continuous functions of the potential, namely

$$|\lambda_j(V_1) - \lambda_j(V_2)| \leq \|V_1 - V_2\|_{L^\infty(\mathbb{T}^d)} \lesssim \|V_1 - V_2\|_{H^s(\mathbb{T}^d)}\,. \tag{12.1.1}$$

**Step 1. The construction of Kappeler-Kuksin [89].** For $J \subset \mathbb{N}$, define the set of potentials $V := V(x)$ in $H^s(\mathbb{T}^d)$ such that the eigenvalues $\lambda_j(V)$, $j \in J$, of $-\Delta + V(x)$ are simple, i.e.

$$E_J := \Big\{V \in H^s(\mathbb{T}^d) \ : \ \lambda_j(V) \text{ is a simple eigenvalue of } -\Delta + V(x)\,,\ \forall j \in J\Big\}\,.$$

The set $E_{[\![0,N]\!]}$ will be simply denoted by $E_N$. Since the eigenvalues $\lambda_j(V)$ of $-\Delta + V(x)$ are simple on $E_J$, it turns out that each function

$$\lambda_j : E_J \to \mathbb{R}\,, \quad j \in J\,,$$

is *real analytic*. Moreover the corresponding eigenfunctions $\Psi_j := \Psi_j(V)$, normalized with $\|\Psi_j\|_{L^2} = 1$, can locally be expressed as real analytic functions of the potential $V \in E_J$.

By [89], Lemma 2.2, we have that, for any $J \subset \mathbb{N}$ finite,

1. $E_J$ is an *open*, *dense* and *connected* subset of $H^s(\mathbb{T}^d)$,

2. $H^s(\mathbb{T}^d)\backslash E_J$ is a real analytic variety. This implies that for all $V \in H^s(\mathbb{T}^d)\backslash E_J$, there are $r > 0$ and real analytic functions $f_1, \dots, f_s$ on the open ball $B(V,r)$ such that

$$(i) \quad B(V,r)\backslash E_J \subset \bigcup_{i=1}^{s} f_i^{-1}(0) \qquad (ii) \quad \forall i \in [\![1,s]\!]\,,\ f_{i|B(V,r)} \not\equiv 0\,. \tag{12.1.2}$$

Recalling Definition 1.2.2 we prove this further lemma.

**Lemma 12.1.3.** *Let $\mathcal{P}$ be the set defined in* (1.2.36)*. The subset $E_J \cap \mathcal{P} \subset H^s(\mathbb{T}^d)$ is connected and $C^\infty$-dense open in $\mathcal{P}$.*

Proof. We first prove that $E_J \cap \mathcal{P}$ is $C^\infty$-dense open in $\mathcal{P}$. By (12.1.2)

$$\bigcap_{i=1}^{s} B(V,r)\backslash f_i^{-1}(0) \subset B(V,r) \cap E_J\,. \tag{12.1.3}$$

By Lemma 12.1.2 each set $B(V,r)\backslash f_i^{-1}(0)$, $i = 1, \dots, s$, is $C^\infty$-dense open in $B(V,r)$ (since $f_{i|B(V,r)} \not\equiv 0$), as well as their intersection, and therefore (12.1.3) implies that $B(V,r) \cap E_J$ is $C^\infty$-dense in $B(V,r)$. Thus $E_J$ is $C^\infty$-dense open in $H^s(\mathbb{T}^d)$. Finally, since $\mathcal{P}$ is an open subset of $H^s(\mathbb{T}^d)$, the set $E_J \cap \mathcal{P}$ is $C^\infty$-dense open in $\mathcal{P}$.

There remains to justify that $E_J \cap \mathcal{P}$ is connected. Let $V_0, V_1 \in E_J \cap \mathcal{P}$. Since $E_J$ is an open connected subset of $H^s(\mathbb{T}^d)$, it is arcwise connected: there is a continuous path $\gamma : [0,1] \to E_J$ such that $\gamma(0) = V_0$, $\gamma(1) = V_1$. Notice that, if $V \in E_J$, then, for all $m \in \mathbb{R}$, the potential $V + m \in E_J$. Let $\lambda_0(t)$ be the smallest eigenvalue of $-\Delta + \gamma(t)$. The map $t \mapsto \lambda_0(t)$ is continuous. Since $V_0, V_1 \in \mathcal{P}$, we have $\lambda_0(0) > 0$, $\lambda_0(1) > 0$. Choose any continuous map $\mu : [0,1] \to ]0,+\infty[$ such that $\mu(0) = \lambda_0(0)$ and $\mu(1) = \lambda_0(1)$ and define $m : [0,1] \to \mathbb{R}$ by $m(t) := \mu(t) - \lambda_0(t)$. Then $\gamma + m$ is a continuous path in $E_J \cap \mathcal{P}$ connecting $V_0$ and $V_1$. In conclusion, $E_J \cap \mathcal{P}$ is arcwise connected. ■

We have the following result, which is Lemma 2.3 in [89], with some new estimates on the eigenfunctions, i.e. items $(iii)$-$(iv)$.

**Lemma 12.1.4.** *Fix $J \subset \mathbb{N}$ , $J$ finite. There is a sequence $(q_n)_{n\in\mathbb{N}}$ of $C^\infty$ positive potentials with the following properties:*

*(i) $\forall n \in \mathbb{N}$, the potential $q_n$ is in $E_J$. More precisely, for each $j \in J$, the sequence of eigenvalues $(\lambda_{n,j})_n := (\lambda_j(q_n))_n$ converges to some $\lambda_j > 0$, with $\lambda_j \neq \lambda_k$ if $j, k \in J$, $j \neq k$.*

*(ii)* *Let* $\pm\Psi_{n,j}$ *be the eigenfunctions of* $-\Delta + q_n(x)$ *with* $\|\Psi_{n,j}\|_{L^2} = 1$*. For each* $j \in J$*, the sequence* $(\Psi_{n,j})_n \to \Psi_j$ *weakly in* $H^1(\mathbb{T}^d)$*, hence strongly in* $L^2(\mathbb{T}^d)$*, where the functions* $\Psi_j \in L^\infty(\mathbb{T}^d)$ *have disjoint essential supports.*

*(iii)* *For each* $j \in J$*, the sequence* $(\Psi_{n,j})_n \to \Psi_j$ *strongly in* $L^q(\mathbb{T}^d)$ *for any* $q \geq 2$*.*

*(iv)* *For any* $\rho \in L^\infty(\mathbb{T}^d)$*,* $\forall j, k \in J$*,*

$$\lim_{n\to\infty} \int_{\mathbb{T}^d} \rho(x)\Psi^2_{n,j}(x)\Psi^2_{n,k}(x)\, dx = \delta^j_k \int_{\mathbb{T}^d} \rho(x)\Psi^4_j(x)\, dx$$

*where* $\delta^j_k$ *is the Kronecker delta with values* $\delta^j_k := 0$*, if* $k \neq j$*, and* $\delta^j_j := 1$*.*

Proof. Let $M \in \mathbb{N}$ be such that $J \subset [\![0, M]\!]$. It is enough to prove the lemma for $J = [\![0, M]\!]$. We recall the construction in Lemma 2.3 of [89]. Choose disjoints open balls $B_j := B(x_j, r_j)$, $0 \leq j \leq M$, of decreasing radii $r_0 > \ldots > r_M$ in such a way that, denoting by $\lambda_j$ the smallest Dirichlet eigenvalue of $-\Delta$ on $B_j$, it results

$$\lambda_0 < \ldots < \lambda_M < \lambda_0^{(2)}$$

where $\lambda_0^{(2)}$ is the second Dirichlet eigenvalue of $-\Delta$ on $B_1$. Define a sequence of $C^\infty$ positive potentials such that

$$q_n(x) := \begin{cases} n\,, & \forall x \in \mathbb{T}^d \backslash \bigcup_{i=0}^M B_i\,, \\ 0\,, & \forall x \in B(x_i, r_i - \epsilon_n)\,, \end{cases} \qquad \text{with} \quad \lim_{n\to\infty} \epsilon_n = 0\,, \quad 0 \leq q_n \leq n\,.$$

It is proved in [89], Lemma 2.3, that properties $(i)-(ii)$ hold with functions $\Psi_j$, $j \in [\![0, M]\!]$, in $H^1(\mathbb{T}^d)$, satisfying suppess$(\Psi_j) = B_j$ and such that $\Psi_{j|B_j} \in H^1_0(B_j)$ is an eigenfunction of $-\Delta$ associated to the eigenvalue $\lambda_j$. Such functions $\Psi_j$ are in $L^\infty(\mathbb{T}^d)$ because they are smooth in each ball $B_j$ and vanish outside.

To prove $(iii)$, it is sufficient to prove the $L^p$ bounds

$$\forall j \in [\![0, M]\!]\,,\ \forall p \geq 2\,, \quad \sup_n \|\Psi_{n,j}\|_{L^p(\mathbb{T}^d)} \leq C_{j,p} < +\infty\,. \tag{12.1.4}$$

Indeed, since $(\Psi_{n,j})_n$ converges to $\Psi_j$ in $L^2(\mathbb{T}^d)$ by item $(ii)$, the bound (12.1.4), the fact that $\Psi_j \in L^\infty(\mathbb{T}^d)$, and Hölder inequality, imply that the sequence $(\Psi_{n,j})_n$ converges to $\Psi_j$ in $L^q(\mathbb{T}^d)$ for any $q$.

We fix $0 \leq j \leq M$ and, for simplicity, we write $\Psi_{n,j} = \Psi_n$ in what follows. To prove (12.1.4) we perform a bootstrap argument for the $L^p$ norms of the solutions of the elliptic eigenvalue equation

$$-\Delta\Psi_n(x) + q_n(x)\Psi_n(x) = \lambda_n\Psi_n(x)\,. \tag{12.1.5}$$

Remark that the eigenfunctions $\Psi_n$ are in $C^\infty(\mathbb{T}^d)$, but we shall not perform Schauder estimates because we want bounds independent of the potentials $q_n(x)$ which are unbounded in Sobolev spaces. We multiply (12.1.5) by $|\Psi_n|^{r-2}\Psi_n$, $r \geq 2$, and integrate by parts on $\mathbb{T}^d$, obtaining

$$\int_{\mathbb{T}^d} \nabla\Psi_n(x) \cdot \nabla(|\Psi_n(x)|^{r-2}\Psi_n(x)) + q_n(x)|\Psi_n(x)|^r \, dx = \lambda_n \int_{\mathbb{T}^d} |\Psi_n(x)|^r \, dx \,. \tag{12.1.6}$$

Now

$$\begin{aligned} \nabla\Psi_n \cdot \nabla(|\Psi_n|^{r-2}\Psi_n) = (r-1)|\Psi_n|^{r-2}|\nabla\Psi_n|^2 &= (r-1)||\Psi_n|^{\frac{r}{2}-1}\nabla\Psi_n|^2 \\ &= K_r|\nabla z_n|^2 \end{aligned}$$

where

$$K_r := 4r^{-2}(r-1)\,, \qquad z_n := z_n(x) := |\Psi_n(x)|^{\frac{r}{2}-1}\Psi_n(x)\,.$$

Hence, by (12.1.6) and since $q_n \geq 0$,

$$K_r \int_{\mathbb{T}^d} |\nabla z_n(x)|^2 \, dx \leq \int_{\mathbb{T}^d} K_r|\nabla z_n(x)|^2 + q_n(x)|z_n(x)|^2 \leq \lambda_n \int_{\mathbb{T}^d} |z_n(x)|^2 \, dx\,,$$

which gives

$$\|z_n\|_{H^1} \leq (K_r^{-1}\lambda_n + 1)^{1/2}\|z_n\|_{L^2} \leq C_r\|z_n\|_{L^2}$$

because the sequence $(\lambda_n)_n$ is bounded, see item $(i)$. The continuous Sobolev embedding $H^1(\mathbb{T}^d) \hookrightarrow L^{d_*}(\mathbb{T}^d)$ with $d_* = \dfrac{2d}{d-2}$, implies that $\|z_n\|_{L^{d_*}} \leq C'_r\|z_n\|_{L^2}$, i.e.

$$\|\Psi_n\|_{L^{\frac{rd}{d-2}}} \leq C_r\|\Psi_n\|_{L^r}\,. \tag{12.1.7}$$

Iterating (12.1.7) (and starting from $r = 2$), we obtain that the sequence $(\|\Psi_n\|_{L^p})_n$ is bounded for any $p$. Note that if $d \leq 2$, we obtain (12.1.4) in one step only since in this case $H^1$ is continuously embedded in $L^p$ for any $p$.

$(iv)$ is a straightforward consequence of the convergence of the sequence $(\Psi_{n,j})_n$ to $\Psi_j$ in $L^4(\mathbb{T}^d)$ for all $j \geq M$ and of the fact that the functions $\Psi_j$ $(0 \leq j \leq M)$ have disjoint essential supports. ■

As a corollary we deduce the following lemma.

**Lemma 12.1.5. (Lemma 2.3 in [89])** *There is a $C^\infty$ potential $q(x)$ such that all the eigenvalues $\lambda_j(q)$, $j \in \mathbb{S}$, are simple (therefore $q$ is in $E_{\mathbb{S}}$), and the corresponding $L^2$-normalized eigenfunctions $\Psi_j(q)$, $j \in \mathbb{S}$, have the property that $(\Psi_j^2(q))_{j\in\mathbb{S}}$ are linearly independent.*

Consider the real analytic map

$$\Lambda : E_{\mathbb{S}} \to \mathbb{R}^{|\mathbb{S}|}\,, \quad \Lambda(V) := (\lambda_j(V))_{j\in\mathbb{S}}\,.$$

**Lemma 12.1.6.** *There is a $|\mathbb{S}|$-dimensional linear subspace $E$ of $C^\infty(\mathbb{T}^d)$ such that*

$$\mathcal{N}_{\mathbb{S}}^{(1)} := \big\{ V \in E_{\mathbb{S}} \ : \ d\Lambda(V)_{|E} \text{ is an isomorphism} \big\} \tag{12.1.8}$$

*is a $C^\infty$-dense open subset of $E_{\mathbb{S}}$, thus of $H^s(\mathbb{T}^d)$.*

Proof. We follow [89]. For any $\widehat{V} \in H^s(\mathbb{T}^d)$ we have that the differential

$$d\Lambda(q)[\widehat{V}] = \big( (\widehat{V}, \Psi_j^2)_{L^2} \big)_{j\in\mathbb{S}} .$$

Since the $\Psi_j^2(q)$ are linearly independent by Lemma 12.1.5, $d\Lambda(q)$ is onto, and there is a $|\mathbb{S}|$-dimensional linear subspace $E$ of $C^\infty(\mathbb{T}^d)$ such that $d\Lambda(q)_{|E}$ is an isomorphism. Let $(g_1, \ldots, g_{|\mathbb{S}|})$ be a basis of $E$. For any $V \in E_{\mathbb{S}}$, denote by $A_V$ the $|\mathbb{S}| \times |\mathbb{S}|$-matrix whose columns are given by $d\Lambda(V)[g_j]$, so that

$$\mathcal{N}_{\mathbb{S}}^{(1)} = \big\{ V \in E_{\mathbb{S}} \ : \ d\Lambda(V)_{|E} \text{ is an isomorphism} \big\} = \big\{ V \in E_{\mathbb{S}} \ : \ \det A_V \neq 0 \big\} .$$

$E_{\mathbb{S}}$ is a connected open subset of $H^s(\mathbb{T}^d)$ and the map $V \mapsto \det A_V$ is real analytic on $E_{\mathbb{S}}$ and does not vanish at $q$. Hence Lemma 12.1.2 implies that $\mathcal{N}_{\mathbb{S}}^{(1)}$ is a $C^\infty$-dense open subset of $E_{\mathbb{S}}$. ■

For any $V \in \mathcal{P}$, defined in (1.2.36), all the eigenvalues $\lambda_j$ of $-\Delta + V(x)$ are strictly positive, and therefore we deduce the following lemma:

**Lemma 12.1.7.** *The map*

$$\bar{\mu} : E_{\mathbb{S}} \cap \mathcal{P} \to \mathbb{R}^{|\mathbb{S}|} , \quad \bar{\mu}(V) := (\mu_j(V))_{j\in\mathbb{S}} = \big( \lambda_j^{\frac{1}{2}}(V) \big)_{j\in\mathbb{S}} , \tag{12.1.9}$$

*is real analytic and, for any $V \in \mathcal{N}_{\mathbb{S}}^{(1)}$, the differential $d\bar{\mu}(V)_{|E}$ is an isomorphism onto $\mathbb{R}^{|\mathbb{S}|}$.*

**Remark 12.1.8.** More generally, for any finite $J \subset \mathbb{N}$, there is a $C^\infty$-dense open subset $\mathcal{N}_J^{(1)}$ of $H^s(\mathbb{T}^d)$ such that, for all $V \in \mathcal{N}_J^{(1)} \cap \mathcal{P}$, the linear map $(d\mu_j(V))_{j\in J} : H^s(\mathbb{T}^d) \to \mathbb{R}^{|J|}$ is onto. ■

The $|\mathbb{S}|$-dimensional linear subspace $E$ of $C^\infty(\mathbb{T}^d)$ defined in Lemma 12.1.6 is the same subspace that appears in the statement of Theorem 1.2.3.

**Step 2. Genericity of the twist condition** (1.2.12)

Our aim is to prove that the twist matrix $\mathscr{A}$ defined in (1.2.9) is invertible for $(V, a)$ belonging to some $C^\infty$-dense open subset of $E_{\mathbb{S}} \times H^s(\mathbb{T}^d)$. Note that $\mathscr{A}$ is invertible if and only if $\det G \neq 0$ where the matrix $G := (G_k^j(V, a))_{j,k\in\mathbb{S}}$ is defined in (1.2.10). The matrix $G$ depends linearly on the function $a(x)$, and nonlinearly on the potential $V(x)$, through the eigenfunctions $\Psi_j := \Psi_j(V)$ defined in (1.1.5). By previous considerations, the functions $\Psi_j^2(x)$, $j \in \mathbb{S}$, where the eigenfunctions $\Psi_j(x)$ are normalized by the condition $\|\Psi_j\|_{L^2} = 1$, depend analytically on the potential $V \in E_{\mathbb{S}}$, and so each map

$$\begin{aligned} &G_k^j : H^s(\mathbb{T}^d) \times H^s(\mathbb{T}^d) \to \mathbb{R}\,, \\ &(V,a) \mapsto G_k^j(V,a) = \frac{3}{4}(2-\delta_k^j)(\Psi_j^2, a\Psi_k^2)_{L^2}\,, \quad \forall j,k \in \mathbb{S}\,, \end{aligned} \tag{12.1.10}$$

is real analytic on $E_{\mathbb{S}} \times H^s(\mathbb{T}^d)$, as well as the map $(V,a) \mapsto \det G(V,a)$.

**Lemma 12.1.9.** *The set*

$$\mathcal{N}_{\mathbb{S}}^{(2)} := \{V \in E_{\mathbb{S}} \ : \ \det G(V,1) \neq 0\}$$

*is a $C^\infty$-dense open subset of $E_{\mathbb{S}}$.*

Proof. Consider the sequence $(q_n)_n$ of $C^\infty$ potentials provided by Lemma 12.1.4 with $J = \mathbb{S}$. By property $(iv)$ of that lemma, taking the limit for $n \to \infty$ in (12.1.10), we get

$$\lim_{n\to\infty} G_k^j(q_n,1) = \frac{3}{4}\delta_k^j(\Psi_j^2, \Psi_k^2)_{L^2}$$

and therefore

$$\lim_{n\to\infty} \det G(q_n,1) = (3/4)^{|\mathbb{S}|}\prod_{j\in\mathbb{S}}\int_{\mathbb{T}^d} \Psi_j^4(x)\,dx =: \rho \neq 0\,. \tag{12.1.11}$$

In particular, there is a potential $q_n(x)$ in $E_{\mathbb{S}}$ such that $\det G(q_n,1) \not\equiv 0$. Since the map $V \mapsto \det G(V,1)$ is real analytic on the open and connected subset $E_{\mathbb{S}}$ of $H^s(\mathbb{T}^d)$, Lemma 12.1.2 implies that $\mathcal{N}_{\mathbb{S}}^{(2)}$ is $C^\infty$-dense open in $E_{\mathbb{S}}$. ■

We deduce the following corollary.

**Corollary 12.1.10.** *The set*

$$\mathcal{G}^{(2)} := \big\{(V,a) \in E_{\mathbb{S}} \times H^s(\mathbb{T}^d) \ : \ \det G(V,a) \neq 0\big\} \tag{12.1.12}$$

*is a $C^\infty$-dense open subset of $E_{\mathbb{S}} \times H^s(\mathbb{T}^d)$, thus of $H^s(\mathbb{T}^d) \times H^s(\mathbb{T}^d)$.*

Proof. By Lemma 12.1.9, for each potential $V \in \mathcal{N}_{\mathbb{S}}^{(2)}$, we have that $\det G(V,1) \neq 0$ and, since the function $(V,a) \mapsto \det G(V,a)$ is real analytic on the open and connected subset $E_{\mathbb{S}} \times H^s(\mathbb{T}^d)$, Lemma 12.1.2 implies that $\mathcal{G}^{(2)}$ is a $C^\infty$-dense open subset of $E_{\mathbb{S}} \times H^s(\mathbb{T}^d)$. ■

**Remark 12.1.11.** *With similar arguments we deduce that, for each potential $V \in \mathcal{N}_{\mathbb{S}}^{(2)}$, the set $\big\{a \in H^s(\mathbb{T}^d) \ : \ \det G(V,a) \neq 0\big\}$ is a $C^\infty$-dense open subset of $H^s(\mathbb{T}^d)$.*

**Step 3. Genericity of the non-degeneracy conditions** (1.2.21)-(1.2.22)

Let $M \in \mathbb{N}$ such that $\mathbb{S} \cup \mathbb{F} \subset [\![0, M]\!]$. We define

$$\mathcal{G}_M := \Big\{ (V, a) \in (E_{\mathbb{S}} \cap \mathcal{P}) \times H^s(\mathbb{T}^d) \ : \ \text{the following conditions hold} \tag{12.1.13}$$

1. $\big(\det \mathscr{A}\, \mu_j - [\mathscr{B}\mathscr{A}^\sharp \bar{\mu}]_j\big) - \big(\det \mathscr{A}\, \mu_k - [\mathscr{B}\mathscr{A}^\sharp \bar{\mu}]_k\big) \neq 0, \forall j, k \in \mathbb{S}^c, j, k \leq M, j \neq k,$

2. $\big(\det \mathscr{A}\, \mu_j - [\mathscr{B}\mathscr{A}^\sharp \bar{\mu}]_j\big) + \big(\det \mathscr{A}\, \mu_k - [\mathscr{B}\mathscr{A}^\sharp \bar{\mu}]_k\big) \neq 0\, \forall j, k \in \mathbb{S}^c\,,\, j, k \leq M \ \Big\}$

where $\mathscr{A} := \mathscr{A}(V, a)$, $\mathscr{B} := \mathscr{B}(V, a)$ are the Birkhoff matrices introduced in (1.2.9) and where $\mathscr{A}^\sharp$ denotes the comatrix of $\mathscr{A}$. Notice that

$$\text{for any } (V, a) \in \mathcal{G}^{(2)}\,, \ \text{defined in (12.1.12)}\,, \ \mathscr{A}^{-1} = \mathscr{A}^\sharp / \det \mathscr{A} \tag{12.1.14}$$

so that conditions 1, 2, above imply the non-degeneracy conditions (1.2.21)-(1.2.22) for $\mathbb{S} \cup \mathbb{F} \subset [\![0, M]\!]$.

**Proposition 12.1.12.** *The set $\mathcal{G}_M$ defined by* (12.1.13) *is a $C^\infty$-dense open subset of $\mathcal{P} \times H^s(\mathbb{T}^d)$. As a result, $\mathcal{G}_M \cap \mathcal{G}^{(2)}$ is a $C^\infty$-dense open subset of $\mathcal{P} \times H^s(\mathbb{T}^d)$ and, for any $(V, a) \in \mathcal{G}_M \cap \mathcal{G}^{(2)}$, the conditions* (1.2.21)-(1.2.22) *hold, provided that $\mathbb{S} \cup \mathbb{F} \subset [\![0, M]\!]$.*

Proof of Proposition 12.1.12. Define

$$\mathcal{I}_M := \big\{ (j, k, \sigma) \in \big([\![0, M]\!] \cap \mathbb{S}^c\big)^2 \times \{\pm 1\} \ : \ j \neq k \text{ or } \sigma = 1 \big\}\,,$$

and, for any $(j, k, \sigma) \in \mathcal{I}_M$, the real function $F_{j,k,\sigma}$ on $\mathcal{P} \times H^s(\mathbb{T}^d)$ by

$$F_{j,k,\sigma}(V, a) := \big(\det \mathscr{A}\, \mu_j - [\mathscr{B}\mathscr{A}^\sharp \bar{\mu}]_j\big) + \sigma\big(\det \mathscr{A}\, \mu_k - [\mathscr{B}\mathscr{A}^\sharp \bar{\mu}]_k\big)\,. \tag{12.1.15}$$

**Lemma 12.1.13.** *The set*

$$\mathcal{N}_{\mathbb{S},M} := \big\{ V \in E_M \cap \mathcal{P} \ : \ \forall (j, k, \sigma) \in \mathcal{I}_M\,, \ F_{j,k,\sigma}(V, 1) \neq 0 \big\} \tag{12.1.16}$$

*is an open and $C^\infty$-dense subset of $\mathcal{P}$.*

Proof. It is enough to prove that

$$\text{for each } (j, k, \sigma) \in \mathcal{I}_M\,, \ \text{there exists } V \in E_M \cap \mathcal{P} \text{ such that } F_{j,k,\sigma}(V, 1) \neq 0\,. \tag{12.1.17}$$

Indeed, since, for all $(j, k, \sigma) \in \mathcal{I}_M$, the function $F_{j,k,\sigma}(\cdot, 1)$ is real analytic on the open connected subset $E_M \cap \mathcal{P}$ of $H^s(\mathbb{T}^d)$, by (12.1.17), Lemma 12.1.2 and the finiteness of $\mathcal{I}_M$, we conclude that $\mathcal{N}_{\mathbb{S},M}$ is $C^\infty$-dense open in $E_M \cap \mathcal{P}$. Hence, since $E_M \cap \mathcal{P}$ is $C^\infty$-dense open in $\mathcal{P}$, the set $\mathcal{N}_{\mathbb{S},M}$ is $C^\infty$-dense open in $\mathcal{P}$.

To prove (12.1.17), we consider the sequence $(q_n)$ of potentials provided by Lemma 12.1.4, with $J = [\![0, M]\!]$. In particular, $q_n \in E_M \cap \mathcal{P}$. By Lemma 12.1.4-(i) we get

$$\forall j \in [\![0, M]\!]\,,\ \mu_{n,j} \overset{n\to+\infty}{\longrightarrow} \mu_j = \sqrt{\lambda_j} > 0 \ \text{ with } \ \mu_j \neq \mu_k \ \text{ for } \ j \neq k\,. \tag{12.1.18}$$

Moreover, by Lemma 12.1.4-($i$)-($ii$), for all $k \in \mathbb{S}$, $j \in [\![0, M]\!] \cap \mathbb{S}^c$, the matrix elements (recall (1.2.9), (1.2.10) and that $j \neq k$)

$$\big[\mathscr{B}(q_n, 1)\big]_j^k = \frac{3}{2}\mu_{n,j}^{-1}(\Psi_{n,j}^2, \Psi_{n,k}^2)_{L^2}\mu_{n,k}^{-1} \to 0 \quad \text{as } n \to +\infty\,. \tag{12.1.19}$$

In addition, by (1.2.9), (12.1.11) and (12.1.18), we have

$$\det \mathscr{A}(q_n, 1) = \Big(\prod_{j\in\mathbb{S}} \mu_{n,j}^{-1}\Big)^2 \det G(q_n, 1) \overset{n\to+\infty}{\longrightarrow} \Big(\prod_{j\in\mathbb{S}} \mu_j^{-1}\Big)^2 \rho =: \rho_1 > 0\,. \tag{12.1.20}$$

Therefore $(F_{j,k,\sigma}(q_n, 1))$ defined in (12.1.15) converges to $\rho_1(\mu_j + \sigma\mu_k)$, and since $(\mu_j)_{j\in[\![0,M]\!]}$ are distinct and strictly positive, $\rho_1(\mu_j + \sigma\mu_k) \neq 0$ for $(j, k, \sigma) \in \mathcal{I}_M$. This implies (12.1.17). ■

We have the following corollary.

**Corollary 12.1.14.** *For each potential $V(x) \in \mathcal{N}_{\mathbb{S},M}$ (defined in* (12.1.16)*), the set*

$$\mathcal{G}_{V,\mathbb{S},M} := \bigcap_{(j,k,\sigma)\in\mathcal{I}_M} \big\{a \in H^s(\mathbb{T}^d) \ : \ F_{j,k,\sigma}(V, a) \neq 0\big\}$$

*is $C^\infty$-dense in $H^s(\mathbb{T}^d)$.*

Proof. By Lemma 12.1.13, for each potential $V \in \mathcal{N}_{\mathbb{S},M}$, for each $(j, k, \sigma) \in \mathcal{I}_M$, we have that $F_{j,k,\sigma}(V, 1) \neq 0$. Since the function $a \mapsto F_{j,k,\sigma}(V, a)$ is real analytic on $H^s(\mathbb{T}^d)$, Lemma 12.1.2, and the fact that $\mathcal{I}_M$ is finite, imply that $\mathcal{G}_{V,\mathbb{S},M}$ is a $C^\infty$-dense open subset of $H^s(\mathbb{T}^d)$. ■

Proof of Proposition 12.1.12 concluded. The set $\mathcal{G}_M$ defined in (12.1.13) is clearly open. Moreover $\mathcal{G}_M$ is $C^\infty$-dense in $(E_\mathbb{S} \cap \mathcal{P}) \times H^s(\mathbb{T}^d)$, by Corollary 12.1.14, and because $\mathcal{N}_{\mathbb{S},M}$ is $C^\infty$-dense in $E_\mathbb{S} \cap \mathcal{P}$. Hence $\mathcal{G}_M$ is a $C^\infty$-dense open subset of $(E_\mathbb{S} \cap \mathcal{P}) \times H^s(\mathbb{T}^d)$. Since $E_\mathbb{S}$ is $C^\infty$-dense open in $\mathcal{P}$, the set $\mathcal{G}_M$ is a $C^\infty$-dense open subset of $\mathcal{P} \times H^s(\mathbb{T}^d)$. By Corollary 12.1.10, so is $\mathcal{G}_M \cap \mathcal{G}^{(2)}$, and by (12.1.14), we deduce the last claim of Proposition 12.1.12. ■

**Step 4. Genericity of finitely many first and second Melnikov conditions** (1.2.7), (1.2.16)-(1.2.19)

Let $L, M \in \mathbb{N}$ with $\mathbb{S} \subset [\![0, M]\!]$. Consider the following Conditions:

$(C1)$ $\bar{\mu}\cdot\ell+\mu_j\neq 0\,,\ \forall\ell\in\mathbb{Z}^{\mathbb{S}},|\ell|\le L,j\in\mathbb{S}^c,$

$(C2)$ $\bar{\mu}\cdot\ell+\mu_j-\mu_k\neq 0\,,\ \forall(\ell,j,k)\neq(0,j,j)\in\mathbb{Z}^{\mathbb{S}}\times([\![0,M]\!]\cap\mathbb{S}^c)\times\mathbb{S}^c\ ,|\ell|\le L\ ,$

$(C3)$ $\bar{\mu}\cdot\ell+\mu_j+\mu_k\neq 0\,,\ \forall(\ell,j,k)\in\mathbb{Z}^{\mathbb{S}}\times([\![0,M]\!]\cap\mathbb{S}^c)\times\mathbb{S}^c\ ,|\ell|\le L\ .$

Note that $(C1)$, $(C2)$, $(C3)$ correspond to the Melnikov conditions (1.2.7), (1.2.16)-(1.2.19).

We denote by $\mathcal{E}^{(1)}_L$, respectively $\mathcal{E}^{(2)}_{L,M}$, $\mathcal{E}^{(3)}_{L,M}$, the set of potentials $V\in E_{\mathbb{S}}\cap\mathcal{P}$ satisfying conditions $(C1)$, respectively $(C2)$, $(C3)$.

**Lemma 12.1.15.** *Let $L,M\in\mathbb{N}$. The set of potentials*

$$\mathcal{E}_{L,M}:=\mathcal{E}^{(1)}_L\cap\mathcal{E}^{(2)}_{L,M}\cap\mathcal{E}^{(3)}_{L,M}=\big\{V\in E_{\mathbb{S}}\cap\mathcal{P}\ :\ (C1),(C2),(C3)\ \textit{hold}\big\}\tag{12.1.21}$$

*is a $C^\infty$-dense open subset of $E_{\mathbb{S}}\cap\mathcal{P}$, thus of $\mathcal{P}$.*

Proof. First note that for any potential $q\in\mathcal{P}$, Weyl asymptotic formula about the distribution of the eigenvalues of $-\Delta+q(x)$ implies that

$$C_1 j^{1/d}\le\mu_j\le C_2 j^{1/d}\,,\quad\forall j\in\mathbb{N}\,,\tag{12.1.22}$$

for some positive constants $C_1,C_2$, which is uniform on some open neighborhood $B_q$ of $q$. Hence Condition $(C1)$ above may be violated by some $V\in B_q$ for $j\le C(|\bar{\mu}|L)^d$ only. Similarly, Conditions $(C2)$ and $(C3)$ may be violated by some $V\in B_q$ for $k\le C(M+(|\bar{\mu}|L)^d)$ only. Hence the inequalities in Conditions $(C1)$-$(C3)$ above are locally finitely many so that it enough to check that for any $\ell\in\mathbb{Z}^{|\mathbb{S}|}$, $j\in\mathbb{S}^c$, $k\in\mathbb{S}^c$, the sets

$$\begin{aligned}\mathcal{E}^{(1)}_{\ell,j}&:=\big\{V\in E_{\mathbb{S}}\cap\mathcal{P}\ :\ \bar{\mu}\cdot\ell+\mu_j\neq 0\big\}\ ,\\ \mathcal{E}^{(2)}_{\ell,j,k}&:=\big\{V\in E_{\mathbb{S}}\cap\mathcal{P}\ :\ \bar{\mu}\cdot\ell+\mu_j-\mu_k\neq 0\big\}\text{ with }\ell\neq 0\text{ or }j\neq k,\\ \mathcal{E}^{(3)}_{\ell,j,k}&:=\big\{V\in E_{\mathbb{S}}\cap\mathcal{P}\ :\ \bar{\mu}\cdot\ell+\mu_j+\mu_k\neq 0\big\}\end{aligned}\tag{12.1.23}$$

are $C^\infty$-dense open in $E_{\mathbb{S}}\cap\mathcal{P}$. The sets in (12.1.23) are open, since each $\mu_j$ depends continuously on $V\in E_{\mathbb{S}}\cap\mathcal{P}$. We now prove the $C^\infty$-density of $\mathcal{E}^{(2)}_{\ell,j,k}$, with $(\ell,j,k)\neq(0,j,j)$. Notice that the map

$$\Upsilon_{j,k}:E_{\mathbb{S}\cup\{j\}\cup\{k\}}\cap\mathcal{P}\to\mathbb{R}^{|\mathbb{S}|+2}\,,\quad V\mapsto\Upsilon_{j,k}(V):=(\bar{\mu},\mu_j,\mu_k)\,,$$

is real analytic and, by remark 12.1.8, the differential $d\Upsilon_{j,k}(V)$ is onto for any $V$ belonging to some $C^\infty$-dense open subset $\mathcal{N}^{(1)}_{\mathbb{S}\cup\{j\}\cup\{k\}}$ of $H^s(\mathbb{T}^d)$. Hence the map $V\mapsto\bar{\mu}\cdot\ell+\mu_j-\mu_k$ is real analytic on the connected open set $E_{\mathbb{S}\cup\{j\}\cup\{k\}}\cap\mathcal{P}$ and does not vanish everywhere for $(\ell,j,k)\neq(0,j,j)$. Therefore Lemma 12.1.2 implies that $\mathcal{E}^{(2)}_{\ell,j,k}$ is $C^\infty$-dense open in

$E_{\mathbb{S}\cup\{j\}\cup\{k\}} \cap \mathcal{P}$ , hence also in $E_{\mathbb{S}} \cap \mathcal{P}$. The same arguments can be applied to $\mathcal{E}^{(1)}_{\ell,j}$ and $\mathcal{E}^{(3)}_{\ell,j,k}$. ■

**Step 5. Genericity of the Diophantine conditions** (1.2.6)**,** (1.2.8)**, of the first Melnikov conditions** (1.2.7) **and the second Melnikov conditions** (1.2.16)**-**(1.2.19)
Consider the set of $(V, a)$ defined by

$$\mathcal{G}^{(3)} := \big( (\mathcal{N}^{(1)}_{\mathbb{S}} \cap \mathcal{P}) \times H^s(\mathbb{T}^d) \big) \bigcap \mathcal{G}^{(2)} \tag{12.1.24}$$

where $\mathcal{N}^{(1)}_{\mathbb{S}}$ is the set of potentials defined in (12.1.8) and $\mathcal{G}^{(2)}$ is the set of $(V, a)$ defined in (12.1.12) (for which the twist condition (1.2.12) holds).

**Lemma 12.1.16.** *$\mathcal{G}^{(3)}$ is a $C^\infty$-dense open subset of $\mathcal{P} \times H^s(\mathbb{T}^d)$.*

Proof. By Lemma 12.1.6 and Corollary 12.1.10. ■

We also remind that the map

$$\bar{\mu} : E_{\mathbb{S}} \cap \mathcal{P} \to \mathbb{R}^{|\mathbb{S}|} , \quad V \mapsto \bar{\mu}(V) := (\mu_j(V))_{j\in\mathbb{S}}$$

defined in (12.1.9) is real analytic, and, by Lemma 12.1.7, for any $V \in \mathcal{N}^{(1)}_{\mathbb{S}}$, the differential $d\bar{\mu}(V)_{|E}$ is an isomorphism, where $E$ is the $|\mathbb{S}|$-dimensional subspace of $C^\infty(\mathbb{T}^d)$ defined in Lemma 12.1.6.

We fix some $(\bar{V}, \bar{a}) \in \mathcal{G}^{(3)}$ and, according to the decomposition

$$H^s(\mathbb{T}^d) = E \oplus F , \quad \text{where} \quad F := E^{\perp_{L^2}} \cap H^s(\mathbb{T}^d) ,$$

we write uniquely

$$\bar{V} = \bar{v}_1 + \bar{v}_2 , \quad \bar{v}_1 \in E , \ \bar{v}_2 \in F ,$$

i.e. $\bar{v}_1$ is the projection of $\bar{V}$ on $E$ and $\bar{v}_2$ is the projection of $\bar{V}$ on $F$.

**Lemma 12.1.17.** *(i) There are open balls $B_1 \subset E$ (the subspace $E \simeq \mathbb{R}^{|\mathbb{S}|}$), $B_2 \subset F \subset H^s(\mathbb{T}^d)$, centered respectively at $\bar{v}_1 \in E$, $\bar{v}_2 \in F$, such that, for all $v_2 \in B_2$, the map*

$$u_{v_2} : B_1 \subset E \to \mathbb{R}^{|\mathbb{S}|} , \quad v_1 \mapsto u_{v_2}(v_1) := \bar{\mu}(v_1 + v_2)$$

*is a $C^1$ diffeomorphism from $B_1$ onto its image $\bar{\mu}(B_1 + v_2) =: \mathcal{O}_{v_2}$ which is an open bounded subset of $\mathbb{R}^{|\mathbb{S}|}$, the closure of which is included in $(0, +\infty)^{|\mathbb{S}|}$.*

*(ii) There is a constant $K_{\bar{V}} > 0$ such that the inverse functions*

$$u^{-1}_{v_2}(\cdot) : \mathcal{O}_{v_2} \subset \mathbb{R}^{|\mathbb{S}|} \to E$$

*are $K_{\bar{V}}$-Lipschitz continuous.*

*(iii) There is an open ball $B \subset H^s(\mathbb{T}^d)$ centered at $\bar{a}$ such that (possibly after reducing $B_1$ and $B_2$) the neighborhood $\mathcal{U}_{\bar{V},\bar{a}} := (B_1 + B_2) \times B$ of $(\bar{V}, \bar{a})$ has closure contained in $\mathcal{G}^{(3)}$, and there is a constant $C_{\bar{V},\bar{a}} > 0$ such that*

$$\|a\|_{L^\infty}(1 + \|\mathscr{A}^{-1}\|) \le C_{\bar{V},\bar{a}}\,, \quad \forall (V, a) \in \mathcal{U}_{\bar{V},\bar{a}}\,, \tag{12.1.25}$$

*where $\mathscr{A}$ is the Birkhoff matrix in* (1.2.9)-(1.2.10).

Proof. Since $\bar{V} = \bar{v}_1 + \bar{v}_2$ is in $\mathcal{N}_{\mathbb{S}}^{(1)}$, the differential $d\bar{\mu}(\bar{v}_1 + v_2)_{|E}$ is an isomorphism and the local inversion theorem implies item $(i)$ of the lemma. Items $(ii)$-$(iii)$ are straightforward taking $B_1$ and $B_2$ small enough. ■

In the sequel we shall always restrict to the neighborhood $\mathcal{U}_{\bar{V},\bar{a}}$ of $(\bar{V}, \bar{a}) \in \mathcal{G}^{(3)}$ provided by Lemma 12.1.17-$(iii)$.

For any $j \in \mathbb{N}$, $v_2 \in B_2$, we consider the $C^1$ function

$$\Xi_j : \mathcal{O}_{v_2} \subset \mathbb{R}^{|\mathbb{S}|} \to \mathbb{R}\,, \quad \omega \mapsto \Xi_j(\omega) := \mu_j(u_{v_2}^{-1}(\omega) + v_2)\,. \tag{12.1.26}$$

Note that, for any $j \in \mathbb{S}$, it results $\Xi_j(\omega) = \omega_j$.

**Lemma 12.1.18.** *There is a constant $\bar{K} := \bar{K}_{\bar{V}} > 0$ such that, for any potential $v_2 \in B_2 \subset F \subset H^s(\mathbb{T}^d)$, each function $\Xi_j$ defined in* (12.1.26) *is $\bar{K}$-Lipschitz continuous. Moreover there are constants $C, C' > 0$, such that, for all $v_2 \in B_2$,*

$$C j^{1/d} \le \Xi_j(\omega) \le C' j^{1/d}\,, \quad \forall j \in \mathbb{N}\,,\ \omega \in \mathcal{O}_{v_2}\,. \tag{12.1.27}$$

Proof. By (12.1.1), $\lambda_j = \mu_j^2$, and the fact that

$$\alpha := \inf_{V \in B_1 + B_2} \inf_{j \in \mathbb{N}} \mu_j(V) > 0$$

we get that

$$\forall V, V' \in B_1 + B_2\,,\ |\mu_j(V) - \mu_j(V')| \le \frac{1}{2\alpha}\|V - V'\|_{H^s}\,. \tag{12.1.28}$$

Since, by Lemma 12.1.17 the functions $u_{v_2}^{-1}$ are $K_{\bar{V}}$-Lipschitz continuous, (12.1.28) implies that the function $\Xi_j$ defined in (12.1.26) is $\dfrac{K_{\bar{V}}}{2\alpha}$-Lipschitz continuous. The bound (12.1.27) follows by (12.1.22). ■

Notice that, in view of (7.1.8), the set $\mathbb{M} := \mathbb{M}_{V,a} \subset \mathbb{N}\backslash\mathbb{S}$ associated to $(V, a)$ in Lemma 7.1.1 has an upper bound $C(V, a)$ that depends only on $\|\mathscr{A}^{-1}\|$ and $\|a\|_{L^\infty}$. Hence $\mathbb{M}$ can be taken constant for all $(V, a)$ in a neighborhood of $(\bar{V}, \bar{a})$, namely, by (12.1.25),

$$\exists \bar{M} \in \mathbb{N} \text{ such that,} \quad \forall (V, a) \in \mathcal{U}_{\bar{V},\bar{a}}\,, \quad \mathbb{M}_{V,a} \cup \mathbb{S} \subset [\![0, \bar{M}]\!]\,. \tag{12.1.29}$$

**Lemma 12.1.19.** *Fix an integer $\bar{L} \geq 4\bar{K}$ where the constant $\bar{K}$ is defined in Lemma 12.1.18 and let $\bar{M} \in \mathbb{N}$ be as in (12.1.29). Given a potential $v_2 \in B_2 \subset F \subset H^s(\mathbb{T}^d)$, we define, for any $\gamma > 0$, the subset of potentials $\mathcal{G}(\gamma, v_2) \subset B_1 \subset E \subset C^\infty$ of all the $v_1 \in B_1$ such that, for $V = v_1 + v_2$, the following Diophantine conditions hold:*

1. $|\bar{\mu} \cdot \ell| \geq \dfrac{\gamma}{\langle \ell \rangle^{\tau_0}}\,, \ \forall \ell \in \mathbb{Z}^{|\mathbb{S}|} \backslash \{0\}\,;$

2. $\Big| n + \sum_{i,j \in \mathbb{S}, i<j} p_{ij} \mu_i \mu_j \Big| \geq \dfrac{\gamma}{\langle p \rangle^{\tau_0}}\,, \quad \forall (n,p) \in \mathbb{Z} \times \mathbb{Z}^{\frac{|\mathbb{S}|(|\mathbb{S}|+1)}{2}} \setminus \{0\}\,;$

3. $|\bar{\mu} \cdot \ell + \mu_j| \geq \dfrac{\gamma}{\langle \ell \rangle^{\tau_0}}\,, \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|}\,,\ |\ell| \geq \bar{L}\,,\ j \in \mathbb{S}^c := \mathbb{N} \setminus \mathbb{S}\,;$

4. 
$$
\begin{aligned}
|\bar{\mu} \cdot \ell + \mu_j - \mu_k| \geq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}\,, \ & \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times (\llbracket 0, \bar{M} \rrbracket \cap \mathbb{S}^c) \times \mathbb{S}^c\,, \\
& (\ell, j, k) \neq (0, j, j)\,,\ |\ell| \geq \bar{L}\,,
\end{aligned}
\tag{12.1.30}
$$
$$
\begin{aligned}
|\bar{\mu} \cdot \ell + \mu_j + \mu_k| \geq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}\,, \ & \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times (\llbracket 0, \bar{M} \rrbracket \cap \mathbb{S}^c) \times \mathbb{S}^c\,, \\
& |\ell| \geq \bar{L}\,.
\end{aligned}
\tag{12.1.31}
$$

*Then the measure (on the finite dimensional subspace $E \simeq \mathbb{R}^{|\mathbb{S}|}$)*

$$
\Big| B_1 \backslash \Big( \bigcup_{\gamma > 0} \mathcal{G}(\gamma, v_2) \Big) \Big| = 0\,. \tag{12.1.32}
$$

Proof. In this lemma we denote by $m_{\rm Leb}$ the Lebesgue measure in $\mathbb{R}^{|\mathbb{S}|}$.
i) Let $\mathcal{F}_{1,\gamma}$ be the set of the Diophantine frequency vectors $\omega \in \mathcal{O}_{v_2}$ such that

$$
|\omega \cdot \ell| \geq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}\,, \ \forall \ell \in \mathbb{Z}^{|\mathbb{S}|} \backslash \{0\}\,. \tag{12.1.33}
$$

It is well known that, for $\tau_0 > |\mathbb{S}| - 1$, $m_{\rm Leb}(\mathcal{O}_{v_2} \backslash \mathcal{F}_{1,\gamma}) = O(\gamma)$.
ii) Let $\mathcal{F}_{2,\gamma}$ be the set of the frequency vectors $\omega \in \mathcal{O}_{v_2}$ such that

$$
\Big| n + \sum_{i,j \in \mathbb{S}, i<j} p_{ij} \omega_i \omega_j \Big| \geq \frac{\gamma}{\langle p \rangle^{\tau_0}}\,, \quad \forall (n,p) \in \mathbb{Z} \times \mathbb{Z}^{\frac{|\mathbb{S}|(|\mathbb{S}|+1)}{2}} \setminus \{0\}\,. \tag{12.1.34}
$$

Arguing as in Lemma 2.3.1 we deduce that $m_{\rm Leb}(\mathcal{O}_{v_2} \backslash \mathcal{F}_{2,\gamma}) = O(\gamma)$, see also Lemma 6.3 in [22].

iii) Let $\mathcal{F}_{3,\gamma}$ be the set of the frequency vectors $\omega \in \mathcal{O}_{v_2}$ such that

$$|\omega \cdot \ell + \Xi_j(\omega)| \geq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}, \quad \forall \ell \in \mathbb{Z}^{|\mathbb{S}|}, \ |\ell| \geq \bar{L}, \ j \in \mathbb{S}^c. \tag{12.1.35}$$

Let us prove that $m_{\text{Leb}}(\mathcal{O}_{v_2} \backslash \mathcal{F}_{3,\gamma}) = O(\gamma)$. Define on $\mathcal{O}_{v_2}$ the map $f_{\ell,j}(\omega) := \omega \cdot \ell + \Xi_j(\omega)$. By (12.1.27), there is a constant $C > 0$ such that if $j \geq C|\ell|^d$, then $f_{\ell,j} \geq 1$ on $\mathcal{O}_{v_2}$. Assume $j \leq C|\ell|^d$. Since $\Xi_j$ is $\bar{K}$-Lipschitz continuous by Lemma 12.1.18, we have, for $|\ell| \geq \bar{L} \geq 4\bar{K}$,

$$\frac{\ell}{|\ell|} \cdot \partial_\omega(\omega \cdot \ell + \Xi_j(\omega)) = |\ell| + \frac{\ell}{|\ell|} \cdot \partial_\omega \Xi_j(\omega) \geq \frac{3|\ell|}{4}.$$

Hence, for $|\ell| \geq \bar{L}$,

$$m_{\text{Leb}}\Big(\Big\{\omega \in \mathcal{O}_{v_2} \ : \ \exists j \in \mathbb{S}^c, \ |f_{\ell,j}(\omega)| \leq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}\Big\}\Big) \leq C|\ell|^d \frac{\gamma}{\langle \ell \rangle^{\tau_0+1}} \leq C \frac{\gamma}{\langle \ell \rangle^{\tau_0+1-d}}.$$

Hence

$$m_{\text{Leb}}(\mathcal{O}_{v_2} \backslash \mathcal{F}_{3,\gamma}) \leq C \sum_{|\ell| \geq \bar{L}} \frac{\gamma}{|\ell|^{\tau_0+1-d}} \leq C\gamma,$$

provided that $\tau_0 > d + |\mathbb{S}| - 1$.

iv) Let $\mathcal{F}_{4,\gamma}$, resp. $\mathcal{F}_{5,\gamma}$, be the set of the frequency vectors $\omega \in \mathcal{O}_{v_2}$ such that

$$\begin{gathered} |\omega \cdot \ell + \Xi_j(\omega) - \Xi_k(\omega)| \geq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}, \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times ([\![0, \bar{M}]\!] \cap \mathbb{S}^c) \times \mathbb{S}^c, \\ (\ell, j, k) \neq (0, j, j), \ |\ell| \geq \bar{L}, \end{gathered} \tag{12.1.36}$$

respectively

$$\begin{gathered} |\bar{\mu} \cdot \ell + \Xi_j(\omega) + \Xi_k(\omega)| \geq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}, \\ \forall (\ell, j, k) \in \mathbb{Z}^{|\mathbb{S}|} \times ([\![0, \bar{M}]\!] \cap \mathbb{S}^c) \times \mathbb{S}^c, \ |\ell| \geq \bar{L}. \end{gathered} \tag{12.1.37}$$

Define on $\mathcal{O}_{v_2}$ the map $f_{\ell,j,k}(\omega) := \omega \cdot \ell + \Xi_j(\omega) - \Xi_k(\omega)$. By (12.1.27), there is a constant $C$ such that, if $j \leq \bar{M}$ and $k \geq C(|\ell|^d + \bar{M})$, then $f_{\ell,j,k} \geq 1$ on $\mathcal{O}_{v_2}$. Moreover, since $\Xi_j$, $\Xi_k$ are $\bar{K}$-Lipschitz continuous by Lemma 12.1.18, we deduce that, for $|\ell| \geq \bar{L} \geq 4\bar{K}$,

$$\frac{\ell}{|\ell|} \cdot \partial_\omega(\omega \cdot \ell + \Xi_j(\omega) - \Xi_k(\omega)) = |\ell| + \frac{\ell}{|\ell|} \cdot \partial_\omega \Xi_j(\omega) - \frac{\ell}{|\ell|} \cdot \partial_\omega \Xi_k(\omega) \geq \frac{|\ell|}{2}.$$

Therefore, for $|\ell| \geq \bar{L}$,

$$\begin{aligned} & m_{\text{Leb}}\Big(\Big\{\omega \in \mathcal{O}_{v_2} \ : \ \exists (j,k) \in ([\![0, \bar{M}]\!] \cap \mathbb{S}^c) \times \mathbb{S}^c, \ |f_{\ell,j,k}(\omega)| \leq \frac{\gamma}{\langle \ell \rangle^{\tau_0}}\Big\}\Big) \\ & \leq C\bar{M}(|\ell|^d + \bar{M}) \frac{\gamma}{\langle \ell \rangle^{\tau_0+1}} \\ & \leq C'(\bar{M}) \frac{\gamma}{\langle \ell \rangle^{\tau_0+1-d}}. \end{aligned}$$

Hence, as in iii), $m_{\rm Leb}(\mathcal{O}_{v_2}\backslash\mathcal{F}_{4,\gamma}) \leq C\gamma$, provided that $\tau_0 > d + |\mathbb{S}| - 1$. We have the similar estimate $m_{\rm Leb}(\mathcal{O}_{v_2}\backslash\mathcal{F}_{5,\gamma}) \leq C\gamma$.

We have proved that $m_{\rm Leb}(\mathcal{O}_{v_2}\backslash\mathcal{F}_{i,\gamma}) = O(\gamma)$ for $i = 1, \ldots, 5$. We conclude that $\mathcal{F}(\gamma) := \cap_{i=1}^5 \mathcal{F}_{i,\gamma}$ satisfies

$$m_{\rm Leb}(\mathcal{O}_{v_2}\backslash\mathcal{F}(\gamma)) \leq \sum_{i=1}^{5} m_{\rm Leb}(\mathcal{O}_{v_2}\backslash\mathcal{F}_{i,\gamma}) = O(\gamma) \quad \Longrightarrow \quad m_{\rm Leb}\big(\mathcal{O}_{v_2}\backslash\bigcup_{\gamma>0}\mathcal{F}(\gamma)\big) = 0\,.$$

Finally, since by Lemma 12.1.17-($i$) the map $u_{v_2}$ is a diffeomorphism between $B_1$ and $\mathcal{O}_{v_2}$, the measure

$$m_{\rm Leb}(B_1\backslash\mathcal{G}(\gamma, v_2, a)) = m_{\rm Leb}(u_{v_2}^{-1}(\mathcal{O}_{v_2}\backslash\mathcal{F}(\gamma))) = 0\,.$$

This completes the proof of the lemma. ■

**Step 6. Conclusion: Proof of Theorem 1.2.3**

For any $(\bar V, \bar a) \in \mathcal{G}^{(3)}$ (see (12.1.24)), set introduced in (12.1.24), we define

$$\mathcal{G}_{\bar V,\bar a} := \big(\mathcal{E}_{\bar L,\bar M} \times H^s(\mathbb{T}^d)\big) \bigcap \mathcal{G}_{\bar M} \bigcap \mathcal{U}_{\bar V,\bar a} \tag{12.1.38}$$

where $\mathcal{U}_{\bar V,\bar a} \subset \mathcal{G}^{(3)}$ is the neighborhood of $(\bar V, \bar a)$ fixed in Lemma 12.1.17, and the sets $\mathcal{E}_{\bar L,\bar M}$, $\mathcal{G}_{\bar M}$ are defined respectively in (12.1.21), (12.1.13) with integers $\bar L$, $\bar M$, associated to $(\bar V, \bar a)$, that are fixed in Lemma 12.1.19 and (12.1.29).

**Lemma 12.1.20.** *The set $\mathcal{G}_{\bar V,\bar a}$ is $C^\infty$-dense open in $\mathcal{U}_{\bar V,\bar a}$.*

Proof. By Lemma 12.1.15 and Proposition 12.1.12. ■

Finally, we define the set $\mathcal{G}$ of Theorem 1.2.3 as

$$\mathcal{G} := \bigcup_{(\bar V,\bar a)\in\mathcal{G}^{(3)}} \mathcal{G}_{\bar V,\bar a}\,, \tag{12.1.39}$$

where $\mathcal{G}^{(3)}$ is defined in (12.1.24).

**Lemma 12.1.21.** *$\mathcal{G}$ is a $C^\infty$-dense open subset of $\mathcal{P} \times H^s(\mathbb{T}^d)$.*

Proof. Since $\mathcal{G}_{\bar V,\bar a}$ is open and $C^\infty$-dense in $\mathcal{U}_{\bar V,\bar a}$ by Lemma 12.1.20, the set $\mathcal{G}$ defined in (12.1.39) is open and $C^\infty$-dense in

$$\bigcup_{(\bar V,\bar a)\in\mathcal{G}^{(3)}} \mathcal{U}_{\bar V,\bar a} = \mathcal{G}^{(3)}\,.$$

(recall that $\mathcal{U}_{\bar V,\bar a} \subset \mathcal{G}^{(3)}$ by Lemma 12.1.17). Now, by Lemma 12.1.16, $\mathcal{G}^{(3)}$ is a $C^\infty$-dense open subset of $\mathcal{P} \times H^s(\mathbb{T}^d)$ and therefore $\mathcal{G}$ is a $C^\infty$-dense open subset of $\mathcal{P} \times H^s(\mathbb{T}^d)$. ■

The next lemma completes the proof of Theorem 1.2.3.

**Lemma 12.1.22.** *Let* $\widetilde{a} \in H^s(\mathbb{T}^d)$, $\widetilde{v}_2 \in E^{\perp_{L^2}} \cap H^s(\mathbb{T}^d)$, *where* $E$ *is the finite dimensional linear subspace of* $C^\infty(\mathbb{T}^d)$ *defined in Lemma 12.1.6. Then*

$$\Big| \big\{ v_1 \in E \,:\, (v_1 + \widetilde{v}_2, \widetilde{a}) \in \mathcal{G} \backslash \widetilde{\mathcal{G}} \big\} \Big| = 0 \tag{12.1.40}$$

*where* $\widetilde{\mathcal{G}}$ *is defined in* (1.2.37).

Proof. We may suppose that

$$\mathcal{W}_{\widetilde{v}_2, \widetilde{a}} := \big\{ v_1 \in E \,:\, (v_1 + \widetilde{v}_2, \widetilde{a}) \in \mathcal{G} \big\} \neq \emptyset \,,$$

otherwise (12.1.40) is trivial. Since $\mathcal{G}$ is open in $H^s(\mathbb{T}^d)$, the set $\mathcal{W}_{\widetilde{v}_2, \widetilde{a}}$ is an open subset of $E$ and, in order to deduce (12.1.40), it is enough to prove that, for any $\widetilde{v}_1 \in \mathcal{W}_{\widetilde{v}_2, \widetilde{a}}$, there is an open neighborhood $\mathcal{W}'_{\widetilde{v}_1} \subset \mathcal{W}_{\widetilde{v}_2, \widetilde{a}}$ of $\widetilde{v}_1$ such that

$$\Big| \big\{ v_1 \in \mathcal{W}'_{\widetilde{v}_1} \,:\, (v_1 + \widetilde{v}_2, \widetilde{a}) \notin \widetilde{\mathcal{G}} \big\} \Big| = 0 \,. \tag{12.1.41}$$

Since $(\widetilde{v}_1 + \widetilde{v}_2, \widetilde{a}) \in \mathcal{G}$, by the definition of $\mathcal{G}$ in (12.1.39), there is $(\bar{V}, \bar{a}) \in \mathcal{G}^{(3)}$ such that

$$(\widetilde{v}_1 + \widetilde{v}_2, \widetilde{a}) \in \mathcal{G}_{\bar{V}, \bar{a}} \stackrel{(12.1.38)}{=} \big( \mathcal{E}_{\bar{L}, \bar{M}} \times H^s(\mathbb{T}^d) \big) \bigcap \mathcal{G}_{\bar{M}} \bigcap \mathcal{U}_{\bar{V}, \bar{a}}$$

where

$$\mathcal{U}_{\bar{V}, \bar{a}} = (B_1 + B_2) \times B$$

is defined in Lemma 12.1.17. Define

$$\mathcal{W}'_{\widetilde{v}_1} := \big\{ v_1 \in E \,:\, (v_1 + \widetilde{v}_2, \widetilde{a}) \in \mathcal{G}_{\bar{V}, \bar{a}} \big\} \subset B_1 \,.$$

Since $\mathcal{G}_{\bar{V}, \bar{a}}$ is open, $\mathcal{W}'_{\widetilde{v}_1} \subset \mathcal{W}_{\widetilde{v}_2, \widetilde{a}}$ is an open neighborhood of $\widetilde{v}_1 \in \mathcal{W}_{\widetilde{v}_2, \widetilde{a}}$.

Now for all $(V, a) \in \mathcal{G}_{\bar{V}, \bar{a}}$, the twist condition (1.2.12) and the non-degeneracy properties (1.2.21)-(1.2.22) hold and, by the definition of $\mathcal{E}_{\bar{L}, \bar{M}}$ in (12.1.21) and by (12.1.29), there is $\gamma_0 = \gamma_0(V, a) > 0$ such that (1.2.7) and (1.2.16)-(1.2.19) for all $|\ell| \leq \bar{L}$. Thus, recalling the definition of the sets $\mathcal{G}(\gamma, \widetilde{v}_2)$ in Lemma 12.1.19 and $\widetilde{\mathcal{G}}$ in (1.2.37), and (12.1.29), we have

$$\mathcal{G}_{\bar{V}, \bar{a}} \bigcap \Big( \bigcup_{\gamma > 0} \mathcal{G}(\gamma, \widetilde{v}_2) \Big) \subset \widetilde{\mathcal{G}} \,,$$

so that

$$\big\{ v_1 \in \mathcal{W}'_{\widetilde{v}_1} \,:\, (v_1 + \widetilde{v}_2, \widetilde{a}) \notin \widetilde{\mathcal{G}} \big\} \subset B_1 \backslash \Big( \bigcup_{\gamma > 0} \mathcal{G}(\gamma, \widetilde{v}_2) \Big) \,.$$

Hence, by (12.1.32), the measure estimate (12.1.41) holds. This completes the proof of (12.1.40). ∎

# Appendix A

# Hamiltonian and Reversible PDEs

In this Appendix we first introduce the concept of Hamiltonian and/or reversible vector field. Then we shortly review the Hamiltonian and/or Reversible structure of some classical PDE.

## A.1 Hamiltonian and Reversible vector fields

Let $E$ be a real Hilbert space with scalar product $\langle\ ,\ \rangle$. Endow $E$ with a constant exact symplectic 2-form

$$\Omega(z, w) = \langle \bar{J} z, w \rangle \,, \quad \forall z, w \in E \,,$$

where $\bar{J} : E \to E$ is a non-degenerate, antisymmetric linear operator. Then, given a Hamiltonian function $H : \mathcal{D}(H) \subset E \to \mathbb{R}$, we associate the Hamiltonian system

$$u_t = X_H(u) \qquad \text{where} \qquad dH(u)[\cdot] = -\Omega(X_H(u), \cdot) \tag{A.1.1}$$

formally defines the Hamiltonian vector field $X_H$. The vector field $X_H : E_1 \subset E \to E$ is, in general, well defined and smooth only on a dense subspace $E_1 \subset E$. A continuous curve $[t_0, t_1] \ni t \mapsto u(t) \in E$ is a solution of (A.1.1) if it is $C^1$ as a map from $[t_0, t_1] \mapsto E_1$ and

$$u_t(t) = X_H(u(t)) \text{ in } E \,, \quad \forall t \in [t_0, t_1] \,.$$

If, for all $u \in E_1$, there is a vector $\nabla_u H(u) \in E$ such that the differential writes

$$dH(u)[h] = \langle \nabla_u H(u), h \rangle \,, \quad \forall h \in E_1 \,, \tag{A.1.2}$$

(since $E_1$ is in general not a Hilbert space with the scalar product $\langle\, ,\, \rangle$ then (A.1.2) does not follow by the Riesz theorem), then the Hamiltonian vector field $X_H : E_1 \mapsto E$ writes

$$X_H = J \nabla_u H \,, \qquad J := -\bar{J}^{-1} \,. \tag{A.1.3}$$

In PDE applications that we shall review below, usually $E = L^2$ and the dense subspace $E_1$ belongs to the Hilbert scale formed by the Sobolev spaces of periodic functions

$$H^s := \Big\{ u(x) = \sum_{j \in \mathbb{Z}^d} u_j e^{\mathrm{i} j \cdot x} \ : \ \|u\|_s^2 := \sum_{j \in \mathbb{Z}^d} |u_j|^2 (1 + |j|^{2s}) < +\infty \Big\}$$

for $s \geq 0$, or by the spaces of analytic functions

$$H^{\sigma,s} := \Big\{ u(x) = \sum_{j \in \mathbb{Z}^d} u_j \, e^{\mathrm{i} j \cdot x} \ : \ \|u\|_{\sigma,s}^2 := \sum_{j \in \mathbb{Z}^d} |u_j|^2 \, e^{2|j|\sigma} (1 + |j|^{2s}) < +\infty \Big\}$$

for $\sigma > 0$. For $s > d/2$ the spaces $H^s$ and $H^{\sigma,s}$ are an algebra with respect to the product of functions. We refer to [95] for a general functional setting of Hamiltonian PDEs on scales of Hilbert spaces.

**Reversible vector field.** A vector field $X$ is *reversible* if there exists an involution $S$ of the phase space, i.e a linear operator of $E$ satisfying $S^2 = \mathrm{Id}$, such that

$$X \circ S = -S \circ X \, . \tag{A.1.4}$$

Such condition is equivalent to the relation

$$\Phi^t \circ S = S \circ \Phi^{-t}$$

for the flow $\Phi^t$ associated to the vector field $X$. For a reversible equation it is natural to look for "reversible" solutions $u(t)$ of $u_t = X(u)$, namely such that

$$u(-t) = S u(t) \, .$$

If $S$ is antisymplectic, i.e. $S^* \Omega = -\Omega$, then a Hamiltonian vector field $X = X_H$ is reversible if and only if the Hamiltonian $H$ satisfies

$$H \circ S = H \, .$$

**Remark A.1.1.** *The possibility of developing KAM theory for reversible systems was first observed for finite dimensional systems by Moser in [99], see [110] for a complete presentation. In infinite dimension, the first KAM results for reversible PDEs have been obtained in [119].*

We now present some examples of Hamiltonian and/or Reversible PDE.

## A.2 Nonlinear wave and Klein-Gordon equations

We consider the Nonlinear wave equation (NLW)

$$y_{tt} - \Delta y + V(x) y = f(x, y) \,, \quad x \in \mathbb{T}^d := (\mathbb{R}/2\pi\mathbb{Z})^d \,, \quad y \in \mathbb{R} \,, \tag{A.2.1}$$

with a real valued multiplicative potential $V(x) \in \mathbb{R}$. If $V(x) = m$ is constant, (A.2.1) is also called a nonlinear Klein-Gordon equation.

The NLW equation (A.2.1) can be written as the first order Hamiltonian system

$$\frac{d}{dt}\begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} p \\ \Delta y - V(x) y + f(x, y) \end{pmatrix} = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}\begin{pmatrix} \nabla_y H(y,p) \\ \nabla_p H(y,p) \end{pmatrix}$$

where $\nabla_y H$, $\nabla_p H$ denote the $L^2(\mathbb{T}^d_x)$-gradient of the Hamiltonian

$$H(y, p) := \int_{\mathbb{T}^d} \frac{p^2}{2} + \frac{1}{2}\big((\nabla_x y)^2 + V(x) y^2\big) + F(x, y)\, dx \tag{A.2.2}$$

with potential density $F(x, y) := -\int_0^y f(x, z)\, dz$ and $\nabla_x y := (\partial_{x_1} y, \ldots, \partial_{x_d} y)$. Thus for the NLW equation $E = L^2 \times L^2$ with $L^2 := L^2(\mathbb{T}^d, \mathbb{R})$, the symplectic matrix

$$J = \bar{J} = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$$

and $\langle\ ,\ \rangle$ is the $L^2$ real scalar product. The variables $(y, p)$ are "Darboux coordinates".

**Remark A.2.1.** *The transposed operator $\bar{J}^\top = -\bar{J}$ (with respect to $\langle\, , \rangle$) and $\bar{J}^{-1} = \bar{J}^\top$.*

The Hamiltonian $H$ is properly defined on the subspaces $E_1 := H^s \times H^s$, $s \geq 2$, so that the Hamiltonian vector field is a map

$$J \nabla_u H : H^s \times H^s \to H^{s-2} \times H^{s-2} \subset L^2 \times L^2 \,.$$

Note that the loss of two derivatives is only due to the Laplace operator $\Delta$ and that the Hamiltonian vector field generated by the nonlinearity is bounded, because the composition operator

$$y(x) \mapsto f(x, y(x)) \tag{A.2.3}$$

is a map of $H^s$ into itself. For such a reason the PDE (A.2.1) is *semi-linear*.

If the nonlinearity $f(x, u)$ is analytic in the variable $u$, and $H^{s_0}$, $s_0 > d/2$, in the space variable $x$, then the composition operator (A.2.3) is analytic from $H^{s_0}$ in itself. It is only finitely many times differentiable on $H^{s_0}$ if the nonlinearity $f(x, u)$ is only (sufficiently) many times differentiable with respect to $u$.

**Remark A.2.2.** *The regularity of the vector field is relevant for KAM theory: for finite dimensional systems, it has been rigorously proved that, if the vector field is not sufficiently smooth, then all the invariant tori could be destroyed and only discontinuous Aubry-Mather invariant sets survive, see e.g. [82].*

If the potential density $F(x, y, \nabla_x y)$ in (A.2.2) depends also on the first order derivative $\nabla_x y$, we obtain a *quasi-linear* wave equation. For simplicity we write explicitly only the equation in dimension $d = 1$. Given the Hamiltonian

$$H(y,p) := \int_{\mathbb{T}} \frac{p^2}{2} + \frac{1}{2}\big(y_x^2 + V(x)y^2\big) + F(x,y,y_x)\, dx$$

we derive the Hamiltonian wave equation

$$y_{tt} - y_{xx} + V(x)y = f(x,y,y_x,y_{xx})$$

with nonlinearity (denoting by $(x, y, \zeta)$ the independent variables of the potential density $(x, y, \zeta) \mapsto F(x, y, \zeta)$) given by

$$\begin{aligned} f(x,y,y_x,y_{xx}) &= -(\partial_y F)(x,y,y_x) + \frac{d}{dx}\big\{(\partial_\zeta F)(x,y,y_x)\big\} \\ &= -(\partial_y F)(x,y,y_x) + (\partial_{\zeta x} F)(x,y,y_x) + (\partial_{\zeta y} F)(x,y,y_x)y_x + (\partial_{\zeta\zeta} F)(x,y,y_x)y_{xx}\,. \end{aligned}$$

Note that $f$ depends on all the derivatives $y, y_x, y_{xx}$ but it is linear in the second order derivatives $y_{xx}$, i.e. $f$ is quasi-linear. The nonlinear composition operator

$$y(x) \mapsto f(x, y(x), y_x(x), y_{xx}(x))$$

maps $H^s \to H^{s-2}$, i.e. loses two derivatives.

Derivative Wave equations. If the nonlinearity $f(x, y, y_x)$ depends only on first order derivatives, then the Hamiltonian structure of the wave equation is lost (at least the usual one). However such equation can admit a reversible structure. Consider the derivative wave, or, better called, derivative Klein-Gordon equation

$$y_{tt} - y_{xx} + my = f(x, y, y_x, y_t)\,, \quad x \in \mathbb{T}\,,$$

where the nonlinearity depends also on the first order space and time derivatives $(y_x, y_t)$, and write it as the first order system

$$\frac{d}{dt}\begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} p \\ y_{xx} - my + f(x,y,y_x,p) \end{pmatrix}.$$

Its vector field $X$ is *reversible* (see (A.1.4)) with respect to the involution

$$S : (y,p) \to (y,-p)\,, \qquad \text{resp.} \ \ S : (y(x), p(x)) \to (y(-x), -p(-x))\,,$$

assuming the reversibility condition

$$f(x,y,y_x,-p) = f(x,y,y_x,p)\,, \qquad \text{resp.} \ \ f(x,y,-y_x,p) = f(-x,y,y_x,-p)\,.$$

KAM results have been obtained for reversible derivative wave equations in [18].

## A.3 Nonlinear Schrödinger equation

Consider the Hamiltonian Schrödinger equation

$$\mathrm{i}u_t - \Delta u + V(x)u = f(x,u)\,, \quad x \in \mathbb{T}^d\,,\ u \in \mathbb{C}\,, \tag{A.3.1}$$

where $f(x,u) = \partial_{\bar{u}} F(x,u)$ and the potential $F(x,u) \in \mathbb{R}$, $\forall u \in \mathbb{C}$, is real valued. The NLS equation (A.3.1) can be written as the infinite dimensional complex Hamiltonian equation

$$u_t = \mathrm{i}\nabla_{\bar{u}} H(u)\,, \quad H(u) := \int_{\mathbb{T}^d} |\nabla u|^2 + V(x)|u|^2 - F(x,u)\,dx\,.$$

Actually (A.3.1) is a real Hamiltonian PDE. In the variables $(a,b) \in \mathbb{R}^2$, real and imaginary part of

$$u = a + \mathrm{i}b\,,$$

denoting the real valued potential $W(a,b) := F(x, a+\mathrm{i}b)$ so that

$$\partial_{\bar{u}} F(x, a+\mathrm{i}b) := \frac{1}{2}(\partial_a + \mathrm{i}\partial_b)W(a,b)\,,$$

the NLS equation (A.3.1) reads

$$\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \Delta b - V(x)b + \frac{1}{2}\partial_b W(a,b) \\ -\Delta a + V(x)a - \frac{1}{2}\partial_a W(a,b) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 & -\mathrm{Id} \\ \mathrm{Id} & 0 \end{pmatrix}\begin{pmatrix} \nabla_a H(a,b) \\ \nabla_b H(a,b) \end{pmatrix}$$

with real valued Hamiltonian $H(a,b) := H(a+\mathrm{i}b)$ and $\nabla_a, \nabla_b$ denote the $L^2$-real gradients. The variables $(a,b)$ are "Darboux coordinates".

A simpler Hamiltonian pseudo-differential model equation often studied is

$$\mathrm{i}u_t - \Delta u + V * u = \partial_{\bar{u}} F(x,u)\,, \quad x \in \mathbb{T}^d\,,\ u \in \mathbb{C}\,, \tag{A.3.2}$$

where the convolution potential $V * u$ is the Fourier multipliers operator

$$u(x) = \sum_{j \in \mathbb{Z}^d} u_j e^{\mathrm{i}j\cdot x} \quad \mapsto \quad V * u := \sum_{j \in \mathbb{Z}^d} V_j u_j e^{\mathrm{i}j\cdot x}$$

with real valued Fourier multipliers $V_j \in \mathbb{R}$. The Hamiltonian of (A.3.2) is

$$H(u) := \int_{\mathbb{T}^d} |\nabla u|^2 + (V * u)\,\bar{u} - F(x,u)\,dx\,.$$

Also for the NLS equation (A.3.1), the Hamiltonian vector field loses two derivatives because of the Laplace operator $\Delta$. On the other hand the nonlinear Hamiltonian vector field $\mathrm{i}\partial_{\bar{u}} F(x,u)$ is bounded, so that the PDE (A.3.1), as well as (A.3.2), is a semi-linear equation.

If the nonlinearity depends also on first and second order derivatives, we have, respectively, derivative NLS (DNLS) and fully-non-linear (or quasi-linear) Schrödinger equations. For simplicity we present the model equations only in dimension $d = 1$. For the DNLS equation

$$\mathrm{i}u_t + u_{xx} = f(x, u, u_x) \tag{A.3.3}$$

the Hamiltonian structure is lost (at least the usual one). However (A.3.3) is reversible with respect to the involution

$$S : u \mapsto \bar{u}$$

(see (A.1.4)) if the nonlinearity $f$ satisfies the condition

$$f(x, \bar{u}, \bar{u}_x) = \overline{f(x, u, u_x)}\,.$$

KAM results for DNLS have been proved in [119]. On the other hand, fully nonlinear, or quasi-linear, perturbed NLS equations like

$$\mathrm{i}u_t = u_{xx} + \varepsilon f(\omega t, x, u, u_x, u_{xx})$$

may be both reversible or Hamiltonian. They have been considered in [62].

## A.4 Perturbed KdV equations

An important class of equations which arises in fluid mechanics concerns nonlinear perturbations

$$u_t + u_{xxx} + \partial_x u^2 + \mathcal{N}(x, u, u_x, u_{xx}, u_{xxx}) = 0\,, \quad x \in \mathbb{T}\,, \tag{A.4.1}$$

of the KdV equation

$$u_t + u_{xxx} + \partial_x u^2 = 0\,. \tag{A.4.2}$$

Here the unknown $u(x) \in \mathbb{R}$ is *real* valued. If the nonlinearity $\mathcal{N}(x, u, u_x, u_{xx}, u_{xxx})$ does not depend on $u_{xxx}$ such equation is semilinear. The most general Hamiltonian (local) nonlinearity is

$$\mathcal{N}(x, u, u_x, u_{xx}, u_{xxx}) := -\partial_x \big[ (\partial_u f)(x, u, u_x) - \partial_x ((\partial_{u_x} f)(x, u, u_x)) \big] \tag{A.4.3}$$

which is quasi-linear. In this case (A.4.1) is the Hamiltonian PDE

$$u_t = \partial_x \nabla_u H(u)\,, \qquad H(u) = \int_{\mathbb{T}} \frac{u_x^2}{2} - \frac{u^3}{3} + f(x, u, u_x)\, dx\,, \tag{A.4.4}$$

where $\nabla_u H$ denotes the $L^2(\mathbb{T}_x)$ gradient.

The "mass" $\int_{\mathbb{T}} u(x)\,dx$ is a prime integral of (A.4.1)-(A.4.3) and a natural phase space is

$$H^1_0(\mathbb{T}_x) := \Big\{ u(x) \in H^1(\mathbb{T}, \mathbb{R}) \ : \ \int_{\mathbb{T}} u(x) dx = 0 \Big\} .$$

Thus for the Hamiltonian PDE (A.4.4) we have

$$E = L^2_0(\mathbb{T}, \mathbb{R}) := \Big\{ u \in L^2(\mathbb{T}, \mathbb{R}) \ : \ \int_{\mathbb{T}} u(x)\,dx = 0 \Big\} , \quad \bar{J} = -\partial_x^{-1} ,$$

and $\langle \ , \ \rangle$ is the $L^2$ real scalar product. Note that $J = -\bar{J}^{-1} = \partial_x$ see (A.1.3). The nonlinear Hamiltonian vector field is defined and smooth on the subspaces $E_1 = H^s_0(\mathbb{T})$, $s \geq 3$, because

$$\mathcal{N}(x, u, u_x, u_{xx}, u_{xxx}) : H^s_0(\mathbb{T}) \mapsto H^{s-3}_0(\mathbb{T}) \subset L^2_0(\mathbb{T}) .$$

Note that, if the Hamiltonian density $f = f(x, u)$ does not depend on the first order derivative $u_x$, the nonlinearity in (A.4.3) reduces to $\mathcal{N} = -\partial_x[(\partial_u f)(x, u)]$ and so the PDE (A.4.1)-(A.4.3) is semilinear.

Birkhoff coordinates. The KdV equation (A.4.2) has very special algebraic properties: it possesses infinitely many analytic prime integrals in involution, i.e. pairwise commuting, and it is integrable in the strongest possible sense, namely it possesses global analytic action-angle variables, called the Birkhoff coordinates. The whole infinite dimensional phase space is foliated by quasi-periodic and almost-periodic solutions. The quasi-periodic solutions are called the "finite gap" solutions. Kappeler and collaborators (see for example [90] and references therein) proved that there exists an analytic symplectic diffeomorphism

$$\Psi^{-1} : H^N_0(\mathbb{T}) \to \ell^2_{N+\frac{1}{2}}(\mathbb{R}) \times \ell^2_{N+\frac{1}{2}}(\mathbb{R}) ,$$

from the Sobolev spaces $H^N_0(\mathbb{T}) := H^N(\mathbb{T}) \cap L^2_0(\mathbb{T})$, $N \geq 0$, to the spaces of coordinates $(x, y) \in \ell^2_{N+\frac{1}{2}}(\mathbb{R}) \times \ell^2_{N+\frac{1}{2}}(\mathbb{R})$ equipped with the canonical symplectic form $\sum_{n \geq 1} dx_n \wedge dy_n$, where

$$\ell^2_s(\mathbb{R}) := \Big\{ x := (x_n)_{n \in \mathbb{N}} , \ x_n \in \mathbb{R} \ : \ \sum_{n \geq 1} x_n^2 n^{2s} < +\infty \Big\} ,$$

such that, for $N = 1$, the KdV Hamiltonian

$$H_{KdV}(u) = \int_{\mathbb{T}} \frac{u_x^2}{2} - \frac{u^3}{3}\,dx ,$$

expressed in the new coordinates, i.e. $K := H_{KdV} \circ \Psi$, depends only on $x_n^2 + y_n^2$, $n \geq 1$ (actions). The KdV equation appears therefore, in these new coordinates, as an infinite chain of anharmonic oscillators, whose frequencies

$$\omega_n(I) := \partial_{I_n} K(I) , \quad I = (I_1, I_2, \ldots) , \quad I_n = (x_n^2 + y_n^2)/2$$

depend on their amplitudes in a nonlinear and real analytic fashion. This is the basis for studying perturbations of the finite gap solutions of KdV.

**Remark A.4.1.** *The existence of Birkhoff coordinates is much more than what is needed for local KAM perturbation theory. As noted by Kuksin in [94]-[95] it is sufficient that the unperturbed torus is reducible, namely that the linearized equation at the quasi-periodic solution have, in a suitable set of coordinates, constant coefficients.*

Other integrable Hamiltonian PDEs which possess Birkhoff coordinates are the mKdV equation, see [91],

$$u_t + u_{xxx} + \partial_x u^3 = 0\,, \quad x \in \mathbb{T}\,, \tag{A.4.5}$$

and the cubic 1-d NLS equation, see [74],

$$\mathrm{i} u_t = u_{xx} + |u|^2 u\,, \quad x \in \mathbb{T}\,. \tag{A.4.6}$$

# Appendix B

# Multiscale Step

In this Appendix we provide the proof of the multiscale Step Proposition B.2.4 proved in [23], which is used to prove Proposition 4.3.4.

- **Notation.** In this Appendix we use the notation of [23], in particular be aware that the Definitions B.2.1, B.2.2, B.2.3 below of $N$-good/bad matrix, regular/singular sites, and $(A, N)$-good/bad site are different with respect to those introduced in section 4.3 which are used in the Monograph.

In order to give a self-contained presentation we first prove the properties of the decay norms introduced in [23], recalled in section 3.3.

## B.1 Matrices with off-diagonal decay

Let $e_i = e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)}$ for $i := (\ell, j) \in \mathbb{Z}^b := \mathbb{Z}^\nu \times \mathbb{Z}^d$. In the vector-space $\mathcal{H}^s = \mathcal{H}^s(\mathbb{T}^\nu \times \mathbb{T}^d; \mathbb{C}^r)$ defined in (3.3.2) with $\nu = |\mathbb{S}|$, we consider the basis

$$e_k = e_i e_{\mathfrak{a}} \,, \quad k := (i, \mathfrak{a}) \in \mathbb{Z}^b \times \mathfrak{I} \,, \tag{B.1.1}$$

where $e_{\mathfrak{a}} := (0, \ldots, \underbrace{1}_{\mathfrak{a}-th}, \ldots, 0) \in \mathbb{C}^r$, $\mathfrak{a} = 1, \ldots, r$, denote the canonical basis of $\mathbb{C}^r$, and

$$\mathfrak{I} := \{1, \ldots, r\} \,.$$

Then we write any $u \in \mathcal{H}^s$ as

$$u = \sum_{k \in \mathbb{Z}^b \times \mathfrak{I}} u_k e_k \,, \quad u_k \in \mathbb{C} \,.$$

For $B \subset \mathbb{Z}^b \times \mathfrak{I}$, we introduce the subspace

$$\mathcal{H}^s_B := \Big\{ u \in \mathcal{H}^s \, : \, u_k = 0 \text{ if } k \notin B \Big\} .$$

When $B$ is finite, the space $\mathcal{H}^s_B$ does not depend on $s$ and will be denoted $\mathcal{H}_B$. We define

$$\Pi_B : \mathcal{H}^s \to \mathcal{H}_B$$

the $L^2$-orthogonal projector onto $\mathcal{H}_B$.

In what follows $B, C, D, E$ are finite subsets of $\mathbb{Z}^b \times \mathfrak{I}$.

We identify the space $\mathcal{L}^B_C$ of the linear maps $L : \mathcal{H}_B \to \mathcal{H}_C$ with the space of matrices

$$\mathcal{M}^B_C := \Big\{ M = (M^{k'}_k)_{k' \in B, k \in C} \, , \ M^{k'}_k \in \mathbb{C} \Big\}$$

according to the following usual definition.

**Definition B.1.1.** *The matrix $M \in \mathcal{M}^B_C$ represents the linear operator $L \in \mathcal{L}^B_C$, if*

$$\forall k' = (i', \mathfrak{a}') \in B, \; k = (i, \mathfrak{a}) \in C \, , \quad \Pi_k L e_{k'} = M^{k'}_k e_k \, ,$$

*where $e_k$ are defined in* (B.1.1) *and $M^{k'}_k \in \mathbb{C}$.*

Example. The multiplication operator for

$$\begin{pmatrix} p(\varphi, x) & q(\varphi, x) \\ \overline{q}(\varphi, x) & p(\varphi, x) \end{pmatrix} ,$$

acting in $\mathcal{H}^s(\mathbb{T}^\nu \times \mathbb{T}^d; \mathbb{C}^2)$, is represented by the matrix

$$T := (T^{i'}_i)_{i, i' \in \mathbb{Z}^b} \qquad \text{where} \qquad T^{i'}_i = \begin{pmatrix} p_{i-i'} & q_{i-i'} \\ (\overline{q})_{i-i'} & p_{i-i'} \end{pmatrix} \tag{B.1.2}$$

and $p_i, q_i$ denote the Fourier coefficients of $p(\varphi, x), q(\varphi, x)$. With the above notation, the set $\mathfrak{I} = \{1, 2\}$ and

$$T^{(i',1)}_{(i,1)} = p_{i-i'} \, , \ T^{(i',2)}_{(i,1)} = q_{i-i'} \, , \ T^{(i',1)}_{(i,2)} = (\overline{q})_{i-i'} \, , \ T^{(i',1)}_{(i,2)} = p_{i-i'} \, .$$

Notation. For any subset $B$ of $\mathbb{Z}^b \times \mathfrak{I}$, we denote by

$$\overline{B} := \mathrm{proj}_{\mathbb{Z}^b} B \tag{B.1.3}$$

the projection of $B$ in $\mathbb{Z}^b$.

Given $B \subset B'$, $C \subset C' \subset \mathbb{Z}^b \times \mathfrak{I}$ and $M \in \mathcal{M}_{C'}^{B'}$ we can introduce the restricted matrices

$$M_C^B := \Pi_C M_{|\mathcal{H}_B}\,, \quad M_C := \Pi_C M\,, \quad M^B := M_{|\mathcal{H}_B}\,. \tag{B.1.4}$$

If $D \subset \mathrm{proj}_{\mathbb{Z}^b} B'$, $E \subset \mathrm{proj}_{\mathbb{Z}^b} C'$, then we define

$$M_E^D \ \text{ as } \ M_C^B \quad \text{where} \quad B := (D \times \mathfrak{I}) \cap B',\ C := (E \times \mathfrak{I}) \cap C'\,. \tag{B.1.5}$$

In the particular case $D = \{i'\}$, $E := \{i\}$, $i, i' \in \mathbb{Z}^b$, we use the simpler notation

$$M_i := M_{\{i\}} \quad \text{(it is either a line or a group of } 2, \ldots, r \text{ lines of } M)\,, \tag{B.1.6}$$

$$M^{i'} := M^{\{i'\}} \quad \text{(it is either a column or a group of } 2, \ldots, r \text{ columns of } M)\,, \tag{B.1.7}$$

and

$$M_i^{i'} := M_{\{i\}}^{\{i'\}}\,, \tag{B.1.8}$$

it is a $m \times m'$-complex matrix, where $m \in \{1, \ldots, r\}$ (resp. $m' \in \{1, \ldots, r\}$) is the cardinality of $C$ (resp. of $B$) defined in (B.1.5) with $E := \{i\}$ (resp. $D = \{i'\}$).

We endow the vector-space of the $m \times m'$, $m, m' \in \{1, \ldots, r\}$, complex matrices with a norm $|\ |$ such that

$$|UW| \le |U||W|\,,$$

whenever the dimensions of the matrices make their multiplication possible, and $|U| \le |V|$ if $U$ is a submatrix of $V$.

**Remark B.1.2.** *The notation in* (B.1.5), (B.1.6), (B.1.7), (B.1.8)*, may be not very specific, but it is deliberate: it is convenient not to distinguish the index* $\mathfrak{a} \in \mathfrak{I}$*, which is irrelevant in the definition of the s-norms, in Definition B.1.3.*

We also set the $L^2$-operatorial norm

$$\|M_C^B\|_0 := \sup_{h \in \mathcal{H}_B, h \neq 0} \frac{\|M_C^B h\|_0}{\|h\|_0} \tag{B.1.9}$$

where $\|\ \|_0 := \|\ \|_{L^2}$.

**Definition B.1.3. ($s$-norm)** *The s-norm of a matrix* $M \in \mathcal{M}_C^B$ *is defined by*

$$|M|_s^2 := \sum_{n \in \mathbb{Z}^b} [M(n)]^2 \langle n \rangle^{2s} \tag{B.1.10}$$

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

$$[M(n)] := \begin{cases} \max_{i - i' = n, i \in \overline{C}, i' \in \overline{B}} |M_i^{i'}| & \text{if } \ n \in \overline{C} - \overline{B} \\ 0 & \text{if } \ n \notin \overline{C} - \overline{B}\,, \end{cases} \tag{B.1.11}$$

*with* $\overline{B} := \mathrm{proj}_{\mathbb{Z}^b} B$*,* $\overline{C} := \mathrm{proj}_{\mathbb{Z}^b} C$ *(see* (B.1.3)*).*

It is easy to check that $| \ |_s$ is a norm on $\mathcal{M}_C^B$. It verifies $| \ |_s \leq | \ |_{s'}$, $\forall s \leq s'$, and

$$\forall M \in \mathcal{M}_C^B\,, \quad \forall B' \subseteq B\ ,\ C' \subseteq C\ , \quad |M_{C'}^{B'}|_s \leq |M|_s\,.$$

The $s$-norm is designed to estimate the off-diagonal decay of matrices similar to the Töplitz matrix which represents the multiplication operator for a Sobolev function.

**Lemma B.1.4.** *The matrix $T$ in* (B.1.2) *with $(p,q) \in \mathcal{H}^s(\mathbb{T}^\nu \times \mathbb{T}^d; \mathbb{C}^2)$ satisfies*

$$|T|_s \lesssim \|(q,p)\|_s\,. \tag{B.1.12}$$

Proof. By (B.1.11), (B.1.2) we get

$$[T(n)] := \max_{i-i'=n} \left| \begin{pmatrix} p_{i-i'} & q_{i-i'} \\ \overline{q_{i-i'}} & \overline{p_{i-i'}} \end{pmatrix} \right| \lesssim |p_n| + |q_n|\,.$$

Hence, the definition in (B.1.10) implies

$$|T|_s^2 = \sum_{n \in \mathbb{Z}^b} [T(n)]^2 \langle n \rangle^{2s} \lesssim \sum_{n \in \mathbb{Z}^b} (|p_n| + |q_n|)^2 \langle n \rangle^{2s} \lesssim \|(p,q)\|_s^2$$

and (B.1.12) follows. ∎

In order to prove that the matrices with finite $s$-norm satisfy the interpolation inequalities (B.1.15), and then the algebra property (B.1.16), the guiding principle is the analogy between these matrices and the Töpliz matrices which represent the multiplication operator for functions. We introduce the set $\mathcal{H}_+$ of the trigonometric polynomials with positive Fourier coefficients

$$\mathcal{H}_+ := \Big\{ h = \sum h_{\ell,j} e^{\mathrm{i}(\ell\cdot\varphi + j\cdot x)} \text{ with } h_{\ell,j} \neq 0 \\ \text{for a finite number of } (\ell,j) \text{ only and } h_{\ell,j} \in \mathbb{R}_+ \Big\}\,.$$

Note that the sum and the product of two functions in $\mathcal{H}_+$ remain in $\mathcal{H}_+$.

**Definition B.1.5.** *Given $M \in \mathcal{M}_C^B$, $h \in \mathcal{H}_+$, we say that $M$ is dominated by $h$, and we write $M \prec h$, if*

$$[M(n)] \leq h_n\,, \quad \forall n \in \mathbb{Z}^b\,, \tag{B.1.13}$$

*in other words if* $|M_i^{i'}| \leq h_{i-i'}$ , $\forall i' \in \mathrm{proj}_{\mathbb{Z}^b} B$, $i \in \mathrm{proj}_{\mathbb{Z}^b} C$.

It is easy to check ($B$ and $C$ being finite) that

$$|M|_s = \min\Big\{ \|h\|_s \ :\ h \in \mathcal{H}_+\ ,\ M \prec h \Big\} \quad \text{and} \\ \exists h \in \mathcal{H}_+\ ,\ \forall s \geq 0\,,\ |M|_s = \|h\|_s\,. \tag{B.1.14}$$

**Lemma B.1.6.** *For $M_1 \in \mathcal{M}_D^C$, $M_2 \in \mathcal{M}_C^B$, $M_3 \in \mathcal{M}_D^C$, we have*

$$\begin{gathered} M_1 \prec h_1\,,\ M_2 \prec h_2\,,\ M_3 \prec h_3 \quad \Longrightarrow \\ M_1 + M_3 \prec h_1 + h_3 \quad \text{and} \quad M_1 M_2 \prec h_1 h_2\,. \end{gathered}$$

Proof. Property $M_1 + M_3 \prec h_1 + h_3$ is straightforward. For $i \in \mathrm{proj}_{\mathbb{Z}^b} D$, $i' \in \mathrm{proj}_{\mathbb{Z}^b} B$, we have

$$\begin{aligned} |(M_1 M_2)_i^{i'}| = \Big| \sum_{q \in \overline{C} := \mathrm{proj}_{\mathbb{Z}^b} C} (M_1)_i^q (M_2)_q^{i'} \Big| \ &\leq \ \sum_{q \in \overline{C}} |(M_1)_i^q| |(M_2)_q^{i'}| \\ &\leq \ \sum_{q \in \overline{C}} (h_1)_{i-q} (h_2)_{q-i'} \\ &\leq \ \sum_{q \in \mathbb{Z}^b} (h_1)_{i-q} (h_2)_{q-i'} = (h_1 h_2)_{i-i'} \end{aligned}$$

implying $M_1 M_2 \prec h_1 h_2$ by Definition B.1.5. ■

We deduce from (B.1.14) and Lemma 3.5.1, the following interpolation estimates.

**Lemma B.1.7. (Interpolation)** *For all $s \geq s_0 > (d+\nu)/2$, there is $C(s) \geq 1$, such that, for any finite subset $B, C, D \subset \mathbb{Z}^b \times \mathfrak{I}$, for all matrices $M_1 \in \mathcal{M}_D^C$, $M_2 \in \mathcal{M}_C^B$,*

$$|M_1 M_2|_s \leq C(s_0) |M_1|_{s_0} |M_2|_s + C(s) |M_1|_s |M_2|_{s_0}\,, \tag{B.1.15}$$

*in particular,*

$$|M_1 M_2|_s \leq C(s) |M_1|_s |M_2|_s\,. \tag{B.1.16}$$

Note that the constant $C(s)$ in Lemma B.1.7 is independent of $B$, $C$, $D$.

**Lemma B.1.8.** *For all $s \geq s_0 > (d+\nu)/2$, there is $C(s) \geq 1$, such that, for any finite subset $B, C, D \subset \mathbb{Z}^b \times \mathfrak{I}$, for all $M_1 \in \mathcal{M}_D^C$, $M_2 \in \mathcal{M}_C^B$, we have*

$$|M_1 M_2|_{s_0} \leq C(s_0) |M_1|_{s_0} |M_2|_{s_0}\,, \tag{B.1.17}$$

*and, $\forall M \in \mathcal{M}_B^B$, $\forall n \geq 1$,*

$$\begin{aligned} &|M^n|_{s_0} \leq (C(s_0))^{n-1} |M|_{s_0}^n\,, \\ &|M^n|_s \leq C(s) (C(s_0))^{n-1} |M|_{s_0}^{n-1} |M|_s\,,\ \forall s \geq s_0\,. \end{aligned} \tag{B.1.18}$$

Proof. The first estimate in (B.1.18) follows by (B.1.16) with $s = s_0$ and the second estimate in (B.1.18) is obtained from (B.1.15), using $C(s) \geq 1$. ■

The $s$-norm of a matrix $M \in \mathcal{M}_C^B$ controls also the Sobolev $\mathcal{H}^s$-norm. Indeed, we identify $\mathcal{H}_B$ with the space $\mathcal{M}_B^{\{0\}}$ of column matrices and the Sobolev norm $\|\ \|_s$ is equal to the $s$-norm $|\ |_s$, i.e.

$$\forall w \in \mathcal{H}_B\,, \quad \|w\|_s = |w|_s\,, \quad \forall s \geq 0\,. \tag{B.1.19}$$

Then $Mw \in \mathcal{H}_C$ and the next lemma is a particular case of Lemma B.1.7.

**Lemma B.1.9. (Sobolev norm)** *$\forall s \geq s_0$ there is $C(s) \geq 1$ such that, for any finite subset $B, C \subset \mathbb{Z}^b \times \mathfrak{I}$, for any $M \in \mathcal{M}_C^B$, $w \in \mathcal{H}_B$,*

$$\|Mw\|_s \leq C(s_0)|M|_{s_0}\|w\|_s + C(s)|M|_s\|w\|_{s_0} \,. \tag{B.1.20}$$

The following lemma is the analogue of the smoothing properties of the projection operators.

**Lemma B.1.10. (Smoothing)** *Let $M \in \mathcal{M}_C^B$. Then, $\forall s' \geq s \geq 0$,*

$$M_i^{i'} = 0 \,, \ \forall |i - i'| < N \quad \Longrightarrow \quad |M|_s \leq N^{-(s'-s)}|M|_{s'} \,, \tag{B.1.21}$$

*and, for $N \geq N_0$,*

$$M_i^{i'} = 0 \,, \ \forall |i - i'| > N \quad \Longrightarrow \quad \begin{cases} |M|_{s'} \leq N^{s'-s}|M|_s \\ |M|_s \leq N^{s+b}\|M\|_0 \,. \end{cases} \tag{B.1.22}$$

Proof. Estimate (B.1.21) and the first bound of (B.1.22) follow from the definition of the norms $| \ |_s$. The second bound of (B.1.22) follows by the first bound in (B.1.22), noting that $|M_i^{i'}| \leq \|M\|_0$, $\forall i, i'$,

$$|M|_s \leq N^s|M|_0 \leq N^s\sqrt{(2N+1)^b}\|M\|_0 \leq N^{s+b}\|M\|_0$$

for $N \geq N_0$. ■

In the next lemma we bound the $s$-norm of a matrix in terms of the $(s+b)$-norms of its lines.

**Lemma B.1.11. (Decay along lines)** *Let $M \in \mathcal{M}_C^B$. Then, $\forall s \geq 0$,*

$$|M|_s \lesssim \max_{i \in \mathrm{proj}_{\mathbb{Z}^b} C} |M_{\{i\}}|_{s+b} \tag{B.1.23}$$

*(we could replace the index $b$ with any $\alpha > b/2$).*

Proof. For all $i \in \overline{C} := \mathrm{proj}_{\mathbb{Z}^b} C$, $i' \in \overline{B} := \mathrm{proj}_{\mathbb{Z}^b} B$, $\forall s \geq 0$,

$$|M_i^{i'}| \leq \frac{|M_{\{i\}}|_{s+b}}{\langle i - i' \rangle^{s+b}} \leq \frac{m(s+b)}{\langle i - i' \rangle^{s+b}}$$

where $m(s+b) := \max_{i \in \overline{C}} |M_{\{i\}}|_{s+b}$. As a consequence

$$|M|_s = \Big( \sum_{n \in \overline{C} - \overline{B}} (M[n])^2 \langle n \rangle^{2s} \Big)^{1/2} \leq m(s+b) \Big( \sum_{n \in \mathbb{Z}^b} \langle n \rangle^{-2b} \Big)^{1/2}$$

implying (B.1.23). ■

The $L^2$-norm and $s_0$-norm of a matrix are related.

**Lemma B.1.12.** *Let $M \in \mathcal{M}_B^C$. Then, for $s_0 > (d+\nu)/2$,*

$$\|M\|_0 \lesssim_{s_0} |M|_{s_0} . \tag{B.1.24}$$

Proof. Let $m \in \mathcal{H}_+$ be such that $M \prec m$ and $|M|_s = \|m\|_s$ for all $s \geq 0$, see (B.1.14). Also for $H \in \mathcal{M}_C^{\{0\}}$, there is $h \in \mathcal{H}_+$ such that $H \prec h$ and $|H|_s = \|h\|_s$, $\forall s \geq 0$. Lemma B.1.6 implies that $MH \prec mh$ and so

$$|MH|_0 \leq \|mh\|_0 \leq |m|_{L^\infty} \|h\|_0 \lesssim_{s_0} \|m\|_{s_0} \|h\|_0 = |M|_{s_0} |H|_0 , \quad \forall H \in \mathcal{M}_C^{\{0\}} .$$

Then (B.1.24) follows (recall (B.1.19)). ■

In the sequel we use the notion of left invertible operators.

**Definition B.1.13. (Left Inverse)** *A matrix $M \in \mathcal{M}_C^B$ is left invertible if there exists $N \in \mathcal{M}_B^C$ such that $NM = \mathrm{Id}_B$. Then $N$ is called a left inverse of $M$.*

Note that $M$ is left invertible if and only if $M$ (considered as a linear map) is injective (then $\dim \mathcal{H}_C \geq \dim \mathcal{H}_B$). The left inverses of $M$ are not unique if $\dim \mathcal{H}_C > \dim \mathcal{H}_B$: they are uniquely defined only on the range of $M$.

We shall often use the following perturbation lemma for left invertible operators. Note that the bound (B.1.25) for the perturbation in $s_0$-norm only, allows to estimate the inverse (B.1.28) also in $s \geq s_0$ norm.

**Lemma B.1.14. (Perturbation of left invertible matrices)** *If $M \in \mathcal{M}_C^B$ has a left inverse $N \in \mathcal{M}_B^C$ , then there exists $\delta(s_0) > 0$ such that,*

$$\forall P \in \mathcal{M}_C^B \qquad \text{with} \qquad |N|_{s_0} |P|_{s_0} \leq \delta(s_0) , \tag{B.1.25}$$

*the matrix $M + P$ has a left inverse $N_P$ that satisfies*

$$|N_P|_{s_0} \leq 2 |N|_{s_0} , \tag{B.1.26}$$

*and, $\forall s \geq s_0$,*

$$|N_P|_s \leq \big(1 + C(s)|N|_{s_0}|P|_{s_0}\big)|N|_s + C(s)|N|_{s_0}^2 |P|_s \tag{B.1.27}$$
$$\lesssim_s |N|_s + |N|_{s_0}^2 |P|_s . \tag{B.1.28}$$

*Moreover,*

$$\forall P \in \mathcal{M}_C^B \qquad \text{with} \qquad \|N\|_0 \|P\|_0 \leq 1/2 , \tag{B.1.29}$$

*the matrix $M + P$ has a left inverse $N_P$ that satisfies*

$$\|N_P\|_0 \leq 2 \|N\|_0 . \tag{B.1.30}$$

PROOF.
*Proof of* (B.1.26). The matrix $N_P = AN$ with $A \in \mathcal{M}^B_B$ is a left inverse of $M+P$ if and only if

$$I_B = AN(M+P) = A(I_B + NP)\,,$$

i.e. if and only if $A$ is the inverse of $I_B + NP \in \mathcal{M}^B_B$. By (B.1.17) and (B.1.25) we have, taking $\delta(s_0) > 0$ small enough,

$$|NP|_{s_0} \le C(s_0)|N|_{s_0}|P|_{s_0} \le C(s_0)\delta(s_0) \le 1/2\,. \tag{B.1.31}$$

Hence the matrix $I_B + NP$ is invertible and

$$N_P = AN = (I_B + NP)^{-1}N = \sum_{p=0}^{\infty}(-1)^p(NP)^pN \tag{B.1.32}$$

is a left inverse of $M+P$. Estimate (B.1.26) follows by (B.1.32) and (B.1.31).
*Proof of (B.1.27).* For all $s \ge s_0$, $\forall p \ge 1$,

$$\begin{aligned} |(NP)^pN|_s &\overset{(B.1.15)}{\lesssim_s} |N|_{s_0}|(NP)^p|_s + |N|_s|(NP)^p|_{s_0} \\ &\overset{(B.1.18)}{\lesssim_s} |N|_{s_0}(C(s_0)|NP|_{s_0})^{p-1}|NP|_s + |N|_s(C(s_0)|NP|_{s_0})^p \\ &\overset{(\mathrm{B.1.31}),(\mathrm{B.1.15})}{\lesssim_s} 2^{-p}(|N|_{s_0}|P|_{s_0}|N|_s + |N|^2_{s_0}|P|_s)\,. \end{aligned} \tag{B.1.33}$$

We derive (B.1.27) by

$$\begin{aligned} |N_P|_s &\overset{(B.1.32)}{\le} |N|_s + \sum_{p\ge1}|(NP)^pN|_s \\ &\overset{(B.1.33)}{\le} |N|_s + C(s)(|N|_{s_0}|P|_{s_0}|N|_s + |N|^2_{s_0}|P|_s)\,. \end{aligned}$$

Finally (B.1.30) follows from (B.1.29) as (B.1.26) because the operatorial $L^2$-norm (see (B.1.9)) satisfies the algebra property $\|NP\|_0 \le \|N\|_0\|P\|_0$. ■

## B.2 Multiscale step Proposition

This section is devoted to prove the multiscale step Proposition B.2.4.

In the whole section $\varsigma \in (0,1)$ is fixed and $\tau' > 0$, $\Theta \ge 1$ are real parameters, on which we shall impose conditions in Proposition B.2.4.

Given $\Omega, \Omega' \subset E \subset \mathbb{Z}^b \times \mathfrak{I}$ we define

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

where, for $k = (i, \mathfrak{a})$, $k' := (i', \mathfrak{a}')$ we set

$$|k - k'| := \begin{cases} 1 & \text{if } i = i' \,, \mathfrak{a} \neq \mathfrak{a}' \,, \\ 0 & \text{if } i = i' \,, \mathfrak{a} = \mathfrak{a}' \,, \\ |i - i'| & \text{if } i \neq i' \,. \end{cases}$$

**Definition B.2.1. ($N$-good/bad matrix [23])** *The matrix $A \in \mathcal{M}_E^E$, with $E \subset \mathbb{Z}^b \times \mathfrak{I}$,* $\text{diam}(E) \leq 4N$*, is $N$-good if $A$ is invertible and*

$$\forall s \in [s_0, s_1] \,, \quad |A^{-1}|_s \leq N^{\tau' + \delta s}. \tag{B.2.1}$$

*Otherwise $A$ is $N$-bad.*

**Definition B.2.2. (Regular/Singular sites [23])** *The index $k := (i, \mathfrak{a}) \in \mathbb{Z}^b \times \mathfrak{I}$ is* REGULAR *for $A$ if $|A_k^k| \geq \Theta$. Otherwise $k$ is* SINGULAR.

**Definition B.2.3. ($(A, N)$-good/bad site [23])** *For $A \in \mathcal{M}_E^E$, we say that $k \in E \subset \mathbb{Z}^b \times \mathfrak{I}$ is*

- $(A, N)$-REGULAR *if there is $F \subset E$ such that* $\text{diam}(F) \leq 4N$, $\text{d}(k, E \backslash F) \geq N$ *and $A_F^F$ is $N$-good.*
- $(A, N)$-GOOD *if it is regular for $A$ or $(A, N)$-regular. Otherwise we say that $k$ is* $(A, N)$-BAD.

Let us consider the new larger scale

$$N' = N^\chi \tag{B.2.2}$$

with $\chi > 1$.

For a matrix $A \in \mathcal{M}_E^E$ we define $\text{Diag}(A) := (\delta_{kk'} A_k^{k'})_{k,k' \in E}$.

**Proposition B.2.4. (Multiscale step [23])** *Assume*

$$\varsigma \in (0, 1/2) \,, \ \tau' > 2\tau + b + 1 \,, \ C_1 \geq 2 \,, \tag{B.2.3}$$

*and, setting $\kappa := \tau' + b + s_0$,*

$$\chi(\tau' - 2\tau - b) > 3(\kappa + (s_0 + b)C_1) \,, \ \chi\varsigma > C_1 \,, \tag{B.2.4}$$

$$s_1 > 3\kappa + \chi(\tau + b) + C_1 s_0 \,. \tag{B.2.5}$$

*For any given $\Upsilon > 0$, there exist $\Theta := \Theta(\Upsilon, s_1) > 0$ large enough (appearing in Definition B.2.2), and $N_0(\Upsilon, \Theta, s_1) \in \mathbb{N}$ such that:*
*$\forall N \geq N_0(\Upsilon, \Theta, s_1)$, $\forall E \subset \mathbb{Z}^b \times \mathfrak{I}$ with* $\text{diam}(E) \leq 4N' = 4N^\chi$ *(see* (B.2.2)*), if $A \in \mathcal{M}_E^E$ satisfies*

- **(H1)** $|A - \mathrm{Diag}(A)|_{s_1} \le \Upsilon$
- **(H2)** $\|A^{-1}\|_0 \le (N')^{\tau}$
- **(H3)** *There is a partition of the* $(A, N)$*-bad sites* $B = \cup_\alpha \Omega_\alpha$ *with*

$$\mathrm{diam}(\Omega_\alpha) \le N^{C_1} , \quad \mathrm{d}(\Omega_\alpha, \Omega_\beta) \ge N^2 \ , \ \forall \alpha \ne \beta \, , \tag{B.2.6}$$

*then* $A$ *is* $N'$*-good. More precisely*

$$\forall s \in [s_0, s_1] \ , \quad |A^{-1}|_s \le \frac{1}{4}(N')^{\tau'} \Big( (N')^{\varsigma s} + |A - \mathrm{Diag}(A)|_s \Big) \, , \tag{B.2.7}$$

*and, for all* $s \ge s_1$*,*

$$|A^{-1}|_s \le C(s)(N')^{\tau'} \big( (N')^{\varsigma s} + |A - \mathrm{Diag}(A)|_s \big) \, . \tag{B.2.8}$$

The above proposition says, roughly, the following. If $A$ has a sufficient off-diagonal decay (assumption (H1) and (B.2.5)), and if the sites that can not be inserted in good "small" submatrices (of size $O(N)$) along the diagonal of $A$ are sufficiently separated (assumption (H3)), then the $L^2$-bound $(H2)$ for $A^{-1}$ implies that the "large" matrix $A$ (of size $N' = N^\chi$ with $\chi$ as in (B.2.4)) is good, and $A^{-1}$ satisfies also the bounds (B.2.7) in $s$-norm for $s > s_1$. Notice that the bounds for $s > s_1$ follow only by informations on the $N$-good submatrices in $s_1$-norm (see Definition B.2.1) plus the $s$-decay of $A$. The link between the various constants is the following:

- According to (B.2.4) the exponent $\chi$, which measures the new scale $N' \gg N$, is large with respect to the size of the bad clusters $\Omega_\alpha$, i.e. with respect to $C_1$. The intuitive meaning is that, for $\chi$ large enough, the "resonance effects" due to the bad clusters are "negligible" at the new larger scale.

- The constant $\Theta \ge 1$ which defines the regular sites (see Definition B.2.2) must be large enough with respect to $\Upsilon$, i.e. with respect to the off diagonal part $\mathcal{T} := A - \mathrm{Diag}(A)$, see (H1) and Lemma B.2.5.

- Note that $\chi$ in (B.2.4) can be taken large independently of $\tau$, choosing, for example, $\tau' := 3\tau + 2b$.

- The Sobolev index $s_1$ has to be large with respect to $\chi$ and $\tau$, according to (B.2.5). This is also natural: if the decay is sufficiently strong, then the "interaction" between different clusters of $N$-bad sites is weak enough.

- In (B.2.6) we have fixed the separation $N^2$ between the bad clusters just for definiteness: any separation $N^\mu$, $\mu > 0$, would be sufficient. Of course, the smaller $\mu > 0$ is, the larger the Sobolev exponent $s_1$ has to be.

The proof of Proposition B.2.4 is divided in several lemmas. In each of them we shall assume that the hypotheses of Proposition B.2.4 are satisfied. We set

$$\mathcal{T} := A - \mathrm{Diag}(A)\,, \qquad |\mathcal{T}|_{s_1} \overset{(H1)}{\leq} \Upsilon\,. \tag{B.2.9}$$

Call $G$ (resp. $B$) the set of the $(A, N)$-good (resp. bad) sites. The partition

$$E = B \cup G$$

induces the orthogonal decomposition

$$\mathcal{H}_E = \mathcal{H}_B \oplus \mathcal{H}_G$$

and we write

$$u = u_B + u_G \qquad \text{where} \qquad u_B := \Pi_B u\,,\ u_G := \Pi_G u\,.$$

We shall denote by $I_G$, resp. $I_B$, the restriction of the identity matrix to $\mathcal{H}_G$, resp. $\mathcal{H}_B$, according to (B.1.4).

The next Lemmas B.2.5 and B.2.6 say that the system $Au = h$ can be nicely reduced along the good sites $G$, giving rise to a (non-square) system $A' u_B = Zh$, with a good control of the $s$-norms of the matrices $A'$ and $Z$. Moreover $A^{-1}$ is a left inverse of $A'$.

**Lemma B.2.5. (Semi-reduction on the good sites)** *Let $\Theta^{-1}\Upsilon \leq c_0(s_1)$ be small enough. There exist $\mathcal{M} \in \mathcal{M}_G^E$, $\mathcal{N} \in \mathcal{M}_G^B$ satisfying, if $N \geq N_1(\Upsilon)$ is large enough,*

$$|\mathcal{M}|_{s_0} \leq cN^{\kappa}\,, \quad |\mathcal{N}|_{s_0} \leq c\,\Theta^{-1}\Upsilon\,, \tag{B.2.10}$$

*for some $c := c(s_1) > 0$, and, $\forall s \geq s_0$,*

$$\begin{aligned} |\mathcal{M}|_s &\leq C(s) N^{2\kappa}\big(N^{s-s_0} + N^{-b}|\mathcal{T}|_{s+b}\big)\,,\\ |\mathcal{N}|_s &\leq C(s) N^{\kappa}\big(N^{s-s_0} + N^{-b}|\mathcal{T}|_{s+b}\big)\,, \end{aligned} \tag{B.2.11}$$

*such that*

$$Au = h \quad \Longrightarrow \quad u_G = \mathcal{N}u_B + \mathcal{M}h\,.$$

*Moreover*

$$u_G = \mathcal{N}u_B + \mathcal{M}h \quad \Longrightarrow \quad \forall k \text{ regular}\,,\ (Au)_k = h_k\,. \tag{B.2.12}$$

Proof. It is based on "resolvent identity" arguments.

**Step I.** *There exist $\Gamma, L \in \mathcal{M}_G^E$ satisfying*

$$|\Gamma|_{s_0} \leq C_0(s_1)\Theta^{-1}\Upsilon\,, \quad |L|_{s_0} \leq N^{\kappa}\,, \tag{B.2.13}$$

*and,* $\forall s \geq s_0$*,*

$$|\Gamma|_s \leq C(s) N^{\kappa}\big(N^{s-s_0} + N^{-b}|\mathcal{T}|_{s+b}\big), \quad |L|_s \leq C(s) N^{\kappa+s-s_0}, \tag{B.2.14}$$

*such that*

$$Au = h \quad \Longrightarrow \quad u_G + \Gamma u = Lh. \tag{B.2.15}$$

Fix any $k \in G$ (see Definition B.2.3). If $k$ is regular, let $F := \{k\}$, and, if $k$ is not regular but $(A,N)$-regular, let $F \subset E$ such that $\mathrm{d}(k, E\backslash F) \geq N$, $\mathrm{diam}(F) \leq 4N$, $A_F^F$ is $N$-good. We have

$$\begin{aligned} Au = h \quad &\Longrightarrow \quad A_F^F u_F + A_F^{E\backslash F} u_{E\backslash F} = h_F \\ &\Longrightarrow \quad u_F + Q u_{E\backslash F} = (A_F^F)^{-1} h_F \end{aligned} \tag{B.2.16}$$

where

$$Q := (A_F^F)^{-1} A_F^{E\backslash F} = (A_F^F)^{-1} \mathcal{T}_F^{E\backslash F} \in \mathcal{M}_F^{E\backslash F}. \tag{B.2.17}$$

The matrix $Q$ satisfies

$$|Q|_{s_1} \overset{(B.1.16)}{\leq} C(s_1)|(A_F^F)^{-1}|_{s_1}|\mathcal{T}|_{s_1} \overset{(B.2.1),(B.2.9)}{\leq} C(s_1) N^{\tau'+\varsigma s_1}\Upsilon \tag{B.2.18}$$

(the matrix $A_F^F$ is $N$-good). Moreover, $\forall s \geq s_0$, using (B.1.15) and $\mathrm{diam}(F) \leq 4N$,

$$\begin{aligned} |Q|_{s+b} &\lesssim_s |(A_F^F)^{-1}|_{s+b}|\mathcal{T}|_{s_0} + |(A_F^F)^{-1}|_{s_0}|\mathcal{T}|_{s+b} \\ &\overset{(B.1.22)}{\lesssim_s} N^{s+b-s_0}|(A_F^F)^{-1}|_{s_0}|\mathcal{T}|_{s_0} + |(A_F^F)^{-1}|_{s_0}|\mathcal{T}|_{s+b} \\ &\overset{(B.2.1),(B.2.9)}{\lesssim_s} N^{(\varsigma-1)s_0}\big(N^{s+b+\tau'}\Upsilon + N^{\tau'+s_0}|\mathcal{T}|_{s+b}\big). \end{aligned} \tag{B.2.19}$$

Applying the projector $\Pi_{\{k\}}$ in (B.2.16), we obtain

$$Au = h \quad \Longrightarrow \quad u_k + \sum_{k' \in E} \Gamma_k^{k'} u_{k'} = \sum_{k' \in E} L_k^{k'} h_{k'} \tag{B.2.20}$$

that is (B.2.15) with

$$\begin{aligned} \Gamma_k^{k'} &:= \begin{cases} 0 & \text{if } \ k' \in F \\ Q_k^{k'} & \text{if } \ k' \in E \setminus F, \end{cases} \\ L_k^{k'} &:= \begin{cases} [(A_F^F)^{-1}]_k^{k'} & \text{if } \ k' \in F \\ 0 & \text{if } \ k' \in E \setminus F. \end{cases} \end{aligned} \tag{B.2.21}$$

If $k$ is regular then $F = \{k\}$, and, by Definition B.2.2,

$$|A_k^k| \geq \Theta. \tag{B.2.22}$$

Therefore, by (B.2.21) and (B.2.17), the $k$-line of $\Gamma$ satisfies

$$|\Gamma_k|_{s_0+b} \leq |(A_k^k)^{-1}\mathcal{T}_k|_{s_0+b} \overset{(B.2.22),(B.2.9)}{\lesssim_{s_0}} \Theta^{-1}\Upsilon\,. \tag{B.2.23}$$

If $k$ is not regular but $(A,N)$-regular, since $\mathrm{d}(k, E\backslash F) \geq N$ we have, by (B.2.21), that $\Gamma_k^{k'} = 0$ for $|k-k'| \leq N$. Hence, by Lemma B.1.10,

$$\begin{aligned}|\Gamma_k|_{s_0+b} &\overset{(B.1.21)}{\leq} N^{-(s_1-s_0-b)}|\Gamma_k|_{s_1} \overset{(B.2.21)}{\leq} N^{-(s_1-s_0-b)}|Q|_{s_1} \\ &\overset{(B.2.18)}{\lesssim_{s_1}} \Upsilon N^{\tau'+s_0+b-(1-\varsigma)s_1} \\ &\lesssim_{s_1} \Theta^{-1}\Upsilon\end{aligned} \tag{B.2.24}$$

for $N \geq N_0(\Theta)$ large enough. Indeed the exponent $\tau' + s_0 + b - (1-\varsigma)s_1 < 0$ because $s_1$ is large enough according to (B.2.5) and $\varsigma \in (0, 1/2)$ (recall $\kappa := \tau' + s_0 + b$). In both cases (B.2.23)-(B.2.24) imply that each line $\Gamma_k$ decays like

$$|\Gamma_k|_{s_0+b} \lesssim_{s_1} \Theta^{-1}\Upsilon\,, \quad \forall k \in G\,.$$

Hence, by Lemma B.1.11, we get

$$|\Gamma|_{s_0} \leq C'(s_1)\Theta^{-1}\Upsilon\,,$$

which is the first inequality in (B.2.13). Likewise we prove the second estimate in (B.2.13). Moreover, $\forall s \geq s_0$, still by Lemma B.1.11,

$$|\Gamma|_s \lesssim \sup_{k\in G} |\Gamma_k|_{s+b} \overset{\text{(B.2.21)}}{\lesssim} |Q|_{s+b} \overset{\text{(B.2.19)}}{\lesssim_s} N^{\kappa}\big(N^{s-s_0} + N^{-b}|\mathcal{T}|_{s+b}\big)$$

where $\kappa := \tau' + s_0 + b$ and for $N \geq N_0(\Upsilon)$.

The second estimate in (B.2.14) follows by $|L|_{s_0} \leq N^{\kappa}$ (see (B.2.13)) and (B.1.22) (note that by (B.2.21), since $\mathrm{diam} F \leq 4N$, we have $L_k^{k'} = 0$ for all $|k-k'| > 4N$).

**Step II.** By (B.2.15) we have

$$Au = h \quad \Longrightarrow \quad (I_G + \Gamma^G)u_G = Lh - \Gamma^B u_B\,. \tag{B.2.25}$$

By (B.2.13), if $\Theta$ is large enough (depending on $\Upsilon$, namely on the potential $V_0$), we have $|\Gamma^G|_{s_0} \leq 1/2$. Hence, by Lemma B.1.14, $I_G + \Gamma^G$ is invertible and

$$|(I_G + \Gamma^G)^{-1}|_{s_0} \overset{(B.1.26)}{\leq} 2\,, \tag{B.2.26}$$

and, $\forall s \geq s_0$

$$|(I_G + \Gamma^G)^{-1}|_s \overset{(B.1.26)}{\lesssim_s} 1 + |\Gamma^G|_s \overset{(B.2.14)}{\lesssim_s} N^{\kappa}\big(N^{s-s_0} + N^{-b}|\mathcal{T}|_{s+b}\big)\,. \tag{B.2.27}$$

By (B.2.25), we have

$$Au = h \quad \Longrightarrow \quad u_G = \mathcal{M}h + \mathcal{N}u_B$$

with

$$\begin{aligned} \mathcal{M} &:= (I_G + \Gamma^G)^{-1} L \,, \\ \mathcal{N} &:= -(I_G + \Gamma^G)^{-1} \Gamma^B \,, \end{aligned} \tag{B.2.28}$$

and estimates (B.2.10)-(B.2.11) follow by Lemma B.1.7, (B.2.26)-(B.2.27) and (B.2.13)-(B.2.14).

Note that

$$u_G + \Gamma u = Lh \quad \Longleftrightarrow \quad u_G = \mathcal{M}h + \mathcal{N}u_B \,. \tag{B.2.29}$$

As a consequence, if $u_G = \mathcal{M}h + \mathcal{N}u_B$ then, by (B.2.21), for $k$ regular,

$$u_k + (A_k^k)^{-1} \sum_{k' \neq k} A_k^{k'} u_{k'} = (A_k^k)^{-1} h_k \,,$$

hence $(Au)_k = h_k$, proving (B.2.12). ■

**Lemma B.2.6. (Reduction on the bad sites)** *We have*

$$Au = h \quad \Longrightarrow \quad A' u_B = Zh$$

*where*

$$\begin{aligned} A' &:= A^B + A^G \mathcal{N} \ \in \mathcal{M}_E^B \,, \\ Z &:= I_E - A^G \mathcal{M} \ \in \mathcal{M}_E^E \,, \end{aligned} \tag{B.2.30}$$

*satisfy*

$$\begin{aligned} |A'|_{s_0} &\leq c(\Theta) \,, \\ |A'|_s &\leq C(s, \Theta) N^{\kappa} (N^{s - s_0} + N^{-b} |\mathcal{T}|_{s+b}) \,, \end{aligned} \tag{B.2.31}$$

*and*

$$\begin{aligned} |Z|_{s_0} &\leq c N^{\kappa} \,, \\ |Z|_s &\leq C(s, \Theta) N^{2\kappa} (N^{s - s_0} + N^{-b} |\mathcal{T}|_{s+b}) \,. \end{aligned} \tag{B.2.32}$$

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

Proof. By Lemma B.2.5,

$$\begin{aligned} Au = h \quad &\Longrightarrow \quad \begin{cases} A^G u_G + A^B u_B = h \\ u_G = \mathcal{N} u_B + \mathcal{M} h \end{cases} \\ &\Longrightarrow \quad (A^G \mathcal{N} + A^B) u_B = h - A^G \mathcal{M} h \,, \end{aligned}$$

*i.e.* $A' u_B = Zh$. Let us prove estimates (B.2.31)-(B.2.32) for $A'$ and $Z$.

**Step I.** $\forall\, k$ *regular we have* $A'_k = 0$, $Z_k = 0$.
By (B.2.12), for all $k$ regular,

$$\forall h\,,\ \forall u_B \in \mathcal{H}_B\,,\ \Big(A^G(\mathcal{N} u_B + \mathcal{M} h) + A^B u_B\Big)_k = h_k\ , \qquad i.e. \quad (A' u_B)_k = (Zh)_k\,,$$

which implies $A'_k = 0$ and $Z_k = 0$.
**Step II.** *Proof of (B.2.31)-(B.2.32).*
Call $R \subset E$ the regular sites in $E$. For all $k \in E\backslash R$, we have $|A^k_k| < \Theta$ (see Definition B.2.2). Then (B.2.9) implies

$$\begin{aligned} |A_{E\backslash R}|_{s_0} &\le \Theta + |\mathcal{T}|_{s_0} \le c(\Theta)\,, \\ |A_{E\backslash R}|_s &\le \Theta + |\mathcal{T}|_s\,, \quad \forall s \ge s_0\,. \end{aligned} \tag{B.2.33}$$

By Step I and the definition of $A'$ in (B.2.30) we get

$$|A'|_s = |A'_{E\backslash R}|_s \le |A^B_{E\backslash R}|_s + |A^G_{E\backslash R}\mathcal{N}|_s\,.$$

Therefore Lemma B.1.7, (B.2.33), (B.2.10), (B.2.11), imply

$$|A'|_s \le C(s,\Theta) N^\kappa \big(N^{s-s_0} + N^{-b}|\mathcal{T}|_{s+b}\big) \quad \text{and} \quad |A'|_{s_0} \le c(\Theta)\,,$$

proving (B.2.31). The bound (B.2.32) follows similarly.
**Step III.** $(A^{-1})_B$ *is a left inverse of* $A'$.
By

$$A^{-1}A' = A^{-1}(A^B + A^G\mathcal{N}) = I^B_E + I^G_E\mathcal{N}$$

we get

$$(A^{-1})_B A' = (A^{-1}A')_B = I^B_B - 0 = I^B_B$$

proving that $(A^{-1})_B$ is a left inverse of $A'$. ■

Now $A' \in \mathcal{M}^B_E$, and the set $B$ is partitioned in clusters $\Omega_\alpha$ of size $O(N^{C_1})$, far enough one from another, see (H3). Then, up to a remainder of very small $s_0$-norm (see (B.2.37)), $A'$ is defined by the submatrices $(A')^{\Omega_\alpha}_{\Omega'_\alpha}$ where $\Omega'_\alpha$ is some neighborhood of $\Omega_\alpha$ (the distance between two distinct $\Omega'_\alpha$ and $\Omega'_\beta$ remains large). Since $A'$ has a left inverse with $L^2$-norm $O(N'^\tau)$, so have the submatrices $(A')^{\Omega_\alpha}_{\Omega'_\alpha}$. Since these submatrices are of size $O(N^{C_1})$, the $s$-norms of their inverse will be estimated as $O(N^{C_1 s}N'^\tau) = O(N'^{\tau+\chi^{-1}C_1 s})$, see (B.2.43). By Lemma B.1.14, provided $\chi$ is chosen large enough, $A'$ has a left inverse $V$ with $s$-norms satisfying (B.2.34). The details are given in the following lemma.

**Lemma B.2.7. (Left inverse with decay)** *The matrix $A'$ defined in Lemma B.2.6 has a left inverse $V$ which satisfies*

$$\forall s \geq s_0 \ , \quad |V|_s \lesssim_s N^{2\chi\tau+\kappa+2(s_0+b)C_1}\big(N^{C_1 s} + |\mathcal{T}|_{s+b}\big) \, . \tag{B.2.34}$$

Proof. Define $\zeta \in \mathcal{M}_E^B$ by

$$\zeta_{k'}^k := \begin{cases} (A')_{k'}^k & \text{if} \quad (k,k') \in \cup_\alpha(\Omega_\alpha \times \Omega'_\alpha) \\ 0 & \text{if} \quad (k,k') \notin \cup_\alpha(\Omega_\alpha \times \Omega'_\alpha) \end{cases} \tag{B.2.35}$$

where

$$\Omega'_\alpha := \big\{k \in E \ : \ \mathrm{d}(k, \Omega_\alpha) \leq N^2/4\big\} \, . \tag{B.2.36}$$

**Step I.** *$\zeta$ has a left inverse $W \in \mathcal{M}_B^E$ with $\|W\|_0 \leq 2(N')^\tau$.*
We define $\mathcal{R} := A' - \zeta$. By the definition (B.2.35)-(B.2.36), if $\mathrm{d}(k', k) < N^2/4$ then $\mathcal{R}_{k'}^k = 0$ and so

$$\begin{aligned} |\mathcal{R}|_{s_0} &\overset{(\mathrm{B.1.21})}{\leq} 4^{s_1} N^{-2(s_1-b-s_0)}|\mathcal{R}|_{s_1-b} \leq 4^{s_1} N^{-2(s_1-b-s_0)}|A'|_{s_1-b} \\ &\overset{(B.2.31),(B.2.9)}{\lesssim_{s_1}} N^{-2(s_1-b-s_0)} N^\kappa (N^{s_1-b-s_0} + N^{-b}\Upsilon) \\ &\lesssim_{s_1} N^{2\kappa - s_1} \end{aligned} \tag{B.2.37}$$

for $N \geq N_0(\Upsilon)$ large enough. Therefore

$$\begin{aligned} \|\mathcal{R}\|_0 \|(A^{-1})_B\|_0 \overset{(\mathrm{B.1.24})}{\lesssim_{s_0}} |\mathcal{R}|_{s_0} \|A^{-1}\|_0 &\overset{(B.2.37),(H2)}{\lesssim_{s_1}} N^{2\kappa - s_1}(N')^\tau \\ &\overset{(B.2.2)}{=} C(s_1) N^{2\kappa - s_1 + \chi\tau} \\ &\overset{(B.2.5)}{\leq} 1/2 \end{aligned} \tag{B.2.38}$$

for $N \geq N(s_1)$. Since $(A^{-1})_B \in \mathcal{M}_B^E$ is a left inverse of $A'$ (see Lemma B.2.6), Lemma B.1.14 and (B.2.38) imply that $\zeta = A' - R$ has a left inverse $W \in \mathcal{M}_B^E$, and

$$\|W\|_0 \overset{(B.1.30)}{\leq} 2\|(A^{-1})_B\|_0 \leq 2\|A^{-1}\|_0 \overset{(H2)}{\leq} 2(N')^\tau \, . \tag{B.2.39}$$

**Step II.** *$W_0 \in \mathcal{M}_B^E$ defined by*

$$(W_0)_k^{k'} := \begin{cases} W_k^{k'} & \text{if} \quad (k,k') \in \cup_\alpha(\Omega_\alpha \times \Omega'_\alpha) \\ 0 & \text{if} \quad (k,k') \notin \cup_\alpha(\Omega_\alpha \times \Omega'_\alpha) \end{cases} \tag{B.2.40}$$

*is a left inverse of $\zeta$ and $|W_0|_s \leq C(s) N^{(s+b)C_1+\chi\tau}$, $\forall s \geq s_0$.*

Since $W\zeta = I_B$, we prove that $W_0$ is a left inverse of $\zeta$ showing that

$$(W - W_0)\zeta = 0\,. \tag{B.2.41}$$

Let us prove (B.2.41). For $k \in B = \cup_\alpha \Omega_\alpha$, there is $\alpha$ such that $k \in \Omega_\alpha$, and

$$\forall k' \in B\,,\ \big((W - W_0)\zeta\big)_k^{k'} = \sum_{q \notin \Omega'_\alpha} (W - W_0)_k^q \zeta_q^{k'} \tag{B.2.42}$$

since $(W - W_0)_k^q = 0$ if $q \in \Omega'_\alpha$, see the Definition (B.2.40).

Case I: $k' \in \Omega_\alpha$. Then $\zeta_q^{k'} = 0$ in (B.2.42) and so $((W - W_0)\zeta)_k^{k'} = 0$.

Case II: $k' \in \Omega_\beta$ for some $\beta \neq \alpha$. Then, since $\zeta_q^{k'} = 0$ if $q \notin \Omega'_\beta$, we obtain by (B.2.42) that

$$\begin{aligned} ((W - W_0)\zeta)_k^{k'} &= \sum_{q \in \Omega'_\beta} (W - W_0)_k^q \zeta_q^{k'} \stackrel{(B.2.40)}{=} \sum_{q \in \Omega'_\beta} W_k^q \zeta_q^{k'} \\ &\stackrel{(B.2.35)}{=} \sum_{k \in E} W_k^q \zeta_q^{k'} = (W\zeta)_k^{k'} = (I_B)_k^{k'} = 0\,. \end{aligned}$$

Since $\mathrm{diam}(\Omega'_\alpha) \leq 2N^{C_1}$, definition (B.2.40) implies $(W_0)_k^{k'} = 0$ for all $|k - k'| \geq 2N^{C_1}$. Hence, $\forall s \geq 0$,

$$|W_0|_s \stackrel{(B.1.22)}{\lesssim_s} N^{(s+b)C_1} \|W_0\|_0 \stackrel{(B.2.39)}{\lesssim_s} N^{(s+b)C_1 + \chi\tau}\,. \tag{B.2.43}$$

**Step III.** *$A'$ has a left inverse $V$ satisfying (B.2.34).*

Now $A' = \zeta + \mathcal{R}$, $W_0$ is a left inverse of $\zeta$, and

$$|W_0|_{s_0} |\mathcal{R}|_{s_0} \stackrel{(B.2.43),(B.2.37)}{\leq} C(s_1) N^{(s_0+b)C_1 + \chi\tau + 2\kappa - s_1} \stackrel{(B.2.5)}{\leq} 1/2$$

(we use also that $\chi > C_1$ by (B.2.4)) for $N \geq N(s_1)$ large enough. Hence, by Lemma B.1.14, $A'$ has a left inverse $V$ with

$$|V|_{s_0} \stackrel{(B.1.26)}{\leq} 2|W_0|_{s_0} \stackrel{(B.2.43)}{\leq} C N^{(s_0+b)C_1 + \chi\tau} \tag{B.2.44}$$

and, $\forall s \geq s_0$,

$$\begin{aligned} |V|_s &\stackrel{(B.1.26)}{\lesssim_s} |W_0|_s + |W_0|_{s_0}^2 |\mathcal{R}|_s \lesssim_s |W_0|_s + |W_0|_{s_0}^2 |A'|_s \\ &\stackrel{(B.2.43),(B.2.31)}{\lesssim_s} N^{2\chi\tau + \kappa + 2(s_0+b)C_1} \big(N^{C_1 s} + |\mathcal{T}|_{s+b}\big) \end{aligned}$$

proving (B.2.34). ■

Proof of Proposition B.2.4 completed. Lemmata B.2.5, B.2.6, B.2.7 imply

$$Au = h \quad \Longrightarrow \quad \begin{cases} u_G = \mathcal{M}h + \mathcal{N}u_B \\ u_B = VZh \end{cases}$$

whence

$$\begin{aligned} &(A^{-1})_B = VZ \\ &(A^{-1})_G = \mathcal{M} + \mathcal{N}VZ = \mathcal{M} + \mathcal{N}(A^{-1})_B \,. \end{aligned} \tag{B.2.45}$$

Therefore, $\forall s \geq s_0$,

$$\begin{aligned} |(A^{-1})_B|_s &\overset{(B.2.45),(B.1.15)}{\lesssim_s} |V|_s|Z|_{s_0} + |V|_{s_0}|Z|_s \\ &\overset{(B.2.34),(B.2.32),(B.2.9),(B.2.44)}{\lesssim_s} N^{2\kappa+2\chi\tau+2(s_0+b)C_1}\big(N^{C_1 s} + |\mathcal{T}|_{s+b}\big) \\ &\lesssim_s (N')^{\alpha_1}\big((N')^{\alpha_2 s} + |\mathcal{T}|_s\big) \end{aligned}$$

using $|\mathcal{T}|_{s+b} \leq C(s)(N')^b|\mathcal{T}|_s$ (by (B.1.22)) and defining

$$\alpha_1 := \, 2\tau + b + 2\chi^{-1}(\kappa + C_1(s_0+b)) \,, \quad \alpha_2 := \chi^{-1}C_1 \,. \tag{B.2.46}$$

We obtain the same bound for $|(A^{-1})_G|_s$. Notice that by (B.2.4) and (B.2.3), the exponents $\alpha_1, \alpha_2$ in (B.2.46) satisfy

$$\alpha_1 < \tau' \,, \quad \alpha_2 < \varsigma \,. \tag{B.2.47}$$

Hence, for all $s \geq s_0$,

$$\begin{aligned} |A^{-1}|_s \leq |(A^{-1})_B|_s + |(A^{-1})_G|_s &\leq C(s)(N')^{\alpha_1}\big((N')^{\alpha_2 s} + |\mathcal{T}|_s\big) \\ &\overset{(B.2.47),(B.2.9)}{\leq} C(s)(N')^{\tau'}\big((N')^{\varsigma s} + |A - \mathrm{Diag}(A)|_s\big) \end{aligned} \tag{B.2.48}$$

which is (B.2.8). Moreover, for $N \geq N(s_1)$ large enough, we have

$$\forall s \in [s_0, s_1] \,, \quad C(s)(N')^{\alpha_1} \leq \frac{1}{4}(N')^{\tau'} \,,$$

and by (B.2.48) we deduce (B.2.7).

# Appendix C

# Normal form close to an isotropic torus

In this Appendix we report the results in [24], which are used in Chapter 6. Theorem C.1.4 provides, in a neighborhood of an isotropic invariant torus for an Hamiltonian vector field $X_K$, symplectic variables in which the Hamiltonian $K$ assumes the normal form (C.1.22). It is a classical result of Herman [81], [64] that an invariant torus, densily filled by a quasi-periodic solution, is isotropic, see Lemma C.1.2 and Lemma C.2.4 for a more quantitative version. In view of the Nash-Moser iteration we need to perform an analogous construction for an only "approximately invariant" torus. The key step is Lemma C.2.5 which constructs near an "approximately invariant" torus an isotropic torus. This Appendix is written with a self-contained character.

## C.1 Symplectic coordinates near an invariant torus

We consider the toroidal phase space

$$\mathcal{P} := \mathbb{T}^\nu \times \mathbb{R}^\nu \times E \qquad \text{where} \qquad \mathbb{T}^\nu := \mathbb{R}^\nu/(2\pi\mathbb{Z})^\nu$$

is the standard flat torus and $E$ is a real Hilbert space with scalar product $\langle\ ,\ \rangle$. We denote by $u := (\theta, y, z)$ the variables of $\mathcal{P}$. We call $(\theta, y)$ the "action-angle" variables and $z$ the "normal" variables. We assume that $E$ is endowed with a constant exact symplectic 2-form

$$\Omega_E(z, w) = \langle \bar{J} z, w\rangle\,, \quad \forall z, w \in E\,, \tag{C.1.1}$$

where $\bar{J} : E \to E$ is an antisymmetric bounded linear operator with trivial kernel. Thus $\mathcal{P}$ is endowed with the symplectic 2-form

$$\mathcal{W} := (dy \wedge d\theta) \,\oplus\, \Omega_E \tag{C.1.2}$$

which is exact, namely

$$\mathcal{W} = d\varkappa \tag{C.1.3}$$

where $d$ denotes the exterior derivative and $\varkappa$ is the Liouville 1-form on $\mathcal{P}$ defined by

$$\begin{aligned} &\varkappa_{(\theta,y,z)} : \mathbb{R}^\nu \times \mathbb{R}^\nu \times E \to \mathbb{R}\,, \\ &\varkappa_{(\theta,y,z)}[\hat\theta, \hat y, \hat z] := y \cdot \hat\theta + \frac{1}{2}\langle \bar J z, \hat z\rangle\,, \quad \forall (\hat\theta, \hat y, \hat z) \in \mathbb{R}^\nu \times \mathbb{R}^\nu \times E\,, \end{aligned} \tag{C.1.4}$$

and the dot " $\cdot$ " denotes the usual scalar product of $\mathbb{R}^\nu$.

Given a Hamiltonian function $K : \mathcal{D} \subset \mathcal{P} \to \mathbb{R}$, we consider the Hamiltonian system

$$u_t = X_K(u) \qquad \text{where} \qquad dK(u)[\cdot] = -\mathcal{W}(X_K(u), \cdot) \tag{C.1.5}$$

formally defines the Hamiltonian vector field $X_K$. For infinite dimensional systems (i.e. PDEs) the Hamiltonian $K$ is, in general, well defined and smooth only on a dense subset $\mathcal{D} = \mathbb{T}^\nu \times \mathbb{R}^\nu \times E_1 \subset \mathcal{P}$ where $E_1 \subset E$ is a dense subspace of $E$. We require that, for all $(\theta, y) \in \mathbb{T}^\nu \times \mathbb{R}^\nu$, $\forall z \in E_1$, the Hamiltonian $K$ admits a gradient $\nabla_z K$, defined by

$$d_z K(\theta, y, z)[h] = \langle \nabla_z K(\theta, y, z), h\rangle\,, \quad \forall h \in E_1\,, \tag{C.1.6}$$

and that $\nabla_z K(\theta, y, z) \in E$ is in the space of definition of the (possibly unbounded) operator $J := -\bar J^{-1}$. Then by (C.1.5), (C.1.1), (C.1.2), (C.1.6) the Hamiltonian vector field $X_K : \mathbb{T}^\nu \times \mathbb{R}^\nu \times E_1 \mapsto \mathbb{R}^\nu \times \mathbb{R}^\nu \times E$ writes

$$X_K = (\partial_y K, -\partial_\theta K, J\nabla_z K)\,, \qquad J := -\bar J^{-1}\,.$$

A continuous curve $[t_0, t_1] \ni t \mapsto u(t) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times E$ is a solution of the Hamiltonian system (C.1.5) if it is $C^1$ as a map from $[t_0, t_1]$ to $\mathbb{T}^\nu \times \mathbb{R}^\nu \times E_1$ and $u_t(t) = X_K(u(t))$, $\forall t \in [t_0, t_1]$. For PDEs, the flow map $\Phi^t_K$ may not be well-defined everywhere. The next arguments, however, will not require to solve the initial value problem, but only a functional equation in order to find quasi-periodic solutions, see (C.1.11).

We suppose that (C.1.5) possesses an embedded invariant torus

$$\varphi \mapsto i(\varphi) := (\underline{\theta}(\varphi), \underline{y}(\varphi), \underline{z}(\varphi)) \tag{C.1.7}$$

$$i \in C^1(\mathbb{T}^\nu, \mathcal{P}) \cap C^0(\mathbb{T}^\nu, \mathcal{P} \cap \mathbb{T}^\nu \times \mathbb{R}^\nu \times E_1)\,, \tag{C.1.8}$$

which supports a quasi-periodic solution with non-resonant frequency vector $\omega \in \mathbb{R}^\nu$, more precisely

$$i \circ \Psi^t_\omega = \Phi^t_K \circ i\,, \quad \forall t \in \mathbb{R}\,, \tag{C.1.9}$$

where $\Phi^t_K$ denotes the flow generated by $X_K$ and

$$\Psi^t_\omega : \mathbb{T}^\nu \to \mathbb{T}^\nu\,, \quad \Psi^t_\omega(\varphi) := \varphi + \omega t\,, \tag{C.1.10}$$

is the translation flow of vector $\omega$ on $\mathbb{T}^\nu$. Since $\omega \in \mathbb{R}^\nu$ is non-resonant, namely

$$\omega \cdot \ell \neq 0\,, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\,,$$

each orbit of $(\Psi^t_\omega)$ is *dense* in $\mathbb{T}^\nu$. Note that (C.1.9) *only* requires that the flow $\Phi^t_K$ is well defined and smooth on the compact manifold $\mathcal{T} := i(\mathbb{T}^\nu) \subset \mathcal{P}$ and $(\Phi^t_K)_{|\mathcal{T}} = i \circ \Psi^t_\omega \circ i^{-1}$. This remark is important because, for PDEs, the flow could be ill-posed in a neighborhood of $\mathcal{T}$. From a functional point of view (C.1.9) is equivalent to the equation

$$\omega \cdot \partial_\varphi i(\varphi) - X_K(i(\varphi)) = 0\,. \tag{C.1.11}$$

**Remark C.1.1.** *In the sequel we will formally differentiate several times the torus embedding $i$, so that we assume more regularity than* (C.1.8). *In the framework of a Nash-Moser scheme, the approximate torus embedding solutions $i$ are indeed regularized at each step.*

We require that $\underline{\theta} : \mathbb{T}^\nu \to \mathbb{T}^\nu$ is a diffeomorphism of $\mathbb{T}^\nu$ isotopic to the identity. Then the embedded torus $\mathcal{T} := i(\mathbb{T}^\nu)$ is a smooth graph over $\mathbb{T}^\nu$. Moreover the lift on $\mathbb{R}^\nu$ of $\underline{\theta}$ is a map

$$\underline{\theta} : \mathbb{R}^\nu \to \mathbb{R}^\nu, \quad \underline{\theta}(\varphi) = \varphi + \underline{\vartheta}(\varphi)\,, \tag{C.1.12}$$

where $\underline{\vartheta}(\varphi)$ is $2\pi$-periodic in each component $\varphi_i$, $i = 1, \ldots, \nu$, with invertible Jacobian matrix $D\underline{\theta}(\varphi) = \mathrm{Id} + D\underline{\vartheta}(\varphi)$, $\forall \varphi \in \mathbb{T}^\nu$. In the usual applications $D\underline{\vartheta}$ is small and $\omega$ is a Diophantine vector, namely

$$|\omega \cdot \ell| \geq \frac{\gamma}{|\ell|^\tau}\,, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\,.$$

The torus $\mathcal{T}$ is the graph of the function (see (C.1.7) and (C.1.12))

$$j := i \circ \underline{\theta}^{-1}\,, \quad j : \mathbb{T}^\nu \to \mathbb{T}^\nu \times \mathbb{R}^\nu \times E\,, \quad j(\theta) := (\theta, \underline{\tilde{y}}(\theta), \underline{\tilde{z}}(\theta))\,, \tag{C.1.13}$$

namely

$$\mathcal{T} = \big\{ (\theta, \underline{\tilde{y}}(\theta), \underline{\tilde{z}}(\theta))\ ;\ \theta \in \mathbb{T}^\nu \big\}\,, \quad \text{where} \quad \underline{\tilde{y}} := \underline{y} \circ \underline{\theta}^{-1}\,,\ \underline{\tilde{z}} := \underline{z} \circ \underline{\theta}^{-1}\,. \tag{C.1.14}$$

We first prove the isotropy of an invariant torus as in [81], [64], i.e. that the 2-form $\mathcal{W}$ vanishes on the tangent space to $i(\mathbb{T}^\nu) \subset \mathcal{P}$,

$$0 = i^*\mathcal{W} = i^* d\varkappa = d(i^*\varkappa)\,, \tag{C.1.15}$$

or equivalently the 1-form $i^*\varkappa$ on $\mathbb{T}^\nu$ is closed.

**Lemma C.1.2.** *The invariant torus $i(\mathbb{T}^\nu)$ is isotropic.*

PROOF. By (C.1.9) the pullback

$$(i \circ \Psi_\omega^t)^* \mathcal{W} = (\Phi_K^t \circ i)^* \mathcal{W} = i^* \mathcal{W} . \tag{C.1.16}$$

For smooth Hamiltonian systems in finite dimension $(C.1.16)$ is true because the 2-form $\mathcal{W}$ is invariant under the flow map $\Phi_K^t$ (i.e. $(\Phi_K^t)^* \mathcal{W} = \mathcal{W}$). In our setting, the flow $(\Phi_K^t)$ may not be defined everywhere, but $\Phi_K^t$ is well defined on $i(\mathbb{T}^\nu)$ by the assumption (C.1.9), and still preserves $\mathcal{W}$ on the manifold $i(\mathbb{T}^\nu)$, see the proof of Lemma C.2.4 for details. Next, denoting the 2-form

$$(i^* \mathcal{W})(\varphi) = \sum_{i<j} A_{ij}(\varphi) d\varphi_i \wedge d\varphi_j ,$$

we have

$$(i \circ \Psi_\omega^t)^* \mathcal{W} = (\Psi_\omega^t)^* \circ i^* \mathcal{W} = \sum\nolimits_{i<j} A_{ij}(\varphi + \omega t) d\varphi_i \wedge d\varphi_j$$

and so (C.1.16) implies that

$$A_{ij}(\varphi + \omega t) = A_{ij}(\varphi) , \quad \forall t \in \mathbb{R} .$$

Since the orbit $\{\varphi + \omega t\}$ is dense on $\mathbb{T}^\nu$ ($\omega$ is non-resonant) and each function $A_{ij}$ is continuous, it implies that

$$A_{ij}(\varphi) = c_{ij} , \ \forall \varphi \in \mathbb{T}^\nu , \quad i.e. \ \ i^* \mathcal{W} = \sum\nolimits_{i<j} c_{ij} d\varphi_i \wedge d\varphi_j$$

is constant. But, by (C.1.3), the 2-form $i^* \mathcal{W} = i^* d\varkappa = d(i^* \varkappa)$ is also exact. Thus each $c_{ij} = 0$ namely $i^* \mathcal{W} = 0$. ■

We now consider the diffeomorphism of the phase space

$$\begin{pmatrix} \theta \\ y \\ z \end{pmatrix} = G \begin{pmatrix} \phi \\ \zeta \\ w \end{pmatrix} := \begin{pmatrix} \underline{\theta}(\phi) \\ \underline{y}(\phi) + [D\underline{\theta}(\phi)]^{-\top} \zeta - \big[D\underline{\tilde{z}}(\underline{\theta}(\phi))\big]^\top \bar{J} w \\ \underline{z}(\phi) + w \end{pmatrix} \tag{C.1.17}$$

where $\underline{\tilde{z}}(\theta) := (\underline{z} \circ \underline{\theta}^{-1})(\theta)$, see (C.1.14). The transposed operator $\big[D\underline{\tilde{z}}(\theta)\big]^\top : E \to \mathbb{R}^\nu$ is defined by the duality relation

$$\big[D\underline{\tilde{z}}(\theta)\big]^\top w \cdot \hat{\theta} = \langle w, D\underline{\tilde{z}}(\theta)[\hat{\theta}] \rangle , \quad \forall w \in E , \ \hat{\theta} \in \mathbb{R}^\nu .$$

**Lemma C.1.3.** *Let $i$ be an isotropic torus embedding. Then $G$ is symplectic.*

*Proof.* We may see $G$ as the composition $G := G_2 \circ G_1$ of the diffeomorphisms

$$\begin{pmatrix} \theta \\ y \\ z \end{pmatrix} = G_1 \begin{pmatrix} \phi \\ \zeta \\ w \end{pmatrix} := \begin{pmatrix} \underline{\theta}(\phi) \\ [D\underline{\theta}(\phi)]^{-\top} \zeta \\ w \end{pmatrix}$$

and

$$\begin{pmatrix} \theta \\ y \\ z \end{pmatrix} \mapsto G_2 \begin{pmatrix} \theta \\ y \\ z \end{pmatrix} := \begin{pmatrix} \theta \\ \tilde{\underline{y}}(\theta) + y - \big[D\tilde{\underline{z}}(\theta)\big]^\top \bar{J} z \\ \tilde{\underline{z}}(\theta) + z \end{pmatrix} \tag{C.1.18}$$

where $\tilde{\underline{y}} := \underline{y} \circ \underline{\theta}^{-1}$, $\tilde{\underline{z}} := \underline{z} \circ \underline{\theta}^{-1}$, see (C.1.14). We claim that both $G_1$, $G_2$ are symplectic, whence the lemma follows.

$G_1$ IS SYMPLECTIC. Since $G_1$ is the identity in the third component, it is sufficient to check that the map

$$(\phi, \zeta) \mapsto \big(\underline{\theta}(\phi), [D\underline{\theta}(\phi)]^{-\top}\zeta\big)$$

is a symplectic diffeomorphism on $\mathbb{T}^\nu \times \mathbb{R}^\nu$, which is a direct calculus.

$G_2$ IS SYMPLECTIC. We prove that $G_2^* \varkappa - \varkappa$ is closed and so (see (C.1.3))

$$G_2^* \mathcal{W} = G_2^* d\varkappa = dG_2^* \varkappa = d\varkappa = \mathcal{W}\,.$$

By (C.1.18) and the definition of pullback we have

$$\begin{aligned}(G_2^*\varkappa)_{(\theta,y,z)}[\hat\theta, \hat y, \hat z] = {}& \big(\tilde{\underline{y}}(\theta) + y - \big[D\tilde{\underline{z}}(\theta)\big]^\top \bar J z\big) \cdot \hat\theta \\ & + \frac12 \langle \bar J(\tilde{\underline{z}}(\theta) + z), \hat z + D\tilde{\underline{z}}(\theta)[\hat\theta]\rangle\,.\end{aligned}$$

Therefore (recall (C.1.4))

$$\begin{aligned}\big((G_2^*\varkappa)_{(\theta,y,z)} - \varkappa_{(\theta,y,z)}\big)[\hat\theta, \hat y, \hat z] &= \big(\tilde{\underline{y}}(\theta) - \big[D\tilde{\underline{z}}(\theta)\big]^\top \bar J z\big) \cdot \hat\theta + \frac12 \langle \bar J \tilde{\underline{z}}(\theta), \hat z\rangle \\ &\quad + \frac12 \langle \bar J \tilde{\underline{z}}(\theta), D\tilde{\underline{z}}(\theta)[\hat\theta]\rangle + \frac12 \langle \bar J z, D\tilde{\underline{z}}(\theta)[\hat\theta]\rangle \\ &= \tilde{\underline{y}}(\theta) \cdot \hat\theta + \frac12 \langle \bar J \tilde{\underline{z}}(\theta), D\tilde{\underline{z}}(\theta)[\hat\theta]\rangle \\ &\quad + \frac12 \langle \bar J \tilde{\underline{z}}(\theta), \hat z\rangle + \frac12 \langle \bar J D\tilde{\underline{z}}(\theta)[\hat\theta], z\rangle\,,\end{aligned} \tag{C.1.19}$$

having used that $\bar J^\top = -\bar J$. We first note that the 1-form

$$(\hat\theta, \hat y, \hat z) \mapsto \langle \bar J \tilde{\underline{z}}(\theta), \hat z\rangle + \langle \bar J D\tilde{\underline{z}}(\theta)[\hat\theta], z\rangle = d(\langle \bar J \tilde{\underline{z}}(\theta), z\rangle)[\hat\theta, \hat y, \hat z] \tag{C.1.20}$$

is exact. Moreover

$$\tilde{\underline{y}}(\theta) \cdot \hat\theta + \frac12 \langle \bar J \tilde{\underline{z}}(\theta), D\tilde{\underline{z}}(\theta)[\hat\theta]\rangle = (j^*\varkappa)_\theta[\hat\theta] \tag{C.1.21}$$

(recall (C.1.4)) where $j := i \circ \underline{\theta}^{-1}$ see (C.1.13). Hence (C.1.19), (C.1.20), (C.1.21) imply

$$(G_2^*\varkappa)_{(\theta,y,z)} - \varkappa_{(\theta,y,z)} = \pi^*(j^*\varkappa)_{(\theta,y,z)} + d\Big(\frac12 \langle \bar J \tilde{\underline{z}}(\theta), z\rangle\Big)$$

where $\pi : \mathbb{T}^\nu \times \mathbb{R}^\nu \times E \to \mathbb{T}^\nu$ is the canonical projection.

Since the torus $j(\mathbb{T}^\nu) = i(\mathbb{T}^\nu)$ is isotropic the 1-form $j^*\varkappa$ on $\mathbb{T}^\nu$ is closed (as $i^*\varkappa$, see (C.1.15)). This concludes the proof. □

Since $G$ is symplectic the transformed Hamiltonian vector field

$$G^* X_K := (DG)^{-1} \circ X_K \circ G = X_{\mathtt{K}}\,, \quad \mathtt{K} := K \circ G\,,$$

is still Hamiltonian. By construction (see (C.1.17)) the torus $\{\zeta = 0, w = 0\}$ is invariant and (C.1.11) implies

$$X_{\mathtt{K}}(\phi, 0, 0) = (\omega, 0, 0)$$

(see also Lemma C.2.6). As a consequence, the Taylor expansion of the transformed Hamiltonian $\mathtt{K}$ in these new coordinates assumes the normal form

$$\mathtt{K} = const + \omega \cdot \zeta + \frac{1}{2} A(\phi)\zeta \cdot \zeta + \langle C(\phi)\zeta, w\rangle + \frac{1}{2}\langle B(\phi)w, w\rangle + O_3(\zeta, w) \tag{C.1.22}$$

where $A(\phi) \in \mathrm{Mat}(\nu \times \nu)$ is a real symmetric matrix, $B(\phi)$ is a self-adjoint operator of $E$, $C(\phi) \in \mathcal{L}(\mathbb{R}^\nu, E)$, and $O_3(\zeta, w)$ collects all the terms at least cubic in $(\zeta, w)$.

We have proved the following theorem:

**Theorem C.1.4. [24] (Normal form close to an invariant isotropic torus)** *Let $\mathcal{T} = i(\mathbb{T}^\nu)$ be an embedded torus, see* (C.1.7)-(C.1.8)*, which is a smooth graph over $\mathbb{T}^\nu$, see* (C.1.13)-(C.1.14)*, invariant for the Hamiltonian vector field $X_K$, and on which the flow is conjugate to the translation flow of vector $\omega$, see* (C.1.9)-(C.1.10)*. Assume moreover that $\mathcal{T}$ is* ISOTROPIC*, a property which is automatically verified if $\omega$ is non-resonant.*

*Then there exist symplectic coordinates $(\phi, \zeta, w)$ in which $\mathcal{T}$ is described by*

$$\mathbb{T}^\nu \times \{0\} \times \{0\}$$

*and the Hamiltonian assumes the normal form* (C.1.22)*, i.e. the torus*

$$\mathcal{T} = G(\mathbb{T}^\nu \times \{0\} \times \{0\})$$

*where $G$ is the symplectic diffeomorphism defined in* (C.1.17)*, and the Hamiltonian $K \circ G$ has the Taylor expansion* (C.1.22) *in a neighborhood of the invariant torus.*

The normal form (C.1.22) is relevant in view of a Nash-Moser approach, because it provides a control of the linearized equations in the normal bundle of the torus. The linearized Hamiltonian system associated to $\mathtt{K}$ at the trivial solution $(\phi, \zeta, w)(t) = (\omega t, 0, 0)$ is

$$\begin{cases} \dot\phi - A(\omega t)\zeta - [C(\omega t)]^\top w = 0 \\ \dot\zeta = 0 \\ \dot w - J\big(B(\omega t)w + C(\omega t)\zeta\big) = 0 \end{cases}$$

and note that the second equation is decoupled from the others. Inserting its constant solution $\zeta(t) = \zeta_0$ in the third equation we are reduced to solve the quasi-periodically forced Hamiltonian linear equation in $w$,

$$w_t - JB(\omega t)w = g(\omega t)\,, \quad g(\omega t) := JC(\omega t)\zeta_0\,.$$

This linear system may be studied with both the reducibility and the multiscale techniques presented in sections 1.3 and 1.4. In particular, if the reducibility approach outlined in subsection 1.3.1 applies, there is a symplectic change of variable which makes $B(\phi)$ constant.

## C.2 Symplectic coordinates near an approximately invariant torus

In this section we report a construction of suitable symplectic coordinates near a torus which is only approximately invariant, analogous to the one in the previous section.

For that, we first report a basic fact about 1-forms on a torus. We regard a 1-form $a = \sum_{i=1}^{\nu} a_i(\varphi)d\varphi_i$ equivalently as the vector field $\mathbf{a}(\varphi) = (a_1(\varphi), \ldots, a_\nu(\varphi))$.

Given a function $g : \mathbb{T}^\nu \to \mathbb{R}$ with zero average, we denote by

$$u := \Delta^{-1} g$$

the unique solution of $\Delta u = g$ with zero average.

**Lemma C.2.1. (Helmotz decomposition)** *A smooth vector field* $\mathbf{a}$ *on* $\mathbb{T}^\nu$ *may be decomposed as the sum of a conservative and a divergence-free vector field:*

$$\mathbf{a} = \nabla U + \mathbf{c} + \rho\,, \quad U : \mathbb{T}^\nu \to \mathbb{R}\,, \quad \mathbf{c} \in \mathbb{R}^\nu\,, \quad \mathrm{div}\rho = 0\,, \quad \int_{\mathbb{T}^\nu} \rho d\varphi = 0\,. \qquad \text{(C.2.1)}$$

*The above decomposition is unique if we impose that the mean value of* $U$ *vanishes. We have that*

$$U = \Delta^{-1}(\mathrm{div}\,\mathbf{a})\,,$$

*the components of* $\rho$ *are*

$$\rho_j(\varphi) = \Delta^{-1} \sum_{k=1}^{\nu} \partial_{\varphi_k} A_{kj}(\varphi)\,, \quad A_{kj} := \partial_{\varphi_k} a_j - \partial_{\varphi_j} a_k\,, \qquad \text{(C.2.2)}$$

*and*

$$\mathbf{c} = (c_j)_{j=1,\ldots,\nu}\,, \quad c_j = (2\pi)^{-\nu} \int_{\mathbb{T}^\nu} a_j(\varphi)\, d\varphi\,.$$

*Proof.* Notice that $\mathrm{div}(\vec{a} - \nabla U) = 0$ if and only if $\mathrm{div}\,\mathbf{a} = \Delta U$. This equation has the solution $U := \Delta^{-1}(\mathrm{div}\,\mathbf{a})$ (note that $\mathrm{div}\,\mathbf{a}$ has zero average). Hence (C.2.1) is achieved with $\rho := \mathbf{a} - \nabla U - \mathbf{c}$. By taking the $\varphi$-average we get that each

$$c_j = (2\pi)^{-\nu} \int_{\mathbb{T}^\nu} a_j(\varphi)\, d\varphi \, .$$

Let us now prove the expression (C.2.2) of $\rho_j$. We have

$$\partial_{\varphi_k} \rho_j - \partial_{\varphi_j} \rho_k = \partial_{\varphi_k} a_j - \partial_{\varphi_j} a_k =: A_{kj}$$

because

$$\partial_{\varphi_j} \partial_{\varphi_k} U - \partial_{\varphi_k} \partial_{\varphi_j} U = 0 \, .$$

For each $j = 1, \ldots, \nu$ we differentiate $\partial_{\varphi_k} \rho_j - \partial_{\varphi_j} \rho_k = A_{kj}$ with respect to $\varphi_k$ and we sum in $k$, obtaining

$$\Delta \rho_j - \sum_{k=1}^{\nu} \partial_{\varphi_k \varphi_j} \rho_k = \sum_{k=1}^{\nu} \partial_{\varphi_k} A_{kj} \, .$$

Since

$$\sum_{k=1}^{\nu} \partial_{\varphi_k \varphi_j} \rho_k = \partial_{\varphi_j} \mathrm{div} \rho = 0$$

then $\Delta \rho_j = \sum_{k=1}^{\nu} \partial_{\varphi_k} A_{kj}$ and (C.2.2) follows. □

**Corollary C.2.2.** *Any closed 1-form on $\mathbb{T}^\nu$ has the form $a(\varphi) = \mathbf{c} + dU$ for some $\mathbf{c} \in \mathbb{R}^\nu$.*

**Corollary C.2.3.** *Let $a(\varphi)$ be a 1-form on $\mathbb{T}^\nu$, and let $\rho$ be defined by (C.2.2). Then $a - \sum_{j=1}^{\nu} \rho_j(\varphi) d\varphi_j$ is closed.*

We quantify how an embedded torus $i(\mathbb{T}^\nu)$ is approximately invariant for the Hamiltonian vector field $X_K$ in terms of the "error function"

$$Z(\varphi) := \mathcal{F}(i) = (\omega \cdot \partial_\varphi i)(\varphi) - X_K(i(\varphi)) \, . \tag{C.2.3}$$

Consider the pullback 1-form on $\mathbb{T}^\nu$ (see (C.1.4))

$$(i^* \varkappa)(\varphi) = \sum_{k=1}^{\nu} a_k(\varphi) d\varphi_k \tag{C.2.4}$$

where

$$\begin{aligned} a_k(\varphi) &:= \Big[ \big[ D\underline{\theta}(\varphi) \big]^\top \underline{y}(\varphi) + \frac{1}{2} [D\underline{z}(\varphi)]^\top \bar{J} \underline{z}(\varphi) \Big]_k \\ &= \underline{y}(\varphi) \cdot \frac{\partial \underline{\theta}}{\partial \varphi_k}(\varphi) + \frac{1}{2} \langle \bar{J} \underline{z}(\varphi), \frac{\partial \underline{z}}{\partial \varphi_k}(\varphi) \rangle \, , \end{aligned} \tag{C.2.5}$$

and the 2-form (recall (C.1.3))

$$i^*\mathcal{W} = d(i^*\varkappa) = \sum\nolimits_{k<j} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j \,, \qquad A_{kj}(\varphi) = \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi) \,. \tag{C.2.6}$$

We call the coefficients $(A_{kj})$ the "lack of isotropy" of the torus embedding $\varphi \mapsto i(\varphi)$. In Lemma C.2.4 below we quantify their size in terms of the error function $Z$ defined in (C.2.3). We first recall that the Lie derivative of a $k$-form $\beta$ with respect to the vector field $Y$ is

$$L_Y \beta := \frac{d}{dt}\big[(\Phi_Y^t)^* \beta\big]_{|t=0}$$

where $\Phi_Y^t$ denotes the flow generated by $Y$.

Given a function $g(\varphi)$ with zero average, we denote by $u := (\omega \cdot \partial_\varphi)^{-1} g$ the unique solution of $\omega \cdot \partial_\varphi u = g$ with zero average.

**Lemma C.2.4.** *The "lack of isotropy" coefficients $A_{kj}$ satisfy, $\forall \varphi \in \mathbb{T}^\nu$,*

$$(\omega \cdot \partial_\varphi) A_{kj}(\varphi) = \mathcal{W}\big(DZ(\varphi)\underline{e}_k, Di(\varphi)\underline{e}_j\big) + \mathcal{W}\big(Di(\varphi)\underline{e}_k, DZ(\varphi)\underline{e}_j\big) \tag{C.2.7}$$

*where $(\underline{e}_1, \ldots, \underline{e}_\nu)$ denotes the canonical basis of $\mathbb{R}^\nu$. Thus, since each $A_{kj}$ has zero mean value, if the frequency vector $\omega \in \mathbb{R}^\nu$ is non-resonant, then*

$$A_{kj}(\varphi) = (\omega \cdot \partial_\varphi)^{-1}\Big(\mathcal{W}\big(DZ(\varphi)\underline{e}_k, Di(\varphi)\underline{e}_j\big) + \mathcal{W}\big(Di(\varphi)\underline{e}_k, DZ(\varphi)\underline{e}_j\big)\Big) \,. \tag{C.2.8}$$

*Proof.* We use Cartan's formula

$$L_\omega(i^*\mathcal{W}) = d\big((i^*\mathcal{W})(\omega, \cdot)\big) + \big(d(i^*\mathcal{W})\big)(\omega, \cdot) \,.$$

Since $d(i^*\mathcal{W}) = i^* d\mathcal{W} = 0$ by (C.1.3) we get

$$L_\omega(i^*\mathcal{W}) = d\big((i^*\mathcal{W})(\omega, \cdot)\big) \,. \tag{C.2.9}$$

Now we compute, for $\hat{\phi} \in \mathbb{R}^\nu$

$$\begin{aligned}(i^*\mathcal{W})(\omega, \hat{\phi}) &= \mathcal{W}(Di(\varphi)\omega, Di(\varphi)\hat{\phi}) = \mathcal{W}(X_K(i(\varphi)) + Z(\varphi), Di(\varphi)\hat{\phi}) \\ &= -dK(i(\varphi))[Di(\varphi)\hat{\phi}] + \mathcal{W}(Z(\varphi), Di(\varphi)\hat{\phi}) \,.\end{aligned}$$

We obtain

$$(i^*\mathcal{W})(\omega, \cdot) = \sum_{j=1}^{\nu} b_j(\varphi) d\varphi_j$$

$$b_j(\varphi) = (i^*\mathcal{W})(\omega, \underline{e}_j) = -\frac{\partial(K \circ i)}{\partial \varphi_j}(\varphi) + \mathcal{W}(Z(\varphi), Di(\varphi)\underline{e}_j) \,.$$

Hence, by (C.2.9), the Lie derivative

$$L_\omega(i^*\mathcal{W}) = \sum_{k<j} B_{kj}(\varphi) d\varphi_k \wedge d\varphi_j \tag{C.2.10}$$

with

$$\begin{aligned} B_{kj}(\varphi) &= \frac{\partial b_j}{\partial \varphi_k}(\varphi) - \frac{\partial b_k}{\partial \varphi_j}(\varphi) \\ &= \frac{\partial}{\partial \varphi_k}(\mathcal{W}(Z(\varphi), Di(\varphi)\underline{e}_j)) - \frac{\partial}{\partial \varphi_j}(\mathcal{W}(Z(\varphi), Di(\varphi)\underline{e}_k)) \\ &= \mathcal{W}(DZ(\varphi)\underline{e}_k, Di(\varphi)\underline{e}_j) + \mathcal{W}(Di(\varphi)\underline{e}_k, DZ(\varphi)\underline{e}_j)\,. \end{aligned} \tag{C.2.11}$$

Recalling (C.1.10) and (C.2.6) we have, $\forall \varphi \in \mathbb{T}^\nu$,

$$(\psi^t_\omega)^*(i^*\mathcal{W})(\varphi) = i^*\mathcal{W}(\varphi + \omega t) = \sum_{k<j} A_{kj}(\varphi + \omega t) d\varphi_k \wedge d\varphi_j\,.$$

Hence the Lie derivative

$$L_\omega(i^*\mathcal{W})(\varphi) = \sum_{k<j} (\omega \cdot \partial_\varphi A_{kj})(\varphi) d\varphi_k \wedge d\varphi_j\,. \tag{C.2.12}$$

Comparing (C.2.10)-(C.2.11) and (C.2.12) we deduce (C.2.7). □

The previous lemma provides another proof of Lemma C.1.2. For an invariant torus embedding $i(\varphi)$ the "error function" $Z(\varphi) = 0$ (see (C.2.3)) and so each $A_{kj} = 0$. We now prove that near an approximate isotropic torus there is an isotropic torus.

**Lemma C.2.5. (Isotropic torus)** *The torus embedding* $i_\delta(\varphi) = (\underline{\theta}(\varphi), y_\delta(\varphi), \underline{z}(\varphi))$ *defined by*

$$y_\delta(\varphi) = \underline{y}(\varphi) - [D\underline{\theta}(\varphi)]^{-\top} \rho(\varphi)\,, \quad \rho_j := \Delta^{-1}\Big( \sum_{k=1}^{\nu} \partial_{\varphi_j} A_{kj}(\varphi) \Big)\,, \tag{C.2.13}$$

*is isotropic.*

*Proof.* By Corollary C.2.3 the 1-form $i^*\varkappa - \rho$ is closed with $\rho_j$ defined in (C.2.13), see also (C.2.2), (C.2.4). Actually

$$i^*\varkappa - \rho = i^*_\delta \varkappa$$

is the pullback of the 1-form $\varkappa$ under the modified torus embedding $i_\delta$ defined in (C.2.13), see (C.2.5). Thus the torus $i_\delta(\mathbb{T}^\nu)$ is isotropic. □

In analogy with Theorem C.1.4 we now introduce a symplectic set of coordinates $(\phi, \zeta, w)$ near the isotropic torus $\mathcal{T}_\delta := i_\delta(\mathbb{T}^\nu)$ via the symplectic diffeomorphism

$$\begin{pmatrix} \theta \\ y \\ z \end{pmatrix} = G_\delta \begin{pmatrix} \phi \\ \zeta \\ w \end{pmatrix} := \begin{pmatrix} \underline{\theta}(\phi) \\ y_\delta(\phi) + [D\underline{\theta}(\phi)]^{-\top}\zeta - \big[D\tilde{\underline{z}}(\underline{\theta}(\phi))\big]^\top \bar{J} w \\ \underline{z}(\phi) + w \end{pmatrix} \tag{C.2.14}$$

where $\tilde{\underline{z}} := \underline{z} \circ \underline{\theta}^{-1}$. The map $G_\delta$ is symplectic by Lemma C.1.3 because $i_\delta$ is isotropic (Lemma C.2.5). In the new coordinates $(\phi, \zeta, w)$ the isotropic torus embedding $i_\delta$ is trivial, namely

$$i_\delta(\phi) = G_\delta(\phi, 0, 0)\,.$$

Under the symplectic change of variable (C.2.14), the Hamiltonian vector field $X_K$ changes into

$$X_{\mathtt{K}} = G_\delta^* X_K = (DG_\delta)^{-1} X_K \circ G_\delta \qquad \text{where} \qquad \mathtt{K} := K \circ G_\delta\,. \tag{C.2.15}$$

The Taylor expansion of the new Hamiltonian $\mathtt{K} : \mathbb{R}^\nu \times \mathbb{R}^\nu \times E \to \mathbb{R}$ at the trivial torus $(\phi, 0, 0)$ is

$$\begin{aligned} \mathtt{K} &= \mathtt{K}_{00}(\phi) + \mathtt{K}_{10}(\phi)\cdot\zeta + \langle \mathtt{K}_{01}(\phi), w\rangle \\ &\quad + \frac{1}{2}\mathtt{K}_{20}(\phi)\zeta\cdot\zeta + \langle \mathtt{K}_{11}(\phi)\zeta, w\rangle + \frac{1}{2}\langle \mathtt{K}_{02}(\phi) w, w\rangle + \mathtt{K}_{\geq 3}(\phi, \zeta, w) \end{aligned} \tag{C.2.16}$$

where $\mathtt{K}_{\geq 3}$ collects all the terms at least cubic in the variables $(\zeta, w)$. The Taylor coefficients of $\mathtt{K}$ are $\mathtt{K}_{00}(\phi) \in \mathbb{R}$, $\mathtt{K}_{10}(\phi) \in \mathbb{R}^\nu$, $\mathtt{K}_{01}(\phi) \in E$, $\mathtt{K}_{20}(\phi) \in \mathrm{Mat}(\nu \times \nu)$ is a real symmetric matrix, $\mathtt{K}_{02}(\phi)$ is a self-adjoint operator of $E$ and $\mathtt{K}_{11}(\phi) \in \mathcal{L}(\mathbb{R}^\nu, E)$.

As seen in Theorem C.1.4, if $i_\delta$ were an invariant torus embedding, the coefficient $\mathtt{K}_{00}(\phi) = \text{const}$, $\mathtt{K}_{10}(\phi) = \omega$ and $\mathtt{K}_{01}(\phi) = 0$. We now express these coefficients in terms of the error function $Z_\delta := \mathcal{F}(i_\delta)$.

**Lemma C.2.6.** *The vector field*

$$X_{\mathtt{K}}(\phi, 0, 0) = \begin{pmatrix} \mathtt{K}_{10}(\phi) \\ -\partial_\phi \mathtt{K}_{00}(\phi) \\ J\mathtt{K}_{01}(\phi) \end{pmatrix} = \begin{pmatrix} \omega \\ 0 \\ 0 \end{pmatrix} - \big(DG_\delta(\phi, 0, 0)\big)^{-1} Z_\delta(\phi)\,. \tag{C.2.17}$$

*Proof.* By (C.2.15) and $i_\delta(\phi) = G_\delta(\phi, 0, 0)$, we have

$$\begin{aligned} X_{\mathtt{K}}(\phi, 0, 0) &= DG_\delta(\phi, 0, 0)^{-1} X_K(i_\delta(\phi)) \\ &= DG_\delta(\phi, 0, 0)^{-1}\big(\omega \cdot \partial_\varphi i_\delta(\phi) - Z_\delta(\phi)\big) \end{aligned}$$

and (C.2.17) follows because $DG_\delta(\phi, 0, 0)^{-1} Di_\delta(\phi)[\omega] = (\omega, 0, 0)$. □

We finally write the expression of the coefficients $\mathtt{K}_{11}(\phi)$, $\mathtt{K}_{20}(\phi)$ in terms of $K$, which is used in Chapter 6.

**Lemma C.2.7.** *The coefficients*

$$\mathtt{K}_{11}(\phi) = D_y \nabla_z K(i_\delta(\phi))[D\theta_0(\phi)]^{-\top} + \bar{J}(D_\theta \tilde{z}_0)(\theta_0(\phi))(D_y^2 K)(i_\delta(\phi))[D\theta_0(\phi)]^{-\top} \tag{C.2.18}$$

$$\mathtt{K}_{20}(\phi) = [D\theta_0(\phi)]^{-1}(D_y^2 K)(i_\delta(\phi)))[D\theta_0(\phi)]^{-\top} \,. \tag{C.2.19}$$

*Proof.* Formulas (C.2.18)(C.2.19) follow differentiating $\mathtt{K} = K \circ G_\delta$. □

# Bibliography

[1] Alazard T., Baldi P., *Gravity capillary standing water waves*, Arch. Ration. Mech. Anal., 217, no 3, 741-830, 2015.

[2] Arnold V.I., *Proof of a theorem of A. N. Kolmogorov on the persistence of quasi-periodic motions under small perturbations of the Hamiltonian*, Russian Math. Surveys, 18, 9-36, 1963.

[3] Arnold, V. I., *Instability of dynamical systems with many degrees of freedom.* Dokl. Akad. Nauk SSSR 156, 9-12, 1964.

[4] Arnold, V. I., *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics, 60. Springer-Verlag, xvi+508, ISBN: 0-387-96890-3, 1989.

[5] Baldi P., Berti M., Haus E., Montalto R., *Time quasi-periodic gravity water waves in finite depth*, Inventiones Math. 214 (2), 739-911, 2018.

[6] Baldi P., Berti M., Montalto R., *KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation.* Math. Annalen, 359, 1-2, 471-536, 2014.

[7] Baldi P., Berti M., Montalto R., *KAM for autonomous quasi-linear perturbations of mKdV*, Bollettino Unione Matematica Italiana, 9, 143-188, 2016.

[8] Baldi P., Berti M., Montalto R. *KAM for autonomous quasi-linear perturbations of KdV.* Ann. I. H. Poincaré (C) Anal. Non Linéaire, AN 33, 1589-1638, 2016.

[9] Bambusi D., *Reducibility of 1-d Schrödinger equation with time quasiperiodic unbounded perturbations. I*, Trans. Amer. Math. Soc. 370, no. 3, 1823-1865, 2018.

[10] Bambusi D., *Reducibility of 1-d Schrödinger equation with time quasiperiodic unbounded perturbations, II.* Comm. Math. Phys. 353, no. 1, 353–378, 2017.

[11] Bambusi D., Berti M., Magistrelli E., *Degenerate KAM theory for partial differential equations*, Journal Diff. Equations, 250, 8, 3379-3397, 2011.

[12] 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, no. 11, 1665-1690, 2007.

[13] Bambusi D., Montalto R., *Reducibility of 1-d Schrödinger equation with unbounded time quasi-periodic perturbations. III*, J. Math. Phys. 59, no. 12, 2018.

[14] Berti M., *Nonlinear Oscillations of Hamiltonian PDEs*, Progr. Nonlinear Differential Equations Appl. 74, Birkhäuser, Boston, 1-181, 2008.

[15] Berti M., *KAM Theory for Partial Differential Equations*, Analysis in Theory and Applications Anal. Theory Appl., Vol. 35, No. 3, pp. 235-267, 2019.

[16] Berti M., Biasco L., *Branching of Cantor manifolds of elliptic tori and applications to PDEs*, Comm. Math. Phys, 305, 3, 741-796, 2011.

[17] Berti M., Biasco L., Procesi M., *KAM theory for the Hamiltonian derivative wave equation.* Ann. Sci. Éc. Norm. Supér. (4), 46(2):301–373, 2013.

[18] Berti M., Biasco L., Procesi M., *KAM for Reversible Derivative Wave Equations.* Arch. Ration. Mech. Anal., 212(3):905–955, 2014.

[19] Berti M., Bolle P., *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Mathematical Journal, 134, issue 2, 359-419, 2006.

[20] Berti M., Bolle P., *Cantor families of periodic solutions of wave equations with $C^k$ nonlinearities*, NoDEA Nonlinear Differential Equations Appl., 15, 247-276, 2008.

[21] Berti M., Bolle P., *Sobolev Periodic Solutions of Nonlinear Wave Equations in Higher Spatial Dimensions*, Arch. Rational Mech. Anal. 195, 609-642, 2010.

[22] Berti M., Bolle P., *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity, 25, 2579-2613, 2012.

[23] Berti M., Bolle P., *Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential*, J. Eur. Math. Soc., Vol. 15, 229-286, 2013.

[24] Berti M., Bolle P., *A Nash-Moser approach to KAM theory*, Fields Institute Communications, 255-284, special volume "Hamiltonian PDEs and Applications", 255-284, 2015.

[25] Berti M., Bolle P., Procesi M., *An abstract Nash-Moser theorem with parameters and applications to PDEs*, Annales de l'Institute H. Poincaré, Analyse nonlinaire, 27, 377-399, 2010.

[26] Berti M., Corsi L., Procesi M., *An abstract Nash-Moser Theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds*, Comm. Math. Phys. 334, no. 3, 1413-1454, 2015.

[27] Berti M., Delort J-M., *Almost global existence for capillarity-gravity water waves equations on the circle*, UMI book prize 2017, UMI Lecture Notes, 24, Springer, X, 269, 2018.

[28] Berti M., Kappeler T., Montalto R., *Large KAM tori for perturbations of the dNLS equation*, Astérisque, 403, viii + 148, 2018.

[29] Berti M., Maspero A., *Long time dynamics of Schrödinger and wave equations on flat tori*, Journal Diff. eq., 267, 2, 5, 1167-1200, 2019.

[30] Berti M., Montalto R., *Quasi-periodic standing wave solutions of gravity-capillary water waves* preprint https://arxiv.org/abs/1602.02411, to appear on Memoires AMS, MEMO 891.

[31] Berti M., Procesi M., *Nonlinear wave and Schrödinger equations on compact Lie groups and homogeneous spaces*, Duke Math. J., 159, 3, 479-538, 2011.

[32] Bobenko A., Kuksin S., *The nonlinear Klein-Gordon equation on an interval as a perturbed sine- Gordon equation*, Comment. Math. Helv. 70, no. 1, 63-112, 1995.

[33] Bolle P., Khayamian C., *Quasi-periodic solutions for wave equations on Zoll manifolds*, preprint 2018.

[34] 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.

[35] Bourgain J., *Construction of periodic solutions of nonlinear wave equations in higher dimension*, GAFA, v.5, n.4, 629-639, 1995.

[36] Bourgain J., *On Melnikov's persistency problem*, Internat. Math. Res. Letters, 4, 445 - 458, 1997.

[37] Bourgain J., *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Annals of Math. 148, 363-439, 1998.

[38] Bourgain J., *Periodic solutions of nonlinear wave equations*, Harmonic analysis and partial differential equations, 69–97, Chicago Lectures in Math., Univ. Chicago Press, 1999.

[39] Bourgain J., *Estimates on Green's functions, localization and the quantum kicked rotor model*, Annals of Math., 156, 1, 249-294, 2002.

[40] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton, 2005.

[41] Bourgain J., Goldstein M., Schlag W., *Anderson localization for Schrödinger operators on* $\mathbb{Z}^2$ *with quasi-periodic potential*, Acta Math., 188, 41-86, 2002.

[42] Bourgain J., Wang W.M., *Anderson localization for time quasi-periodic random Schrödinger and wave equations*, Comm. Math. Phys. 248, 429 - 466, 2004.

[43] Bourgain J., Wang W.M., *Quasi-periodic solutions for nonlinear random Schrödinger*, J. European Math. Society, 10, 1 - 45, 2008.

[44] Broer H.; Huitema G., Sevryuk M., *Quasi-periodic motions in families of dynamical systems* Order amidst chaos. Lecture Notes in Mathematics, 1645. Springer-Verlag, xii+196 pp. ISBN: 3-540-62025-7, 1996.

[45] Chierchia L., You J., *KAM tori for 1D nonlinear wave equations with periodic boundary conditions*, Comm. Math. Phys. 211, 497-525, 2000.

[46] Colliander J., Keel M., Staffilani G., Takaoka H., Tao T., *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, Invent. Math. 181, no. 1, 39–113, 2010.

[47] Corsi L., Feola R., Procesi M., *Finite dimensional invariant KAM tori for tame vector fields*, preprint, to appear on Transactions AMS.

[48] Corsi L., Montalto R., *Quasi-periodic solutions for the forced Kirchoff equation on* $\mathbb{T}^d$, Nonlinearity 31, no. 11, 5075-5109, 2018.

[49] Craig W., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses, 9, Société Mathématique de France, Paris, 2000.

[50] Craig W., Sulem C., *Numerical simulation of gravity waves.* J. Comput. Phys., 108(1):73-83, 1993.

[51] Craig W., Wayne C. E., *Newton's method and periodic solutions of nonlinear wave equation*, Comm. Pure Appl. Math. 46, 1409-1498, 1993.

[52] Delort J.M., *Periodic solutions of nonlinear Schrödinger equations: a para-differential approach*, Analysis and PDEs, 4, no. 5, 639-676, 2011.

[53] Eliasson H., *Perturbations of stable invariant tori for hamiltonian systems*, Annali della Scuola Normale Superiore di Pisa, v.15, 1, 115-147, 1988.

[54] Eliasson H., *Almost Reducibility for the Quasi-Periodic Linear Wave Equation.* The conference "In Memory of Jean-Christophe Yoccoz", 2017. `https://www.college-de-france.fr/site/en-pierre-louis-lions/symposium-2017-05-29-10h00.htm`

[55] Eliasson H., Grébert B., Kuksin S., *Almost reducibility of the linear wave equation*, in preparation.

[56] Eliasson H., Grébert B., Kuksin S., *KAM for the nonlinear beam equation*, Geom. Funct. Anal., 26, 1588-1715, 2016.

[57] Eliasson L. H., Kuksin S., *On reducibility of Schrödinger equations with quasiperiodic in time potentials*, Comm. Math. Phys, 286, 125-135, 2009.

[58] Eliasson H., Kuksin S., *KAM for non-linear Schrödinger equation*, Annals of Math., 172, 371-435, 2010.

[59] Ekeland I., *Convexity Methods in Hamiltonian Mechanics*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) 19. Springer-Verlag, Berlin, x+247, 1990.

[60] Ekeland I., Seré E., *A surjection theorem for singular perturbations with loss of derivatives*, arXiv:1811.07568v2, 2019.

[61] Feldman J., Knörrer H., Trubowitz E., *Perturbatively unstable eigenvalues of a periodic Schrödinger operator*, Comment. Math. Helv., no. 4, 557-579, 1991.

[62] Feola R., Procesi M., *Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations.* J. Diff. Eq., 259, no. 7, 3389-3447, 2015.

[63] Feola R., Procesi M., *KAM for quasi-linear autonomous NLS*, preprint 2018, arXiv:1705.07287.

[64] Fejoz J., *Démonstration du théorème d' Arnold sur la stabilité du système planétaire (d' après M. Herman)*, Ergod. Th. Dynam. Sys., 24, 162, 2004.

[65] Fontich E., De la Llave R., Sire Y., *Construction of invariant whiskered tori by a parameterization method. Part II: Quasi-periodic and almost periodic breathers in coupled map lattices.* J. Differential Equations 259, no. 6, 2180-2279, 2015.

[66] Fröhlich J., Spencer T., *Absence of diffusion in the Anderson tight binding model for large disorder or low energy*, Comm. Math. Phys. 88, no. 2, 151-184, 1983.

[67] Geng J., You J., *A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions*, J. Differential Equations 209, no. 1, 1-56, 2005.

[68] Geng J., You J., *KAM tori for higher dimensional beam equations with constant potentials*, Nonlinearity 19, no. 10, 2405–2423, 2006.

[69] Geng J., You J., *A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces*, Comm. Math. Phys. 262, 343-372, 2006.

[70] 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.

[71] Gentile G., Procesi M., *Periodic solutions for a class of nonlinear partial differential equations in higher dimension*, Comm. Math. Phys., 3, 863-906, 2009.

[72] Giuliani F. *Quasi-periodic solutions for quasi-linear generalized KdV equations*, J. Differential Equations 262, no. 10, 5052–5132, 2017.

[73] Graff S., *On the conservation of hyperbolic invariant tori for Hamiltonian systems*, J. Differential Equations 15, 1-69, 1974.

[74] Grébert B., Kappeler T., *The defocusing NLS equation and its normal form*, Springer, 2014.

[75] Grébert B., Paturel E., *KAM for the Klein-Gordon equation on $S^d$*, Bollettino dell'Unione Matematica Italiana (BUMI), 9(2), 237-288, 2016.

[76] Grébert B., Paturel E., *On reducibility of quantum harmonic oscillators on $\mathbb{R}^d$ with a quasi-periodic in time potential* , preprint.

[77] Grébert B., Thomann L., *KAM for the quantum harmonic oscillator*, Comm. Math. Phys. 307, 2, 383-427, 2011.

[78] Guardia M., Haus E., Procesi M., *Growth of Sobolev norms for the analytic NLS on $\mathbb{T}^2$*, Adv. Math. 301, 615–692, 2016.

[79] Guardia M., Kaloshin V., *Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation*, J. Eur. Math. Soc. (JEMS) 17, no. 1, 71–149, 2015.

[80] Hamilton, R.S., *The inverse function theorem of Nash and Moser*, Bull. A.M.S., 7, 65-222, 1982.

[81] Herman, M. *Inégalités "a priori" pour des tores lagrangiens invariants par des difféomorphismes symplectiques*, Publ. Math. Inst. Hautes Étud. Sci. **70**, 47–101, 1989.

[82] Herman M., *Non existence of Lagrangian graphs*, available online in Archive Michel Herman: www.college-de-france.fr .

[83] Ito, H., *Convergence of Birkhoff normal forms for integrable systems*, Comment. Math. Helv. 64, no. 3, 412-46, 1989.

[84] Iooss G., Plotnikov P., Toland J., *Standing waves on an infinitely deep perfect fluid under gravity.* Arch. Ration. Mech. Anal., 177(3):367–478, 2005.

[85] Iooss G., Plotnikov P., *Small divisor problem in the theory of three-dimensional water gravity waves.* Mem. Amer. Math. Soc., 200(940):viii+128, 2009.

[86] Liu J., Yuan X., *A KAM theorem for Hamiltonian partial differential equations with unbounded perturbations.* Comm. Math. Phys., 307(3), 629–673, 2011.

[87] Lojasiewicz S., Zehnder E., *An inverse function theorem in Fréchet-spaces*, J. Funct. Anal. 33, 165-174, 1979.

[88] Kappeler T., Liang Z., *A KAM theorem for the defocusing NLS equation*, J. Diff. Equ. 252, no. 6, 4068 - 4113, 2012.

[89] Kappeler T., Kuksin S., *Strong non-resonance of Schrödinger operators and an averaging theorem*, Physica D, 86, 349-362, 1995.

[90] Kappeler T., Pöschel J., *KAM and KdV*, Springer, 2003.

[91] Kappeler, T., Topalov, P., *Global well-posedness of mKdV in* $L^2(\mathbb{T}, \mathbb{R})$, Comm. Partial Differential Equations 30, no. 1-3, 435-449, 2005.

[92] Kolmogorov A. N., *Preservation of conditionally perodic motions for a small change in Hamilton's function* (Russian) Dokl. Akad. Nauk SSSR 98:4 (1954) 527-530.

[93] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional. Anal. i Prilozhen. 21, no. 3, 22-37, 95, 1987.

[94] Kuksin S., *A KAM theorem for equations of the Korteweg-de Vries type*, Rev. Math. Phys., 10, 3, 1-64, 1998.

[95] Kuksin S., *Analysis of Hamiltonian PDEs*, Oxford University Press, Oxford, 2000.

[96] Kuksin S., Pöschel J., *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*, Ann. of Math. 143, 149 - 179, 1996.

[97] Moser J., *On invariant curves of area-preserving mappings of an annulus*, Nachr. Akad. Wiss. Gott. Math.-Phys. Kl. II, 1-20, 1962.

[98] Moser J., *A rapidly convergent iteration method and non-linear partial differential equations I-II*, Ann. Scuola Norm. Sup. Pisa (3) 20, 265-315, 499-535, 1966.

[99] Moser J., *Convergent series expansions for quasi-periodic motions*, Math. Ann. 169, 136-176, 1967.

[100] Nirenberg L., *Topics in Nonlinear Functional Analysis*, Courant Institute of Mathematical Sciences, New York University.

[101] Plotnikov P., Toland J., *Nash-Moser theory for standing water waves.* Arch. Ration. Mech. Anal., 159(1):1–83, 2001.

[102] Pöschel J., *On Elliptic Lower Dimensional Tori in Hamiltonian Systems*, Math. Z. 202, 559-608, 1989.

[103] Pöschel J., *A KAM-Theorem for some nonlinear PDEs*, Ann. Sc. Norm. Pisa, 23, 119-148, 1996.

[104] Pöschel J., *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv., 71, no. 2, 269-296, 1996.

[105] Procesi M., Procesi C., *A normal form for the Schrödinger equation with analytic non-linearities*, Comm. Math. Phys. 312, 501-557, 2012.

[106] Procesi C., Procesi M., *A KAM algorithm for the completely resonant nonlinear Schrödinger equation*, Advances in Mathematics, volume 272, 399-470, 2015.

[107] Procesi C., Procesi M., *Reducible quasi-periodic solutions for the Non Linear Schrödinger equation*, Boll. Unione Mat. Ital. 9, no. 2, 189–236, 2016.

[108] Procesi M., Xu X., *Quasi-Töplitz Functions in KAM Theorem*, SIAM Math. Analysis, Vol. 45, No. 4, 2148-2181, 2013.

[109] Rüssmann, H., *Über das Verhalten analytischer Hamiltonscher Differentialgleichungen in der Nähe einer Gleichgewichtslsung*, Math. Ann. 154, 285-300, 1964.

[110] Sevryuk M.B., *Reversible Systems*, Lecture Notes in Math, 1211, Springer-Verlag, 1986.

[111] Vey, J., *Sur certains systèmes dynamiques séparables*, Amer. J. Math. 100, no. 3, 591-614, 1978.

[112] Wang W. M., *Eigenfunction Localization for the 2D Periodic Schrodinger Operator*, International Mathematics Research Notices, n. 8, 1804-1838, 2011.

[113] Wang W. M., *Energy supercritical nonlinear Schrödinger equations: quasi-periodic solutions*, Duke Math. J. 165, no. 6, 1129-1192, 2016.

[114] Wang W. M., *Quasi-periodic solutions for nonlinear Klein-Gordon equations*, arxiv.org/abs/1609.00309, 2017.

[115] Wayne E., *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. 127, 479-528, 1990.

[116] Zakharov V., *Stability of periodic waves of finite amplitude on the surface of a deep fluid* Journal of Applied Mechanics and Technical Physics, 9(2):190–194, 1968.

[117] Zehnder E., *Generalized implicit function theorems with applications to some small divisors problems I*, Comm. Pure Appl. Math., 28, 91-140, 1975.

[118] Zehnder E., *Generalized implicit function theorems with applications to some small divisors problems II*, Comm. Pure Appl. Math., 29, 49-111, 1976.

[119] Zhang J., Gao M., Yuan X., *KAM tori for reversible partial differential equations*, Nonlinearity 24, 1189-1228, 2011.

# Index