

# QUASI-PERIODIC TRAVELING WAVES ON AN INFINITELY DEEP PERFECT FLUID UNDER GRAVITY

ROBERTO FEOLA AND FILIPPO GIULIANI

**ABSTRACT.** We consider the gravity water waves system with a periodic one-dimensional interface in infinite depth and we establish the existence and the linear stability of small amplitude, quasi-periodic in time, traveling waves. This provides the first existence result of quasi-periodic water waves solutions bifurcating from a *completely resonant* elliptic fixed point. The proof is based on a Nash-Moser scheme, Birkhoff normal form methods and pseudo differential calculus techniques. We deal with the combined problems of *small divisors* and the *fully-nonlinear* nature of the equations.

The lack of parameters, like the capillarity or the depth of the ocean, demands a refined *nonlinear* bifurcation analysis involving several non-trivial resonant wave interactions, as the well-known “Benjamin-Feir resonances”. We develop a novel normal form approach to deal with that. Moreover, by making full use of the Hamiltonian structure, we are able to provide the existence of a wide class of solutions which are free from restrictions of parity in the time and space variables.

## CONTENTS

<b>1</b>	<b>Introduction</b>	2
1.1.	Formulation of the problem	3
1.2.	Scheme of the proof	8
1.3.	Comparison with previous KAM results	12
<b>2</b>	<b>Functional setting</b>	14
2.1.	Function spaces	14
2.2.	Linear operators	15
2.3.	Hamiltonian and translation invariant operators	22
<b>3</b>	<b>Normal forms and integrability properties of the pure gravity water waves</b>	26
3.1.	Hamiltonian structure of water waves	26
3.2.	Comparison among Birkhoff normal form procedures	28
3.3.	Approximate constant of motion for the gravity water waves	30
3.4.	Identification of normal forms	32
<b>4</b>	<b>Weak Birkhoff Normal Form</b>	34
4.1.	Resonances	35
4.2.	Proof of Proposition 4.3	36
<b>5</b>	<b>The nonlinear functional setting</b>	37
5.1.	Action-angle variables	37
5.2.	The functional equation	39
<b>6</b>	<b>Approximate inverse</b>	42
<b>7</b>	<b>The linearized operator in the normal directions</b>	47
7.1.	Homogeneity expansions	48
7.2.	Hamiltonian structure of the linearized operator	51
7.3.	Algebraic properties of the linearized operator	52
<b>8</b>	<b>Symmetrization of the linearized operator at the highest order</b>	52
8.1.	Good unknown of Alinhac	53
8.2.	Complex formulation of Water waves	54
<b>9</b>	<b>Block-diagonalization</b>	55
9.1.	Block-diagonalization at order $1/2$	55

---

Roberto Feola has been supported of the Centre Henri Lebesgue ANR-11-LABX-0020-01 and by ANR-15-CE40-0001-02 “BEKAM” of the Agence Nationale de la Recherche. Filippo Giuliani has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 757802.

9.2. Block-diagonalization at negative orders . . . . .	57
<b>10 Reduction at the highest orders</b> . . . . .	59
10.1. Integrability at order 1 . . . . .	61
10.2. Integrability at order 1/2 . . . . .	66
10.3. Integrability at order zero . . . . .	71
<b>11 Linear Birkhoff normal form</b> . . . . .	75
11.1. Preliminary technical results . . . . .	75
11.2. Terms $O(\varepsilon)$ , $O(\varepsilon^2)$ and identification of normal forms . . . . .	79
11.3. Terms $O(\varepsilon^{\geq 3})$ : analysis of high modes . . . . .	84
11.4. Terms $O(\varepsilon^{\geq 3})$ : analysis of low modes . . . . .	85
<b>12 Inversion of the linearized operator</b> . . . . .	87
<b>13 The Nash-Moser nonlinear iteration</b> . . . . .	88
13.1. Measure estimates . . . . .	90
13.2. Conclusion of the Proof of Theorem 5.6 . . . . .	93
<b>Appendix A Flows and conjugations</b> . . . . .	93
A.1. Flows of pseudo differential PDEs . . . . .	93
A.2. Egorov Theory and conjugation rules . . . . .	96
<b>Appendix B Technical Lemmata</b> . . . . .	99
B.1. Dirichlet-Neumann operator . . . . .	99
B.2. Technical Lemmata of Section 11 . . . . .	101
<b>References</b> . . . . .	105

## 1. INTRODUCTION

We consider a two dimensional (1d interface) incompressible and irrotational perfect fluid with periodic boundary conditions under the action of gravity. The fluid layer is assumed to be infinitely deep and the motion is governed by the free surface Euler equations. This paper is concerned with the existence and the stability of small amplitude, quasi-periodic in time, traveling waves on the surface of the fluid. By a quasi-periodic traveling wave we mean a motion that, at the first orders of amplitude, is a superposition of an arbitrarily large number of periodic traveling waves with rationally independent frequencies. The irrationality of the frequencies of oscillations excludes the existence of a moving frame for which such motions are stationary. As a consequence the traveling quasi-periodic waves are “truly” time dependent solutions and the search for them is a *small divisors* problem. A natural approach to deal with it is to implement a Nash-Moser implicit function scheme.

Besides the small divisors difficulty, the pure gravity water waves system with infinite depth presents several issues regarding the bifurcation of small amplitude solutions. Indeed the absence of parameters, like the capillarity of the fluid or the depth of the ocean, makes the linearized problem at the rest surface *completely resonant*, in the sense that the kernel of the linearized operator has infinite dimension.

In the search for periodic waves, the same issues have been tackled by Iooss-Plotnikov-Toland for proving the existence of standing periodic unimodal waves [51] and, later, by Iooss-Plotnikov [48] for the multimodal generalization.

In the quasi-periodic case, the complexity of the resonant waves interactions and the analysis of the linearized operator in a neighborhood of the equilibrium demand more refined techniques.

The first results on the existence of quasi-periodic waves have been provided only recently and are due to Berti-Montalto [18] for the gravity-capillary case with infinite depth and by Baldi-Berti-Haus-Montalto [6] for the pure gravity case with finite depth for 2d oceans. In both cases the existence of quasi-periodic solutions is provided for some asymptotically full-measure set of the parameters of the problem, respectively capillarity and depth (or equivalently wavelength). We remark that such solutions are standing, i.e. even in space, and reversible (they enjoy an additional symmetry in the space-time variable).

As far as we know, all the previous results on periodic and quasi-periodic in time water waves take advantage from the presence of physical parameters and / or assumptions of parity conditions.

The purpose of the present paper is to address two natural questions: (i) we work on a fixed equation for which the only possible parameters to modulate are the initial data of the solutions.

(ii) we look for a general class of quasi-periodic traveling waves which are free from restrictions of parity in the spatial and time variables.

We deal with these issues by using Birkhoff normal form methods and fully exploiting the (approximate) constants of motion of the pure gravity water waves system.

The theorem that we prove is given below and it has been announced in [28].

**Theorem 1.** *There exist non-trivial small amplitude, linearly stable quasi-periodic traveling waves solutions of the two dimensional pure gravity water waves problem in infinite depth.*

To the best of our knowledge the above theorem is the first existence result concerning quasi-periodic solutions of the water waves equations bifurcating from a completely resonant elliptic fixed point.

As it is well known one of the main difficulties in the search for quasi-periodic solutions is to deal with the resonances between the frequencies of the expected solutions and the eigenvalues of the linearized operator.

This requires a complete control on the frequencies and a good knowledge of the spectrum in a whole neighborhood of the equilibrium. Especially in the infinite dimensional context, this demands the use of some parameters that modulate the frequencies.

In a completely resonant case such parameters are to be found through a *nonlinear* bifurcation analysis. In the small amplitude regime this can be successfully done by performing Birkhoff normal form methods. The bifurcation parameters appear naturally as the amplitudes of an appropriate approximate solution, which is obtained by exciting a finite number of modes that we call *tangential sites*. It turns out that the tangential sites have to be chosen in a suitable way and such choice is a fundamental ingredient of the proof of Theorem 1.

A serious difficulty is to prove that the amplitudes provide a sufficiently good modulation to impose *infinitely many* non-resonance conditions. This is usually achieved by exploiting some non-degeneracy of the Birkhoff normal form at order four, which in turn depends on the choice of the tangential sites. Proving the non-degeneracy of the normal form is tantamount to prove a "twist" condition between the bifurcation parameters and the first order corrections of the eigenvalues of the linearized operator at the equilibrium.

Due to the quasi-linear nature of the equations and the presence of non-trivial resonances at order four (called *Benjamin-Feir resonances*) the proof of the aforementioned "twist" condition requires the implementation of a new set of ideas.

The key ingredient to overcome these problems is a novel identification argument of normal forms. We refer to section 3 for a detailed discussion. We remark that, by the lack of parameters, classical uniqueness arguments concerning *non-resonant* Birkhoff normal forms (see for instance [52]) cannot be applied. Actually our approach relies on the presence of approximate conserved quantities and the formal integrability of the Hamiltonian of order four proved in [70], [25], [22].

**1.1. Formulation of the problem.** We consider an incompressible and irrotational perfect fluid, under the action of gravity occupying, at time  $t$ , a two dimensional domain with infinite depth, periodic in the horizontal variable, given by

$$\mathcal{D}_\eta := \{(x, y) \in \mathbb{T} \times \mathbb{R}; -\infty < y < \eta(t, x)\}, \quad \mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z}), \quad (1.1)$$

where  $\eta$  is a smooth function. The velocity field in the time dependent domain  $\mathcal{D}_\eta$  is the gradient of a harmonic function  $\Phi$ , called the velocity potential. The time-evolution of the fluid is determined by a system of equations for the two functions  $(t, x) \rightarrow \eta(t, x)$ ,  $(t, x, y) \rightarrow \Phi(t, x, y)$ . Following Zakharov [69] and Craig-Sulem [23] we denote by  $\psi(t, x) = \Phi(t, x, \eta(t, x))$  the restriction of the velocity potential to the free interface. Given the shape  $\eta(t, x)$  of the domain  $\mathcal{D}_\eta$  and the Dirichlet value  $\psi(t, x)$  of the velocity potential at the top boundary, one can recover  $\Phi(t, x, y)$  as the unique solution of the elliptic problem

$$\Delta\Phi = 0 \text{ in } \mathcal{D}_\eta, \quad \partial_y\Phi \rightarrow 0 \text{ as } y \rightarrow -\infty, \quad \Phi = \psi \text{ on } \{y = \eta(t, x)\}. \quad (1.2)$$

The  $(\eta, \psi)$  variables then satisfy the gravity water waves system

$$\begin{cases} \partial_t\eta = G(\eta)\psi \\ \partial_t\psi = -g\eta - \frac{1}{2}\psi_x^2 + \frac{1}{2} \frac{(\eta_x\psi_x + G(\eta)\psi)^2}{1 + \eta_x^2} \end{cases} \quad (1.3)$$

where  $G(\eta)\psi$  is the Dirichlet-Neumann operator

$$G(\eta)\psi := \sqrt{1 + \eta_x^2}(\partial_n\Phi)|_{y=\eta(t, x)} = (\partial_y\Phi - \eta_x\partial_x\Phi)(t, x, \eta(t, x)) \quad (1.4)$$

and  $n$  is the outward unit normal at the free interface  $y = \eta(t, x)$ .  $G(\eta)$  is a pseudo differential operator with principal symbol  $|D| := -i\partial_x$ , self-adjoint with respect to the  $L^2$  scalar product, positive-semidefinite, and its kernel contains only the constant functions. Without loss of generality, we set the gravity constant to  $g = 1$ .

It was first observed by Zakharov [69] that (1.3) is a Hamiltonian system with respect to the symplectic form  $d\psi \wedge d\eta$  and it can be written as

$$\partial_t \eta = \nabla_\psi H(\eta, \psi), \quad \partial_t \psi = -\nabla_\eta H(\eta, \psi), \quad (1.5)$$

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

$$H(\eta, \psi) := \frac{1}{2} \int_{\mathbb{T}} \psi G(\eta) \psi \, dx + \frac{1}{2} \int_{\mathbb{T}} \eta^2 \, dx \quad (1.6)$$

given by the sum of the kinetic and potential energy of the fluid. The invariance of the system (1.3) in the  $y$  and  $x$  variable implies the existence of two prime integrals, respectively the ‘‘mass’’  $\int_{\mathbb{T}} \eta \, dx$  and the momentum

$$\mathbf{M} := \int_{\mathbb{T}} \eta_x(x) \psi(x) \, dx. \quad (1.7)$$

The Hamiltonian (1.6) is defined on the spaces

$$(\eta, \psi) \in H_0^s(\mathbb{T}; \mathbb{R}) \times \dot{H}^s(\mathbb{T}; \mathbb{R}) \quad (1.8)$$

where  $H^s(\mathbb{T}; \mathbb{R})$ ,  $s \in \mathbb{R}$ , denotes the Sobolev space of  $2\pi$ -periodic functions of  $x$ ,  $\dot{H}^s(\mathbb{T}; \mathbb{R}) := H^s(\mathbb{T}; \mathbb{R}) / \sim$  is the homogeneous Sobolev space obtained by the equivalence relation  $\psi_1(x) \sim \psi_2(x)$  if and only if  $\psi_1(x) - \psi_2(x)$  is a constant<sup>1</sup>, and  $H_0^s(\mathbb{T}; \mathbb{R})$  is the subspace of  $H^s(\mathbb{T}; \mathbb{R})$  of zero average functions.

Since  $\widehat{\eta}_0(t) := \frac{1}{2\pi} \int_{\mathbb{T}} \eta(t, x) \, dx$ ,  $\widehat{\psi}_0(t) := \frac{1}{2\pi} \int_{\mathbb{T}} \psi(t, x) \, dx$  evolve according to the decoupled equations<sup>2</sup>

$$\partial_t \widehat{\eta}_0(t) = 0, \quad \partial_t \widehat{\psi}_0(t) = -g \widehat{\eta}_0(t), \quad (1.9)$$

we may restrict the study of the dynamics to the invariant subspace of  $(\eta, \psi)$  such that

$$\int_{\mathbb{T}} \eta \, dx = \int_{\mathbb{T}} \psi \, dx = 0. \quad (1.10)$$

**Linear water waves system.** Small amplitude solutions are close to the solutions of the linearized system of (1.3) at the equilibrium  $(\eta, \psi) = (0, 0)$ , namely

$$\begin{cases} \partial_t \eta = G(0) \psi \\ \partial_t \psi = -\eta \end{cases} \quad (1.11)$$

where the Dirichlet-Neumann operator at the flat surface  $\eta = 0$  is the Fourier multiplier  $G(0) = |D|$ . The solution of the linear system (1.11) restricted to the subspace (1.10) are

$$\begin{aligned} \eta(t, x) &= \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z} \setminus \{0\}} \left( \eta_j(0) \cos(\sqrt{|j|}t) + |j|^{\frac{1}{2}} \psi_j(0) \sin(\sqrt{|j|}t) \right) e^{ijx}, \\ \psi(t, x) &= \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z} \setminus \{0\}} \left( \psi_j(0) \cos(\sqrt{|j|}t) - |j|^{-\frac{1}{2}} \eta_j(0) \sin(\sqrt{|j|}t) \right) e^{ijx} \end{aligned} \quad (1.12)$$

with coefficients satisfying  $\overline{\eta_{-j}(0)} = \eta_j(0)$  and  $\overline{\psi_{-j}(0)} = \psi_j(0)$ . The frequency of oscillation of the  $j$ -th mode is  $\sqrt{|j|}$  and we refer to the map

$$j \rightarrow \sqrt{|j|}, \quad j \in \mathbb{Z} \setminus \{0\}, \quad (1.13)$$

as the *dispersion law* of (1.11). We note that the dispersion law is even in  $j$ , then there are infinitely many multiple eigenvalues.

<sup>1</sup>The fact that  $\psi \in \dot{H}^s$  is coherent with the fact that only the velocity field  $\nabla_{x,y} \Phi$  has physical meaning, and the velocity potential  $\Phi$  is defined up to a constant. For simplicity of notation we denote the equivalence class  $[\psi]$  by  $\psi$  and, since the quotient map induces an isometry of  $\dot{H}^s(\mathbb{T}; \mathbb{R})$  onto  $H_0^s(\mathbb{T}; \mathbb{R})$ , we will conveniently identify  $\psi$  with a function with zero average.

<sup>2</sup>Since the ocean has infinite depth, if  $\Phi$  solves (1.2), then  $\Phi_c(x, y) := \Phi(x, y - c)$  solves the same problem in  $\mathcal{D}_{\eta+c}$  assuming the Dirichlet datum  $\psi$  at the free boundary  $\eta + c$ . Therefore  $G(\eta + c) = G(\eta)$ ,  $\forall c \in \mathbb{R}$ , and  $\int_{\mathbb{T}} \nabla_\eta K \, dx = 0$  where  $K := \frac{1}{2} \int_{\mathbb{T}} \psi G(\eta) \psi \, dx$  denotes the kinetic energy.

Passing to the complex coordinates, the system (1.11) is equivalent to the *completely resonant* equation

$$\partial_t u = -i|D|^{\frac{1}{2}}u, \quad u = \frac{1}{\sqrt{2}}(|D|^{-\frac{1}{4}}\eta + i|D|^{\frac{1}{4}}\psi). \quad (1.14)$$

The Fourier multipliers in (1.14) are well-defined thanks to the choice (1.10). The solutions (1.12) assume the form

$$u(t, x) = \sum_{j \in \mathbb{Z} \setminus \{0\}} u_j(0) e^{-i\sqrt{|j|}t + ijx}, \quad u_j(0) := \frac{1}{\sqrt{2}}(|j|^{-\frac{1}{4}}\eta_j(0) + i|j|^{\frac{1}{4}}\psi_j(0)) \quad (1.15)$$

and it is clear that they can be either periodic or quasi-periodic depending on the Fourier support.

**Traveling quasi-periodic solutions.** A *quasi-periodic* solution for the system (1.3) with an irrational frequency vector  $\omega \in \mathbb{R}^\nu$ ,  $\nu \geq 1$ , i.e.  $\omega \cdot \ell \neq 0$  for all  $\ell \in \mathbb{Z}^\nu \setminus \{0\}$ , is defined by a smooth embedding

$$\begin{aligned} \mathbb{T}^\nu &\rightarrow H_0^1(\mathbb{T}; \mathbb{R}) \times H_0^1(\mathbb{T}; \mathbb{R}) \\ \varphi &\mapsto U(\varphi, x) := (\tilde{\eta}(\varphi, x), \tilde{\psi}(\varphi, x)) \end{aligned} \quad (1.16)$$

such that

$$U \circ \Psi_\omega^t = \Phi_H^t \circ U, \quad \Psi_\omega^t(\varphi_0) := \varphi_0 + \omega t, \quad (1.17)$$

where  $\Phi_H^t$  is the flow of (1.6). By differentiating (1.17) at  $t = 0$  we get that  $U$  has to be solution of the following partial differential equation

$$\omega \cdot \partial_\varphi U - X_H(U) = 0, \quad (1.18)$$

where  $X_H$  corresponds to the r.h.s of the system (1.3).

Since the Hamiltonians  $H$  and  $M$  commute, then  $s \mapsto \Phi_M^s \circ \Phi_H^t \circ U$ , where  $\Phi_M^t$  is the flow of (1.7), is a one-parameter group of solutions of (1.18). An embedding  $U$  as in (1.16), with frequency  $\mathbf{v} \in \mathbb{R}^\nu$ , which is a solution of the Hamiltonian  $M$  in (1.7) is such that

$$U \circ \Psi_{-\mathbf{v}}^s = \Phi_M^s \circ U, \quad \leftrightarrow \quad U(\varphi_0 - \mathbf{v}s, x) = U(\varphi_0, x + s). \quad (1.19)$$

We refer to  $\mathbf{v}$  as its *velocity vector*. By differentiating (1.19) at  $s = 0$  we find the following quasi-periodic transport equation with constant coefficients

$$\mathbf{v} \cdot \partial_\varphi U + X_M(U) = 0, \quad X_M(U) = (\tilde{\eta}_x, \tilde{\psi}_x). \quad (1.20)$$

If  $U$  satisfies (1.17), (1.19) then

$$U(\varphi_0 + \omega t, x + s) = \Phi_H^t \circ \Phi_M^s \circ U = U \circ \Psi_{-\mathbf{v}}^s \circ \Psi_\omega^t(\varphi_0) = U(\varphi_0 + \omega t - \mathbf{v}s, x).^3 \quad (1.21)$$

The *quasi-periodic traveling waves* we look for are solutions of the form (1.21). We point out that the linear solutions (1.12) are of this type with velocity vector depending on the Fourier support.

**Definition 1. (Quasi-periodic traveling waves).** (i) We say that a function  $(\eta(t, x), \psi(t, x)) : \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{R}^2$  is a *quasi-periodic solution* of (1.3) with irrational frequency vector  $\omega \in \mathbb{R}^\nu$ , if there is an embedding  $U : \mathbb{T}^\nu \rightarrow \mathbb{R}^2$  as in (1.16) such that  $(\eta(t, x), \psi(t, x)) = U(\omega t, x)$  solves (1.3).

(ii) A *quasi-periodic solution* is *traveling* with velocity vector  $\mathbf{v} \in \mathbb{Z}^\nu$  if there is a function  $\tilde{U} : \mathbb{T}^\nu \rightarrow \mathbb{R}^2$  such that

$$(\eta(t, x), \psi(t, x)) = U(\omega t, x) = \tilde{U}(\omega t - \mathbf{v}x). \quad (1.22)$$

We remark that an embedding  $U$  satisfies (1.22), i.e.  $U(\varphi, x) = \tilde{U}(\varphi - \mathbf{v}x)$  for all  $\varphi \in \mathbb{T}^\nu$  (because of the irrationality of  $\omega$ ), if and only if  $U(\varphi, x)$  solves the transport equation (1.20).

We shall construct such solutions localized in Fourier space at  $\nu$  distinct *tangential sites*

$$S := S^+ \cup S^-, \quad S^+ := \{\bar{j}_1, \dots, \bar{j}_m\} \subset \mathbb{N} \setminus \{0\}, \quad S^- := \{\bar{j}_{m+1}, \dots, \bar{j}_\nu\} \subset -\mathbb{N} \setminus \{0\}, \quad (1.23)$$

for some  $1 \leq m \leq \nu$  and where

$$k \neq -j, \quad \forall j \in S^+, \quad \forall k \in S^-. \quad (1.24)$$

Let us denote

$$\max(S) := \max\{|j| : j \in S\}. \quad (1.25)$$

<sup>3</sup>Note that the freedom in the choice of the phase  $\varphi_0$  comes from the autonomous nature of the systems involved.

The solutions of (1.11) that originate by exciting the tangential modes are superpositions of periodic traveling linear waves with velocity  $\bar{j}_i$  and frequency  $\sqrt{|\bar{j}_i|}$ . Such motions are quasi-periodic (or periodic) traveling waves of the form (1.22) with frequency vector

$$\bar{\omega} := \left( \sqrt{|\bar{j}_1|}, \dots, \sqrt{|\bar{j}_\nu|} \right) \in \mathbb{R}^\nu \quad (1.26)$$

and velocity vector

$$\mathbf{v} := (\bar{j}_1, \dots, \bar{j}_\nu) \in \mathbb{Z}^\nu. \quad (1.27)$$

We construct quasi-periodic traveling waves solutions of (1.3) which are “close” to the linear ones, namely they will be of the form

$$\begin{aligned} \eta(t, x) &= \sum_{j \in S} \sqrt{2\zeta_j} |j|^{1/4} \cos(\omega(j)t - jx) + o(\sqrt{|\zeta|}), \\ \psi(t, x) &= - \sum_{j \in S} \sqrt{2\zeta_j} |j|^{1/4} \sin(\omega(j)t - jx) + o(\sqrt{|\zeta|}), \\ \omega &= \bar{\omega} + O(|\zeta|), \end{aligned} \quad (1.28)$$

where  $o(\sqrt{|\zeta|})$  is meant in the  $H^s$ -topology with  $s$  large. The vectors  $(\sqrt{\zeta_j})_{j \in S} \in \mathbb{R}^\nu$  represent the *amplitudes* of the approximate solutions from which we have the bifurcation.

Due to the absence of physical (or external) parameters in the system (1.3), we shall use these unperturbed amplitudes to modulate the frequencies of the expected solutions. It turns out that such strategy is not doable for any choice of the tangential sites  $\bar{j}_i$ , but we will prove that *generically* it is. When we refer to a generic choice of the tangential sites we mean that the  $\bar{j}_i$ 's are chosen such that the vector  $(\bar{j}_1, \dots, \bar{j}_\nu)$  is not a zero of a certain non-trivial polynomial  $\mathbb{C}^\nu \rightarrow \mathbb{C}$ . We remark that such choice is equivalent to the choice of the velocity vector  $\mathbf{v}$  in (1.27).

*Comment on assumption (1.24).* The genericity arguments are used to impose non-resonance conditions between the tangential linear frequencies of oscillations (1.26). The assumption (1.24) guarantees the simplicity of the tangential linear eigenvalues, excluding some delicate non-trivial resonant interactions. In section 1.2-(iv) we explain where this assumption comes in handy in the analysis of the linearized operator.

The (1.24) simplifies also the computation of the first order corrections of the frequency of the expected quasi-periodic solutions, which are important to impose non-degeneracy conditions. See also section 1.2-(i).

Denoting by  $B(0, \varrho)$  the ball centered at the origin of  $\mathbb{R}^\nu$  of radius  $\varrho > 0$ , our result can be stated as follows.

**Theorem 2. (Quasi-periodic traveling gravity waves).** *Let  $\nu \geq 1$ . For a generic choice of the velocity vector  $\mathbf{v}$  as in (1.27) there exist  $s \gg 1$ ,  $0 < \varrho \ll 1$  and a positive measure Cantor-like set  $\mathfrak{A} \subseteq B(0, \varrho)$  such that the following holds. For any  $\zeta \in \mathfrak{A}$ , the equation (1.3) possesses a small amplitude quasi-periodic solution  $(\eta, \psi)(t, x; \zeta) = U(\omega t, x; \zeta)$  of the form (1.28) which is a traveling wave with velocity vector  $\mathbf{v}$ ,  $U(\varphi, x) \in H^s(\mathbb{T}^{\nu+1}, \mathbb{R}^2)$  and  $\omega := \omega(\zeta) \in \mathbb{R}^\nu$  is a diophantine frequency vector. For  $0 < \varepsilon \leq \sqrt{\varrho}$ , the set  $\mathfrak{A}$  has asymptotically full relative measure in  $[\varepsilon^2, 2\varepsilon^2]^\nu$ . Moreover these solutions are **linearly stable**.*

Now we discuss the main issues and the novelties of the paper.

- The general form of the linear frequencies of oscillations for the water waves equations is the following

$$\sqrt{|j| \tanh(\mathfrak{h}|j|)(g + \kappa j^2)},$$

where  $\mathfrak{h}$  and  $\kappa$  are respectively the depth and the capillarity of the fluid. If  $\mathfrak{h} < \infty$  or  $\kappa \neq 0$  such parameters may be used to impose non-resonance conditions (see for instance [6], [18], [62]). In our case  $\mathfrak{h} = \infty$  and  $\kappa = 0$ , thus the linear frequencies of oscillations are  $\sqrt{g|j|}$  and the elements of the infinite dimensional space  $\text{span}\{e^{i|n|t} e^{in^2x} : n \in \mathbb{Z}\}$  are *periodic* solutions of the linearized problem at the origin (completely resonant case).

The physical parameter  $g$  clearly does not modulate the frequencies, hence if we look for quasi-periodic solutions we need to extract parameters directly from the nonlinearities of the equation. We do that by means of *Birkhoff normal form* (BNF) techniques (see section 3 for a detailed discussion). In this way the bifurcation parameters are essentially the “initial data” or the amplitudes of an appropriate approximate solution (see (1.28)) from which the bifurcation occurs. The choice of the Fourier support of such approximate solution plays a fundamental role in proving some non-degeneracy conditions. Roughly speaking, both the amplitudes  $\sqrt{\zeta_j}$  and the tangential sites  $\bar{j}_i$  will be “parameters” of our problem.



- The frequency of the expected quasi-periodic solutions are close to resonant vectors, then the diophantine constant  $\gamma$  appearing in non-resonance conditions, see for instance (5.22), has small size as the amplitudes. This clearly produces difficulties in the application of perturbative methods.
- In performing BNF procedures we shall deal with resonances among linear frequencies. It is known that the pure gravity case in infinite depth has no 3-waves resonant interactions. On the other hand, there are many *non-trivial* 4-wave interactions, called *Benjamin-Feir* resonances (see (3.30)). We then exploit a fundamental property of the pure gravity waver waves Hamiltonian (1.6) in infinite depth: the formal integrability, up to order four, of the Birkhoff normal form. This has been proved in [70], [25], [22] by showing explicit key algebraic cancellations occurring for the coefficients of the Hamiltonian.
- In order to show that the “initial data” of the expected solutions tune in an efficient way the frequencies we shall find the explicit expression of the first order corrections of the tangential frequencies and of the spectrum of the linearized operator in the normal directions. To do that we use an identification argument of normal forms based on the presence of approximate constants of motion (see section 3.4 and Proposition 11.18). Actually, even after this procedure, the Hamiltonian is still partially degenerate. Indeed it turns out that there is a finite number of eigenvalues which are still in resonance. We overcome this difficulty by passing to suitable rotating coordinates, see Lemma 11.19.
- We exhibit the existence of a wide class of traveling quasi-periodic solutions with no parity restrictions in time and space by using the Hamiltonian structure and the  $x$ -translation invariance of (1.3). It is well known that the water waves system (1.3) exhibits additional symmetries. For instance the vector field  $X_H$  in (1.5) is

(i) *reversible* with respect to the involution

$$S : \begin{pmatrix} \eta(x) \\ \psi(x) \end{pmatrix} \mapsto \begin{pmatrix} \eta(-x) \\ -\psi(-x) \end{pmatrix} \quad (1.29)$$

i.e. it satisfies  $X_H \circ S = -S \circ X_H$ ;

(ii) *even-to-even*, i.e. maps  $x$ -even functions into  $x$ -even functions.

In several papers (for instance [6], [18], [51], [62]) such symmetries are adopted to remove degeneracies due to translation invariance in  $x$  and  $t$ . In the aforementioned works the authors look for solutions which are

- standing, namely even in  $x$

$$(\eta, \psi)(t, x) = (\eta, \psi)(t, -x),$$

- reversible, namely

$$(\eta, \psi)(t, x) = (\eta, -\psi)(-t, -x) = S(\eta, \psi)(-t, x). \quad (1.30)$$

We observe that combining these properties, the solutions have to satisfy the following parity conditions

$$\eta \text{ is even in } x \text{ and even in } t, \quad \psi \text{ is odd in } t \text{ and even in } x.$$

Since we do not look for solutions in the subspace of reversible functions, the existence of a solution  $U(t, x)$  of (1.3) implies the existence of a possibly different solution  $SU(-t, x)$ .

- The vector field in (1.3) is a *singular perturbation* of the linearized system at the origin (1.11) since the nonlinearity contains derivatives of the first order, while (1.14) contains only derivatives of order  $1/2$ . We remark that this is not that case when the capillarity  $\kappa \neq 0$ .
- The dispersion law (1.13) is *sub-linear*. This is a major difference in developing KAM theory for equations with *super-linear* dispersion law, such as in the gravity-capillary case. Indeed a weaker dispersion law implies *bad* separation properties of the eigenvalues. The main issue concerns the verification of non-resonance conditions between the tangential frequencies and the differences of the normal ones, called “second order Melnikov conditions”. The bad separation properties force to impose very weak conditions. In this paper we adopt a different strategy with respect to [6] where the same problem occurs.

Our approach is well adapted to the completely resonant case since it allows to reduce the number of steps of nonlinear bifurcation. It is based on the conservation of momentum, the algebraic structure of the equations and a careful analysis of the resonant regimes. We refer to section 1.2 for a detailed discussion.

- The stability result is the same given in Theorem 1.2 in [6]. This is of course an interesting dynamical information itself. We also remark that it is a consequence of an important ingredient of the proof, a reducibility argument of the linearized equation at the quasi-periodic solution.

By a linearly stable  $U(\omega t, x)$  we mean that the linearized operator at the embedded torus  $U(\varphi, x)$  has purely imaginary spectrum. In particular we are able to provide a set of coordinates in which the linear problem is diagonal in the

directions normal to the torus. As a consequence the Cauchy problem of the linearized equation is stable, i.e. the Sobolev norms are uniformly bounded in  $t$ .

**Literature.** We present some results on the water waves systems.

*Euclidean case.* In the Euclidean case, i.e. when  $x \in \mathbb{R}^d$ , the problem of local/global well-posedness has been addressed by several authors. Without trying to be exhaustive we mention Coutand-Shkoller [20], Lindblad [59], Lannes [56] Alazard-Burq-Zuily [2, 3] for local well-posedness results. We refer to [46] for a survey on the subject (and reference therein). In the Euclidean case it is also possible to construct global in time solution by exploiting the dispersive effect of the linearized problem. We quote Germain-Masmoudi-Shatah [39], Wu [68], Ionescu-Pusateri [44], Alazard-Delort [4, 5], Ifrim-Tataru [42]. See also [67], [45], [43].

We now briefly discuss some known results about the the existence of *special* solutions for the water waves equations.

*Traveling waves.* Early results about *traveling* waves date back to 19-th century. In [65] Stokes provided the first nonlinear analysis of the two dimensional traveling gravity waves, by computing a nonlinear approximation of the flow. The rigorous bifurcation of small amplitude bi-dimensional traveling gravity water waves solutions has been obtained by Levi-Civita [58], and Struik [66]. The three-dimensional case has been successfully approached more recently. We quote the paper [21] (and reference therein) by Craig-Nicholls where the existence of traveling wave solutions is proved in the gravity-capillary case with space periodic boundary conditions. In absence of capillarity the existence of periodic traveling waves is a small divisor problem. This case has been treated by Iooss and Plotnikov in [49, 50].

*Standing waves.* The time periodic *standing waves* are not stationary with respect to a moving reference frame. We quote Plotnikov and Toland [62] for the gravity case in finite depth, Iooss-Plotnikov-Toland [51] and Iooss-Plotnikov [48] in infinite depth. In [1] Alazard-Baldi show the existence of periodic gravity-capillary standing waves.

We remark that the above quoted results concern *periodic* in time solutions.

*Time quasi-periodic waves.* The existence of time quasi-periodic solutions for PDEs has been widely studied from the '80s. We mention the pioneering works by Kuksin [53], Craig-Wayne [24], Bourgain [19], Kuksin-Pöschel [55]. The KAM theory for PDEs with *unbounded* perturbation is quite more recent, see for instance [54, 60, 10]. The problem of dealing with *quasi-linear* equations has been tackled first by Baldi-Berti-Montalto in [7, 8] for the KdV equation. Their method has been extended to deal with many models of interest in hydrodynamics, such as the Schrödinger equation [35], generalized KdV [40], Kirchoff [61], Degasperis-Procesi equation [32]. Regarding the water waves problem we quote the novel results of existence of standing waves by Berti-Montalto [18] for the gravity-capillary case and by Baldi-Berti-Haus-Montalto [6] for the gravity case in finite depth. Eventually we mention the recent work [16] and the preprint [17] by Berti-Franzoi-Maspero on the existence of quasi-periodic traveling waves for fluids with constant vorticity, and [9] by Baldi-Montalto regarding time quasi-periodic solutions for the incompressible Euler equation on  $\mathbb{T}^3$ .

*Birkhoff normal form and long time stability on tori.* The normal form theory plays a fundamental role in our approach. In particular we refer to the pioneering works [70, 27, 25] (see also [22]) where is proved the integrability (at the formal level) of the Birkhoff normal form at order four of the pure gravity water waves in infinite depth. The paper quoted above provide just a “formal” result in the sense that no actual relation can be established between the flow generated by  $H$  in (1.6) and the one generated by the Hamiltonian in Birkhoff normal form.

Rigorous long time existence results for water waves, based on a normal form approach, are quite more recent. We mention [12] by Berti-Delort, where the authors provide the existence of solutions evolving from  $\varepsilon$ -small data up to time of order  $\varepsilon^{-M}$  for the 2-d gravity-capillary water waves equations. Ionescu-Pusateri in [47] prove a life span of order  $\varepsilon^{-5/3+}$  for the gravity-capillary waves in 3-d. These results hold for almost all values of the gravity and surface tension parameters.

Without the use of any external parameter, we mention the paper by Berti-Feola-Pusateri [14] on the pure gravity case (1.3) (see also [15]). Using a novel identification of normal form argument introduced in [32], the authors proved an  $\varepsilon^{-3}$  existence result providing a rigorous proof of the Zakharov-Dyachenko conjecture (see [70], [27]). With a similar idea Berti-Feola-Franzoi proved in [13] a quadratic life span for gravity-capillary water waves for any values of the gravity and surface tension.

**1.2. Scheme of the proof.** Here we summarize the steps of the proof of Theorem 2 highlighting the key ingredients. For further details see [28].



(i) *Nash-Moser theorem of hypothetical conjugation.* The quasi-periodic solutions are found as zeros of the nonlinear functional equation (6.1), (5.37). We apply a Nash-Moser scheme (see Theorem 13.2) that provides the zeros of such functional as limit of a sequence of approximate solutions convergent in some Sobolev space.

The main issues concern the invertibility of the linearized operator in a neighborhood of the equilibrium and the search for a good approximate solution that initializes the scheme.

In sections 4 and 5 we construct the first nonlinear approximate solution from which the expected solutions bifurcate. From a geometrical point of view, we determine the "unperturbed" embedded torus that we want to continue to a torus which is invariant for the full Hamiltonian system (1.6).

In a suitable set of coordinates the linear dynamics of the tangential (to the torus) variables is decoupled by the dynamics of the normal ones (we follow the Berti-Bolle method [11]). In section 6 we solve the equations corresponding to the tangential part and in sections 7-12 we deal with the linearized operator in the normal directions.

(ii) *Weak Birkhoff normal form.* In order to find the first nonlinear approximation we implement a Birkhoff normal form method, more precisely we construct a map, which is close to the identity up to a finite rank operator, such that the Hamiltonian  $H$  in (1.6) assumes the form (see (4.7))  $H_{\text{BirK}} + R$  with  $R$  a small remainder and

1. there exists a finite dimensional subspace  $U_S$  invariant for  $H_{\text{BirK}}$ ;
2. the Hamiltonian restricted to  $U_S$  is integrable and non-degenerate in the sense that the "frequency-to-amplitude" map (see (5.4)) is invertible.

We refer to this procedure as a *weak* Birkhoff normal form, see the discussion in section 3.

At the  $n$ -th step of the BNF procedure one has to deal with  $n$ -waves resonant interactions, which are zeros of the algebraic equations

$$\sum_{i=1}^n \sigma_i j_i = 0, \quad \sum_{i=1}^n \sigma_i \sqrt{|j_i|} = 0, \quad \sigma_i = \pm 1, \quad i = 1, \dots, n. \quad (1.31)$$

We say that a  $n$ -tuple  $(j_1, \dots, j_n)$  is a *trivial* resonance if  $n$  is even,  $\sigma_i = -\sigma_{i+1}$ ,  $i = 1, \dots, n-1$  (up to permutations), and the  $n$ -tuple has the form  $(j, j, k, k, \dots)$ . It is easy to note that monomials  $u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}$ , with  $u_j^{\pm} := u_j, u_j^- := \overline{u_j}$ , supported on trivial resonances are *integrable* or *action preserving*, meaning that depend only on the actions  $|u_j|^2$ . Unfortunately, even for  $n = 4$ , there are infinitely many non-trivial solutions of (1.31). For  $n = 4$  they are called Benjamin-Feir resonances (see (3.30)). To deal with these resonances we reason as follows. In [70], [25], [22] it has been proved that the coefficients of the normalized Hamiltonian (obtained by a *full* Birkhoff normal form procedure, see section 3) at order four of the monomials corresponding to the Benjamin-Feir resonances vanish. In subsections 3.3, 3.4, by using suitable algebraic arguments, we actually prove that such cancellations of [70], [25], [22] occur also performing a *weak* version of BNF which involves only *finitely* many *tangential sites* (see (1.23)). This is the content of Proposition 3.3. In this way we conclude the *integrability* of the weak BNF Hamiltonian at degree 4 and we obtain its explicit formula (4.8). We remark that formula (4.8) is fundamental to prove the non-degeneracy of the frequency shift. Without assumption (1.24) this formula would be more complicated and in principle it would be not clear whether the twist condition (see Lemma 5.2) is satisfied.

In order to deal with higher order resonances we use the genericity argument of section 4.1. We remark that this argument exploits the conservation of momentum, as it is evident from the proof of Proposition 4.4.

(iii) *Invertibility of the linearized operator.* In section 7 we compute the linearized operator in the normal directions  $\mathcal{L}_\omega$  (see (7.4)).

The invertibility is obtained by a *reducibility argument* that consists into two main steps:

- (a) A pseudo differential reduction in decreasing order of  $\mathcal{L}_\omega$  which conjugates the linearized operator to a pseudo differential one with constant coefficients up to a bounded remainder;
- (b) A reduction of bounded operators and a KAM scheme which completes the diagonalization;

The main new issues are the following:

- (I) By the complete resonance of the linear equation (1.14) the frequency  $\omega$  is close to a resonant vector, then the diophantine constant  $\gamma$  in (5.21) (see also (5.22)) is small with the size  $\varepsilon$  of the amplitudes. This implies that many terms in  $\mathcal{L}_\omega$  are not perturbative, in the sense that, roughly speaking, the size of their coefficients divided by  $\gamma$  is big ( $\varepsilon\gamma^{-1} \gg 1$ ).
- (II) We consider the linearization on a quasi-periodic traveling function  $U(\varphi, x)$  without any assumption on the parity of  $\varphi$  and  $x$ . Usually such conditions provide some algebraic cancellations in performing steps (a)-(b) and reduce the multiplicity of the eigenvalues simplifying the proof.

The regularization procedure (a) consists in applying several changes of coordinates. Each change of variables makes constant the coefficient of an unbounded pseudo differential operator and it is constructed in two steps. This is because the coefficients of the linearized operator are sum of two parts, one is non-perturbative in size, in the sense of (I), and the other one is just small enough. Then in the first step, that we call *preliminary step*, we treat the non-perturbative part. In the second one we make constant the whole coefficient. In the regularization procedure (a) the pseudo differential structure is fundamental, because it guarantees that the linearized operator is diagonal up to a remainder whose order is getting smaller and smaller at any step. This fact relies on the property that the commutator between two pseudo differential operators gains one derivative (see Lemma 2.19).

The procedure (b) completes the reduction of the linearized operator. It is again a transformation method but it deals only with bounded terms. Also in this case the changes of coordinates are constructed in two steps, for the non-perturbative and perturbative terms. The steps that deal with the non-perturbative terms are called *linear Birkhoff normal form steps*. The combination of the preliminary steps, the Linear BNF and the identification argument of the Proposition 11.18 allows to construct a normal form around the approximately invariant embedded tori. Once we obtained this normal form we can apply a KAM scheme as in a semilinear case.

In both procedures (a) and (b) the fact that the coefficients can be made constant is guaranteed by the assumption (1.24) and certain symmetries. Since we do not assume parity conditions on the approximate solutions on which we linearize, as explained in (II), we need to exploit the Hamiltonian structure and the conservation of momentum. At the linear level the presence of such symmetries comes from the fact that we linearize a Hamiltonian operator on a quasi-periodic traveling function (see section 7.3 for more details). Along the reducibility procedure we need to ensure that this structure is left invariant by all the changes of coordinates we perform.

One of the main issues that we deal with by using the conservation of momentum is when we need to reduce to constant coefficient the transport term

$$\omega \cdot \partial_\varphi + V(\varphi, x) \partial_x \quad (1.32)$$

appearing in  $\mathcal{L}_\omega$ . This is the most delicate point of procedure (a) and we describe it below. Actually this problem is equivalent to straighten the degenerate vector field  $\omega \cdot \frac{\partial}{\partial \varphi} + V(\varphi, x) \frac{\partial}{\partial x}$  on the  $\nu + 1$ -dimensional torus. By using the fact that we linearized on a quasi-periodic traveling wave (or using momentum conservation) we can reduce it to the study of the following *non-degenerate* vector field (on a lower dimensional torus)

$$(\omega - \mathbf{V}(\Theta)_\nu) \cdot \frac{\partial}{\partial \Theta}, \quad \Theta \in \mathbb{T}^\nu,$$

where  $\mathbf{V}(\varphi - \nu x) = V(\varphi, x)$ . Then we can apply a result of straightening of weakly perturbed constant vector fields on tori given in [29].

(iv) *Non-resonance conditions.* Since the dispersion law is sub-linear, at each step of items (a) and (b) small divisors problems arise. We discuss the non-resonance conditions that we shall require on the frequencies.

In the preliminary steps described above we encounter the same small divisors appearing in the weak BNF procedure, namely the combinations of the tangential linear frequencies of oscillations

$$\sum_{i=1}^n \sigma_i \sqrt{|j_i|}, \quad j_i \in S, \quad \sigma_i = \pm. \quad (1.33)$$

In the weak BNF we use momentum conservation

$$\sum_{i=1}^n \sigma_i j_i = 0 \quad j_i \in S, \quad \sigma_i = \pm \quad (1.34)$$

to apply genericity arguments ensuring that the above relations vanish just in the trivial case. Indeed if a non-trivial  $(\sigma_i, j_i)_{i=1}^n$  makes (1.33) vanish, then one can choose the tangential sites such that (1.34) is not satisfied (see the proof of Proposition 4.4).

This corresponds to a generic choice because the conditions (1.34) provide a system of **finitely many** polynomial equations. However the same reasoning cannot be applied in the preliminary steps, where the conservation of momentum reads as

$$\sum_{i=1}^n \sigma_i j_i + j - k = 0 \quad \forall j_i \in S, \quad j, k \in S^c,$$

and so it involves infinitely many polynomial equations. In this case we are forced to choose the tangential sites out of the zero set of the following finitely many algebraic functions

$$\sum_{i=1}^n \sigma_i \sqrt{|j_i|} = 0, \quad j_i \in S, \quad \sigma_i = \pm.$$

This is certainly a generic choice, but the problem is that the above functions are identically zero if  $(\sigma_i, |j_i|)_{i=1}^n$  is a trivial resonance, and not just  $(\sigma_i, j_i)_{i=1}^n$ . This is the typical issue that arises in a case with multiple eigenvalues. Then by assuming the condition (1.24) on the tangential set we get rid of this problem.

In order to deal with the operator in (1.32) we shall impose the following ‘‘zero order Melnikov conditions’’:

$$|(\omega - \mathbf{m}_1 \mathbf{v}) \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\} \quad (1.35)$$

where  $\mathbf{m}_1 \approx (2\pi)^{-\nu} \int_{\mathbb{T}^\nu} \mathbf{V}(\Theta) d\Theta$ ,  $\tau > \nu + 1$ .

In section 11 we compute the first order corrections to the eigenvalues of the linearized operator in the normal directions. This step is fundamental in order to impose the non-resonance conditions required for the KAM reducibility scheme of the procedure (b). Thanks to the identification argument of section 11.2 it turns out that we need to prove the non-degeneracy of the following map:

$$\omega \mapsto \mathbb{A} \zeta(\omega) \cdot \ell + m_1(\omega)(j - k) + c_j(\omega)j - c_k(\omega)k, \quad \mathbf{v} \cdot \ell + j - k = 0, \quad \forall \ell \in \mathbb{Z}^\nu, \quad j, k \in S^c,$$

where  $\mathbb{A}$  is the matrix in (5.3) (which only depends on the choice of the tangential sites),  $m_1$  is in (10.3) and  $c_j$  is in (11.61). In principle the presence of the  $c_j$ 's produces serious difficulties in imposing the non-resonance conditions. However only a finite number of such corrections are different from zero. We show that in an appropriate set of rotating coordinates those corrections disappear. See Lemma 11.19.

In the KAM reducibility scheme in the procedure (b) we shall impose suitable lower bounds on the functions

$$\psi_{\ell, j, k} = \omega \cdot \ell + m_1(j - k) + (1 + m_{\frac{1}{2}})(\sqrt{|j|} - \sqrt{|k|}) + m_0(\text{sign}(j) - \text{sign}(k)) + r_j - r_k \quad (1.36)$$

$\ell \in \mathbb{Z}^\nu$ ,  $j, k \in \mathbb{Z} \setminus \{0\}$ , with  $m_{1/2}$ ,  $m_0$ ,  $r_j$  depending on  $\omega$  and satisfying

$$(\sup_j |j|^{-1/2} |r_j| + |m_{1/2}| + |m_0|) \gamma^{-1} \ll 1.$$

These conditions are called ‘‘second order Melnikov conditions’’ and are of the form

$$|\psi_{\ell, j, k}| \geq \frac{\eta}{\langle \ell \rangle^\tau}, \quad \ell \in \mathbb{Z}^\nu, \quad j, k \in \mathbb{Z} \setminus \{0\}, \quad (\ell, j, k) \neq (0, j, j), \quad (1.37)$$

where  $\eta, \tau > 0$  are some constants to be fixed.

We prove that the (1.37) hold with

$$\eta := \gamma^3 \ll \gamma, \quad (1.38)$$

where  $\gamma$  is the constant appearing in (1.35). Actually we show that the measure of the complementary of the set of frequencies  $\omega$  such that (1.37) holds goes to zero as  $\varepsilon \rightarrow 0$ . The crucial problem is the summability in the indexes  $\ell, j, k$ . Since in this case the linear frequencies grow at infinity only sublinearly, the differences  $\sqrt{|j|} - \sqrt{|k|}$  accumulate everywhere in  $\mathbb{R}$ . This means that for any fixed  $\ell$  there are *infinitely many* indexes  $j, k$  to be taken into account. The key idea that we use is the following. By the conservation of momentum we need to impose the conditions (1.37) only for  $(\ell, j, k)$  satisfying

$$\mathbf{v} \cdot \ell + j - k = 0, \quad (1.39)$$

where  $\mathbf{v}$  is given in (1.27). This allows us to show (see Lemma 13.8) that, if  $|j|, |k|$  are much larger than  $|\ell|$ , then the conditions (1.37) are implied by the (1.35) and then we are left to control the small divisors only for *finitely many* indexes  $j, k$ .

We point out that, despite the multiplicity of the linear eigenvalues  $\sqrt{|j|}$ , the linearized operator is diagonal in the Fourier basis, and not just block-diagonal. This is usually the case because one cannot impose the (1.37) for  $\ell = 0, j = \pm k$ . Thanks to the conservation of momentum, when  $k = -j$  in (1.36) the relation (1.39) implies that

$$\mathbf{v} \cdot \ell = -2j$$

and

$$\psi_{\ell, j, -j} = \omega \cdot \ell + 2m_1 j + 2m_0 \text{sign}(j) + r_j - r_{-j} = \bar{\omega} \cdot \ell + \varepsilon^2 (\mathbb{A} \zeta(\omega) - m_1(\omega) \mathbf{v}) \cdot \ell + o(\varepsilon^2).$$

By choosing  $\omega$  in the set  $\mathcal{G}_0^{(1)}$  in (5.22), the first order expansion (the  $\varepsilon^2$ -terms) of  $\psi_{\ell, j, -j}$  does not vanish and we are able to impose a lower bound as (1.37).

We mention that an alternative approach to impose the non-resonance conditions required by the KAM scheme is to prove Melnikov conditions which *lose* derivatives by setting

$$\eta := \frac{\gamma}{\langle j \rangle^d \langle k \rangle^d}, \quad d > 1. \quad (1.40)$$

With this choice “many” frequencies  $\omega$  satisfy the (1.37), but the small divisors creates a true loss of space derivatives.

The class of non perturbative terms is enlarged by the choice (1.38) but such terms are compactly time-Fourier supported ( $|\ell| \leq C$  for some constant  $C > 0$ ). It turns out that if at least one between  $|j|, |k|$  is large enough then there are no small divisors (see Lemma 11.3). Otherwise we choose to impose the second Melnikov conditions with  $\eta$  as in (1.40). Actually this does not create any loss of derivatives since  $|j|, |k|$  are taken into a ball with finite radius.

Another possibility would be to use  $\eta$  as in (1.40) in any regimes  $\ell, j, k$  as done in [6]. As a drawback this choice would require several steps of reduction to constant coefficients up to smoothing remainders before applying the KAM scheme of the procedure (b) above. Since we are dealing with a resonant case, it turns out that the accuracy necessary for the bifurcation point depends on the number of reduction steps. In other words the number of weak Birkhoff normal form steps in section 4 increases rapidly as the number of reduction steps increases. Choosing to reduce to constant coefficients only the pseudo differential operators of non-negative orders (up to remainders which are only 1/2-smoothing) provides the optimal balancing of the two procedures.

**1.3. Comparison with previous KAM results.** In this section we discuss the main differences between the present work and related KAM results for PDEs.

*KAM results without external or physical parameters.* As it is well-known the implementation of KAM schemes in the PDE context requires the use of parameters to impose (infinitely many) non-resonance conditions.

Then it is common to use physical or external parameters to modulate the linear frequencies of oscillation. Excluding the cases of the water waves systems that we discussed before, this can be done artificially by adding convolution or multiplicative potentials, as for instance in nonlinear Schrödinger equations, or more naturally by considering mass terms, as in the case of Wave/Klein-Gordon equation.

However many important physical phenomena are modeled by equations without parameters. In these cases the parameters are extracted directly from the equation by using Birkhoff normal form methods. This approach has been implemented successfully in [55, 37, 38, 36, 64, 32]. In all these works the authors exploit also the conservation of momentum. Indeed usually the modulation of the inner parameters is not strong enough to impose non-resonance conditions.

As we discussed in section 1.2, in our case the conservation of momentum is fundamental to deal with the multiplicity of the linear eigenvalues and the sublinear dispersion law. We remark that in some papers, for instance [8, 32], the preservation of momentum is used just up to some order. In fact in [8] the authors may deal with non- $x$ -translation invariant nonlinearities. However in that case the dispersion law is super-linear (stronger separation property of the eigenvalues) and the linear spectrum is simple. In [32] the dispersion law is linear and this fact is used to prove that many second order Melnikov conditions are implied by first order Melnikov conditions. Thanks to this argument one does not need to fully exploit the conservation of momentum at any order.

*Comparison with previous KAM results for fluid dynamics models .* Here we make a brief comparison between the present paper and the works [8, 18, 6, 32] (mentioning also [40], [16]) concerning results of existence and stability of small amplitude quasi-periodic solutions for nonlinear shallow waters and water waves systems in spatial periodic domains of dimension one. The overall strategy for these KAM results for quasi-linear PDEs, based on Nash-Moser theory, has been developed in [11] and nowadays it is well-established. However, as usual in the PDEs context, every equation presents different issues. We believe that the main ingredients to take into account are

- The (non-) resonant nature of the equilibrium and the presence of parameters;
- The presence of non-trivial resonances of order 4;
- The use of the (approximate) integrability of the *unperturbed* system;
- The dispersion law: super-linear, linear and sub-linear cases;
- The analysis of the linearized operator: invertibility in a whole neighborhood of the equilibrium.

The main distinction among the aforementioned papers relies is the fact that the works [8, 32] are parameter-free and completely resonant problems, while for [18, 6] the elliptic equilibrium is non-resonant for the set of the chosen parameters. We start by commenting the former group of results.

In [8] the authors consider the KdV equation with a high order perturbation (a smooth nonlinearity of order 4 at the

origin), hence very small close to the origin. The linear frequencies of oscillations are  $j^3$ ,  $j \in \mathbb{Z}$  and there are no non-trivial resonances at order four. This fact depends on the particular form of the linear eigenvalues and the conservation of momentum at order four, and it does not rely on the integrability of the KdV equation, which is not exploited to prove this result. The first order corrections of the eigenvalues, which are fundamental to prove a non-degeneracy condition, it is not affected by the presence of the perturbation. This makes easier the nonlinear bifurcation analysis. Following the strategy of [8], in [40] the second author of the present paper deals with quadratic quasi-linear perturbations of KdV <sup>4</sup>. The lower order of the nonlinearity implies a stronger perturbative effect on the linear dynamics and one of the main problems is to keep track of the corrections to the eigenvalues also in the regularization procedure described in section 1.2-(iii) (*preliminary steps*). All these computations are made without the use of identification arguments. This is doable thanks to the classical differential structure of the equation.

Similar difficulties are encountered in [32] to prove existence of reducible KAM tori for the Degasperis-Procesi equation and its (high order) Hamiltonian perturbations. In this case the bifurcation analysis is challenging due to the complicated structure of the resonances. Indeed non-trivial resonances are present also at order four, as the Benjamin-Feir resonances for the pure gravity system with infinite depth. The authors fully exploit the integrability of the equation and develop a novel argument of identification of normal forms to provide non-degeneracy conditions. Another fundamental difference with the works on KdV equations is the fact that the dispersion law is asymptotically linear  $j \rightarrow j + O(1/j)$ , instead of super-linear  $j \rightarrow j^3$ . This complicates both the imposition of non-resonance conditions, since there is a weaker separation of the eigenvalues, and the analysis of the linearized operator. Concerning the latter, in the KdV case the linearized operator is purely differential, while in the DP case this has a pseudo differential structure. A KAM result for the pure gravity water waves system in infinite depth collects all the difficulties previously mentioned: a weak dispersion law, absence of parameters, presence of many non-trivial resonances of lower order, pseudo differential structure of the linearized operator. The main differences with the DP case rely on the sub-linear dispersion law  $j \rightarrow \sqrt{|j|}$ , the lack of constants of motion (except momentum) and the fact that the linear analysis is a *singular* perturbation problem. In section 3 we explain in details how to use the *formal* approximate integrability of the pure gravity system in finite depth and the characterization of the Benjamin-Feir resonances to provide the identification argument of normal forms.

Concerning the works [18, 6], the dispersion law of the gravity-capillary case in infinite depth and pure gravity with finite depth is respectively super-linear and sub-linear. In both cases the measure estimates for the set of the good parameters (for which the existence of quasi-periodic water waves is ensured) are provided by using degenerate KAM theory. The present work and [6] share the difficulty of the weak separation of the linear eigenvalues, but we deal with this issue in a different way, as it has been explained in section 1.2-(iv). In [6] the authors impose Melnikov conditions that imply a possible loss of (spatial) derivatives of the KAM transformations. This is compensated by a preliminary long regularization procedure that makes constant coefficients the linearized operator up to very smooth terms. In a resonant case such procedure is not convenient, because the number of steps of the normal form methods that we need to perform increases with the number of steps of the regularization procedure. Hence we do not impose Melnikov conditions with loss of derivatives thanks to a careful analysis of the small divisors (we refer to section 11 for more details).

Concerning the role of integrability, we remark that in non-resonant cases the integrability is usually used just at the linear level. Roughly speaking the equation is seen as a small perturbation of the linear part, that is integrable, and the *unperturbed tori* are linear solutions obtained by exciting a finite number of modes.

Regarding the analysis of the linearized operator, except for the number of steps of the regularization procedure, an important difference with respect to [18, 6] is the fact that we look for traveling quasi-periodic waves, instead of standing and reversible. The main novelty is providing that the terms that cannot be eliminated or normalized in the reducibility procedure of the linearized operator are constant coefficients operators. In the standing/reversible case this is guaranteed by parity conditions on the solutions that one look for, which reflect on parity properties of the non-constant coefficients of the linearized operator. In the present work we use the Hamiltonian structure and the momentum conservation to verify that such coefficients have nice properties, as for instance that the averages in time or space of some of them are constant. This clearly requires that all the changes of coordinates that we perform leave invariant those symmetries along the reducibility procedure. In other words, the reducibility of the linearized operator has to be performed within a class of Hamiltonian operators,  $x$ -translation invariant (to which it belongs the linearized operator at a quasi-periodic traveling function), see section 2.3.

---

<sup>4</sup>The equations considered in [40] cannot be considered as small perturbations of the integrable KdV equation. This can be seen as an evidence that the integrability does not play a role in [8].



To conclude, we mention the work [16]. The authors provide the existence of traveling quasi-periodic water waves for the gravity-capillary case with constant vorticity. Also in this case the solutions are reversible (parity conditions in the time-space variable) and the vorticity is used as a parameter to modulate the linear frequencies of oscillations.

*Possible extensions to other water waves problems.* It would be interesting to investigate the existence of quasi-periodic solutions for the models considered in [18, 6, 16] keeping *fixed* the parameters of the problems, respectively the capillarity, the depth of the ocean and the vorticity, and looking for solutions free from parity restrictions. Depending on the value of the fixed parameter, these equations present the same issues of a resonant case, like the pure gravity system in infinite depth. Our method could be applied to these cases provided that: (i) the normal form is formally integrable at order four, namely the coefficients of the Hamiltonian vanishes on non-trivial resonances (if they exist); (ii) we have some information on non-trivial resonances at order four; (iii) the twist condition holds, namely the frequency-amplitude map is a local diffeomorphism. All these ones are algebraic conditions that have to be checked for each choice of the fixed parameter. At the moment it is not clear to us whether these conditions are satisfied for some (large) set of the physical parameters.

Concerning the analysis of the linearized operator: the specific changes of coordinates that reduce the linear operator depend strongly on the considered model. However we believe that the reducibility strategy that we implement in the present work, made of an analysis of perturbative and non-perturbative terms, is quite general. Regarding the parity conditions, one should apply the reducibility argument in a class of linear operators with some algebraic structure dictated by the equation considered. We think that in water waves models is natural to exploit the Hamiltonian structure and the conservation of momentum. In this paper we provide a guideline to take advantage of these symmetries which are present also in the finite depth case and the gravity-capillary system.

**Plan of the paper.** In section 2 we introduce the functional setting and the classes of linear operators that we shall consider in the reducibility argument.

In section 3 we discuss the formal integrability of the pure gravity water waves at order four and we show how to exploit it to prove the normal form identification argument of section 3.4.

In section 4 we apply a “weak” version of a Birkhoff normal form procedure to the Hamiltonian (1.6) (see Proposition 4.3) and we show that there exists a finite dimensional approximately invariant subspace foliated by embedded tori. Then, in section 5, we prove that the dynamics on such subspace is *non-isochronous* (see the Hamiltonian system (5.2) and Lemma 5.2). In this way we find a first good nonlinear approximate solution. In section 5.2 we state the Theorem 5.6 from which we deduce the main result. This theorem is proved through a Nash-Moser nonlinear iteration, see Theorem 13.2.

In section 6 we prove Theorem 6.8 that concerns the inversion of the linearized operator at an approximate solution. Sections 8-12 are devoted to the verification of the assumptions of Theorem 6.8, in particular we prove the invertibility of the linearized operator in the normal directions, whose properties are discussed in section 7. The key results of sections 8-9 are Propositions 8.3, 9.1, 9.5. In section 10 we show how to reduce the linearized operator to a diagonal one up to a smoothing remainder, see Proposition 10.3. In section 12 we prove the reducibility Theorem 12.1 which implies the invertibility result in Proposition 12.2.

In section 11 we deal with the terms which still are not *perturbative* for the KAM scheme of Theorem 12.1. By Lemma 11.17 and Proposition 11.18 we find the first order corrections of the normal eigenvalues. Then we split the reduction of the remaining non perturbative terms by performing a *low / high* modes analysis.

We show that the analysis of high modes in section 11.3 does not involve small divisors. In section 11.4 we perform the analysis for low modes. In section 13 we conclude the proof of the main theorem by providing the measure estimates of the set of “good” frequencies.

**Acknowledgements.** The authors wish to thank Michela Procesi, Marcel Guardia and Raffaele Scandone for many useful discussions and comments. This paper has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 757802.

## 2. FUNCTIONAL SETTING

**2.1. Function spaces.** We consider functions  $u(\varphi, x)$  defined on  $\mathbb{T}^\nu \times \mathbb{T}$  with zero average in  $x \in \mathbb{T}$ . Passing to the Fourier representation we write

$$u(\varphi, x) = \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z} \setminus \{0\}} u_j(\varphi) e^{ijx} = \frac{1}{\sqrt{2\pi}} \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z} \setminus \{0\}} u_{\ell j} e^{i(\ell \cdot \varphi + jx)}. \quad (2.1)$$



For  $\varphi$ -independent Fourier coefficients we also use the notation

$$u_n^+ := u_n := \widehat{u}(n) \quad \text{and} \quad u_n^- := \overline{u_n} := \overline{\widehat{u}(n)}.$$

Let  $s \in \mathbb{R}$ . We define the scale of Sobolev spaces

$$H^s := \left\{ u(\varphi, x) \in L^2(\mathbb{T}^{\nu+1}; \mathbb{C}) : \|u\|_s^2 := \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z} \setminus \{0\}} |u_{\ell j}|^2 \langle \ell, j \rangle^{2s} < \infty \right\}, \quad (2.2)$$

where  $\langle \ell, j \rangle := \max\{1, |\ell|, |j|\}$ ,  $|\ell| := \sum_{i=1}^\nu |\ell_i|$ .

In the following we shall consider the scale of Sobolev spaces  $(H^s)_{s=s_0}^{\mathcal{S}}$  where

$$\mathcal{S} \gg s_0 := [\nu/2] + 2. \quad (2.3)$$

We remark that  $H^{s_0}$  is continuously embedded in  $L^\infty(\mathbb{T}^{\nu+1}; \mathbb{C})$ .

**Lipschitz norm.** Fix  $\nu \in \mathbb{N} \setminus \{0\}$  and let  $\mathcal{O}$  be a compact subset of  $\mathbb{R}^\nu$ . For a function  $u: \mathcal{O} \rightarrow E$ , where  $(E, \|\cdot\|_E)$  is a Banach space, we define the sup-norm and the lip-seminorm of  $u$  as

$$\|u\|_E^{sup} := \|u\|_E^{sup, \mathcal{O}} := \sup_{\omega \in \mathcal{O}} \|u(\omega)\|_E, \quad \|u\|_E^{lip} := \|u\|_E^{lip, \mathcal{O}} := \sup_{\substack{\omega_1, \omega_2 \in \mathcal{O}, \\ \omega_1 \neq \omega_2}} \frac{\|u(\omega_1) - u(\omega_2)\|_E}{|\omega_1 - \omega_2|}.$$

If  $E$  is finite dimensional, for any  $\gamma > 0$  we introduce the weighted Lipschitz norm

$$\|u\|_E^{\gamma, \mathcal{O}} := \|u\|_E^{sup, \mathcal{O}} + \gamma \|u\|_E^{lip, \mathcal{O}}.$$

If  $E$  is a scale of Banach spaces, say  $E = H^s$ , for  $\gamma > 0$  we introduce the weighted Lipschitz norm

$$\|u\|_s^{\gamma, \mathcal{O}} := \|u\|_s^{sup, \mathcal{O}} + \gamma \|u\|_{s-1}^{lip, \mathcal{O}}, \quad \forall s \geq s_0 \quad (2.4)$$

where we denoted by  $[r]$  the integer part of  $r \in \mathbb{R}$ .

**2.2. Linear operators.** We introduce general classes of linear operators which will be used in the paper following [32]. We consider  $\varphi$ -dependent families of operators  $A: \mathbb{T}^\nu \mapsto \mathcal{L}(L^2(\mathbb{T}; \mathbb{R}))$  (or  $\mathcal{L}(L^2(\mathbb{T}; \mathbb{C}))$ )  $\varphi \mapsto A(\varphi)$  acting on functions  $u(x) \in L^2(\mathbb{T}; \mathbb{R})$  (or  $L^2(\mathbb{T}; \mathbb{C})$ ). We also regard  $A$  as an operator acting on functions  $u(\varphi, x) \in L^2(\mathbb{T}^{\nu+1}; \mathbb{R})$  (or  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C})$ ) of the space-time. In other words the action of  $A \in \mathcal{L}(L^2(\mathbb{T}^{\nu+1}; \mathbb{R}))$  is defined by

$$(Au)(\varphi, x) = (A(\varphi)u(\varphi, \cdot))(x).$$

We represent a linear operator  $\mathcal{Q}$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{R}^2)$  by a matrix

$$\mathcal{Q} := \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad (2.5)$$

where  $A, B, C, D$  are operators acting on the scalar valued components  $\eta, \psi \in L^2(\mathbb{T}^{\nu+1}; \mathbb{R})$ . The action of  $A$  on periodic functions as in (2.1) is given by

$$(Au)(\varphi, x) = \frac{1}{\sqrt{2\pi}} \sum_{l \in \mathbb{Z}^\nu, j \in \mathbb{Z}} e^{il \cdot \varphi} e^{ijx} \left( \sum_{p \in \mathbb{Z}^\nu, k \in \mathbb{Z}} A_j^k(l-p) \widehat{u}(p, k) \right). \quad (2.6)$$

We shall identify the operator  $A$  with the matrix  $(A_j^k(\ell))_{\ell \in \mathbb{Z}^\nu, j, k \in \mathbb{Z}}$ . We say that  $A$  is a *real* operator if it maps real valued functions in real valued functions. For the matrix coefficients this means that

$$\overline{A_j^k(\ell)} = A_{-j}^{-k}(-\ell).$$

It will be convenient to work with the the complex variables  $(u, \bar{u}) = \mathcal{C}(\eta, \psi)$  introduced in (3.3). Thus we study how the linear operators  $\mathcal{Q}$  as in (2.5) transform under the map  $\mathcal{C}$ . We have

$$\begin{aligned} \mathcal{T} &:= \mathcal{C}\mathcal{Q}\mathcal{C}^{-1} := (\mathcal{T}_{\sigma'}^{\sigma'})_{\sigma, \sigma' = \pm} := \begin{pmatrix} \mathcal{T}_+^+ & \mathcal{T}_+^- \\ \mathcal{T}_-^+ & \mathcal{T}_-^- \end{pmatrix}, \quad \mathcal{T}_{\sigma'}^{\sigma'} = \overline{\mathcal{T}_{-\sigma'}^{-\sigma'}}, \quad \sigma, \sigma' = \pm \\ \mathcal{T}_+^+ &:= \frac{1}{2} \{(A+D) - i(B-C)\}, \quad \mathcal{T}_+^- := \frac{1}{2} \{(A-D) + i(B+C)\}, \end{aligned} \quad (2.7)$$

where the *conjugate* operator  $\overline{\mathcal{T}_{\sigma'}^{\sigma'}}$  is defined as

$$\overline{\mathcal{T}_{\sigma'}^{\sigma'}[h]} := \overline{\mathcal{T}_{\sigma'}^{\sigma'}[\overline{h}]}. \quad (2.8)$$

By setting  $u^+ = u$ ,  $u^- = \bar{u}$  for any  $u \in L^2(\mathbb{T}^{\nu+1}; \mathbb{C})$ , the action of an operator  $\mathcal{T}$  of the form (2.7) is given by

$$(\mathcal{T}[\frac{u}{\bar{u}}])(\varphi, x) = \begin{pmatrix} \mathcal{T}_+^+ u^+ + \mathcal{T}_+^- u^- \\ \mathcal{T}_-^+ u^+ + \mathcal{T}_-^- u^- \end{pmatrix} \stackrel{(2.7)}{=} \begin{pmatrix} \mathcal{T}_+^+ u^+ + \mathcal{T}_+^- u^- \\ \mathcal{T}_-^+ u^+ + \mathcal{T}_-^- u^- \end{pmatrix}. \quad (2.9)$$

By formulæ (2.6), (2.9) we shall identify the operator  $\mathcal{T}$  with the matrix  $(\mathcal{T}_{\sigma,j}^{\sigma',k}(\ell))_{\ell \in \mathbb{Z}^\nu, j,k \in \mathbb{Z}, \sigma, \sigma' = \pm}$ , where

$$\mathcal{T}_{\sigma,j}^{\sigma',k}(\ell) := \int_{\mathbb{T}^{\nu+1}} \mathcal{T}_\sigma^{\sigma'} [e^{i(\sigma'k - \sigma j)x} e^{-i\ell \cdot \varphi}] dx d\varphi.$$

**Definition 2.1. (Real-to-real).** We say that an operator  $\mathcal{T}$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2)$  of the form (2.7) is real-to-real. Equivalently an operator is real-to-real if it preserves the following subspace

$$\mathcal{U} := \{ (u, v) \in L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2) : v = \bar{u} \}.$$

**Remark 2.2.** Notice that the conjugate of a real operator as in (2.5) under the map  $\Lambda$  in (3.4) has the form (2.7) and hence it is real-to-real.

2.2.1. *Tame operators.* We use the notation  $A \lesssim B$  to denote  $A \leq CB$  where  $C$  is a positive constant possibly depending on fixed parameters given by the problem. We use the notation  $A \lesssim_y B$  to denote  $A \leq C(y)B$  if we wish to highlight the dependence on the variable  $y$  of the constant  $C(y) > 0$ .

We use the notation  $\mathcal{O} \Subset \mathbb{R}^\nu$  to denote a compact subset of  $\mathbb{R}^\nu$ .

**Linear Tame operators.** Here we introduce rigorously the spaces and the classes of operators on which we work.

**Definition 2.3. ( $\sigma$ -Tame operators).** Let  $0 < s_0 < \mathcal{S} \leq \infty$  be two integer numbers and  $\sigma \geq 0$ . We say that a linear operator  $A$  is  $\sigma$ -tame w.r.t. a non-decreasing sequence of positive real numbers  $\{\mathfrak{M}_A(\sigma, s)\}_{s=s_0}^{\mathcal{S}}$  if (recall (2.2))

$$\|Au\|_s \leq \mathfrak{M}_A(\sigma, s) \|u\|_{s_0+\sigma} + \mathfrak{M}_A(\sigma, s_0) \|u\|_{s+\sigma} \quad u \in H^s,$$

for any  $s_0 \leq s \leq \mathcal{S}$ . We call  $\mathfrak{M}_A(\sigma, s)$  a TAME CONSTANT for the operator  $A$ . When the index  $\sigma$  is not relevant we write  $\mathfrak{M}_A(\sigma, s) = \mathfrak{M}_A(s)$ . For an operator  $\mathcal{T}$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2)$  of the form (2.7) we set

$$\mathfrak{M}_{\mathcal{T}}(\sigma, s) := \max\{\mathfrak{M}_{\mathcal{T}_+}(\sigma, s), \mathfrak{M}_{\mathcal{T}_-}(\sigma, s)\}.$$

**Definition 2.4. (Lip- $\sigma$ -Tame operators).** Let  $\sigma \geq 0$  and  $A = A(\omega)$  be a linear operator defined for  $\omega \in \mathcal{O} \Subset \mathbb{R}^\nu$ . Let us define

$$\Delta_{\omega, \omega'} A := \frac{A(\omega) - A(\omega')}{|\omega - \omega'|}, \quad \omega, \omega' \in \mathcal{O}, \quad \omega \neq \omega'. \quad (2.10)$$

Then  $A$  is Lip- $\sigma$ -tame w.r.t. a non-decreasing sequence  $\{\mathfrak{M}_A^\gamma(\sigma, s)\}_{s=s_0}^{\mathcal{S}}$  if the following estimate holds

$$\sup_{\omega \in \mathcal{O}} \|Au\|_s, \quad \gamma \sup_{\omega \neq \omega'} \|(\Delta_{\omega, \omega'} A)\|_{s-1} \leq_s \mathfrak{M}_A^\gamma(\sigma, s) \|u\|_{s_0+\sigma} + \mathfrak{M}_A^\gamma(\sigma, s) \|u\|_{s+\sigma}, \quad u \in H^s.$$

We call  $\mathfrak{M}_A^\gamma(\sigma, s)$  a LIP-TAME CONSTANT of the operator  $A$ . When the index  $\sigma$  is not relevant we write  $\mathfrak{M}_A^\gamma(\sigma, s) = \mathfrak{M}_A^\gamma(s)$ .

**Modulo-tame operators and majorant norms.** We now introduce a class of operators which are tame respect to a norm stronger than the operator norm.

**Definition 2.5. (Majorant operator).** Let  $s \in \mathbb{R}$  and  $u \in H^s$ , we define the majorant function

$$\underline{u}(\varphi, x) := (2\pi)^{-1/2} \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z} \setminus \{0\}} |u_{\ell j}| e^{i(\ell \cdot \varphi + jx)}.$$

Note that  $\|u\|_s = \|\underline{u}\|_s$ . Let  $A \in \mathcal{L}(H^s)$  and recall its matrix representation (2.6). We define the majorant matrix  $\underline{A}$  as the matrix with entries

$$(\underline{A})_j^k(\ell) := |(A)_j^k(\ell)| \quad j, k \in \mathbb{Z} \setminus \{0\}, \quad \ell \in \mathbb{Z}^\nu.$$

We consider the majorant operator norms

$$\|\underline{A}\|_{\mathcal{L}(H^s)} := \sup_{\|u\|_s \leq 1} \|\underline{A}u\|_s.$$

We have a partial ordering relation in the set of the infinite dimensional matrices, i.e. if

$$A \preceq B \Leftrightarrow |A_j^k(\ell)| \leq |B_j^k(\ell)| \quad \forall j, k, \ell \Rightarrow \|\underline{A}\|_{\mathcal{L}(H^s)} \leq \|\underline{B}\|_{\mathcal{L}(H^s)}, \quad \|Au\|_s \leq \|\underline{A}u\|_s \leq \|\underline{B}u\|_s. \quad (2.11)$$

Since we are working on a majorant norm we have the continuity of the projections on monomial subspace, in particular we define the following functor acting on the matrices

$$\Pi_K A := \begin{cases} A_j^k(\ell) & \text{if } |\ell| \leq K, \\ 0 & \text{otherwise} \end{cases}, \quad \Pi_K^\perp := I - \Pi_K. \quad (2.12)$$

Finally we define for  $\mathbf{b} \in \mathbb{N}$  the operator  $\langle \partial_\varphi \rangle^{\mathbf{b}} A$  whose matrix element are

$$\langle \langle \partial_\varphi \rangle^{\mathbf{b}} A \rangle_j^k(\ell) := \langle \ell \rangle^{\mathbf{b}} A_j^k(\ell). \quad (2.13)$$

**Definition 2.6. ( $\sigma$ -modulo tame operators).** Let  $0 < s_0 < \mathcal{S} \leq \infty$  be two integer numbers and  $\sigma \geq 0$ . We say that a linear operator  $A$  is  $\sigma$ -modulo tame w.r.t. a non-decreasing sequence of positive real numbers  $\{\mathfrak{M}_A^\sharp(\sigma, s)\}_{s=s_0}^{\mathcal{S}}$  if (recall (2.2))

$$\|\underline{A}u\|_s \leq \mathfrak{M}_A^\sharp(\sigma, s)\|u\|_{s_0+\sigma} + \mathfrak{M}_A^\sharp(\sigma, s_0)\|u\|_{s+\sigma} \quad u \in H^s,$$

for any  $s_0 \leq s \leq \mathcal{S}$ . We call  $\mathfrak{M}_A^\sharp(\sigma, s)$  a MODULO TAME CONSTANT for the operator  $A$ . When the index  $\sigma$  is not relevant we write  $\mathfrak{M}_A^\sharp(\sigma, s) = \mathfrak{M}_A^\sharp(s)$ . For an operator  $\mathcal{T}$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2)$  of the form (2.7) we set

$$\mathfrak{M}_{\mathcal{T}}^\sharp(\sigma, s) := \max\{\mathfrak{M}_{\mathcal{T}^+}^\sharp(\sigma, s), \mathfrak{M}_{\mathcal{T}^-}^\sharp(\sigma, s)\}.$$

**Definition 2.7.** We say that  $A$  is  $(-1/2)$ -modulo tame if  $\langle D \rangle^{1/4} A \langle D \rangle^{1/4}$  is 0-modulo tame, where  $\langle D \rangle$  is the Fourier multiplier with symbol  $\langle j \rangle$ . We denote

$$\begin{aligned} \mathfrak{M}_A^\sharp(-1/2, s) &:= \mathfrak{M}_{\langle D \rangle^{1/4} A \langle D \rangle^{1/4}}^\sharp(0, s), \\ \mathfrak{M}_A^\sharp(-1/2, s, a) &:= \mathfrak{M}_{\langle \partial_\varphi \rangle^a \langle D \rangle^{1/4} A \langle D \rangle^{1/4}}^\sharp(0, s), \quad a \geq 0. \end{aligned}$$

**Definition 2.8. (Lip- $\sigma$ -modulo tame).** Let  $\sigma \geq 0$ . A linear operator  $A := A(\omega)$ ,  $\omega \in \mathcal{O} \Subset \mathbb{R}^\nu$ , is Lip- $\sigma$ -modulo-tame w.r.t. a non-decreasing sequence  $\{\mathfrak{M}_A^{\sharp, \gamma}(\sigma, s)\}_{s=s_0}^{\mathcal{S}}$  if the majorant operators  $\underline{A}$ ,  $\underline{\Delta_{\omega, \omega'} A}$  (see (2.10)) are Lip- $\sigma$ -tame w.r.t. these constants, i.e. they satisfy the following weighted tame estimates: for all  $s \geq s_0$  and for any  $u \in H^s$ ,

$$\sup_{\omega \in \mathcal{O}} \|\underline{A}u\|_s, \quad \sup_{\omega \neq \omega' \in \mathcal{O}} \gamma \|\underline{\Delta_{\omega, \omega'} A}u\|_s \leq \mathfrak{M}_A^{\sharp, \gamma}(\sigma, s_0)\|u\|_{s+\sigma} + \mathfrak{M}_A^{\sharp, \gamma}(\sigma, s)\|u\|_{s_0+\sigma}.$$

The constant  $\mathfrak{M}_A^{\sharp, \gamma}(\sigma, s)$  is called the MODULO-TAME CONSTANT of the operator  $A$ . When the index  $\sigma$  is not relevant we write  $\mathfrak{M}_A^{\sharp, \gamma}(\sigma, s) = \mathfrak{M}_A^{\sharp, \gamma}(s)$ .

**Definition 2.9.** We say that  $A$  is Lip- $(-1/2)$ -modulo tame if  $\langle D \rangle^{1/4} A \langle D \rangle^{1/4}$  is Lip-0-modulo tame, where  $\langle D \rangle$  is the Fourier multiplier with symbol  $\langle j \rangle$ . We denote

$$\begin{aligned} \mathfrak{M}_A^{\sharp, \gamma}(-1/2, s) &:= \mathfrak{M}_{\langle D \rangle^{1/4} A \langle D \rangle^{1/4}}^{\sharp, \gamma}(0, s), \\ \mathfrak{M}_A^{\sharp, \gamma}(-1/2, s, a) &:= \mathfrak{M}_{\langle \partial_\varphi \rangle^a \langle D \rangle^{1/4} A \langle D \rangle^{1/4}}^{\sharp, \gamma}(0, s), \quad a \geq 0. \end{aligned}$$

We refer the reader to the Appendix of [31], [32] for the properties of tame and modulo-tame operators. In particular we shall apply several times the results of Lemma A.5 in [31]. We remark that in order to estimate the tame constants of operators  $A$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2)$  we shall use these results for each entry of the matrix  $A$ .

**Classes of “smoothing” operators.** The following class of smoothing (in space) operators has been introduced in [31], [32].

**Definition 2.10.** Let  $p \in \mathbb{N}$  with  $s_0 \leq p < \mathcal{S} \leq \infty$  (see (2.3)). Fix  $\rho \in \mathbb{N}$ , with  $\rho \geq 3$  and consider  $\mathcal{O} \Subset \mathbb{R}^\nu$ . We denote by  $\mathfrak{L}_{\rho, p} = \mathfrak{L}_{\rho, p}(\mathcal{O})$  the set of the linear operators  $A = A(\omega): H^s \rightarrow H^s$ ,  $\omega \in \mathcal{O}$  with the following properties:

- the operator  $A$  is Lipschitz in  $\omega$ ,
- the operators  $\partial_\varphi^{\vec{\mathbf{b}}} A$ ,  $[\partial_\varphi^{\vec{\mathbf{b}}} A, \partial_x]$ , for all  $\vec{\mathbf{b}} = (\mathbf{b}_1, \dots, \mathbf{b}_\nu) \in \mathbb{N}^\nu$  with  $0 \leq |\vec{\mathbf{b}}| \leq \rho - 2$  have the following properties, for any  $s_0 \leq s < \mathcal{S}$ :

- for any  $m_1, m_2 \in \mathbb{R}$ ,  $m_1, m_2 \geq 0$  and  $m_1 + m_2 = \rho - |\vec{\mathbf{b}}|$  one has that  $\langle D \rangle^{m_1} \partial_\varphi^{\vec{\mathbf{b}}} A \langle D \rangle^{m_2}$  is Lip-0-tame according to Definition 2.4 and we set

$$\mathfrak{M}_{\partial_\varphi^{\vec{\mathbf{b}}} A}^\gamma(-\rho + |\vec{\mathbf{b}}|, s) := \sup_{\substack{m_1 + m_2 = \rho - |\vec{\mathbf{b}}| \\ m_1, m_2 \geq 0}} \mathfrak{M}_{\langle D \rangle^{m_1} \partial_\varphi^{\vec{\mathbf{b}}} A \langle D \rangle^{m_2}}^\gamma(0, s);$$

(ii) for any  $m_1, m_2 \in \mathbb{R}$ ,  $m_1, m_2 \geq 0$  and  $m_1 + m_2 = \rho - |\vec{\mathbf{b}}| - 1$  one has that  $\langle D \rangle^{m_1} [\partial_{\varphi}^{\vec{\mathbf{b}}} A, \partial_x] \langle D \rangle^{m_2}$  is Lip-0-tame and we set

$$\mathfrak{M}_{[\partial_{\varphi}^{\vec{\mathbf{b}}} A, \partial_x]}^{\gamma}(-\rho + |\vec{\mathbf{b}}| + 1, s) := \sup_{\substack{m_1 + m_2 = \rho - |\vec{\mathbf{b}}| - 1 \\ m_1, m_2 \geq 0}} \mathfrak{M}_{\langle D \rangle^{m_1} [\partial_{\varphi}^{\vec{\mathbf{b}}} A, \partial_x] \langle D \rangle^{m_2}}^{\gamma}(0, s).$$

We define for  $0 \leq \mathbf{b} \leq \rho - 2$

$$\mathbb{M}_A^{\gamma}(s, \mathbf{b}) := \max_{0 \leq |\vec{\mathbf{b}}| \leq \mathbf{b}} \max \left( \mathfrak{M}_{\partial_{\varphi}^{\vec{\mathbf{b}}} A}^{\gamma}(-\rho + |\vec{\mathbf{b}}|, s), \mathfrak{M}_{\partial_{\varphi}^{\vec{\mathbf{b}}} [A, \partial_x]}^{\gamma}(-\rho + |\vec{\mathbf{b}}| + 1, s) \right).$$

Let  $s \geq p$  and consider  $i_1 = i_1(\omega)$ ,  $i_2 = i_2(\omega)$ , defined for  $\omega \in \mathcal{O}_1$  and  $\mathcal{O}_2$  respectively, such that  $\|i_k\|_s^{\gamma \cdot \mathcal{O}_k} \lesssim 1$ ,  $k = 1, 2$ . Consider an operator  $A = A(\omega, i_k(\omega))$  defined for  $\omega \in \mathcal{O}(i_k) \subseteq \mathcal{O}_k$ ,  $k = 1, 2$ , and we define

$$\Delta_{12} A = \Delta_{12} A(\omega, i_1(\omega), i_2(\omega)) := A(i_1) - A(i_2), \quad \text{for } \omega \in \mathcal{O}(i_1) \cap \mathcal{O}(i_2).$$

We require the following:

• The operators  $\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A$ ,  $[\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A, \partial_x]$ , for  $0 \leq |\vec{\mathbf{b}}| \leq \rho - 3$ , have the following properties:

(iii) for any  $m_1, m_2 \in \mathbb{R}$ ,  $m_1, m_2 \geq 0$  and  $m_1 + m_2 = \rho - |\vec{\mathbf{b}}| - 1$  one has that  $\langle D \rangle^{m_1} \partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A \langle D \rangle^{m_2}$  is bounded on  $H^p$ . More precisely there is a positive constant  $\mathfrak{N}_{\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A}(-\rho + |\vec{\mathbf{b}}| + 1, p)$  such that, for any  $h \in H^p$ , we have

$$\sup_{\substack{m_1 + m_2 = \rho - |\vec{\mathbf{b}}| - 1 \\ m_1, m_2 \geq 0}} \|\langle D \rangle^{m_1} \partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A \langle D \rangle^{m_2} h\|_p \leq \mathfrak{N}_{\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A}(-\rho + |\vec{\mathbf{b}}| + 1, p) \|h\|_p$$

uniformly for  $\omega \in \mathcal{O}(i_1) \cap \mathcal{O}(i_2)$ .

(iv) for any  $m_1, m_2 \in \mathbb{R}$ ,  $m_1, m_2 \geq 0$  and  $m_1 + m_2 = \rho - |\vec{\mathbf{b}}| - 2$  one has that  $\langle D \rangle^{m_1} [\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A, \partial_x] \langle D \rangle^{m_2}$  is bounded on  $H^p$ . More precisely there is a positive constant  $\mathfrak{N}_{[\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A, \partial_x]}(-\rho + |\vec{\mathbf{b}}| + 2, p)$  such that for any  $h \in H^p$  one has

$$\sup_{\substack{m_1 + m_2 = \rho - |\vec{\mathbf{b}}| - 2 \\ m_1, m_2 \geq 0}} \|\langle D \rangle^{m_1} [\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A, \partial_x] \langle D \rangle^{m_2} h\|_p \leq \mathfrak{N}_{[\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A, \partial_x]}(-\rho + |\vec{\mathbf{b}}| + 2, p) \|h\|_p$$

uniformly for  $\omega \in \mathcal{O}(i_1) \cap \mathcal{O}(i_2)$ .

We define for  $0 \leq \mathbf{b} \leq \rho - 3$

$$\mathbb{M}_{\Delta_{12} A}(p, \mathbf{b}) := \max_{0 \leq |\vec{\mathbf{b}}| \leq \mathbf{b}} \max \left( \mathfrak{N}_{\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A}(-\rho + |\vec{\mathbf{b}}| + 1, p), \mathfrak{N}_{[\partial_{\varphi}^{\vec{\mathbf{b}}} \Delta_{12} A, \partial_x]}(-\rho + |\vec{\mathbf{b}}| + 2, p) \right).$$

By construction one has that  $\mathbb{M}_A^{\gamma}(s, \mathbf{b}_1) \leq \mathbb{M}_A^{\gamma}(s, \mathbf{b}_2)$  if  $0 \leq \mathbf{b}_1 \leq \mathbf{b}_2 \leq \rho - 2$  and  $\mathbb{M}_{\Delta_{12} A}(p, \mathbf{b}_1) \leq \mathbb{M}_{\Delta_{12} A}(p, \mathbf{b}_2)$  if  $0 \leq \mathbf{b}_1 \leq \mathbf{b}_2 \leq \rho - 3$ .

For the properties of operators in the class  $\mathfrak{L}_{\rho, p}$  we refer the reader to Appendix B in [31].

2.2.2. *pseudo differential operators and symbolic calculus.* Following [18] we give the following definitions.

**Definition 2.11. (Symbols).** Let  $m \in \mathbb{R}$ . We define the class  $S^m$  as the set of symbols  $a(x, j)$  which are the restriction to  $\mathbb{T} \times \mathbb{Z}$  of a complex valued function  $a(x, \xi)$  which is  $C^\infty$  smooth on  $\mathbb{T} \times \mathbb{R}$ ,  $2\pi$ -periodic in  $x$  and satisfies

$$|\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \leq C_{\alpha, \beta} \langle \xi \rangle^{m - \beta}, \quad \forall \alpha, \beta \in \mathbb{N}. \quad (2.14)$$

**Definition 2.12. (Pseudo differential operators).** Let  $m \in \mathbb{R}$ . A linear operator  $A$  is called pseudo differential of order  $\leq m$  if its action on any  $H^s(\mathbb{T}; \mathbb{C})$  with  $s \geq m$  is given by

$$Au := Op^W(a(x, \xi))[u] := \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \left( \sum_{j \in \mathbb{Z} \setminus \{0\}} \widehat{a}\left(k - j, \frac{k + j}{2}\right) u_j \right) \frac{e^{ikx}}{\sqrt{2\pi}}, \quad u = \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z} \setminus \{0\}} u_j e^{ijx}, \quad (2.15)$$

where  $a(x, \xi)$  is a symbol in the class  $S^m$  (see Def. 2.11) and  $\widehat{a}(\eta, \xi)$  denotes the Fourier transform of  $a(x, \xi)$  in the variable  $x \in \mathbb{T}$ . We shall also write  $A[\cdot] = Op^W(a)[\cdot] = Op^W(a(x, \xi))[\cdot]$ . We call  $OPS^m$  the class of the pseudo differential operator of order less or equal to  $m$  and  $OPS^{-\infty} := \bigcap_m OPS^m$ .

We will consider mainly operators acting on  $H^s(\mathbb{T}; \mathbb{C})$  with a quasi-periodic time dependence. In the case of pseudo differential operators this corresponds<sup>5</sup> to considering symbols  $a(\varphi, x, \xi)$  with  $\varphi \in \mathbb{T}^\nu$ . Clearly these operators can be thought as acting on functions  $u(\varphi, x) = \sum_{j \in \mathbb{Z}} u_j(\varphi) e^{ijx}$  in  $H^s(\mathbb{T}^{\nu+1}; \mathbb{C})$  in the following sense:

$$(Au)(\varphi, x) = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \left( \sum_{j \in \mathbb{Z}} \widehat{a}\left(\varphi, k-j, \frac{k+j}{2}\right) u_j(\varphi) \right) \frac{e^{ikx}}{\sqrt{2\pi}}, \quad a(\varphi, x, j) \in S^m.$$

The symbol  $a(\varphi, x, \xi)$  is  $C^\infty$  smooth also in the variable  $\varphi$ . We still denote  $A := A(\varphi) = Op^W(a(\varphi, \cdot)) = Op^W(a)$ .

**Definition 2.13.** Let  $a := a(\varphi, x, \xi) \in S^m$  and set  $A := Op^W(a) \in OPS^m$ ,

$$|A|_{m,s,\alpha} := \max_{0 \leq \beta \leq \alpha} \sup_{\xi \in \mathbb{R}} \|\partial_\xi^\beta a(\cdot, \cdot, \xi)\|_s \langle \xi \rangle^{-m+\beta}. \quad (2.16)$$

We will use also the notation  $|a|_{m,s,\alpha} := |A|_{m,s,\alpha}$ .

Note that the norm  $|\cdot|_{m,s,\alpha}$  is non-decreasing in  $s$  and  $\alpha$ . Moreover given a symbol  $a(\varphi, x)$  independent of  $\xi \in \mathbb{R}$ , the norm of the associated multiplication operator  $Op^W(a)$  is just the  $H^s$  norm of the function  $a$ . If on the contrary the symbol  $a(\xi)$  depends only on  $\xi$ , then the norm of the corresponding Fourier multipliers  $Op^W(a(\xi))$  is just controlled by a constant. We shall use the following notation, used also in [1]. For any  $m \in \mathbb{R} \setminus \{0\}$  we set

$$|D|^m := Op^W(\chi(\xi)|\xi|^m), \quad (2.17)$$

where  $\chi \in C^\infty(\mathbb{R}, \mathbb{R})$  is an even and positive cut-off function such that

$$\chi(\xi) = \begin{cases} 0 & \text{if } |\xi| \leq 1/3 \\ 1 & \text{if } |\xi| \geq 2/3 \end{cases} \quad \partial_\xi \chi(\xi) > 0 \quad \forall \xi \in \left(\frac{1}{3}, \frac{2}{3}\right). \quad (2.18)$$

To simplify the notation we shall also write

$$Op^W(\text{sign}(\xi)) \quad \text{to denote} \quad Op^W(\chi(\xi)\text{sign}(\xi)), \quad \text{where} \quad \text{sign}(\xi) := \xi/|\xi|. \quad (2.19)$$

As in formula (2.4), if  $A = Op^W(a(\omega, \varphi, x, \xi)) \in OPS^m$  is a family of pseudo differential operators with symbols  $a(\omega, \varphi, x, \xi)$  belonging to  $S^m$  and depending in a Lipschitz way on some parameter  $\omega \in \mathcal{O} \subset \mathbb{R}^\nu$ , we set

$$|A|_{m,s,\alpha}^{\mathcal{O}} := \sup_{\omega \in \mathcal{O}} |A|_{m,s,\alpha} + \gamma \sup_{\omega_1, \omega_2 \in \mathcal{O}} \frac{|Op^W(a(\omega_1, \varphi, x, \xi) - a(\omega_2, \varphi, x, \xi))|_{m,s-1,\alpha}}{|\omega_1 - \omega_2|}.$$

For the properties of compositions, adjointness and quantitative estimates of the actions on the Sobolev spaces  $H^s$  of pseudo differential operators we refer to Appendix B of [32].

From now on we consider symbols  $a(\varphi, x, \xi) = a(\omega, i(\omega); \varphi, x, \xi)$  where  $\omega \in \mathcal{O} \Subset \mathbb{R}^\nu$  and  $i = i(\omega)$  is a Lipschitz function such that  $\|i\|_{s_0}^{\gamma, \mathcal{O}} \leq 1$  for some  $s_0 \geq [\nu/2] + 2$ .

Let us consider two subsets  $\mathcal{O}_1, \mathcal{O}_2 \Subset \mathbb{R}^\nu$  and two Lipschitz functions  $i_1(\omega), i_2(\omega)$  defined respectively on  $\mathcal{O}_1, \mathcal{O}_2$ . We denote by  $\Delta_{12}a = \Delta_{12}a(\varphi, x, \xi)$  the symbol

$$\Delta_{12}a = \Delta_{12}a(\omega, i_1(\omega), i_2(\omega)) := a(\omega, i_1(\omega)) - a(\omega, i_2(\omega))$$

defined for  $\omega \in \mathcal{O}_1 \cap \mathcal{O}_2$ . To simplify the notation we shall omit the dependence on  $(\omega, i(\omega))$  in symbols and operators.

**Weyl/Standard quantizations.** The quantization of a symbol in  $S^m$  given in (2.15) is called *Weyl* quantization. We shall compare this quantization with the *standard one*, that we now recall: given a symbol  $b(x, j) \in S^m$ , the associated pseudo differential operator is defined as  $B := Op(b)$  with

$$Op(b) \sum_{j \in \mathbb{Z} \setminus \{0\}} u_j \frac{e^{ijx}}{\sqrt{2\pi}} = \sum_{j \in \mathbb{Z} \setminus \{0\}} b(x, j) u_j \frac{e^{ijx}}{\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z}} \left( \sum_{j \in \mathbb{Z} \setminus \{0\}} \widehat{b}(k-j, j) u_j \right) \frac{e^{ikx}}{\sqrt{2\pi}}. \quad (2.20)$$

It is known (see paragraph 18.5 in [41]) that the two quantizations are equivalent. Moreover one can transform the symbols between different quantizations by using the formulæ

$$Op^W(a) = Op(b) \quad \widehat{a}(k, j) = \widehat{b}\left(k, j - \frac{k}{2}\right). \quad (2.21)$$

<sup>5</sup>Since  $\omega$  is diophantine we can replace the time variable with angles  $\varphi \in \mathbb{T}^\nu$ . The time dependence is recovered by setting  $\varphi = \omega t$ .

Along the paper we shall use only the Weyl quantization, which is convenient in a Hamiltonian setting. We use the equivalence with the standard one to recover some known results of pseudo differential calculus on tori that have been proved in the standard case. We mostly refer to the results in [31, 32] where symbolic calculus is provided with sharp estimates on the seminorms of the symbols.

We shall need an asymptotic expansion of the symbol  $a$  in terms of the symbol  $b$ . The distributional kernel of the operator  $Op^W(a)$  is given by the oscillatory integral

$$K(x, y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i(x-y)\xi} a\left(\frac{x+y}{2}, \xi\right) d\xi. \quad (2.22)$$

Hence we may recover the Weyl symbol  $a$  from the kernel by the inverse formula

$$a(x, \xi) = \int_{\mathbb{R}} K\left(x + \frac{t}{2}, x - \frac{t}{2}\right) e^{-it\xi} dt. \quad (2.23)$$

If we consider the standard quantization  $Op(b) = Op^W(a)$  we can express the distributional kernel in terms of  $b$

$$K(x, y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{i(x-y)\xi} b(x, \xi) d\xi.$$

Then by (2.22) and (2.23) we obtain an expression of the symbol  $a$  in terms of  $b$

$$a(x, \xi) = \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{-iz\xi} b\left(x + \frac{z}{2}, \xi - \zeta\right) dz d\zeta \quad (2.24)$$

and viceversa

$$b(x, \xi) = \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{-iz\xi} a\left(x - \frac{z}{2}, \xi - \zeta\right) dz d\zeta. \quad (2.25)$$

We have the following important Lemma (see e.g. Lemma 3.5 in [12]).

**Lemma 2.14.** *Let  $b(\varphi, x, \xi) \in S^m$  and consider the function  $a(\varphi, x, \xi)$  in (2.24). Then for any  $N \geq 1$  we have the expression*

$$a(\varphi, x, \xi) = \sum_{p=0}^{N-1} \frac{(-1)^p}{2^p p!} (\partial_x^p D_\xi^p b)(\varphi, x, \xi) + \tilde{a}(\varphi, x, \xi)$$

where  $D_\xi = -i\partial_\xi$  and  $\tilde{a}$  satisfies the following bounds

$$\begin{aligned} |\tilde{a}|_{m-N, s, \alpha}^{\gamma, \mathcal{O}} &\lesssim |b|_{m, s+s_0+N, \alpha+2+2N}^{\gamma, \mathcal{O}}, \\ |\Delta_{12}\tilde{a}|_{m-N, p, \alpha} &\lesssim |\Delta_{12}b|_{m, p+s_0+N, \alpha+2+2N}. \end{aligned} \quad (2.26)$$

*Proof.* The proof follows word by word the proof of Lemma 3.5 in [12] taking into account the norm (2.16) to estimate the regularity of the symbol. The bound for the variations  $\Delta_{\omega, \omega'} \tilde{a}$ ,  $\Delta_{12} \tilde{a}$  follows by repeating the proof with  $a \rightsquigarrow \Delta_{\omega, \omega'} a$ ,  $\Delta_{12} a$ ,  $b \rightsquigarrow \Delta_{\omega, \omega'} b$ ,  $\Delta_{12} b$ .  $\square$

As a consequence of Lemma 2.14 we have the following.

**Lemma 2.15.** *Fix  $\rho, p$  as in Definition 2.10. Let  $m \in \mathbb{R}$ ,  $N := m + \rho$  and consider a symbol  $b(\varphi, x, \xi)$  in  $S^m$ . Then there exists a remainder  $R_\rho \in \mathfrak{L}_{\rho, p}$  such that*

$$Op(b) = Op^W(c) + R_\rho, \quad c(\varphi, x, \xi) := \sum_{p=0}^{N-1} \frac{(-1)^p}{2^p p!} (\partial_x^p D_\xi^p b)(\varphi, x, \xi) \quad (2.27)$$

with

$$|c|_{m, s, \alpha}^{\gamma, \mathcal{O}} \leq_{s, \rho} |b|_{m, s+N, \alpha+N}^{\gamma, \mathcal{O}}, \quad \mathbb{M}_{R_\rho}^{\tilde{\gamma}}(s, \mathbf{b}) \leq_{s, \rho} |b|_{m, s+N, N}^{\gamma, \mathcal{O}}, \quad (2.28)$$

for all  $0 \leq \mathbf{b} \leq \rho - 2$  and  $s_0 \leq s \leq \mathcal{S}$ . Moreover one has

$$|\Delta_{12}c|_{m+m', p, \alpha} \leq_{p, \rho} |\Delta_{12}b|_{m, s+N, \alpha+N}, \quad \mathbb{M}_{\Delta_{12}R_\rho}(p, \mathbf{b}) \leq_{p, \rho} |\Delta_{12}b|_{m, s+N, N}, \quad (2.29)$$

for all  $0 \leq \mathbf{b} \leq \rho - 3$  and where  $p$  is the constant given in Definition 2.10.

*Proof.* Let  $a(\varphi, x, \xi)$  be a symbol such that  $Op^W(a) = Op(b)$ . We apply Lemma 2.14. and we set  $R_\rho := Op^W(\tilde{a})$ . The bounds (2.28) and (2.29) on the symbol  $c(\varphi, x, \xi)$  follow by an explicit computation. The estimates on the remainder  $R_\rho$  follows by (2.26) and Lemma B.2 in [31].  $\square$



**Remark 2.16.** *The formula (2.25) gives an explicit expression of a standard symbol  $b(x, \xi)$  as a function of a Weyl symbol  $a(x, \xi)$ . Then one can prove Lemmata 2.14, 2.15 inverting the role of the two symbols. More precisely one can obtain*

$$b(\varphi, x, \xi) = f(\varphi, x, \xi) + \tilde{b}(\varphi, x, \xi), \quad f(\varphi, x, \xi) := \sum_{p=0}^{N-1} \frac{1}{2^p p!} (\partial_x^p D_\xi^p a)(\varphi, x, \xi) \quad (2.30)$$

with  $\tilde{b}$  satisfying estimates like (2.26). Hence  $Op^W(a) = Op(f) + R_\rho$  for some remainder  $R_\rho \in \mathfrak{L}_{\rho,p}$  satisfying bounds like (2.28), (2.29).

It is known that the  $L^2$ -adjoint of a pseudo differential operator is pseudo differential. In the Weyl quantization it is easier, w.r.t. the standard quantization, to determine the symbol of the adjoint operator. In particular we have the following:

- if  $A := Op^W(a(x, \xi))$ , with  $a(x, \xi) \in S^m$  then the operator  $\bar{A}$  defined in (2.8) and the  $L^2$ -adjoint have the form

$$\bar{A} := Op^W(\overline{a(x, -\xi)}), \quad A^* := Op^W(\overline{a(x, \xi)}). \quad (2.31)$$

**Composition of pseudo differential operators.** Let

$$\sigma(D_x, D_\xi, D_y, D_\eta) := D_\xi D_y - D_x D_\eta$$

where  $D_x := -i\partial_x$  and  $D_\xi, D_y, D_\eta$  are similarly defined.

**Definition 2.17. (Asymptotic expansion of composition symbol).** *Let  $\rho \geq 0$ . Consider symbols  $a(x, \xi) \in S^m$  and  $b(x, \xi) \in S^{m'}$ . We define*

$$(a \#_\rho^W b)(x, \xi) := \sum_{k=0}^{\rho} \frac{1}{k!} \left( \frac{i}{2} \sigma(D_x, D_\xi, D_y, D_\eta) \right)^k \left[ a(x, \xi) b(y, \eta) \right]_{|x=y, \xi=\eta} \quad (2.32)$$

modulo symbols in  $S^{m+m'-\rho}$ .

- We have the expansion

$$a \#_\rho^W b = ab + \frac{1}{2i} \{a, b\} - \frac{1}{8} \left( \partial_{\xi\xi} a \partial_{xx} b - 2\partial_{\xi x} a \partial_{x\xi} b + \partial_{xx} a \partial_{\xi\xi} b \right)$$

up to a symbol in  $S^{m+m'-3}$ , where

$$\{a, b\} := \partial_\xi a \partial_x b - \partial_x a \partial_\xi b$$

denotes the Poisson bracket. In the following Lemma we prove that the composition of pseudo differential operators is pseudo differential up to up to a remainder which is smoothing in the  $x$ -variable. We provide precise estimates on the  $|\cdot|_{m,s,\alpha}^{\gamma,\mathcal{O}}$ -norm of the symbol of the composition operator.

**Lemma 2.18. (Composition).** *Let  $m, m' \in \mathbb{R}$ . Fix  $\rho, p$  as in Definition 2.10, such that  $\rho \geq \max\{-(m+m'+1), 3\}$  and define  $N := m+m'+\rho \geq 1$ . Consider two symbols  $a(\varphi, x, \xi) \in S^m$ ,  $b(\varphi, x, \xi) \in S^{m'}$ . There exist an operator  $R_\rho \in \mathfrak{L}_{\rho,p}$  and a constant  $\sigma = \sigma(N) \sim N$  such that (recall Def. 2.17)*

$$Op^W(a) \circ Op^W(b) = Op^W(c) + R_\rho, \quad c := a \#_N^W b \in S^{m+m'} \quad (2.33)$$

where

$$|c|_{m+m',s,\alpha}^{\gamma,\mathcal{O}} \leq_{s,\rho,\alpha,m,m'} |a|_{m,s+\sigma,\alpha}^{\gamma,\mathcal{O}} |b|_{m',s_0+\sigma,\alpha+\sigma}^{\gamma,\mathcal{O}} |a|_{m,s_0+\sigma,\sigma+\alpha}^{\gamma,\mathcal{O}} + |b|_{m',s+\sigma,\alpha+\sigma}^{\gamma,\mathcal{O}}, \quad (2.34)$$

$$\mathbb{M}_{R_\rho}^\gamma(s, \mathbf{b}) \leq_{s,\rho,m,m'} |a|_{m,s+\sigma,\sigma}^{\gamma,\mathcal{O}} |b|_{m',s_0+\sigma,\sigma}^{\gamma,\mathcal{O}} + |a|_{m,s_0+\sigma,\sigma}^{\gamma,\mathcal{O}} |b|_{m',s+\rho+\sigma,\sigma}^{\gamma,\mathcal{O}}, \quad (2.35)$$

for all  $0 \leq \mathbf{b} \leq \rho - 2$  and  $s_0 \leq s \leq \mathcal{S}$ . Moreover one has

$$|\Delta_{12} c|_{m+m',p+\sigma,\alpha+\sigma} \leq_{p,\alpha,\rho,m,m'} |\Delta_{12} a|_{m,p+\sigma,\sigma+\alpha} |b|_{m',p+\sigma,\alpha+\sigma} + |a|_{m,p+\sigma,\sigma+\alpha} |\Delta_{12} b|_{m',p+\sigma,\alpha+\sigma}, \quad (2.36)$$

$$\mathbb{M}_{\Delta_{12} R_\rho}(p, \mathbf{b}) \leq_{p,\rho,m,m'} |\Delta_{12} a|_{m+1,p+\sigma,\sigma} |b|_{m',p+\sigma,\sigma} + |a|_{m,p+\sigma,\sigma} |\Delta_{12} b|_{m'+1,p+\sigma,\sigma}, \quad (2.37)$$

for all  $0 \leq \mathbf{b} \leq \rho - 3$  and where  $p$  is the constant given in Definition 2.10.

*Proof.* By Lemma 2.15 and Remark 2.16 we can write

$$Op^W(a) \circ Op^W(b) = Op(a_N) \circ Op(b_N) + Q_\rho \quad (2.38)$$

where the symbols  $a_N$  and  $b_N$  are defined as (recall (2.27), (2.30))

$$a_N = \sum_{p=0}^{N-1} \frac{1}{2^p p!} (\partial_x^p D_\xi^p a)(\varphi, x, \xi), \quad b_N = \sum_{p=0}^{N-1} \frac{1}{2^p p!} (\partial_x^p D_\xi^p b)(\varphi, x, \xi). \quad (2.39)$$

Moreover the remainder  $Q_\rho$  satisfies bounds like (2.28),(2.29). Lemma B.4 in [31] implies that

$$Op(a_N) \circ Op(b_N) = Op(c_N) + L_\rho \quad (2.40)$$

for some  $L_\rho \in \mathfrak{L}_{\rho,p}$  and where the symbol  $c_N \in S^m$  has the form

$$c_N(x, \xi) := \sum_{n=0}^{N-1} \frac{1}{n! i^n} (\partial_\xi^n a_N)(x, \xi) \cdot (\partial_x^n b_N)(x, \xi) \quad (2.41)$$

By Lemma 2.15 applied to the operator  $Op(c_N)$  in (2.40) we get  $Op(c_N) = Op^W(c) + \tilde{R}_\rho$  where  $\tilde{R}_\rho \in \mathfrak{L}_{\rho,p}$  and

$$c = \sum_{p=0}^{N-1} \frac{(-1)^p}{2^p p!} (\partial_x^p D_\xi^p c_N)(\varphi, x, \xi). \quad (2.42)$$

By an explicit computation using (2.41), (2.39) we deduce that the symbol  $c$  in (2.42) has the form  $c = a \#_N^W b$  (see (2.32)). Using (2.38), (2.40) we get the (2.33) with  $R_\rho := \tilde{R}_\rho + L_\rho + Q_\rho$ . The estimates (2.34)-(2.37) follow by combining the estimates in Lemmata B.2 and B.4 in [31] and Lemmata 2.14, 2.15 and Remark 2.16.  $\square$

**Lemma 2.19. (Commutator).** *Let  $m, m' \in \mathbb{R}$ . Fix  $\rho, p$  as in Definition 2.10, such that  $\rho \geq \max\{-(m+m'+1), 3\}$  and define  $N := m + m' + \rho \geq 1$ . Consider two symbols  $a(\varphi, x, \xi) \in S^m, b(\varphi, x, \xi) \in S^{m'}$ . There exist an operator  $R_\rho \in \mathfrak{L}_{\rho,p}$  and  $\sigma = \sigma(N) \sim N$  such that the commutator between  $Op^W(a)$  and  $Op^W(b)$  has the form*

$$[Op^W(a), Op^W(b)] = Op^W(c) + R_\rho, \quad c := a \star_N b \in S^{m+m'-1} \quad (2.43)$$

where (recall (2.32))

$$a \star_N b := a \#_N^W b - b \#_N^W a = \frac{1}{i} \{a, b\} + r, \quad r \in S^{m+m'-3}. \quad (2.44)$$

Moreover the symbols  $\{a, b\}, r$  satisfy bounds as (2.34), (2.36) with  $m + m'$  replaced by  $m + m' - 1, m + m' - 3$  respectively. Finally the smoothing operator  $R_\rho$  satisfies bounds as (2.35) and (2.37).

*Proof.* It is a direct application of Lemma 2.18. The expansion (2.43), (2.44) follows using formula (2.32).  $\square$

#### Notation.

- Let  $a, b, c, d$  be symbols in the class  $S^m, m \in \mathbb{R}$  (see Def. 2.12) depending in a Lipschitz way on  $\omega \in \mathcal{O}$ . We shall write

$$A := A(\varphi, x, \xi) := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in S^m \otimes \mathcal{M}_2(\mathbb{C}) \quad (2.45)$$

to denote  $2 \times 2$  matrices of symbols. With abuse of notation we write  $|A|_{m,s,\alpha}^{\gamma,\mathcal{O}}$  to denote  $\max\{|f|_{m,s,\alpha}^{\gamma,\mathcal{O}}, f = a, b, c, d\}$ .

- Similarly we shall write  $\mathcal{Q} \in \mathfrak{L}_{\rho,p} \otimes \mathcal{M}_2(\mathbb{C})$  to denote a  $2 \times 2$ -matrix whose entries are smoothing operators in  $\mathfrak{L}_{\rho,p}$ .

**2.3. Hamiltonian and translation invariant operators.** Let  $\mathbf{v} \in \mathbb{Z}^\nu$  as in (1.27). We define the subspace

$$S_{\mathbf{v}} := \{u(\varphi, x) \in L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2) : u(\varphi, x) = U(\varphi - \mathbf{v}x), U(\Theta) \in L^2(\mathbb{T}^\nu; \mathbb{C}^2)\}. \quad (2.46)$$

With abuse of notation we denote by  $S_{\mathbf{v}}$  also the subspace of  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C})$  of scalar functions  $u(\varphi, x)$  of the form  $u(\varphi, x) = U(\varphi - \mathbf{v}x), U(\Theta) \in L^2(\mathbb{T}^\nu; \mathbb{C})$ . We consider the symplectic form in the extended phase space  $(\varphi, Q, h, \bar{h}) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H^s(\mathbb{T}; \mathbb{C}) \times H^s(\mathbb{T}; \mathbb{C})$

$$\Omega_e(\varphi, Q, h, \bar{h}) = dQ \wedge d\varphi - idh \wedge d\bar{h} \quad (2.47)$$

<sup>6</sup>namely, for any  $U_1 := (\varphi_1, Q_1, h_1, \bar{h}_1), U_2 := (\varphi_2, Q_2, h_2, \bar{h}_2)$ , one has  $\Omega_e(U_1, U_2) = J \begin{bmatrix} \varphi_1^1 \\ Q_1^1 \end{bmatrix} \cdot \begin{bmatrix} \varphi_2^2 \\ Q_2^2 \end{bmatrix} - \int_{\mathbb{T}} \begin{bmatrix} h_1 \\ \bar{h}_1 \end{bmatrix} \cdot iJ \begin{bmatrix} h_2 \\ \bar{h}_2 \end{bmatrix} dx$ ,  $J := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ .

where  $\cdot$  denotes the standard scalar product on  $\mathbb{R}^2$ . Given a function  $F : \mathbb{T}^\nu \times \mathbb{R}^\nu \times H^s(\mathbb{T}; \mathbb{C}) \times H^s(\mathbb{T}; \mathbb{C}) \rightarrow \mathbb{R}$  its Hamiltonian vector field is given by

$$X_F = \left( \partial_Q F, -\partial_\varphi F, -i\partial_{\bar{h}} F, i\partial_h F \right),$$

since

$$dH(U)[\hat{\eta}] = \Omega_e(\hat{\eta}, X_F(U)), \quad \forall U, \hat{\eta} \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H^s(\mathbb{T}; \mathbb{C}) \times H^s(\mathbb{T}; \mathbb{C}).$$

Here  $\partial_h = (\partial_{\text{Re}(h)} - i\partial_{\text{Im}(h)})/\sqrt{2}$ ,  $\partial_{\bar{h}} = (\partial_{\text{Re}(h)} + i\partial_{\text{Im}(h)})/\sqrt{2}$  denote the  $L^2$ -gradient. The Poisson brackets between two Hamiltonians  $F, G$  are defined as

$$\{G, F\}_e := \Omega_e(X_G, X_F) = \partial_Q G \partial_\varphi F - \partial_\varphi G \partial_Q F + \frac{1}{i} \int_{\mathbb{T}} (\partial_h F \partial_{\bar{h}} G - \partial_{\bar{h}} F \partial_h G) dx. \quad (2.48)$$

We give the following Definition.

**Definition 2.20.** Consider an operator  $\mathcal{T} = \mathcal{T}(\varphi)$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2)$  of the form (2.7). We say that  $\mathcal{T}(\varphi)$  is

(i) **Hamiltonian** if it has the form

$$\mathcal{T} = iE\tilde{\mathcal{T}}, \quad E = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (2.49)$$

where  $\tilde{\mathcal{T}}$  has the form (2.7) and

$$(\tilde{\mathcal{T}}_+^+)^* = \tilde{\mathcal{T}}_+^+, \quad (\tilde{\mathcal{T}}_+^-)^* = \overline{\tilde{\mathcal{T}}_+^-} \quad (2.50)$$

where  $(\mathcal{T}_\sigma^{\sigma'})^*$ ,  $\sigma, \sigma' = \pm$  denotes the adjoint operator with respect to the complex scalar product of  $L^2(\mathbb{T}; \mathbb{C})$ ;

(ii)  **$x$ -translation invariant** if (recall (2.46))

$$\mathcal{T} : S_{\mathbf{v}} \rightarrow S_{\mathbf{v}}. \quad (2.51)$$

Consider a real operator  $\mathcal{Q} = \mathcal{Q}(\varphi)$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{R}^2)$  of the form (2.5). We say that  $\mathcal{Q}$  is, respectively, Hamiltonian,  $x$ -translation invariant, if the operator  $\mathcal{T} = \mathcal{C}\mathcal{Q}\mathcal{C}^{-1}$  defined in (2.7) is, respectively, Hamiltonian,  $x$ -translation invariant. We say that an operator  $\mathcal{L} = \mathcal{L}(\varphi)$  of the form

$$\mathcal{L} := \omega \cdot \partial_\varphi + \mathcal{T}(\varphi) \quad (2.52)$$

is respectively Hamiltonian,  $x$ -translation invariant if  $\mathcal{T}(\varphi)$  is respectively Hamiltonian,  $x$ -translation invariant. Similar for  $\mathcal{L} := \omega \cdot \partial_\varphi + \mathcal{Q}(\varphi)$  where  $\mathcal{Q}(\varphi)$  has the form (2.5).

Let  $\mathcal{T} = \mathcal{T}(\varphi)$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2)$  of the form (2.7) be an Hamiltonian operator according to Definition 2.20. We associate to  $\mathcal{T}$  a real valued Hamiltonian function

$$\mathcal{H}(U) = \omega \cdot Q + \frac{1}{2} \int_{\mathbb{T}} \tilde{\mathcal{T}}(\varphi) \left[ \frac{h}{h} \right] \cdot \left[ \frac{\bar{h}}{h} \right] dx, \quad \omega \in \mathbb{R}^\nu, \quad U = (\varphi, Q, h, \bar{h}) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H^s(\mathbb{T}; \mathbb{C}) \times H^s(\mathbb{T}; \mathbb{C}) \quad (2.53)$$

with  $\tilde{\mathcal{T}}$  as in (2.49), (2.50). Its Hamiltonian vector field  $X_{\mathcal{H}}$ , w.r.t. the symplectic form (2.47), as the form

$$\partial_t \varphi = \omega, \quad \partial_t Q = -\frac{1}{2} \partial_\varphi \mathcal{H}(U), \quad \partial_t \left[ \frac{h}{h} \right] = -iE\tilde{\mathcal{T}}(\varphi) \left[ \frac{h}{h} \right] = -\mathcal{T}(\varphi) \left[ \frac{h}{h} \right].$$

The second equation is not relevant for the dynamics since it is completely determined from the equation on the variables  $(\varphi, h, \bar{h})$ . Moreover, setting  $\varphi = \omega t$ ,  $z(\omega t) := h(t)$ ,  $z(\varphi) \in H^s(\mathbb{T}^{\nu+1}; \mathbb{C})$ , we can write the third equation as

$$0 = \omega \cdot \partial_\varphi \left[ \frac{z}{z} \right](\varphi) + \mathcal{T}(\varphi) \left[ \frac{z}{z} \right](\varphi) = \mathcal{L}(\varphi) \left[ \frac{z}{z} \right](\varphi),$$

where  $\mathcal{L}(\varphi)$  is the operator defined in (2.52) acting on functions in  $H^s(\mathbb{T}^{\nu+1}; \mathbb{C})$ . This justifies Definition 2.20. In the following we shall characterize Hamiltonian operators which are  $x$ -translation invariant, namely (2.51) holds, by giving a condition on the associated Hamiltonian function  $\mathcal{H}$  in (2.53).

**Lemma 2.21.** The following conditions are equivalent:

- (a) the function  $h = h(\varphi, x)$  belong to  $S_{\mathbf{v}}$  in (2.46);
- (b) one has that

$$(\mathbf{v} \cdot \partial_\varphi + \partial_x)h = 0; \quad (2.54)$$

- (c)  $h(\varphi, x) = (2\pi)^{-1/2} \sum_{\ell, j} h_{\ell, j} e^{i\ell \cdot \varphi + i j x}$  is such that

$$\text{if } h_{\ell, j} \neq 0 \Rightarrow \mathbf{v} \cdot \ell + j = 0, \quad \forall \ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}. \quad (2.55)$$

*Proof.* (a)  $\Rightarrow$  (b). If  $h \in S_{\mathbf{v}}$  then we can write  $h(\varphi, x) = G(\varphi - \mathbf{v}x)$  for some  $G(\Theta)$ ,  $\Theta \in \mathbb{T}^{\nu}$ . Hence  $(\mathbf{v} \cdot \partial_{\varphi} + \partial_x)h(\varphi, x) = (\mathbf{v} \cdot \partial_{\Theta}G)(\varphi - \mathbf{v}x) - (\mathbf{v} \cdot \partial_{\Theta}G)(\varphi - \mathbf{v}x) = 0$  which is the (2.54).

(b)  $\Rightarrow$  (c). By (2.54) and passing to the Fourier basis we have

$$\sum_{\ell \in \mathbb{Z}^{\nu}, j \in \mathbb{Z}} (\mathbf{i}\mathbf{v} \cdot \ell + \mathbf{i}j)h_{\ell j}e^{\mathbf{i}\ell \cdot \varphi + \mathbf{i}jx} = 0,$$

and hence we get the (2.55). The implications (c)  $\Rightarrow$  (b) and (c)  $\Rightarrow$  (a) are trivial.  $\square$

**Lemma 2.22.** (*x*-translation invariant operators). Consider an operator  $\tilde{\mathcal{T}}$  acting on  $L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2)$  of the form (2.7) satisfying (2.50). Consider the Hamiltonian  $\mathcal{H}$  in (2.53), the operator  $\mathcal{T} := \mathbf{i}E\tilde{\mathcal{T}}$  (see (2.49)) and the momentum Hamiltonian

$$\mathcal{M} = \mathcal{M}(\varphi, Q, h, \bar{h}) = -\mathbf{v} \cdot Q + \frac{1}{2} \int_{\mathbb{T}} \mathbf{i}E \left[ \frac{h_x}{h_x} \right] \cdot \left[ \frac{\bar{h}}{h} \right] dx. \quad (2.56)$$

The following conditions are equivalent:

- (a) the Hamiltonian operator  $\mathcal{T}$  is *x*-translation invariant, namely satisfies (2.51);
- (b) the Hamiltonians  $\mathcal{H}, \mathcal{M}$  Poisson commutes, i.e.

$$\{\mathcal{M}, \mathcal{H}\}_e = 0, \quad (2.57)$$

where  $\{\cdot, \cdot\}_e$  are the Poisson brackets in (2.48);

- (c) by identifying the operator  $\tilde{\mathcal{T}}$  with the matrix coefficients  $(\tilde{\mathcal{T}})_{\sigma, j}^{\sigma', k}(\ell)$ ,  $\ell \in \mathbb{Z}^{\nu}$ ,  $j, k \in \mathbb{Z}$ ,  $\sigma, \sigma' = \pm$  (recall (2.6), (2.9)) one has that

$$(\tilde{\mathcal{T}})_{\sigma, j}^{\sigma', k}(\ell) \neq 0 \Rightarrow \mathbf{v} \cdot \ell + j - k = 0, \quad \forall \ell \in \mathbb{Z}^{\nu}, j, k \in \mathbb{Z}, \sigma, \sigma' = \pm. \quad (2.58)$$

*Proof.* Consider the condition (2.57) in item (b). By using the (2.48), (2.53), (2.56) we have that

$$0 = \{\mathcal{M}, \mathcal{H}\}_e = \Omega_e(X_{\mathcal{M}}, X_{\mathcal{H}}) = \frac{1}{2} \int_{\mathbb{T}} \left( -\mathbf{v} \cdot \partial_{\varphi} \tilde{\mathcal{T}}(\varphi) + \tilde{\mathcal{T}}(\varphi) \partial_x - \partial_x \tilde{\mathcal{T}}(\varphi) \right) \left[ \frac{h}{h} \right] \cdot \left[ \frac{\bar{h}}{h} \right] dx.$$

The equation above is equivalent, passing to the Fourier representation, to

$$\mathbf{i}(-\mathbf{v} \cdot \ell + k - j)(\tilde{\mathcal{T}})_{\sigma, j}^{\sigma', k} = 0, \quad \forall \ell \in \mathbb{Z}^{\nu}, j, k \in \mathbb{Z}, \sigma, \sigma' = \pm.$$

This condition holds if and only if (2.58) is satisfied. This proves (b)  $\Leftrightarrow$  (c).

Consider the operator  $\tilde{\mathcal{T}}_+^+$  and let  $h \in S_{\mathbf{v}}$ . We have

$$\tilde{\mathcal{T}}_+^+ h = \sum_{\ell \in \mathbb{Z}^{\nu}, j \in \mathbb{Z}} e^{\mathbf{i}\ell \cdot \varphi + \mathbf{i}jx} \left( \sum_{p \in \mathbb{Z}^{\nu}, k \in \mathbb{Z}} (\tilde{\mathcal{T}})_{+, j}^{+, k}(\ell - p) h_{pk} \right).$$

We also recall that  $\mathbf{v} \cdot p + k = 0$ , for any  $p, k$  since  $h \in S_{\mathbf{v}}$ . Then if  $\tilde{\mathcal{T}}_+^+ h \in S_{\mathbf{v}}$  we must have (recall (2.55))  $\mathbf{v} \cdot \ell + j = 0$  which implies (2.58) on the coefficients  $(\tilde{\mathcal{T}})_{+, j}^{+, k}(\ell - p)$ . This proves (a)  $\Rightarrow$  (c). The converse is similar.  $\square$

**Lemma 2.23.** Let  $\mathcal{A} = (\mathcal{A}_{\sigma, \sigma'}^{\sigma'})_{\sigma, \sigma'}$  be an operator of the form (2.7) satisfying (2.50), i.e.  $(\mathcal{A}_{\sigma}^{\sigma})^* = \mathcal{A}_{\sigma}^{\sigma}$  and  $(\mathcal{A}_{\sigma}^{-\sigma})^* = \overline{\mathcal{A}_{\sigma}^{-\sigma}}$ , and let  $((\mathcal{A})_{\sigma, j}^{\sigma', k}(\ell))$ ,  $i = 1, 2$ ,  $\ell \in \mathbb{Z}^{\nu}$ ,  $j, k \in \mathbb{Z}$  be the matrices representing the operators  $\mathcal{A}_{\sigma}^{\sigma'}$ . Then

$$\overline{\mathcal{A}_{\sigma}^{\sigma'}} := \left( (\overline{\mathcal{A}})_{\sigma, j}^{\sigma', k}(\ell) \right), \quad (\overline{\mathcal{A}})_{\sigma, j}^{\sigma', k}(\ell) = \overline{(\mathcal{A})_{\sigma, -j}^{\sigma', -k}(-\ell)}, \quad (2.59)$$

$$(\mathcal{A}_{\sigma}^{\sigma'})^* := (\overline{\mathcal{A}_{\sigma}^{\sigma'}})^T := \left( (\mathcal{A}^*)_{\sigma, j}^{\sigma', k}(\ell) \right), \quad (\mathcal{A}^*)_{\sigma, j}^{\sigma', k}(\ell) = \overline{(\mathcal{A})_{\sigma, k}^{\sigma', j}(-\ell)}, \quad (2.60)$$

In particular, since  $\mathcal{A}$  is self-adjoint (see (2.50)) and real-to real, one has

$$(\mathcal{A})_{\sigma, j}^{\sigma, k}(\ell) = \overline{(\mathcal{A})_{\sigma, k}^{\sigma, j}(-\ell)}, \quad (\mathcal{A})_{\sigma, -j}^{-\sigma, -k}(\ell) = ((\mathcal{A}_{\sigma}^{-\sigma})^T)_k^j(\ell) = (\mathcal{A})_{\sigma, k}^{-\sigma, j}(\ell) \quad (2.61)$$

and

$$\mathcal{A}_{-\sigma, j}^{-\sigma', k}(\ell) = \overline{(\mathcal{A})_{\sigma, -j}^{\sigma', -k}(-\ell)} \quad (2.62)$$

*Proof.* We prove (2.59), (2.60) for  $\mathcal{A}_1 := \mathcal{A}_+^\dagger$ . Consider a function  $h = (\sqrt{2\pi})^{-1} \sum_k h_k e^{ikx}$ . Then

$$\overline{\mathcal{A}_1(\varphi)[h]} \stackrel{(2.8)}{=} \overline{\mathcal{A}_1(\varphi)[\bar{h}]} = \sum_j e^{ijx} \sum_{k,\ell} e^{i\ell \cdot \varphi} (\mathcal{A}_1)_j^k(\ell) \bar{h}_{-k} = \sum_{j,k,\ell} (\mathcal{A}_1)_{-j}^{-k}(-\ell) h_k e^{ijx} e^{i\ell \cdot \varphi}.$$

This implies the (2.59). Moreover we have

$$\sum_{j,k,\ell} e^{i\ell \cdot \varphi} (\mathcal{T}_1)_j^k(\ell) h_k \bar{v}_j = (\mathcal{T}_1(\varphi)h, v)_{L^2} = (h, \mathcal{T}_1^*(\varphi)v)_{L^2} = \int_{\mathbb{T}} h \cdot \overline{\mathcal{T}_1^*(\varphi)v} dx = \sum_{j,k,\ell} (\mathcal{T}_1^*)_{-j}^k(\ell) h_k \bar{v}_j e^{i\ell \cdot \varphi}$$

This implies the (2.60). Then the (2.61) follows by conditions (2.50).  $\square$

**Lemma 2.24.** (*x-translation invariant symbols*). Consider a symbol  $a(\varphi, x, \xi)$  in  $S^m$ ,  $m \in \mathbb{R}$ . We have that  $a$  has the form

$$a(\varphi, x, \xi) = A(\varphi - \mathbf{v}x, \xi) \quad (2.63)$$

for some  $A(\Theta, \xi)$ ,  $\Theta \in \mathbb{T}^\nu$ ,  $\xi \in \mathbb{R}$  satisfying

$$|\partial_\Theta^\alpha \partial_\xi^\beta A(\Theta, \xi)| \leq C_{\alpha,\beta} \langle \xi \rangle^{m-\beta}, \quad \forall \alpha, \beta \in \mathbb{N} \quad (2.64)$$

if and only if

$$\{\mathcal{M}, \mathcal{H}\}_e = 0, \quad \mathcal{H}(\varphi, Q, h, \bar{h}) = \omega \cdot Q + \int_{\mathbb{T}} Op^W(a(\varphi, x, \xi)) h \cdot \bar{h} dx, \quad (2.65)$$

*Proof.* By (2.48), (2.65), (2.56) we have

$$0 = \int_{\mathbb{T}} Op^W \left( -(\mathbf{v} \cdot \partial_\varphi a)(\varphi, x, \xi) + \frac{1}{i} \{a(\varphi, x, \xi), i\xi\} \right) h \cdot \bar{h} dx = \int_{\mathbb{T}} Op^W \left( -(\mathbf{v} \cdot \partial_\varphi a)(\varphi, x, \xi) - (\partial_x a)(\varphi, x, \xi) \right) h \cdot \bar{h} dx.$$

Therefore we must have

$$0 = \left( -\mathbf{v} \cdot \partial_\varphi a - \partial_x a \right)(\varphi, x, \xi) = \sum_{\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}} (-i)(\mathbf{v} \cdot \ell + j) \tilde{a}(\ell, j, \xi) e^{i\ell \cdot \varphi + ijx}$$

where  $\tilde{a}(\ell, j, \xi)$  denotes the Fourier transform of  $a(\varphi, x, \xi)$  in the variables  $(\varphi, x) \in \mathbb{T}^\nu \times \mathbb{T}$ . Then the (2.65) is verified if and only if  $\mathbf{v} \cdot \ell + j = 0$  for any  $\ell \in \mathbb{Z}^\nu, j \in \mathbb{Z}$ . Then we can write the symbol  $a(\varphi, x, \xi)$  as in (2.63). The bound (2.64) follows by the estimates (2.14) on  $a(\varphi, x, \xi)$ .  $\square$

If condition (2.63) holds we shall say that the symbol  $a(\varphi, x, \xi)$  is *x-translation invariant*.

**Lemma 2.25.** (*Real-to-real/Self-adjoint matrices of symbols*). Consider the operator (recall (2.45))

$$L = Op^W(A(\varphi, x, \xi)), \quad A \in S^m \otimes \mathcal{M}_2(\mathbb{C}).$$

We have that  $L$  is real-to-real according to Definition 2.1 if and only if the matrix  $A$  has the form

$$A(\varphi, x, \xi) := \begin{pmatrix} a(\varphi, x, \xi) & b(\varphi, x, \xi) \\ b(\varphi, x, -\xi) & a(\varphi, x, -\xi) \end{pmatrix}, \quad (2.66)$$

and  $L$  is self-adjoint, i.e. satisfies (2.50), if and only if the matrix of symbols  $A$  satisfies

$$\overline{(a(\varphi, x, \xi))} = a(\varphi, x, \xi), \quad b(\varphi, x, -\xi) = b(\varphi, x, \xi). \quad (2.67)$$

Finally  $L$  is Hamiltonian, i.e. satisfies (2.49) if and only if the matrix  $A$  satisfies

$$\overline{(a(\varphi, x, \xi))} = -a(\varphi, x, \xi), \quad b(\varphi, x, -\xi) = b(\varphi, x, \xi). \quad (2.68)$$

*Proof.* The Lemma follows by usign formulæ (2.7), (2.31).  $\square$

We conclude this section with the definition of a special class of operator we shall use in sections 8-10.3.

**Definition 2.26.** We say that a linear operator  $\mathcal{T}(\varphi) \in \mathcal{L}(L^2(\mathbb{T}^{\nu+1}; \mathbb{C}^2))$  belongs to the class  $\mathfrak{S}_0$  if it is

- (i) real-to-real, i.e. satisfies (2.7);
- (ii) Hamiltonian, i.e. satisfies (2.49);
- (iii) *x-translation invariant*, i.e. satisfies the (2.51) with  $\mathbf{v}$  in (5.24).

We say that a linear operator of the form

$$\mathcal{L} = \omega \cdot \partial_\varphi + Op^W(A(\varphi, x, \xi)) + \mathcal{R}$$

for some  $A \in S^m \otimes \mathcal{M}_2(\mathbb{C})$  and  $\mathcal{R} \in \mathcal{L}_{p,p} \otimes \mathcal{M}_2(\mathbb{C})$  belongs to  $\mathfrak{S}_1$  if  $\mathcal{L}$  and  $Op^W(A(\varphi, x, \xi))$  belong to  $\mathfrak{S}_0$ .

**Remark 2.27.** Consider  $\mathcal{L} \in \mathfrak{S}_1$ . By Lemmata 2.24, 2.25 we deduce that the operator  $A(\varphi, x, \xi)$  satisfies (2.66), (2.68) and (2.63). Moreover we deduce that also the remainder  $\mathcal{R}$  belongs to  $\mathfrak{S}_0$ .

In the following we shall look for transformations of coordinates which preserves the structure the class  $\mathfrak{S}_1$ . In particular we give the following definition.

**Definition 2.28.** We say that a map  $\Phi = \Phi(\varphi)$  belongs to  $\mathfrak{T}_1$  if it is symplectic and  $x$ -translation invariant.

### 3. NORMAL FORMS AND INTEGRABILITY PROPERTIES OF THE PURE GRAVITY WATER WAVES

In this section we recall some properties of the water waves system (1.3), we discuss its Hamiltonian structure and normal form. In particular we focus on the formal integrability of the Hamiltonian at order four which has been proved in [25], [22], [70]. Due to the quasi-linear nature of the water waves equations the Birkhoff normal form procedure turns out to be not well defined. To overcome this problem we perform a "weaker" but *rigorous* normal form algorithm: first we normalize the dynamics on a finite dimensional subspace to find an approximately invariant torus; secondly we normalize the dynamics in the normal directions around the embedded torus. These procedures are discussed in details in section 3.2. In sections 3.3, 3.4 we prove, thanks to an argument of identification of normal forms, that the formal (approximate) integrability implies the (approximate) integrability of the linearized problem at the torus.

**3.1. Hamiltonian structure of water waves.** Let  $W_1 = \begin{bmatrix} \eta_1 \\ \psi_1 \end{bmatrix}$ ,  $W_2 = \begin{bmatrix} \eta_2 \\ \psi_2 \end{bmatrix}$ . We consider the the symplectic form

$$\tilde{\Omega}(W_1, W_2) := \int_{\mathbb{T}} W_1 \cdot J^{-1} W_2 dx = \int_{\mathbb{T}} -(\eta_1 \psi_2 - \psi_1 \eta_2) dx, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (3.1)$$

The vector field (see (1.5))

$$X(\eta, \psi) = X_H(\eta, \psi) = J \nabla H(\eta, \psi) = \begin{pmatrix} \nabla_\psi H(\eta, \psi) \\ -\nabla_\eta H(\eta, \psi) \end{pmatrix}$$

is the Hamiltonian vector field of  $H$  in (1.6) w.r.t. the symplectic form (3.1), i.e.

$$-\tilde{\Omega}(X_H(\eta, \psi), h) = dH((\eta, \psi))[h] \quad h = \begin{bmatrix} \hat{\eta} \\ \hat{\psi} \end{bmatrix},$$

while the Poisson bracket between functions  $F(\eta, \psi), H(\eta, \psi)$  are defined as

$$\{F, H\} = \tilde{\Omega}(X_F, X_H) = \int_{\mathbb{T}} (\nabla_\eta H \nabla_\psi F - \nabla_\psi H \nabla_\eta F) dx. \quad (3.2)$$

Consider the maps

$$\mathcal{C} := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix}, \quad \mathfrak{F} := \begin{bmatrix} |D|^{-\frac{1}{4}} & 0 \\ 0 & |D|^{\frac{1}{4}} \end{bmatrix}, \quad \Lambda := \mathcal{C} \circ \mathfrak{F} := \frac{1}{\sqrt{2}} \begin{bmatrix} |D|^{-\frac{1}{4}} & i|D|^{\frac{1}{4}} \\ |D|^{-\frac{1}{4}} & -i|D|^{\frac{1}{4}} \end{bmatrix}. \quad (3.3)$$

We note that the operator  $|D|^{-1/4}$  is well defined on the space of functions with zero spatial average. We introduce the complex symplectic variables

$$\begin{pmatrix} u \\ \bar{u} \end{pmatrix} = \Lambda \begin{pmatrix} \eta \\ \psi \end{pmatrix} := \frac{1}{\sqrt{2}} \begin{pmatrix} |D|^{-\frac{1}{4}} \eta + i|D|^{\frac{1}{4}} \psi \\ |D|^{-\frac{1}{4}} \eta - i|D|^{\frac{1}{4}} \psi \end{pmatrix}, \quad \begin{pmatrix} \eta \\ \psi \end{pmatrix} = \Lambda^{-1} \begin{pmatrix} u \\ \bar{u} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} |D|^{\frac{1}{4}} (u + \bar{u}) \\ -i|D|^{-\frac{1}{4}} (u - \bar{u}) \end{pmatrix}. \quad (3.4)$$

The symplectic form in (3.1) transforms, for  $U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix}$ ,  $V = \begin{bmatrix} v \\ \bar{v} \end{bmatrix}$ , into

$$\Omega(U, V) := - \int_{\mathbb{T}} U \cdot iJV dx = - \int_{\mathbb{T}} i(u\bar{v} - \bar{u}v) dx. \quad (3.5)$$

The Poisson bracket in (3.2) assumes the form

$$\{F, H\} = \frac{1}{i} \int_{\mathbb{T}} (\nabla_u H \nabla_{\bar{u}} F - \nabla_{\bar{u}} H \nabla_u F) dx = \frac{1}{i} \sum_{k \in \mathbb{Z} \setminus \{0\}} (\partial_{u_k} H \partial_{\bar{u}_k} F - \partial_{\bar{u}_k} H \partial_{u_k} F), \quad (3.6)$$



where

$$\partial_{u_k} = \frac{1}{\sqrt{2}}(|k|^{-\frac{1}{4}}\partial_{\eta_k} - i|k|^{\frac{1}{4}}\partial_{\psi_k}), \quad \partial_{\overline{u-k}} = \frac{1}{\sqrt{2}}(|k|^{-\frac{1}{4}}\partial_{\eta_k} + i|k|^{\frac{1}{4}}\partial_{\psi_k}).$$

In these coordinates the vector field  $X = X_H$  in (1.3) assumes the form (setting  $H_C := H \circ \Lambda^{-1}$ )

$$X^C := X_{H_C} = \begin{pmatrix} -i\partial_{\overline{u}}H_C \\ i\partial_u H_C \end{pmatrix} = \frac{1}{\sqrt{2\pi}} \sum_{k \in \mathbb{Z} \setminus \{0\}} \begin{pmatrix} -i\partial_{\overline{u_k}}H_C e^{ikx} \\ i\partial_{u_k}H_C e^{-ikx} \end{pmatrix}, \quad (3.7)$$

that we also identify, using the standard vector field notation, with

$$X^C = \sum_{k \in \mathbb{Z} \setminus \{0\}, \sigma = \pm} -i\sigma \partial_{\overline{u_k}^{-\sigma}} H_C \partial_{u_k^\sigma}.$$

The Hamiltonian of the momentum (recall (1.7)) reads as

$$M_C := M \circ \Lambda^{-1} = \int_{\mathbb{T}} i u_x \cdot \overline{u} dx. \quad (3.8)$$

**Taylor expansion at the origin.** Consider the Dirichlet-Neumann operator  $G(\eta)$  in (1.4). The map  $(\eta, \psi) \rightarrow G(\eta)\psi$  is linear with respect to  $\psi$  and nonlinear with respect to the profile  $\eta$ . The derivative with respect to  $\eta$  (which is called ‘‘shape derivative’’) is given by the formula (see for instance [56])

$$G'(\eta)[\widehat{\eta}]\psi = \lim_{\epsilon \rightarrow 0} \frac{(G(\eta + \epsilon\widehat{\eta})\psi - G(\eta)\psi)}{\epsilon} = -G(\eta)(B\widehat{\eta}) - \partial_x(V\widehat{\eta}), \quad (3.9)$$

where we denoted the horizontal and vertical components of the velocity field at the free interface by

$$V = V(\eta, \psi) := (\partial_x \Phi)(x, \eta(x)) = \psi_x - \eta_x B, \quad (3.10)$$

$$B = B(\eta, \psi) := (\partial_y \Phi)(x, \eta(x)) = \frac{G(\eta)\psi + \eta_x \psi_x}{1 + \eta_x^2}. \quad (3.11)$$

It is also known that  $\eta \mapsto G(\eta)$  is analytic and admits the Taylor expansion of the Hamiltonian near equilibrium ( $\eta = 0, \psi = 0$ )

$$G(\eta) = \sum_{m=0}^{\infty} G^{(m)}(\eta). \quad (3.12)$$

Each term  $G^{(m)}$  in the Taylor expansion is homogeneous of degree  $m$  and can be computed explicitly. For instance (see e.g. formula (2.5) of [22]) we have

$$\begin{aligned} G^{(0)} &:= |D|, & G^{(1)}(\eta) &:= -\partial_x \eta \partial_x - |D|\eta|D|, \\ G^{(2)}(\eta) &:= -\frac{1}{2} \left( D^2 \eta^2 |D| + |D|\eta^2 D^2 - 2|D|\eta|D|\eta|D \right). \end{aligned} \quad (3.13)$$

For further properties about the Dirichlet-Neumann operator we refer the reader to Appendix B.1.

The Hamiltonian  $H$  in (1.6) has a convergent Taylor expansion

$$H = H^{(2)} + H^{(3)} + H^{(4)} + \dots + H^{(m)} + R^{(m+1)}.$$

In the complex coordinates given by (3.3)-(3.4) the Hamiltonian reads as

$$\begin{aligned} H_C &:= H \circ \Lambda^{-1} = \sum_{p=2}^N H_C^{(p)} + H_C^{(\geq N+1)}, \\ H_C^{(2)} &= \sum_{j \in \mathbb{Z} \setminus \{0\}} \sqrt{|j|} u_j \overline{u_j}, & H_C^{(p)} &= \sum_{\sigma_1 j_1 + \dots + \sigma_p j_p = 0} H_{j_1, \dots, j_p}^{\sigma_1, \dots, \sigma_p} u_{j_1}^{\sigma_1} \dots u_{j_p}^{\sigma_p} \end{aligned} \quad (3.14)$$

where  $N \geq 2$ ,  $H_{j_1, \dots, j_p}^{\sigma_1, \dots, \sigma_p} \in \mathbb{C}$  and the Hamiltonian  $H_C^{(\geq N+1)}$  collects all the monomials of homogeneity  $\geq N+1$ .

**Splitting of the phase space.** Recall  $S$  in (1.23) and define  $S^c := \mathbb{Z} \setminus (S \cup \{0\})$ . We decompose the phase space as

$$\begin{aligned} H_0^1(\mathbb{T}) \times H_0^1(\mathbb{T}) &:= H_S \oplus H_S^\perp, \\ H_S &:= \left\{ \begin{pmatrix} \eta_S \\ \psi_S \end{pmatrix} := \sum_{j \in S} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} e^{ijx} \right\}, & H_S^\perp &:= \left\{ \begin{pmatrix} \tilde{\eta} \\ \tilde{\psi} \end{pmatrix} := \sum_{j \in S^c} \begin{pmatrix} \eta_j \\ \psi_j \end{pmatrix} e^{ijx} \right\}, \end{aligned} \quad (3.15)$$

and we denote by  $\Pi_S, \Pi_S^\perp$  the corresponding orthogonal projectors. The subspaces  $H_S$  and  $H_S^\perp$  are symplectic orthogonal respect to the 2-form  $\Omega$  (see (3.5)). We use the following notations for the complex variables,

$$\begin{pmatrix} \eta \\ \psi \end{pmatrix} = \begin{pmatrix} \eta_S \\ \psi_S \end{pmatrix} + \begin{pmatrix} \tilde{\eta} \\ \tilde{\psi} \end{pmatrix}, \quad \Lambda \begin{pmatrix} \eta_S \\ \psi_S \end{pmatrix} =: \begin{pmatrix} v \\ \bar{v} \end{pmatrix}, \quad \Lambda \begin{pmatrix} \tilde{\eta} \\ \tilde{\psi} \end{pmatrix} =: \begin{pmatrix} z \\ \bar{z} \end{pmatrix}. \quad (3.16)$$

The notation  $R_k(v^{\alpha_1} \bar{v}^{\beta_1} z^{\alpha_2} \bar{z}^{\beta_2})$  indicates a homogeneous polynomial of degree  $k$  in  $(v, \bar{v}, z, \bar{z})$  of the form

$$R_k(v^{\alpha_1} \bar{v}^{\beta_1} z^{\alpha_2} \bar{z}^{\beta_2}) = M[\underbrace{v, \dots, v}_{\alpha_1\text{-times}}, \underbrace{\bar{v}, \dots, \bar{v}}_{\beta_1\text{-times}}, \underbrace{z, \dots, z}_{\alpha_2\text{-times}}, \underbrace{\bar{z}, \dots, \bar{z}}_{\beta_2\text{-times}}], \quad (3.17)$$

$$M = k\text{-linear} \quad k := \alpha_1 + \alpha_2 + \beta_1 + \beta_2.$$

We denote with  $H^{(n, \geq k)}, H^{(n, k)}, H^{(n, \leq k)}$  the terms of type  $R_n(v^{\alpha_1} \bar{v}^{\beta_1} z^{\alpha_2} \bar{z}^{\beta_2})$ , where, respectively,  $\alpha_2 + \beta_2 \geq k, \alpha_2 + \beta_2 = k, \alpha_2 + \beta_2 \leq k$ , that appear in the homogeneous polynomial  $H^{(n)}$  of degree  $n$  in the variables  $(v, \bar{v}, z, \bar{z})$ . We denote by  $\Pi^{d_z \leq k}$ , respectively  $\Pi^{d_z = k}$ , the projector of a homogeneous Hamiltonian of degree  $n$  on the monomials with degree less or equal than  $k$ , respectively equal  $k$ , in the normal variable  $z$ , i.e.

$$\Pi^{d_z \leq k} H^{(n)} := H^{(n, \leq k)}, \quad \Pi^{d_z = k} H^{(n)} := H^{(n, k)}.$$

Given two Hamiltonians  $Q, K$  we shall define the adjoint action  $\text{ad}_Q(H) := \{Q, H\}$ . We denote by  $\Pi_{\text{Ker}(Q)}$ , the projection on the kernel of the adjoint action. We define the projector on the range of the adjoint action as  $\Pi_{\text{Rg}(Q)} := \text{I} - \Pi_{\text{Ker}(Q)}$ .

**3.2. Comparison among Birkhoff normal form procedures.** In this section we present three kinds of Birkhoff normal form procedures and we compare them. All these procedures are *formal* and they differ by the terms of the Hamiltonian that are supposed to be normalized along the process.

In this paper we shall implement a, not just formally defined, modification of the third one, the ‘‘Weak’’ plus ‘‘Linear’’ Birkhoff normal form.

In the following we use the notations introduced above related to the Hamiltonian of the pure gravity water waves system (3.14), but we point out that the procedures that we describe below apply to general analytic Hamiltonians close to an elliptic fixed point of the form  $H = H^{(2)} + H^{(3)} + \dots$  commuting with momentum and such that there are no 3-waves resonant interactions.

**‘‘Full’’ Birkhoff normal form.** We refer to (formal) full Birkhoff normal form method as the normalization procedure of the cubic and quartic terms  $H_{\mathbb{C}}^{(3)}$  and  $H_{\mathbb{C}}^{(4)}$  of the Hamiltonian in (3.14). This is the strongest normal form method and, in the PDEs context, it is usually hard to implement. Indeed the normalizing transformations are constructed as time-one flow maps of certain nonlinear PDEs whose vector field may lose derivatives, making the equations possibly ill-posed. There are two main sources of this loss of regularity:

- (1) The presence of small divisors - in the full procedure one has to get lower bounds for infinitely many combinations of the linear eigenvalues;
- (2) The unboundness of the vector field of the PDE that we want to normalize.

In the pure gravity case with infinite depth one has to deal with both issues. The normalization is done by applying the change of coordinates  $\Phi_{FB} := \Phi_{F_3} \circ \Phi_{F_4}$  where  $\Phi_{F_i}$  are the time one (formal) flow map generated by the Hamiltonians

$$F_3 := \text{ad}_{H_{\mathbb{C}}^{(2)}}^{-1} H^{(3)}, \quad F_4 := \text{ad}_{H_{\mathbb{C}}^{(2)}}^{-1} \Pi_{\text{Rg}(H_{\mathbb{C}}^{(2)})} \left( H_{\mathbb{C}}^{(4)} + \frac{1}{2} \{F_3, H_{\mathbb{C}}^{(3)}\} \right). \quad (3.18)$$

By using the fact that there are no 3-waves interactions it is easy to check (using Lie series) that one obtains

$$H_{\mathbb{C}} \circ \Phi_{FB} = H_{\mathbb{C}}^{(2)} + H_{FB}^{(4)} + \text{quintic terms}, \quad (3.19)$$

$$H_{FB}^{(4)} := \Pi_{\text{Ker}(H_{\mathbb{C}}^{(2)})} \left( H_{\mathbb{C}}^{(4)} + \frac{1}{2} \{F_3, H_{\mathbb{C}}^{(3)}\} \right). \quad (3.20)$$

**‘‘Partial’’ Birkhoff normal form.** We refer to (formal) partial Birkhoff normal form method as the normalization procedure of the cubic and quartic terms with at most two wave numbers outside  $S$ , i.e. the terms  $H_{\mathbb{C}}^{(3, \leq 2)}$  and  $H_{\mathbb{C}}^{(4, \leq 2)}$  of the Hamiltonian in (3.14). Such method has been successfully applied for many KAM results for *semilinear* PDEs, starting from the pioneering work [55]. Roughly speaking it is the optimal normal form procedure that provides a control on the tangent bundle of the expected invariant torus. However, by the quasi-linear nature of the equations, this method cannot be applied for the water waves system (see discussions below for more details).

The normalization procedure is done by applying the change of coordinates  $\Phi_{PB} := \Phi_{F_3^{(3,\leq 2)}} \circ \Phi_{\widehat{F}_4^{(4,\leq 2)}}$  where  $\Phi_{F_3^{(3,\leq 2)}}$ ,  $\Phi_{\widehat{F}_4^{(4,\leq 2)}}$  are the time one (formal) flow map generated by the Hamiltonians

$$\begin{aligned} F_3^{(3,\leq 2)} &:= ad_{H_C^{(2)}}^{-1} H^{(3,\leq 2)}, & \widehat{F}_4^{(4,\leq 2)} &:= ad_{H_C^{(2)}}^{-1} \Pi_{Rg(H_C^{(2)})} \Pi^{d_z \leq 2} H_1^{(4)}, \\ H_1^{(4)} &:= H_C^{(4)} + \frac{1}{2} \{F_3^{(3,\leq 2)}, H_C^{(3,\leq 2)}\} + \{F_3^{(3,\leq 2)}, H_C^{(3,3)}\}. \end{aligned} \quad (3.21)$$

It is easy to check (using Lie series) that one obtains

$$\begin{aligned} H_C \circ \Phi_{PB} &= H_C^{(2)} + H_C^{(3,3)} + H_{PB}^{(4)} + H_1^{(4,\geq 3)} + \text{quintic terms}, \\ H_{PB}^{(4)} &:= \Pi_{Ker(H_C^2)} \Pi^{d_z \leq 2} H_1^{(4)}. \end{aligned} \quad (3.22)$$

**“Weak” plus “Linear” Birkhoff normal form.** In this case the normalization procedure is split into two steps: (i) a weak BNF, which is always rigorously defined if the Hamiltonian terms that we want to normalize commute with the momentum (this assumption holds naturally for many physical PDE models); (ii) a linear BNF, which is defined just at the formal level, which aims to normalize the tangent bundle of the expected invariant torus.

(i) We first define the map  $\Phi_{WB} := \Phi_{F_3^{(3,\leq 1)}} \circ \Phi_{\widehat{F}_4^{(4,\leq 1)}}$  where  $\Phi_{F_3^{(3,\leq 1)}}$ ,  $\Phi_{\widehat{F}_4^{(4,\leq 1)}}$  are the time-one flow maps generated by the Hamiltonians

$$\begin{aligned} F_3^{(3,\leq 1)} &:= ad_{H_C^{(2)}}^{-1} H^{(3,\leq 1)}, & \widehat{F}_4^{(4,\leq 1)} &:= ad_{H_C^{(2)}}^{-1} \Pi_{Rg(H_C^{(2)})} \Pi^{d_z \leq 1} \widehat{H}_1^{(4)}, \\ \widehat{H}_1^{(4)} &:= H_C^{(4)} + \frac{1}{2} \{F_3^{(3,\leq 1)}, H_C^{(3,\leq 1)}\} + \{F_3^{(3,\leq 1)}, H_C^{(3,\geq 2)}\}. \end{aligned} \quad (3.23)$$

Observe that homogenous Hamiltonians of the form  $F^{(n,\leq 1)}$  that commute with momentum generate a finite dimensional vector field. This comes from the fact that the conservation of momentum for a monomial  $u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}$  reads as  $\sum_{i=1}^n \sigma_i j_i = 0$ . Since there is at most one index  $j_i \notin S$  this implies that all the  $j_i$ 's are contained in the ball centered at the origin with radius  $(n-1) \max(S)$  (recall (1.25)). Then the flow maps  $\Phi_{F_3^{(3,\leq 1)}}$ ,  $\Phi_{\widehat{F}_4^{(4,\leq 1)}}$  are well-defined.

One obtains

$$H_C \circ \Phi_{WB} = H_C^{(2)} + H_C^{(3,\geq 2)} + H_{WB}^{(4)} + \widehat{H}_1^{(4,\geq 2)} + \text{quintic terms}, \quad (3.24)$$

$$H_{WB}^{(4)} := \Pi_{Ker(H_C^2)} \Pi^{d_z \leq 1} \widehat{H}_1^{(4)}. \quad (3.25)$$

(ii) As a second step we define the map  $\Phi_{LB} := \Phi_{F_3^{(3,2)}} \circ \Phi_{\mathcal{F}_4^{(4,2)}}$  where  $\Phi_{F_3^{(3,\leq 1)}}$ ,  $\Phi_{\mathcal{F}_4^{(4,\leq 1)}}$  are the flows generated by the Hamiltonians

$$\begin{aligned} F_3^{(3,2)} &:= ad_{H_C^{(2)}}^{-1} H^{(3,2)}, & \mathcal{F}_4^{(4,2)} &:= ad_{H_C^{(2)}}^{-1} \Pi_{Rg(H_C^{(2)})} \Pi^{d_z=2} \mathcal{H}^{(4)} \\ \mathcal{H}^{(4)} &:= H_C^{(4)} + \frac{1}{2} \{F_3^{(3,\leq 1)}, H_C^{(3,\leq 1)}\} + \frac{1}{2} \{F_3^{(3,2)}, H_C^{(3,2)}\} + \{F_3^{(3,\leq 1)}, H_C^{(3,2)}\} + \{F_3^{(3,\leq 2)}, H_C^{(3,3)}\}. \end{aligned} \quad (3.26)$$

One obtains

$$H_C \circ \Phi_{WB} \circ \Phi_{LB} = H_C^{(2)} + H_C^{(3,3)} + H_{WB}^{(4)} + H_{LB}^{(4)} + \mathcal{H}^{(4,\geq 3)} + \text{quintic terms} \quad (3.27)$$

$$H_{LB}^{(4)} := \Pi_{Ker(H_C^2)} \Pi^{d_z=2} \mathcal{H}^{(4)}. \quad (3.28)$$

We remark again that the procedures described above are “formal” since the maps we used are not *a priori* well-defined. Indeed they are flows at time one of possibly ill-posed PDEs. This is due to two main reasons: the presence of “small divisors” in the inversion of the adjoint action  $ad_{H_C^{(2)}}$  and the fact that the vector field of  $H_C$  is *quasi-linear*.

**Remark 3.1.** In [14] the authors provide a result of long-time stability of the pure gravity system. One may wonder if there is a connection between this result and the strongest normalization procedures described above (partial and full BNF) and whether this could be applied in the search for quasi-periodic solutions for the pure gravity system.

In [14] the authors use a modified energies method which is based on a normal form approach (and the integrability at order four), but no genuine Birkhoff maps, meant as real transformations of the phase space, are provided. This is usually requested in KAM theory for PDEs.

At the best of our knowledge, full or partial Birkhoff transformations are still not available for water waves systems. In any case we believe that our approach is well-suited for proving the existence of finite-dimensional quasi-periodic invariant tori, because in some sense it requires minimal assumptions.

Notice that the map  $\Phi_{WB}$  generated by Hamiltonians in (3.23) is actually well-posed because is the flow of an *ODE*. This is true since  $\Pi^{d_z \leq 1}$ , combined with the conservation of momentum, gives the projection over a finite number of modes. In the paper we shall follow a procedure similar to the *Weak plus Linear* Birkhoff normal form. The aim of the *Weak* procedure is to find the first nonlinear approximate solution  $v_I$  of (1.3). Such approximate solution will be the starting point for the Nash-Moser scheme and the leading term of the expected quasi-periodic solution (see (1.28)). Actually, in section 11, we will construct a map admitting the same Taylor expansion (at low degree) of the map  $\Phi_{LB}$  but with rigorous estimates on the Sobolev spaces  $H^s$ . Therefore the main order of corrections of the eigenvalues of the linearized operator at  $v_I$  will be given by the vector field of the Hamiltonian  $H_{LB}^{(4)}$  in (3.28).

To compute such corrections one needs to study the kernel of the adjoint action of  $H_C^{(2)}$ . By an explicit computation one has that a fourth order monomial  $u_{j_1}^{\sigma_1} \dots u_{j_4}^{\sigma_4}$  belongs to  $Ker(H_C^{(2)})$  if it is Fourier supported on a *4-waves resonances* that is, non-trivial integer solutions of

$$\sigma_1 \sqrt{|j_1|} + \sigma_2 \sqrt{|j_2|} + \sigma_3 \sqrt{|j_3|} + \sigma_4 \sqrt{|j_4|} = 0, \quad \sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 + \sigma_4 j_4 = 0. \quad (3.29)$$

We know that (see for instance [70], [25], [22]) the quartic resonances are given by two families of quartets of wave numbers:

- (i) the *trivial* ones, which have the form  $j_1 = j_2, \sigma_1 = -\sigma_2$  and  $j_3 = j_4, \sigma_3 = -\sigma_4$ ;
- (ii) the *Benjamin-Feir* resonances, which consist in the two parameter family of solutions

$$\bigcup_{\lambda \in \mathbb{Z} \setminus \{0\}, b \in \mathbb{N}} \left\{ j_1 = -\lambda b^2, j_2 = \lambda(b+1)^2, j_3 = \lambda(b^2 + b + 1)^2, j_4 = \lambda(b+1)^2 b^2 \right\}, \quad (3.30)$$

with  $\sigma_1 = \sigma_3 = -\sigma_2 = -\sigma_4$ ,

In [70] (see also [25], [22]) the authors show that the coefficients of the Hamiltonian  $H_{FB}^{(4)}$  in (3.20) related to monomials supported on Benjamin-Feir resonances are zero. The consequence of this “null condition” of the gravity water waves system in infinite depth is that  $H_{FB}^{(4)}$  is *integrable* and in particular

$$\begin{aligned} H_{FB}^{(4)} &= \Pi_{\text{triv}} \left( H_C^{(4)} + \frac{1}{2} \{F_3, H_C^{(3)}\} \right) = \frac{1}{4\pi} \sum_{k \in \mathbb{Z}} |k|^3 (|z_k|^4 - 2|z_k|^2 |z_{-k}|^2) \\ &\quad + \frac{1}{\pi} \sum_{\substack{k_1, k_2 \in \mathbb{Z}, \text{sign}(k_1) = \text{sign}(k_2) \\ |k_2| < |k_1|}} |k_1| |k_2|^2 ( -|z_{-k_1}|^2 |z_{k_2}|^2 + |z_{k_1}|^2 |z_{k_2}|^2 ), \end{aligned} \quad (3.31)$$

where  $\Pi_{\text{triv}}$  denotes the projection on *trivial* resonances. To prove that the corrections to the eigenvalues are integrable is equivalent to show that

$$H_{WB}^{(4)} = \Pi_{\text{triv}} H_{WB}^{(4)}, \quad (3.32)$$

$$H_{LB}^{(4)} = \Pi_{\text{triv}} H_{LB}^{(4)}. \quad (3.33)$$

The condition in (3.32) could be proved using a “generic” choice of the tangential sites. On the other hand it seems that the same argument does not fit for the proof of the equality (3.33). So, to prove the (3.33) we use an argument of identification of normal forms based on the ideas developed in [32]. In that paper the authors exploit the complete integrability of the Degasperis-Procesi equation (see for instance [26], [30]) using a sequence of constants of motion in order to deal with non-trivial 4-waves interactions. As it is well known the pure gravity water waves equation is not integrable, as it has been proved in [25]; furthermore no other constants of motion are known, except for the energy and the momentum.

Actually the identification argument of [32] works with *approximate* constants of motion and for our purpose it is sufficient to find one of it. In section 3.3 we construct such approximate conserved quantity by exploiting the formal integrability at order four.

**3.3. Approximate constant of motion for the gravity water waves.** The idea of the construction of an approximate conserved quantity  $K$  is the following: we know that the full BNF is integrable at order four, this implies that the Sobolev norms of the solutions of the truncated system  $H_C^{(2)} + H_{FB}^{(4)}$  (recall (3.19)) are preserved. Then quadratic

Hamiltonians of the form  $\sum_j j^{2n} |u_j|^2$ ,  $n \geq 0$ , are constants of motion for the truncated system. Written in formula this means that, setting  $K^{(2)}(u) = \sum_j j^{2n} |u_j|^2$ , we have

$$\{K^{(2)} \circ \Phi_{F^{(4)}}^{-1} \circ \Phi_{F^{(3)}}^{-1}, H\} \circ \Phi_{BF} = \{K^{(2)}, H \circ \Phi_{BF}\} = O(u^5).$$

It is easy to check that  $\Phi_{F^{(4)}}^{-1} \circ \Phi_{F^{(3)}}^{-1}$  differs from  $\Phi_{F^{(3)}+F^{(4)}}^{-1}$  by terms of order  $O(u^5)$ . Then by Taylor expanding we get

$$K^{(2)} \circ \Phi_{F^{(3)}+F^{(4)}}^{-1} = K^{(2)} + \{K^{(2)}, F^{(3)}\} + \{K^{(2)}, F^{(4)}\} + \frac{1}{2} \{\{K^{(2)}, F^{(3)}\}, F^{(3)}\} + O(u^5).$$

This justifies the choice of the Hamiltonian  $K$  constructed in the following proposition.

**Proposition 3.2.** *Consider the formal polynomial  $K = K^{(2)} + K^{(3)} + K^{(4)}$  where*

$$K^{(2)} := \sum_{j \in \mathbb{Z} \setminus \{0\}} j^2 |u_j|^2, \quad K^{(3)} := ad_{K^{(2)}} F_3, \quad K^{(4)} := ad_{K^{(2)}} F_4 + \frac{1}{2} \{ad_{K^{(2)}} F_3, F_3\} \quad (3.34)$$

and  $F_3, F_4$  are given in (3.18). Then we have

$$\{H_C, K\} = O(u^5) \quad (3.35)$$

where  $O(u^5)$  denotes a formal power series with a zero at the origin of order 5.

*Proof.* We first observe that  $K^{(2)}$  is the Hamiltonian of the linear Schrödinger equation  $i u_t = u_{xx}$  restricted to the invariant subspace of zero-average functions  $u(x)$ ,  $x \in \mathbb{T}$ . It is well known that  $ad_{K^{(2)}}$  has only trivial resonances of order 3 and 4. In particular, this implies that cubic Hamiltonians, as for instance  $F_3$ , are in the range of the adjoint action of  $K^{(2)}$ .

The equation (3.35) holds if and only if

$$\{K^{(2)}, H_C^{(2)}\} = 0, \quad (3.36)$$

$$\{K^{(2)}, H_C^{(3)}\} + \{K^{(3)}, H_C^{(2)}\} = 0, \quad (3.37)$$

$$\{K^{(2)}, H_C^{(4)}\} + \{K^{(4)}, H_C^{(2)}\} + \{K^{(3)}, H_C^{(3)}\} = 0. \quad (3.38)$$

We note that  $K^{(2)}$  is diagonal, hence (3.36) holds. From this fact we deduce that the adjoint actions  $ad_{K^{(2)}}$  and  $ad_{H_C^{(2)}}$  commutes on the intersection of their range  $Rg(H_C^{(2)}) \cap Rg(K^{(2)})$ .

The equation (3.37) holds since

$$ad_{H_C^{(2)}} [K^{(3)}] \stackrel{(3.34)}{=} ad_{K^{(2)}} ad_{H_C^{(2)}} [F_3] \stackrel{(3.18)}{=} \{K^{(2)}, H_C^{(3)}\},$$

where we used that the adjoint actions of  $H_C^{(2)}$  and  $K^{(2)}$  commute and  $F_3 \in Rg(H_C^{(2)}) \cap Rg(K^{(2)})$ .

The equality (3.38) is equivalent to

$$K^{(4)} = ad_{H_C^{(2)}}^{-1} \Pi_{Rg(H_C^{(2)})} \left( \{K^{(2)}, H_C^{(4)}\} + \{K^{(3)}, H_C^{(3)}\} \right). \quad (3.39)$$

We need to check that the choice of  $K^{(4)}$  in (3.34) implies the (3.39). First of all notice that, by (3.37), (3.18), (3.34) and the Jacobi identity, we have

$$\{K^{(3)}, H_C^{(3)}\} \stackrel{(3.34)}{=} \{\{K^{(2)}, F_3\}, H_C^{(3)}\} = -\{\{H_C^{(3)}, K^{(2)}\}, F_3\} - \{\{F_3, H_C^{(3)}\}, K^{(2)}\},$$

$$\{F_3, \{H_C^{(3)}, K^{(2)}\}\} - \{H_C^{(3)}, \{K^{(2)}, F_3\}\} \stackrel{(3.37)}{=} -\{F_3, ad_{H_C^{(2)}}(K^{(3)})\} - \{H_C^{(3)}, K^{(3)}\} = ad_{H_C^{(2)}} \{K^{(3)}, F_3\}.$$

Therefore

$$\begin{aligned} \{K^{(2)}, H_C^{(4)}\} + \{K^{(3)}, H_C^{(3)}\} &= ad_{K^{(2)}}(H_C^{(4)}) + ad_{K^{(2)}}(\{F_3, H_C^{(3)}\}) - \{\{H_C^{(3)}, K^{(2)}\}, F_3\} \\ &= ad_{K^{(2)}}(H_C^{(4)} + \frac{1}{2} \{F_3, H_C^{(3)}\}) + \frac{1}{2} \left( \{F_3, \{H_C^{(3)}, K^{(2)}\}\} - \{H_C^{(3)}, \{K^{(2)}, F_3\}\} \right) \\ &\stackrel{(3.18), (3.34)}{=} ad_{K^{(2)}}(H_C^{(4)} + \frac{1}{2} \{F_3, H_C^{(3)}\}) + \frac{1}{2} ad_{H_C^{(2)}} \{ad_{K^{(2)}} F_3, F_3\}. \end{aligned} \quad (3.40)$$

Moreover using (3.36) we have

$$\begin{aligned} \Pi_{\text{Ker}(H_C^{(2)})} ad_{K^{(2)}}(H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}) &= \Pi_{\text{Ker}(H_C^{(2)})} ad_{K^{(2)}} \Pi_{\text{Ker}(H_C^{(2)})}(H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}), \\ &\stackrel{(3.31)}{=} ad_{K^{(2)}} \Pi_{\text{triv}}(H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}) = 0, \end{aligned} \quad (3.41)$$

since  $K^{(2)}$  commutes with terms supported on trivial resonances. Then we have

$$(3.39) \stackrel{(3.40)}{=} ad_{K^{(2)}} ad_{H_C^{(2)}}^{-1} \Pi_{\text{Rg}(H_C^{(2)})}(H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}) + \frac{1}{2}\{ad_{K^{(2)}} F_3, F_3\} \stackrel{(3.18),(3.34)}{=} K^{(4)}.$$

This proves that (3.38) holds with  $K^{(4)}$  as in (3.34).  $\square$

**3.4. Identification of normal forms.** In this section we prove that the Hamiltonians  $H_{WB}^{(4)}, H_{LB}^{(4)}$  in (3.25), (3.28) are actually given in terms of the projections  $\Pi^{d_z=0}$  and  $\Pi^{d_z=2}$  of the full Birkhoff normal form Hamiltonian  $H_{FB}^{(4)}$  in (3.20). This will be done into two steps. First in Propositions 3.3 and 3.4 we show that (3.32) and (3.33) hold, i.e. the coefficients of monomials in  $H_{WB}^{(4)}, H_{LB}^{(4)}$  supported on Benjamin-Feir resonances (see (3.30)) are zero. Then in Proposition 3.6 we conclude the identification.

The next two propositions are based on the fact that two commuting Hamiltonians can be put simultaneously in Birkhoff normal form and the normalized terms have to be supported in the intersection of the kernels of *both* adjoint actions.

Then the main point is that the adjoint action of the constructed Hamiltonian  $K$  possesses only *trivial* 4-waves resonant interactions with two plus signs and two minus signs and the Benjamin-Feir are of this type (see (3.30)). Hence the common 4-waves resonances are just the trivial ones.

**Notation:** Following [32], given a Hamiltonian  $H$  we use the sub-index  $H_n$  to denote the transformed Hamiltonian after  $n$  steps of a normal form procedure.

**Proposition 3.3.** *Consider  $H_C$  in (3.14) and its approximate constant of motion  $K$  given in Proposition 3.2. Recalling (3.23) let us call  $\Phi := \Phi_{F_3^{(3, \leq 1)}} \circ \Phi_{\widehat{F}_4^{(4, \leq 1)}}$ . Then*

$$H_C \circ \Phi = H^{(2)} + H_C^{(3, \geq 2)} + H_{WB}^{(4)} + \widehat{H}_1^{(4, 2)} + \widehat{H}_1^{(4, \geq 3)} + R_2^{(\geq 5)}, \quad (3.42)$$

$$K \circ \Phi = K^{(2)} + K^{(3, \geq 2)} + W_2^{(4, 0)} + Q_2^{(\geq 5, \leq 1)} + K_1^{(\geq 4, \geq 2)}, \quad (3.43)$$

where  $\widehat{H}_1^{(4)}$  is given in (3.23),  $H_{WB}^{(4)}$  is in (3.25), and  $H_{WB}^{(4)}, W_2^{(4, 0)} \in \text{Ker}(K^{(2)}) \cap \text{Ker}(H_C^{(2)})$ .

*Proof.* One can reason as in Proposition 3.6 in [32]. The explicit form (3.42) comes from (3.23), (3.24), (3.25).  $\square$

**Proposition 3.4.** *Consider  $H_C$  in (3.14) and  $K$  given in Proposition 3.2. Recalling (3.26) let us call  $\Psi := \Phi_{F_3^{(3, 2)}} \circ \Phi_{\mathcal{F}_4^{(4, 2)}}$ . Then*

$$\begin{aligned} H_C \circ \Phi \circ \Psi &= H_C^{(2)} + H_{WB}^{(4)} + H_{LB}^{(4)} + R_4^{(\geq 5, \leq 2)} + H_4^{(\geq 3, \geq 3)}, \\ K \circ \Phi \circ \Psi &= K^{(2)} + W_2^{(4, 0)} + W_2^{(4, 2)} + Q_4^{(\geq 5, \leq 2)} + K_4^{(\geq 3, \geq 3)}, \end{aligned} \quad (3.44)$$

where  $H_{LB}^{(4)}$  is given in (3.28) and  $H_{LB}^{(4)}, W_2^{(4, 2)} \in \text{Ker}(K^{(2)}) \cap \text{Ker}(H_C^{(2)})$ .

*Proof.* Recall the definition of  $F^{(3, 2)}$  in (3.26). The equality (3.37) projected on  $d_z \leq 2$  gives

$$ad_{K^{(2)}}[F^{(3, 2)}] = K^{(3, 2)}. \quad (3.45)$$

Hence

$$\begin{aligned} H \circ \Phi \circ \Phi_{F^{(3, 2)}} &= H^{(2)} + H_C^{(3, 3)} + H_{WB}^{(4)} + \mathcal{H}^{(4, 2)} + \mathcal{H}^{(4, \geq 3)} + R_3^{(\geq 5, \leq 2)} + H_3^{(\geq 3, \geq 3)}, \\ K \circ \Phi \circ \Phi_{F^{(3, 2)}} &= K^{(2)} + W_2^{(4, 0)} + Q_3^{(\geq 4, \leq 2)} + K_3^{(\geq 3, \geq 3)}, \end{aligned} \quad (3.46)$$

where  $\mathcal{H}^{(4)}$  is in (3.26). Since (3.35) holds and the map  $\Phi \circ \Phi_{F^{(3, 2)}}$  is close to the identity we have that  $\{H_C, K\} \circ \Phi \circ \Phi_{F^{(3, 2)}} = O(u^5)$  then

$$\{\mathcal{H}^{(4, 2)}, K^{(2)}\} = \{Q_3^{(4, 2)}, H_C^{(2)}\}.$$

We note the following fact, which derives from the Jacobi identity and (3.36): if  $f \in \text{Ker}(H^{(2)})$  then  $\{f, K^{(2)}\} \in \text{Ker}(H^{(2)})$ , if  $f \in \text{Rg}(H^{(2)})$  then  $\{f, K^{(2)}\} \in \text{Rg}(H^{(2)})$ .



Then we have that  $\{\Pi_{\text{Ker}(H^{(2)})} \mathcal{H}^{(4,2)}, K^{(2)}\} \in \text{Ker}(H^{(2)})$

$$\{\Pi_{\text{Ker}(H^{(2)})} \mathcal{H}^{(4,2)}, K^{(2)}\} = -\{\Pi_{\text{Rg}(H^{(2)})} \mathcal{H}^{(4,2)}, K^{(2)}\} + \{H^{(2)}, Q_3^{(4,2)}\} \in \text{Rg}(H^{(2)}).$$

Thus  $\{\Pi_{\text{Ker}(H^{(2)})} \mathcal{H}^{(4,2)}, K^{(2)}\} = 0$  and

$$\Pi_{\text{Ker}(H^{(2)})} \mathcal{H}^{(4,2)} = \Pi_{\text{Ker}(H^{(2)})} \Pi_{\text{Ker}(K^{(2)})} \mathcal{H}^{(4,2)}.$$

By symmetry  $\Pi_{\text{Ker}(K^{(2)})} Q_3^{(4,2)} = \Pi_{\text{Ker}(H^{(2)})} \Pi_{\text{Ker}(K^{(2)})} Q_3^{(4,2)}$ . Hence

$$\Pi_{\text{Rg}(H^{(2)})} \Pi_{\text{Ker}(K^{(2)})} Q_3^{(4,2)} = \Pi_{\text{Rg}(K^{(2)})} \Pi_{\text{Ker}(H^{(2)})} \mathcal{H}^{(4,2)} = 0. \quad (3.47)$$

The function  $\mathcal{F}_4^{(4,2)}$  solves the homological equation

$$\{H^{(2)}, \mathcal{F}_4^{(4,2)}\} = \Pi_{\text{Rg}(H^{(2)})} \mathcal{H}^{(4,2)} \stackrel{(3.47)}{=} \Pi_{\text{Rg}(K^{(2)})} \Pi_{\text{Rg}(H^{(2)})} \mathcal{H}^{(4,2)}.$$

We now show that  $\mathcal{F}_4^{(4,2)}$  solves also the homological equation for the commuting Hamiltonian  $K \circ \Phi \circ \Phi_{F^{(3,2)}}$ . Indeed, by the fact that  $\text{ad}_{H^{(2)}}^{-1}$  commutes with  $\text{ad}_{K^{(2)}}$  on the intersection  $\text{Rg}(H^{(2)}) \cap \text{Rg}(K^{(2)})$ , we have

$$\{K^{(2)}, \mathcal{F}_4^{(4,2)}\} = \text{ad}_{H^{(2)}}^{-1} \{K^{(2)}, \Pi_{\text{Rg}(K^{(2)})} \Pi_{\text{Rg}(H^{(2)})} \mathcal{H}^{(4,2)}\},$$

and we get

$$\{K^{(2)}, \Pi_{\text{Rg}(K^{(2)})} \Pi_{\text{Rg}(H^{(2)})} \mathcal{H}^{(4,2)}\} = \{H^{(2)}, \Pi_{\text{Rg}(K^{(2)})} \Pi_{\text{Rg}(H^{(2)})} Q_1^{(4,2)}\}.$$

By (3.47) we have that the resonant term  $H_{LB}^{(4)} := \Pi_{\text{Ker}(H^{(2)})} \mathcal{H}^{(4,2)}$  belongs to the intersection of the kernels. This concludes the proof.  $\square$

**Remark 3.5.** Notice that the monomials of degree 4 with two bars (possibly supported on the Benjamin-Feir resonances) that belong to the kernel of  $\text{ad}_{K^{(2)}}$  (see (3.34)) are action-preserving (they are supported just on trivial resonances). Therefore Propositions 3.3, 3.4 imply the (3.32), (3.33).

**Proposition 3.6.** One has that (see (3.25), (3.28))

$$H_{WB}^{(4)} = \Pi_{\text{triv}} H_{WB}^{(4)} = \Pi^{d_z=0} H_{FB}^{(4)}, \quad H_{LB}^{(4)} = \Pi_{\text{triv}} H_{LB}^{(4)} = \Pi^{d_z=2} H_{FB}^{(4)} \quad (3.48)$$

where  $H_{FB}^{(4)}$  is in (3.20) (see also (3.31)).

*Proof.* We have already remarked that in [70], [25], [22] it has been proved that

$$(\Pi_{\text{Ker}(H_C^{(2)})} - \Pi_{\text{triv}})(H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}) = 0.$$

Hence

$$\Pi^{d_z=2} \Pi_{\text{Ker}(H_C^{(2)})} (H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}) = \Pi^{d_z=2} \Pi_{\text{triv}} (H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}). \quad (3.49)$$

We claim that

$$\Pi_{\text{triv}} \{F_3^{(3,1)}, H_C^{(3,3)}\} = \Pi_{\text{triv}} \{F_3^{(3,3)}, H_C^{(3,1)}\} = \Pi_{\text{triv}} \{F_3^{(3,2)}, H_C^{(3,0)}\} = \Pi_{\text{triv}} \{F_3^{(3,0)}, H_C^{(3,2)}\} = 0. \quad (3.50)$$

This implies that

$$\begin{aligned} \Pi_{\text{Ker}(H_C^{(2)})} \Pi^{d_z=2} (H_C^{(4)} + \frac{1}{2}\{F_3, H_C^{(3)}\}) &= \Pi_{\text{triv}} \Pi^{d_z=2} (H_C^{(4,2)} + \{F_3^{(3,\leq 1)}, H_C^{(3,\leq 1)}\} + \{F_3^{(3,\leq 1)}, H_C^{(3,\geq 2)}\}) \\ &+ \Pi_{\text{triv}} \Pi^{d_z=2} (H_C^{(4,2)} + \{F_3^{(3,2)}, H_C^{(3,0)}\} + \{F_3^{(3,2)}, H_C^{(3,2)}\} + \{F_3^{(3,3)}, H_C^{(3,1)}\}) \\ &= \Pi_{\text{triv}} \Pi^{d_z=2} (H_C^{(4,2)} + \{F_3^{(3,\leq 1)}, H_C^{(3,\leq 1)}\} + \{F_3^{(3,2)}, H_C^{(3,2)}\}). \end{aligned} \quad (3.51)$$

*Proof of the claim (3.50).* Let us consider the term  $\{F_3^{(3,1)}, H_C^{(3,3)}\}$ , where

$$F_3^{(3,1)} = \sum_{\substack{j_1, j_2 \in S, j_3 \in S^c \\ \sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 = 0 \\ \sigma_i = \pm}} C_{j_1, j_2, j_3}^{\sigma_1, \sigma_2, \sigma_3} u_{j_1}^{\sigma_1} u_{j_2}^{\sigma_2} u_{j_3}^{\sigma_3}, \quad H_C^{(3,3)} = \sum_{\substack{j_1, j_2, j_3 \in S^c \\ \sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 j_3 = 0 \\ \sigma_i = \pm}} \tilde{C}_{j_1, j_2, j_3}^{\sigma_1, \sigma_2, \sigma_3} u_{j_1}^{\sigma_1} u_{j_2}^{\sigma_2} u_{j_3}^{\sigma_3}, \quad (3.52)$$

for some coefficients  $C_{j_1, j_2, j_3}^{\sigma_1, \sigma_2, \sigma_3}, \tilde{C}_{j_1, j_2, j_3}^{\sigma_1, \sigma_2, \sigma_3} \in \mathbb{C}$ . Using (3.6) one gets

$$\{F_3^{(3,1)}, H_C^{(3,3)}\} = \sum_{\substack{j_1, j_2 \in S, k_1, k_2 \in S^c \\ \sigma_1 j_1 + \sigma_2 j_2 + \sigma_3 k_1 + \sigma_4 k_2 = 0 \\ \sigma_i = \pm}} P_{j_1, j_2, k_1, k_2}^{\sigma_1, \sigma_2, \sigma_3, \sigma_4} u_{j_1}^{\sigma_1} u_{j_2}^{\sigma_2} u_{k_1}^{\sigma_3} u_{k_2}^{\sigma_4}$$

for some  $P_{j_1, j_2, k_1, k_2}^{\sigma_1, \sigma_2, \sigma_3, \sigma_4} \in \mathbb{C}$ . A monomial of degree 4  $u_{j_1}^{\sigma_1} u_{j_2}^{\sigma_2} u_{k_1}^{\sigma_3} u_{k_2}^{\sigma_4}$  is supported on trivial resonances if  $\sigma_1 = -\sigma_2$ ,  $\sigma_3 = -\sigma_4$ ,  $j_1 = j_2$  and  $k_1 = k_2$ , since  $j_i \in S$ ,  $k_i \in S^c$ ,  $i = 1, 2$ . We note that by applying (3.6) the coefficients  $P_{j_1, j_1, k_1, k_1}^{\sigma_1, -\sigma_1, \sigma_3, -\sigma_3}$  should be given by combinations of coefficients of the form  $C_{j, j, 0}^{\sigma, -\sigma, \pm} \tilde{C}_{k, k, 0}^{\sigma', -\sigma', \mp}$  (up to permutation of the indexes). However, due to the restrictions in the sums (3.52) and the fact that  $0 \notin S \cup S^c$  (see (1.23)) such coefficients do not appear. Hence the (3.50) holds. The others equalities in (3.50) follow by reasoning in the same way.  $\square$

Now consider the Hamiltonian  $H_{LB}^{(4)}$  in (3.28) obtained by two steps of Weak normal form plus two steps of Linear Birkhoff normal form (see Propositions 3.3, 3.4). Then, recalling the explicit formula (3.26) we get

$$\begin{aligned} Z_2^{(4,2)} \stackrel{Rmk. 3.5}{=} \Pi^{d_z=2} \Pi_{\text{triv}} \left( \frac{1}{2} \{F^{(3,2)}, H_C^{(3,2)}\} + \frac{1}{2} \{F^{(3, \leq 1)}, H_C^{(3, \leq 1)}\} + \{F^{(3,1)}, H_C^{(3,3)}\} + H_C^{(4,2)} \right) \\ \stackrel{(3.50)}{=} \Pi^{d_z=2} \Pi_{\text{triv}} \left( H_C^{(4,2)} + \frac{1}{2} \{F^{(3,2)}, H_2^{(3,2)}\} + \frac{1}{2} \{F^{(3, \leq 1)}, H^{(3, \leq 1)}\} \right) \\ \stackrel{(3.51), (3.49)}{=} \Pi^{d_z=2} \Pi_{\text{triv}} \left( H_C^{(4)} + \frac{1}{2} \{F^{(3)}, H^{(3)}\} \right). \end{aligned}$$

This implies, recalling the (3.31), the second equality in the claim (3.48). The first one in (3.48) follows by using the projector  $\Pi^{d_z=0}$  and reasoning in the same way.  $\square$

#### 4. WEAK BIRKHOFF NORMAL FORM

The aim of this section is to construct a  $\zeta$ -parameter family of approximately invariant, finite dimensional tori supporting quasi-periodic motions with frequency  $\omega(\zeta)$ . We will impose the map  $\zeta \mapsto \omega(\zeta)$  to be a diffeomorphism and we will consider such approximate solutions as the starting point for the Nash-Moser algorithm. For this purpose we apply a Weak Birkhoff normal form procedure as explained in section 3.2.

**Definition 4.1. (Resonance).** *Let  $n \geq 3$  be an integer. We say that  $\{(j_i, \sigma_i)\}_{i=1}^n \subset (\mathbb{Z} \times \{\pm\})^n$  is a resonance of order  $n$  if*

$$\mathcal{M}_{\vec{\sigma}}(j_1, \dots, j_n) := \sum_{i=1}^n \sigma_i j_i = 0, \quad \mathcal{R}_{\vec{\sigma}}(j_1, \dots, j_n) := \sum_{i=1}^n \sigma_i \sqrt{|j_i|} = 0 \quad (4.1)$$

where  $\vec{\sigma} := (\sigma_1, \dots, \sigma_n)$ . Let us denote by  $n_+$  the number of  $\sigma_i = +$  and, up to reordering in the index  $i$ , let us assume that  $\sigma_i = +$  for  $i = 1, \dots, n_+$ .

A resonance of order  $n$  is said to be trivial if  $n$  is even,  $n_+ = n/2$  and

$$\{j_1, \dots, j_{n_+}\} = \{j_{n_++1}, \dots, j_n\}. \quad (4.2)$$

Otherwise we say that it is non-trivial.

**Remark 4.2.** *Note that resonances of odd order are non-trivial. We also remark that if  $\{(j_i, \sigma_i)\}_{i=1}^n$  is a trivial resonance then the functions  $\mathcal{M}_{\vec{\sigma}}$  and  $\mathcal{R}_{\vec{\sigma}}$  defined in (4.1) are identically zero.*

Consider a monomial of the form  $u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}$  and the Hamiltonian  $H_C^{(2)}$  in (3.14). Then we have (recall (3.8), (3.6))

$$\{M_C, u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}\} = -i \mathcal{M}_{\vec{\sigma}}(j_1, \dots, j_n) u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}, \quad (4.3)$$

$$\{H_C^{(2)}, u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}\} = -i \mathcal{R}_{\vec{\sigma}}(j_1, \dots, j_n) u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}. \quad (4.4)$$

We point out that the Hamiltonian (3.14) preserves the momentum and, in particular, by (4.3), each term  $H_C^{(p)}$  is sum of monomials Fourier supported on  $j_1, \dots, j_n$  such that  $\mathcal{M}_{\vec{\sigma}}(j_1, \dots, j_n) = 0$ .

We say that a monomial  $u_{j_1}^{\sigma_1} \dots u_{j_n}^{\sigma_n}$  is *resonant* if it belongs to the  $\text{Ker}(H_C^{(2)})$ , namely, by (4.4), it is Fourier supported on  $j_1, \dots, j_n$  such that  $\mathcal{R}_{\vec{\sigma}}(j_1, \dots, j_n) = 0$ .

For a finite dimensional subspace of  $H_0^1(\mathbb{T}) \times H_0^1(\mathbb{T})$

$$E := E_C := \text{span} \{e^{ijx} : 0 < |j| \leq C\}, \quad C > 0, \quad (4.5)$$

let  $\Pi_E$  denotes the corresponding  $L^2$ -projector on  $E$ .

The aim of this section is to prove the following.

**Proposition 4.3.** *Let  $s \geq 0$ . For any generic choice of the set  $S$  of tangential sites of the form (1.23) there exist  $r > 0$ , depending on  $S$ , and an analytic symplectic change of coordinates*

$$\Phi_B: \mathcal{B}_r(0, H_0^s(\mathbb{T}) \times H_0^s(\mathbb{T})) \rightarrow H_0^s(\mathbb{T}) \times H_0^s(\mathbb{T}), \quad \Phi_B = \text{I} + \Psi, \quad \Psi = \Pi_E \circ \Psi \circ \Pi_E \quad (4.6)$$

where  $E$  is a finite dimensional space as in (4.5), such that  $M_{\mathbb{C}} \circ \Phi_B = M_{\mathbb{C}}$  where  $M_{\mathbb{C}}$  is the momentum Hamiltonian in (3.8) and the Hamiltonian  $H_{\mathbb{C}}$  in (3.14) transforms into

$$\mathcal{H}_{\mathbb{C}} := H_{\mathbb{C}} \circ \Phi_B = H_{\mathbb{C}}^{(2)} + \mathcal{H}_{\mathbb{C}}^{(4,0)} + \sum_{k=3}^7 \mathcal{H}_{\mathbb{C}}^{(2k,0)} + \mathcal{H}_{\mathbb{C}}^{(\geq 16, \leq 1)} + \mathcal{H}_{\mathbb{C}}^{(\geq 3, \geq 2)}, \quad (4.7)$$

where

$$\mathcal{H}_{\mathbb{C}}^{(4,0)} := \frac{1}{4\pi} \sum_{k \in S} |k|^3 |u_k|^4 + \frac{1}{\pi} \sum_{\substack{k_1, k_2 \in S, \text{sign}(k_1) = \text{sign}(k_2) \\ |k_2| < |k_1|}} |k_1| |k_2|^2 |u_{k_1}|^2 |u_{k_2}|^2 \quad (4.8)$$

is the fourth order Hamiltonian provided by Zakharov-Dyachenko in [70] and Craig-Worfolk in [25] restricted to monomials supported on the tangential set  $S$ , and  $\mathcal{H}_{\mathbb{C}}^{(2k,0)} = \Pi_{\text{Ker}(H_{\mathbb{C}}^{(2)})} \mathcal{H}_{\mathbb{C}}^{(2k,0)}$  with  $k = 3, \dots, 7$  are supported only on trivial resonances.

In the following subsection we show that one can choose “generically” the set of indexes  $S$  in such a way that in the Hamiltonian in (3.14) there are no monomials  $u_{j_1}^{\sigma_1} \cdots u_{j_p}^{\sigma_p}$ , with at most one index  $j_i \in S^c$ , supported on a non-trivial resonance. In turn this implies the existence of a finite dimensional submanifold that is invariant for the truncated normal form. The integrability of the normalized Hamiltonian gives that such manifold is foliated by families of tori. In subsection 4.2 we prove Proposition 4.3.

**4.1. Resonances.** The aim of this subsection is to prove the following result.

**Proposition 4.4.** *If  $\{(j_i, \sigma_i)\}_{i=1}^n$  is a non-trivial resonance of order  $n$  with at most one index  $j_i$  outside  $S$  then there exists a non-identically zero polynomial  $\mathcal{P}: \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $\mathcal{P}(j_1, \dots, j_n) = 0$ .*

We remark that there are just a finite number of non-trivial resonances with one index out of  $S$ , since  $S$  is finite and the momentum is preserved. Then by Proposition 4.4 there exists a generic choice of tangential sites such that in the Hamiltonian (3.14) there are no non-trivial resonant monomials  $u_{j_1}^{\sigma_1} \cdots u_{j_p}^{\sigma_p}$ , with at most one index  $j_i \in S^c$ .

Before giving the proof of the proposition above we need the following algebraic lemma.

**Lemma 4.5.** *Let  $f: \mathbb{R}^{n-1} \rightarrow \mathbb{R}$  be an algebraic function and let  $(\bar{x}_1, \dots, \bar{x}_{n-1}) \in \mathbb{R}^{n-1}$  be a zero of  $f$ .*

$$\nabla f(\bar{x}_1, \dots, \bar{x}_{n-1}) \neq 0 \quad (4.9)$$

then there exists a non-identically zero polynomial  $Q: \mathbb{R}^{n-1} \rightarrow \mathbb{R}$  such that  $Q(\bar{x}_1, \dots, \bar{x}_{n-1}) = 0$ .

*Proof.* Since  $f$  is an algebraic function there exists a non-identically zero polynomial  $P: \mathbb{R}^n \rightarrow \mathbb{R}$  such that

$$P(x_1, \dots, x_{n-1}, f(x_1, \dots, x_{n-1})) = 0 \quad \forall (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}. \quad (4.10)$$

We can write the polynomial  $P$  in the following form

$$P(x_1, \dots, x_n) = p_0(x_1, \dots, x_{n-1}) + \sum_{k \geq m} p_k(x_1, \dots, x_{n-1}) x_n^k \quad (4.11)$$

where  $m \geq 1$  and  $p_k$  are polynomials of  $n-1$  real variables. Without loss of generality we can assume that  $p_m$  is non-identically zero. By (4.10) we have that

$$P(\bar{x}_1, \dots, \bar{x}_{n-1}, 0) = 0.$$

If  $p_0$  is non-identically zero then we conclude by considering  $Q := p_0$ . Otherwise, by (4.9) there exists  $j \in \{1, \dots, n-1\}$  such that  $\partial_{x_j} f(\bar{x}_1, \dots, \bar{x}_{n-1}) \neq 0$ . By (4.10), (4.11) we have that for all  $(x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$

$$\begin{aligned} 0 &= \partial_{x_j}^m P(x_1, \dots, x_{n-1}, f(x_1, \dots, x_{n-1})) \\ &= \sum_{k \geq m} \sum_{\alpha=0}^m \binom{m}{\alpha} \partial_{x_j}^{m-\alpha} p_k(x_1, \dots, x_{n-1}) \partial_{x_j}^{\alpha} (f(x_1, \dots, x_{n-1})^k). \end{aligned} \quad (4.12)$$

Now we consider the above relation when it is evaluated at  $(\bar{x}_1, \dots, \bar{x}_{n-1})$ . Since  $k \geq m$  there is just one term in (4.12) which is not zero, the one involving just derivatives of  $f$ , but not  $f$  itself (recall that  $f(\bar{x}_1, \dots, \bar{x}_{n-1}) = 0$ ). Such term corresponds to  $k = m = \alpha$  and gives

$$0 = \partial_{x_j}^m P(\bar{x}_1, \dots, \bar{x}_{n-1}, 0) = m! \left( \partial_{x_j} f(\bar{x}_1, \dots, \bar{x}_{n-1}) \right)^m p_m(\bar{x}_1, \dots, \bar{x}_{n-1}).$$

Hence we deduce that  $p_m(\bar{x}_1, \dots, \bar{x}_{n-1}) = 0$  and we can choose  $Q := p_m$ .  $\square$

**Proof of Proposition 4.4.** Note that if  $\{(j_i, \sigma_i)\}_{i=1}^n$  is a non-trivial resonance and all the  $j_i$ 's belong to  $S$  then, thanks to the momentum conservation (recall  $\mathcal{M}_{\vec{\sigma}}$  in (4.1)), we can consider

$$\mathcal{P}(x_1, \dots, x_n) := \sum_{i=1}^n \sigma_i x_i.$$

Hence let us consider the case in which there is one  $j_i \notin S$  and, to simplify the notation, suppose that  $i = n$ . We can assume that  $0 < n_+ < n$  otherwise  $\mathcal{R}(j_1, \dots, j_n)$  in (4.1) is strictly positive.

By using the first equation in (4.1) we have that  $\sigma_n j_n = -\sum_{i \neq n} \sigma_i j_i$ . Then, by using the second equation in (4.1), we have that  $(j_1, \dots, j_{n-1})$  is a zero of the algebraic function

$$f(x_1, \dots, x_{n-1}) := \sum_{i=1}^{n-1} \sigma_i \sqrt{|x_i|} + \sigma_n \sqrt{\left| \sum_{i \neq n} \sigma_i x_i \right|}.$$

We have

$$\partial_{x_1} f(j_1, \dots, j_{n-1}) = \frac{\sigma_1 \operatorname{sgn}(j_1)}{2\sqrt{|j_1|}} + \frac{\sigma_1 \sigma_n \operatorname{sgn}(\sum_{i \neq n} \sigma_i j_i)}{2\sqrt{|j_n|}}.$$

Thus  $\partial_{x_1} f(j_1, \dots, j_{n-1}) = 0$  if and only if  $|j_1| = |j_n|$  and  $\operatorname{sgn}(j_1) = \sigma_n \operatorname{sgn}(\sigma_n j_n)$ , which implies  $j_1 = j_n$ . Since we assumed that  $j_n \in S^c$  and  $j_1 \in S$  we have that  $\partial_{x_1} f(j_1, \dots, j_{n-1}) \neq 0$ . By Lemma 4.5 this means that  $(j_1, \dots, j_{n-1})$  is the zero of some non-identically zero polynomial  $Q(x_1, \dots, x_{n-1})$ . We conclude by considering  $\mathcal{P}(x_1, \dots, x_n) = x_n Q(x_1, \dots, x_{n-1})$ .  $\square$

**4.2. Proof of Proposition 4.3.** The change of coordinates  $\Phi_B$  is constructed via a weak version of the Birkhoff normal form algorithm as described in section 3.2. Such method has been implemented in several papers (see for instance [8], [32]), so here we just resume the scheme that we follow to obtain the result in Proposition 4.3.

The map  $\Phi_B$  is defined as the composition of several symplectic changes of variables  $\Phi_k$  that are time-one flow maps of homogeneous Hamiltonians  $F^{(k+2, \leq 1)}$  such that  $\Pi^{d_z \leq 1} F^{(k+2, \leq 1)} = F^{(k+2, \leq 1)}$ , namely  $F^{(k+2, \leq 1)}$  is Fourier supported on set of indexes with at most one index out of  $S$ . Since the Hamiltonian  $H_{\mathbb{C}}$  Poisson commutes with the momentum  $\mathcal{M}_{\mathbb{C}}$  we look for generators  $F^{(k+2, \leq 1)}$  with the same property. By momentum conservation and the fact that  $\Pi^{d \leq 1} F^{(k+2, \leq 1)} = F^{(k+2, \leq 1)}$ , we have that the Fourier support of  $F^{(k+2, \leq 1)}$  is given by a finite set of integer indexes. So  $\Phi_k$  it is a well defined, analytic and invertible map of the phase space as time-one flow map of an ordinary differential equation with polynomial vector field.

Now we describe a generic step of weak Birkhoff normal form and we see how we choose the generators  $F^{(k+2, \leq 1)}$ . Let  $k \geq 1$ , then after  $k$ -steps of weak Birkhoff normal form the Hamiltonian has the form

$$\begin{cases} H_k := H_{k-1} \circ \Phi_k^{-1} = H_{\mathbb{C}}^{(2)} + Z_k^{(\leq k+2, \leq 1)} + R_k^{(\geq k+3, \leq 1)} + H_{k-1}^{(\geq 3, \geq 2)} \\ H_0 = H_{\mathbb{C}}, \quad Z_0^{(\leq 2, \leq 1)} = 0, \quad R_0^{(3, \leq 1)} = H_{\mathbb{C}}^{(3, \leq 1)}, \end{cases} \quad (4.13)$$

where  $\Pi_{\operatorname{Ker}(H^{(2)})} Z_k^{(\leq k+2, \leq 1)} = Z_k^{(\leq k+2, \leq 1)}$  is given by

$$Z_k^{(\leq k+2, \leq 1)} := Z_{k-1}^{(\leq k+1, \leq 1)} + \Pi_{\operatorname{Ker}(H^{(2)})} R_{k-1}^{(k+2, \leq 1)}.$$

By Proposition 4.4, for a generic choice of the set  $S$  we have

- $Z_k^{(\leq k+2, 1)} = 0$  for any  $k$ ,
- $Z_k^{(k+2, \leq 1)} = 0$  if  $k$  is odd,
- $Z_k^{(k+2, 0)}$  depends only on the actions  $|u_j|^2$  with  $j \in \mathbb{Z}$ ,  $j \neq 0$  if  $k$  is even.

The Hamiltonian  $H_k$  in (4.13) is obtained by choosing  $F^{(k+2, \leq 1)}$  as the solution of the following equation

$$\{H_{\mathbb{C}}^{(2)}, F^{(k+2, \leq 1)}\} = \Pi_{\mathbf{Rg}(H^{(2)})} R_{k-1}^{(k+2, \leq 1)}, \quad (4.14)$$

and Taylor expanding  $H_{k-1} \circ \Phi_k^{-1}$ . We perform  $N = 13$  steps of the Weak BNF procedure described above. Then we consider  $\Phi_B = \Phi_{13}^{-1} \circ \dots \circ \Phi_1^{-1}$  and we set

$$\mathcal{H}_{\mathbb{C}}^{(2k, 0)} = Z_{13}^{(2k, 0)} \quad 2 \leq k \leq 7, \quad \mathcal{H}_{\mathbb{C}}^{(\geq 15, \leq 1)} = R_{13}^{(\leq 15, \leq 1)}, \quad \mathcal{H}_{\mathbb{C}}^{(\geq 3, \geq 2)} = H_{13}^{(\geq 3, \geq 2)}.$$

By the discussion of section 3 we conclude that the expression of the normalized Hamiltonian  $\mathcal{H}^{(4, 0)}$  is the one obtained by ([14], [25], [70]) through a formal procedure and then restricted to monomials supported on the tangential sites  $S$ . Indeed the map  $\Phi_B$  coincides with the map  $\Phi$  in Proposition 3.3 up to order 5, and hence  $\mathcal{H}^{(4, 0)}$  is equal to  $H_{WB}^{(4)}$  appearing in (3.42). By Proposition 3.6 (see the first equality in (3.48)) and the expression (3.31) we get the (4.8).

## 5. THE NONLINEAR FUNCTIONAL SETTING

In this section we consider the Hamiltonian  $\mathcal{H}_{\mathbb{C}}$  in (4.7) obtained after the weak Birkhoff normal form procedure of section 4. We show that actually  $\mathcal{H}_{\mathbb{C}}$  possesses *approximately* invariant manifolds foliated by tori supporting quasi-periodic motions. To do this, we first introduce action-angle coordinates in the neighborhood of such manifolds. Lemma 5.2 guarantees that the dynamics on the manifold is non-isochronous. In subsection 5.2 we introduce a nonlinear functional whose zeros are quasi-periodic solutions for the whole Hamiltonian  $\mathcal{H}_{\mathbb{C}}$ . This is stated in Theorem 5.6 which will imply our main result.

**5.1. Action-angle variables.** On the submanifold  $\{z = 0\}$  we put the following action-angle variables

$$\mathbb{T}^{\nu} \times \mathbb{R}_+^{\nu} \longrightarrow \{z = 0\}, \quad (\theta, I) \longmapsto v = \sum_{j \in S} \sqrt{I_j} e^{-i\theta_j} e^{ijx}, \quad (5.1)$$

with  $\mathbb{R}_+^{\nu} := (0, \infty)^{\nu}$ . The symplectic form in (3.1) (or equivalently (3.5) in the complex coordinates) restricted to the subspace  $H_S$  (see (3.15)) transforms into the 2-form  $\sum_{j \in S} dI_j \wedge d\theta_j$ . Recalling (4.7), we have that the Hamiltonian  $\mathcal{H}_{\mathbb{C}}^{(\leq 14, 0)}(\theta, I, 0) = \sum_{j \in S} |j|^{\frac{1}{2}} I_j + \mathcal{H}_{\mathbb{C}}^{(4, 0)}(I) + \sum_{k=3}^7 \mathcal{H}_{\mathbb{C}}^{(2k, 0)}(I)$  depends only by the actions  $I$  and its equations of motion read as

$$\begin{cases} \dot{\theta}_j = \partial_{I_j} \mathcal{H}_{\mathbb{C}}^{(\leq 14, 0)}(\theta, I, 0) = |j|^{\frac{1}{2}} + \sum_{k \in S} \mathbb{A}_j^k I_k + O(I^2), \\ \dot{I}_j = -\partial_{\theta_j} \mathcal{H}_{\mathbb{C}}^{(\leq 14, 0)}(\theta, I, 0) = 0, \end{cases} \quad j \in S, \quad (5.2)$$

where the matrix  $\mathbb{A}$  is the  $\nu \times \nu$  symmetric matrix associated to the quadratic form (see (4.8))

$$\frac{1}{2} \mathbb{A} I \cdot I := \mathcal{H}_{\mathbb{C}}^{(4, 0)}(I) := \frac{1}{4\pi} \sum_{k \in S} |k|^3 I_k^2 + \frac{1}{\pi} \sum_{\substack{k_1, k_2 \in S, \text{sign}(k_1) = \text{sign}(k_2) \\ |k_2| < |k_1|}} |k_1| |k_2|^2 I_{k_1} I_{k_2}. \quad (5.3)$$

Let  $0 < \varepsilon \ll 1$  and rescale  $I \mapsto \varepsilon^2 I$ , so that the *frequency-amplitude map* can be written as

$$\alpha(I) = \bar{\omega} + \varepsilon^2 \mathbb{A} I + O(\varepsilon^4), \quad O(\varepsilon^4) = \sum_{k=2}^6 \varepsilon^{2k} \mathbf{b}_k(\underbrace{I, \dots, I}_{k\text{-times}}), \quad (5.4)$$

where  $\bar{\omega}$  is the linear frequencies vector in (1.26), and where  $\mathbf{b}_k$  are  $k$ -linear functions of the variables  $I$ .

In order to work in a small neighborhood of the prefixed torus  $\{I \equiv \zeta\}$ , with  $\zeta \in \mathbb{R}_+^{\nu}$ , it is advantageous to introduce a set of coordinates  $(\theta, y, z, \bar{z}) \in \mathbb{T}^{\nu} \times \mathbb{R}^{\nu} \times H_S^{\frac{1}{2}} \times H_S^{\frac{1}{2}}$  adapted to it, defined by

$$u = A_{\varepsilon}(\theta, y, z) = \varepsilon v_{\varepsilon}(\theta, y) + \varepsilon^b z \iff \begin{cases} u_j := \varepsilon \sqrt{\zeta_j + \varepsilon^{2b-2} y_j} e^{-i\theta_j} e^{ijx}, & j \in S, \\ u_j := \varepsilon^b z_j, & j \in S^c, \end{cases} \quad (5.5)$$

with  $b > 1$  chosen as

$$b := 1 + \frac{a}{2}, \quad 0 < a \ll 1. \quad (5.6)$$

We define

$$\left[ \frac{u}{\bar{u}} \right] := \mathbf{A}_\varepsilon(\theta, y, z, \bar{z}) := \left( \frac{A_\varepsilon(\theta, y, z)}{A_\varepsilon(\theta, y, z)} \right). \quad (5.7)$$

For the tangential sites  $S := \{\bar{j}_1, \dots, \bar{j}_\nu\}$  (see (1.23)) we will also denote  $\theta_{\bar{j}_i} := \theta_i$ ,  $y_{\bar{j}_i} := y_i$ ,  $\zeta_{\bar{j}_i} := \zeta_i$ ,  $i = 1, \dots, \nu$ . After the time-rescaling  $t \mapsto \varepsilon^{2b}t$  we can consider the symplectic 2-form

$$\mathcal{W} := \sum_{i=1}^{\nu} dy_i \wedge d\theta_i - i \sum_{j \in S^c} dz_j \wedge d\bar{z}_j = \left( \sum_{i=1}^{\nu} dy_i \wedge d\theta_i \right) \oplus \Omega_{S^\perp}, \quad (5.8)$$

where  $\Omega_{S^\perp}$  is the symplectic form  $\Omega$  restricted to the subspace  $H_S^\perp$  in (3.15).

**Remark 5.1.** Note that  $\mathcal{W}$  in (5.8) has the form (2.47) when  $(h, \bar{h})$  is restricted to the subspace  $H_S^\perp$ . Moreover, in the real coordinates, the symplectic form  $\mathcal{W}$  in (5.8) reads

$$\widetilde{\mathcal{W}} := \sum_{i=1}^{\nu} dy_i \wedge d\theta_i + \sum_{j \in S^c} d\psi_j \wedge d\eta_{-j} = \left( \sum_{i=1}^{\nu} dy_i \wedge d\theta_i \right) \oplus \widetilde{\Omega}_{S^\perp} = d\mathcal{V}, \quad (5.9)$$

where  $\widetilde{\Omega}_{S^\perp}$  is the symplectic form  $\widetilde{\Omega}$  in (3.1) restricted to the subspace  $H_S^\perp$  in (3.15) and  $\mathcal{V}$  is the contact 1-form on  $\mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$  defined by  $\mathcal{V}_{(\theta, y, W)} : \mathbb{R}^\nu \times \mathbb{R}^\nu \times H_S^\perp \rightarrow \mathbb{R}$ ,

$$\mathcal{V}_{(\theta, y, W)}[\widehat{\theta}, \widehat{y}, \widehat{W}] := y \cdot \widehat{\theta} + \frac{1}{2}(JW, \widehat{W})_{L^2}. \quad (5.10)$$

The Hamiltonian  $\mathcal{H}_C$  in (4.7) becomes

$$H_{C, \varepsilon} := \varepsilon^{-2b} \mathcal{H}_C \circ \mathbf{A}_\varepsilon. \quad (5.11)$$

In the following lemma we prove that, under an appropriate choice of the tangential set (1.23), the function (5.4) is a diffeomorphism for  $\varepsilon$  small enough and then the system (5.2) is integrable and *non-isochronous*.

**Lemma 5.2. (Twist condition).** For a generic choice of the set  $S$  one has  $|\det \mathbb{A}| \geq \frac{1}{(4\pi)^\nu}$ .

*Proof.* Recall (5.3). The entries of  $\mathbb{A}$  are homogeneous polynomials of the tangential sites, and so it is the  $\det \mathbb{A}$ . We claim that  $\det \mathbb{A}$  is a *not identically zero* polynomial. Hence the result follows since  $4\pi\mathbb{A}$  is an integer matrix and then  $(4\pi)^\nu \det \mathbb{A}$  is a non-zero integer number. To prove the claim we consider the matrix  $\mathbb{A}(\bar{j}_i)$  restricted at  $\bar{j}_i = \lambda$ ,  $i = 1, \dots, \nu$ , for some  $\lambda > 0$ . Such matrix has the form

$$\mathbb{A}(\lambda, \dots, \lambda) = \lambda^3 \pi^{-1} D \quad (5.12)$$

where  $D$  is a  $\nu \times \nu$  matrix whose diagonal elements are all  $1/4$  and the appearing out of diagonal are just 0 and  $1/2$ . Then the matrix

$$4\pi\lambda^{-3} \mathbb{A}(\lambda, \dots, \lambda) = 4D \quad (5.13)$$

is the identity in  $\mathbb{Z}/2\mathbb{Z}$  and in particular its determinant is 1. This means that the determinant of the matrix in (5.13) is an odd number, in particular it is different from zero. This concludes the proof.  $\square$

As a consequence of Lemma 5.2 the map in (5.4) is invertible and we denote

$$\zeta := \zeta(\omega) := \alpha^{(-1)}(\omega) = \varepsilon^{-2} \mathbb{A}^{-1}(\omega - \bar{\omega}) + O(\varepsilon^2). \quad (5.14)$$

In the following it will be convenient to consider the Hamiltonian  $H_{C, \varepsilon}$  in (5.11) expressed in terms of real variables. Recall the splitting in (3.16). Then the variables  $(\eta, \psi)$  in the action-angle variables in (5.5), (5.7) read

$$\begin{pmatrix} \eta \\ \psi \end{pmatrix} := \Lambda^{-1} \mathbf{A}_\varepsilon \left( \theta, y, \Lambda \left[ \begin{smallmatrix} \tilde{\eta} \\ \tilde{\psi} \end{smallmatrix} \right] \right) \quad (5.15)$$

Using this notation, the Hamiltonian  $H_{C, \varepsilon}(\theta, y, z, \bar{z})$  in (5.11) expressed in terms of real variables is

$$H_\varepsilon(\theta, y, \tilde{\eta}, \tilde{\psi}) := \varepsilon^{-2b} \mathcal{H}_C \circ \mathbf{A}_\varepsilon \left( \theta, y, \Lambda \left[ \begin{smallmatrix} \tilde{\eta} \\ \tilde{\psi} \end{smallmatrix} \right] \right) = \varepsilon^{-2b} H \circ \Lambda^{-1} \circ \Phi_B \circ \mathbf{A}_\varepsilon \left( \theta, y, \Lambda \left[ \begin{smallmatrix} \tilde{\eta} \\ \tilde{\psi} \end{smallmatrix} \right] \right) \quad (5.16)$$

where we used also (4.7) and (3.14). In the following we shall need a homogeneous expansion of the Hamiltonian  $H_\varepsilon(\theta, y, \tilde{\eta}, \tilde{\psi})$  in (5.16) in the variables  $y, \tilde{\eta}, \tilde{\psi}$ . Notice that the map  $\Lambda$  in (3.4) does not affect the homogeneity degree



in the normal variables. For simplicity we expand the Hamiltonian  $H_{\mathbb{C},\varepsilon}$  in degree of homogeneity in  $y$  and in the complex variables  $z, \bar{z}$  (see (3.16)). In order to lighten the notation (3.17) we shall write

$$R(v^{k-q}z^q) = R_k(v^{\alpha_1}\bar{v}^{\beta_1}z^{\alpha_2}\bar{z}^{\beta_2}), \quad \alpha_1 + \beta_1 = k - q, \quad \alpha_2 + \beta_2 = q.$$

We also define (see (5.4))

$$\mathbb{B}(\zeta)y := \sum_{k=2}^7 \varepsilon^{2(k-1)} \mathfrak{b}_k(\underbrace{\zeta, \dots, \zeta}_{(k-1)\text{-times}}, y). \quad (5.17)$$

Hence, by (3.14), (4.7), and writing  $v_\varepsilon := v_\varepsilon(\theta, y)$ , we have that

$$\begin{aligned} H_{\mathbb{C},\varepsilon}(\theta, y, z, \bar{z}) &= e(\zeta) + \alpha(\zeta) \cdot y + \int_{\mathbb{T}} |D|^{\frac{1}{2}} z \cdot \bar{z} dx + \frac{\varepsilon^{2b}}{2} (\mathbb{A} + \mathbb{B}(\zeta)) y \cdot y + \varepsilon^{4b} O(y^3) \\ &+ \sum_{k=1}^{15} \varepsilon^k R(v_\varepsilon^k z^2) + \sum_{k=3}^{15} \varepsilon^{k-3+b} \sum_{q=3}^k \varepsilon^{(b-1)(q-3)} R(v^{k-q} z^q) + \varepsilon^{-2b} \mathcal{H}^{(\geq 16)}(\varepsilon v_\varepsilon + \varepsilon^b z) \end{aligned} \quad (5.18)$$

where  $e(\zeta)$  is a constant,  $\alpha(\zeta)$  is the rescaled frequency-amplitude map (5.4) and  $\mathbb{A}$  is in (5.3).

**5.2. The functional equation.** We write the Hamiltonian in (5.16) (possibly eliminating constant terms depending only on  $\zeta$  which are irrelevant for the dynamics) as

$$H_\varepsilon = \mathcal{N} + P, \quad W := \begin{pmatrix} \tilde{\eta} \\ \tilde{\psi} \end{pmatrix} \quad (5.19)$$

$$\mathcal{N}(\theta, y, W) = \alpha(\zeta) \cdot y + \frac{1}{2}(N(\theta)W, W)_{L^2}, \quad \frac{1}{2}(N(\theta)W, W)_{L^2} := \frac{1}{2}((d\nabla_W H_\varepsilon)(\theta, 0, 0)[W], W)_{L^2},$$

where  $\mathcal{N}$  describes the linear dynamics normal to the torus, and  $P := H_\varepsilon - \mathcal{N}$  collects the nonlinear perturbative effects. Note that both  $\mathcal{N}$  and  $P$  depend on  $\omega$  through the map  $\omega \mapsto \zeta(\omega)$ .

We consider  $H_\varepsilon$  as a  $(\omega, \varepsilon)$ -parameter family of Hamiltonians and we note that, for  $P = 0$ ,  $H_\varepsilon$  possess an invariant torus at the origin with frequency  $\omega$ , which we want to continue to an invariant torus for the full system.

We will select the frequency parameters from the following set (recall (5.14))

$$\Omega_\varepsilon := \{\omega \in \mathbb{R}^\nu : \zeta(\omega) \in [1, 2]^\nu\}. \quad (5.20)$$

Set (see (5.6))

$$\gamma = \varepsilon^{2b}, \quad \tau := 3\nu + 7 \quad (5.21)$$

and consider two constants  $\mathbb{C}_1, \mathbb{C}_2 > 0$  depending on  $S$ . We define the non-resonant sets

$$\begin{aligned} \mathcal{G}_0^{(0)} &:= \{\omega \in \Omega_\varepsilon : |\omega \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\}, \\ \mathcal{G}_0^{(1)} &:= \{\omega \in \Omega_\varepsilon : |(\bar{\omega} - \varepsilon^2 m_1 \mathbf{v}) \cdot \ell + \varepsilon^2 \mathbb{A} \zeta(\omega) \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\}, \end{aligned} \quad (5.22)$$

$$\begin{aligned} \mathcal{G}_0^{(2)}(\mathbb{C}_1, \mathbb{C}_2) &:= \left\{ \omega \in \Omega_\varepsilon : |(\bar{\omega} - \varepsilon^2 m_1 \mathbf{v} + \varepsilon^2 \mathbb{A} \zeta(\omega)) \cdot \ell + \sigma \sqrt{|j|} - \sigma' \sqrt{|k|}| \geq \frac{\gamma}{\langle \ell \rangle^\tau}, \right. \\ &\quad \left. |\ell| \leq \mathbb{C}_1, \max\{|j|, |k|\} \leq \mathbb{C}_2, \mathbf{v} \cdot \ell + j - k = 0, \sigma, \sigma' = \pm 1 \right\} \end{aligned} \quad (5.23)$$

where  $\mathbb{A}$  is defined in (5.3) and

$$m_1 = m_1(\omega) := w \cdot \zeta(\omega), \quad w := (w_j)_{j \in S}, \quad w_j := |j|j, \quad \mathbf{v} := (\mathbf{v}_j)_{j \in S}, \quad \mathbf{v}_j := j. \quad (5.24)$$

We require that

$$\omega \in \mathcal{G}_0 := \mathcal{G}_0^{(0)} \cap \mathcal{G}_0^{(1)} \cap \mathcal{G}_0^{(2)}. \quad (5.25)$$

**Lemma 5.3.** *For generic choices of  $S$  and any constants  $\mathbb{C}_1, \mathbb{C}_2 > 0$  depending only on  $S$ , we have that  $|\Omega_\varepsilon \setminus \mathcal{G}_0| \leq C_* \varepsilon^{2(\nu-1)\gamma}$  for some  $C_* = C_*(S) > 0$ .*

*Proof.* The estimate for  $\Omega_\varepsilon \setminus \mathcal{G}_0^{(0)}$  follows using (5.4), (5.14) and Lemma 5.2. Consider now the set  $\mathcal{G}_0^{(1)}$ . We have that

$$\Omega_\varepsilon \setminus \mathcal{G}_0^{(1)} = \bigcup_{\ell \in \mathbb{Z}^\nu} T_\ell, \quad T_\ell := \{\omega \in \Omega_\varepsilon : |\phi(\omega, \ell)| < \gamma \langle \ell \rangle^{-\tau}\}, \quad \phi(\omega, \ell) := (\bar{\omega} - \varepsilon^2 m_1 \mathbf{v}) \cdot \ell + \varepsilon^2 \mathbb{A} \zeta(\omega) \cdot \ell. \quad (5.26)$$

First we write

$$m_1 \mathbf{v} = \mathbb{V} \zeta, \quad \mathbb{V} := \mathbf{v} w^T. \quad (5.27)$$

Then  $\nabla_\zeta \phi(\omega, \ell) = (\mathbb{A}^T - \mathbb{V}^T) \ell$ . Now we prove that, for generic choices of  $S$ , we have

$$\det(\mathbb{A} - \mathbb{V}) \neq 0. \quad (5.28)$$

Since  $\mathbb{V}$  and  $\mathbb{A}$  are integer matrices and  $\ell \in \mathbb{Z}^\nu$  we will have that  $\min\{|\partial_{\zeta_i} \phi(\omega, \ell)| : i = 1, \dots, \nu\} \geq 1$ . Now we compute the matrix  $\mathbb{A} - \mathbb{V}$  evaluated at the point  $(\lambda, \dots, \lambda)$  for some  $\lambda \in \mathbb{N}$ . At such point  $\mathbb{V}$  has all entries equal to  $\lambda^3$ . We write

$$4\pi\lambda^{-3}(\mathbb{A} - \mathbb{V}) = \mathbb{A} - \pi\mathbb{V}, \quad \mathbb{A} := 4\pi\lambda^{-3}\mathbb{A}, \quad \mathbb{V} := 4\lambda^{-3}\mathbb{V}.$$

Notice that  $\mathbb{V}$ ,  $\mathbb{A}$  are integer matrices (recall (5.12), (5.13) and the definition of  $D$ ) and  $\mathbb{A}$  is invertible (see proof of Lemma 5.2). Then  $\mathbb{A} - \pi\mathbb{V} = \mathbb{A}(\mathbb{I}_N - \pi\mathbb{A}^{-1}\mathbb{V})$ . The eigenvalues of  $\mathbb{A}^{-1}\mathbb{V}$  are algebraic numbers, hence 1 cannot be an eigenvalue of  $\pi\mathbb{A}^{-1}\mathbb{V}$ . This proves that also  $\mathbb{A} - \pi\mathbb{V}$  is invertible. Therefore the determinant of  $\mathbb{A} - \pi\mathbb{V}$  is not an identically zero polynomial. Hence

$$|\Omega_\varepsilon \setminus \mathcal{G}_0^{(1)}| \lesssim \sum_{\ell \in \mathbb{Z}^\nu} |T_\ell| \lesssim \gamma \sum_{\ell \in \mathbb{Z}^\nu} \langle \ell \rangle^{-\tau} \lesssim \gamma,$$

since  $\tau > \nu/2$  (see (5.21)). Now consider the set in (5.23) and write

$$\Omega_\varepsilon \setminus \mathcal{G}_0^{(2)}(\mathbf{C}_1, \mathbf{C}_2) = \bigcup_{\substack{\ell \in \mathbb{Z}^\nu, \sigma, \sigma' = \pm, \\ |\ell| < \mathbf{C}_1}} \bigcup_{\substack{j, k \in S^c, \\ \max\{|j|, |k|\} \leq \mathbf{C}_2}} R_{\ell, j, k}^{\sigma, \sigma'}, \quad (5.29)$$

$$R_{\ell, j, k}^{\sigma, \sigma'} := \{\omega \in \Omega_\varepsilon : |\psi(\omega, \ell, j, k)| < \gamma \langle \ell \rangle^{-\tau}, \mathbf{v} \cdot \ell + j - k = 0\},$$

$$\psi(\omega, \ell, j, k) := (\bar{\omega} - \varepsilon^2 m_1 \mathbf{v} + \varepsilon^2 \mathbb{A} \zeta) \cdot \ell + \sigma \sqrt{|j|} - \sigma' \sqrt{|k|}.$$

Reasoning as done for the function  $\phi$  in (5.26) we deduce that  $\min\{|\partial_{\zeta_i} \psi(\omega, \ell, j, k)| : i = 1, \dots, \nu\} \geq 1$ . Hence the measure of a single bad set is bounded as  $|R_{\ell, j, k}^{\sigma, \sigma'}| \lesssim \gamma \langle \ell \rangle^{-\tau}$ . By (5.29) we deduce

$$|\Omega_\varepsilon \setminus \mathcal{G}_0^{(2)}(\mathbf{C}_1, \mathbf{C}_2)| \lesssim \mathbf{C}_2 \sum_{\ell \in \mathbb{Z}^\nu} \gamma \langle \ell \rangle^{-\tau} \lesssim \mathbf{C}_2 \gamma.$$

This implies the thesis.  $\square$

We look for an embedded invariant torus

$$i: \mathbb{T}^\nu \rightarrow \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp, \quad \varphi \mapsto i(\varphi) := (\theta(\varphi), y(\varphi), W(\varphi)) \quad (5.30)$$

$$\mathfrak{J}(\varphi) := i(\varphi) - (\varphi, 0, 0) := (\Theta(\varphi), y(\varphi), W(\varphi))$$

of the Hamiltonian vector field  $X_{H_\varepsilon}$  (see (5.19), (5.16)) supporting quasi-periodic solutions with diophantine frequency  $\omega \in \mathcal{G}_0$  (recall (5.25)), satisfying the following condition:

- **(Travelling wave):** the torus embedding  $i(\varphi) = (\varphi, 0, 0) + \mathfrak{J}(\varphi)$  satisfies

$$\mathbf{v} \cdot \partial_\varphi \Theta(\varphi) = 0, \quad \mathbf{v} \cdot \partial_\varphi y(\varphi) = 0, \quad (\mathbf{v} \cdot \partial_\varphi + \partial_x) W(\varphi, x) = 0, \quad (5.31)$$

where  $\mathbf{v}$  is the vector velocity given in (5.24).

We remark that the embedding  $i(\varphi)$  in the original coordinates (see (3.4), (4.6), (5.7)), reads

$$(\eta(t, x), \psi(t, x)) = \left( \Lambda^{-1} \Phi_B \mathbf{A}_\varepsilon \right) (\theta(\omega t), y(\omega t), \Lambda W(\omega t)). \quad (5.32)$$

**Remark 5.4.** The condition (5.31) on the embedding  $i(\varphi)$  is equivalent to require that  $(\eta, \psi)$  in (5.32) belongs to the subspace  $S_\nu$  defined in (2.46).

Consider the momentum Hamiltonian in (1.7) expressed in the action-angle variables (5.1) (see also (3.16))

$$M = M(\theta, y, W) = -\mathbf{v} \cdot y + \int_{\mathbb{T}} \tilde{\eta}_x \tilde{\psi} dx, \quad W = \begin{pmatrix} \tilde{\eta} \\ \tilde{\psi} \end{pmatrix}. \quad (5.33)$$

Then the condition (5.31) can be written as

$$\mathbf{v} \cdot \partial_\varphi i(\varphi) + X_M(i(\varphi)) = 0, \quad X_M(i(\varphi)) = (-\mathbf{v}, 0, W_x), \quad (5.34)$$

or equivalently

$$\mathbf{v} \cdot \partial_\varphi \mathfrak{I}(\varphi) + (0, 0, W_x) = 0. \quad (5.35)$$

For technical reason, it is useful to consider the modified Hamiltonian

$$H_{\varepsilon, \Xi}(\theta, y, W) := H_\varepsilon(\theta, y, W) + \Xi \cdot \theta, \quad \Xi \in \mathbb{R}^\nu. \quad (5.36)$$

More precisely, we introduce  $\Xi$  in order to control the average in the  $y$ -component in our Nash Moser scheme. The vector  $\Xi$  has no dynamical consequences since an invariant torus for the Hamiltonian vector field  $X_{H_{\varepsilon, \Xi}}$  is actually invariant for  $X_{H_\varepsilon}$  itself (see Lemma 6.1).

Thus, we look for zeros of the nonlinear operator  $\mathcal{F}(i, \Xi) \equiv \mathcal{F}(i, \Xi, \omega, \varepsilon) := \omega \cdot \partial_\varphi i(\varphi) - X_{\mathcal{N}}(i(\varphi)) - X_P(i(\varphi)) + (0, \Xi, 0)$  defined as

$$\mathcal{F}(i, \Xi) = \begin{pmatrix} \omega \cdot \partial_\varphi \Theta(\varphi) - \partial_y P(i(\varphi)) \\ \omega \cdot \partial_\varphi y(\varphi) + \frac{1}{2} \partial_\theta (N(\theta(\varphi)) W(\varphi), W(\varphi))_{L^2(\mathbb{T})} + \partial_\theta P(i(\varphi)) + \Xi \\ \omega \cdot \partial_\varphi W(\varphi) - JN(\theta(\varphi)) W(\varphi) - J \nabla_W P(i(\varphi)) \end{pmatrix} \quad (5.37)$$

where  $\Theta(\varphi) := \theta(\varphi) - \varphi$  is  $(2\pi)^\nu$ -periodic and  $J$  is defined in (3.1).

**Lemma 5.5.** *If  $i(\varphi)$  is traveling then  $\mathcal{F}(i(\varphi), \Xi)$  is traveling.*

*Proof.* We observe that  $\mathcal{F}(i, \Xi) = \omega \cdot \partial_\varphi i - X_{H_\varepsilon}(i) + X_{\Xi \cdot \theta}(i)$ . We show that  $\mathcal{F}(i, \Xi)$  solves (5.35). We have  $\mathbf{v} \cdot \partial_\varphi X_{H_\varepsilon}(i) = -dX_{H_\varepsilon}(i)[(-\mathbf{v}, 0, W_x)] = -dX_{H_\varepsilon}(i)[(0, 0, W_x)] + \mathbf{v} \cdot \partial_\theta X_{H_\varepsilon}(i) = -d_W X_{H_\varepsilon}(i)[W_x] + \mathbf{v} \cdot \partial_\theta X_{H_\varepsilon}(i)$

$$(0, 0, \partial_x)(X_{H_\varepsilon}(i)) = (0, 0, d_W X_{H_\varepsilon}^{(W)}(i)[W_x])$$

then

$$(\mathbf{v} \cdot \partial_\varphi + (0, 0, \partial_x)) X_{H_\varepsilon}(i) = dX_M(i)[X_{H_\varepsilon}(i)] - dX_{H_\varepsilon}(i)[X_M(i)] = [X_M, X_{H_\varepsilon}](i) = X_{\{M, H_\varepsilon\}}(i) = 0,$$

where the Poisson brackets are with respect to the symplectic form (5.9). Eventually it is easy to see that

$$(\mathbf{v} \cdot \partial_\varphi + (0, 0, \partial_x))(X_{\Xi \cdot \theta}(i)) = X_{\{M, \Xi \cdot \theta\}} = X_{-\mathbf{v} \cdot \Xi} = 0.$$

Since  $\omega \cdot \partial_\varphi$  commutes with  $\mathbf{v} \cdot \partial_\varphi$  and  $X_M$ , then  $\omega \cdot \partial_\varphi$  satisfies the (5.35). □

We define the Sobolev norm of the periodic component of the embedded torus

$$\|\mathfrak{I}\|_s := \|\Theta\|_s + \|y\|_s + \|W\|_s, \quad (5.38)$$

where  $W \in H_{S^\perp}^s := H^s \cap H_S^\perp$  (recall (3.15)) with norm defined in (2.2) and with abuse of notation, we are denoting by  $\|\cdot\|_s$  the Sobolev norms of functions in  $H^s(\mathbb{T}^\nu; \mathbb{R}^\nu)$ . From now on we fix

$$s_0 := \lceil \nu/2 \rceil + 4.$$

We say that  $\mathfrak{I}$  is traveling if  $\mathfrak{I} + (\varphi, 0, 0)$  satisfies the (5.31).

Notice that in the coordinates (5.5), a quasi-periodic solution corresponds to an embedded invariant torus (5.30). Therefore we can reformulate the main Theorem 2 as follows.

**Theorem 5.6.** *Let  $\nu \geq 1$ . For any generic choice of  $S$  (see (1.23)) there exists  $\varepsilon_0 > 0$  small enough, such that the following holds. For all  $\varepsilon \in (0, \varepsilon_0)$  there exist positive constants  $C = C(\nu)$ ,  $\mu_1 = \mu_1(\nu)$  and a Cantor-like set  $\mathcal{C}_\varepsilon \subseteq \Omega_\varepsilon$  (see (5.20)), with asymptotically full measure as  $\varepsilon \rightarrow 0$ , namely*

$$\lim_{\varepsilon \rightarrow 0} \frac{|\mathcal{C}_\varepsilon|}{|\Omega_\varepsilon|} = 1, \quad (5.39)$$

*such that, for all  $\omega \in \mathcal{C}_\varepsilon$ , there exists a solution  $i_\infty(\varphi) := i_\infty(\omega, \varepsilon)(\varphi)$  of the equation  $\mathcal{F}(i_\infty, 0, \omega, \varepsilon) = 0$ . Hence the embedded torus  $\varphi \mapsto i_\infty(\varphi)$  is invariant for the Hamiltonian vector field  $X_{H_\varepsilon}$ , and it is filled by quasi-periodic solutions with frequency  $\omega$ . The torus  $i_\infty$  satisfies*

$$\|i_\infty(\varphi) - (\varphi, 0, 0)\|_{s_0 + \mu_1}^{\gamma, \mathcal{C}_\varepsilon} \leq C \varepsilon^{16-2b} \gamma^{-1}.$$

*Moreover the torus  $i_\infty$  satisfies condition (5.31) and it is linearly stable.*

We can deduce Theorem 2 from Theorem 5.6, indeed the quasi-periodic solution  $(\eta, \psi)$  in (1.28) is

$$(\eta(t, x), \psi(t, x)) = \left( \Lambda^{-1} \circ \Phi_B \circ \mathbf{A}_\varepsilon \right) (\theta_\infty(\omega t), y_\infty(\omega t), \Lambda W_\infty(\omega t))$$

for  $\omega = \omega(\zeta) \in \mathcal{C}_\varepsilon$ , where  $\omega(\zeta)$  is the frequency amplitude map (5.4). The rest of the paper is devoted to the proof of Theorem 5.6.

**5.2.1. Tame estimates of the nonlinear vector field.** We give tame estimates for the composition operator induced by the Hamiltonian vector fields  $X_{\mathcal{N}}$  and  $X_P$  in (5.37). Since the functions  $y \rightarrow \sqrt{\zeta + \varepsilon^{2(b-1)}y}$  and  $\theta \rightarrow e^{i\theta}$  are analytic for  $\varepsilon$  small enough and  $|y| \leq C$ , classical composition results (see for instance Lemma 2.2 in [8]) imply that, for all  $\|\mathcal{J}\|_{s_0}^{\gamma, \mathcal{O}} \leq 1$ ,

$$\|A_\varepsilon(\theta(\varphi), y(\varphi), W(\varphi))\|_s^{\gamma, \mathcal{O}} \lesssim_s \varepsilon(1 + \|\mathcal{J}\|_s^{\gamma, \mathcal{O}}).$$

In the following lemma we collect tame estimates for the Hamiltonian vector fields  $X_{\mathcal{N}}, X_P, X_{H_\varepsilon}$  (see (5.19)).

**Lemma 5.7.** *Let  $\mathcal{O} \in \Omega_\varepsilon$  and  $\mathcal{J}(\varphi)$  in (5.38) satisfying  $\|\mathcal{J}\|_{s_0+1}^{\gamma, \mathcal{O}} \lesssim \varepsilon^{16-2b}\gamma^{-1}$ . Then we have (recall (5.4), (5.17))*

$$\begin{aligned} \|\partial_y P(i)\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{14} + \varepsilon^{2b} \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}}, & \|\partial_\theta P(i)\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{16-2b}(1 + \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}}), \\ \|\nabla_W P(i)\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{15-b} + \varepsilon^{16-b}\gamma^{-1} \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}}, & \|X_P(i)\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{16-2b} + \varepsilon^{2b} \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}}, \\ \|\partial_\theta \partial_y P(i)\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{14} + \varepsilon^{15}\gamma^{-1} \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}}, & \|\partial_y \nabla_W P(i)\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{13+b} + \varepsilon^{2b-1} \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}}, \\ \|\partial_{yy} P(i) - \frac{\varepsilon^{2b}}{2} (\mathbb{A} + \mathbb{B}(\zeta))\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{12+2b} + \varepsilon^{13+2b}\gamma^{-1} \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}}, \end{aligned}$$

and for all  $\widehat{v} := (\widehat{\Theta}, \widehat{y}, \widehat{W})$ ,

$$\begin{aligned} \|\partial_y d_i X_P(i)[\widehat{v}]\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon^{2b-1} (\|\widehat{v}\|_{s+1}^{\gamma, \mathcal{O}} + \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}} \|\widehat{v}\|_{s_0+1}^{\gamma, \mathcal{O}}), \\ \|d_i X_{H_\varepsilon}(i)[\widehat{v}] + (0, 0, J\widehat{W})\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon (\|\widehat{v}\|_{s+1}^{\gamma, \mathcal{O}} + \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}} \|\widehat{v}\|_{s_0+1}^{\gamma, \mathcal{O}}), \\ \|d_i^2 X_{H_\varepsilon}(i)[\widehat{v}, \widehat{v}]\|_s^{\gamma, \mathcal{O}} &\lesssim_s \varepsilon (\|\widehat{v}\|_{s+1}^{\gamma, \mathcal{O}} \|\widehat{v}\|_{s_0+1}^{\gamma, \mathcal{O}} + \|\mathcal{J}\|_{s+1}^{\gamma, \mathcal{O}} (\|\widehat{v}\|_{s_0+1}^{\gamma, \mathcal{O}})^2). \end{aligned}$$

In the sequel we will use that, by the diophantine condition (5.25), the operator  $(\omega \cdot \partial_\varphi)^{-1}$  is defined for all functions  $u$  with zero  $\varphi$ -average, and satisfies

$$\|(\omega \cdot \partial_\varphi)^{-1} u\|_s \lesssim_s \gamma^{-1} \|u\|_{s+\tau}, \quad \|(\omega \cdot \partial_\varphi)^{-1} u\|_s^{\gamma, \mathcal{O}} \lesssim_s \gamma^{-1} \|u\|_{s+2\tau+1}^{\gamma, \mathcal{O}}.$$

## 6. APPROXIMATE INVERSE

We want to solve the nonlinear functional equation

$$\mathcal{F}(i, \Xi) = 0 \tag{6.1}$$

by applying a Nash-Moser scheme. It is well known that the main issue in implementing this algorithm concerns the approximate inversion of the linearized operator of  $\mathcal{F}$  at any approximate solution  $(i_n, \Xi_n)$ , namely  $d\mathcal{F}(i_n, \Xi_n)$ . Note that  $d\mathcal{F}(i_n, \Xi_n)$  is independent of  $\Xi_n$  (see (5.37)). One of the main problems is that the  $(\theta, y, W)$ -components of  $d\mathcal{F}(i_n, \Xi_n)$  are coupled and then the linear system

$$d\mathcal{F}(i_n, \Xi_n)[\widehat{v}, \widehat{\Xi}] = \omega \cdot \partial_\varphi \widehat{v} - d_i X_{H_\varepsilon}(i_n)[\widehat{v}] - (0, \widehat{\Xi}, 0) = g = (g^{(\theta)}, g^{(y)}, g^{(W)}) \tag{6.2}$$

is quite involved. In order to approximately solve (6.2) we follow the scheme developed by Berti-Bolle in [11] which describes a way to approximately triangularize (6.2). We recall here the main steps of the strategy.

We study the solvability of equation (6.2) at an approximate solution, which we denote by  $(i_0, \Xi_0)$ ,  $i_0(\varphi) = (\theta_0(\varphi), y_0(\varphi), W_0(\varphi))$ . We assume the following hypothesis, which we shall verify at any step of the Nash-Moser iteration:

- **Assumption.** The map  $\omega \mapsto i_0(\omega)$  is a Lipschitz function of the form (5.30), satisfying (5.31), defined on some subset  $\mathcal{O}_0 \in \mathcal{G}_0 \subseteq \Omega_\varepsilon$  (recall (5.25), (5.20)) and, for some  $\mathfrak{p}_0 := \mathfrak{p}_0(\nu) > 0$ ,

$$\|\mathcal{J}_0\|_{s_0+\mathfrak{p}_0}^{\gamma, \mathcal{O}_0} \leq \varepsilon^{16-2b}\gamma^{-1}, \quad \|Z\|_{s_0+\mathfrak{p}_0}^{\gamma, \mathcal{O}_0} \leq \varepsilon^{16-2b}, \tag{6.3}$$

where  $\mathcal{J}_0(\varphi) := i_0(\varphi) - (\varphi, 0, 0)$  and  $Z$  is the error function

$$Z(\varphi) := (Z_1, Z_2, Z_3)(\varphi) := \mathcal{F}(i_0, \Xi_0)(\varphi) = \omega \cdot \partial_\varphi i_0(\varphi) - X_{H_\varepsilon, \Xi_0}(i_0(\varphi)).$$

This assumption will be verified inductively in Theorem 13.2, see (13.4).

By estimating the Sobolev norm of the function  $Z$  we can measure how the embedding  $i_0$  is close to being invariant for  $X_{H_{\varepsilon, \Xi_0}}$ . It is well known that an invariant torus  $i_0$  with diophantine flow is isotropic (see e.g. [11]), namely the pull-back 1-form  $i_0^* \mathcal{V}$  is closed, where  $\mathcal{V}$  is the Liouville 1-form in (5.10). This is tantamount to say that the 2-form  $\widetilde{\mathcal{W}}$  in (5.9) vanishes on the torus  $i_0(\mathbb{T}^\nu)$ , because  $i_0^* \widetilde{\mathcal{W}} = i_0^* d\mathcal{V} = d i_0^* \mathcal{V}$ . For an ‘‘approximately invariant’’ embedded torus  $i_0$  the 1-form  $i_0^* \mathcal{V}$  is only ‘‘approximately closed’’. In order to make this statement quantitative we consider

$$i_0^* \mathcal{V} = \sum_{k=1}^{\nu} a_k(\varphi) d\varphi_k, \quad a_k(\varphi) := ([\partial_\varphi \theta_0(\varphi)]^T y_0(\varphi))_k + \frac{1}{2} (\partial_{\varphi_k} W_0(\varphi), JW_0(\varphi))_{L^2}$$

and we quantify how small is

$$i_0^* \widetilde{\mathcal{W}} = d i_0^* \mathcal{V} = \sum_{1 \leq k < j \leq \nu} A_{kj}(\varphi) d\varphi_k \wedge d\varphi_j, \quad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi).$$

The next lemma proves that if  $i_0$  is a solution of the equation (6.1), then the parameter  $\Xi$  has to be naught, hence the embedded torus  $i_0$  supports a quasi-periodic solution of the ‘‘original’’ system with Hamiltonian  $H_\varepsilon$ .

**Lemma 6.1.** (Lemma 6.1 in [8]) *We have  $|\Xi_0|^{\gamma, \mathcal{O}_0} \lesssim \|Z\|_{s_0}^{\gamma, \mathcal{O}_0}$ . In particular, if  $\mathcal{F}(i_0, \Xi_0) = 0$  then  $\Xi_0 = 0$  and the torus  $i_0(\varphi)$  is invariant for the vector field  $X_{H_\varepsilon}$ .*

By [11] we know that it is possible to construct an embedded torus  $i_\delta(\varphi) = (\theta_0(\varphi), y_\delta(\varphi), W_0(\varphi))$ , which differs from  $i_0$  only for a small modification of the  $y$ -component, such that the 2-form  $\mathcal{W}$  (recall (5.9)) vanishes on the torus  $i_\delta(\mathbb{T}^\nu)$ , namely  $i_\delta$  is isotropic. In particular  $i_\delta(\varphi)$  is approximately invariant up to order  $O(Z)$  (see Lemma 7 in [11]). More precisely we have the following. In the paper we denote equivalently the differential  $\partial_i$  or  $d_i$ . We use the notation  $\Delta_\varphi := \sum_{k=1}^{\nu} \partial_{\varphi_k}^2$ .

**Lemma 6.2. (Isotropic torus).** (Lemma 6.3 in [8]) *The torus  $i_\delta = (\theta_0(\varphi), y_\delta(\varphi), z_0(\varphi))$  defined by*

$$y_\delta := y_0 + [\partial_\varphi \theta_0(\varphi)]^{-T} \rho(\varphi), \quad \rho_j(\varphi) := \Delta_\varphi^{-1} \sum_{k=1}^{\nu} \partial_{\varphi_j} A_{kj}(\varphi),$$

*is isotropic. If (6.3) holds, then, for some  $\mathfrak{d} := \mathfrak{d}(\nu, \tau)$ ,*

$$\begin{aligned} \|y_\delta - y_0\|_s^{\gamma, \mathcal{O}_0} &\leq_s \|\mathfrak{J}_0\|_{s+\mathfrak{d}}^{\gamma, \mathcal{O}_0}, \\ \|y_\delta - y_0\|_s^{\gamma, \mathcal{O}_0} &\leq_s \gamma^{-1} (\|Z\|_{s+\mathfrak{d}}^{\gamma, \mathcal{O}_0} \|\mathfrak{J}_0\|_{s_0+\mathfrak{d}}^{\gamma, \mathcal{O}_0} + \|Z\|_{s_0+\mathfrak{d}}^{\gamma, \mathcal{O}_0} \|\mathfrak{J}_0\|_{s+\mathfrak{d}}^{\gamma, \mathcal{O}_0}), \\ \|\mathcal{F}(i_\delta, \zeta_0)\|_s^{\gamma, \mathcal{O}_0} &\leq_s \|Z\|_{s+\mathfrak{d}}^{\gamma, \mathcal{O}_0} + \|Z\|_{s_0+\mathfrak{d}}^{\gamma, \mathcal{O}_0} \|\mathfrak{J}_0\|_{s+\mathfrak{d}}^{\gamma, \mathcal{O}_0}, \quad \|\partial_i i_\delta \widehat{|\mathfrak{I}}\|_s \leq_s \|\widehat{|\mathfrak{I}}\|_s + \|\mathfrak{J}_0\|_{s+\mathfrak{d}} \|\widehat{|\mathfrak{I}}\|_s. \end{aligned}$$

The strategy now is to construct an approximate inverse for  $d\mathcal{F}(i_0, \Xi_0)$  by starting from an approximate inverse for the linear operator  $d\mathcal{F}(i_\delta, \Xi_0)$ . The advantage of analyzing the linearized problem at  $i_\delta$  is that, thanks to the isotropicity of  $i_\delta$ , it is possible to construct a *symplectic* change of variables which approximately triangularizes the linear system. We define the symplectic change of coordinates

$$\begin{pmatrix} \theta \\ y \\ W \end{pmatrix} := G_\delta \begin{pmatrix} \vartheta \\ Y \\ U \end{pmatrix} := \begin{pmatrix} \theta_0(\vartheta) \\ y_\delta(\vartheta) + [\partial_\vartheta \theta_0(\vartheta)]^{-T} Y + [(\partial_\vartheta \widetilde{W}_0)(\theta_0(\vartheta))]^T J^{-1} U \\ W_0(\vartheta) + U \end{pmatrix}, \quad (6.4)$$

where  $\widetilde{W}_0 := W_0(\theta_0^{-1}(\theta))$ . We show that the map  $G_\delta$  ‘‘preserves’’ the subspace of traveling embeddings, namely satisfying (5.31).

**Lemma 6.3.** *Assume that  $i_\delta(\varphi)$  satisfies the condition (5.31). Then  $\mathcal{I}(\varphi) := (\vartheta(\varphi), Y(\varphi), U(\varphi))$  satisfies (5.31) if and only if  $i(\varphi) = (\theta(\varphi), y(\varphi), W(\varphi)) := G_\delta(\mathcal{I}(\varphi))$  satisfies (5.31).*

*Proof.* We use the equivalent condition (5.34). We note that

$$\mathbf{v} \cdot \partial_\varphi i(\varphi) + X_M(i(\varphi)) = 0 \quad \Leftrightarrow \quad \mathbf{v} \cdot \partial_\varphi \mathcal{I}(\varphi) + (dG_\delta)^{-1} X_M G_\delta(\mathcal{I}(\varphi)) = 0.$$

Then, since  $G_\delta$  is symplectic one has  $(dG_\delta)^{-1} X_M G_\delta = X_{M \circ G_\delta}$ . Therefore it is sufficient to prove that  $M \circ G_\delta = M$ . For the variables  $W_0, W, U$  in  $H_S^\perp$  (see (3.15)) we shall use the notation  $W_0 = (\eta_0, \psi_0)^T$ ,  $W = (\tilde{\eta}, \tilde{\psi})^T$ ,

$U = (\eta_1, \psi_1)^T$ . Hence, using (6.4) and (5.33), we have

$$M(G_\delta(\vartheta, Y, U)) = -\mathbf{v} \cdot y_\delta(\vartheta) - [\partial_\vartheta \theta_0(\vartheta)]^{-1} \mathbf{v} \cdot Y - \mathbf{v} \cdot B J^{-1} U + \int_{\mathbb{T}} (\eta_0 + \eta_1)_x (\psi_0 + \psi_1) dx,$$

where  $B := [(\partial_\vartheta \tilde{W}_0)(\theta_0(\vartheta))]^T$ . Using  $\tilde{W}_0 := W_0(\theta_0^{-1}(\vartheta))$  and that  $i_\delta(\varphi)$  satisfies (5.31) we deduce

$$\begin{aligned} -\mathbf{v} \cdot B J^{-1} U &= \int_{\mathbb{T}} (d\eta_0)[\mathbf{v}] \psi_1 dx - \int_{\mathbb{T}} (d\psi_0)[\mathbf{v}] \eta_1 dx, \\ (d\eta_0)[\mathbf{v}] &= -(\eta_0)_x, \quad (d\psi_0)[\mathbf{v}] = -(\psi_0)_x, \quad [d\theta_0(\vartheta)]^{-1} \mathbf{v} = \mathbf{v}. \end{aligned}$$

Therefore, by integrating by parts, we get

$$\begin{aligned} M(G_\delta(\vartheta, Y, U)) &= -\mathbf{v} \cdot Y - \mathbf{v} \cdot y_\delta(\vartheta) + \int_{\mathbb{T}} [(\psi_0)_x \eta_1 - (\eta_0)_x \psi_1] dx \\ &\quad + \int_{\mathbb{T}} [(\eta_0)_x \psi_0 + (\eta_0)_x \psi_1 + (\eta_1)_x \psi_0 + (\eta_1)_x \psi_1] dx \\ &= -\mathbf{v} \cdot Y + \int_{\mathbb{T}} (\eta_1)_x \psi_1 dx - \mathbf{v} \cdot y_\delta(\vartheta) + \int_{\mathbb{T}} (\eta_0)_x \psi_0 dx. \end{aligned}$$

Since  $i_\delta$  satisfies (5.31) the term  $-\mathbf{v} \cdot y_\delta(\vartheta) + \int_{\mathbb{T}} (\eta_0(\vartheta))_x \psi_0(\vartheta) dx$  is independent of  $\vartheta$ , hence it does not contribute to the vector field  $X_{M \circ G_\delta}$ . Then we have the thesis.  $\square$

The transformed Hamiltonian  $K := K(\vartheta, Y, U, \Xi_0)$  is (recall (5.36))

$$\begin{aligned} K &:= H_{\varepsilon, \Xi_0} \circ G_\delta = \theta_0(\vartheta) \cdot \Xi_0 + K_{00}(\vartheta) + K_{10}(\vartheta) \cdot Y + (K_{01}(\vartheta), U)_{L^2} + \frac{1}{2} K_{20}(\vartheta) Y \cdot Y + \\ &\quad + (K_{11}(\vartheta) Y, U)_{L^2} + \frac{1}{2} (K_{02}(\vartheta) U, U)_{L^2} + K_{\geq 3}(\vartheta, Y, U) \end{aligned} \quad (6.5)$$

where  $K_{\geq 3}$  collects the terms at least cubic in the variables  $(Y, U)$ . At any fixed  $\vartheta$ , the Taylor coefficient  $K_{00}(\vartheta) \in \mathbb{R}$ ,  $K_{10}(\vartheta) \in \mathbb{R}^\nu$ ,  $K_{01}(\vartheta) \in H_S^\perp$ ,  $K_{20}(\vartheta)$  is a  $\nu \times \nu$  real matrix,  $K_{02}(\vartheta)$  is a linear self-adjoint operator of  $H_S^\perp$  and finally  $K_{11}(\vartheta): \mathbb{R}^\nu \rightarrow H_S^\perp$ . The above Taylor coefficients do not depend on the parameter  $\Xi_0$ . The Hamilton equations associated to (6.5) are

$$\begin{cases} \dot{\vartheta} = K_{10}(\vartheta) + K_{20}(\vartheta) Y + K_{11}^T(\vartheta) U + \partial_\eta K_{\geq 3}(\vartheta, Y, U) \\ \dot{Y} = -[\partial_\vartheta \theta_0(\vartheta)]^T \Xi_0 - \partial_\vartheta K_{00}(\vartheta) - [\partial_\vartheta K_{10}(\vartheta)]^T Y - [\partial_\vartheta K_{01}(\vartheta)]^T U - \\ \quad - \partial_\vartheta \left( \frac{1}{2} K_{20}(\vartheta) Y \cdot Y + (K_{11}(\vartheta) Y, U)_{L^2} + \frac{1}{2} (K_{02}(\vartheta) U, U)_{L^2} + K_{\geq 3}(\vartheta, Y, U) \right) \\ \dot{U} = J(K_{01}(\vartheta) + K_{11}(\vartheta) Y + K_{02}(\vartheta) U) + \nabla_U K_{\geq 3}(\vartheta, Y, U) \end{cases} \quad (6.6)$$

where  $[\partial_\vartheta K_{10}(\vartheta)]^T$  is the  $\nu \times \nu$  transposed matrix and  $[\partial_\vartheta K_{01}(\vartheta)]^T, K_{11}^T(\vartheta): H_S^\perp \rightarrow \mathbb{R}^\nu$  are defined by the duality relation

$$(\partial_\vartheta K_{01}(\vartheta)[\hat{\vartheta}], U)_{L^2} = \hat{\vartheta} \cdot [\partial_\vartheta K_{01}(\vartheta)]^T U, \quad \forall \hat{\vartheta} \in \mathbb{R}^\nu, U \in H_S^\perp,$$

and similarly for  $K_{11}$ . Explicitly, for all  $U \in H_S^\perp$ , and denoting  $\underline{e}_k$  the  $k$ -th versor of  $\mathbb{R}^\nu$ ,

$$K_{11}^T(\vartheta) U = \sum_{k=1}^{\nu} (K_{11}^T(\vartheta) U \cdot \underline{e}_k) \underline{e}_k = \sum_{k=1}^{\nu} (U, K_{11}(\vartheta) \underline{e}_k)_{L^2} \underline{e}_k \in \mathbb{R}^\nu.$$

In the next lemma we estimate the coefficients  $K_{00}, K_{10}, K_{01}$  in the Taylor expansion (6.5). Note that on an exact solution  $(i_0, \Xi_0)$  we have  $K_{00}(\vartheta) = \text{const}$ ,  $K_{10} = \omega$  and  $K_{01} = 0$ .

**Lemma 6.4.** (Lemma 6.4 in [8]) *Assume (6.3). Then there is  $\mu := \mu(\nu)$  such that*

$$\|\partial_\vartheta K_{00}\|_s^{\gamma, \mathcal{O}_0} + \|K_{10} - \omega\|_s^{\gamma, \mathcal{O}_0} + \|K_{01}\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \|Z\|_{s+\mu}^{\gamma, \mathcal{O}_0} + \|Z\|_{s_0+\mu}^{\gamma, \mathcal{O}_0} \|\mathfrak{J}_0\|_{s+\mu}^{\gamma, \mathcal{O}_0}.$$

In the next Lemma we estimate  $K_{20}, K_{11}$  in (6.5). This result can be obtained reasoning as in Lemma 6.6 in [8] taking into account the bounds in Lemma 5.7.



**Lemma 6.5.** *Assume (6.3). Then for some  $\mu := \mu(\nu)$  we have*

$$\begin{aligned} \|K_{20} - \frac{\varepsilon^{2b}}{2}\mathbb{A}\|_s^{\gamma, \mathcal{O}_0} &\lesssim_s \varepsilon^{2b+2} + \varepsilon^{2b}\|\mathfrak{J}_0\|_{s+\mu}^{\gamma, \mathcal{O}_0}, \\ \|K_{11}Y\|_s^{\gamma, \mathcal{O}_0} &\lesssim_s \varepsilon^{15}\gamma^{-1}\|Y\|_s^{\gamma, \mathcal{O}_0} + \varepsilon^{2b-1}\|\mathfrak{J}_0\|_{s+\mu}^{\gamma, \mathcal{O}_0}\|Y\|_{s_0}^{\gamma, \mathcal{O}_0}, \\ \|K_{11}^T U\|_s^{\gamma, \mathcal{O}_0} &\lesssim_s \varepsilon^{15}\gamma^{-1}\|U\|_{s+2}^{\gamma, \mathcal{O}_0} + \varepsilon^{2b-1}\|\mathfrak{J}_0\|_{s+\mu}^{\gamma, \mathcal{O}_0}\|U\|_{s_0+2}^{\gamma, \mathcal{O}_0}. \end{aligned}$$

*In particular*

$$\|K_{20} - \frac{\varepsilon^{2b}}{2}\mathbb{A}\|_{s_0}^{\gamma, \mathcal{O}_0} \lesssim \varepsilon^{2b+2}, \quad \|K_{11}Y\|_{s_0}^{\gamma, \mathcal{O}_0} \lesssim \varepsilon^{15}\gamma^{-1}\|Y\|_{s_0}^{\gamma, \mathcal{O}_0}, \quad \|K_{11}^T U\|_{s_0}^{\gamma, \mathcal{O}_0} \lesssim \varepsilon^{15}\gamma^{-1}\|U\|_{s_0+2}^{\gamma, \mathcal{O}_0}.$$

**Remark 6.6.** *Notice that, if  $(\vartheta, Y, U)$  is such that*

$$\mathbf{v} \cdot \partial_\varphi \vartheta = \mathbf{v}, \quad \mathbf{v} \cdot \partial_\varphi Y = 0, \quad (\mathbf{v} \cdot \partial_\varphi + \partial_x)U = 0,$$

*then, using the equations (6.6), we deduce that*

$$\begin{aligned} \mathbf{v} \cdot \partial_\varphi K_{10}(\vartheta) &= 0, & (\mathbf{v} \cdot \partial_\varphi + \partial_x)(JK_{11}(\vartheta)Y) &= 0, & \mathbf{v} \cdot \partial_\varphi K_{20}(\vartheta)Y &= 0, \\ \mathbf{v} \cdot \partial_\varphi K_{11}^T(\vartheta)U &= 0, & (\mathbf{v} \cdot \partial_\varphi + \partial_x)(K_{02}(\vartheta)U) &= 0. \end{aligned} \quad (6.7)$$

We apply the linear change of variables

$$dG_\delta(\varphi, 0, 0) \begin{pmatrix} \widehat{\vartheta} \\ \widehat{Y} \\ \widehat{U} \end{pmatrix} := \begin{pmatrix} \partial_\vartheta \theta_0(\varphi) & 0 & 0 \\ \partial_\vartheta y_\delta(\varphi) & [\partial_\vartheta \theta_0(\varphi)]^{-T} & [(\partial_\vartheta \tilde{W}_0)(\theta_0(\varphi))]^T J^{-1} \\ \partial_\vartheta W_0(\varphi) & 0 & \text{Id} \end{pmatrix} \begin{pmatrix} \widehat{\vartheta} \\ \widehat{Y} \\ \widehat{U} \end{pmatrix}. \quad (6.8)$$

In these new coordinates the linearized operator  $d_{i, \Xi} \mathcal{F}(i_\delta, \Xi_0)$  is ‘‘approximately’’ the operator obtained linearizing (6.6) at  $(\vartheta, Y, U, \Xi) = (\varphi, 0, 0, \zeta_0)$  with  $\omega \cdot \partial_\varphi$  instead of  $\partial_t$ , namely

$$\begin{pmatrix} \omega \cdot \partial_\varphi \widehat{\vartheta} - \partial_\vartheta K_{10}(\varphi)[\widehat{\vartheta}] - K_{20}(\varphi)\widehat{Y} - K_{11}^T(\varphi)\widehat{U} \\ \omega \cdot \partial_\varphi \widehat{Y} + [\partial_\vartheta \theta_0(\varphi)]^T \widehat{\Xi} + \partial_\vartheta [\partial_\vartheta \theta_0(\varphi)]^T [\widehat{\vartheta}, \widehat{\Xi}_0] + \partial_{\vartheta\vartheta} K_{00}(\varphi)[\widehat{\vartheta}] + [\partial_\vartheta K_{10}(\varphi)]^T \widehat{Y} + [\partial_\vartheta K_{01}(\varphi)]^T \widehat{U} \\ \omega \cdot \partial_\varphi \widehat{U} - J\{\partial_\vartheta K_{01}(\varphi)[\widehat{\vartheta}] + K_{11}(\varphi)\widehat{Y} + K_{02}(\varphi)\widehat{U}\} \end{pmatrix}. \quad (6.9)$$

We give estimates on the composition operator induced by the transformation (6.8).

**Lemma 6.7.** *(Lemma 6.7 in [8]) Assume (6.3) and let  $\widehat{\mathbf{v}} := (\widehat{\vartheta}, \widehat{Y}, \widehat{U})$ . Then, for some  $\mu := \mu(\nu)$ , we have*

$$\begin{aligned} \|dG_\delta(\varphi, 0, 0)[\widehat{\mathbf{v}}]\|_s + \|dG_\delta(\varphi, 0, 0)^{-1}[\widehat{\mathbf{v}}]\|_s &\leq_s \|\widehat{\mathbf{v}}\|_s + \|\mathfrak{J}_0\|_{s+\mu}\|\widehat{\mathbf{v}}\|_{s_0} \\ \|d^2 G_\delta(\varphi, 0, 0)[\widehat{\mathbf{v}}_1, \widehat{\mathbf{v}}_2]\|_s &\leq_s \|\widehat{\mathbf{v}}_1\|_s \|\widehat{\mathbf{v}}_2\|_{s_0} + \|\widehat{\mathbf{v}}_1\|_{s_0} \|\widehat{\mathbf{v}}_2\|_s + \|\mathfrak{J}_0\|_{s+\mu}\|\widehat{\mathbf{v}}\|_{s_0}\|\widehat{\mathbf{v}}_2\|_{s_0}. \end{aligned}$$

*Moreover the same estimates hold if we replace  $\|\cdot\|_s$  with  $\|\cdot\|_s^{\gamma, \mathcal{O}_0}$ .*

In order to construct an approximate inverse of (6.9) it is sufficient to solve the system of equations

$$\mathbb{D}[\widehat{\vartheta}, \widehat{Y}, \widehat{U}, \widehat{\Xi}] := \begin{pmatrix} \omega \cdot \partial_\varphi \widehat{\vartheta} - K_{20}(\varphi)\widehat{Y} - K_{11}^T(\varphi)\widehat{U} \\ \omega \cdot \partial_\varphi \widehat{Y} + [\partial_\vartheta \theta_0(\varphi)]^T \widehat{\Xi} \\ \omega \cdot \partial_\varphi \widehat{U} - JK_{11}(\varphi)\widehat{Y} - JK_{02}(\varphi)\widehat{U} \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \quad (6.10)$$

which is obtained by (6.9) neglecting the terms that are naught at a solution, namely, by Lemmata 6.1 and 6.4,  $\partial_\vartheta K_{10}$ ,  $\partial_{\vartheta\vartheta} K_{00}$ ,  $\partial_\vartheta K_{00}$ ,  $\partial_\vartheta K_{01}$  and  $\partial_\vartheta [\partial_\vartheta \theta_0(\varphi)]^T[\cdot, \zeta_0]$ . The term  $(g_1, g_2, g_3)^T$  is assumed to satisfies the (5.31) and we look for a solution  $(\widehat{\vartheta}, \widehat{Y}, \widehat{U})$  of (6.10) with the same property.

First, we solve the second equation, namely

$$\omega \cdot \partial_\varphi \widehat{Y} = g_2 - [\partial_\vartheta \theta_0(\varphi)]^T \widehat{\Xi}. \quad (6.11)$$

We choose  $\widehat{\Xi} := \langle g_2 \rangle_\varphi$  (where  $\langle \cdot \rangle_\varphi$  denotes the  $\varphi$ -average) so that the  $\varphi$ -average of the right hand side of (6.11) is zero. Note that the  $\varphi$ -averaged matrix  $\langle (\partial_\vartheta \theta_0)^T \rangle_\varphi = \langle \text{Id} + (\partial_\vartheta \Theta_0)^T \rangle_\varphi = \text{Id}$  because  $\theta_0(\varphi) = \varphi + \Theta_0(\varphi)$  and  $\Theta_0(\varphi)$  is periodic in each component. Therefore

$$\widehat{Y} = (\omega \cdot \partial_\varphi)^{-1}(g_2 - [\partial_\vartheta \theta_0(\varphi)]^T \langle g_2 \rangle_\varphi) + \langle \widehat{Y} \rangle_\varphi, \quad \langle \widehat{Y} \rangle_\varphi \in \mathbb{R}^\nu, \quad (6.12)$$

where the average  $\langle \widehat{Y} \rangle_\varphi$  will be fix when we deal with the first equation. We remark that, by assumption on  $i_0$  and  $(g_1, g_2, g_3)^T$  one has (recall  $\theta_0(\varphi) = \varphi + \Theta_0(\varphi)$ )

$$\mathbf{v} \cdot \partial_\varphi \Theta_0(\varphi) = \mathbf{v} \cdot \partial_\varphi g_2(\varphi) = 0 \quad \Rightarrow \quad \mathbf{v} \cdot \partial_\varphi \widehat{Y}(\varphi) = 0. \quad (6.13)$$

We now analyze the third equation, namely

$$\mathcal{L}_\omega \widehat{U} = g_3 + JK_{11}(\varphi) \widehat{Y}, \quad \mathcal{L}_\omega := \omega \cdot \partial_\varphi - JK_{02}(\varphi). \quad (6.14)$$

Since  $\widehat{Y}$  has been fixed in (6.12) (up to a constant), then solving the equation (6.14) is tantamount to invert the operator  $\mathcal{L}_\omega$ . For the moment we make the following assumption (that will be proved in section 12):

- **Inversion Assumption:** There exist  $\mathbf{p}_1 := \mathbf{p}_1(\nu) > 0$  and a set  $\Omega_\infty \subset \mathcal{O}_0 \subseteq \Omega_\varepsilon$  such that for all  $\omega \in \Omega_\infty$  and every  $G \in \left( H_{S_\pm}^{s+2\tau+1} \cap S_\mathbf{v} \right)^2$  (see (2.46)), there exists a solution  $h := \mathcal{L}_\omega^{-1} G \in \left( H_{S_\pm}^s \cap S_\mathbf{v} \right)^2$  of the linear equation  $\mathcal{L}_\omega h = G$  and satisfies

$$\|\mathcal{L}_\omega^{-1} G\|_{s, \Omega_\infty}^{\gamma, \Omega_\infty} \lesssim_s \gamma^{-1} (\|G\|_{s+2\tau+1}^{\gamma, \Omega_\infty} + \varepsilon \gamma^{-6} \|\mathfrak{J}_\delta\|_{s+\mathbf{p}_1}^{\gamma, \mathcal{O}_0} \|G\|_{s_0}^{\gamma, \Omega_\infty}). \quad (6.15)$$

By the above assumption, there exists a solution of (6.14)

$$\widehat{U} = \mathcal{L}_\omega^{-1} [g_3 + JK_{11}(\varphi) \widehat{Y}]. \quad (6.16)$$

Moreover, recalling the assumption on  $g_3$  and using Remark 6.6, we have

$$(\mathbf{v} \cdot \partial_\varphi + \partial_x) \widehat{U} = 0. \quad (6.17)$$

Now consider the first equation

$$\omega \cdot \partial_\varphi \widehat{\vartheta} = g_1 + K_{20} \widehat{Y} - K_{11}^T(\varphi) \widehat{U}. \quad (6.18)$$

Substituting (6.12), (6.16) in the equation (6.18), we get

$$\omega \cdot \partial_\varphi \widehat{\vartheta} = g_1 + M_1(\varphi) \langle \widehat{Y} \rangle_\varphi + M_2(\varphi) g_2 + M_3(\varphi) g_3 - M_2(\varphi) [\partial_\vartheta \theta_0]^T \langle g_2 \rangle_\varphi, \quad (6.19)$$

where

$$M_1(\varphi) := K_{20}(\varphi) - K_{11}^T(\varphi) \mathcal{L}_\omega^{-1} JK_{11}(\varphi), \quad M_2(\varphi) := M_1(\varphi) (\omega \cdot \partial_\varphi)^{-1}, \quad M_3(\varphi) := -K_{11}^T(\varphi) \mathcal{L}_\omega^{-1}. \quad (6.20)$$

In order to solve the equation (6.19) we have to choose  $\langle \widehat{Y} \rangle_\varphi$  such that the right hand side in (6.19) has zero  $\varphi$ -average. By Lemma 6.5 and (6.3), the  $\varphi$ -averaged matrix  $\langle M_1 \rangle_\varphi$  is invertible, for  $\varepsilon$  small, and  $\langle M_1 \rangle_\varphi^{-1} = O(\varepsilon^{-2b}) = O(\gamma^{-1})$ . Thus we define

$$\langle \widehat{Y} \rangle_\varphi = -(\langle M_1 \rangle_\varphi)^{-1} \{ \langle g_1 \rangle_\varphi + \langle M_2 g_2 \rangle_\varphi + \langle M_3 g_3 \rangle_\varphi - \langle M_2 (\partial_\vartheta \theta_0)^T \rangle_\varphi \langle g_2 \rangle_\varphi \}.$$

With this choice of  $\langle \widehat{Y} \rangle_\varphi$  the equation (6.19) has the solution

$$\widehat{\vartheta} := (\omega \cdot \partial_\varphi)^{-1} \left( g_1 + M_1(\varphi) \langle \widehat{\eta} \rangle_\varphi + M_2(\varphi) g_2 + M_3(\varphi) g_3 - M_2(\varphi) [\partial_\vartheta \theta_0]^T \langle g_2 \rangle_\varphi \right).$$

Again, using Remark 6.6 and (6.18), we get

$$\mathbf{v} \cdot \partial_\varphi \widehat{\vartheta} = 0. \quad (6.21)$$

In conclusion, we have constructed a solution  $(\widehat{\vartheta}, \widehat{Y}, \widehat{U}, \widehat{\Xi})$  of the linear system (6.10) which satisfies (5.31). Consider the operator

$$\mathbf{T}_0 := (d\tilde{G}_\delta)(\varphi, 0, 0) \circ \mathbb{D}^{-1} \circ (dG_\delta(\varphi, 0, 0))^{-1},$$

where  $\tilde{G}_\delta((\vartheta, Y, U), \Xi)$  is the identity on the  $\Xi$ -component. By (6.13), (6.17), (6.21) and using Lemma 6.3 one can check that  $\mathbf{T}_0 g$  satisfies (5.31) if  $g$  satisfies (5.31). We denote the norm  $\|(\vartheta, Y, U, \Xi)\|_{s, \mathcal{O}}^{\gamma, \mathcal{O}} := \max\{\|(\vartheta, Y, U)\|_s, \|\Xi\|_{\gamma, \mathcal{O}}\}$ . In [11] (see also [8],[40]) the following result is proved.

**Theorem 6.8.** *Assume (6.3) and the inversion assumption. Then there exists  $\mu_1 := \mu_1(\nu)$  such that, for all  $\omega \in \Omega_\infty$ , the following holds:*

1. *for all  $g := (g^{(\theta)}, g^{(y)}, g^{(W)})$  satisfying (5.31)  $\mathbf{T}_0 g$  satisfies (5.31) and*

$$\|\mathbf{T}_0 g\|_{s, \Omega_\infty}^{\gamma, \Omega_\infty} \lesssim_s \gamma^{-1} (\|g\|_{s+\mu_1}^{\gamma, \Omega_\infty} + \varepsilon \gamma^{-6} \|\mathfrak{J}_0\|_{s+\mu_1}^{\gamma, \mathcal{O}_0} \|g\|_{s_0+\mu_1}^{\gamma, \Omega_\infty}).$$

2.  $\mathbf{T}_0$  is an approximate inverse of  $d\mathcal{F}(i_0)$ , namely

$$\begin{aligned} \|(d\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathbf{I})g\|_s^{\gamma, \Omega_\infty} &\lesssim_s \varepsilon^{2b-1} \gamma^{-2} \left( \|\mathcal{F}(i_0, \Xi_0)\|_{s_0+\mu_1}^{\gamma, \mathcal{O}_0} \|g\|_{s+\mu_1}^{\gamma, \Omega_\infty} \right. \\ &\quad \left. + \{ \|\mathcal{F}(i_0, \Xi_0)\|_{s+\mu_1}^{\gamma, \mathcal{O}_0} + \varepsilon \gamma^{-6} \|\mathcal{F}(i_0, \Xi_0)\|_{s_0+\mu_1}^{\gamma, \mathcal{O}_0} \|\mathfrak{J}_0\|_{s+\mu_1}^{\gamma, \mathcal{O}_0} \} \|g\|_{s_0+\mu_1}^{\gamma, \Omega_\infty} \right). \end{aligned}$$

The proof of the above theorem follows word by word the one of Theorem 6.10 in [8]. It is based on tame bounds for the map  $G_\delta$  and the coefficients of the Taylor expansion of the Hamiltonian  $K$  at the trivial torus  $(\varphi, 0, 0)$ , see Lemmata 6.4, 6.5, 6.2, 6.7.

## 7. THE LINEARIZED OPERATOR IN THE NORMAL DIRECTIONS

Recalling the assumption (6.3), in the sequel we assume that  $\mathfrak{J}_\delta := \mathfrak{J}_\delta(\varphi; \omega) = i_\delta(\varphi; \omega) - (\varphi, 0, 0)$  satisfies, for some  $\mathfrak{p}_1 = \mathfrak{p}_1(\nu) > 0$ ,

$$\|\mathfrak{J}_\delta\|_{s_0+\mathfrak{p}_1}^{\gamma, \mathcal{O}_0} \lesssim \varepsilon^{16-2b} \gamma^{-1}, \quad (7.1)$$

and the momentum condition (5.31). We note that  $G_\delta$  in (6.4) is the identity plus a translation plus a finite rank linear operator and it is  $O(\varepsilon^{16-2b} \gamma^{-1})$ -close to the identity in low norm  $s_0 + \mathfrak{p}_1$ . Returning to the initial variables we define (see (5.5), (6.4))  $T_\delta := T_\delta(\varphi)$  as

$$T_\delta := \mathbf{A}_\varepsilon(\Lambda G_\delta(\varphi, 0, 0)) = \varepsilon \left[ \frac{v_\delta}{v_\delta} \right] + \varepsilon^b W_0, \quad v_\delta = \sum_{j \in S} \sqrt{\zeta_j + \varepsilon^{2b-2} y_{\delta j}(\varphi)} e^{i(jx - \theta_{0j}(\varphi))} \quad (7.2)$$

and we have, for some  $\mu := \mu(\nu) > 0$ ,

$$\|\Phi_B(T_\delta)\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \varepsilon (1 + \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad \|d_i \Phi_B(T_\delta)[\widehat{v}]\|_s \lesssim_s \varepsilon (\|\widehat{v}\|_{s+\mu} + \|\mathfrak{J}_\delta\|_{s+\mu} \|\widehat{v}\|_{s_0+\mu}).$$

By following section 7 in [8] (see Lemma 7.1),  $K_{02}$  in (6.14) has rather explicit estimates (see also Proposition 6.2 in [32]). In the next proposition we will give a more explicit formulation of  $K_{02}(\varphi)$ . Notice that, by the shape derivative formula (3.9), the linearized operator of (1.3) at  $(\eta, \psi)(\varphi, x)$  is given by

$$\mathcal{L} := \mathcal{L}(\varphi) := \omega \cdot \partial_\varphi + \begin{pmatrix} \partial_x V + G(\eta)B & -G(\eta) \\ (1 + BV_x) + BG(\eta)B & V\partial_x - BG(\eta) \end{pmatrix}. \quad (7.3)$$

where  $V, B$  are given in (3.10) and (3.11). Hence we have the following.

**Proposition 7.1.** *Assume (7.1). Then there exists  $\mu_0 = \mu_0(\nu) > 0$  such that the following holds. The Hamiltonian operator  $\mathcal{L}_\omega$  in (6.14) has the form*

$$\mathcal{L}_\omega = \Pi_S^\perp (\omega \cdot \partial_\varphi - J \partial_{(\eta, \psi)} \nabla_{(\eta, \psi)} H(\Lambda^{-1} \Phi_B(T_\delta)) + \mathcal{Q}_0) = \Pi_S^\perp (\mathcal{L} + \mathcal{Q}_0), \quad (7.4)$$

where  $T_\delta$  is defined in (7.2),  $\Phi_B$  is the Birkhoff map given in Proposition 4.3,  $H$  is the Hamiltonian in (1.6) and  $\mathcal{L}$  is the operator in (7.3). The operator  $\mathcal{Q}_0$  is finite rank and has the form

$$\mathcal{Q}_0(\varphi)w = \sum_{|j| \leq C} \int_0^1 (w, g_j(\tau, \varphi))_{L^2} \chi_j(\tau, \varphi) d\tau, \quad (7.5)$$

for some functions  $g_j(\tau, \cdot), \chi_j(\tau, \cdot) \in H^s \cap H_S^\perp$ . In particular we consider the expansion  $\mathcal{Q}_0 = \sum_{i=1}^{12} \varepsilon^i \mathcal{R}_i + \mathcal{R}_{>12}$ , where the remainders  $\mathcal{R}_i, i = 1, \dots, 12$  have the form (7.5), they do not depend on  $\mathfrak{J}_\delta$  and satisfy

$$\|g_j^{(i)}\|_s^{\gamma, \mathcal{O}_0} + \|\chi_j^{(i)}\|_s^{\gamma, \mathcal{O}_0} \lesssim_s 1, \quad i = 1, \dots, 12, \quad (7.6)$$

while  $\mathcal{R}_{>12}$  satisfies

$$\begin{aligned} \|g_j^{>12}\|_s^{\gamma, \mathcal{O}_0} \|\chi_j^{>12}\|_{s_0}^{\gamma, \mathcal{O}_0} + \|g_j^{>12}\|_{s_0}^{\gamma, \mathcal{O}_0} \|\chi_j^{>12}\|_s^{\gamma, \mathcal{O}_0} &\lesssim_s \varepsilon^{13} + \varepsilon^2 \|\mathfrak{J}_\delta\|_{s+\mu_0}^{\gamma, \mathcal{O}_0}, \\ \|d_i g_j^{>12}[\widehat{v}]\|_s \|\chi_j^{>12}\|_{s_0} + \|d_i g_j^{>12}[\widehat{v}]\|_{s_0} \|\chi_j^{>12}\|_s &+ \|g_j^{>12}\|_{s_0} \|d_i \chi_j^{>12}[\widehat{v}]\|_s + \|g_j^{>12}\|_s \|d_i \chi_j^{>12}[\widehat{v}]\|_{s_0} \\ &\lesssim_s \varepsilon^2 \|\widehat{v}\|_{s+\mu_0} + \varepsilon^b \|\mathfrak{J}_\delta\|_{s+\mu_0} \|\widehat{v}\|_{s_0+\mu_0}. \end{aligned} \quad (7.7)$$

Finally, recalling the Definition 2.4, we have

$$\begin{aligned} \mathfrak{M}_{\mathcal{Q}_0}^\gamma(0, s) &\lesssim_s \varepsilon^2 (1 + \|\mathfrak{J}_\delta\|_{s+\mu_0}^{\gamma, \mathcal{O}_0}), \\ \mathfrak{M}_{d_i \mathcal{Q}_0[\widehat{v}]}(0, s) &\lesssim_s \varepsilon^2 \|\widehat{v}\|_{s+\sigma_0} + \varepsilon^{2b-1} \|\mathfrak{J}_\delta\|_{s+\mu_0} \|\widehat{v}\|_{s_0+\mu_0}. \end{aligned} \quad (7.8)$$

*Proof.* The expression (7.4) follows from the Definition (6.14) and (6.5) by remarking that  $G_\delta$  (up to a translation) and the weak BNF transformation  $\Phi_B$  are the identity plus a finite rank operator, while the action angle change of coordinates is a rescaling plus a finite rank operator. Then by chain rule we get

$$\begin{aligned} d_{\bar{W}} \nabla_{\bar{W}} (H_\varepsilon \circ G_\delta) &\stackrel{(5.16)}{=} \varepsilon^{-2b} (d_W \nabla_W (\mathcal{H}_\mathbb{C} \circ \mathbf{A}_\varepsilon)) \circ G_\delta + R_1 = (d_z \nabla_z \mathcal{H}_\mathbb{C}) \circ \mathbf{A}_\varepsilon \circ G_\delta + R_1 \\ &\stackrel{(4.7)}{=} (d_W \nabla_W (H \circ \Lambda^{-1} \circ \Phi_B)) \circ \mathbf{A}_\varepsilon \circ G_\delta + R_1 \\ &= (d_W \nabla_W H) \circ \Lambda^{-1} \circ \Phi_B \circ \mathbf{A}_\varepsilon \circ G_\delta + R_1 + R_2 \end{aligned}$$

where the finite rank part contains all the terms where a derivative falls on  $\Phi_B \circ \mathbf{A}_\varepsilon \circ G_\delta$ . Then (7.4) follows from the definition of  $H$  in (1.6) (see (7.3)). Regarding the bounds (7.6), (7.7), we split the finite rank part  $R_1 + R_2$  as follows. The operator  $R_1$  contains all terms arising from derivatives of  $G_\delta$ . By tame estimates on the map  $G_\delta$  (see for instance Lemma 6.7 in [8]), it satisfies the bounds (7.7) and we put it in  $\mathcal{R}_{>12}$ . The finite rank term  $R_2$  comes from the Birkhoff map. This is an analytic map so we consider the Taylor expansion

$$\Phi_B(u, \bar{u}) = \left[ \frac{u}{\bar{u}} \right] + \sum_{i=2}^{12} \Psi_i(u, \bar{u}) + \Psi_{\geq 13}(u, \bar{u}),$$

where each  $\Psi_i(u)$  is homogeneous of degree  $i$  in  $u$ , while  $\Psi_{\geq 13} = O(u^{13})$  and they all map  $H_0^1(\mathbb{T}) \times H_0^1(\mathbb{T})$  in itself. We have to evaluate  $\Phi_B$  and its derivatives (up to order two) at  $u = T_\delta$  given in (7.2). We denote by  $v_I$  the traveling wave

$$v_I(\varphi, x) := \sum_{j \in S} \sqrt{\zeta_j} e^{i(jx + 1(j) \cdot \varphi)} = \sum_{j \in S} \sqrt{\zeta_j} e^{i(\varphi - vx) \cdot 1(j)} = A_\varepsilon(\varphi, 0, 0) \quad (7.9)$$

where  $1(\bar{j}_i) := -e_i$ ,  $e_i$  being the  $i$ -th vector of the canonical basis of  $\mathbb{Z}^\nu$ . We observe that<sup>7</sup>

$$\|v_\delta - v_I\|_s^{\gamma, \mathcal{O}_0} \lesssim \|\mathfrak{I}_\delta\|_s^{\gamma, \mathcal{O}_0},$$

and hence we can expand

$$\Phi_B(T_\delta) = \varepsilon \left[ \frac{v_I}{\bar{v}_I} \right] + \sum_{i=2}^{13} \varepsilon^i \Psi_i(v_I, \bar{v}_I) + \tilde{q} = \Phi_B^{\leq 12} + \tilde{q}, \quad (7.10)$$

where  $\tilde{q}$  is a remainder which satisfies

$$\|\tilde{q}\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \varepsilon^{13} + \varepsilon \|\mathfrak{I}_\delta\|_s^{\gamma, \mathcal{O}_0}, \quad \|d_i \tilde{q}[\tilde{v}]\|_s \lesssim_s \varepsilon (\|\tilde{v}\|_s + \|\mathfrak{I}_\delta\|_s \|\tilde{v}\|_{s_0}). \quad (7.11)$$

Then in  $\mathcal{R}_i$  we include all the terms homogeneous of degree  $i$  coming from derivatives of  $\Phi_B - \mathbf{I}$ , evaluated at  $\tilde{q} = 0$ ; we put in  $\mathcal{R}_{>12}$  all the rest. The (7.8) follows by (7.6), (7.7).  $\square$

**Remark 7.2.** *The motivation for separating the  $\mathcal{R}_i$  and  $\mathcal{R}_{>12}$  is the following. Consider the Hamiltonian  $H_\varepsilon$  as a function of  $\zeta$  instead of  $\omega$ . Then in all our expressions we can, and shall, evidence a purely polynomial term  $\sum_{i=0}^{12} \varepsilon^i f_i(\zeta)$  (where the  $f_i$  are  $\varepsilon$  independent) plus a remainder, which is not analytic in  $\varepsilon$ , of size  $\varepsilon^{13} + \varepsilon \|\mathfrak{I}_\delta\|_s^{\gamma, \mathcal{O}_0}$ . By the assumption (6.3), this means that in low norm  $s = s_0 + \mathfrak{p}_1$  all these remainders are negligible w.r.t. terms of order  $\varepsilon^{12}$ . In this framework  $\mathcal{R}_{>12}$  is purely a remainder, while the  $\mathcal{R}_i$  are homogeneous polynomial terms.*

**7.1. Homogeneity expansions.** In the following we shall assume that the assumption (7.1) holds true for some  $\mathfrak{p}_1$  large enough. The constant  $\mathfrak{p}_1$  represents the *loss of derivatives* accumulated along the reduction procedure of subsequent sections. In order to estimate the variation of the eigenvalues with respect to the approximate invariant torus, we need also to estimate the variation with respect to the torus embedding  $i(\varphi)$  in a low norm  $\|\cdot\|_p$  such that

$$s_0 \leq p \ll s_0 + \mathfrak{p}_1. \quad (7.12)$$

From now on we denote with  $\mu$  an increasing running index that represents the loss of derivatives appearing at each step of the reducibility procedure.

We need the following definition.

<sup>7</sup>The function  $\varepsilon \Lambda^{-1} \left[ \frac{v_I}{\bar{v}_I} \right]$  represents a torus supporting a quasi-periodic motion which is invariant for the system (5.19) with  $P = 0$ , namely it is the approximate solution from which we bifurcate.

**Definition 7.3.** Let  $1 \leq k \leq 6$ . Assume that (7.1) holds and let  $a \in S^m$  for some  $m \in \mathbb{R}$ ,  $\mathbf{R} \in \mathfrak{L}_{\rho,p}$  (see Def. 2.12, 2.10), depending in a Lipschitz way on  $\omega \in \mathcal{O} \subseteq \mathcal{O}_0$  and  $\mathfrak{J}_\delta$ . We say that  $a \in S_k^m$  if it can be written as

$$a(\varphi, x, \xi) = \sum_{i=1}^{14-2k} \varepsilon^i a_i(\varphi, x, \xi) + \tilde{a}(\varphi, x, \xi) \quad (7.13)$$

where  $a_i$  is a  $i$ -homogeneous symbol of the form

$$a_i(\varphi, x, \xi) := \frac{1}{(\sqrt{2\pi})^i} \sum_{\substack{j_1, \dots, j_i \in S \\ \sigma_i = \pm}} (\mathbf{a}_i)^{\sigma_1, \dots, \sigma_i}(\xi) \sqrt{\zeta_{j_1} \cdots \zeta_{j_i}} e^{i(\sigma_1 j_1 + \dots + \sigma_i j_i)x} e^{i(\sigma_1 \mathbf{1}(j_1) + \dots + \sigma_i \mathbf{1}(j_i)) \cdot \varphi} \quad (7.14)$$

for some Fourier multipliers  $(\mathbf{a}_i)^{\sigma_1, \dots, \sigma_i}(\xi) \in S^m$ , the symbol  $\tilde{a} \in S^m$  satisfies

$$\begin{aligned} |\tilde{a}|_{m,s,\alpha}^{\gamma, \mathcal{O}} &\lesssim_{m,s,\alpha} \gamma^{1-k} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \\ |\Delta_{12} \tilde{a}|_{m,p,\alpha} &\lesssim_{m,p,\alpha} \varepsilon \gamma^{1-k} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \end{aligned} \quad (7.15)$$

for some  $\mu = \mu(\nu) > 0$ . Similarly, we write  $\mathbf{R} \in \mathfrak{L}_{\rho,p}^k \otimes \mathcal{M}_2(\mathbb{C})$  to denote a remainder  $\mathbf{R}$  which has the form

$$\mathbf{R} = \sum_{i=1}^{14-2k} \varepsilon^i \mathbf{R}_i + \tilde{\mathbf{R}} \quad (7.16)$$

where  $\mathbf{R}_i$  is a  $i$ -homogeneous smoothing operator of the form

$$\begin{aligned} \mathbf{R}_i &= \begin{pmatrix} (\mathbf{R}_i)_+^+ & (\mathbf{R}_i)_+^- \\ (\mathbf{R}_i)_-^+ & (\mathbf{R}_i)_-^- \end{pmatrix}, \quad (\mathbf{R}_i)_\sigma^{\sigma'} = \overline{(\mathbf{R}_i)_{-\sigma}^{-\sigma'}}, \\ (\mathbf{R}_i)_\sigma^{\sigma'} z^{\sigma'} &:= \frac{1}{\sqrt{2\pi}} \sum_{j \in \mathbb{Z} \setminus \{0\}} e^{i\sigma j x} \left( \sum_{k \in \mathbb{Z}} (\mathbf{R}_i)_{\sigma,j}^{\sigma',k} z_k^{\sigma'} \right), \\ (\mathbf{R}_i)_{\sigma,j}^{\sigma',k} &:= \frac{1}{(\sqrt{2\pi})^i} \sum_{\sum_{q=1}^i \sigma_q j_q = \sigma j - \sigma' k} ((\mathbf{R}_i)_{\sigma,j}^{\sigma',k})_{j_1, \dots, j_i}^{\sigma_1, \dots, \sigma_i} \sqrt{\zeta_{j_1} \cdots \zeta_{j_i}} e^{i(\sigma_1 \mathbf{1}(j_1) + \dots + \sigma_i \mathbf{1}(j_i)) \cdot \varphi} \end{aligned} \quad (7.17)$$

for some coefficients  $((\mathbf{R}_i)_{\sigma,j}^{\sigma',k})_{j_1, \dots, j_i}^{\sigma_1, \dots, \sigma_i} \in \mathbb{C}$  and  $(\mathbf{R}_i)_\sigma^{\sigma'} \in \mathfrak{L}_{\rho,p}$ , the operator  $\tilde{\mathbf{R}} \in \mathfrak{L}_{\rho,p}(\mathcal{O})$  satisfies

$$\begin{aligned} \mathbb{M}_{\tilde{\mathbf{R}}}^\gamma(s, \mathbf{b}) &\lesssim_{s,\rho} \gamma^{1-k} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad 0 \leq \mathbf{b} \leq \rho - 2, \\ \mathbb{M}_{\Delta_{12} \tilde{\mathbf{R}}}(s, \mathbf{b}) &\lesssim_{p,\rho} \varepsilon \gamma^{1-k} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \quad 0 \leq \mathbf{b} \leq \rho - 3, \end{aligned} \quad (7.18)$$

for some  $\mu = \mu(\nu) > 0$ .

We now prove that the results of section 2.2.2 extends to the classes of symbols introduced above.

**Lemma 7.4.** Fix  $\rho, p$  as in Definition 2.10. Let  $m \in \mathbb{R}$ ,  $N := m + \rho$ ,  $1 \leq k \leq 6$ , and consider a symbol  $b(\varphi, x, \xi)$  in  $S_k^m$ . Assume the (7.1). Then there exists a remainder  $R_\rho \in \mathfrak{L}_{\rho,p}^k$  such that

$$Op(b) = Op^W(c) + R_\rho, \quad (7.19)$$

where  $c$  has the form (2.27), belongs to  $S_k^m$  and satisfies (7.15) with  $\mu \rightsquigarrow \mu + N$ .

*Proof.* One deduce the lemma on the symbol  $c$  by using the explicit expression (2.27) which is linear in  $b$ . So the homogeneous expansion of  $c$  follows from the one of  $b$ . The remainder  $R_\rho$  is equal to  $Op^W(\tilde{a})$  where  $\tilde{a}$  is in Lemma 2.14. The expression of  $\tilde{a}$  can be computed explicitly as one can deduce from the proof of Lemma 3.5 in [12]. In particular it is given in formulæ (3.2.11), (3.2.12), (3.2.13) therein. Such expressions are linear in  $b$ . Therefore  $\tilde{a}$  admits the expansion as in (7.13). The non homogeneous terms in the expansions of  $c$  and  $R_\rho$  satisfy estimates (7.15), (7.18) using the (2.28), (2.29) and the fact that, by hypothesis, the symbol  $b$  belongs to  $S_k^m$  and hence satisfies (7.15), (7.18).  $\square$

**Lemma 7.5. (Composition).** Let  $m, m' \in \mathbb{R}$ ,  $k, k' \in \mathbb{N}$ ,  $k' \leq k$ . Fix  $\rho, p$  as in Definition 2.10, such that  $\rho \geq \max\{-(m+m'+1), 3\}$  and define  $N := m+m'+\rho \geq 1$ . Consider two symbols  $a(\varphi, x, \xi) \in S_k^m$ ,  $b(\varphi, x, \xi) \in S_{k'}^{m'}$ . Assume the (7.1). There exist an operator  $R_\rho \in \mathfrak{L}_{\rho,p}^{k'}$  and a constant  $\mu = \mu(N) \sim N$  such that (recall Def. 2.17)

$$Op^W(a) \circ Op^W(b) = Op^W(c) + R_\rho, \quad c := a \#_N^W b \in S_{k'}^{m+m'}. \quad (7.20)$$

*Proof.* In Lemma 2.18 the symbol  $c$  is constructed by using the expansion  $a \#_N^W b$  (defined in (2.32)). Hence, by linearity, this expansion is a sum of  $a_i \#_N^W b_j$ , where  $a_i$  and  $b_j$  are the homogeneous terms of  $a$  and  $b$ , plus the term  $\tilde{a} \#_N^W \tilde{b}$  where  $\tilde{a}$  and  $\tilde{b}$  are respectively the non homogeneous symbols in the expansions for  $a$  and  $b$ . If  $i+j \geq 15-2k'$  then  $\varepsilon^{i+j} a_i \#_N^W b_j$  satisfies the (7.15) with  $k \rightsquigarrow k'$ , and then can be considered in  $\tilde{c}$ . The term  $\tilde{a} \#_N^W \tilde{b}$  satisfies (7.15) since  $\tilde{a}, \tilde{b}$  satisfies the same estimates. One can deduce from the proof of Lemma 2.18 that the remainder  $R_\rho$  is the sum of three operators  $L_\rho, \tilde{R}_\rho, Q_\rho$  belonging to  $\mathfrak{L}_{\rho,p}^{k'}$ . Indeed  $\tilde{R}_\rho, Q_\rho$  are obtained by applying Lemma 7.4 (which is the counterpart of Lemma 2.15). The operator  $L_\rho$  has an explicit expression (in terms of  $a, b$ ) given by formula (2.30) in [18]. The estimates (7.18) on the non-homogeneous term of  $R_\rho$  follow by the (2.35), (2.37) and the estimates on the symbols  $a, b$ .  $\square$

Recall the function  $v_I(\varphi, x)$  defined in (7.9). Notice that, by setting

$$v_j := (v_I)_j(\varphi) := \sqrt{\zeta_j} e^{i1(j)\cdot\varphi}, \quad j \in S, \quad (7.21)$$

we can write

$$v_I(\varphi, x) = \sum_{j \in S} \sqrt{\zeta_j} e^{i(jx+1(j)\cdot\varphi)} = \sum_{j \in S} v_j \frac{e^{ijx}}{\sqrt{2\pi}}. \quad (7.22)$$

We shall also use the notation  $v_j^\sigma$ ,  $\sigma = \pm$  where  $v_j^+ = v_j$ ,  $v_j^- = \overline{v_j}$ ,  $j \in S$ . Let  $a_k(\varphi, x, \xi)$  be a  $k$ -homogeneous symbol of the form (7.14). Of course such symbol depends only on the function  $v_I$  in (7.22). In particular we can write

$$a_k(\varphi, x, \xi) = \sum_{\substack{j_1, \dots, j_k \in S \\ \sigma_i = \pm}} (\mathbf{a}_k)_{j_1, \dots, j_k}^{\sigma_1, \dots, \sigma_k}(\xi) v_{j_1}^{\sigma_1} v_{j_2}^{\sigma_2} \dots v_{j_k}^{\sigma_k} e^{i(\sigma_1 j_1 + \dots + \sigma_k j_k)x}. \quad (7.23)$$

Notice that homogeneous symbols and operators of Definition 7.3 do not depend on  $\mathfrak{I}_\delta$ .

**Lemma 7.6.** *Consider a homogeneous symbol  $a_k(\varphi, x, \xi)$  with coefficients  $(\mathbf{a}_k)_{j_1, \dots, j_k}^{\sigma_1, \dots, \sigma_k}(\xi)$  as in (7.23). Then **Reality:** the symbol  $a_k(\varphi, x, \xi)$  is real valued if and only if*

$$\overline{(\mathbf{a}_k)_{j_1, \dots, j_k}^{\sigma_1, \dots, \sigma_k}(\xi)} = (\mathbf{a}_k)_{j_1, \dots, j_k}^{-\sigma_1, \dots, -\sigma_k}(\xi). \quad (7.24)$$

**Anti-reality:** the symbol  $a_k(\varphi, x, \xi)$  is purely imaginary if and only if

$$\overline{(\mathbf{a}_k)_{j_1, \dots, j_k}^{\sigma_1, \dots, \sigma_k}(\xi)} = -(\mathbf{a}_k)_{j_1, \dots, j_k}^{-\sigma_1, \dots, -\sigma_k}(\xi). \quad (7.25)$$

Moreover  $a_k \in S_k^m$  and  $\mathbf{R}_k \in \mathfrak{L}_{\rho,p}^k$  are  $x$ -translation invariant, i.e. satisfy respectively the (2.63) and (2.58).

*Proof.* It follows by explicit computations.  $\square$

In view of Remark 7.2 we show that the operator  $\mathcal{L}$  in (7.3) can be written in terms of symbols and operators belonging to the classes introduced in Definition 7.3.

**Lemma 7.7.** *The functions  $V, B$  in (3.10) and (3.11) belongs to the class  $S_1^0$  of Definition 7.3. In particular*

$$V(\Lambda^{-1} \Phi_B \circ T_\delta \left[ \frac{v_I}{\overline{v_I}} \right]) = \sum_{k=1}^{12} \varepsilon^k \mathbf{V}_k \left( \left[ \frac{v_I}{\overline{v_I}} \right] \right) + \mathbf{V}_{\geq 13} \quad (7.26)$$

with

$$\mathbf{V}_k \left( \left[ \frac{v_I}{\overline{v_I}} \right] \right) := \sum_{\substack{j_1, \dots, j_p \in S \\ \sigma_i = \pm}} (\mathbf{V}_k)_{j_1, \dots, j_p}^{\sigma_1, \dots, \sigma_k} \sqrt{\zeta_{j_1} \dots \zeta_{j_k}} e^{i(\sigma_1 j_1 + \dots + \sigma_k j_k)x} e^{i(\sigma_1 1(j_1) + \dots + \sigma_k 1(j_k))\cdot\varphi}, \quad (7.27)$$

and

$$(\mathbf{V}_1)_n^+ = (\mathbf{V}_1)_n^- = \frac{1}{\sqrt{2}} n |n|^{-1/4}, \quad (\mathbf{V}_2)_{n,n}^{+-} = (\mathbf{V}_2)_{n,n}^{-+} = \frac{1}{2} n |n|. \quad (7.28)$$

The Dirichlet-Neumann operator has the form  $G(\eta) = Op^W(|\xi|) + \mathcal{R}$  for some  $\mathcal{R} \in \mathfrak{L}_{\rho,p}^1$ , for any  $\rho \geq 3$ .



*Proof.* In view of (3.12), (3.13), and using (3.10), (3.11), the function  $V$  admits the expansion

$$V(\eta, \psi) = V(\Lambda^{-1}[\frac{u}{\bar{u}}]) = \sum_{k=1}^{12} \varepsilon^k \tilde{V}_k([\frac{u}{\bar{u}}]) + \tilde{V}_{\geq 13}([\frac{u}{\bar{u}}]). \quad (7.29)$$

The function  $\tilde{V}_{\geq 13}$  satisfies (7.11) thanks to Proposition B.2, and so it satisfies (7.15). The functions  $\tilde{V}_k$  is a homogeneous function of  $u, \bar{u}$  (of degree  $k$ ). The functions  $\tilde{V}_1, \tilde{V}_2$  are given in Lemma B.4. Moreover, since the Hamiltonian in (1.6) poisson commutes with the momentum in (1.7) (i.e.  $G(\eta)$  depends on  $x$  only through the variable  $\eta$ ) we deduce that  $\tilde{V}_k$  is supported, in Fourier, on monomials  $u_{j_1}^{\sigma_1} \cdots u_{j_k}^{\sigma_k}$  such that

$$\sum_{p=1}^k \sigma_p j_p = 0.$$

The (7.26), (7.27) follow by (7.21), (7.22), (7.10) and by using the (7.29). The estimates (7.15),  $k = 1$ , for  $V_{\geq 13}$  follow by the estimates on  $\tilde{V}_{\geq 13}$ , the estimates on  $\Phi_B$  and  $G_\delta$ . For the function  $B$  one can reason similarly.

The function  $V_1 = \tilde{V}_1$  and  $V_2 = \tilde{V}_2 + \tilde{V}_1 \circ \Psi_2$  (recall (7.10)). Then (7.28) follows by Lemma B.5 and the fact that the  $x$ -average of  $\tilde{V}_1 \circ \Psi_2$  is zero.

By Proposition B.1 we have that  $G(\eta) := Op^W(|\xi|) + R_G(\eta)$  with  $R_G(\eta)$  a pseudo differential operator satisfying (B.2). By Lemma B.2 in [31] we have that  $R_G(\eta)$  belongs to  $\mathcal{L}_{\rho,p}$ . By Taylor expanding  $R_G(\eta)$  in  $\eta$  one can deduce that actually  $R_G(\eta) \in \mathcal{L}_{\rho,p}^1$ .  $\square$

**7.2. Hamiltonian structure of the linearized operator.** Following Remark 7.2, we evidence the homogeneous terms in the Hamiltonian of  $\mathcal{L}_\omega$  whose Hamiltonian vector fields have degree  $\leq 13$ , since they are NOT perturbative. As explained in (7.10) this entails expanding the map  $\Phi_B(T_\delta)$  in powers of  $\varepsilon$  up to order five plus a *small remainder*  $\tilde{q}$ . We consider the symplectic form in the extended phase space  $(\varphi, Q, W) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$ , (recall the definition of  $W$  in (5.19), (3.16))

$$\Omega_e(\varphi, Q, W) := dQ \wedge d\varphi + d\tilde{\psi} \wedge d\tilde{\eta} \quad (7.30)$$

with the Poisson brackets (recalling  $\{\cdot, \cdot\}$  defined in (3.2))

$$\begin{aligned} \{F, G\}_e &:= \partial_Q F \partial_\varphi G - \partial_\varphi F \partial_Q G + \{F, G\} \\ &= \partial_Q F \partial_\varphi G - \partial_\varphi F \partial_Q G + \int_{\mathbb{T}} (\nabla_{\tilde{\eta}} G \nabla_{\tilde{\psi}} F - \nabla_{\tilde{\psi}} G \nabla_{\tilde{\eta}} F) dx. \end{aligned} \quad (7.31)$$

Passing to the complex variables  $(\varphi, Q, z, \bar{z})$  in (3.16), we have that, with abuse of notation, the extended symplectic form (7.30) and the Poisson brackets in (7.31) reads (recalling (3.6))

$$\begin{aligned} \Omega_e(\varphi, Q, W) &:= dQ \wedge d\varphi - idz \wedge d\bar{z}, \\ \{F_C, G_C\}_e &:= \partial_Q F_C \partial_\varphi G_C - \partial_\varphi F_C \partial_Q G_C + \{F_C, G_C\} \\ &= \partial_\varphi F_C \partial_\eta G_C - \partial_\eta F_C \partial_\varphi G_C - i \int_{\mathbb{T}} (\nabla_z G_C \nabla_{\bar{z}} F_C - \nabla_{\bar{z}} G_C \nabla_z F_C) dx \end{aligned} \quad (7.32)$$

where  $F_C = F \circ \Lambda^{-1}$  and  $G_C = G \circ \Lambda^{-1}$ .

We denote by  $H$  the Hamiltonian of the operator (7.4) with respect to the symplectic form (7.30). In the complex variables, we have

$$H \circ \Lambda^{-1} := H_0 + \sum_{i=1}^{12} \varepsilon^i H_i + H_{>12} + \sum_{i=2}^{12} \varepsilon^i H_{\mathcal{R}_i} + H_{\mathcal{R}_{>12}}, \quad (7.33)$$

$$H_0 = \bar{\omega} \cdot Q + \int_{\mathbb{T}} |D|^{\frac{1}{2}} z \cdot \bar{z} dx, \quad \text{and} \quad \|X_{H_{>12}}\|_s^{\gamma, \mathcal{O}_0}, \|X_{H_{\mathcal{R}_{>12}}}\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}, \quad (7.34)$$

for some  $\mu > 0$  and the Hamiltonians  $H_i$  are homogeneous in the variables  $v_I, \bar{v}_I$  given in (7.9). The functions  $H_{\mathcal{R}_i}, H_{\mathcal{R}_{>12}}$  are the quadratic forms associated to the linear operators  $\mathcal{R}_i, \mathcal{R}_{>12}$ , thus the estimates on the Hamiltonian vector fields can be deduced from (7.6), (7.7).

**Remark 7.8.** We note that  $\mathcal{L}_\omega$  is the linearized operator in the normal directions of the Hamiltonian  $\mathcal{H}_C = H_C \circ \Phi_B$  given in Proposition 4.3 written in the real variables. Since the map  $\Phi_B$  coincides, up to degree 2, with  $\Phi_{WB}$ , the

Taylor expansion (up to degree 2 in  $\varepsilon$ ) of  $H_C \circ \Phi_B$  coincides with the Hamiltonian  $H_C \circ \Phi_{WB}$  constructed in section 3.2. Therefore the  $\varepsilon^2$ -terms of  $\mathcal{L}_\omega$  are given by the Hamiltonian vector field of

$$\varepsilon^2(H_2 + H_{\mathcal{R}_2}) = (\Pi^{d_z=2} \widehat{H}_1^{(4)} \circ \mathbf{A}_{\varepsilon|y=0, \theta=\varphi})$$

where  $\widehat{H}_1^{(4)}$  is in (3.23).

**7.3. Algebraic properties of the linearized operator.** The linearized operator  $\mathcal{L}$  in (7.3) satisfies several algebraic properties which are consequence of the symmetries of the water waves vector field. The next lemma is fundamental for our scope.

**Lemma 7.9.** *Consider functions  $(\eta, \psi) \in S_v$  (see (2.46)) and the linearized operator  $\mathcal{L}$  in (7.3). Then  $\mathcal{L}$  is Hamiltonian and  $x$ -translation invariant. The same holds true for the operator  $\mathcal{L}_\omega$  in (6.14).*

*Proof.* The operator  $\omega \cdot \partial_\varphi - \mathcal{L}$  is the linearized hamiltonian vector field (1.3) at  $(\eta, \psi)$ . Hence  $J^{-1}(\omega \cdot \partial_\varphi - \mathcal{L})$  is symmetric. An explicit computation shows that  $\mathcal{C}J^{-1}(\omega \cdot \partial_\varphi - \mathcal{L})\mathcal{C}^{-1}$  is self-adjoint, i.e. satisfies (2.49)-(2.50). Now the Hamiltonian of the operator  $(\omega \cdot \partial_\varphi - \mathcal{L})$  is given by

$$\mathcal{Q}(\varphi, \widehat{\eta}, \widehat{\psi}) = \omega \cdot Q + \frac{1}{2} \int_{\mathbb{T}} J^{-1}(\omega \cdot \partial_\varphi - \mathcal{L}) \begin{bmatrix} \widehat{\eta} \\ \widehat{\psi} \end{bmatrix} \cdot \begin{bmatrix} \widehat{\eta} \\ \widehat{\psi} \end{bmatrix} dx.$$

Consider also the Momentum Hamiltonian

$$\mathcal{M}_{real}(\varphi, \widehat{\eta}, \widehat{\psi}) = -\mathbf{v} \cdot Q + \int_{\mathbb{T}} \widehat{\eta}_x \cdot \widehat{\psi} dx.$$

By an explicit computation (recall (7.31)) we have

$$\{\mathcal{M}_{real}, \mathcal{Q}\}_e = \frac{1}{2} \int_{\mathbb{T}} A_1(\varphi) \widehat{\psi} \cdot \widehat{\psi} dx + \frac{1}{2} \int_{\mathbb{T}} A_2(\varphi) \widehat{\eta} \cdot \widehat{\eta} dx - \frac{1}{2} \int_{\mathbb{T}} A_3(\varphi) \widehat{\eta} \cdot \widehat{\psi} dx,$$

where  $A_i(\varphi)$ ,  $i = 1, 2, 3$  are defined as

$$\begin{aligned} A_1(\varphi) &:= -(\mathbf{v} \cdot \partial_\varphi G)(\eta) + G(\eta) \partial_x - \partial_x G(\eta), \\ A_2(\varphi) &:= -\left( \mathbf{v} \cdot \partial_\varphi (BG(\eta)B + BV_x) \right) + (BG(\eta)B + BV_x) \partial_x - \partial_x (BG(\eta)B + BV_x), \\ A_3(\varphi) &:= -\left( \mathbf{v} \cdot \partial_\varphi (\partial_x V + G(\eta)B) \right) + (\partial_x V + G(\eta)B) \partial_x - \partial_x (\partial_x V + G(\eta)B). \end{aligned}$$

By Lemma B.3 we have that  $A_1(\varphi) \equiv 0$ . This is true since  $\eta \in S_v$ . Consider the operator  $A_3(\varphi)$ . First of all we note that

$$-(\mathbf{v} \cdot \partial_\varphi (\partial_x V)) + \partial_x V \partial_x - \partial_{xx} V = -(\mathbf{v} \cdot \partial_\varphi V + V_x) \partial_x - \mathbf{v} \cdot \partial_\varphi V_x - V_{xx} = 0$$

since the function  $V(\eta, \psi) \in S_v$  (see item (ii) of Lemma B.3). Moreover (recall (3.9))

$$-(\mathbf{v} \cdot \partial_\varphi (G(\eta)B)) + G(\eta)B \partial_x - \partial_x G(\eta)B = -G'(\eta) [\mathbf{v} \cdot \partial_\varphi \eta + \eta_x] B - G(\eta) (\mathbf{v} \cdot \partial_\varphi B + B_x) = 0,$$

since (recall Lemmata 2.21, B.3)  $\eta, B(\eta, \psi) \in S_v$ . Then we have  $A_3(\varphi) \equiv 0$ . Reasoning similarly one can check that also the operator  $A_2(\varphi)$  is identically zero. Then we proved that  $\{\mathcal{M}_{real}, \mathcal{Q}\}_e = 0$ , which implies (writing the Hamiltonians  $\mathcal{M}_{real}, \mathcal{Q}$  in complex variables) the (2.57). Hence, by Lemma 2.22, we have that  $\omega \cdot \partial_\varphi - \mathcal{L}$  is  $x$ -translation invariant, i.e.  $\mathcal{C}(\omega \cdot \partial_\varphi - \mathcal{L})\mathcal{C}^{-1}$  satisfies (2.51). The operator  $\mathcal{L}_\omega$  in (6.14) is Hamiltonian by the construction of section 6. Moreover, by Remark 6.6 (see equation (6.7)), we deduce that

$$\mathbf{v} \cdot \partial_\varphi K_{02}(\varphi) + \partial_x K_{02}(\varphi) - K_{02}(\varphi) \partial_x = 0.$$

This implies that the coefficients of  $K_{02}(\varphi)$  satisfy (2.58). By Lemma 2.22 we have that the operator  $K_{02}(\varphi)$ , and hence  $\mathcal{L}_\omega$ , is  $x$ -translation invariant.  $\square$

## 8. SYMMETRIZATION OF THE LINEARIZED OPERATOR AT THE HIGHEST ORDER

The aim of the following sections is to conjugate the linearized operator  $\mathcal{L}_\omega$  in (7.4) to a constant coefficients operator, up to a regularizing remainder. This will be achieved by applying several transformations which clearly depends nonlinearly on the point  $(\eta, \psi)$  in (5.32) on which we linearized.

**8.1. Good unknown of Alinhac.** The aim of this section is to rewrite the operator  $\mathcal{L}_\omega$  in (7.4) in terms of the so called ‘‘good unknown’’ of Alinhac (more precisely its symplectic correction). This will be done in Proposition 8.2. As we will see, these coordinates are the correct ones in order to diagonalize, at the highest order, the operator  $\mathcal{L}_\omega$ . We shall first prove some preliminary results. Following [1], [18] we conjugate the linearized operator  $\mathcal{L}$  in (7.3) by the operator

$$\Phi_{\mathbb{B}} = \Pi_S^\perp \mathcal{G} \Pi_S^\perp = \text{Id} + \mathbb{B}, \quad \mathbb{B} := \Pi_S^\perp \begin{pmatrix} 0 & 0 \\ B & 0 \end{pmatrix} \Pi_S^\perp, \quad \mathcal{G} := \begin{pmatrix} 1 & 0 \\ B & 1 \end{pmatrix}, \quad (8.1)$$

where  $B$  is the real valued function in (3.11). Define the function

$$a = a(\varphi, x) := (\omega \cdot \partial_\varphi B) + V B_x. \quad (8.2)$$

In the following lemma we study some properties of the map  $\Phi_{\mathbb{B}}$ .

**Lemma 8.1.** *The function  $a(\varphi, x) \in S_1^0$  and it satisfies (2.63). Moreover (recall Definition 2.4) the maps  $\mathcal{G}^{\pm 1} - \text{Id}$  and  $(\mathcal{G}^{\pm 1} - \text{Id})^*$  satisfy*

$$\begin{aligned} \mathfrak{M}_{(\mathcal{G}^{\pm 1} - \text{Id})}^\gamma(s) + \mathfrak{M}_{(\mathcal{G}^{\pm 1} - \text{Id})^*}^\gamma(s) &\lesssim_s \varepsilon (1 + \|\mathfrak{I}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \\ \|\Delta_{12}(\mathcal{G}^{\pm 1} - \text{Id})h\|_p + \|\Delta_{12}(\mathcal{G}^{\pm 1} - \text{Id})^*h\|_p &\lesssim_p \varepsilon \|h\|_p \|i_1 - i_2\|_{p+\mu}. \end{aligned} \quad (8.3)$$

The map  $\Phi_{\mathbb{B}}$  is symplectic w.r.t the symplectic form (7.30) and  $x$ -translation invariant, i.e. is in  $\mathfrak{T}_1$  (see Def. 2.28). One has that  $\Phi_{\mathbb{B}}^{-1} = \text{Id} - \mathbb{B}$ . Moreover  $\Phi_{\mathbb{B}}^{-1}$ ,  $\Phi_{\mathbb{B}}$  satisfy the estimates (8.3).

*Proof.* The homogeneity expansion of  $a$  (with estimates (7.15)) follows from the ones of  $V$  and  $B$  (given in Lemma 7.7) and by formula (8.2). The functions  $a, V, B$  satisfy the (2.63) by Lemma B.3. Let us now consider the Hamiltonian

$$H_{\mathbb{B}} = \frac{1}{2} \int_{\mathbb{T}} J^{-1} \mathbb{B} \begin{bmatrix} \tilde{\eta} \\ \tilde{\psi} \end{bmatrix} \cdot \begin{bmatrix} \tilde{\eta} \\ \tilde{\psi} \end{bmatrix} dx,$$

where  $J$  is in (3.1) and denote by  $\Phi_{\mathbb{B}}^\tau$ ,  $\tau \in [0, 1]$ , the flow generated by the Hamiltonian  $H_{\mathbb{B}}$ . We have that

$$\Phi_{\mathbb{B}} = \Phi_{\mathbb{B}}^1 = \exp(\mathbb{B}) = \sum_{n \geq 0} \frac{1}{n!} \mathbb{B}^n = \text{Id} + \mathbb{B} + \mathcal{R}_1, \quad \mathcal{R}_1 := \sum_{n=2}^{\infty} \frac{1}{n!} \mathbb{B}^n.$$

Using that the matrix  $\begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix}$  is nilpotent, we deduce that

$$\mathbb{B}^2 = -\Pi_S^\perp \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix} \Pi_S \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix} \Pi_S^\perp, \quad \mathbb{B}^n = -(-1)^n \Pi_S^\perp \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix} (\Pi_S \begin{bmatrix} 0 & 0 \\ B & 0 \end{bmatrix})^{n-1} \Pi_S^\perp.$$

We note that, given  $W = \Pi_S^\perp W = \begin{bmatrix} \tilde{\eta} \\ \tilde{\psi} \end{bmatrix}$ , we have  $\mathbb{B}^2 W \equiv 0$ . Therefore  $\mathcal{R}_1 = 0$  and formula (8.1) follows. By Lemma B.3, the function  $B$  belongs to  $S_v$  hence  $\mathcal{G}$  is  $x$ -translation invariant.  $\square$

We now study the conjugate of the operator  $\mathcal{L}_\omega$ .

**Proposition 8.2. (Symplectic good unknown).** *The conjugate the operator  $\mathcal{L}_\omega$  in (7.4) under the map  $\Phi_{\mathbb{B}}$  in (8.1) has the form*

$$\mathcal{L}_0 = \Phi_{\mathbb{B}}^{-1} \mathcal{L}_\omega \Phi_{\mathbb{B}} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} i\xi V + \frac{V_x}{2} & -|\xi| \\ 1+a & i\xi V - \frac{V_x}{2} \end{pmatrix} + \tilde{\mathcal{Q}}_0 \right), \quad (8.4)$$

where  $\tilde{\mathcal{Q}}_0 \in \mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$  (see Def. 2.10, 7.3) for any  $\rho \geq 3$ . Finally the operator  $\mathcal{L}_0$  is real-to-real, Hamiltonian and  $x$ -translation invariant.

*Proof.* Notice that

$$\begin{aligned} \Phi_{\mathbb{B}}^{-1} \mathcal{L}_\omega \Phi_{\mathbb{B}} &= \Pi_S^\perp \mathcal{G}^{-1} \mathcal{L} \mathcal{G} \Pi_S^\perp + \tilde{\mathcal{Q}}_0, \\ \tilde{\mathcal{Q}}_0 &:= \Pi_S^\perp \mathcal{G}^{-1} \mathcal{Q}_0 \mathcal{G} \Pi_S^\perp - \Pi_S^\perp \mathcal{G}^{-1} \Pi_S (\mathcal{L} + \mathcal{Q}_0) \Pi_S^\perp \mathcal{G} \Pi_S^\perp - \Pi_S^\perp \mathcal{G}^{-1} (\mathcal{L} + \mathcal{Q}_0) \Pi_S \mathcal{G} \Pi_S^\perp. \end{aligned}$$

Recalling (7.3) we have that

$$\mathcal{G}^{-1} \mathcal{L} \mathcal{G} = \omega \cdot \partial_\varphi + \begin{pmatrix} \partial_x V & -G(\eta) \\ 1+a & V \partial_x \end{pmatrix} \quad (8.5)$$

where  $a$  is the function defined in (8.2) By applying Lemma 7.5 we have

$$\begin{aligned}\partial_x V &= Op^W(i\xi) \circ Op^W(V) = Op^W\left(i\xi V + \frac{1}{2i}\{i\xi, V\}\right) = Op^W\left(i\xi V + \frac{V_x}{2}\right), \\ V\partial_x &= Op^W(V) \circ Op^W(i\xi) = Op^W\left(i\xi V + \frac{1}{2i}\{V, i\xi\}\right) = Op^W\left(i\xi V - \frac{V_x}{2}\right),\end{aligned}$$

up to smoothing terms in  $\mathfrak{L}_{\rho,p}^1$ . Moreover, by Lemma 7.7, we also have that  $G(\eta)[\cdot] := |D| = Op^W(|\xi|)$  up to smoothing term  $\mathfrak{L}_{\rho,p}^1$ . Then the formula for the pseudo differential operator in (8.4) follows. The remainder  $\tilde{\mathcal{Q}}_0$  is finite rank of the form (7.5). Notice that the remainder  $\mathcal{Q}_0$  in Proposition 7.1 admits expansions in homogeneous remainders, i.e. belongs to  $\mathfrak{L}_{\rho,p}^1$  also using Lemma C.7 in [32]. Using Lemma 7.7 we conclude that  $\tilde{\mathcal{Q}}_0 \in \mathfrak{L}_{\rho,p}^1$ . The operator  $\mathcal{L}_0$  is Hamiltonian and  $x$ -translation invariant thanks to the properties of  $\Phi_{\mathbb{B}}$  and Lemma 7.9.  $\square$

**8.2. Complex formulation of Water waves.** We want to rewrite the operator  $\mathcal{L}_0(\varphi)$  in (8.4) in the complex coordinates (3.4). Following the strategy used in Proposition 3.3 in [14] we prove the following result.

**Proposition 8.3. (Linearized operator in complex variables).** *There is  $\mu = \mu(\nu) > 0$  and a matrix of symbols  $A_{-1} \in S_1^{-1} \otimes \mathcal{M}_2(\mathbb{C})$  (see Def. 7.3), satisfying (2.63), such that (recall (3.4), (8.4))*

$$\begin{aligned}\mathcal{L}_1(\varphi) &:= \Lambda \mathcal{L}_0(\varphi) \Lambda^{-1} := \\ &\Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \left( iA_1(\varphi, x)\xi + iA_{1/2}(\varphi, x)|\xi|^{\frac{1}{2}} + A_0(\varphi, x) + A_{-1}(\varphi, x, \xi) \right) + \mathcal{R}^{(1)} \right),\end{aligned}\tag{8.6}$$

where  $\mathcal{R}^{(1)} = \mathcal{R}^{(1)}(\varphi)$  is in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$  and

$$A_1(\varphi, x) := \begin{pmatrix} V(\varphi, x) & 0 \\ 0 & V(\varphi, x) \end{pmatrix}\tag{8.7}$$

$$A_{1/2}(\varphi, x) := \begin{pmatrix} 1 + \tilde{a}(\varphi, x) & \tilde{a}(\varphi, x) \\ -\tilde{a}(\varphi, x) & -(1 + \tilde{a}(\varphi, x)) \end{pmatrix}, \quad \tilde{a} := \frac{1}{2}(\omega \cdot \partial_\varphi B + VB_x),\tag{8.8}$$

$$A_0(\varphi, x) := \frac{1}{4} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} V_x(\varphi, x).\tag{8.9}$$

*Proof.* We start by applying the change of variables  $\mathfrak{F}$  in (3.3). Using Lemma 7.5 and the (8.4) we get

$$\mathfrak{F} \mathcal{L}_0 \mathfrak{F}^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \left( \begin{bmatrix} |\xi|^{-1/4} & 0 \\ 0 & |\xi|^{1/4} \end{bmatrix} \#_\rho^W \begin{bmatrix} iV\xi + \frac{V_x}{2} & -|\xi| \\ 1+a & iV\xi - \frac{V_x}{2} \end{bmatrix} \#_\rho^W \begin{bmatrix} |\xi|^{1/4} & 0 \\ 0 & |\xi|^{-1/4} \end{bmatrix} \right) \left[ \begin{smallmatrix} \tilde{\eta} \\ \tilde{\omega} \end{smallmatrix} \right] \right)$$

up to a remainder in  $\mathfrak{L}_{\rho,p}^1$ , i.e. a remainder which admits an expansion as (7.16), (7.17) with estimates (7.18). By expanding the symbols on the diagonal, using formula (2.32), we get

$$|\xi|^{-1/4} \#_\rho^W \left( iV\xi + \frac{V_x}{2} \right) \#_\rho^W |\xi|^{1/4} = iV\xi + \frac{V_x}{4}, \quad |\xi|^{1/4} \#_\rho^W \left( iV\xi - \frac{V_x}{2} \right) \#_\rho^W |\xi|^{-1/4} = iV\xi - \frac{V_x}{4},$$

up to symbols in  $S_1^{-1}$ . Similarly we get

$$|\xi|^{-1/4} \#_\rho^W (-|\xi|) \#_\rho^W |\xi|^{-1/4} = -|\xi|^{1/2}, \quad |\xi|^{1/4} \#_\rho^W (1+a_0) \#_\rho^W |\xi|^{1/4} = (1+a)|\xi|^{1/2},$$

up to symbols in  $S_1^{-3/2}$ . All this new symbols of order less or equal  $-1$  and belongs to  $S_1^{-1}$ . By applying the change of coordinates  $\mathcal{C}$  in (3.3) we get that  $\mathcal{C} \mathfrak{F} \mathcal{L}_0 \mathfrak{F}^{-1} \mathcal{C}^{-1} = \Lambda \mathcal{L}_0 \Lambda^{-1}$  has the form (8.6) with matrices of symbols in (8.7)-(8.9).  $\square$

**Lemma 8.4.** *The operator  $\mathcal{L}_1$  in (8.6) belongs to the class  $\mathfrak{S}_1$  given in Definition 2.26.*

*Proof.* First notice that, by Proposition 8.2, the operator  $\mathcal{L}_0$  in (8.4) is real-to-real, Hamiltonian and  $x$ -translation invariant. Moreover the map  $\Lambda$  in (3.4) is *symplectic,  $x$ -translation invariant*. Hence also  $\mathcal{L}_1$  satisfies the same properties of  $\mathcal{L}_0$ . To prove that  $\mathcal{L}_1 \in \mathfrak{S}_1$  we need to prove that the pseudo differential operator in (8.6) is itself Hamiltonian and  $x$ -translation invariant (it is clearly real-to-real). By Lemmata 2.24, 2.25 we just have to show that the matrix of symbols

$$A(\varphi, x, \xi) := iA_1(\varphi, x)\xi + iA_{1/2}(\varphi, x)|\xi|^{\frac{1}{2}} + A_0(\varphi, x) + A_{-1}(\varphi, x, \xi)$$

satisfies (2.68) and (2.63). The condition (2.63) follows trivially by the fact that the functions  $V, B$  belongs to  $S_\nu$  and that the expansion (2.32) preserves this property. The (2.68) follows by the explicit computations performed in Proposition 8.3 since the functions  $V, B$  in (3.10), (3.11) and  $\tilde{a}$  in (8.8) are real valued. Therefore  $\mathcal{L}_1 \in \mathfrak{S}_1$ .  $\square$

## 9. BLOCK-DIAGONALIZATION

In this section we block-diagonalize the operator  $\mathcal{L}_1$  in (8.6) up to a  $\rho$ -smoothing remainders in  $\mathfrak{L}_{\rho,p}$ .

**9.1. Block-diagonalization at order 1/2.** In this subsection we diagonalize the matrix  $A_{1/2}(\varphi, x)|\xi|^{1/2}$  in (8.8). We consider the multiplication operator  $\mathcal{M}$  defined as

$$\begin{aligned} \mathcal{M} &:= \mathcal{M}(\varphi) := \begin{pmatrix} f & g \\ g & f \end{pmatrix}, \\ f &:= f(\varphi, x) := \frac{1 + \tilde{a} + \lambda_+}{\sqrt{(1 + \tilde{a} + \lambda_+)^2 - \tilde{a}^2}}, \quad g := g(\varphi, x) := \frac{-\tilde{a}}{\sqrt{(1 + \tilde{a} + \lambda_+)^2 - \tilde{a}^2}}, \\ \lambda_{\pm} &:= \lambda_{\pm}(\varphi, x) := \pm \sqrt{(1 + \tilde{a})^2 - \tilde{a}^2} \end{aligned} \quad (9.1)$$

where the functions  $\lambda_{\pm}$  are the eigenvalues of  $A_{1/2}$  in (8.8). Notice that  $f \geq 0$ . Notice also that

$$\det(\mathcal{M}) = f^2 - g^2 = 1, \quad \mathcal{M}^{-1} = \begin{pmatrix} f & -g \\ -g & f \end{pmatrix}, \quad (9.2)$$

$$A_{1/2}^{(2)} := \mathcal{M}^{-1} A_{1/2} \mathcal{M} = \begin{pmatrix} \lambda_+ & 0 \\ 0 & -\lambda_+ \end{pmatrix} = \begin{pmatrix} 1 + a^{(0)} & 0 \\ 0 & -(1 + a^{(0)}) \end{pmatrix}, \quad a^{(0)} := \lambda_+ - 1. \quad (9.3)$$

The main result of this section is the following.

**Proposition 9.1. (Block-diagonalization at order 1/2).** *There exists an invertible map  $\Phi_{\mathbb{M}}(\varphi) : H_{S^{\perp}}^s(\mathbb{T}) \times H_{S^{\perp}}^s(\mathbb{T}) \rightarrow H_{S^{\perp}}^s(\mathbb{T}) \times H_{S^{\perp}}^s(\mathbb{T})$  such that (recall (8.6))*

$$\begin{aligned} \mathcal{L}_2 &:= \Phi_{\mathbb{M}} \mathcal{L}_1 \Phi_{\mathbb{M}}^{-1} := \Pi_{S^{\perp}}^{\perp} \omega \cdot \partial_{\varphi} \\ &+ \Pi_{S^{\perp}}^{\perp} \left( O_p^W \left( i A_1(\varphi, x) \xi + i A_{1/2}^{(2)}(\varphi, x) |\xi|^{\frac{1}{2}} + A_0^{(2)}(\varphi, x) + A_{-1/2}^{(2)}(\varphi, x, \xi) \right) + \mathcal{R}^{(2)} \right) \end{aligned} \quad (9.4)$$

where  $\mathcal{R}^{(2)} = \mathcal{R}^{(2)}(\varphi)$  is in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ ,  $A_{-1/2}^{(2)}$  is in  $S_1^{-\frac{1}{2}} \otimes \mathcal{M}_2(\mathbb{C})$  and satisfies (2.63),  $A_1$  is the matrix in (8.7),  $A_{1/2}^{(2)}$  is in (9.3) and

$$A_0^{(2)}(\varphi, x) := \begin{pmatrix} 0 & b^{(2)} \\ b^{(2)} & 0 \end{pmatrix}, \quad b^{(2)} := \frac{V_x}{4} - V(f_x g - g_x f), \quad (9.5)$$

where the functions  $f, g$  are given in (9.1). Moreover

$$\begin{aligned} \mathfrak{M}_{(\Phi_{\mathbb{M}}^{\pm} - \text{Id})}^{\gamma}(s) + \mathfrak{M}_{(\Phi_{\mathbb{M}}^{\pm} - \text{Id})^*}^{\gamma}(s) &\lesssim_s \varepsilon (1 + \|\mathfrak{J}_{\delta}\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \\ \|\Delta_{12}(\Phi_{\mathbb{M}}^{\pm} - \text{Id})h\|_p + \|\Delta_{12}(\Phi_{\mathbb{M}}^{\pm} - \text{Id})^*h\|_p &\lesssim_p \varepsilon \|h\|_p \|i_1 - i_2\|_{p+\mu}. \end{aligned} \quad (9.6)$$

We divide the proof of the proposition above into several steps. We have the following.

**Lemma 9.2.** *The functions  $a^{(0)}$  in (9.3),  $f, g$  in (9.1) belong to the class  $S_1^0$  of Def. 7.3. In particular the functions  $\tilde{a}, f, g$  are  $x$ -translation invariant, i.e. satisfies (2.63). Moreover (recall Definition 2.4) the maps  $\mathcal{M}^{\pm} - \text{Id}$  and  $(\mathcal{M}^{\pm} - \text{Id})^*$  (see (9.1)) satisfy*

$$\begin{aligned} \mathfrak{M}_{(\mathcal{M}^{\pm} - \text{Id})}^{\gamma}(s) + \mathfrak{M}_{(\mathcal{M}^{\pm} - \text{Id})^*}^{\gamma}(s) &\lesssim_s \varepsilon (1 + \|\mathfrak{J}_{\delta}\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \\ \|\Delta_{12}(\mathcal{M}^{\pm} - \text{Id})h\|_p + \|\Delta_{12}(\mathcal{M}^{\pm} - \text{Id})^*h\|_p &\lesssim_p \varepsilon \|h\|_p \|i_1 - i_2\|_{p+\mu}. \end{aligned} \quad (9.7)$$

*Proof.* One has  $a^{(0)}, f, g \in S_1^0$  since they are regular functions of  $\tilde{a}$  (see (9.3), (9.1)) and the fact that  $\tilde{a}$  in (8.8) belongs to  $S_1^0$  by Lemma 8.1. The (9.7) follows by Lemma 2.13 in [18].  $\square$

In order to conjugate the operator  $\mathcal{L}_1$  in (8.6) we shall consider the map  $\mathcal{M}^{\perp} := \Pi_{S^{\perp}}^{\perp} \mathcal{M} \Pi_{S^{\perp}}^{\perp}$ . Unfortunately the map  $\mathcal{M}^{\perp}$  is not symplectic with respect to the symplectic form (7.30). We first construct a symplectic correction.

**Lemma 9.3.** *Consider the map  $\mathcal{M}$  in (9.1). There is a function  $c = c(\varphi, x) \in S_1^0$  such that  $\mathcal{M}^{-1} := \Phi_C^1$  where  $\Phi_C^{\tau}$  is the flow of*

$$\partial_{\tau} \Phi_C^{\tau} = C \Phi_C^{\tau}, \quad \Phi_C^0 = \text{Id}, \quad C := C(\varphi, x) := \begin{pmatrix} 0 & c(\varphi, x) \\ c(\varphi, x) & 0 \end{pmatrix}. \quad (9.8)$$

*Proof.* The proof is based on ideas used in Lemma 5.2 in [33] (see also [34]). One has that

$$\Phi_C^1 := S(\varphi) := \exp\{C(\varphi, x)\} := \begin{pmatrix} c_1(\varphi, x) & c_2(\varphi, x) \\ c_2(\varphi, x) & c_1(\varphi, x) \end{pmatrix}, \quad \begin{aligned} c_1 &= c_1(\varphi, x) := \cosh(|c(\varphi, x)|), \\ c_2 &= c_2(\varphi, x) := \frac{c(\varphi, x)}{|c(\varphi, x)|} \sinh(|c(\varphi, x)|). \end{aligned}$$

Note that the function  $c_2(\varphi, x)$  above is not singular indeed

$$c_2 = \frac{c}{|c|} \sinh(|c|) = \frac{c}{|c|} \sum_{k=0}^{\infty} \frac{(|c|)^{2k+1}}{(2k+1)!} = c \sum_{k=0}^{\infty} \frac{(c\bar{c})^k}{(2k+1)!}.$$

We note moreover that for any  $x \in \mathbb{T}$  one has  $\det(S(\varphi, x)) = 1$ , hence its inverse  $S^{-1}(\varphi, x)$  is well defined. We choose  $c(\varphi, x)$  in such a way that  $S^{-1}(\varphi, x) := \mathcal{M}$  (see (9.1)). Therefore we have to solve the following equations

$$\cosh(|c(\varphi, x)|) = f(\varphi, x), \quad \frac{c(\varphi, x)}{|c(\varphi, x)|} \sinh(|c(\varphi, x)|) = -g(\varphi, x).$$

Concerning the first one we note that  $f^2 - 1 = \frac{|\bar{a}|^2}{2\lambda_+(1+\bar{a}+\lambda_+)} \geq 0$ . Therefore

$$|c(\varphi, x)| := \operatorname{arccosh}(f(\varphi, x)) = \ln \left( f(\varphi, x) + \sqrt{(f(\varphi, x))^2 - 1} \right),$$

is well-defined. For the second equation one observes that the function

$$\frac{\sinh(|c(\varphi, x)|)}{|c(\varphi, x)|} = 1 + \sum_{k \geq 0} \frac{(c(\varphi, x)\bar{c}(\varphi, x))^k}{(2k+1)!} \geq 1,$$

hence we set

$$c(\varphi, x) := g(\varphi, x) \frac{|c(\varphi, x)|}{\sinh(|c(\varphi, x)|)}. \quad (9.9)$$

Using that  $f, g \in S_1^0$  (see Lemma 9.2), and using the (9.9) we deduce that  $c \in S_1^0$ .  $\square$

**Proof of Proposition 9.1.** Let us define  $\Phi_{\mathcal{M}} = \Phi_{\mathcal{M}}^1$  where  $\Phi_{\mathcal{M}}^\tau$  is the flow of

$$\partial_\tau \Phi_{\mathcal{M}}^\tau = \Pi_S^\perp C \Pi_S^\perp \Phi_{\mathcal{M}}^\tau, \quad \Phi_{\mathcal{M}}^0 = \operatorname{Id}, \quad (9.10)$$

where  $C$  is the matrix in (9.8) with symbol  $c(\varphi, x)$  given by Lemma 9.3. By Lemma A.4 (see also Remark A.6) we have that

$$\Phi_{\mathcal{M}}^1 = \Pi_S^\perp \Phi_C^1 \Pi_S^\perp \circ (\operatorname{Id} + \mathcal{R})$$

where  $\Phi_C^1 = \mathcal{M}^{-1}$  is the flow of (9.8) (see also (9.1)) and  $\mathcal{R}$  is a matrix of finite rank operators of the form (A.14) satisfying estimates (A.15). Hence  $\Phi_{\mathcal{M}}$  is well-defined and satisfies bounds (9.6) thanks to (9.7). By Lemma A.5 we have that the conjugate  $\mathcal{L}_2 := \Phi_{\mathcal{M}} \mathcal{L}_1 \Phi_{\mathcal{M}}^{-1}$  is given by  $\Pi_S^\perp \mathcal{M}^{-1} \mathcal{L}_1 \mathcal{M} \Pi_S^\perp$  up to finite rank remainders belonging to  $\mathcal{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . We have that

$$\mathcal{M}^{-1} \mathcal{L}_1 \mathcal{M} = \mathcal{M}^{-1} \omega \cdot \partial_\varphi \mathcal{M} + \mathcal{M}^{-1} O p^W (iA_1 \xi + iA_{1/2} |\xi|^{\frac{1}{2}} + A_0 + A_{-1}) \mathcal{M} + \mathcal{M}^{-1} \mathcal{R}^{(1)} \mathcal{M}.$$

Hence Lemma 7.5 implies that

$$\mathcal{M}^{-1} \mathcal{L}_1 \mathcal{M} = O p^W \left( \mathcal{M}^{-1} \#_\rho^W \omega \cdot \partial_\varphi \mathcal{M} + \mathcal{M}^{-1} \#_\rho^W (iA_1 \xi + iA_{1/2} |\xi|^{\frac{1}{2}} + A_0 + A_{-1}) \#_\rho^W \mathcal{M} \right)$$

up to terms in  $\mathcal{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . In the following we apply several times Lemma 7.5 on the composition of pseudo differential operators. By (9.1), (9.2) we deduce that  $(\omega \cdot \partial_\varphi f) f - (\omega \cdot \partial_\varphi g) g = 0$ . Hence we have

$$\mathcal{M}^{-1} \#_\rho^W \omega \cdot \partial_\varphi \mathcal{M} = \begin{pmatrix} 0 & -(\omega \cdot \partial_\varphi f) g + (\omega \cdot \partial_\varphi g) f \\ -(\omega \cdot \partial_\varphi f) g + (\omega \cdot \partial_\varphi g) f & 0 \end{pmatrix}.$$

By (8.7), using symbolic calculus and  $f^2 - g^2 = 1$  (see (9.2)), we obtain the exact expansion

$$\mathcal{M}^{-1} \#_\rho^W (iA_1 \xi) \#_\rho^W \mathcal{M} = \begin{pmatrix} iV\xi & -V(f_x g - g_x f) \\ -V(f_x g - g_x f) & iV\xi \end{pmatrix}.$$

By (9.3) we have

$$\mathcal{M}^{-1} \#_\rho^W (iA_{1/2} |\xi|^{\frac{1}{2}}) \#_\rho^W \mathcal{M} = i \begin{pmatrix} (1 + a^{(0)}) |\xi|^{\frac{1}{2}} & 0 \\ 0 & -(1 + a^{(0)}) |\xi|^{\frac{1}{2}} \end{pmatrix} + r_{-1/2},$$



for some  $r_{-1/2} \in S_1^{-\frac{1}{2}} \otimes \mathcal{M}_2(\mathbb{C})$ . Moreover, recalling (8.9) and that  $A_{-1} \in S_1^{-1} \otimes \mathcal{M}_2(\mathbb{C})$ , we have

$$\mathcal{M}^{-1} \#_{\rho}^W A_0 \#_{\rho}^W \mathcal{M} = A_0 = \frac{1}{4} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} V_x, \quad \mathcal{M}^{-1} \#_{\rho} A_{-1} \#_{\rho}^W \mathcal{M} =: r_{-1},$$

with  $r_{-1} \in S_1^{-1} \otimes \mathcal{M}_2(\mathbb{C})$ . By the discussion above we obtain the (9.4) and (9.5).  $\square$

We conclude this section with the following result.

**Lemma 9.4.** *The operator  $\mathcal{L}_2$  in (9.4) is in  $\mathfrak{S}_1$  and the map  $\Phi_{\mathbb{M}}$  in Proposition 9.1 is in  $\mathfrak{T}_1$  (see Def. 2.26, 2.28).*

*Proof.* We first note the following. The functions  $f, g$  in (9.1) are real valued and satisfy (2.63) since, by Lemma 8.4, the operator  $\mathcal{L}_1$  is in  $\mathfrak{S}_1$ . Therefore the function  $c(\varphi, x)$  given by Lemma 9.3 is in  $S_v$ . As a consequence the flow of (9.10) is in  $\mathfrak{T}_1$ , i.e. is symplectic and  $x$ -translation invariant. Then also the operator  $\mathcal{L}_2$  is Hamiltonian and  $x$ -translation invariant. The matrices  $iA_1, iA_{1/2}, A_j^{(2)}, j = 0, -1/2$ , in (9.4) satisfy (2.68), (2.63) by construction.  $\square$

**9.2. Block-diagonalization at negative orders.** In this subsection we iteratively block-diagonalize the operator (9.4) (which is already block-diagonal at the orders 1 and  $1/2$ ) up to smoothing remainders. The main result of the section is the following.

**Proposition 9.5. (Block-diagonalization at lower orders).** *There exists an invertible map  $\Psi(\varphi): H_{S^{\pm}}^s(\mathbb{T}) \times H_{S^{\pm}}^s(\mathbb{T}) \rightarrow H_{S^{\pm}}^s(\mathbb{T}) \times H_{S^{\pm}}^s(\mathbb{T})$  such that (recall (9.4))*

$$\mathcal{L}_3 := \Psi \mathcal{L}_2 \Psi^{-1} = \Pi_S^{\perp} \left( \omega \cdot \partial_{\varphi} + Op^W \begin{pmatrix} d(\varphi, x, \xi) & 0 \\ 0 & d(\varphi, x, -\xi) \end{pmatrix} + \mathcal{R}^{(3)} \right), \quad (9.11)$$

where  $\mathcal{R}^{(3)} = \mathcal{R}^{(3)}(\varphi)$  is a smoothing remainder in  $\mathfrak{L}_{\rho, p}^1 \otimes \mathcal{M}_2(\mathbb{C})$  and

$$d(\varphi, x, \xi) := iV(\varphi, x)\xi + i(1 + a^{(0)}(\varphi, x))|\xi|^{\frac{1}{2}} + c^{(3)}(\varphi, x, \xi), \quad (9.12)$$

for some  $c^{(3)} \in S_1^{-\frac{1}{2}}$ . Finally, the map  $\Psi$  satisfies

$$\|\Psi^{\pm 1} u\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \|u\|_s^{\gamma, \mathcal{O}_0} + \varepsilon \|\mathcal{J}_{\delta}\|_{s+\mu}^{\gamma, \mathcal{O}_0} \|u\|_{s_0}^{\gamma, \mathcal{O}_0}. \quad (9.13)$$

The rest of the section is devoted to the proof of the proposition above. We shall prove inductively that for  $j = 0, \dots, 2\rho$ , there are

- symbols  $c_j \in S_1^{-1/2}, b_j \in S_1^{-j/2}$ ;
- smoothing operators  $\mathcal{Q}^{(j)} \in \mathfrak{L}_{\rho, p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ ,

such that the following holds true. Let  $\Psi_j^{\theta} := \Psi_j^{\theta}(\varphi)$  be the flow at time  $\theta \in [0, 1]$  of

$$\partial_{\theta} \Psi_j^{\theta}(\varphi) = \Pi_S^{\perp} Op^W(iM_j(\varphi, x, \xi)) \Pi_S^{\perp} \Psi_j^{\theta}(\varphi), \quad \Psi_j^0(\varphi) = \text{Id}, \quad (9.14)$$

with

$$M_j(\varphi, x, \xi) := \begin{pmatrix} 0 & -im_j(\varphi, x, \xi) \\ -im_j(\varphi, x, -\xi) & 0 \end{pmatrix}, \quad m_j = \frac{\chi(\xi)b_j(\varphi, x, \xi)}{2i(1 + a^{(0)}(\varphi, x))|\xi|^{\frac{1}{2}}} \in S_1^{-\frac{j+1}{2}}, \quad (9.15)$$

where  $\chi \in C^{\infty}(\mathbb{R}; \mathbb{R})$  is an even and positive cut-off function as in (2.18), and

$$\mathcal{Y}^{(j)} := \mathcal{Y}^{(j)}(\varphi) := \omega \cdot \partial_{\varphi} + Op^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)}) + \mathcal{Q}^{(j)}(\varphi) \quad (9.16)$$

where (recall (8.7), (9.3))

$$\mathcal{D}(\varphi, x, \xi) := iA_1(\varphi, x)\xi + iA_{1/2}^{(2)}(\varphi, x)|\xi|^{\frac{1}{2}}, \quad (9.17)$$

and  $B^{(j)} = B^{(j)}(\varphi, x, \xi)$  is the matrix of symbols of the form

$$B^{(j)} = \begin{pmatrix} c_j(\varphi, x, \xi) & b_j(\varphi, x, \xi) \\ b_j(\varphi, x, -\xi) & c_j(\varphi, x, -\xi) \end{pmatrix}, \quad c_j \in S_1^{-\frac{1}{2}}, \quad b_j \in S_1^{-\frac{j}{2}}. \quad (9.18)$$

Then one has that

$$\mathcal{Y}^{(j+1)} := \left( \Psi_j^{\theta} \mathcal{Y}^{(j)} \Psi_j^{-\theta} \right)_{|\theta=1} \quad (9.19)$$

has the form (9.16) with  $j + 1$  instead of  $j$  and  $B^{(j+1)}, \mathcal{Q}^{(j+1)}$ .

**Inizialization.** The operator  $\mathcal{L}_2$  in (9.4) has the form of  $\mathcal{Y}^{(j)}$  in (9.16) for  $j = 0$  with matrix of symbols  $B^{(0)} \rightsquigarrow A_0^{(2)} + A_{-1/2}^{(2)}$  and smoothing operator  $\mathcal{Q}^{(0)} \rightsquigarrow \mathcal{R}^{(2)}$ .

**Iteration.** We now argue by induction. First of all notice that the symbols of the matrix  $M_j$  defined in (9.15) have negative order for any  $j \geq 0$ . We prove that the flow map  $\Psi_j^\theta$  in (9.14) is well-posed.

**Lemma 9.6.** *Assume that  $b_j$  in (9.15) is in  $S_1^{-j/2}$ . Then  $m_j \in S_1^{-(j+1)/2}$ . Moreover the map  $\Psi_j^\theta$  has the form*

$$\Psi_j^\theta = \Pi_S^\perp \tilde{\Psi}_j^\theta \Pi_S^\perp \circ (\text{Id} + \mathcal{R}_j(\theta)), \quad \forall \theta \in [0, 1], \quad (9.20)$$

with  $\mathcal{R}_j \in \mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$  and  $\tilde{\Psi}_j^\theta$  is the flow of

$$\partial_\theta \tilde{\Psi}_j^\theta(\varphi) = \text{Op}^W(iM_j(\varphi, x, \xi)) \tilde{\Psi}_j^\theta(\varphi), \quad \tilde{\Psi}_j^0(\varphi) = \text{Id}. \quad (9.21)$$

Moreover  $\tilde{\Psi}_j^\theta(\varphi)$  has the form

$$\tilde{\Psi}_j^\theta(\varphi) = \text{Id} + \text{Op}^W(\tilde{M}_j(\theta; \varphi, x, \xi)) + \tilde{\mathcal{R}}_j(\theta) \quad (9.22)$$

where  $\tilde{M}_j \in S_1^{-\frac{j+1}{2}} \otimes \mathcal{M}_2(\mathbb{C})$  and  $\tilde{\mathcal{R}}_j \in \mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . In particular the non homogeneous terms of  $\tilde{M}_j$ ,  $\tilde{\mathcal{R}}_j$  and  $\tilde{\mathcal{R}}_j$  satisfy the bounds (7.15), (7.18) uniformly in  $\theta \in [0, 1]$ .

*Proof.* The first assertion of the matrix  $M_j$  follows by the (9.15) and by the fact that  $b_j \in S_1^{-j/2}$  and  $a^{(0)} \in S_1^0$  by Lemma 9.2. The (9.20) follows by Lemma A.4 (see also Remark A.6). The (9.22) follows by applying Lemma B.6 in the appendix of [31] and by Lemma 7.4 to pass to the Weyl quantization.  $\square$

By Lemma A.5 the conjugated operator  $\mathcal{Y}^{(j+1)}$  in (9.19) under the map  $\Psi_j^\theta$  in (9.20) is given by

$$\Psi_j^1 \mathcal{Y}^{(j)} \Psi_j^{-1} = \tilde{\Psi}_j^1(\omega \cdot \partial_\varphi \tilde{\Psi}_j^{-1}) + \tilde{\Psi}_j^1 \text{Op}^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)}) \tilde{\Psi}_j^{-1} + \tilde{\Psi}_j^1 \mathcal{Q}_j \tilde{\Psi}_j^{-1}, \quad (9.23)$$

up to finite rank remainders in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . The third summand in (9.23) is in  $\mathfrak{L}_{\rho,p}^1$  and satisfies the required estimates by Lemma 9.6 and the estimates on the smoothing remainder  $\mathcal{Q}_j$ .

Consider now the first summand in (9.23). We have that the time contribution admits the Lie expansion

$$\tilde{\Psi}_j^1(\omega \cdot \partial_\varphi \tilde{\Psi}_j^{-1}) = - \sum_{q=1}^L \frac{1}{q!} \text{ad}_{i\mathbf{A}}^{q-1} [i\omega \cdot \partial_\varphi \mathbf{A}] - \frac{1}{L!} \int_0^1 (1-\theta)^L \tilde{\Psi}^\theta \text{ad}_{i\mathbf{A}}^L [i\omega \cdot \partial_\varphi \mathbf{A}] (\tilde{\Psi}^\theta)^{-1} d\theta$$

specified for  $\mathbf{A} := \text{Op}^W(M_j(U))$  and where  $\text{ad}_{i\mathbf{A}}[X] = [i\mathbf{A}, X]$ . We recall (see (2.32)) that

$$M_j \#_\rho^W \omega \cdot \partial_\varphi M_j - \omega \cdot \partial_\varphi M_j \#_\rho^W M_j = \{M_j, \omega \cdot \partial_\varphi M_j\} \in S^{-(j+1)-1}$$

up to a symbol in  $S_1^{-(j+1)-3}$ . By Lemma 7.5 we have that  $\text{ad}_{i\text{Op}^W(M_j)} [i\text{Op}^W(\omega \cdot \partial_\varphi M_j)]$  is a pseudo differential operator with symbol in  $S_1^{-(j+1)-1} \otimes \mathcal{M}_2(\mathbb{C})$  plus a smoothing remainder in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . As a consequence we deduce (using also Lemma B.3 in [31]), for  $q \geq 2$ ,

$$\text{ad}_{i\text{Op}^W(M_j)}^q [i\text{Op}^W(\omega \cdot \partial_\varphi M_j)] = \text{Op}^W(B_q) + R_q, \quad B_q \in S_1^{-\frac{j+1}{2}(q+1)-q} \otimes \mathcal{M}_2(\mathbb{C}),$$

and  $R_q \in \mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . By taking  $L$  large enough with respect to  $\rho$ , we obtain that  $\tilde{\Psi}_j^1(\omega \cdot \partial_\varphi \tilde{\Psi}_j^{-1})$  is a pseudo differential operator with symbol in  $S_1^{-(j+1)/2} \otimes \mathcal{M}_2(\mathbb{C})$  plus a smoothing operator in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ .

We now study the space contribution, which is the second summand in (9.23). It admits the Lie expansion

$$\tilde{\Psi}_j^1 X (\tilde{\Psi}_j^1)^{-1} = X + \sum_{q=1}^L \frac{1}{q!} \text{ad}_{i\mathbf{A}}^q [X] + \frac{1}{L!} \int_0^1 (1-\theta)^L \tilde{\Psi}^\theta \text{ad}_{i\mathbf{A}}^{L+1} [X] (\tilde{\Psi}^\theta)^{-1} d\theta \quad (9.24)$$

where  $X = \text{Op}^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)})$ . We claim that

$$\begin{aligned} \tilde{\Psi}_j^1 \text{Op}^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)}) \tilde{\Psi}_j^{-1} &= \text{Op}^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)}) \\ &+ [\text{Op}^W(iM_j), \text{Op}^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)})] \end{aligned} \quad (9.25)$$

plus a pseudo differential operator with symbol in  $S_1^{-\frac{j+1}{2}} \otimes \mathcal{M}_2(\mathbb{C})$  and a smoothing operator in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . We first give the expansion of the commutator terms in (9.25) using the expression of  $\mathcal{Y}^{(j)}$  in (9.16). By Lemma 7.5 we have

$$\begin{aligned} \left[ Op^W(iM_j), Op^W(\mathcal{D}(\varphi, x, \xi)) \right] &:= Op^W \left( \begin{bmatrix} 0 & p_j(\varphi, x, \xi) \\ p_j(\varphi, x, -\xi) & 0 \end{bmatrix} \right) + Op^W(B) + R \\ p_j &:= -2im_j(\varphi, x, \xi)(1 + a^{(0)}(\varphi, x))|\xi|^{\frac{1}{2}} \end{aligned} \quad (9.26)$$

for some matrices of symbols  $B \in S_1^{-\frac{j+1}{2}} \otimes \mathcal{M}_2(\mathbb{C})$  and remainder  $R \in \mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . Moreover, since  $B^{(j)}$ , is a matrix of symbols of order  $-1/2$ , for  $j \geq 1$ , respectively 0 for  $j = 0$  (see (9.18)), we have that

$$\left[ Op^W(iM_j), Op^W(B^{(j)}) \right] \in \begin{cases} OPS^{-\frac{j+2}{2}} \otimes \mathcal{M}_2(\mathbb{C}) & \text{for } j \geq 1 \\ OPS^{-\frac{1}{2}} \otimes \mathcal{M}_2(\mathbb{C}) & \text{for } j = 0 \end{cases}$$

up to a smoothing operator in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . It follows that the off-diagonal symbols of order  $-j/2$  in (9.25) are of the form  $\begin{bmatrix} 0 & q_j(\varphi, x, \xi) \\ q_j(\varphi, x, -\xi) & 0 \end{bmatrix}$  with

$$q_j(\varphi, x, \xi) \stackrel{(9.26)}{:=} b_j(\varphi, x, \xi) - 2im_j(\varphi, x, \xi)(1 + a^{(0)}(\varphi, x))|\xi|^{\frac{1}{2}}.$$

By the definition of  $\chi$  in (2.18) we have that the operator  $Op^W((1 - \chi(\xi))b_j(\varphi, x, \xi))$  is in  $\mathfrak{L}_{\rho,p}^1$  for any  $\rho \geq 3$ . This follows by applying Lemma B.2 in [31] and by the fact that the symbol  $(1 - \chi(\xi))b_j(\varphi, x, \xi)$  is in  $S_1^{-m}$  for any  $m > 0$ . Moreover, by the choice of  $m_j(\varphi, x, \xi)$  in (9.15) we have that

$$\chi(\xi)b_j(\varphi, x, \xi) + p_j(\varphi, x, \xi) = 0.$$

This implies that  $[iOp^W(M_j), Op^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)})]$  is a pseudo differential operator with symbol in  $S_1^{-\frac{j+1}{2}} \otimes \mathcal{M}_2(\mathbb{C})$  plus a remainder in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . Now, using again Lemma 7.5, we deduce, for  $q \geq 2$ ,

$$\text{ad}_{iOp^W(M_j)}^q [Op^W(\mathcal{D}(\varphi, x, \xi) + B^{(j)})] = Op^W(\tilde{B}_q) + \tilde{R}_q, \quad \tilde{B}_q \in S_1^{-\frac{j+1}{2}q} \otimes \mathcal{M}_2(\mathbb{C}),$$

where  $\tilde{R}_q$  is in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ . Using formula (9.24) with  $L$  large enough (w.r.t.  $\rho$ ) one obtains the claim in (9.25).

We conclude that the operator  $\mathcal{Y}^{(j+1)}$  in (9.23) has the form (9.16) for some matrix of symbol  $B^{(j+1)}$  of the form (9.18) with  $j \rightsquigarrow j+1$  and smoothing operators  $\mathcal{Q}^{(j+1)}$  in  $\mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ .

We are in position to conclude the proof of Proposition 9.5.

**Proof of Proposition 9.5.** We set

$$\Psi := (\Psi_{2\rho-1}^\theta \circ \dots \circ \Psi_0^\theta)|_{\theta=1}. \quad (9.27)$$

Then, recalling (9.4), (9.3), (8.7), (9.16), (9.17), we have that  $\mathcal{L}_3 := \mathcal{Y}^{(2\rho)}$  has the form (9.11), (9.12).  $\square$

The following result regards the algebraic properties of the map  $\Psi$  and of the new operator  $\mathcal{L}_3$  in (9.11).

**Lemma 9.7.** *The operator  $\mathcal{L}_3$  is in  $\mathfrak{S}_1$  and the map  $\Psi$  is in  $\mathfrak{T}_1$  (recall Def. 2.26, 2.28).*

*Proof.* We prove inductively that  $\mathcal{Y}^{(j)} \in \mathfrak{S}_1$  for  $j \geq 0$ . Let us start by the case  $j = 0$ . For  $j = 0$  we have that  $\mathcal{Y}^{(0)} = \mathcal{L}_2 \in \mathfrak{S}_1$  by Lemma 9.4. Assume that  $\mathcal{Y}^{(j)} \in \mathfrak{S}_1$ . We show that  $\mathcal{Y}^{(j+1)} \in \mathfrak{S}_1$ . First of all the symbol  $m_j$  in (9.15) satisfies (2.63) as well as  $a^{(0)}$  and  $b_j$ . Moreover the matrix  $M_j$  has the form (2.66) with symbols satisfying (2.68) and (2.63). Hence, by Lemmata 2.25, 2.24  $\Psi_j^\theta \in \mathfrak{T}_1$ . As a consequence  $\mathcal{Y}^{(j+1)}$  is Hamiltonian and  $x$ -translation invariant and the matrix of symbols  $B^{(j+1)}$  in (9.18) satisfies (2.68) and (2.63) by construction. Thus  $\mathcal{Y}^{(j+1)}$  is in  $\mathfrak{S}_1$ . By the fact that  $\mathcal{Y}^{(j)} \in \mathfrak{S}_1$  for  $j = 0, \dots, 2\rho$ , implies that  $\Psi_j^\theta \in \mathfrak{T}_1$  for  $j = 0, \dots, 2\rho - 1$ , we deduce that  $\Psi \in \mathfrak{T}_1$ .  $\square$

## 10. REDUCTION AT THE HIGHEST ORDERS

In this section we want to eliminate the  $(\varphi, x)$ -dependence from the unbounded symbols in (9.12). In particular we shall prove that the symbols transforms into *integrable* ones according to the following definition.

**Definition 10.1. (Integrable symbols).** Let  $k = 2p$  for some  $p \in \mathbb{N}$ . We say that a  $k$ -homogeneous symbol  $a_k \in S^m$  as in (7.14) is integrable if it does not depend on  $(\varphi, x)$  and can be written as

$$a_k(\xi) := \frac{1}{(\sqrt{2\pi})^k} \sum_{j_1, \dots, j_p \in S} (\mathbf{a}_k)_{j_1, \dots, j_p}(\xi) \zeta_{j_1} \cdots \zeta_{j_p}.$$

Equivalently a symbol  $a_k \in S^m$  is integrable if  $(\mathbf{a}_k)_{j_1, \dots, j_k}^{\sigma_1, \dots, \sigma_k} \neq 0$  implies that  $\{(j_i, \sigma_i)\}_{i=1}^k \in S \times \{\pm\}$  is such that, up to permutation of the indexes,

$$\sigma_1 = \cdots = \sigma_p = +, \quad \sigma_{p+1} = \cdots = \sigma_{2p} = -, \quad \{j_1, \dots, j_p\} = \{j_{p+1}, \dots, j_{2p}\}. \quad (10.1)$$

We denote by  $\mathcal{S}_k$  the set of such  $\{(j_i, \sigma_i)\}_{i=1}^k$ . We adopt the convention to write  $\mathcal{S}_k = \emptyset$  if  $k = 2p + 1$  (recall Remark 4.2).

**Remark 10.2.** It is easy to note that, if  $a_k \in S^m$  is integrable, then it is a Fourier multiplier. We remark that the function  $\mathcal{R}_{\vec{\sigma}}(j_1, \dots, j_k)$  in (4.1) is identically zero only when it is restricted to the set  $\mathcal{S}_k$ . Indeed if  $j \in S$  then  $-j \notin S$  (see formula (1.23), (1.24)). Notice that, for generic choices of  $S$ , we have that  $\{(j_i, \sigma_i)\}_{i=1}^k \notin \mathcal{S}_k$  implies  $|\mathcal{R}_{\vec{\sigma}}(j_1, \dots, j_k)| \geq c > 0$  with  $c$  depending only on  $S$  and  $k$ .

The key result of this section is the following.

**Proposition 10.3. (Reduction to constant coefficients).** Assume (7.1). For generic choices of the set of tangential sites  $S$  there exist constants  $\mathbf{m}_i : \Omega_\varepsilon \rightarrow \mathbb{R}$ ,  $i = 1, 1/2, 0$  of the form

$$\mathbf{m}_i = \mathbf{m}_i(\omega) = \varepsilon^2 m_i(\omega) + \mathbf{m}_i^{(\geq 4)}(\omega), \quad \mathbf{m}_i^{(\geq 4)} := \sum_{k=2}^{4+2i} \varepsilon^{2k} m_i^{(2k)}(\omega) + \mathbf{r}_i(\omega), \quad (10.2)$$

where  $m_i, m_i^{(2k)}$ ,  $k \geq 2$  are integrable (and  $2k$ -homogeneous) according to Definition 10.1, and

$$m_1 := \frac{1}{\pi} \sum_{n \in S} n |n| \zeta_n, \quad (10.3)$$

$$|\mathbf{r}_i|_{\gamma, \Omega_\varepsilon} \lesssim \varepsilon^{17-2b} \gamma^{-3+2i}, \quad |\Delta_{12} \mathbf{r}_i| \lesssim_p \varepsilon (1 + \|\mathcal{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \quad i = 1, 1/2, 0 \quad (10.4)$$

such that the following holds. For all  $\omega$  belonging to the set

$$\Omega_\infty^{2\gamma} := \left\{ \omega \in \mathcal{O}_0 : |\omega \cdot \ell + \mathbf{m}_1 j| \geq \frac{\gamma}{\langle \ell \rangle^\tau}, \quad \forall \ell \in \mathbb{Z}^\nu, \quad \mathbf{v} \cdot \ell + j = 0, \quad j \in S^c, \quad (\ell, j) \neq (0, 0) \right\}, \quad (10.5)$$

there is a map  $\Phi = \Phi(\omega, \varphi) : H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$  such that (recall (9.11))

$$\mathcal{L}_6 := \Phi \mathcal{L}_3 \Phi^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + \mathbf{D} + \mathbf{Q} \right), \quad (10.6)$$

where

$$\mathbf{D} := Op^W \begin{pmatrix} i\mathbf{m}(\xi) & 0 \\ 0 & -i\mathbf{m}(-\xi) \end{pmatrix}, \quad i\mathbf{m}(\xi) := i\mathbf{m}_1 \xi + i(1 + \mathbf{m}_{\frac{1}{2}}) |\xi|^{\frac{1}{2}} + i\mathbf{m}_0 \text{sign}(\xi). \quad (10.7)$$

The remainder  $\mathbf{Q}$  has the form

$$\mathbf{Q} := Op^W \begin{pmatrix} i\mathbf{q}(\varphi, x, \xi) & 0 \\ 0 & -i\mathbf{q}(\varphi, x, -\xi) \end{pmatrix} + \mathcal{Q}(\varphi), \quad (10.8)$$

where  $\mathbf{q} \in S_4^{-1/2}$  and  $\mathcal{Q} \in \mathcal{L}_{\rho,p}^4 \otimes \mathcal{M}_2(\mathbb{C})$  (see Def. 7.3). Moreover the operator  $\mathcal{L}_6$  is in  $\mathfrak{S}_1$  and the map  $\Phi$  is in  $\mathfrak{T}_1$  and satisfies

$$\|(\Phi)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}, \quad (10.9)$$

for some  $\mu > 0$  depending on  $\nu$ .

The proof of Proposition 10.3 is divided into several steps performed in subsections 10.1, 10.2, 10.3.

**10.1. Integrability at order 1.** The aim of this subsection is to eliminate the  $(\varphi, x)$ -dependence from the symbol  $V(\varphi, x)\xi$  in (9.12). More precisely we shall prove the following result.

**Proposition 10.4.** *Let  $\rho \geq 3$ ,  $p \geq s_0$  and assume that (7.1) holds. Then there exist  $\mu = \mu(\nu)$  and  $\mathfrak{m}_1: \Omega_\varepsilon \rightarrow \mathbb{R}$  of the form (10.2) with  $i = 1$ , and  $m_1$  defined as in (10.3) such that, for all  $\omega \in \Omega_\infty^{2\gamma}$  (see (10.5)) there exists a map  $\Phi_1(\omega, \varphi): H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$  such that (recall (9.11))*

$$\mathcal{L}_4 := \Phi_1 \mathcal{L}_3 \Phi_1^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} d^{(4)}(\varphi, x, \xi) & 0 \\ 0 & \frac{0}{d^{(4)}(\varphi, x, -\xi)} \end{pmatrix} + \mathcal{R}^{(4)} \right), \quad (10.10)$$

where  $\mathcal{R}^{(4)} \in \mathfrak{L}_{\rho, p}^2 \otimes \mathcal{M}_2(\mathbb{C})$ , the symbol  $d^{(4)}(\varphi, x, \xi)$  has the form

$$d^{(4)}(\varphi, x, \xi) := \text{im}_1 \xi + i(1 + a^{(4)}(\varphi, x))|\xi|^{\frac{1}{2}} + c^{(4)}(\varphi, x, \xi), \quad (10.11)$$

with  $a^{(4)} \in S_2^0$ ,  $c^{(4)} \in S_2^{-1/2}$  (recall Def. 7.3). Moreover, for any  $\omega \in \Omega_\infty^{2\gamma}$ , we have

$$\|(\Phi_1)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \gamma^{-1} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}. \quad (10.12)$$

Finally the operator  $\mathcal{L}_4$  is in  $\mathfrak{S}_1$  and the map  $\Phi_1$  is in  $\mathfrak{T}_1$  (recall Def. 2.26, 2.28).

The rest of the section 10.1 is devoted to the proof of Proposition 10.4. It is based on two steps which we perform in subsections 10.1.1 and 10.1.2.

**10.1.1. Preliminary steps.** By Lemma 7.7 we know that the function  $V(\varphi, x)$  in (9.12) admits an expansion in  $\varepsilon$  as in (7.13) with  $k = 1$ . In this subsection we show how to reduce the terms of order less or equal to  $\varepsilon^{12}$ .

**Lemma 10.5. (Preliminary steps).** *There exists a function  $\beta^{(1)}(\varphi, x)$  of the form*

$$\beta^{(1)}(\varphi, x) = \sum_{i=1}^{12} \varepsilon^i \beta_i^{(1)}(\varphi, x), \quad (10.13)$$

where  $\beta_i^{(1)}$  is a real valued  $i$ -homogeneous symbol (as in (7.14)) independent of  $\xi \in \mathbb{R}$  for  $i = 1, \dots, 12$  satisfying

$$\|\beta^{(1)}\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \varepsilon, \quad (10.14)$$

such that the following holds. Let

$$\mathbf{T}_1^\tau := \begin{pmatrix} \mathcal{T}_{\beta^{(1)}}^\tau(\varphi) & 0 \\ 0 & \overline{\mathcal{T}_{\beta^{(1)}}^\tau(\varphi)} \end{pmatrix}, \quad \tau \in [0, 1], \quad (10.15)$$

where  $\mathcal{T}_{\beta^{(1)}}^\tau$  is the flow of (A.1) with  $f$  as in (A.2) and  $\beta \rightsquigarrow \beta^{(1)}$ , then

$$\mathcal{L}_{3,*} := \mathbf{T}_1^1 \mathcal{L}_3 (\mathbf{T}_1^1)^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} d_*^{(3)}(\varphi, x, \xi) & 0 \\ 0 & \frac{0}{d_*^{(3)}(\varphi, x, -\xi)} \end{pmatrix} + \mathcal{R}_*^{(3)} \right) \quad (10.16)$$

where  $\mathcal{R}_*^{(3)} \in \mathfrak{L}_{\rho, p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ , the symbols  $d_*^{(3)}$  has the form

$$d_*^{(3)}(\varphi, x, \xi) := iV^+(\varphi, x)\xi + i(1 + a_*^{(3)}(\varphi, x))|\xi|^{\frac{1}{2}} + c_*^{(3)}(\varphi, x, \xi), \quad (10.17)$$

$$V^+(\varphi, x) := \varepsilon^2 m_1 + \sum_{k=1}^6 \varepsilon^{2k} m_1^{(2k)} + V_{\geq 13}^+(\varphi, x), \quad (10.18)$$

$$1 + a_*^{(3)}(\varphi, x) := (1 + a^{(3)}(\varphi, x + \beta^{(1)}(x))) \left| (1 + \tilde{\beta}_y^{(1)}(1, y)) \Big|_{y=x+\beta^{(1)}(x)} \right|^{\frac{1}{2}}, \quad (10.19)$$

where  $m_1$  is in (10.3) and  $m_1^{(2k)}$  are integrable (according to Def. 10.1) and

$$\begin{aligned} \|V_{\geq 13}^+\|_s^{\gamma, \mathcal{O}_0} &\lesssim_s (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \\ \|\Delta_{12} V_{\geq 13}^+\|_p &\lesssim_p \varepsilon (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \end{aligned} \quad (10.20)$$

for some  $\mu = \mu(\nu) > 0$ , and  $c_*^{(3)}$  is in  $S_1^{-1/2}$ . The operator  $\mathcal{L}_{3,*}$  is in  $\mathfrak{S}_1$ , the map  $\mathcal{T}_{\beta^{(1)}}^\tau$  is in  $\mathfrak{T}_1$  and satisfies

$$\|(\mathcal{T}_{\beta^{(1)}}^\tau)^{\pm 1} u\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \|u\|_s^{\gamma, \mathcal{O}_0} + \varepsilon \|u\|_{s_0}^{\gamma, \mathcal{O}_0}. \quad (10.21)$$

*Proof.* We start by proving the last assertions on the map  $\mathbf{T}_1^\tau$ . If  $\beta^{(1)}$  is real valued then we have that the symbol  $b(\tau, \varphi, x)\xi$  defined in (A.2) with  $\beta \rightsquigarrow \beta^{(1)}$  in (10.13) is real and  $x$ -translation invariant. Then the flow map generated by  $\Pi_S^\perp Op^W(ib(\tau, \varphi, x)\xi)\Pi_S^\perp$  (see (A.1)) is symplectic and  $x$ -translation invariant, i.e. it belongs to  $\mathfrak{T}_1$ . Moreover, if  $\beta^{(1)}$  satisfies (10.14), by Lemma A.4, we can write

$$\mathcal{T}_{\beta^{(1)}}^\tau = \Pi_S^\perp \Phi_{\beta^{(1)}}^\tau \Pi_S^\perp \circ (\text{Id} + \mathcal{R}),$$

where  $\Phi_{\beta^{(1)}}^\tau$  is the flow of (A.5) with  $f \rightsquigarrow b\xi$  with  $b$  as in (A.2) with  $\beta \rightsquigarrow \beta^{(1)}$  and where  $\mathcal{R}$  is a finite rank operator of the form (A.14) satisfying (A.15). Therefore (see Lemma A.2) we have that the flow  $\mathcal{T}_{\beta^{(1)}}^\tau$  is well posed and satisfies estimates like (A.9)-(A.11). This implies, using also (10.14), the (10.21).

By Lemma A.5 (see item (ii)) we have that the conjugate of  $\mathcal{L}_3$  in (9.11) under the map  $\mathbf{T}_1^1$  is given by

$$\Pi_S^\perp \tilde{\mathbf{T}}_1^1 \mathcal{L}_3 (\tilde{\mathbf{T}}_1^1)^{-1} \Pi_S^\perp, \quad \tilde{\mathbf{T}}_1^\tau := \begin{pmatrix} \Phi_{\beta^{(1)}}^\tau(\varphi) & 0 \\ 0 & \overline{\Phi_{\beta^{(1)}}^\tau(\varphi)} \end{pmatrix}, \quad \tau \in [0, 1], \quad (10.22)$$

up to finite rank operators belonging to  $\mathfrak{L}_{\rho,p}^1$ . In order to study the operator in (10.22), we apply Theorem A.7 and Lemmata A.8, A.11 to the operator  $\mathcal{L}_3$  in (9.11). Then formulæ (10.16)-(10.17) follow for some  $\mathcal{R}_*^{(3)} \in \mathfrak{L}_{\rho,p}^1 \otimes \mathcal{M}_2(\mathbb{C})$ ,  $c_*^{(3)} \in S_1^{-1/2}$ . Formula (10.19) on  $a_*^{(3)}$  follows by (A.21) applied with  $w(x, \xi) \rightsquigarrow (1 + a^{(0)}(x))|\xi|^{1/2}$ . Similarly, by a direct computation, we have that the new symbol at order 1 is given by

$$V^+(\varphi, x) := -g(\varphi, x) + V(\varphi, x + \beta^{(1)}(x))(1 + \tilde{\beta}_y^{(1)}(1, y))|_{y=x+\beta^{(1)}(x)}, \quad (10.23)$$

where  $g(\varphi, x)$  is given by Lemma A.8. We need to prove that, for some suitable  $\beta^{(1)}$  the symbol in (10.23) has the form (10.18).

By formula (A.26) in Lemma A.8 we get the following Taylor expansion for  $g(\varphi, x)$ :

$$\begin{aligned} g(\varphi, x) &= \varepsilon \omega \cdot \partial_\varphi \beta_1^{(1)} + \varepsilon^2 \omega \cdot \partial_\varphi \beta_2^{(1)} - \varepsilon^2 (\beta_1^{(1)})_{x\omega} \cdot \partial_\varphi \beta_1^{(1)} + \sum_{i=3}^{12} \varepsilon^i (\omega \cdot \partial_\varphi \beta_i^{(1)} + \mathbf{g}_i) + \tilde{g}(\varphi, x) \\ &\stackrel{(5.4), (1.26)}{=} \varepsilon \bar{\omega} \cdot \partial_\varphi \beta_1^{(1)} + \varepsilon^2 \bar{\omega} \cdot \partial_\varphi \beta_2^{(1)} - \varepsilon^2 (\beta_1^{(1)})_{x\bar{\omega}} \cdot \partial_\varphi \beta_1^{(1)} + \sum_{i=3}^{12} \varepsilon^i (\bar{\omega} \cdot \partial_\varphi \beta_i^{(1)} + \tilde{\mathbf{g}}_i) + \tilde{g}(\varphi, x), \end{aligned} \quad (10.24)$$

where  $\tilde{g}(\varphi, x)$  satisfies (7.15) (actually is independent of  $\mathfrak{J}_\delta$ ) and  $\mathbf{g}_i, \tilde{\mathbf{g}}_i(\varphi, x)$  are  $i$ -homogeneous symbols of the form (7.14) whose coefficients depend, for any  $i = 1, \dots, 12$ , only on  $\beta_j^{(1)}$  with  $j < i$ .

By using formula (A.22) in Theorem A.7 and recalling that the symbol  $V$  has an expansion as in (7.13)-(7.14) for some homogeneous symbols  $\mathbf{V}_i$  (see Lemma 7.7), we can Taylor expand the second summand in (10.23). We have

$$V(\varphi, x + \beta^{(1)}(x))(1 + \tilde{\beta}_y^{(1)}(1, y))|_{y=x+\beta^{(1)}(x)} = \varepsilon \mathbf{V}_1 + \varepsilon^2 \left( \mathbf{V}_2 + (\mathbf{V}_1)_x \beta_1^{(1)} - \mathbf{V}_1 (\beta_1^{(1)})_x \right) + \sum_{i=3}^{12} \varepsilon^i \mathbf{r}_i + \tilde{V}(\varphi, x), \quad (10.25)$$

where  $\tilde{V}(\varphi, x)$  satisfies (7.15) and  $\mathbf{r}_i$  are  $i$ -homogeneous symbols of the form (7.14) whose coefficients depend, for any  $i = 1, \dots, 12$ , only on  $\beta_j^{(1)}$  with  $j < i$ . Hence, more precisely, we have

$$V^+(\varphi, x) = \sum_{i=1}^{12} \varepsilon^i \mathbf{V}_i^+(\varphi, x) + \tilde{V}^+(\varphi, x), \quad (10.26)$$

$$\mathbf{V}_1^+ = -\bar{\omega} \cdot \partial_\varphi \beta_1^{(1)} + \mathbf{V}_1, \quad (10.27)$$

$$\mathbf{V}_2^+ = -\bar{\omega} \cdot \partial_\varphi \beta_2^{(1)} + F_2, \quad F_2 := \mathbf{V}_2 + \mathbf{f}_2, \quad (10.28)$$

$$\mathbf{f}_2 = \mathbf{r}_2 - \tilde{g}_2 = (\mathbf{V}_1)_x \beta_1^{(1)} - \mathbf{V}_1 (\beta_1^{(1)})_x + (\beta_1^{(1)})_{x\bar{\omega}} \cdot \partial_\varphi \beta_1^{(1)}$$

$$\mathbf{V}_i^+ = -\bar{\omega} \cdot \partial_\varphi \beta_i^{(1)} + F_i, \quad F_i := \mathbf{V}_i + \mathbf{f}_i, \quad i = 3, \dots, 12, \quad (10.29)$$

$$\tilde{V}^+ := \tilde{V} - \tilde{g}, \quad (10.30)$$

where  $\mathbf{f}_i = \mathbf{r}_i - \tilde{g}_i$  are  $i$ -homogeneous symbols of the form (7.14) whose coefficients are sums and products of derivatives of  $\beta_j^{(1)}$ ,  $\mathbf{V}_j$  with  $j < i$ . The function  $\tilde{V}^+$  satisfies (7.15) with  $k = 1$ . We shall verify that the coefficients



$(F_i)_{j_1, \dots, j_i}^{\sigma_1 \dots \sigma_i}$ ,  $i = 2, \dots, 12$  satisfy the conditions (7.24) by showing, iteratively, that the  $\beta_j^{(1)}$  are real valued. Indeed we observe that to construct the coefficients  $\mathbf{f}_i$  we just used the equations (A.26) and (A.22) to Taylor expand the new symbol.

We now show how to choose the function  $\beta^{(1)}$  in (10.13) in such a way that the function  $V^+$  in (10.26) is *integrable* (according to Def. 10.1) up to terms of “high” degree of homogeneity.

All the functions  $\mathbf{f}_i$  can be actually computed explicitly but this is not necessary for our scope. Our aim is to find, iteratively, functions  $\beta_i^{(1)}$  in order to reduce the functions  $V_i^+$  to integrable symbols (see Def. 10.1). We shall denote by

$$(\beta_i^{(1)})_{j_1, \dots, j_i}^{\sigma_1 \dots \sigma_i}, \quad (V_i)_{j_1, \dots, j_i}^{\sigma_1 \dots \sigma_i}, \quad (F_i)_{j_1, \dots, j_i}^{\sigma_1 \dots \sigma_i}, \quad i = 1, \dots, 12$$

respectively the coefficients in the expansion (7.14) of the functions  $\beta_i^{(1)}$ ,  $V_i$  and  $F_i$  in (10.27)-(10.29). By Lemma 7.7, Lemma 7.6 applied on the symbols  $i\xi V_i$  the coefficients of the functions  $V_i$  satisfy the conditions (7.24).

**Step  $\varepsilon$ .** By (7.14) we have the expansion

$$V_1 := \sum_{\sigma=\pm, n \in S} (V_1)_n^\sigma \sqrt{\zeta_n} e^{i\sigma 1(n) \cdot \varphi} e^{inx}, \quad (V_1)_n^\sigma \in \mathbb{C}.$$

Since the function  $V_1$  is real-valued, the coefficients  $(V_1)_n^\sigma$  satisfy (7.24) with  $k = 1$ . We set

$$(\beta_1^{(1)})_n^\sigma := -\frac{\sigma(V_1)_n^\sigma}{i\sqrt{|n|}}, \quad \sigma = \pm, n \in S, \quad n \neq 0. \quad (10.31)$$

One can check that

$$V_1^+ \stackrel{(10.27)}{=} -\bar{\omega} \cdot \partial_\varphi \beta_1^{(1)} + V_1 \stackrel{(10.31)}{=} 0 \quad (10.32)$$

and that the coefficients of  $\beta_1^{(1)}$  satisfy (7.24), i.e.  $\beta_1^{(1)}$  is real valued.

**Step  $\varepsilon^2$ .** Now we want to eliminate the function in (10.28). We define, for  $n_1, n_2 \in S$ ,

$$(\beta_2^{(1)})_{n_1, n_2}^{\sigma\sigma} := \frac{-(F_2)_{n_1, n_2}^{\sigma\sigma}}{i\sigma(\omega_{n_1} + \omega_{n_2})}, \quad \sigma = \pm, \quad (\beta_2^{(1)})_{n_1, n_2}^{\sigma(-\sigma)} := \frac{-\sigma(F_2)_{n_1, n_2}^{\sigma(-\sigma)}}{i(\omega_{n_1} - \omega_{n_2})}, \quad n_1 \neq \pm n_2, \quad (10.33)$$

and  $(\beta_2^{(1)})_{0,0}^{\sigma\sigma} := 0$ ,  $(\beta_2^{(1)})_{n, \pm n}^{\sigma(-\sigma)} := 0$ . In this way we have

$$\begin{aligned} V_2^+ &= m_1 + \sum_{n \in S^+ \cup (-S^+)} ((F_2)_{n, -n}^{+-} + (F_2)_{-n, n}^{-+}) e^{i2nx} \zeta_n, \\ m_1 &:= \frac{1}{2\pi} \sum_{n \in S} ((F_2)_{n, n}^{+-} + (F_2)_{n, n}^{-+}) \zeta_n. \end{aligned} \quad (10.34)$$

By the form of  $S$  in (1.23) we note that the second sum in (10.34) is actually zero since, if  $n \in S$  then  $-n \in S^c$ . Hence  $V_2^+$  is integrable according to Definition 10.1. Again we can check explicitly that  $\beta_2^{(1)}$ , defined by (10.33), satisfies (7.24), i.e.  $\beta_2^{(1)}$  is real. This is a consequence of the fact that  $F_2$  in (10.28) is real since  $\beta_1^{(1)}$  is real. It remains to show that  $m_1$  in (10.34) has the form (10.3). To do this we compute the coefficients of the function  $F_2$  in (10.28). Using the expansion (7.23) with  $k = 1$  for  $V_1$  and  $\beta_1^{(1)}$ , and recalling (7.22), we get

$$\begin{aligned} F_2 \stackrel{(10.32), (10.28)}{=} V_2 + (V_1)_x \beta_1^{(1)} &= \sum_{j_1, j_2 \in S, \sigma = \pm} (F_2)_{j_1, j_2}^{\sigma\sigma} v_{j_1}^\sigma v_{j_2}^\sigma e^{i\sigma(j_1 + j_2)x} \\ &+ \frac{1}{2\pi} \sum_{j_1, j_2 \in S} v_{j_1}^+ v_{j_2}^- [(V_2)_{j_1, j_2}^{+-} + (V_2)_{j_2, j_1}^{-+} + (V_1)_{j_1}^+ (\beta_1^{(1)})_{j_2}^- i j_1 - (V_1)_{j_2}^- (\beta_1^{(1)})_{j_1}^+ i j_2] e^{i(j_1 - j_2)x}. \end{aligned} \quad (10.35)$$

We are not interested in computing the coefficients  $(F_2)_{j_1, j_2}^{\sigma\sigma}$ . By Lemma 7.7 we have

$$(F_2)_{n, n}^{+-} + (F_2)_{n, n}^{-+} \stackrel{(10.35), (10.31)}{=} (V_2)_{n, n}^{+-} + (V_2)_{n, n}^{-+} + (V_1)_n^+ (V_1)_n^- \frac{2n}{\sqrt{|n|}} \stackrel{(7.28)}{=} n|n| + \frac{1}{2} n^2 |n|^{-1/2} 2n |n|^{-1/2} = 2n|n|$$

which implies the (10.3).

**Step  $\varepsilon^{\geq 3}$ .** Consider now the functions (10.29). We recall that for any  $3 \leq i \leq 12$ , the functions  $V_i^+$  depends only on  $\beta_j^{(1)}, \mathbf{V}_j, \mathbf{V}_i$  with  $j < i$ . For  $i = 3, \dots, 12$  we define

$$[V_i^+] := \sum_{\mathcal{S}_i} (V_i^+)_{j_1, \dots, j_i}^{\sigma_1, \dots, \sigma_i}(\xi) \sqrt{\xi_{j_1} \dots \xi_{j_i}} e^{i(\sum_{k=1}^i \sigma_k j_k)x} e^{i(\sum_{k=1}^i \sigma_k 1(j_k)) \cdot \varphi}, \quad (10.36)$$

where  $\sum_{\mathcal{S}_i}$  denotes the sum over indexes restricted to the set  $\mathcal{S}_i$  defined in Definition 10.1. We also set, for  $\vec{\sigma} = (\sigma_1, \dots, \sigma_i), \vec{j} = (j_1, \dots, j_i)$ ,

$$(\beta_i^{(1)})_{\vec{j}}^{\vec{\sigma}} := -\frac{(\mathbf{F}_i)_{\vec{j}}^{\vec{\sigma}}}{i\mathcal{R}_{\vec{\sigma}}(\vec{j})}, \quad \{(j_k, \sigma_k)\}_{k=1}^i \notin \mathcal{S}_i, \quad (\beta_i^{(1)})_{\vec{j}}^{\vec{\sigma}} = 0, \quad \{(j_k, \sigma_k)\}_{k=1}^i \in \mathcal{S}_i, \quad (10.37)$$

and  $\mathcal{R}_{\vec{\sigma}}(\vec{j})$  is the function defined in (4.1). Again, by induction, the coefficients in (10.37) satisfy (7.24), hence the functions  $\beta_i^{(1)}$  are real valued. Moreover, by using the estimates on  $\mathcal{R}_{\vec{\sigma}}$  in Remark 10.2, the function  $\beta^{(1)}$  in (10.13) satisfies (10.14). By Remark 10.2 we have that the symbols in (10.36) are integrable according to Definition 10.1. Therefore the symbol  $V^+$  in (10.17) has the form (10.18) by setting  $m_1^{(2k)} := [V_{2k}^+]$  (see (10.36)) and  $V_{\geq 13}^+ := \tilde{V}^+$  in (10.30). The estimates (10.20) follow by (7.15) with  $k = 1$  on  $\tilde{V}^+$ . Since  $\beta^{(1)}$  is  $x$ -translation invariant (i.e. satisfies (2.63)) it is easy to check that the symbols  $V_{\geq 13}^+, a_*^{(3)}, c_*^{(3)}$  constructed above satisfy the same condition. Hence, since the map  $\mathcal{T}_{\beta^{(1)}}^\tau$  is in  $\mathfrak{T}_1$ , one has that  $\mathcal{L}_{3,*}$  belong to  $\mathfrak{S}_1$ .  $\square$

**10.1.2. Straightening theorem.** In this subsection we conclude the proof of Proposition 10.4. More precisely we eliminate the dependence on  $(\varphi, x)$  in the symbol  $V_{\geq 13}^+$  in (10.18). We first need a preliminary result.

**Lemma 10.6. (Straightening Lemma).** *For all  $\omega \in \Omega_\infty^{2\gamma}$  defined in (10.5) there exists  $\beta^{(2)}(\omega) := \beta^{(2)}(\omega, \cdot) \in S_\nu$ , (see (2.46)), satisfying*

$$\begin{aligned} \|\beta^{(2)}\|_s^{\gamma, \Omega_\infty^{2\gamma}} &\lesssim_s \gamma^{-1}(\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \\ \|\Delta_{12}\beta^{(2)}\|_p &\lesssim_p \varepsilon \gamma^{-1}(1 + \|\mathfrak{J}_\delta\|_{s+\mu}) \|i_1 - i_2\|_{p+\mu} \end{aligned} \quad (10.38)$$

for some  $\mu = \mu(\nu) > 0$ , such that the vector field on  $\mathbb{T}^{\nu+1}$

$$Y := \omega \cdot \frac{\partial}{\partial \varphi} + (c + V_{\geq 13}^+(\varphi, x)) \frac{\partial}{\partial x}, \quad c := c(\omega) := \varepsilon^2 m_1 + \sum_{k=1}^6 \varepsilon^{2k} m_1^{(2k)}, \quad (10.39)$$

transforms, under the diffeomorphism  $\Gamma: (\varphi, x) \mapsto (\varphi, y), y := x + \beta^{(2)}(\omega; \varphi, x)$  of  $\mathbb{T}^{\nu+1}$ , into

$$\Gamma_* Y = \omega \cdot \frac{\partial}{\partial \varphi} + \mathbf{m}_1 \frac{\partial}{\partial y}, \quad (10.40)$$

where  $\mathbf{m}_1 = \mathbf{m}_1(\omega)$  is defined for  $\omega \in \Omega_\varepsilon$  (see (5.20)), and  $\mathbf{r}_1 := \mathbf{m}_1 - c$  satisfies

$$\begin{aligned} |\mathbf{r}_1|^{\gamma, \Omega_\varepsilon} &\lesssim \|V_{\geq 13}^+\|_{s_0}^{\gamma, \mathcal{O}_0} (1 + \gamma^{-1} \|V_{\geq 13}^+\|_{s_0+2\tau+5}^{\gamma, \mathcal{O}_0}), \\ |\Delta_{12}\mathbf{r}_1| &\lesssim_p \varepsilon (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}. \end{aligned} \quad (10.41)$$

*Proof.* First of all we recall that, by Lemma 10.5, the function  $V_{\geq 13}^+$  is in  $S_\nu$ . Then we write  $c + V_{\geq 13}^+(\varphi, x) = f(\Theta)$  with  $\Theta = \varphi - \nu x$  for some smooth function  $f: \mathbb{T}^\nu \rightarrow \mathbb{R}$ . Then the vector field in (10.39) can be written as a non-degenerate vector field on  $\mathbb{T}^\nu$

$$X := (\omega - \nu f(\Theta)) \cdot \frac{\partial}{\partial \Theta}.$$

Theorem 3.1 in [29] (see also Proposition 3.6 and Lemma 3.7 in [31]) provides the existence of a Cantor set  $\Omega_\infty^{2\gamma} \subset \mathbb{R}^\nu$  (of the form (10.5)) such that for  $\omega \in \Omega_\infty^{2\gamma}$  the following holds true. There exists a diffeomorphism  $\Psi := \Psi(\omega): \mathbb{T}^\nu \rightarrow \mathbb{T}^\nu$  of the form  $\Theta \mapsto \tilde{\Theta} := \Theta + \alpha(\omega; \Theta)$ , with inverse  $\Psi^{-1} = \Psi^{-1}(\omega): \mathbb{T}^\nu \rightarrow \mathbb{T}^\nu, \Theta = \tilde{\Theta} + \tilde{\alpha}(\tilde{\Theta})$  such that

$$\Psi_* X = \Psi^{-1} \left( \omega - \nu f(\Theta) + (\omega - \nu f(\Theta)) \cdot \partial_\Theta \alpha \right) \cdot \frac{\partial}{\partial \tilde{\Theta}} = \left( \omega - \nu(c + \mathbf{r}_1) \right) \cdot \frac{\partial}{\partial \tilde{\Theta}} \quad (10.42)$$

for some  $\mathbf{r}_1 = \mathbf{r}_1(\omega) \in \mathbb{R}$  such that

$$|\mathbf{r}_1|^{\gamma, \mathcal{O}_0} \leq \|V_{\geq 13}^+\|_{s_1}^{\gamma, \mathcal{O}_0} \stackrel{(10.20), (7.1)}{\lesssim} \varepsilon^{17-2b} \gamma^{-1}.$$

The assumptions of this theorem are satisfied if we consider  $\mathbf{p}_1$  in (7.1) large enough (more precisely  $s_0 + \mathbf{p}_1 \geq s_1$  where  $s_1$  is an index provided by the theorem) and  $\varepsilon$  small enough (in particular such that  $\varepsilon^{13} \leq \eta_*$  where  $\eta_*$  is provided by the theorem). Moreover the following estimates hold:

$$\|\alpha\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-1} \|V_{\geq 13}^+\|_{s+2\tau+4}^{\gamma, \Omega_\infty^{2\gamma}}, \quad \|\Delta_{12}\alpha\|_p \lesssim_p \varepsilon \gamma^{-1} \|\Delta_{12}V_{\geq 13}^+\|_{p+\mu}. \quad (10.43)$$

Similar for the inverse  $\tilde{\alpha}$ . The (10.42) implies that the function  $\alpha(\Theta)$  solves the equation

$$(\omega - \mathbf{v} f(\Theta)) \cdot \partial_\Theta \alpha = \mathbf{v}(f(\Theta) - \mathbf{m}_1), \quad \mathbf{m}_1 = \mathbf{c} + \mathbf{r}_1. \quad (10.44)$$

Writing (10.44) for the components of the vector  $\alpha(\Theta)$ , we observe that  $\alpha_i(\Theta)/\tilde{j}_i$  solves the same equation for any  $i = 1, \dots, \nu$ . Hence there exists a smooth  $F : \mathbb{T}^\nu \rightarrow \mathbb{R}$  such that  $\alpha(\Theta)$  is of the form  $-\mathbf{v}F(\Theta)$  where

$$(\omega - \mathbf{v} f(\Theta)) \cdot \partial_\Theta F(\Theta) + f(\Theta) = \mathbf{m}_1. \quad (10.45)$$

Using that  $\alpha(\Theta) + \tilde{\alpha}(\Theta + \alpha(\Theta)) = 0$ , we deduce that the function  $G(\tilde{\Theta})$ , defined by the relation  $-\mathbf{v}G(\tilde{\Theta}) = \tilde{\alpha}(\tilde{\Theta})$  is such that

$$F(\Theta) + G(\Theta - \mathbf{v}F(\Theta)) = 0. \quad (10.46)$$

We consider  $\beta^{(2)} := \beta^{(2)}(\omega; \varphi, x) := G(\varphi - \mathbf{v}x)$  and  $\tilde{\beta}^{(2)} := \tilde{\beta}^{(2)}(\omega; \varphi, y) := F(\varphi - \mathbf{v}y)$ . By using (10.46) one can check that the map  $y \mapsto x = y + \tilde{\beta}^{(2)}(\omega; \varphi, y)$  is the inverse of the diffeomorphism  $x \mapsto y := x + \beta^{(2)}(\omega; \varphi, x)$  of  $\mathbb{T}^{\nu+1}$ . By (10.20), (10.43) and (7.1) we deduce the (10.38). By (10.45) we also deduce that

$$\omega \cdot \partial_\varphi \tilde{\beta}^{(2)}(\varphi, y) + (\mathbf{c} + V_{\geq 13}^+(\varphi, y))(1 + \partial_y \tilde{\beta}^{(2)}(\varphi, y)) = \mathbf{m}_1, \quad \forall (\varphi, y) \in \mathbb{T}^{\nu+1}. \quad (10.47)$$

This implies the (10.40). Finally, by  $|\mathbf{r}_1|^{\gamma, \mathcal{O}_0} \lesssim \|V_{\geq 13}^+\|_{s_0}^{\gamma, \mathcal{O}_0} (1 + \|\alpha\|_{s_0+1}^{\gamma, \Omega_\infty^{2\gamma}})$  and (10.43), we get the first bound in (10.41). Indeed by Kirszbraun Theorem we can extend the function  $\mathbf{r}(\omega)$  to  $\Omega_\varepsilon$  with the same bound on the norm. The bound on  $\Delta_{12}\mathbf{r}_1$  in (10.41) can be deduced using Lemma 3.7 in [31] and the estimates on  $\Delta_{12}V_{\geq 13}^+$  in (10.20).  $\square$

**Proof of Proposition 10.4.** We consider the map

$$\mathbf{T}_2^1 := \begin{pmatrix} \mathcal{T}_{\beta^{(2)}}^1(\varphi) & 0 \\ 0 & \mathcal{T}_{\beta^{(2)}}^1(\varphi) \end{pmatrix} \quad (10.48)$$

where  $\mathcal{T}_{\beta^{(2)}}^\tau(\varphi)$ ,  $\tau \in [0, 1]$  is the flow given by (A.1) with  $f$  as in (A.2) with  $\beta \rightsquigarrow \beta^{(2)}$  given by Lemma 10.6. First of all by Lemma A.4 we can write

$$\mathcal{T}_{\beta^{(2)}}^1 = \Pi_S^\perp \Phi_{\beta^{(2)}}^\tau \Pi_S^\perp \circ (\text{Id} + \mathcal{R})$$

where  $\Phi_{\beta^{(2)}}^\tau$  is the flow of (A.5) with  $f \rightsquigarrow b\xi$  with  $b$  as in (A.2) with  $\beta \rightsquigarrow \beta^{(2)}$  and where  $\mathcal{R}$  is a finite rank operator of the form (A.14) satisfying (A.15). Therefore (see Lemma A.2) we have that the flow  $\mathcal{T}_{\beta^{(2)}}^\tau$  is well posed and satisfies estimates like (A.9)-(A.11) This implies, using also (10.38), the (10.12). Notice that, by estimates (10.38), we could say that  $\beta^{(2)}$  is in  $S_2^0$ . Actually all the homogeneous terms in the expansion (7.13) of  $\beta^{(2)}$  are zero.

We now conjugate the operator  $\mathcal{L}_{3,*}$  in (10.16) with the map  $\mathbf{T}_2^1$ . We follow the strategy of Lemma 10.5. By Lemma A.5 we have that the conjugate of  $\mathcal{L}_{3,*}$  is given by

$$\Pi_S^\perp \tilde{\mathbf{T}}_2^1 \mathcal{L}_{3,*} (\tilde{\mathbf{T}}_2^1)^{-1} \Pi_S^\perp, \quad \tilde{\mathbf{T}}_2^\tau := \begin{pmatrix} \Phi_{\beta^{(2)}}^\tau(\varphi) & 0 \\ 0 & \Phi_{\beta^{(2)}}^\tau(\varphi) \end{pmatrix}, \quad \tau \in [0, 1],$$

up to finite rank operators belonging to  $\mathcal{L}_{\rho,p}^2 \otimes \mathcal{M}_2(\mathbb{C})$ . Then we shall apply Theorem A.7 and Lemmata A.8, A.11.

Recall that the map  $\Phi_{\beta^{(2)}}^\tau$  has the explicit expressions in (A.6). By a direct computation we have that

$$\begin{aligned} \Phi_{\beta^{(2)}}^1 \omega \cdot \partial_\varphi (\Phi_{\beta^{(2)}}^1)^{-1} &= \omega \cdot \partial_\varphi + \frac{\omega \cdot \partial_\varphi \tilde{\beta}_y^{(2)}(1, \varphi, y + \beta^{(2)}(\varphi, y))}{2(1 + \tilde{\beta}_y^{(2)}(1, \varphi, y + \beta^{(2)}(\varphi, y)))} + \omega \cdot \partial_\varphi \tilde{\beta}^{(2)}(1, \varphi, y + \beta^{(2)}(\varphi, y)) \partial_y, \\ \Phi_{\beta^{(2)}}^1 V^+(\varphi, y) \partial_y (\Phi_{\beta^{(2)}}^1)^{-1} &= \frac{V^+(\varphi, y + \beta^{(2)}(\varphi, y)) \tilde{\beta}_{yy}^{(2)}(1, \varphi, y + \beta^{(2)}(\varphi, y))}{2(1 + \tilde{\beta}_y^{(2)}(1, \varphi, y + \beta^{(2)}(\varphi, y)))} \\ &\quad + V^+(\varphi, y + \beta^{(2)}(\varphi, y)) (1 + \tilde{\beta}_y^{(2)}(1, \varphi, y + \tau \beta^{(2)}(\varphi, y))) \partial_y. \end{aligned} \quad (10.49)$$

Moreover, by Lemma 7.4 we can write (up to smoothing reminders in  $\mathfrak{L}_{\rho,p}^2$ )

$$Op^W(iV^+(\varphi)\xi) = iV^+(\varphi, x)\partial_x + \frac{1}{2}iV_x^+(\varphi, x). \quad (10.50)$$

Now we consider the unbounded part of the operator  $\mathcal{L}_{3,*}$ , namely

$$\begin{aligned} & \omega \cdot \partial_\varphi + iOp^W(V^+(\varphi, x)\xi + (1 + a_*^{(3)}(\varphi, x))|\xi|^{\frac{1}{2}} + c_*^{(3)}(\varphi, x, \xi)) \\ & \stackrel{(10.50)}{=} \omega \cdot \partial_\varphi + iV^+(\varphi, x)\partial_x + \frac{i}{2}V_x^+(\varphi, x) + Op^W(i(1 + a_*^{(3)}(\varphi, x))|\xi|^{\frac{1}{2}} + c_*^{(3)}(\varphi, x, \xi)). \end{aligned} \quad (10.51)$$

Using the explicit computation in (10.49) and Theorem A.7, we have that the conjugate of (10.51) is, up to smoothing operators in  $\mathfrak{L}_{\rho,p}^2$ ,

$$\omega \cdot \partial_\varphi + iOp^W(V^{++}(\varphi, x)\xi + (1 + a^{(4)}(\varphi, x))|\xi|^{\frac{1}{2}} + c^{(4)}(\varphi, x, \xi) + r_0((\varphi, x))),$$

where

$$\begin{aligned} V^{++}(\varphi, x) &:= \omega \cdot \partial_\varphi \tilde{\beta}^{(2)}(\varphi, x) + V^+(\varphi, x)(1 + \tilde{\beta}_y^{(2)}(\varphi, x))|_{x=y+\beta^{(2)}(\varphi, y)}, \\ r_0(\varphi, x) &:= \frac{\partial_y(V^{++}(\varphi, x)|_{x=y+\beta^{(2)}(\varphi, y)})}{2(1 + \tilde{\beta}_y^{(2)}(1, \varphi, y + \beta^{(2)}(\varphi, y))),} \\ (1 + a^{(4)}(\varphi, x)) &:= (1 + a_*^{(3)}(\varphi, x + \beta^{(2)}(x)))|(1 + \tilde{\beta}_y^{(2)}(1, y))|_{y=x+\beta^{(2)}(x)}|^{\frac{1}{2}}, \end{aligned} \quad (10.52)$$

and  $c^{(4)}(\varphi, x, \xi)$  is some symbol in  $S_2^{-1/2}$ . Recall that the function  $V^+$  constructed in Lemma 10.5 is in  $S_\nu$  (see (2.46)). Therefore, by using Lemma 10.6 (see eq. (10.47)), we deduce that  $V^{++} = \mathfrak{m}_1$  and  $\mathfrak{m}_1$  has the form (10.2) thanks to (10.39), (10.44). As a consequence  $r_0(\varphi, x)$  is identically zero. By item (ii) of Theorem A.7 we have that  $a^{(4)} \in S_2^0$ . The discussion above implies that the conjugate of  $\mathcal{L}_{3,*}$  (10.16) has the form (10.10) with  $\mathcal{R}^{(4)} \in \mathfrak{L}_{\rho,p}^2 \otimes \mathcal{M}_2(\mathbb{C})$ . The bounds in (10.4) with  $i = 1$  on the constant  $\mathfrak{r}_1$  follows by (10.41), (10.20) and using (7.1). To conclude we set (recall (10.15), (10.48))

$$\Phi_1 := \mathbf{T}_2^1 \circ \mathbf{T}_1^1. \quad (10.53)$$

Lemma A.2 with  $\beta \rightsquigarrow \beta^{(2)}$ , estimate (10.12) and (10.21) imply the (10.12). To conclude the proof of Proposition 10.4 it remains to show that  $\mathcal{L}_4 \in \mathfrak{S}_1$  and  $\Phi_1 \in \mathfrak{T}_1$ .

First of all recall that the functions  $\beta^{(i)}$ ,  $i = 1, 2$  are real and belongs to  $S_\nu$  (see (2.46)). The map  $\Phi_1 \in \mathfrak{T}_1$  because  $\mathbf{T}_1^1, \mathbf{T}_2^1 \in \mathfrak{T}_1$ .

Hence  $\mathcal{L}_4$  defined in (10.10) is Hamiltonian and  $x$ -translation invariant. The symbol  $d^{(4)}(\varphi, x, \xi)$  satisfies (2.66), (2.68), (2.63) by construction. Hence, by Lemmata 2.24, 2.25, we deduce that  $\mathcal{L}_4$  is in  $\mathfrak{S}_1$ .  $\square$

**10.2. Integrability at order 1/2.** The aim of this section is to eliminate the  $(\varphi, x)$  dependence from the symbol  $a^{(4)}$  in (10.11) appearing in the operator  $\mathcal{L}_4$  in (10.10).

**Proposition 10.7.** *Let  $\rho \geq 3$ ,  $p \geq s_0$  and assume that (7.1) holds. Then there exist  $\mu = \mu(\nu)$  and  $\mathfrak{m}_{1/2}: \Omega_\varepsilon \rightarrow \mathbb{R}$  of the form (10.2) with  $i = 1/2$  such that, for all  $\omega \in \Omega_\infty^{2\gamma}$  (see (10.5)), there exists a map  $\Phi_2(\omega): H_{S^\pm}^{s_\pm}(\mathbb{T}) \times H_{S^\pm}^s(\mathbb{T}) \rightarrow H_{S^\pm}^s(\mathbb{T}) \times H_{S^\pm}^s(\mathbb{T})$  such that (recall (10.10))*

$$\mathcal{L}_5 := \Phi_2 \mathcal{L}_4 \Phi_2^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} d^{(5)}(\varphi, x, \xi) & 0 \\ 0 & d^{(5)}(\varphi, x, -\xi) \end{pmatrix} + \mathcal{R}^{(5)} \right) \quad (10.54)$$

where  $\mathcal{R}^{(5)} \in \mathfrak{L}_{\rho,p}^3 \otimes \mathcal{M}_2(\mathbb{C})$ , the symbol  $d^{(5)}(\varphi, x, \xi)$  has the form (recall (9.12))

$$d^{(5)}(\varphi, x, \xi) := \mathfrak{im}_1 \xi + i(1 + \mathfrak{m}_{\frac{1}{2}})|\xi|^{\frac{1}{2}} + ib^{(5)}(\varphi, x)\text{sign}(\xi) + c^{(5)}(\varphi, x, \xi), \quad (10.55)$$

with  $b^{(5)} \in S_3^0$  independent of  $\xi$ ,  $c^{(5)} \in S_3^{-\frac{1}{2}}$ . Moreover

$$\|(\Phi_2)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \gamma^{-2} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}. \quad (10.56)$$

Finally the operator  $\mathcal{L}_5$  is in  $\mathfrak{S}_1$  and the map  $\Phi_2$  is in  $\mathfrak{T}_1$  (recall Def. 2.26, 2.28).

Following the strategy used in section 10.1 we divide the proof of Proposition 10.7 into two steps.

## 10.2.1. Preliminary steps.

**Lemma 10.8. (Preliminary steps).** *There exists a function  $\beta^{(3)}(\varphi, x)$  of the form*

$$\beta^{(3)}(\varphi, x) = \sum_{i=1}^{10} \varepsilon^i \beta_i^{(3)}(\varphi, x), \quad (10.57)$$

where  $\beta_i^{(3)}$  is a real valued  $i$ -homogeneous symbol as in (7.14) for  $i = 1, \dots, 10$  independent of  $\xi \in \mathbb{R}$  satisfying

$$\|\beta^{(3)}\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \varepsilon, \quad (10.58)$$

such that the following holds. If

$$\mathbf{T}_3^\tau := \begin{pmatrix} \mathcal{T}_{\beta^{(3)}}^\tau(\varphi) & 0 \\ 0 & \mathcal{T}_{\beta^{(3)}}^\tau(\varphi) \end{pmatrix}, \quad \tau \in [0, 1], \quad (10.59)$$

where  $\mathcal{T}_{\beta^{(3)}}^\tau$  is the flow of (A.1) with generator as in (A.3) with  $\beta \rightsquigarrow \beta^{(3)}$ , then

$$\mathcal{L}_{4,*} := \mathbf{T}_3^1 \mathcal{L}_4 (\mathbf{T}_3^1)^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} d_*^{(4)}(\varphi, x, \xi) & 0 \\ 0 & d_*^{(4)}(\varphi, x, -\xi) \end{pmatrix} + \mathcal{R}_*^{(4)} \right) \quad (10.60)$$

where  $\mathcal{R}_*^{(4)} \in \mathfrak{L}_{\rho,p}^2 \otimes \mathcal{M}_2(\mathbb{C})$  and

$$d_*^{(4)}(\varphi, x, \xi) := \text{im}_1 \xi + i(1 + a_*^{(4)}(\varphi, x)) |\xi|^{\frac{1}{2}} + i b_*^{(4)}(\varphi, x) \text{sign}(\xi) + c_*^{(4)}(\varphi, x, \xi), \quad (10.61)$$

$$a_*^{(4)}(\varphi, x) := \varepsilon^2 m_{\frac{1}{2}} + \sum_{k=2}^5 \varepsilon^{2k} m_{\frac{1}{2}}^{(2k)} + a_{\geq 11}^{(4,*)}(\varphi, x), \quad (10.62)$$

where  $m_1$  is in (10.2),  $m_{1/2}$ ,  $m_1^{(2k)}$  are respectively 2,  $2k$ -homogeneous and integrable (according to Def. 10.1). Moreover one has  $c_*^{(4)} \in S_2^{-1/2}$  and  $b_*^{(4)} \in S_2^0$  (recall Def. 7.3) independent of  $\xi \in \mathbb{R}$  and

$$\|a_{\geq 11}^{(4,*)}\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-1} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad (10.63)$$

$$\|\Delta_{12} a_{\geq 11}^{(4,*)}\|_p \lesssim_p \gamma^{-1} \varepsilon (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}. \quad (10.64)$$

Finally, the operator  $\mathcal{L}_{4,*}$  is in  $\mathfrak{S}_1$ , the map  $\mathcal{T}_{\beta^{(3)}}^\tau$  is in  $\mathfrak{S}_1$  and satisfies

$$\|(\mathcal{T}_{\beta^{(3)}}^\tau)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \varepsilon \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}. \quad (10.65)$$

*Proof.* By Lemma A.3 we have that, if  $\beta^{(3)}$  is real valued and satisfies (10.58), then the flow  $\mathcal{T}_{\beta^{(3)}}^\tau$  is well posed and satisfies bounds like (A.9)-(A.11). By Lemma A.4 we have

$$\mathcal{T}_{\beta^{(3)}}^1 = \Pi_S^\perp \Phi_{\beta^{(3)}}^\tau \Pi_S^\perp \circ (\text{Id} + \mathcal{R})$$

where  $\Phi_{\beta^{(3)}}^\tau$  is the flow of (A.5) with  $f$  as in (A.3) with  $\beta \rightsquigarrow \beta^{(3)}$  and where  $\mathcal{R}$  is a finite rank operator of the form (A.14) satisfying (A.15). By Lemma A.5 we have that the conjugate of  $\mathcal{L}_4$  in (10.10) under the map  $\mathbf{T}_3^1$  is given by

$$\Pi_S^\perp \tilde{\mathbf{T}}_3^1 \mathcal{L}_4 (\tilde{\mathbf{T}}_3^1)^{-1} \Pi_S^\perp, \quad \tilde{\mathbf{T}}_3^\tau := \begin{pmatrix} \Phi_{\beta^{(3)}}^\tau(\varphi) & 0 \\ 0 & \Phi_{\beta^{(3)}}^\tau(\varphi) \end{pmatrix}, \quad \tau \in [0, 1],$$

up to finite rank operators belonging to  $\mathfrak{L}_{\rho,p}^2 \otimes \mathcal{M}_2(\mathbb{C})$ . We apply Lemmata A.9, A.10. We start by studying the time contribution. By Lemma A.10 (see in particular item (ii)) we have that (see (A.35), (A.39))

$$\begin{aligned} \Phi_{\beta^{(3)}}^1 \omega \cdot \partial_\varphi (\Phi_{\beta^{(3)}}^1)^{-1} &= \omega \cdot \partial_\varphi - iOp^W \left( \omega \cdot \partial_\varphi \beta^{(3)} |\xi|^{\frac{1}{2}} + \frac{1}{2} \{ \beta^{(3)} |\xi|^{\frac{1}{2}}, \omega \cdot \partial_\varphi \beta^{(3)} |\xi|^{\frac{1}{2}} \} + r_1 \right) \\ &= \omega \cdot \partial_\varphi - iOp^W \left( \omega \cdot \partial_\varphi \beta^{(3)} |\xi|^{\frac{1}{2}} + \frac{1}{4} (\beta^{(3)} \omega \cdot \partial_\varphi \beta_x^{(3)} - \beta_x^{(3)} \omega \cdot \partial_\varphi \beta^{(3)}) \text{sign}(\xi) + r_1 \right), \end{aligned} \quad (10.66)$$

up to smoothing remainder  $\mathcal{R}_1$  in the class  $\mathfrak{L}_{\rho,p}^2$  and where  $r_1$  is some symbol in  $S_2^{-3/2}$ . We remark that the symbol of the operator above is not singular at  $\xi = 0$ . Indeed, recalling the notation (2.19), we are writing  $\text{sign}(\xi)$  instead of  $\chi(\xi) \text{sign}(\xi)$  (with  $\chi$  as in (2.18)). In the following we will use systematically such notation.

We now consider the contribution coming from the conjugation of the spatial operator by using Lemma A.9 applied to the pseudo differential operator with symbol in (10.11). First we notice that (see (A.30), (A.33))

$$\begin{aligned} \Phi_{\beta^{(3)}}^1 Op^W(\mathfrak{m}_1 \xi)(\Phi_{\beta^{(3)}}^1)^{-1} &= iOp^W(\mathfrak{m}_1 \xi + \{\beta^{(3)}|\xi|^{\frac{1}{2}}, \mathfrak{m}_1 \xi\}) \\ &\quad + iOp^W\left(\frac{1}{2}\{\beta^{(3)}|\xi|^{\frac{1}{2}}, \{\beta^{(3)}|\xi|^{\frac{1}{2}}, \mathfrak{m}_1 \xi\}\} + r_2\right) \\ &= iOp^W\left(\mathfrak{m}_1 \xi - \mathfrak{m}_1 \beta_x^{(3)}|\xi|^{\frac{1}{2}} - \frac{\mathfrak{m}_1 \text{sign}(\xi)}{4}(\beta^{(3)}\beta_{xx}^{(3)} - (\beta_x^{(3)})^2) + r_2\right) \end{aligned} \quad (10.67)$$

up to smoothing remainder  $\mathcal{R}_2 \in \mathfrak{L}_{\rho,p}^2$  and  $r_2 \in S_2^{-1/2}$ . Moreover we have

$$\begin{aligned} \Phi_{\beta^{(3)}}^1 Op^W(i(1+a^{(4)})|\xi|^{\frac{1}{2}})(\Phi_{\beta^{(3)}}^1)^{-1} &= iOp^W\left((1+a^{(4)})|\xi|^{\frac{1}{2}} + \{\beta^{(3)}|\xi|^{\frac{1}{2}}, (1+a^{(4)})|\xi|^{\frac{1}{2}}\} + r_3\right) \\ &= iOp^W\left((1+a^{(4)})|\xi|^{\frac{1}{2}} + \frac{\text{sign}(\xi)}{2}(\beta^{(3)}a_x^{(4)} - (1+a^{(4)})\beta_x^{(3)}) + r_3\right) \end{aligned} \quad (10.68)$$

up to smoothing remainder  $\mathcal{R}_3 \in \mathfrak{L}_{\rho,p}^2$  and where  $r_3 \in S_2^{-1/2}$ . Similarly (using Lemmata A.9, A.12) we conclude

$$\Phi_{\beta^{(3)}}^1 [Op^W(c^{(4)}) + \mathcal{R}^{(4)}](\Phi_{\beta^{(3)}}^1)^{-1} = Op^W(c^{(4)} + r_4) \quad (10.69)$$

up to smoothing remainder  $\mathcal{R}_4 \in \mathfrak{L}_{\rho,p}^2$  and where  $r_4 \in S_2^{-1}$ . By collecting the (10.66)-(10.69) we conclude that the conjugate of  $\mathcal{L}_4$  in (10.10) under the map in (10.59) has the form (10.60) with symbol  $d_*^{(4)}(\varphi, x, \xi)$  as in (10.61) where  $b_*^{(4)} \in S_2^0$ ,  $c_*^{(4)} \in S_2^{-1/2}$  and

$$\begin{aligned} a_*^{(4)} &= a^{(4)} - \omega \cdot \partial_\varphi \beta^{(3)} - \mathfrak{m}_1 \beta_x^{(3)}, \\ b_*^{(4)} &= -\frac{1}{4}(\beta^{(3)}\omega \cdot \partial_\varphi \beta_x^{(3)} - \beta_x^{(3)}\omega \cdot \partial_\varphi \beta^{(3)}) - \frac{\mathfrak{m}_1}{4}(\beta^{(3)}\beta_{xx}^{(3)} - (\beta_x^{(3)})^2) + \frac{1}{2}(\beta^{(3)}a_x^{(4)} - (1+a^{(4)})\beta_x^{(3)}). \end{aligned}$$

We shall prove that it is possible to choose  $\beta^{(3)}$  in (10.57) in such a way that the symbol  $a_*^{(4)}$  has the form (10.62). Recall that  $a^{(4)}$  is a symbol (independent of  $\xi \in \mathbb{R}$ ) in the class  $S_2^0$  and hence it admits an expansion as (7.13) with  $k=2$  for some  $i$ -homogeneous symbols  $a_i^{(4)}$  and a non-homogeneous symbol  $\tilde{a}^{(4)}$ . Therefore, recalling the expansion of  $\mathfrak{m}_1$  in (10.2) and of  $\omega$  in (5.4), we have

$$\begin{aligned} a_*^{(4)} &= \sum_{i=1}^{10} \varepsilon^i a_i^{(4)} - \sum_{i=1}^{10} \varepsilon^i \bar{\omega} \cdot \partial_\varphi \beta_i^{(3)} - \sum_{i=1}^{10} \varepsilon^i \left( \varepsilon^2 \mathbb{A} \xi + \sum_{k=2}^6 \varepsilon^{2k} \mathfrak{b}_k(\xi, \dots, \xi) \right) \cdot \partial_\varphi \beta_i^{(3)} \\ &\quad - \left( \varepsilon^2 m_1 + \sum_{k=2}^6 \varepsilon^{2k} m_1^{2k} + \mathfrak{r} \right) \sum_{i=1}^{10} \varepsilon^i (\beta_i^{(3)})_x + \tilde{a}^{(4)}. \end{aligned} \quad (10.70)$$

This implies that the symbol  $a_*^{(4)}$  has an expansion as in (7.13) with  $k=2$  for some non-homogeneous symbol  $\tilde{a}_*^{(4)}$  satisfying (7.15), with  $k=2$ , and some  $i$ -homogeneous symbols  $a_{i,*}^{(4)}$ ,  $i=1, \dots, 10$ . In particular we have

$$a_{1,*}^{(4)} := a_1^{(4)} - \bar{\omega} \cdot \partial_\varphi \beta_1^{(3)}, \quad (10.71)$$

$$a_{2,*}^{(4)} := a_2^{(4)} - \bar{\omega} \cdot \partial_\varphi \beta_2^{(3)}, \quad (10.72)$$

$$a_{3,*}^{(4)} := \mathbb{F}_3 - \bar{\omega} \cdot \partial_\varphi \beta_3^{(3)}, \quad \mathbb{F}_3 := a_3^{(4)} + \mathfrak{f}_3, \quad \mathfrak{f}_3 := -\mathbb{A} \xi \cdot \partial_\varphi \beta_1^{(3)} - m_1(\beta_1^{(3)})_x, \quad (10.73)$$

$$a_{j,*}^{(4)} := \mathbb{F}_j - \bar{\omega} \cdot \partial_\varphi \beta_j^{(3)}, \quad \mathbb{F}_j := a_j^{(4)} + \mathfrak{f}_j, \quad 4 \leq j \leq 10, \quad (10.74)$$

where  $\mathfrak{f}_j$  are some  $j$ -homogeneous symbols of the form (7.14) independent of  $\xi \in \mathbb{R}$  and depending only on  $\beta_k^{(3)}$  with  $1 \leq k \leq j-1$ .

**Step  $\varepsilon$ .** By (7.14) we have the expansion

$$a_1^{(4)} := \sum_{\sigma=\pm, n \in S} (\mathfrak{a}_1^{(4)})_\sigma^n \sqrt{\zeta_n} e^{i\sigma 1(n) \cdot \varphi} e^{inx}, \quad (\mathfrak{a}_1^{(4)})_\sigma^n \in \mathbb{C}.$$



Since the function  $a_1^{(4)}$  is real-valued, the coefficients  $(a_1^{(4)})_n^\sigma$  satisfy (7.24) with  $k = 1$ . We set

$$(\beta_1^{(3)})_n^\sigma := \frac{-\sigma(\mathbf{V}_1)_n^\sigma}{i\sqrt{|n|}}, \quad \sigma = \pm, n \in S, \quad n \neq 0.$$

One can check that (see (10.71))  $a_{*,1}^{(4)} \equiv 0$  and that the function  $\beta_1^{(3)}$  satisfies (7.24), i.e. it is real valued. In particular the estimate (10.58) holds.

**Step  $\varepsilon^2$ .** Now we want to eliminate the function in (10.72). We define, for  $n_1, n_2 \in S$ ,

$$\begin{aligned} (\beta_2^{(3)})_{n_1, n_2}^{\sigma\sigma} &:= \frac{-(a_2^{(4)})_{n_1, n_2}^{\sigma\sigma}}{i\sigma(\omega_{n_1} + \omega_{n_2})}, \quad \sigma = \pm, & (\beta_2^{(3)})_{n_1, n_2}^{\sigma(-\sigma)} &:= \frac{-\sigma(a_2^{(4)})_{n_1, n_2}^{\sigma(-\sigma)}}{i(\omega_{n_1} - \omega_{n_2})}, \quad n_1 \neq \pm n_2, \\ (\beta_2^{(3)})_{0,0}^{\sigma\sigma} &:= 0, & (\beta_2^{(3)})_{n, \pm n}^{\sigma(-\sigma)} &:= 0, \end{aligned} \quad (10.75)$$

where  $(a_2^{(4)})_{n_1, n_2}^{\sigma_1 \sigma_2}$  are the coefficients of  $a_2^{(4)}$  in the corresponding expansion (7.14). Notice that, by the form of  $S$  in (1.23), the condition  $n_1 \neq \pm n_2$  in (10.75) reduces to  $n_1 \neq n_2$ . Using that  $a_2^{(4)}$  is real valued one has that  $(\beta_2^{(3)})_{n_1, n_2}^{\sigma_1 \sigma_2}$  satisfy (7.24), hence  $\beta_2^{(3)}$  is real valued. In particular it satisfies the estimate (10.58). Moreover, recalling Definition 10.1, we define  $m_{\frac{1}{2}}$  as the restriction of  $a_2^{(4)}$  to the set  $S_2$ , i.e.

$$m_{\frac{1}{2}} := (a_2^{(4)})|_{S_2} := \frac{1}{2\pi} \sum_{n \in S} ((a_2^{(4)})_{n,n}^{+-} + (a_2^{(4)})_{n,n}^{-+}) \zeta_n,$$

which is integrable according to Definition 10.1 (recall Remark 10.2). By (10.75), one can check that  $a_{2,*}^{(4)} \equiv m_{1/2}$ .

**Step  $\varepsilon \geq 3$ .** We prove inductively that, for  $j \geq 3$ , there are functions  $\beta_j^{(3)}$  of the form (7.14) with  $i = j$  and coefficients independent of  $\xi \in \mathbb{R}$  which are real valued and satisfying (10.58) such that (recall (10.73), (10.74))

$$a_{j,*}^{(4)} \equiv \begin{cases} m_{\frac{1}{2}}^{(j)} & j \text{ even,} \\ 0 & j \text{ odd,} \end{cases} \quad (10.76)$$

with

$$m_{\frac{1}{2}}^{(j)} := \sum_{S_j} (\mathbf{F}_j)_{n_1, \dots, n_j}^{\sigma_1, \dots, \sigma_j} \sqrt{\zeta_{j_1} \cdots \zeta_{j_j}} e^{i(\sigma_1 j_1 + \dots + \sigma_j j_j)x} e^{i(\sigma_1 1(j_1) + \dots + \sigma_j 1(j_j)) \cdot \varphi}$$

where  $\sum_{S_j}$  denotes the sum restricted over indexes in  $S_j$ . Notice that  $m_{1/2}^{(j)}$  is integrable according to Definition 10.1 (recall Remark 10.2).

Assume by induction that the assertion above is true for  $3 \leq k \leq j$ . By construction (see the (10.70)) the function  $\mathbf{F}_{j+1}$  depends only on  $\beta_k^{(3)}$ ,  $k \leq j$ , hence, by the inductive hypothesis it is real valued. We set, for  $\vec{\sigma} = (\sigma_1, \dots, \sigma_{j+1})$ ,  $\vec{j} = (j_1, \dots, j_{j+1})$ ,

$$(\beta_{j+1}^{(3)})_{\vec{j}}^{\vec{\sigma}} := \frac{-(\mathbf{F}_{j+1})_{\vec{j}}^{\vec{\sigma}}}{i\mathcal{R}_{\vec{\sigma}}(\vec{j})}, \quad \{(j_k, \sigma_k)\}_{k=1}^{j+1} \notin S_{j+1}, \quad (\beta_{j+1}^{(3)})_{\vec{j}}^{\vec{\sigma}} = 0, \quad \{(j_k, \sigma_k)\}_{k=1}^{j+1} \in S_{j+1}, \quad (10.77)$$

and  $\mathcal{R}_{\vec{\sigma}}(\vec{j})$  is the function defined in (4.1). Since  $\mathbf{F}_{j+1}$  is real, then the coefficients  $(\mathbf{F}_{j+1})_{n_1, \dots, n_{j+1}}^{\sigma_1, \dots, \sigma_{j+1}}$  satisfies (7.24). By an explicit computation one can check that also  $(\beta_{j+1}^{(3)})_{\vec{j}}^{\vec{\sigma}}$  satisfies (7.24), so that  $\beta_{j+1}^{(3)}$  is real valued. The estimate (10.58) follows by using the bound on  $\mathcal{R}_{\vec{\sigma}}$  in Remark 10.2. The (10.65) follows by Lemma A.3. Using (10.77) one verifies the (10.76). By the construction above, one can check that the operator  $\mathcal{L}_{4,*}$  is in  $\mathfrak{S}_1$  and the map  $\mathcal{T}_{\beta^{(3)}}$  is in  $\mathfrak{T}_1$  (i.e. is symplectic and  $x$ -translation invariant).  $\square$

**10.2.2. Reduction to constant coefficients at order 1/2.** The aim of this section is to eliminate the  $(\varphi, x)$ -dependence in the symbol  $a_*^{(4)}$  in (10.62) appearing in the operator  $\mathcal{L}_{4,*}$  in (10.60). Then we conclude the proof of Proposition 10.10. We first need a preliminary result.

**Lemma 10.9.** *For any  $\omega \in \Omega_\infty^{2\gamma}$  (see (10.5)) there exists a real valued function  $\beta^{(4)}(\varphi, x) \in S_\gamma$  (see (2.46)) such that*

$$a_*^{(4)}(\varphi, x) - (\omega \cdot \partial_\varphi + \mathbf{m}_1 \partial_x) \beta^{(4)}(\varphi, x) = \mathbf{m}_{\frac{1}{2}}, \quad (10.78)$$

where  $a_*^{(4)}$  is in (10.62), the constant  $\mathfrak{m}_{\frac{1}{2}}$  is in (10.2), with  $i = 1/2$  and

$$\mathfrak{r}_{\frac{1}{2}} := \int_{\mathbb{T}^\nu} a_{\geq 11}^{(4,*)}(\varphi, x) d\varphi. \quad (10.79)$$

The function  $\beta^{(4)}(\varphi, x)$  satisfies

$$\|\beta^{(4)}\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-2} (\varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad (10.80)$$

$$\|\Delta_{12}\beta^{(4)}\|_p \lesssim_p \varepsilon \gamma^{-2} (1 + \|\mathcal{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \quad (10.81)$$

for some  $\mu > 0$  depending on  $\nu$ . Moreover the map  $\mathcal{T}_{\beta^{(4)}}^\tau(\varphi)$ ,  $\tau \in [0, 1]$ , defined as the flow of (A.1) with generator as in (A.3) with  $\beta \rightsquigarrow \beta^{(4)}$  is well posed and satisfies

$$\|(\mathcal{T}_{\beta^{(4)}}^\tau)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \gamma^{-2} (\varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}, \quad (10.82)$$

uniformly in  $\tau \in [0, 1]$ . Finally  $\mathcal{T}_{\beta^{(4)}}^\tau$  belongs to  $\mathfrak{T}_1$ .

*Proof.* Passing to the Fourier basis (recall (2.1)) we have that equation (10.78) reads (recall (10.62))

$$(a_{\geq 11}^{(4,*)})_{lj} - i(\omega \cdot l + \mathfrak{m}_1 j)(\beta^{(4)})_{lj} = 0, \quad \forall l \in \mathbb{Z}^\nu, j \in \mathbb{Z} \setminus \{0\}, \quad (l, j) \neq (0, 0). \quad (10.83)$$

Using the diophantine condition in (10.5) one can prove

$$\|\beta^{(4)}\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-1} \|a_{\geq 11}^{(4,*)}\|_{s+2\tau+1}^{\gamma, \Omega_\infty^{2\gamma}} \stackrel{(10.63)}{\lesssim_s} \gamma^{-2} (\varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0})$$

which is the (10.80). The (10.81) follows similarly using (10.64). Since  $a_{\geq 11}^{(4,*)}$  is in  $S_\nu$  (recall that  $\mathcal{L}_{4,*} \in \mathfrak{S}_1$ ) then, by (10.83), we deduce that also  $\beta^{(4)} \in S_\nu$ . By Lemma A.4 we have

$$\mathcal{T}_{\beta^{(4)}}^1 = \Pi_S^\perp \Phi_{\beta^{(4)}}^\tau \Pi_S^\perp \circ (\text{Id} + \mathcal{R}) \quad (10.84)$$

where  $\Phi_{\beta^{(4)}}^\tau$  is the flow of (A.5) with  $f$  as in (A.3) with  $\beta \rightsquigarrow \beta^{(4)}$  and where  $\mathcal{R}$  is a finite rank operator of the form (A.14) satisfying (A.15). The estimate (10.82) follows by Lemma A.3 and estimates (A.9)-(A.11). Moreover the map  $\mathcal{T}_{\beta^{(4)}}^\tau(\varphi)$  is in  $\mathfrak{T}_1$ .  $\square$

**Proof of Proposition 10.7.** We consider the map

$$\mathbf{T}_4 := \begin{pmatrix} \mathcal{T}_{\beta^{(4)}}^1(\varphi) & 0 \\ 0 & \mathcal{T}_{\beta^{(4)}}^1(\varphi) \end{pmatrix} \quad (10.85)$$

where  $\mathcal{T}_{\beta^{(4)}}^\tau(\varphi)$ ,  $\tau \in [0, 1]$  is given by Lemma 10.9. We now conjugate the operator  $\mathcal{L}_{4,*}$  in (10.60) with the map  $\mathbf{T}_4$ . First of all we remark that the symbol

$$f(\varphi, x, \xi) := \beta^{(4)}(\varphi, x) |\xi|^{\frac{1}{2}} \quad (10.86)$$

belongs to the class  $S_3^{1/2}$  (see Def. 7.3). This is true because estimates (10.80), (10.81) imply the (7.15) with  $k = 3$ . Moreover (recall (10.84)), by Lemma A.5 we have that the conjugate of  $\mathcal{L}_{4,*}$  under the map  $\mathbf{T}_4$  is given by

$$\Pi_S^\perp \tilde{\mathbf{T}}_4 \mathcal{L}_{4,*} (\tilde{\mathbf{T}}_4^{-1})^{-1} \Pi_S^\perp, \quad \tilde{\mathbf{T}}_4 := \begin{pmatrix} \Phi_{\beta^{(4)}}^\tau(\varphi) & 0 \\ 0 & \Phi_{\beta^{(4)}}^\tau(\varphi) \end{pmatrix}, \quad \tau \in [0, 1],$$

up to finite rank operators belonging to  $\mathfrak{L}_{\rho,p}^3 \otimes \mathcal{M}_2(\mathbb{C})$ . Hence we apply Lemmata A.9, A.10 with  $f(\varphi, x, \xi)$  as in (10.86). By explicit computations using (A.30), (A.33), (A.35), (A.39) (as done in (10.66)-(10.69)) we obtain that the operator

$$\mathcal{L}_5 := \mathbf{T}_4 \mathcal{L}_{4,*} (\mathbf{T}_4^{-1})^{-1}$$

has the form (10.54) with some remainder  $\mathcal{R}^{(5)} \in \mathfrak{L}_{\rho,p}^3 \otimes \mathcal{M}_2(\mathbb{C})$ , some symbol  $d^{(5)}(\varphi, x, \xi)$  of the form

$$d^{(5)}(\varphi, x, \xi) := \text{im}_1 \xi + i(1 + A(\varphi, x)) |\xi|^{\frac{1}{2}} + i b^{(5)}(\varphi, x) \text{sign}(\xi) + c^{(5)}(\varphi, x, \xi),$$

where

$$A = a_*^{(4)} - (\omega \cdot \partial_\varphi + \mathfrak{m}_1 \partial_x) \beta^{(4)},$$

$$b^{(5)} = b_*^{(4)} - \frac{1}{4} (\beta^{(4)} \omega \cdot \partial_\varphi \beta_x^{(4)} - \beta_x^{(4)} \omega \cdot \partial_\varphi \beta^{(4)}) - \frac{\mathfrak{m}_1}{4} (\beta^{(4)} \beta_{xx}^{(4)} - (\beta_x^{(4)})^2) + \frac{1}{2} (\beta^{(4)} (a_*^{(4)})_x - (1 + a_*^{(4)}) \beta_x^{(4)})$$

and  $c^{(5)} \in S_3^{-1/2}$ . By equation (10.78) in Lemma 10.9 we actually have that  $A(\varphi, x) \equiv \mathbf{m}_{1/2}$  in (10.2),  $i = 1/2$ . The constant  $\mathbf{r}_{1/2}$  in (10.79) satisfies the bounds (10.4) for  $i = 1/2$  by (10.63), (10.64) and (7.1). To conclude we set (recall (10.59), (10.85))

$$\Phi_2 := \mathbf{T}_4 \circ \mathbf{T}_3.$$

The estimate (10.56) follows by composition using (10.65) and (10.82). We conclude the proof by studying the algebraic properties of  $\Phi_2$  and  $\mathcal{L}_5$ .

The map  $\Phi_2$  is in  $\mathfrak{S}_1$  (recall Def. 2.26) since  $\mathbf{T}_3, \mathbf{T}_4 \in \mathfrak{S}_1$  (see Lemmata 10.5, 10.9). Since, by Prop. 10.4,  $\mathcal{L}_4$  is in  $\mathfrak{S}_1$  we have that  $\mathcal{L}_5$  in (10.54) is Hamiltonian and  $x$ -translation invariant. The symbol  $d^{(5)}(\varphi, x, \xi)$  satisfies (2.63), (2.68) by construction, and so the operator  $\mathcal{L}_5 \in \mathfrak{S}_1$ .  $\square$

**10.3. Integrability at order zero.** The aim of this section is to conjugate the operator  $\mathcal{L}_5$  in (10.54) to constant coefficients up to a smoothing remainder of order  $-1/2$ . The key result of the section is the following.

**Proposition 10.10.** *Let  $\rho \geq 3$ ,  $p \geq s_0$  and assume that (7.1) holds. Then there exist  $\mu = \mu(\nu)$  and  $\mathbf{m}_0: \Omega_\varepsilon \rightarrow \mathbb{R}$  of the form (10.2) with  $i = 0$  such that, for all  $\omega \in \Omega_\infty^{2\gamma}$  (see (10.5)), there exists a map  $\Phi_3(\omega): H_{S^\pm}^s(\mathbb{T}) \times H_{S^\pm}^s(\mathbb{T}) \rightarrow H_{S^\pm}^s(\mathbb{T}) \times H_{S^\pm}^s(\mathbb{T})$  such that (recall (10.54))*

$$\mathcal{L}_6 := \Phi_3 \mathcal{L}_5 \Phi_3^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} d^{(6)}(\varphi, x, \xi) & 0 \\ 0 & d^{(6)}(\varphi, x, -\xi) \end{pmatrix} + \mathcal{Q}^{(6)} \right) \quad (10.87)$$

where the remainder  $\mathcal{Q}^{(6)} \in \mathfrak{L}_{\rho,p}^4 \otimes \mathcal{M}_2(\mathbb{C})$ , the symbol  $d^{(6)}(\varphi, x, \xi)$  has the form

$$d^{(6)}(\varphi, x, \xi) := \mathbf{im}_1 \xi + i(1 + \mathbf{m}_{\frac{1}{2}})|\xi|^{\frac{1}{2}} + \mathbf{im}_0 \text{sign}(\xi) + iq^{(6)}(\varphi, x, \xi), \quad (10.88)$$

with  $\mathbf{m}_1, \mathbf{m}_{1/2}$  in (10.2) with  $i = 1, 1/2$ , and  $q^{(6)}(\varphi, x, \xi) \in S_4^{-1/2}$ . Moreover for any  $\omega \in \Omega_\infty^{2\gamma}$  we have

$$\|(\Phi_3)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}. \quad (10.89)$$

Finally the map  $\Phi_3$  are in  $\mathfrak{S}_1$  and the operator  $\mathcal{L}_6$  are in  $\mathfrak{S}_1$  (see Def. 2.26, 2.28).

As done in sections 10.1, 10.2, we divide the construction of the map  $\Phi_j$  into two steps.

**10.3.1. Preliminary steps.** We prove the following result.

**Lemma 10.11. (Preliminary steps).** *There exists a real valued symbol  $\beta^{(5)}(\varphi, x, \xi) \in S^0$  of the form*

$$\beta^{(5)}(\varphi, x, \xi) = \sum_{i=1}^8 \varepsilon^i \beta_i^{(5)}(\varphi, x) \text{sign}(\xi), \quad (10.90)$$

(recall the notation (2.19)) where  $\beta_i^{(5)}$  is a real valued  $i$ -homogeneous functions in  $S^0$  as in (7.14) for  $i = 1, \dots, 8$  satisfying

$$\|\beta^{(5)}\|_s^{\gamma, \mathcal{O}_0} \lesssim_s \varepsilon, \quad (10.91)$$

such that the following holds. If

$$\mathbf{T}_5^\tau := \begin{pmatrix} \mathcal{T}_{\beta^{(5)}}^\tau(\varphi) & 0 \\ 0 & \mathcal{T}_{\beta^{(5)}}^\tau(\varphi) \end{pmatrix}, \quad \tau \in [0, 1], \quad (10.92)$$

where  $\mathcal{T}_{\beta^{(5)}}^\tau$  is the flow of (A.1) with generator as in (A.4) with  $f(\varphi, x, \xi) \rightsquigarrow \beta^{(5)}$ , then

$$\mathcal{L}_{5,*} := \mathbf{T}_5^1 \mathcal{L}_5 (\mathbf{T}_5^1)^{-1} = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} d_*^{(5)}(\varphi, x, \xi) & 0 \\ 0 & d_*^{(5)}(\varphi, x, -\xi) \end{pmatrix} + \mathcal{Q}_*^{(5)} \right) \quad (10.93)$$

where  $\mathcal{Q}_*^{(5)} \in \mathfrak{L}_{\rho,p}^3 \otimes \mathcal{M}_2(\mathbb{C})$  and (recall (10.88))

$$d_*^{(5)}(\varphi, x, \xi) := \mathbf{im}_1 \xi + i(1 + \mathbf{m}_{\frac{1}{2}})|\xi|^{\frac{1}{2}} + iq_*(\varphi, x) \text{sign}(\xi) + iq_*^{(5)}(\varphi, x, \xi), \quad (10.94)$$

$$q_*(\varphi, x) := \left( \varepsilon^2 m_0 + \sum_{k=1}^4 \varepsilon^{2k} m_0^{(2k)} \right) + q_{\geq 9}^{(*)}(\varphi, x), \quad (10.95)$$

where  $m_0, m_0^{(2k)}$  are respectively 2,  $2k$ -homogeneous in  $S^0$  and integrable (according to Def. 10.1). Moreover  $q_* \in S_3^0$  (recall Def. 7.3), independent of  $\xi \in \mathbb{R}$  and

$$\|q_{\geq 9}^{(*)}\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-2} (\varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad (10.96)$$

$$\|\Delta_{12} q_{\geq 9}^{(*)}\|_p \lesssim_p \gamma^{-2} \varepsilon (1 + \|\mathcal{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \quad (10.97)$$

for some  $\mu = \mu(\nu) > 0$ . The symbol  $q_*^{(5)}(\varphi, x, \xi)$  is in  $S_3^{-\frac{1}{2}}$ . Finally the map  $\mathcal{T}_{\beta^{(5)}}^\tau$  is in  $\mathfrak{T}_1$  and satisfies

$$\|(\mathcal{T}_{\beta^{(5)}}^\tau)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \varepsilon \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}, \quad \forall \tau \in [0, 1]. \quad (10.98)$$

The operator  $\mathcal{L}_{5,*}$  is in  $\mathfrak{S}_1$ .

*Proof.* We follow the strategy used in the proof of Lemma 10.8. By Lemma A.4 we have

$$\mathcal{T}_{\beta^{(5)}}^1 = \Pi_S^\perp \Phi_{\beta^{(5)}}^\tau \Pi_S^\perp \circ (\text{Id} + \mathcal{R})$$

where  $\Phi_{\beta^{(5)}}^\tau$  is the flow of (A.5) with  $f$  as in (A.4) with  $\beta \rightsquigarrow \beta^{(5)}$  and where  $\mathcal{R}$  is a finite rank operator of the form (A.14) satisfying (A.15). Then by Lemma A.3 we have that the flow  $\mathcal{T}_{\beta^{(5)}}^\tau$  is well posed and satisfies bounds like (A.9)-(A.11).

By Lemma A.5 we have that the conjugate of  $\mathcal{L}_5$  in (10.54) under the map  $\mathbf{T}_5^1$  is given by

$$\Pi_S^\perp \tilde{\mathbf{T}}_5^1 \mathcal{L}_5 (\tilde{\mathbf{T}}_5^1)^{-1} \Pi_S^\perp, \quad \tilde{\mathbf{T}}_5^\tau := \begin{pmatrix} \Phi_{\beta^{(5)}}^\tau(\varphi) & 0 \\ 0 & \frac{0}{\Phi_{\beta^{(5)}}^\tau(\varphi)} \end{pmatrix}, \quad \tau \in [0, 1],$$

up to finite rank operators belonging to  $\mathfrak{L}_{\rho,p}^3 \otimes \mathcal{M}_2(\mathbb{C})$ . Hence we shall apply Lemmata A.9, A.10. We start by studying the time contribution. By Lemma A.10 we have that (see (A.35), (A.39))

$$\Phi_{\beta^{(5)}}^1 \omega \cdot \partial_\varphi (\Phi_{\beta^{(5)}}^1)^{-1} = \omega \cdot \partial_\varphi - iOp^W \left( \omega \cdot \partial_\varphi \beta^{(5)}(\varphi, x, \xi) + r_1(\varphi, x, \xi) \right), \quad (10.99)$$

up to smoothing remainder  $\mathcal{R}_1$  in the class  $\mathfrak{L}_{\rho,p}^3$  and where  $r_1$  is some symbol in  $S_3^{-1}$ .

We now consider the contribution coming from the conjugation of the spatial operator by using Lemma A.9 applied to the pseudo differential operator with symbol in (10.88). First we notice that (see (A.30), (A.33))

$$\Phi_{\beta^{(5)}}^1 Op^W(\text{im}_1 \xi) (\Phi_{\beta^{(5)}}^1)^{-1} = iOp^W(\mathbf{m}_1 \xi + \{\beta^{(5)}, \mathbf{m}_1 \xi\} + r_2) = Op^W(\text{im}_1 \xi - \mathbf{m}_1 \partial_x(\beta^{(5)}) + r_2) \quad (10.100)$$

up to smoothing remainder  $\mathcal{R}_2$  in the class  $\mathfrak{L}_{\rho,p}^3$  and where  $r_2$  is some symbol in  $S_3^{-1}$ .

Finally we have (recall (10.54))

$$\begin{aligned} \Phi_{\beta^{(5)}}^1 (d^{(j)} - \text{im}_1 \xi) (\Phi_{\beta^{(5)}}^1)^{-1} &= iOp^W(d^{(5)} - \text{im}_1 \xi + r_3 + r_4) \\ r_3 &:= \{\beta^{(5)}, d^{(5)} - \text{im}_1 \xi\} \in S_3^{-\frac{1}{2}}, \end{aligned} \quad (10.101)$$

up to smoothing remainder  $\mathcal{R}_3$  in the class  $\mathfrak{L}_{\rho,p}^3$  and where  $r_4$  is some symbol in  $S_3^{-3/2}$ . By collecting the (10.99)-(10.101) we conclude that the conjugate of  $\mathcal{L}_5$  in (10.54) (see also (10.55)) under the map in (10.92) has the form (10.93) with symbol  $d_*^{(5)}(\varphi, x, \xi)$  as in (10.94) where  $q_*^{(5)} \in S_3^{-\frac{1}{2}}$ , and

$$q_* \text{sign}(\xi) = b^{(5)}(\varphi, x) \text{sign}(\xi) - (\omega \cdot \partial_\varphi + \mathbf{m}_1 \partial_x) \beta^{(5)}. \quad (10.102)$$

We shall prove that it is possible to choose  $\beta^{(5)}$  in (10.90) in such a way the symbol  $q_*$  in (10.102) has the form (10.95). Recall that  $b^{(5)}$  is a symbol in the class  $S_3^0$  (see Prop. 10.7) and hence it admits an expansion as (7.13) with  $k = 3$  for some  $i$ -homogeneous symbols  $b_i^{(5)}$  and a non-homogeneous symbol  $\tilde{b}^{(5)}$ . Therefore, recalling the expansion of  $\mathbf{m}_1$  in (10.2) and of  $\omega$  in (5.4), we have

$$\begin{aligned} q_* &= \sum_{i=1}^8 \varepsilon^i b_i^{(5)} - \sum_{i=1}^8 \varepsilon^i \bar{\omega} \cdot \partial_\varphi \beta_i^{(5)} - \sum_{i=1}^8 \varepsilon^i \left( \varepsilon^2 \mathbb{A} \xi + \sum_{k=2}^6 \varepsilon^{2k} \mathbf{b}_k(\xi, \dots, \xi) \right) \cdot \partial_\varphi \beta_i^{(5)} \\ &\quad - \left( \varepsilon^2 m_1 + \sum_{k=2}^6 \varepsilon^{2k} m_1^{2k} + \mathbf{r} \right) \sum_{i=1}^8 \varepsilon^i (\beta_i^{(5)})_x + \tilde{b}^{(5)}. \end{aligned} \quad (10.103)$$

This implies that the symbol  $q_*$  has an expansion as in (7.13) with  $k = 3$  for some non-homogeneous symbol  $\tilde{q}_*$  satisfying (7.15), with  $k = 3$ , and some  $i$ -homogeneous symbols  $q_{i,*}$ ,  $i = 1, \dots, 8$ . In particular we have

$$q_{1,*} := F_1 - \bar{\omega} \cdot \partial_\varphi \beta_1^{(5)}, \quad F_1 := b_1^{(5)} \quad (10.104)$$

$$q_{2,*} := F_2 - \bar{\omega} \cdot \partial_\varphi \beta_2^{(5)}, \quad F_2 := b_2^{(5)} \quad (10.105)$$

$$q_{3,*} := F_3 - \bar{\omega} \cdot \partial_\varphi \beta_3^{(5)}, \quad F_3 := b_3^{(5)} + \mathbf{f}_3, \quad \mathbf{f}_3 := -\mathbb{A}\xi \cdot \partial_\varphi \beta_1^{(5)} - m_1(\beta_1^{(5)})_x, \quad (10.106)$$

$$q_{i,*} := F_i - \bar{\omega} \cdot \partial_\varphi \beta_i^{(5)}, \quad F_i := b_i^{(5)} + \mathbf{f}_i, \quad 4 \leq i \leq 8, \quad (10.107)$$

where  $\mathbf{f}_i$  are some  $i$ -homogeneous symbols in  $S^0$ , independent of  $\xi \in \mathbb{R}$ , of the form (7.14) and depending only on  $\beta_k^{(5)}$  with  $1 \leq k \leq i - 1$ .

We now prove inductively that, for  $i \geq 1$ , there are functions  $\beta_i^{(5)}$  of the form (7.14) which are real valued and satisfying (10.91) such that (recall (10.106),(10.107))

$$q_{i,*} \equiv \begin{cases} m_0^{(i)}(\xi) & i \text{ even}, \\ 0 & i \text{ odd}, \end{cases} \quad (10.108)$$

with

$$m_0^{(i)} := \sum_{\mathcal{S}_i} (F_i)_{n_1, \dots, n_i}^{\sigma_1, \dots, \sigma_i}(\xi) \sqrt{\xi_{j_1} \dots \xi_{j_i}} e^{i(\sigma_1 j_1 + \dots + \sigma_i j_i)x} e^{i(\sigma_1 1(j_1) + \dots + \sigma_i 1(j_i)) \cdot \varphi}$$

where  $\sum_{\mathcal{S}_i}$  denotes the sum restricted over indexes in  $\mathcal{S}_i$ . Notice that  $m_0^{(i)}$  is integrable according to Definition 10.1 (recall Remark 10.2).

Assume by induction that the assertion above is true for  $1 \leq k \leq j$ . By construction (see the (10.103)) the function  $F_{j+1}$  depends only on  $b_k^{(5)} \text{sign}(\xi)$  (which is real) and  $\beta_k^{(5)}$ ,  $k \leq j$ , hence, by the inductive hypothesis it is real valued. We set, for  $\vec{\sigma} = (\sigma_1, \dots, \sigma_{i+1})$ ,  $\vec{j} = (j_1, \dots, j_{i+1})$ ,

$$(\beta_{i+1}^{(5)})_{\vec{j}}^{\vec{\sigma}} := \frac{-(F_{i+1})_{\vec{j}}^{\vec{\sigma}}}{i \mathcal{R}_{\vec{\sigma}}(\vec{j})}, \quad \{(j_k, \sigma_k)\}_{k=1}^{i+1} \notin \mathcal{S}_{i+1}, \quad (\beta_{i+1}^{(5)})_{\vec{j}}^{\vec{\sigma}} = 0, \quad \{(j_k, \sigma_k)\}_{k=1}^{i+1} \in \mathcal{S}_{i+1}, \quad (10.109)$$

and  $\mathcal{R}_{\vec{\sigma}}(\vec{j})$  is the function defined in (4.1). Since  $F_{j+1}$  is real, then the coefficients  $(F_{i+1})_{n_1, \dots, n_{i+1}}^{\sigma_1, \dots, \sigma_{i+1}}(\xi)$  satisfies (7.24). By an explicit computation one can check that also  $(\beta_{i+1}^{(5)})_{\vec{j}}^{\vec{\sigma}}$  satisfies (7.24), so that  $\beta_{i+1}^{(5)}$  is real valued. The estimate (10.91) follows by using Remark 10.2. The (10.98) follows by Lemma A.3. Using (10.109) one verifies the (10.108). By construction  $\mathcal{L}_{5,*} \in \mathfrak{S}_1$  and the map  $\mathcal{T}_{\beta^{(5)}}^\tau$  is in  $\mathfrak{X}_1$  (i.e. is symplectic and  $x$ -translation invariant). This concludes the proof of the lemma.  $\square$

**10.3.2. Reduction to constant coefficients at order zero.** The aim of this section is to eliminate the  $(\varphi, x)$  dependence in the symbol  $q_*$  in (10.95) appearing in the operator  $\mathcal{L}_{5,*}$  in (10.93). Then we conclude the proof of Proposition 10.7. We first need a preliminary result.

**Lemma 10.12.** *For any  $\omega \in \Omega_\infty^{2\gamma}$  (see (10.5)) there exists a real valued symbol  $\beta^{(6)}(\varphi, x, \xi) \in S_4^0$  satisfying (2.63) and such that*

$$q_*(\varphi, x) \text{sign}(\xi) - (\omega \cdot \partial_\varphi + m_1 \partial_x) \beta^{(6)}(\varphi, x, \xi) = m_0 \text{sign}(\xi), \quad (10.110)$$

where  $q_*$  is in (10.95), the constant  $m_0$  is in (10.2),  $i = 0$ , with

$$\mathbf{r}_0 := \int_{\mathbb{T}^\nu} q_*(\varphi, x) d\varphi. \quad (10.111)$$

The symbol  $\beta^{(6)}(\varphi, x, \xi)$  satisfies

$$|\beta^{(6)}|_{0,s,\alpha}^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad (10.112)$$

$$|\Delta_{12} \beta^{(6)}|_{0,p,\alpha} \lesssim_p \varepsilon \gamma^{-3} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \quad (10.113)$$

for some  $\mu > 0$  depending on  $\nu, \alpha$ . Moreover the map  $\mathcal{T}_{\beta^{(6)}}^\tau(\varphi)$ ,  $\tau \in [0, 1]$ , defined as the flow of (A.1) with generator as in (A.4) with  $f \rightsquigarrow \beta^{(6)}$  is well posed and satisfies

$$\|(\mathcal{T}_{\beta^{(6)}}^\tau)^{\pm 1} u\|_s^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \|u\|_s^{\gamma, \Omega_\infty^{2\gamma}} + \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \|u\|_{s_0}^{\gamma, \Omega_\infty^{2\gamma}}, \quad (10.114)$$

uniformly in  $\tau \in [0, 1]$ . Finally  $\mathcal{T}_{\beta^{(6)}}^\tau$  belongs to  $\mathfrak{T}_1$ .

*Proof.* We reason as done in Lemma 10.9. Due to the special form of the symbol of order zero in (10.94)-(10.95) we look for a symbol  $\beta^{(6)}(\varphi, x, \xi)$  of the form

$$\beta^{(6)}(\varphi, x, \xi) := \sum_{l \in \mathbb{Z}^\nu, j \in \mathbb{Z}} (\beta^{(6)})_{lj} e^{i(l \cdot \varphi + jx)} \text{sign}(\xi).$$

Therefore, passing to the Fourier basis (recall (2.1)), we have that equation (10.110) reads (recall (10.95))

$$(q_{\geq 9}^{(*)})_{lj} - i(\omega \cdot l + \mathbf{m}_1 l)(\beta^{(6)})_{lj} = 0, \quad \forall l \in \mathbb{Z}^\nu, j \in \mathbb{Z} \setminus \{0\}, (l, j) \neq (0, 0). \quad (10.115)$$

Using the diophantine condition in (10.5) one can prove

$$|\beta^{(6)}|_{0,s,\alpha}^{\gamma, \Omega_\infty^{2\gamma}} \lesssim_s \gamma^{-1} |q_{\geq 9}^{(*)}|_{0,s+2\tau+1,\alpha}^{\gamma, \Omega_\infty^{2\gamma}} \stackrel{(10.96)}{\lesssim_s} \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathcal{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}})$$

which is the (10.112). The (10.113) follows similarly using (10.97). Since  $q_{\geq 9}^{(*)}$  is in  $S_\nu$  (recall that  $\mathcal{L}_5$  in (10.54) is in  $\mathfrak{S}_1$  and the fact that  $\mathcal{T}_{\beta^{(5)}}^\tau$  is in  $\mathfrak{T}_1$  is in  $\mathfrak{T}_1$  by Lemma 10.11) then, by (10.83), we deduce that also  $\beta^{(6)} \in S_\nu$ . By Lemma A.4 we have

$$\mathcal{T}_{\beta^{(6)}}^1 = \Pi_S^\perp \Phi_{\beta^{(6)}}^\tau \Pi_S^\perp \circ (\text{Id} + \mathcal{R})$$

where  $\Phi_{\beta^{(6)}}^\tau$  is the flow of (A.5) with  $f$  as in (A.4) with  $f \rightsquigarrow \beta^{(6)}$  and where  $\mathcal{R}$  is a finite rank operator of the form (A.14) satisfying (A.15). Then by Lemma A.3 we have that the flow  $\mathcal{T}_{\beta^{(6)}}^\tau$  is well posed and the estimate (10.114) follows by estimates (A.9)-(A.11). Moreover we have that the map  $\mathcal{T}_{\beta^{(6)}}^\tau$  is in  $\mathfrak{T}_1$ .  $\square$

**Proof of Proposition 10.10.** We consider the map

$$\mathbf{T}_6^\tau := \begin{pmatrix} \mathcal{T}_{\beta^{(6)}}^\tau(\varphi) & 0 \\ 0 & \overline{\mathcal{T}_{\beta^{(6)}}^\tau(\varphi)} \end{pmatrix}, \quad \tau \in [0, 1], \quad (10.116)$$

where  $\mathcal{T}_{\beta^{(6)}}^\tau(\varphi)$ ,  $\tau \in [0, 1]$  is given by Lemma 10.12. We now conjugate the operator  $\mathcal{L}_{5,*}$  in (10.93) with the map  $\mathbf{T}_6^1$ . First of all we remark that the symbol  $\beta^{(6)}(\varphi, x, \xi)$  belongs to the class  $S_4^0$  (see Def. 7.3). This is true because estimates (10.112), (10.113) imply the (7.15) with  $k = 4$ . We follow the strategy of section 10.2.1. By Lemma A.5 we have that the conjugate of  $\mathcal{L}_{5,*}$  in (10.93) is given by

$$\Pi_S^\perp \tilde{\mathbf{T}}_6^1 \mathcal{L}_{5,*} (\tilde{\mathbf{T}}_6^1)^{-1} \Pi_S^\perp, \quad \tilde{\mathbf{T}}_6^\tau := \begin{pmatrix} \Phi_{\beta^{(6)}}^\tau(\varphi) & 0 \\ 0 & \overline{\Phi_{\beta^{(6)}}^\tau(\varphi)} \end{pmatrix}, \quad \tau \in [0, 1],$$

up to finite rank operators belonging to  $\mathfrak{L}_{\rho,p}^4 \otimes \mathcal{M}_2(\mathbb{C})$ . Hence we shall apply Lemmata A.9, A.10 with  $f(\varphi, x, \xi)$  as in (10.86). By explicit computations using (A.30), (A.33), (A.35), (A.39) (as done in (10.66)-(10.69)) we obtain that the operator

$$\mathcal{L}_6 := \mathbf{T}_6^1 \mathcal{L}_{5,*} (\mathbf{T}_6^1)^{-1}$$

has the form (10.87), with some remainder  $\mathcal{Q}^{(6)} \in \mathfrak{L}_{\rho,p}^4$ , some symbol  $\mathbf{d}^{(6)}(\varphi, x, \xi)$  of the form

$$\mathbf{d}^{(6)}(\varphi, x, \xi) := \text{im}_1 \xi + i(1 + \mathbf{m}_{\frac{1}{2}})|\xi|^{\frac{1}{2}} + iA(\varphi, x, \xi) + iq^{(6)}(\varphi, x, \xi),$$

where

$$A = q_* \text{sign}(\xi) - (\omega \cdot \partial_\varphi + \mathbf{m}_1 \partial_x) \beta^{(6)}, \quad (10.117)$$

and  $q^{(6)} \in S_4^{-1/2}$ . By equation (10.110) in Lemma 10.12 we actually have that  $A(\varphi, x, \xi)$  in (10.117) is equal to  $\mathbf{m}_0$  in (10.2). By estimates (10.112), (10.113), and the fact that  $q_* \in S_3^0$  we have that  $q^{(6)}(\varphi, x, \xi)$  is in  $S_4^{-1/2}$  because it satisfies estimates like (7.15) with  $k = 4$ .

The constant  $\mathbf{r}_0$  in (10.111) satisfies the bounds (10.4),  $i = 0$ , by (10.96), (10.97) and (7.1). To conclude we set (recall (10.92), (10.116))

$$\Phi_3 := \mathbf{T}_6^1 \circ \mathbf{T}_5^1.$$

The estimate (10.89) follows by composition using (10.98) and (10.114). The map  $\Phi_3 \in \mathfrak{T}_1$  because  $\mathbf{T}_6^1, \mathbf{T}_5^1 \in \mathfrak{T}_1$ .  $\square$



**Proof of Proposition 10.3.** We set

$$\Phi := \Phi_3 \circ \Phi_2 \circ \Phi_1 \quad (10.118)$$

where the maps  $\Phi_j, j = 1, 2, 3$ , are given by Propositions 10.4, 10.7 and 10.10. The map  $\Phi$  is defined for  $\omega \in \Omega_\infty^{2\gamma}$  and belongs to  $\mathfrak{T}_1$  since  $\Phi_j \in \mathfrak{T}_1, j = 1, 2, 3$ . The (10.9) follows by (7.1) and (10.12), (10.56), (10.89). By construction we have that (see (10.87), (10.88))  $\mathcal{L}_6 := \Phi \mathcal{L}_3 \Phi^{-1}$  has the form

$$\mathcal{L}_6 = \Pi_S^\perp \left( \omega \cdot \partial_\varphi + Op^W \begin{pmatrix} iM(\xi) + iq^{(6)}(\varphi, x, \xi) & 0 \\ 0 & -iM(\xi) - iq^{(6)}(\varphi, x, -\xi) \end{pmatrix} + \mathcal{Q}^{(6)} \right) \quad (10.119)$$

with  $M(\xi)$  as in (10.7), and where  $q^{(6)}(\varphi, x, \xi) \in S_4^{-\frac{1}{2}}$  and  $\mathcal{Q}^{(6)} \in \mathfrak{L}_{\rho,p}^4 \otimes \mathcal{M}_2(\mathbb{C})$  (recall Def. 7.3). The (10.8) follows by setting  $q = q^{(6)}$  and  $Q = \mathcal{Q}^{(6)}$ .  $\square$

## 11. LINEAR BIRKHOFF NORMAL FORM

In this section the main goal is to normalize all the terms in the remainder  $Q$  in (10.6) which are not “*perturbative*” for the KAM procedure of section 12. This will be obtained by applying several changes of coordinates involving “small divisors” (see subsections 11.2, 11.3). In particular, thanks to the abstract result stated in Proposition 11.18, we will compute explicitly the corrections of size  $O(\varepsilon^2)$  to the eigenvalues of the operator  $\mathcal{L}_6$  in (10.6). Such corrections will play a fundamental role in the estimates of the measure of the resonant sets of frequencies.

The key result of the section is the following.

**Proposition 11.1. (Linear Birkhoff normal form).** *There are constants  $r_j : \Omega_\varepsilon \rightarrow \mathbb{R}$ , independent of  $\mathfrak{J}_\delta$ , and constants  $C_1, C_2 > 0$  depending only on  $S$  such that for any  $\omega \in \mathcal{G}_0^{(2)}(C_1, C_2)$  (see (5.23)) there exists a symplectic and  $x$ -translation invariant map  $\Upsilon(\varphi) : H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$  such that (recall 10.6)*

$$\mathcal{L}_{14} := \Upsilon \mathcal{L}_6 \Upsilon^{-1} = \Pi_S^\perp (\omega \cdot \partial_\varphi + \mathfrak{D} + \mathfrak{R}), \quad (11.1)$$

where

$$\mathfrak{D} := (\mathfrak{D}_\sigma^{\sigma'})_{\sigma, \sigma' = \pm}, \quad \mathfrak{D}_\sigma^{-\sigma} \equiv 0, \quad \overline{\mathfrak{D}_\pm^\pm} = \mathfrak{D}_\pm^\pm, \quad \mathfrak{D}_\pm^\pm := \text{diag}_{j \in S^c} (i d_j), \quad (11.2)$$

$$d_j := m_1(\omega) j + (1 + m_1^{(\geq 4)}(\omega)) \sqrt{|j|} + m_0^{(\geq 4)}(\omega) \text{sign}(j) + r_j(\omega), \quad d_j \in \mathbb{R}, \quad (11.3)$$

$$\sup_{j \in S^c} \langle j \rangle^{1/2} |r_j|^{\gamma, \Omega_\varepsilon} \lesssim \varepsilon^3, \quad r_j \in \mathbb{R} \quad (11.4)$$

where  $m_1, m_{1/2}^{(\geq 4)}, m_0^{(\geq 4)}$  are given in (10.2). The operator  $\mathfrak{R}$  is in  $\mathfrak{S}_0$  (see Def. 2.26) and is  $Lip(-1/2)$ -modulo tame (see Def. 2.7, 2.9) with the following estimates

$$\begin{aligned} \mathfrak{M}_{\mathfrak{R}}^{\sharp, \gamma}(-1/2, s, \mathfrak{b}_0) &\lesssim_s \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \\ \mathfrak{M}_{\Delta_{12} \mathfrak{R}}^{\sharp}(-1/2, p, \mathfrak{b}_0) &\lesssim_p \varepsilon \gamma^{-3} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \end{aligned} \quad (11.5)$$

for some  $\mu = \mu(\nu) > 0$ . The map  $\Upsilon$  does not depend on  $\mathfrak{J}_\delta$  and satisfies

$$\|\Upsilon^{\pm 1} u\|_s \lesssim_{s,S} \|u\|_s. \quad (11.6)$$

The proof of the proposition above involves many arguments and will be divided into several steps. The basic idea is to implement a “linear” Birkhoff procedure as explained in subsection 3.2. This will require some “non-resonance” conditions which are studied in subsection 11.1.1. A fundamental argument in the proof is the identification of normal forms result contained in Proposition 11.18. This result implies that the corrections at order  $\varepsilon^2$  in the eigenvalues (11.3) depend only on the constant  $m_1$  in (10.3).

**11.1. Preliminary technical results.** In the following subsection we shall prove some technical results which we need to implement the *Linear Birkhoff normal form* procedure.

We shall study the “non-resonance” conditions required on the small divisors (see subsection 11.1.1); then we will introduce the class of *almost-diagonal* operators (see subsection 11.1.2).

11.1.1. *Non-resonance conditions.* Recalling the vectors  $\bar{\omega}$  in (1.26) and  $\mathbf{v}$  is in (5.24) we define for  $p = 1, \dots, 6$

$$\delta_{\sigma, \sigma', j, k}^{(p)}(\ell) := \bar{\omega} \cdot \ell + \sigma \sqrt{|j|} - \sigma' \sqrt{|k|}, \quad \forall \sigma, \sigma' = \pm, j, k \in S^c, \ell \in \mathbb{Z}^\nu, \quad (11.7)$$

$$\text{with } \mathbf{v} \cdot \ell + j - k = 0, \quad |\ell| \leq p. \quad (11.8)$$

**Remark 11.2.** *Given a set of indexes  $j_1, \dots, j_N$ , we denote by  $\max(1)$  and  $\max(2)$  the first and the second largest values among the moduli of the  $j_i$ 's.*

*If  $\sum_{i=1}^N \sigma_i j_i = 0$  for some  $N > 0$ ,  $j_i \in \mathbb{Z} \setminus \{0\}$ ,  $\sigma_i \in \{\pm 1\}$ , then  $\max(2) \geq N^{-1} \max(1)$ . Indeed, suppose that  $|j_1| = \max(1)$  and  $|j_2| = \max(2) < \max(1) / N$ , then*

$$\max(1) = |\sigma_1 j_1| \leq \sum_{i=2}^N |\sigma_i j_i| \leq \frac{(N-1)}{N} \max(1),$$

which is a contradiction.

**Lemma 11.3. (Lower bounds).** *(i) There exist constants  $C_1(S), C_2(S) > 0$  such that*

$$|\delta_{\sigma, \sigma', j, k}^{(1)}(\ell)| \geq C_1(S), \quad \forall \sigma, \sigma' \in \{\pm 1\}, j, k \in \mathbb{Z} \setminus \{0\}, \ell \in \mathbb{Z}^\nu, |\ell| = 1, \quad (11.9)$$

except for  $\sigma = \sigma', j = k, \ell = 0$  and

$$|\delta_{\sigma, \sigma', j, k}^{(2)}(\ell)| \geq C_2(S), \quad \forall \sigma, \sigma' \in \{\pm 1\}, j, k \in \mathbb{Z} \setminus \{0\}, \ell \in \mathbb{Z}^\nu, |\ell| = 2, \quad (11.10)$$

except for  $\sigma = \sigma', j = k, \ell = 0$ .

*(ii) For  $3 \leq p \leq 6$  and a “generic” choice of the tangential sites  $S$  in (1.23) there exist constants  $C(S) \gg 1$ ,  $0 < K(S) \ll 1$  depending only on the  $\max(S)$  (see (1.25)) such that (recall (11.7))*

$$|\delta_{\sigma, \sigma', j, k}^{(p)}(\ell)| \geq K(S), \quad 0 < |\ell| \leq p, \quad |j|, |k| \geq C(S), \quad (11.11)$$

except for  $\sigma = \sigma', p$  even,  $j = k, \ell = 0$ .

*Proof.* See Appendix B.2. □

11.1.2. *Almost-diagonal operators.* We need further definitions.

**Definition 11.4. (Almost diagonal operators).** *We say that  $\mathbf{A}(\varphi): H_{S^\pm}^s(\mathbb{T}) \times H_{S^\pm}^s(\mathbb{T}) \rightarrow H_{S^\pm}^s(\mathbb{T}) \times H_{S^\pm}^s(\mathbb{T})$  is  $N$ -almost-diagonal if,  $\forall j, k \in S^c$ ,  $\sigma, \sigma' = \pm$ , one has*

$$(\mathbf{A})_{\sigma, j}^{\sigma', k}(\ell) \neq 0 \quad \Rightarrow \quad |\ell| \leq N.$$

*We say that  $\mathbf{A}(\varphi)$  is almost-diagonal if it is  $N$ -almost-diagonal for some  $N > 0$ .*

**Remark 11.5.** *Notice that the following facts hold:*

- *An  $i$ -homogeneous operator of the form (7.17) is  $i$ -almost-diagonal and  $x$ -translation invariant;*
- *if  $\mathbf{A}(\varphi)$  is  $x$ -translation invariant (according to Def. 2.20) and almost-diagonal, then (recall Lemma 2.22)*

$$\text{if } (\mathbf{A})_{\sigma, j}^{\sigma', k}(\ell) \neq 0 \quad \Rightarrow \quad \exists C = C(S) > 0 \quad \text{such that } |\ell| \leq C, |j - k| \leq C;$$

- *if  $\mathbf{A}(\varphi)$  is almost-diagonal and bounded on  $H_{S^\pm}^s(\mathbb{T}) \times H_{S^\pm}^s(\mathbb{T})$ , then it is tame according with Def. 2.3. This follows since  $\mathbf{A}$  has “decay” off-diagonal.*

**Lemma 11.6.** *Let  $m \geq 0$  and  $\mathbf{A}$  be an almost-diagonal,  $x$ -translational invariant operator. Then the operator  $\mathbf{B} := \langle D \rangle^m \mathbf{A} \langle D \rangle^m$  is bounded in the majorant norm (see Def. 2.5) topology if and only if*

$$|(\mathbf{A})_{\sigma, j}^{\sigma', k}(\ell)| \leq \frac{C}{\langle j, k \rangle^{2m}}, \quad \ell \in \mathbb{Z}^\nu, \quad j, k \in S^c, \quad (11.12)$$

for some constant  $C > 0$  depending only on  $S$ .

*Proof.* See Appendix B.2. □

**Lemma 11.7.** Fix  $\mathbf{b}_0 > 0$ . If  $\rho \gg s_0 + \mathbf{b}_0$  then the operator  $\mathbf{Q}$  in (10.8) has the form

$$\mathbf{Q} = \sum_{i=1}^6 \varepsilon^i \mathbf{Q}_i + \mathbf{Q}_{\geq 7} \quad (11.13)$$

with  $\mathbf{Q}_i$ ,  $i = 1, \dots, 6$ ,  $i$ -homogeneous as in (7.17) (see Def. 7.3). In particular one has

$$|(\mathbf{Q}_i)_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C}{\langle j, k \rangle^{1/2}}, \quad |(\mathbf{Q}_i)_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C}{\langle j, k \rangle^\rho}, \quad (11.14)$$

for some  $C > 0$  dependent on the tangential sites  $S$ . Moreover, for any  $\omega \in \Omega_\infty^{2\gamma}$  the operator  $\mathbf{Q}_{\geq 7}$  is Lip- $(-1/2)$ -modulo tame (see Def. 2.8)

$$\begin{aligned} \mathfrak{M}_{\mathbf{Q}_{\geq 7}}^{\sharp, \gamma}(-1/2, s, \mathbf{b}_0) &\lesssim_s \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \\ \mathfrak{M}_{\Delta_{12} \mathbf{Q}_{\geq 7}}^{\sharp}(-1/2, p, \mathbf{b}_0) &\lesssim_p \varepsilon \gamma^{-3} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \end{aligned} \quad (11.15)$$

for some  $\mu = \mu(\nu) > 0$ .

*Proof.* See Appendix B.2. □

**Definition 11.8.** Let  $\mathbf{M} = \mathbf{M}(\varphi)$  be a matrix of operators. (i) Given a constant  $C > 0$  we set

$$\mathbf{M} = \pi_C \mathbf{M} + \pi_C^\perp \mathbf{M},$$

where, for any  $\ell \in \mathbb{Z}^\nu$ ,  $j, k \in S^c$ , we define

$$(\pi_C^\perp \mathbf{M})_{\sigma,j}^{\sigma',k}(\ell) := \begin{cases} \mathbf{M}_{\sigma,j}^{\sigma',k}(\ell), & \max\{|j|, |k|\} \geq C \\ 0 & \text{otherwise.} \end{cases}$$

The operator  $\pi_C \mathbf{M}$  is defined by difference.

(ii) We define the operator  $\llbracket \mathbf{M} \rrbracket$  as

$$(\llbracket \mathbf{M} \rrbracket)_{\sigma,j}^{\sigma',k}(\ell) := \begin{cases} (\mathbf{M})_{\sigma,j}^{\sigma',k}(0) & \text{if } \sigma = \sigma', j = k, \ell = 0, \\ 0 & \text{otherwise.} \end{cases} \quad (11.16)$$

**Remark 11.9.** Notice that, if the operator  $\mathbf{M}$  is in  $\mathfrak{S}_0$  (see Def. 2.26), then  $\pi_C \mathbf{M}$ ,  $\pi_C^\perp \mathbf{M}$  belong to  $\mathfrak{S}_0$ .

We shall use the following notation: given some operator  $M = M(\varphi)$  we write

$$\text{ad}_{\mathbf{A}_i} [M] := [\mathbf{A}_i, M] = \mathbf{A}_i M - M \mathbf{A}_i. \quad (11.17)$$

The following result is fundamental for our scope.

**Lemma 11.10.** Let  $\mathbf{B}$  be a smoothing operator in  $\mathfrak{S}_0$  and  $p$ -almost-diagonal satisfying estimates like (11.14). Consider the function  $\delta_{\sigma,\sigma',j,k}^{(p)}(\ell)$  in (11.7) and let

$$\mathbf{C} := \mathbf{C}(S) := \max\{12|\nu|, 2C(S), 2\max(S)\} > 0, \quad (11.18)$$

where  $C(S)$  is the constant given by item (ii) of Lemma 11.3. We define the operator  $\mathbf{A} = \mathbf{A}(\varphi) = \pi_C^\perp \mathbf{A}(\varphi)$  as (see Def. 11.8)

$$(\pi_C^\perp \mathbf{A})_{\sigma,j}^{\sigma',k}(\ell) := \begin{cases} \mathbf{B}_{\sigma,j}^{\sigma',k}(\ell) / i \delta_{\sigma,\sigma',j,k}^{(p)}(\ell), & \max\{|j|, |k|\} \geq \mathbf{C}, \ell \neq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (11.19)$$

Then  $\mathbf{A}(\varphi)$  is in  $\mathfrak{S}_0$  and  $p$ -almost-diagonal and solves the equation

$$-\text{ad}_{\bar{\omega} \cdot \partial_\varphi + \mathbf{D}_0} [\mathbf{A}] + \mathbf{B} = \llbracket \pi_C^\perp \mathbf{B} \rrbracket + \pi_C \mathbf{B}, \quad \mathbf{D}_0 := Op^W \begin{pmatrix} i|\xi|^{1/2} & 0 \\ 0 & -i|\xi|^{1/2} \end{pmatrix}. \quad (11.20)$$

Moreover one has

$$|\mathbf{A}_{\sigma,j}^{\sigma,k}(\ell)| \leq (K(S))^{-1} \frac{C}{\langle j, k \rangle^{1/2}}, \quad |\mathbf{A}_{\sigma,j}^{-\sigma,k}(\ell)| \leq (K(S))^{-1} \frac{C}{\langle j, k \rangle^\rho} \quad (11.21)$$

where  $K(S)$  is in (11.11) and  $C > 0$  is a constant depending on the tangential sites.

*Proof.* See Appendix B.2. □

**Lemma 11.11.** *Let  $\mathbf{A} \in \mathfrak{S}_0$  be  $i$ -almost-diagonal satisfying*

$$|(\mathbf{A})_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C(\mathbf{A})}{\langle j, k \rangle^{1/2}}, \quad |(\mathbf{A})_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C(\mathbf{A})}{\langle j, k \rangle^\rho}, \quad (11.22)$$

for some constant  $C(\mathbf{A}) > 0$  and some  $\rho > 0$ . Then the following holds.

(i) *The operator  $\mathbf{A}$  is Lip- $(-1/2)$ -modulo tame with*

$$\mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) \lesssim_s C(\mathbf{A}). \quad (11.23)$$

(ii) *Let  $\mathbf{B} \in \mathfrak{S}_0$  be a Lip- $(-1/2)$ -modulo tame operator (see Def. 2.8). The operator  $\mathbf{C}_k := \text{ad}_{\mathbf{A}}^k[\mathbf{B}]$ ,  $k \geq 1$  is in  $\mathfrak{S}_0$ , Lip- $(-1/2)$ -modulo tame and*

$$\begin{aligned} \mathfrak{M}_{\mathbf{C}_k}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) &\lesssim_s \mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) (\mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0))^{k-1} \mathfrak{M}_{\mathbf{B}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) \\ &\quad + (\mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0))^k \mathfrak{M}_{\mathbf{B}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0), \\ \mathfrak{M}_{\Delta_{12}\mathbf{C}_k}^{\sharp}(-1/2, p, \mathbf{b}_0) &\lesssim_p \mathfrak{M}_{\mathbf{B}(i_1)}^{\sharp}(-1/2, p, \mathbf{b}_0) \times \\ &\quad \times \sum_{j_1+j_2=k-1} (\mathfrak{M}_{\mathbf{A}(i_1)}^{\sharp,\gamma}(-1/2, p, \mathbf{b}_0))^{j_1} \mathfrak{M}_{\Delta_{12}\mathbf{A}}^{\sharp}(-1/2, p, \mathbf{b}_0) (\mathfrak{M}_{\mathbf{A}(i_2)}^{\sharp,\gamma}(-1/2, p, \mathbf{b}_0))^{j_2} \\ &\quad + (\mathfrak{M}_{\mathbf{A}(i_2)}^{\sharp,\gamma}(-1/2, p, \mathbf{b}_0))^k \mathfrak{M}_{\Delta_{12}\mathbf{B}}^{\sharp}(-1/2, p, \mathbf{b}_0). \end{aligned}$$

(iii) *Let  $n \geq 1$  and consider  $\mathbf{F} := \sum_{k \geq n} (\varepsilon^k/k!) \mathbf{C}_k$ . For  $\varepsilon$  small enough we have that  $\mathbf{F} \in \mathfrak{S}_0$  and Lip- $(-1/2)$ -modulo tame with the following bounds*

$$\begin{aligned} \mathfrak{M}_{\mathbf{F}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) &\lesssim_s \mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) \mathfrak{M}_{\mathbf{B}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) \varepsilon^n (\mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0))^{n-1} \\ &\quad + \mathfrak{M}_{\mathbf{B}}^{\sharp,\gamma}(s, \mathbf{b}_0) \varepsilon^n (\mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0))^n, \end{aligned} \quad (11.24)$$

$$\mathfrak{M}_{\Delta_{12}\mathbf{F}}^{\sharp}(-1/2, p, \mathbf{b}_0) \lesssim_p \mathfrak{M}_{\mathbf{B}(i_1)}^{\sharp}(-1/2, p, \mathbf{b}_0) \mathfrak{M}_{\Delta_{12}\mathbf{A}}^{\sharp}(-1/2, p, \mathbf{b}_0) + \mathfrak{M}_{\Delta_{12}\mathbf{B}}^{\sharp}(-1/2, p, \mathbf{b}_0). \quad (11.25)$$

*Proof.* See Appendix B.2.  $\square$

Now consider the operator  $\mathbf{D}$  in (10.7). We write (recall (10.2))

$$\begin{aligned} \mathbf{D} &:= \mathbf{D}_0 + \varepsilon^2 \mathbf{D}_2 + \varepsilon^4 \mathbf{D}_4 + \varepsilon^6 \mathbf{D}_6 + \mathbf{D}_{\geq 8}, \quad \mathbf{D}_0 := Op^W \begin{pmatrix} i|\xi|^{1/2} & 0 \\ 0 & -i|\xi|^{1/2} \end{pmatrix}, \\ \mathbf{D}_2 &:= Op^W \begin{pmatrix} i\mathbf{M}_2(\xi) & 0 \\ 0 & -i\mathbf{M}_2(-\xi) \end{pmatrix}, \quad i\mathbf{M}_2(\xi) := im_1 \xi + im_{\frac{1}{2}} |\xi|^{\frac{1}{2}} + im_0 \text{sign}(\xi), \\ \mathbf{D}_{2k} &:= Op^W \begin{pmatrix} i\mathbf{M}_{2k}(\xi) & 0 \\ 0 & -i\mathbf{M}_{2k}(-\xi) \end{pmatrix}, \quad i\mathbf{M}_{2k}(\xi) := im_1^{(2k)} \xi + im_{\frac{1}{2}}^{(2k)} |\xi|^{\frac{1}{2}} + im_0^{(2k)} \text{sign}(\xi), \end{aligned} \quad (11.26)$$

with  $k = 2, 3$ , and  $\mathbf{D}_{\geq 8} := \mathbf{D} - \mathbf{D}_0 - \varepsilon^2 \mathbf{D}_2 - \varepsilon^4 \mathbf{D}_4 - \varepsilon^6 \mathbf{D}_6$ . In the following Lemma we study how the operator  $\mathbf{D}$  conjugates under a map generated by an almost-diagonal operator.

**Lemma 11.12.** *Let  $\mathbf{A}, \mathbf{B} \in \mathfrak{S}_0$  and assume that  $\mathbf{A}$  is  $i_1$ -almost-diagonal,  $\mathbf{B}$  is  $i_2$ -almost-diagonal and that they satisfy estimates like (11.22). Then we have:*

(i) *the operator  $\mathbf{C} := \text{ad}_{\mathbf{A}}[\mathbf{B}]$  is in  $\mathfrak{S}_0$ , is  $(i_1 + i_2)$ -almost-diagonal and*

$$|\mathbf{C}_{\sigma,j}^{\sigma,k}(\ell)| \lesssim \frac{C(\mathbf{C})}{\langle j, k \rangle^{\frac{1}{2}}}, \quad |\mathbf{C}_{\sigma,j}^{-\sigma,k}(\ell)| \lesssim \frac{C(\mathbf{C})}{\langle j, k \rangle^\rho}, \quad (11.27)$$

for any  $j, k \in S^c$ ,  $\ell \in \mathbb{Z}^\nu$ ,  $\sigma = \pm$  for some constant  $C(\mathbf{C})$  depending on  $S$  and the constant  $C(\mathbf{A})$ ,  $C(\mathbf{B})$ .

(ii) *Let  $\mathcal{D} := \mathbf{D} - \mathbf{D}_0 - \mathbf{D}_{\geq 8}$ . Define  $\mathbf{M} := \text{ad}_{\mathbf{A}}[\mathcal{D}]$ . Then  $\mathbf{M}$  belongs to  $\mathfrak{S}_0$  and*

$$|(\mathbf{M})_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{\varepsilon^2 C}{\langle j, k \rangle^{1/2}}, \quad |(\mathbf{M})_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{\varepsilon^2 C}{\langle j, k \rangle^{\rho-1}}. \quad (11.28)$$

(iii) *Let  $\mathbf{N} := \text{ad}_{\mathbf{A}}[\mathbf{D}_{\geq 8}]$ . Then  $\mathbf{N} \in \mathfrak{S}_0$  is Lip- $(-1/2)$ -modulo tame with*

$$\begin{aligned} \mathfrak{M}_{\mathbf{N}}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) &\lesssim_s \varepsilon^{17-2b} \gamma^{-3}, \\ \mathfrak{M}_{\Delta_{12}\mathbf{N}}^{\sharp}(-1/2, p, \mathbf{b}_0) &\lesssim_p \varepsilon^{17-2b} \gamma^{-3} + C(\mathbf{A}) \varepsilon \gamma^{-2} (1 + \varepsilon \|\mathcal{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}. \end{aligned}$$

*Proof.* See Appendix B.2.  $\square$

**11.2. Terms  $O(\varepsilon)$ ,  $O(\varepsilon^2)$  and identification of normal forms.** In this section we perform the first two steps of linear Birkhoff normal form in order to normalize the terms  $O(\varepsilon)$  and  $O(\varepsilon^2)$  in (10.6).

We look for symplectic linear changes of variable  $\Upsilon_i(\varphi): H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$ ,  $i = 1, \dots, 6$  of the form

$$\Upsilon_i := \exp(\varepsilon^i \mathbf{A}_i) \quad (11.29)$$

where  $\mathbf{A}_i(\varphi)$  is a Hamiltonian operator (see Definition 2.20).

**Remark 11.13.** Notice that if the operator  $\mathbf{A}_i \in \mathfrak{S}_0$  then  $\Upsilon_i$  belongs to  $\mathfrak{T}_1$  (see Def. 2.28).

Recalling (11.17), we have that the conjugate of  $M$  under the map  $\Upsilon_i$  is given (formally) by the Lie series

$$\Upsilon_i M \Upsilon_i^{-1} = \sum_{k \geq 0} \frac{1}{k!} \text{ad}_{\mathbf{A}_i}^k M, \quad \text{ad}_{\mathbf{A}_i}^k [M] = \text{ad}_{\mathbf{A}_i} [\text{ad}_{\mathbf{A}_i}^{k-1} M], \quad k \geq 1, \quad \text{ad}_{\mathbf{A}_i}^0 M = M. \quad (11.30)$$

We define  $\Pi_{\text{Ker}(\mathbf{A})}$  and  $\Pi_{\text{Rg}(\mathbf{A})}$  the projection on the kernel and the Range of  $\text{ad}_{\mathbf{A}}$  respectively.

**Step one (order  $\varepsilon$ ).** Let us define the matrix  $\mathbf{A}_1$  as

$$(\mathbf{A}_1)_{\sigma,j}^{\sigma',k}(\ell) = \frac{(\mathbf{Q}_1)_{\sigma,j}^{\sigma',k}(\ell)}{\text{i}\delta_{\sigma,\sigma',j,k}^{(1)}(\ell)} \quad j, k \in S^c, \ell \in \mathbb{Z}^\nu \quad (11.31)$$

where  $\mathbf{Q}_1$  is in (11.13) and  $\delta^{(1)}$  is defined in (11.7). Note that by Lemma 11.3 the denominator in (11.31) never vanishes. We have the following result.

**Lemma 11.14.** The operator  $\mathbf{A}_1$  is in  $\mathfrak{S}_0$  and 1-almost-diagonal and satisfies

$$|(\mathbf{A}_1)_{\sigma,j}^{\sigma',k}(\ell)| \leq (C_1(S))^{-1} |(\mathbf{Q}_1)_{\sigma,j}^{\sigma',k}(\ell)|, \quad (11.32)$$

with  $C_1(S) > 0$  in (11.9). Moreover it solves the equation

$$\text{ad}_{\overline{\omega} \cdot \partial_\varphi + \text{D}_0} [\mathbf{A}_1] = \mathbf{Q}_1. \quad (11.33)$$

*Proof.* Estimate (11.32) follows by Lemma 11.3. Thanks to the definition of  $\mathbf{A}_1$  in (11.31) one can check that equation (11.33) is satisfied. The operator  $\mathbf{A}_1$  is almost-diagonal by definition. Recall that  $\mathbf{Q}_1$  is Hamiltonian and  $x$ -translation invariant. By using the expression of  $\delta_{\sigma,\sigma',j,k}^{(1)}(\ell)$  in (11.7) we have

$$\overline{\delta_{\sigma,\sigma',j,k}^{(1)}(\ell)} = \delta_{\sigma,\sigma',k,j}^{(1)}(-\ell), \quad \overline{\delta_{\sigma,\sigma',-j,-k}^{(1)}(-\ell)} = \delta_{-\sigma,-\sigma',k,j}^{(1)}(\ell). \quad (11.34)$$

Thus by Lemmata 2.23, 2.22  $\mathbf{A}_1$  is Hamiltonian and  $x$ -translation invariant.  $\square$

**Lemma 11.15.** The conjugate of the operator  $\mathcal{L}_6$  in (10.6) has the form

$$\mathcal{L}_7 := \Upsilon_1 \mathcal{L}_6 \Upsilon_1^{-1} = \omega \cdot \partial_\varphi + \text{D} + \sum_{i=2}^6 \varepsilon^i \mathbf{Q}_i^{(1)} + \Pi_S^\perp \mathfrak{R}_1 \quad (11.35)$$

where  $\mathbf{Q}_i^{(1)}$ ,  $i = 2, \dots, 6$ , are in  $\mathfrak{S}_0$ , are  $i$ -almost-diagonal and satisfy

$$|(\mathbf{Q}_i^{(1)})_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C}{\langle j, k \rangle^{1/2}}, \quad |(\mathbf{Q}_i^{(1)})_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C}{\langle j, k \rangle^{p-1}}, \quad \mathbf{Q}_2^{(1)} = \mathbf{Q}_2 + \frac{1}{2} \text{ad}_{\mathbf{A}_1} [\mathbf{Q}_1] \quad (11.36)$$

and  $\mathfrak{R}_1$  is in  $\mathfrak{S}_0$  and is Lip- $(-1/2)$ -modulo tame with the following estimates

$$\begin{aligned} \mathfrak{M}_{\mathfrak{R}_1}^{\sharp,\gamma}(-1/2, s, \mathbf{b}_0) &\lesssim_s \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^\gamma), \\ \mathfrak{M}_{\Delta_{12}\mathfrak{R}_1}^{\sharp}(-1/2, p, \mathbf{b}_0) &\lesssim_p \varepsilon \gamma^{-3} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \end{aligned} \quad (11.37)$$

for some  $\mu = \mu(\nu) > 0$ .

*Proof.* We Taylor expand in  $\varepsilon$  the conjugate  $\mathcal{L}_7$  by applying formula (11.30) (recall (11.13),(11.26)). Thus formula (11.35), with  $\mathbf{Q}_2^{(1)}$  as in (11.36) follows by Lemma 11.14. Estimates (11.37), (11.36) follow by using Lemmata 11.14, 11.11, 11.12.  $\square$

**Step two (order  $\varepsilon^2$ ).** Let us define the matrix  $\mathbf{A}_2$  as

$$(\mathbf{A}_2)_{\sigma',j}^{\sigma',k}(\ell) = \begin{cases} \frac{(\mathbf{Q}_2^{(1)})_{\sigma,j}^{\sigma',k}(\ell)}{i\delta_{\sigma,\sigma',j,k}^{(2)}(\ell)} & \text{if } \delta_{\sigma,\sigma',j,k}^{(2)}(\ell) \neq 0 \\ 0 & \text{if } \delta_{\sigma,\sigma',j,k}^{(2)}(\ell) = 0 \end{cases} \quad (11.38)$$

where  $\mathbf{Q}_2^{(1)}$  is in (11.36) and  $\delta^{(2)}$  is defined in (11.7). We have the following result.

**Lemma 11.16.** *The operator  $\mathbf{A}_2$  is in  $\mathfrak{S}_0$  and 2-almost-diagonal and satisfies*

$$|(\mathbf{A}_2)_{\sigma',j}^{\sigma',k}(\ell)| \leq (C_2(S))^{-1} |(\mathbf{Q}_2^{(1)})_{\sigma,j}^{\sigma',k}(\ell)|, \quad (11.39)$$

with  $C_2(S) > 0$  in (11.10). Moreover it solves the equation

$$ad_{\bar{\omega} \cdot \partial_\varphi + D_0}[\mathbf{A}_2] = \Pi_{Rg(\bar{\omega} \cdot \partial_\varphi + D_0)} \mathbf{Q}_1^{(2)}. \quad (11.40)$$

*Proof.* Estimate (11.39) follows by Lemma 11.3. Thanks to the definition of  $\mathbf{A}_2$  in (11.38) one can check that equation (11.40) is satisfied. The operator  $\mathbf{A}_2$  is almost-diagonal by definition. Recall that  $\mathbf{Q}_2^{(1)}$  is  $\mathfrak{S}_0$ . Thus  $\mathbf{A}_2$  is in  $\mathfrak{S}_0$  because the denominator  $\delta_{\sigma,\sigma',j,k}^{(2)}(\ell)$  satisfies the algebraic relations (11.34) (using Lemmata Lemma 2.23, 2.22).  $\square$

**Lemma 11.17.** *The conjugate of the operator  $\mathcal{L}_7$  in (11.35) has the form*

$$\mathcal{L}_8 := \Upsilon_2 \mathcal{L}_7 \Upsilon_2^{-1} = \omega \cdot \partial_\varphi + D + \varepsilon^2 \llbracket \mathbf{Q}_2^{(1)} \rrbracket + \sum_{i=3}^6 \varepsilon^i \mathbf{Q}_i^{(2)} + \Pi_{\mathfrak{S}}^\perp \mathfrak{R}_2 \quad (11.41)$$

where  $\mathbf{Q}_i^{(2)}$ ,  $i = 3, \dots, 6$ , are in  $\mathfrak{S}_0$ , are  $i$ -almost-diagonal and satisfy

$$|(\mathbf{Q}_i^{(2)})_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C}{\langle j, k \rangle^{1/2}}, \quad |(\mathbf{Q}_i^{(2)})_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C}{\langle j, k \rangle^{\rho-2}}.$$

Moreover  $\mathfrak{R}_2$  is in  $\mathfrak{S}_0$  and is Lip- $(-1/2)$ -modulo tame with the following estimates

$$\begin{aligned} \mathfrak{M}_{\mathfrak{R}_2}^{\sharp, \gamma}(-1/2, s, \mathbf{b}_0) &\lesssim_s \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^\gamma), \\ \mathfrak{M}_{\Delta_{12}\mathfrak{R}_2}^{\sharp}(-1/2, p, \mathbf{b}_0) &\lesssim_p \varepsilon \gamma^{-3} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \end{aligned} \quad (11.42)$$

for some  $\mu = \mu(\nu) > 0$  possibly larger than the one in Lemma 11.15.

*Proof.* One can follow word by word the proof of Lemma 11.15 using that, by Lemma 11.3 (recall (11.16)),

$$\Pi_{Ker(\bar{\omega} \cdot \partial_\varphi + D_0)} \mathbf{Q}_2^{(1)} = \llbracket \mathbf{Q}_2^{(1)} \rrbracket. \quad \square$$

The following proposition allows to compute explicitly the first order corrections of the eigenvalues

$$D_2 + \llbracket \mathbf{Q}_2^{(1)} \rrbracket.$$

**Proposition 11.18. (Identification of normal forms).** *Recall (11.26), (3.31). Consider the symplectic change of coordinates  $\mathbf{A}_\varepsilon$  in (5.7) and let us denote by  $\mathbf{A}_1 := \mathbf{A}_{\varepsilon|\varepsilon=1}$ . Then*

$$D_2 + \llbracket \mathbf{Q}_2^{(1)} \rrbracket = X_H, \quad H := \left[ \Pi_{\text{triv}} \Pi^{d_z=2} \left( H_{FB}^{(4)} \right) \right] \circ \mathbf{A}_1 \Big|_{\substack{y=0 \\ \theta=\varphi}}. \quad (11.43)$$

*Proof.* By Proposition 3.6 the Hamiltonian  $H$  in (11.43) is equal to  $H_{LB}^{(4)} \circ \mathbf{A}_1$  at  $y = 0$ ,  $\theta = \varphi$ , which in turn by (3.28), (3.48) is equivalent to

$$K := \Pi_{\text{triv}} \Pi^{d_z=2} \mathcal{H}^{(4)} \circ \mathbf{A}_1 \Big|_{\substack{y=0 \\ \theta=\varphi}}$$

with  $\mathcal{H}^{(4)}$  in (3.26). Hence we have to prove that

$$D_2 + \llbracket \mathbf{Q}_2^{(1)} \rrbracket = \begin{pmatrix} \partial_{\bar{z}} K \\ -\partial_z K \end{pmatrix}. \quad (11.44)$$

We recall that  $\mathcal{H}^{(4)}$  is the 4-degree Hamiltonian, quadratic in the normal variables, obtained by the *formal* “Weak” plus “Linear” BNF procedure; on the other hand  $D_2 + \llbracket \mathbf{Q}_2^{(1)} \rrbracket$  are the terms of size  $\varepsilon^2$  of  $\mathcal{L}_8$  in (11.41). In turn  $\mathcal{L}_8$  is



the Hamiltonian operator associated to the Hamiltonian, quadratic in the normal variables, obtained by the *rigorous* procedure: weak BNF plus the regularization procedure of section 10 plus two steps of linear BNF. We now show that the two procedures provide the same output, namely the (11.44) holds.

In other words we prove that the Taylor expansion up to degree 2 in  $\varepsilon$  of the operator  $\mathcal{L}_8$  in (11.41) coincides with the vector field of the Hamiltonian  $H_{\mathbb{C}}^{(2)} + H_{LB}^{(4)}$  (see (3.14), (3.28)) written in the action-angle variables defined in (5.7).

We recall that, by construction, the operator  $\mathcal{L}_8$  in (11.41) has the form

$$\mathcal{L}_8 = \mathcal{T} \circ \mathcal{L}_\omega \circ \mathcal{T}^{-1}, \quad (11.45)$$

where  $\mathcal{L}_\omega$  is in (7.4) and the map  $\mathcal{T}$  is the composition

$$\mathcal{T} := \Upsilon_2 \Upsilon_1 \Phi \Psi \Phi_M \Lambda \Phi_B^{-1}, \quad (11.46)$$

where  $\Lambda$  is in (3.4) and the other maps above are given in Propositions 8.2, 8.3, 9.1, 9.5, 10.3 and Lemmata 11.15, 11.17. Now we compute the Taylor expansion of  $\mathcal{L}_8$  in terms of the generators of the maps composing  $\mathcal{T}$  and in terms of the map  $\Phi_B$  of the weak BNF.

First of all, recalling the arguments in section 7, we note that  $\mathcal{L}_\omega$  is the linearized operator in the normal directions of the Hamiltonian  $\mathcal{H}_{\mathbb{C}} = H_{\mathbb{C}} \circ \Phi_B$  (see Proposition 4.3) written in the *real* variables. In particular, recalling subsection 7.2 (see (7.33)), we have in the complex variables

$$\Lambda \mathcal{L}_\omega \Lambda^{-1} = \bar{\omega} \cdot \partial_\varphi \left[ \frac{z}{\bar{z}} \right] + iE|D|^{\frac{1}{2}} \left[ \frac{z}{\bar{z}} \right] + \varepsilon \mathbf{S}_1 \left[ \frac{z}{\bar{z}} \right] + \varepsilon^2 \mathbf{S}_2 \left[ \frac{z}{\bar{z}} \right] + \varepsilon^2 (\mathbb{A}\zeta) \cdot \partial_\varphi \left[ \frac{z}{\bar{z}} \right] + O(\varepsilon^3) \quad (11.47)$$

where

$$\varepsilon^2 (\mathbb{A}\zeta) = \partial_Q (H_2 + H_{\mathcal{R}_2})|_{z=0}, \quad \varepsilon \mathbf{S}_1 \left[ \frac{z}{\bar{z}} \right] = i\varepsilon \begin{pmatrix} \partial_{\bar{z}} H_1 \\ -\partial_z H_1 \end{pmatrix}, \quad \varepsilon^2 \mathbf{S}_2 \left[ \frac{z}{\bar{z}} \right] = i\varepsilon^2 \begin{pmatrix} \partial_{\bar{z}} (H_2 + H_{\mathcal{R}_2}) \\ -\partial_z (H_2 + H_{\mathcal{R}_2}) \end{pmatrix}$$

and  $\mathbf{S}_i = \mathbf{S}_i(v_I)$ ,  $i = 1, 2$  are some  $2 \times 2$  matrix of operators of the form (7.17) with, respectively,  $i = 1, i = 2$ . In (11.47) we denoted by  $O(\varepsilon^3)$  all the terms which are at least cubic in  $v_I, \bar{v}_I$ .

Since the map  $\Phi_B$  coincides, up to degree 2, with  $\Phi_{WB}$ , the Taylor expansion (up to degree 2 in  $\varepsilon$ ) of  $\mathcal{L}_\omega$  coincides with the Hamiltonian vector field, in the normal directions, of the Hamiltonian  $H_{\mathbb{C}} \circ \Phi_{WB}$ . This implies, recalling (3.23), (3.24), that

$$\varepsilon \mathbf{S}_1 \left[ \frac{z}{\bar{z}} \right] \equiv i\varepsilon J \nabla_{(z, \bar{z})} (H_{\mathbb{C}}^{(3,2)} \circ \mathbf{A}_1|_{y=0, \theta=\varphi}), \quad \varepsilon^2 \mathbf{S}_2 \left[ \frac{z}{\bar{z}} \right] := i\varepsilon^2 J \nabla_{(z, \bar{z})} (\widehat{H}_1^{(4)} \circ \mathbf{A}_1|_{y=0, \theta=\varphi}). \quad (11.48)$$

Secondly we note that the map  $\mathcal{T}$  in (11.46) admits a Taylor expansion in  $\varepsilon$ . More precisely we claim that

$$\mathcal{T} = \text{Id} + \varepsilon \mathbf{T}_1 + \varepsilon^2 \mathbf{T}_2 + O(\varepsilon^3), \quad (11.49)$$

where  $\mathbf{T}_i = \mathbf{T}_i(v_I)$  are  $2 \times 2$  matrix of operators of the form (7.17) with, respectively,  $i = 1$  and  $i = 2$ . Indeed the maps  $\Phi_B, \Phi_M, \Psi, \Phi$  appearing in (11.46) has been constructed as the time one flow map of a Hamiltonian PDE whose generator is a pseudo differential operator with symbol in the class  $S_k^m$  for some  $m \in \mathbb{R}$  and  $1 \leq k \leq 6$ . The maps  $\Upsilon_i$ ,  $i = 1, 2$ , are given by (11.29). Hence the (11.49) follows.

By (11.47) and (11.49) (see also (11.26)) we get

$$\begin{aligned} \mathcal{L}_8 &= \bar{\omega} \cdot \partial_\varphi \left[ \frac{z}{\bar{z}} \right] + \mathbf{D}_0 \left[ \frac{z}{\bar{z}} \right] + \varepsilon \mathbf{S}_1 \left[ \frac{z}{\bar{z}} \right] + \varepsilon \left[ \mathbf{T}_1, \bar{\omega} \cdot \partial_\varphi + \mathbf{D}_0 \right] \left[ \frac{z}{\bar{z}} \right] + \varepsilon^2 \left[ \mathbf{T}_1, \mathbf{S}_1 \right] \left[ \frac{z}{\bar{z}} \right] \\ &+ \frac{\varepsilon^2}{2} \left[ \mathbf{T}_1, \left[ \mathbf{T}_1, \bar{\omega} \cdot \partial_\varphi + \mathbf{D}_0 \right] \right] \left[ \frac{z}{\bar{z}} \right] + \varepsilon^2 \mathbf{S}_2 \left[ \frac{z}{\bar{z}} \right] + \varepsilon^2 (\mathbb{A}\zeta) \cdot \partial_\varphi \left[ \frac{z}{\bar{z}} \right] + \varepsilon^2 \left[ \mathbf{T}_2, \bar{\omega} \cdot \partial_\varphi + \mathbf{D}_0 \right] \left[ \frac{z}{\bar{z}} \right] + O(\varepsilon^3). \end{aligned} \quad (11.50)$$

In section 10 and Lemma 11.15 we proved that

$$\mathbf{S}_1 \left[ \frac{z}{\bar{z}} \right] + \left[ \mathbf{T}_1, \bar{\omega} \cdot \partial_\varphi + \mathbf{D}_0 \right] \left[ \frac{z}{\bar{z}} \right] \equiv 0. \quad (11.51)$$

Our aim is to express the operator  $\mathbf{T}$  in terms of the vector field of the Hamiltonian  $F_3^{(3,2)}$  appearing in (3.26). First of all notice that  $F_3^{(3,2)}$  solves (recall (3.7))

$$H_{\mathbb{C}}^{(3,2)} + \{F_3^{(3,2)}, H_{\mathbb{C}}^{(2)}\} = 0 \quad \Leftrightarrow \quad X_{H_{\mathbb{C}}^{(3,2)}} + X_{\{F_3^{(3,2)}, H_{\mathbb{C}}^{(2)}\}} = X_{H_{\mathbb{C}}^{(3,2)}} + [X_{F_3^{(3,2)}}, X_{H_{\mathbb{C}}^{(2)}}] = 0, \quad (11.52)$$

where  $[\cdot, \cdot]$  denotes the *nonlinear* commutator between two vector fields defined as

$$\begin{aligned} X &= \sum_{\sigma=\pm} X^{(u^\sigma)} \partial_{u^\sigma}, \quad Y = \sum_{\sigma=\pm} Y^{(u^\sigma)} \partial_{u^\sigma}, \\ [X, Y] &= \sum_{\sigma=\pm} \left( \sum_{\sigma'} d_{u^{\sigma'}} Y^{(u^\sigma)} [X^{(u^{\sigma'})}] - d_{u^{\sigma'}} X^{(u^\sigma)} [Y^{(u^{\sigma'})}] \right) \partial_{u^\sigma}. \end{aligned}$$

Define the vector field (recall that  $T$  has the form (7.17))

$$R := \sum_{\sigma=\pm} R^\sigma \partial_{z^\sigma}, \quad R^\sigma := \sum_{\sigma'} T_{\sigma'}^{\sigma'} z^{\sigma'}.$$

We define

$$G^{(3,2)} := H_C^{(3,2)} \circ \mathbf{A}_1|_{\substack{y=0 \\ \theta=\varphi}}, \quad \mathcal{F}^{(3,2)} := F_3^{(3,2)} \circ \mathbf{A}_1|_{\substack{y=0 \\ \theta=\varphi}}. \quad (11.53)$$

Then, using that  $H_0 = H_C^{(2)} \circ \mathbf{A}_1$  at  $y = 0, \theta = \varphi$  (recall (7.34)) and the (11.48), we have that equations (11.51), (11.52) read

$$\begin{aligned} X_{G^{(3,2)}} + [R, X_{H_0}] &= 0, \\ X_{\mathcal{F}^{(3,2)}} + [X_{\mathcal{F}^{(3,2)}}, X_{H_0}] &= 0, \end{aligned} \quad (11.54)$$

where the vector fields are computed with respect to the extended symplectic form (7.32).

Due to the absence of 3-waves interactions we have that the operator  $[\cdot, X_{H_0}]$  is invertible. Therefore

$$R \equiv X_{\mathcal{F}^{(3,2)}}. \quad (11.55)$$

The terms  $O(\varepsilon^2)$  of  $\mathcal{L}_8$  in (11.50) are given by the  $O(\varepsilon^2)$ -terms of  $\mathcal{L}_\omega$ , given by (11.48), plus

$$\frac{\varepsilon^2}{2} [\mathbf{T}_1, \mathbf{S}_1] \left[ \frac{z}{z} \right] + \varepsilon^2 \left[ \mathbf{T}_2, \bar{\omega} \cdot \partial_\varphi + \mathbf{D}_0 \right] \left[ \frac{z}{z} \right], \quad (11.56)$$

where we used (11.51). Notice that the second summand belongs to the range of the adjoint action of  $X_{H_0}$ . Using (11.48), (11.53) and (11.55) we note also that the first summand in (11.56) is equal to

$$\frac{\varepsilon^2}{2} [X_{\mathcal{F}^{(3,2)}}, X_{G^{(3,2)}}]. \quad (11.57)$$

By (3.23), (3.26) we have

$$\Pi^{d_z=2} \mathcal{H}^{(4)} = \Pi^{d_z=2} \widehat{H}_1^{(4)} + \frac{1}{2} \{F_3^{(3,2)}, H_C^{(3,2)}\}.$$

By (11.53) we have that the vector field of the Hamiltonian  $\frac{1}{2} \{F_3^{(3,2)}, H_C^{(3,2)}\}$  is equal to the one in (11.57). Hence, by Remark 7.8, we get the (11.44).  $\square$

By Lemma 7.1 in [14] (see also formulæ (7.21), (7.22) therein) we have that the Hamiltonian vector field corresponding to the Hamiltonian function  $H_{FB} := H_C^{(2)} + H_{FB}^{(4)}$  (see (3.14) and (4.8)) is given by (see (3.7))

$$\dot{z}_n = -i\omega_n z_n - \frac{i}{\pi} \left( \sum_{|j| < |n|} j |j| |z_j|^2 \right) n z_n + [R(z)]_n, \quad (11.58)$$

$$[R(z)]_n := -\frac{i}{2\pi} |n|^3 (|z_n|^2 - 2|z_{-n}|^2) z_n + \frac{i}{\pi} \sum_{\substack{|n| < |k_1|, \\ \text{sign}(k_1) = \text{sign}(n)}} |k_1| |n|^2 z_n (|z_{-k_1}|^2 - |z_{k_1}|^2). \quad (11.59)$$

Then it is easy to note that Proposition 11.18 implies

$$\mathbf{D}_2 + \llbracket \mathbf{Q}_2^{(1)} \rrbracket = \begin{bmatrix} \mathbb{Z} & 0 \\ 0 & \mathbb{Z} \end{bmatrix} \quad \mathbf{Z} := \text{diag}_{j \in S^c} (i\kappa_j), \quad \kappa_j = (m_1(\omega) + c_j(\omega)) j, \quad (11.60)$$

where (recall (1.25))

$$c_j := \begin{cases} \frac{1}{\pi} \sum_{\substack{k \in S \\ |k| > |j|}} k (|j| - |k|) \zeta_k, & |j| < \max(S), \\ 0, & |j| \geq \max(S). \end{cases} \quad (11.61)$$

Note that  $Z = \bar{Z}$  since  $c_{-j} = c_j$  for all  $j \in S^c$ . The presence of the finitely many corrections  $c_j j$  makes hard to impose the second Melnikov conditions.

Indeed instead of considering  $\omega \in \mathcal{G}_1^{(0)}$  in (5.22) we should consider more complicated sets like

$$\tilde{\mathcal{G}}_{0,j,k}^{(1)} := \left\{ \omega \in \Omega_\varepsilon : |(\bar{\omega} - \varepsilon^2 m_1(\omega) \mathbf{v}) \cdot \ell + \varepsilon^2 \mathbb{A} \zeta(\omega) \cdot \ell + c_j(\omega) j - c_k(\omega) k| \geq \gamma \langle \ell \rangle^{-\tau}, \forall \ell \in \mathbb{Z}^\nu \setminus \{0\} \right\}$$

for  $j, k$  such that  $|j|, |k| < \max(S)$ . We remark that in contrast to the correction  $m_1(j - k)$ , appearing in the second order Melnikov conditions, we cannot write  $c_j j - c_k k$  as a function of the index  $\ell$  by using the momentum law  $\mathbf{v} \cdot \ell + j - k = 0$ . Then if one wants to prove a twist condition like (5.28) in Lemma 5.3 to prove that the sets  $\tilde{\mathcal{G}}_{0,j,k}^{(1)}$  have large measure, one should have precise information on the  $c_j$ 's. The definition of the  $c_j$ 's depends strongly on the form of the set  $S$ . Since we want to provide the existence of quasi-periodic solutions with an arbitrary number of frequencies (the cardinality of  $S$  is arbitrary) and we want the set  $S$  to have the most general form we prefer to approach this problem in a different way.

To overcome this problem we introduce *rotating coordinates* for a finite number of modes. We consider the map  $\Upsilon_*(\varphi) : H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$  defined as  $\Upsilon_* = \exp(\varepsilon^2 \mathbf{A}_*)$  where

$$\mathbf{A}_*(\varphi) = \begin{bmatrix} \mathbf{A}_* & 0 \\ 0 & \bar{\mathbf{A}}_* \end{bmatrix}, \quad \mathbf{A}_*(\varphi) := \text{diag}_{j \in S^c} (i\alpha_j \cdot \varphi), \quad \alpha_j := \frac{\mathbf{v}^\perp}{\omega \cdot \mathbf{v}^\perp} c_j j, \quad (11.62)$$

where  $\mathbf{v}^\perp \in \mathbb{Z}^\nu$  is orthogonal to  $\mathbf{v}$  and note that  $\mathbf{A}_* = \bar{\mathbf{A}}_*$ . For concreteness we choose (recall (1.27))

$$\mathbf{v}^\perp = (\bar{j}_2, -\bar{j}_1, 0, \dots, 0).$$

In this way it is easy to note that  $|\omega \cdot \mathbf{v}^\perp| \geq C(S)$  for some  $C(S) > 0$  depending only on the set  $S$ . We observe that such change of coordinates differs from the identity for a finite rank operator. Moreover  $\Upsilon_*$  is symplectic and  $x$ -translation invariant thanks to Lemma 2.22 and the fact that  $\mathbf{v} \cdot \alpha_j = 0$  for all  $j \in S^c$ . Since

$$\exp(\varepsilon^2 \mathbf{A}_*) = \begin{bmatrix} \exp(\varepsilon^2 \mathbf{A}_*) & 0 \\ 0 & \exp(\varepsilon^2 \bar{\mathbf{A}}_*) \end{bmatrix}, \quad \exp(\varepsilon^2 \mathbf{A}_*) = \text{diag}_{j \in S^c} (e^{i\varepsilon^2 \alpha_j \cdot \varphi})$$

we have that

$$\|\Upsilon_*^{\pm 1} u - u\|_s \lesssim_s \varepsilon^2 \|u\|_s. \quad (11.63)$$

Recalling the definition of  $\mathbb{D}$  in (10.7) and (11.26) we define

$$\widehat{\mathbb{D}}_2 := Op^W \begin{pmatrix} im_1 \xi & 0 \\ 0 & im_1 \xi \end{pmatrix}, \quad \mathcal{D}_{\geq 4} := \varepsilon^4 \mathbb{D}_4 + \varepsilon^6 \mathbb{D}_6 + \mathbb{D}_{\geq 8} \quad (11.64)$$

We have the following.

**Lemma 11.19. (Rotating coordinates).** *Consider the operator  $\mathcal{L}_8$  in (11.41). We have that*

$$\tilde{\mathcal{L}}_8 := \Upsilon_* \mathcal{L}_8 \Upsilon_*^{-1} = \omega \cdot \partial_\varphi + \mathbb{D}_0 + \varepsilon^2 \widehat{\mathbb{D}}_2 + \mathcal{D}_{\geq 4} + \sum_{i=3}^6 \varepsilon^i (\pi_{\mathbb{C}} \tilde{\mathbf{q}}_i^{(2)} + \pi_{\mathbb{C}}^\perp \mathbf{q}_i^{(2)}) + \Pi_S^\perp \tilde{\mathfrak{R}}_2 \quad (11.65)$$

where  $\mathbb{C}$  is given in (11.18),  $\pi_{\mathbb{C}}^\perp$  is defined in Lemma 11.10,  $\tilde{\mathbf{q}}_i^{(2)}$ ,  $i = 3, \dots, 6$  are in  $\mathfrak{S}_0$  (see Def. 2.26), almost-diagonal (see Def. 11.4) and satisfy

$$\tilde{\mathbf{q}}_i^{(2)} = \pi_{\mathbb{C}} \tilde{\mathbf{q}}_i^{(2)}, \quad |(\tilde{\mathbf{q}}_i^{(2)})_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C_i}{\langle j, k \rangle^{1/2}}, \quad |(\tilde{\mathbf{q}}_i^{(2)})_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C_i}{\langle j, k \rangle^{\rho-2}}, \quad (\tilde{\mathbf{q}}_i^{(2)})_{\sigma,j}^{\sigma,k}(\ell) \neq 0 \Rightarrow |\ell| \leq C, \quad (11.66)$$

for some constants  $C_i, C$ ,  $i = 3, \dots, 6$ , depending on  $S$ . The operator  $\tilde{\mathfrak{R}}_2$  is in  $\mathfrak{S}_0$  and is Lip- $(-1/2)$ -modulo tame with estimates

$$\mathfrak{M}_{\tilde{\mathfrak{R}}_2}^{\sharp, \gamma}(s, \mathbf{b}_0) \lesssim_s \gamma^{-3} (\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad \mathfrak{M}_{\Delta_{12} \tilde{\mathfrak{R}}_2}^\sharp(p) \lesssim_p \varepsilon \gamma^{-3} (1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu} \quad (11.67)$$

for some  $\mu = \mu(\nu) > 0$  possibly larger than the one in Lemma 11.17.

*Proof.* We have that (recall (11.60), (11.64))

$$\Upsilon_* \omega \cdot \partial_\varphi \Upsilon_*^{-1} = \omega \cdot \partial_\varphi - \varepsilon^2 \mathbf{A}_*(\omega) \stackrel{(11.62)}{=} \omega \cdot \partial_\varphi - \varepsilon^2 \mathbb{D}_2 - \varepsilon^2 [\mathbf{Q}_2^{(1)}] + \varepsilon^2 \widehat{\mathbb{D}}_2, \quad (11.68)$$

and, for any  $i = 2, \dots, 6$ ,

$$(\Upsilon_* \mathbf{Q}_i^{(2)} \Upsilon_*^{-1})_{\sigma,j}^{\sigma',k}(\ell) = \sum_{\ell_1, \ell_2} (\Upsilon_*)_{\sigma,j}^{\sigma',k}(\ell_1) (\Upsilon_*^{-1})_{\sigma',k}^{\sigma',k}(\ell_2) (\mathbf{Q}_i^{(2)})_{\sigma,j}^{\sigma',k}(\ell - \ell_1 - \ell_2). \quad (11.69)$$

Using (11.69), (11.62), (11.61),  $\mathbf{C} \gg \max(S)$  and the fact that  $\Upsilon_*, \mathbf{Q}_i^{(2)}$  are  $x$ -translation invariant, we deduce

$$\pi_{\mathbf{C}}^\perp(\Upsilon_* \mathbf{Q}_i^{(2)} \Upsilon_*^{-1}) = \pi_{\mathbf{C}}^\perp \mathbf{Q}_i^{(2)}. \quad (11.70)$$

In order to prove that  $\Upsilon_* \mathfrak{A}_2 \Upsilon_*^{-1}$  is  $(-1/2)$ -modulo tame we shall prove that (recall Definition 2.9) the operator  $R := \langle \partial_\varphi \rangle^{\mathbf{b}_0} \langle D_x \rangle^{\frac{1}{4}} \Upsilon_* \mathfrak{A}_2 \Upsilon_*^{-1} \langle D_x \rangle^{\frac{1}{4}}$  is modulo tame (see Def. 2.8). Recalling (2.13) and reasoning as in (11.69) we get

$$|(R)_{\sigma,j}^{\sigma',k}(\ell)| \lesssim_S \sum_{\ell_1, \ell_2} |(\Upsilon_*)_{\sigma,j}^{\sigma',k}(\ell_1) (\Upsilon_*^{-1})_{\sigma',k}^{\sigma',k}(\ell_2) (\mathfrak{A}_2)_{\sigma,j}^{\sigma',k}(\ell - \ell_1 - \ell_2)| \langle j \rangle^{\frac{1}{4}} \langle k \rangle^{\frac{1}{4}} \langle \ell - \ell_1 - \ell_2 \rangle^{\mathbf{b}_0}, \quad (11.71)$$

where we used that, since  $\Upsilon_*$  is almost diagonal,  $|\ell| \lesssim_S |\ell - \ell_1 - \ell_2|$ . The coefficients in (11.71) are the ones of the operator  $\tilde{R}$  (see Definition 2.5) where

$$\tilde{R} := \Upsilon_* \langle \partial_\varphi \rangle^{\mathbf{b}_0} \langle D_x \rangle^{\frac{1}{4}} \mathfrak{A}_2 \langle D_x \rangle^{\frac{1}{4}} \Upsilon_*^{-1}.$$

Therefore, using (11.42) and (11.63) we deduce the (11.67). Using formula (11.69) it is easy to check that

$$\pi_{\mathbf{C}}(\Upsilon_* \mathbf{Q}_i^{(2)} \Upsilon_*^{-1})$$

satisfies the (11.66) by using that  $\mathbf{Q}_i^{(2)}$  is  $i$ -almost-diagonal and it satisfies the bounds given in Lemma 11.17. All the operators appearing in (11.65) are in  $\mathfrak{S}_0$  since the map  $\Upsilon_*$  is symplectic and  $x$ -translation invariant.  $\square$

**11.3. Terms  $O(\varepsilon^{\geq 3})$ : analysis of high modes.** In this section we shall normalize the terms  $\pi_{\mathbf{C}}^\perp \mathbf{Q}_i^{(2)}$  appearing in (11.65).

**Step three (order  $\varepsilon^3$ ).** Let us define the matrix  $\mathbf{A}_3 = \pi_{\mathbf{C}}^\perp \mathbf{A}_3$  as

$$(\mathbf{A}_3)_{\sigma,j}^{\sigma',k}(\ell) := \begin{cases} (\mathbf{Q}_3^{(2)})_{\sigma,j}^{\sigma',k}(\ell) / i \delta_{\sigma,\sigma',j,k}^{(3)}(\ell), & \max\{|j|, |k|\} \geq \mathbf{C}, \ell \neq 0, |\ell| \leq 3 \\ 0 & \text{otherwise} \end{cases} \quad (11.72)$$

where  $\mathbf{Q}_3^{(2)}$  is in (11.65) and  $\delta_{\sigma,\sigma',j,k}^{(3)}(\ell)$  is defined in (11.7). We have the following result.

**Lemma 11.20.** *The operator  $\mathbf{A}_3$  is in  $\mathfrak{S}_0$  and 3-almost-diagonal and satisfies*

$$|(\mathbf{A}_3)_{\sigma,j}^{\sigma',k}(\ell)| \leq (K(S))^{-1} |(\mathbf{Q}_3^{(2)})_{\sigma,j}^{\sigma',k}(\ell)|, \quad (11.73)$$

with  $K(S) > 0$  in (11.11). Moreover it solves the equation

$$ad_{\overline{\omega} \cdot \partial_\varphi + \mathbf{D}_0}[\mathbf{A}_3] = \pi_{\mathbf{C}}^\perp \mathbf{Q}_3^{(2)}. \quad (11.74)$$

*Proof.* Estimate (11.73) follows by item (ii) in Lemma 11.3 Thanks to the definition of  $\mathbf{A}_3$  in (11.72) one can check that the equation (11.74) is satisfied. Recall that  $\mathbf{Q}_2^{(3)}$  is  $\mathfrak{S}_0$ . By using the expression of  $\delta_{\sigma,\sigma',j,k}^{(3)}(\ell)$  in (11.7) and Lemmata Lemma 2.23, 2.22, one can check that  $\mathbf{A}_3$  is  $\mathfrak{S}_0$ .  $\square$

**Lemma 11.21.** *The conjugate of the operator  $\tilde{\mathcal{L}}_8$  in (11.65) has the form*

$$\mathcal{L}_9 := \Upsilon_3 \tilde{\mathcal{L}}_8 \Upsilon_3^{-1} = \omega \cdot \partial_\varphi + \mathbf{D}_0 + \varepsilon^2 \widehat{\mathbf{D}}_2 + \mathcal{D}_{\geq 4} + \sum_{i=3}^6 \varepsilon^i \pi_{\mathbf{C}} \tilde{\mathbf{Q}}_i^{(3)} + \sum_{i=4}^6 \varepsilon^i \pi_{\mathbf{C}}^\perp \mathbf{Q}_i^{(3)} + \Pi_S^\perp \mathfrak{A}_3 \quad (11.75)$$

where  $\Upsilon_3$  is the map in (11.29) with  $\mathbf{A}_3$  in (11.72),  $\widehat{\mathbf{D}}_2, \mathcal{D}_{\geq 4}$  are in (11.64),  $\tilde{\mathbf{Q}}_i^{(3)} = \pi_{\mathbf{C}} \tilde{\mathbf{Q}}_i^{(3)}$ ,  $i = 3, \dots, 6$  are in  $\mathfrak{S}_0$  (see Def. 2.26), almost-diagonal (see Def. 11.4) and satisfy

$$|(\tilde{\mathbf{Q}}_i^{(3)})_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C_i}{\langle j, k \rangle^{1/2}}, \quad |(\tilde{\mathbf{Q}}_i^{(3)})_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C_i}{\langle j, k \rangle^{\rho-3}}, \quad (\tilde{\mathbf{Q}}_i^{(3)})_{\sigma,j}^{\sigma,k}(\ell) \neq 0 \Rightarrow |\ell| \leq C, \quad (11.76)$$

for some constants  $C_i, C$ ,  $i = 3, \dots, 6$ , depending on  $S$ . Moreover the operators  $Q_i^{(3)}$ ,  $i = 4, \dots, 6$ , are in  $\mathfrak{S}_0$ , are  $i$ -almost-diagonal and satisfy the estimates (11.76). Finally the operator  $\mathfrak{R}_3$  is in  $\mathfrak{S}_0$  and is  $Lip(-1/2)$ -modulo tame with the following estimates

$$\mathfrak{M}_{\mathfrak{R}_3}^{\sharp, \gamma}(-1/2, s, \mathbf{b}_0) \lesssim_s \gamma^{-3}(\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad (11.77)$$

$$\mathfrak{M}_{\Delta_{12}\mathfrak{R}_3}^{\sharp}(-1/2, p, \mathbf{b}_0) \lesssim_p \varepsilon \gamma^{-3}(1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \quad (11.78)$$

for some  $\mu = \mu(\nu) > 0$  possibly larger than the one in Lemma 11.19.

*Proof.* We Taylor expand in  $\varepsilon$  the conjugate  $\mathcal{L}_9$  by applying formula (11.30) (recall the expansion in  $\varepsilon$  in (11.65)). Thus formula (11.75) follows by Lemma 11.20, (11.74). Estimates (11.77), (11.78), (11.76) follow by using Lemmata 11.20, 11.11, 11.12 and bounds (11.67).  $\square$

In the following lemma we collect the result of conjugating the operator  $\mathcal{L}_9$  in (11.75) by three maps of the same form of  $\Upsilon_3$  whose aim is to normalize the terms of order  $\varepsilon^i$  with  $i = 4, 5, 6$  such that  $\pi_c^\perp$  leaves them invariant.

**Lemma 11.22.** *For any  $\omega \in \Omega_\infty^{2\gamma}$  there exists a map  $\Upsilon_{\geq 4}(\omega): H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$  such that*

$$\mathcal{L}_{10} := \Upsilon_{\geq 4} \mathcal{L}_9 \Upsilon_{\geq 4}^{-1} = \omega \cdot \partial_\varphi + \mathcal{D}_0 + \varepsilon^2 \widehat{\mathcal{D}}_2 + \mathcal{D}_{\geq 4} + \pi_c \mathcal{Q} + \varepsilon^4 \llbracket \pi_c^\perp Q_4^{(4)} \rrbracket + \varepsilon^6 \llbracket \pi_c^\perp Q_6^{(4)} \rrbracket + \Pi_S^\perp \mathfrak{R}_4 \quad (11.79)$$

where  $\widehat{\mathcal{D}}_2, \mathcal{D}_{\geq 4}$  are in (11.64),  $\mathcal{Q} = \pi_c \mathcal{Q}$  is in  $\mathfrak{S}_0$  (see Def. 2.26), almost-diagonal (see Def. 11.4) and satisfies

$$|(\mathcal{Q})_{\sigma, j}^{\sigma, k}(\ell)| \leq \frac{C_i \varepsilon^3}{\langle j, k \rangle^{1/2}}, \quad |(\mathcal{Q})_{\sigma, j}^{-\sigma, k}(\ell)| \leq \frac{C_i \varepsilon^3}{\langle j, k \rangle^{\rho-6}}, \quad (\mathcal{Q})_{\sigma, j}^{\sigma, k}(\ell) \neq 0 \Rightarrow |\ell| \leq C_{\mathcal{Q}}, \quad (11.80)$$

for some constants  $C_i, C_{\mathcal{Q}}$ ,  $i = 3, \dots, 6$ , depending on  $S$ . Finally the operator  $\mathfrak{R}_4$  is in  $\mathfrak{S}_0$  and is  $Lip(-1/2)$ -modulo tame with the following estimates

$$\mathfrak{M}_{\mathfrak{R}_4}^{\sharp, \gamma}(-1/2, s, \mathbf{b}_0) \lesssim_s \gamma^{-3}(\varepsilon^{13} + \varepsilon \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}), \quad (11.81)$$

$$\mathfrak{M}_{\Delta_{12}\mathfrak{R}_4}^{\sharp}(-1/2, p, \mathbf{b}_0) \lesssim_p \varepsilon \gamma^{-3}(1 + \|\mathfrak{J}_\delta\|_{p+\mu}) \|i_1 - i_2\|_{p+\mu}, \quad (11.82)$$

for some  $\mu = \mu(\nu) > 0$  possibly large than the one in Lemma 11.19.

*Proof.* The map  $\Upsilon_{\geq 4}$  is the composition of three maps  $\Upsilon_4, \Upsilon_5, \Upsilon_6$  defined in the following way: the  $\Upsilon_i$ ,  $i = 4, 5, 6$ , are the time one flow map of a Hamiltonian system with generator

$$(\mathbf{A}_i)_{\sigma, j}^{\sigma', k}(\ell) := \begin{cases} (Q_i^{(i-1)})_{\sigma, j}^{\sigma', k}(\ell) / i \delta_{\sigma, \sigma', j, k}^{(i)}(\ell), & \max\{|j|, |k|\} \geq \mathbb{C}, \ell \neq 0, |\ell| \leq i \\ 0 & \text{otherwise,} \end{cases} \quad (11.83)$$

where  $Q_4^{(3)}$  is in (11.75),  $Q_5^{(4)} = \pi_{c^\perp} Q_5^{(4)}$  is the term of order  $\varepsilon^5$  in the Taylor expansion of  $\Upsilon_4 \mathcal{L}_9 \Upsilon_4^{-1}$  and  $Q_6^{(5)} = \pi_{c^\perp} Q_6^{(5)}$  is the term of order  $\varepsilon^6$  in the Taylor expansion of  $\Upsilon_5 \Upsilon_4 \mathcal{L}_9 \Upsilon_4^{-1} \Upsilon_5^{-1}$ .

Then the proof follows by following the same proof of Lemma 11.21 for each conjugation through the maps  $\Upsilon_i$ ,  $i = 4, 5, 6$ .  $\square$

**11.4. Terms  $O(\varepsilon^{\geq 3})$ : analysis of low modes.** In this subsection we conclude the proof of Proposition 11.1. In particular we normalize the term  $\pi_c \mathcal{Q}$  in (11.79), which is finite rank *but* not perturbative for the KAM scheme.

**Lemma 11.23.** *Recall the constant  $\mathbb{C}$  in Lemma 11.10 and set  $\mathbb{C}_2 := \mathbb{C}$ . There exists a constant  $\mathbb{C}_1 > 0$  such that for any  $\omega \in \mathcal{G}_0^{(2)}(\mathbb{C}_1, \mathbb{C}_2)$  (see (5.23)) there exists a map  $\Psi(\varphi): H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$  such that*

$$\mathcal{L}_{14} := \Psi \mathcal{L}_{10} \Psi^{-1} = \omega \cdot \partial_\varphi + \mathcal{D} + \Pi_S^\perp \mathfrak{R} \quad (11.84)$$

where  $\mathcal{D}$  is given in (11.2). The operator  $\mathfrak{R}$  is in  $\mathfrak{S}_0$  and is  $Lip(-1/2)$ -modulo tame. Moreover  $\mathfrak{R}$  satisfies the estimates (11.5).

*Proof.* The aim of the procedure is to reduce the size of the remainder  $\mathcal{Q}$  which has the property  $\pi_c \mathcal{Q} = \mathcal{Q}$  up to diagonal (in space and time) terms. This is done iteratively by applying a finite number (five) of changes of coordinates. Such maps will be constructed as flows of operators  $\mathfrak{A}_i$  such that  $\pi_c \mathfrak{A}_i = \mathfrak{A}_i$ . As a consequence the flow map  $\Psi_i = \exp(\mathfrak{A}_i)$  leaves invariant the diagonal operators  $R$  such that  $\pi_c^\perp R = R$  (see (11.89)). Then we set  $\Psi = \Psi_5 \circ \dots \circ \Psi_1$ . The choice of  $\mathfrak{A}_i$  requires to solve an homological equation involving, at each step, only the adjoint action of the operator

$$(\bar{\omega} + \varepsilon^2 \mathbb{A} \zeta) \cdot \partial_\varphi + \mathcal{D}_0 + \varepsilon^2 \widehat{\mathcal{D}}_2$$

(see (11.88)). We remark that the adjoint action of the above operator acting on  $x$ -translation invariant operators is diagonal with eigenvalues

$$i(\bar{\omega} + \varepsilon^2 \mathbb{A}\zeta \cdot \ell) + i\sigma\sqrt{|j|} - i\sigma'\sqrt{|k|} + im_1(j-k) = i((\bar{\omega} + \varepsilon^2(\mathbb{A}\zeta - m_1 \mathbf{v})) \cdot \ell + \sigma\sqrt{|j|} - \sigma'\sqrt{|k|})$$

for  $\ell \in \mathbb{Z}^\nu$ ,  $j, k \in S^c$  with  $\mathbf{v} \cdot \ell + j - k = 0$ . Since  $\omega$  is chosen in  $\mathcal{G}_0^{(2)}(\mathcal{C}_1, \mathcal{C}_2)$  (see (5.23)) the kernel of the adjoint action, restricted to  $\mathcal{C}_1$ -almost diagonal operators, contains just diagonal matrices. By the (11.80) the size of these new diagonal corrections will be  $O(\varepsilon^3)$ . As a consequence, at each step, the size of the remainder we want to cancel out will be  $O(\varepsilon^{3+3k}\gamma^{-k})$ . It is easy to check that for  $k = 5$  the size of the last remainder is less than  $\varepsilon^{13}\gamma^{-3}$ , i.e. it is comparable with the remainder  $\mathfrak{R}_4$  which is already *perturbative* for the KAM scheme.

We now provide the proof of the first step. The others will follow in the same way. The first map is defined as the time one flow map of the Hamiltonian generated by (recall (11.79))

$$(\mathfrak{A}_1)_{\sigma,j}^{\sigma',k}(\ell) := \begin{cases} \frac{(\mathfrak{Q})_{\sigma,j}^{\sigma',k}(\ell)}{i((\bar{\omega} + \varepsilon^2(\mathbb{A}\zeta - m_1 \mathbf{v})) \cdot \ell + \sigma\sqrt{|j|} - \sigma'\sqrt{|k|})} & \text{if } (\ell, j, k) \neq (0, j, j), \\ 0 & \text{otherwise.} \end{cases} \quad (11.85)$$

We observe that  $\mathfrak{A}_1$  shares the same properties of  $\mathfrak{Q}$ , namely it is  $C_\Omega$ -almost-diagonal (recall (11.80)) and it belongs to  $\mathfrak{S}_0$ . By (11.80) and by the fact that  $\omega \in \mathcal{G}_0^{(2)}(\mathcal{C}_1, \mathcal{C}_2)$ , by choosing  $\mathcal{C}_1 > C_\Omega$ , we have that

$$|(\mathfrak{A}_1)_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C_i \varepsilon^3 \gamma^{-1}}{\langle j, k \rangle^{1/2}}, \quad |(\mathfrak{A}_1)_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C_i \varepsilon^3 \gamma^{-1}}{\langle j, k \rangle^{\rho-6}}. \quad (11.86)$$

Then by item (i) of Lemma 11.11 we have that  $\mathfrak{A}_1$  is  $(-1/2)$ -Lip-modulo tame with

$$\mathfrak{M}_{\mathfrak{A}_1}^{\sharp, \gamma}(-1/2, s, \mathbf{b}_0) \lesssim_s \varepsilon^3 \gamma^{-1} \stackrel{(5.21), (5.6)}{\lesssim_s} \varepsilon^{1-a}. \quad (11.87)$$

Note that, thanks to (11.85), the operator  $\mathfrak{A}_1$  satisfies (recall (11.16))

$$\pi_C \mathfrak{A}_1 = \mathfrak{A}_1, \quad \text{ad}_{(\bar{\omega} + \varepsilon^2 \mathbb{A}\zeta) \cdot \partial_\varphi + \mathbb{D}_0 + \varepsilon^2 \widehat{\mathbb{D}}_2} [\mathfrak{A}_1] = \mathfrak{Q} - \llbracket \mathfrak{Q} \rrbracket. \quad (11.88)$$

By Definition 11.8 we also note that, given an operator  $R$ ,

$$[\mathfrak{A}_1, \pi_C R] = \pi_C [\mathfrak{A}_1, R], \quad [\mathfrak{A}_1, \llbracket \pi_C^\perp R \rrbracket] = 0. \quad (11.89)$$

Setting  $\Psi_1 = \exp(\mathfrak{A}_1)$ , we remark that the conjugate  $\Psi_1 \mathfrak{R}_4 \Psi_1^{-1}$  is Lip- $(-1/2)$ -modulo tame and satisfies estimates like (11.5) by (11.86), (11.87), (11.81), (11.82) and using Lemma 11.11. Reasoning as in item (ii) of Lemma 11.12 we deduce that (recall (11.79), (11.64))

$$\text{ad}_{\mathfrak{A}_1}(\mathcal{D}_{\geq 4} + \mathfrak{Q}) = B$$

is almost-diagonal and

$$|B_{\sigma,j}^{\sigma,k}(\ell)| \leq \frac{C \varepsilon^6 \gamma^{-1}}{\langle j, k \rangle^{1/2}}, \quad |B_{\sigma,j}^{-\sigma,k}(\ell)| \leq \frac{C \varepsilon^6 \gamma^{-1}}{\langle j, k \rangle^{\rho-7}}, \quad (11.90)$$

where we used estimates (11.86), (11.80) and Lemma 11.12. Here  $C > 0$  is some constant depending on  $S$  possibly larger than  $C_\Omega$ . Therefore, applying Lemmata 11.11, 11.12 and using the Lie expansion (11.30), the (11.88) and the (11.89), we get

$$\mathcal{L}_{11} := \Psi_1 \mathcal{L}_{10} \Psi_1^{-1} = \omega \cdot \partial_\varphi + \mathbb{D}_0 + \varepsilon^2 \widehat{\mathbb{D}}_2 + \mathcal{D}_{\geq 4} + \llbracket \pi_C \mathfrak{Q} \rrbracket + \pi_C \mathfrak{Q}_2 + \varepsilon^4 \llbracket \pi_C^\perp \mathfrak{Q}_4^{(4)} \rrbracket + \varepsilon^6 \llbracket \pi_C^\perp \mathfrak{Q}_6^{(4)} \rrbracket + \mathfrak{R}_5 \quad (11.91)$$

where

$$\pi_C \mathfrak{Q}_2 = \mathfrak{Q}_2 := \sum_{k=1}^4 \frac{1}{k!} \text{ad}_{\mathfrak{A}_1}^k(\mathcal{D}_{\geq 4} + \mathfrak{Q}) + \sum_{k=2}^4 \frac{1}{k!} \text{ad}_{\mathfrak{A}_1}^k(\omega \cdot \partial_\varphi + \mathbb{D}_0 + \varepsilon^2 \widehat{\mathbb{D}}_2).$$

Moreover the remainder  $\mathfrak{R}_5$  is Lip- $(-1/2)$ -modulo tame and satisfies estimates like (11.5), the operator  $\mathfrak{Q}_2$  satisfies estimates (11.90). In particular  $\mathfrak{Q}_2, \mathfrak{R}_5$  are Hamiltonian and  $x$ -translation invariant. The remainder  $\mathfrak{Q}_2$  is the one that will be normalized in the second step. One iterates this procedure by reasoning as in the step above, taking  $\omega \in \mathcal{G}_2^{(0)}(\mathcal{C}_1, \mathcal{C}_2)$  where  $\mathcal{C}_1$  is chosen large enough in such a way that the terms we have to normalize are again  $\mathcal{C}_1$ -almost diagonal. After five steps one gets the (11.84) for some remainder  $\mathfrak{R}$ , with the same properties of  $\mathfrak{R}_5$ , and where

$$\mathfrak{D} := \mathbb{D}_0 + \varepsilon^2 \widehat{\mathbb{D}}_2 + \mathcal{D}_{\geq 4} + \llbracket \pi_C \widetilde{\mathfrak{Q}} \rrbracket + \varepsilon^4 \llbracket \pi_C^\perp \mathfrak{Q}_4^{(4)} \rrbracket + \varepsilon^6 \llbracket \pi_C^\perp \mathfrak{Q}_6^{(4)} \rrbracket$$



for some  $\tilde{\mathcal{Q}}$   $C_1$ -almost diagonal and satisfying (11.80). By Definition 11.8, (11.64) and (11.26) one has that  $\mathfrak{D} = \text{diag}(id_j)$  where  $d_j$  have the form (11.3).  $\square$

**Proof of Proposition 11.1.** Let us define (recall (11.29))

$$\Upsilon := \Psi \circ \Upsilon_6 \circ \Upsilon_5 \circ \Upsilon_4 \circ \Upsilon_3 \circ \Upsilon_* \circ \Upsilon_2 \circ \Upsilon_1$$

Then the result follows by Lemmata 11.15, 11.17, Proposition 11.18, Lemmata 11.19, 11.21, 11.22 and 11.23.  $\square$

## 12. INVERSION OF THE LINEARIZED OPERATOR

In this section we conclude the inversion of the linearized operator in the normal directions. We diagonalize the operator  $\mathcal{L}_{14}$  in (11.1) through a KAM reducibility scheme. Then we will prove the inversion assumption 6.15.

We state the KAM reducibility theorem. We remark that the small divisors in (12.4) do not lose derivatives as in [6]. Hence the proof shall follow the same line of Theorem 7.13 in [32]. The only difference relies on the fact that the KAM scheme has to run into the class of  $x$ -translation invariant matrices. However it is easy to see that this is a direct consequence of the fact that the remainder  $\Pi_S^{\perp} \mathfrak{R} \in \mathfrak{S}_0$  is  $x$ -translation invariant since, by induction, each KAM transformation turns out to be  $x$ -translation invariant.

**Theorem 12.1. (Reducibility).** Fix  $\tilde{\gamma} \in [\gamma^3/4, 4\gamma^3]$  and  $\tau = 3\nu + 7$ ,  $b_0 := 6\tau + 6$ . Assume that  $\omega \mapsto i_\delta(\omega)$  is a Lipschitz function defined on  $\mathcal{O}_0 \Subset \Omega_\varepsilon$  (recall (6.3)), satisfying (7.1) with  $\mathfrak{p}_1 \geq \mu$  where  $\mu = \mu(\nu)$  is given in Proposition 11.1. There exist  $\delta_0 \in (0, 1)$ ,  $N_0 > 0$ ,  $C_0 > 0$ , such that, if (recall that by (5.6), (5.21)  $\gamma = \varepsilon^{2+\alpha}$ )

$$N_0^{C_0} \varepsilon^{13} \gamma^{-6} \leq \delta_0, \quad (12.1)$$

then the following holds.

(i) (**Eigenvalues**). For all  $\omega \in \Omega_\varepsilon$  there exists a sequence <sup>8</sup>

$$d_j^\infty(\omega) := d_j(\omega, i_\delta(\omega)) + r_j^{(\infty)}(\omega, i_\delta(\omega)), \quad j \in S^c, \quad (12.2)$$

with  $d_j$  given in (11.3) Furthermore, for all  $j \in S^c$

$$\sup_j \langle j \rangle^{\frac{1}{2}} |r_j^{(\infty)}| \gamma^3 \lesssim \varepsilon^{13} \gamma^{-3}, \quad r_j^{(\infty)} \in \mathbb{R}. \quad (12.3)$$

(ii) (**Conjugacy**). For all  $\omega$  in the set (see (10.5))

$$S_\infty^{2\gamma^3} := S_\infty^{2\gamma^3}(i_\delta) := \left\{ \omega \in \Omega_\infty^{2\gamma} : |\omega \cdot \ell + \sigma d_{\sigma j}^\infty(\omega) - \sigma' d_{\sigma' k}^\infty(\omega)| \geq \frac{2\gamma^3}{\langle \ell \rangle^\tau}, \right. \\ \left. \mathbf{v} \cdot \ell + j - k = 0, \forall \ell \in \mathbb{Z}^\nu, \forall j, k \in S^c, j \neq k, \sigma, \sigma' = \pm \right\} \quad (12.4)$$

there is a real-to-real, bounded, invertible, linear operator  $\Phi_\infty(\omega): H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T}) \rightarrow H_{S^\perp}^s(\mathbb{T}) \times H_{S^\perp}^s(\mathbb{T})$ , with bounded inverse  $\Phi_\infty^{-1}(\omega)$ , that conjugates  $\mathcal{L}_{14}$  in (11.1) to constant coefficients, namely

$$\mathcal{L}_\infty(\omega) := \Phi_\infty(\omega) \circ \mathcal{L}_{14} \circ \Phi_\infty^{-1}(\omega) = \omega \cdot \partial_\varphi + \mathfrak{D}_\infty(\omega), \quad (12.5) \\ \mathfrak{D}_\infty := ((\mathfrak{D}_\infty)_\sigma^{\sigma'})_{\sigma, \sigma' = \pm}, \quad (\mathfrak{D}_\infty)_\sigma^{-\sigma} \equiv 0, \quad \overline{(\mathfrak{D}_\infty)_+^\pm} = (\mathfrak{D}_\infty)_-^\pm, \quad (\mathfrak{D}_\infty)_+^\pm := \text{diag}_{j \in S^c} \{id_j^\infty(\omega)\}.$$

The transformations  $\Phi_\infty, \Phi_\infty^{-1}$  are tame and they satisfy for  $s_0 \leq s \leq \mathcal{S}$

$$\|(\Phi_\infty^{\pm 1} - \text{I})h\|_s^{\gamma^3, S_\infty^{2\gamma^3}} \lesssim_s (\varepsilon^{13} \gamma^{-6} + \varepsilon \gamma^{-6} \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0}) \|h\|_{s_0}^{\gamma^3, S_\infty^{2\gamma^3}} + \varepsilon^{13} \gamma^{-6} \|h\|_s^{\gamma^3, S_\infty^{2\gamma^3}}. \quad (12.6)$$

Moreover  $\Phi_\infty, \Phi_\infty^{-1}$  are symplectic and  $x$ -translation invariant, and  $\mathcal{L}_\infty$  is a Hamiltonian and  $x$ -translation invariant operator.

(iii) (**Dependence on  $i_\delta(\omega)$** ). Let  $i_1(\omega)$  and  $i_2(\omega)$  be two Lipschitz maps satisfying (7.1) with  $\mathfrak{J}_\delta \rightsquigarrow i_k(\varphi) - (\varphi, 0, 0)$ ,  $k = 1, 2$ , and such that

$$\|i_1 - i_2\|_{s_0+\mu} \lesssim \rho N^{-(\tau+1)} \quad (12.7)$$

<sup>8</sup>Whenever it is not strictly necessary, we shall drop the dependence on  $i_\delta$ .

for  $N$  sufficiently large and  $0 < \rho < \gamma^3/4$ . Fix  $\gamma_1 \in [\gamma^3/2, 2\gamma^3]$  and  $\gamma_2 := \gamma_1 - \rho$ . Let  $r_j^{(\infty)}(\omega, i_k(\omega))$  be the sequence in (12.2) with  $\tilde{\gamma} \rightsquigarrow \gamma_k$  for  $k = 1, 2$ . Then for all  $\omega \in S_{\infty}^{\gamma_1}(i_1)$  we have, for some  $\kappa_* > (3/2)\tau$ ,

$$\sup_j \langle j \rangle^{1/2} |\Delta_{12} r_j^{(\infty)}| \leq \varepsilon \gamma^{-3} \|i_1 - i_2\|_{s_0+\mu} + \varepsilon^{13} \gamma^{-3} N^{-\kappa_*}. \quad (12.8)$$

We are in position to give estimates on the inverse of the operator  $\mathcal{L}_\omega$  in (7.4) and give the proof of the inversion assumption (6.15) Let us now define (recall (5.21))

$$\mathcal{F}_{\infty}^{2\gamma}(i_\delta) := \{\omega \in \mathcal{O}_0 : |\omega \cdot \ell + \sigma d_{\sigma j}^{\infty}(\omega)| \geq \frac{2\gamma}{\langle \ell \rangle^\tau}, \mathbf{v} \cdot \ell + j = 0, \forall \ell \in \mathbb{Z}^\nu, \forall j \in S^c, \sigma = \pm\}. \quad (12.9)$$

We deduce the inversion assumption (6.15) by the following result.

**Proposition 12.2. (Invertibility).** *Assume the hypotheses of Theorem 12.1, (7.1) with  $\mathfrak{p}_1 \geq \mu + 2\tau + 1$ , where  $\mu$  is given in Proposition 11.1. Then for all  $\omega \in \mathcal{O}_\infty := S_{\infty}^{2\gamma^3}(i_\delta) \cap \mathcal{F}_{\infty}^{2\gamma}(i_\delta)$  (see (12.4)), for any  $G := (g, \bar{g}) \in \left(H_{S^\perp}^{s+2\tau+1} \cap S_\nu\right)^2$  (see (2.46)) the equation  $\mathcal{L}_\omega h = G$  has a solution  $h = \mathcal{L}_\omega^{-1} G \in \left(H_{S^\perp}^s \cap S_\nu\right)^2$ , satisfying*

$$\|\mathcal{L}_\omega^{-1} G\|_s^{\gamma, \mathcal{O}_\infty} \lesssim_s \gamma^{-1} (\|G\|_{s+2\tau+1}^{\gamma, \mathcal{O}_\infty} + \varepsilon \gamma^{-6} \|\mathfrak{J}_\delta\|_{s+\mathfrak{p}_1}^{\gamma, \mathcal{O}_0} \|G\|_{s_0}^{\gamma, \mathcal{O}_\infty}). \quad (12.10)$$

*Proof.* By Propositions 8.2, 8.3, 9.1, 9.5, 10.3, 11.1 and Theorem 12.1 we have that, recall (7.4),

$$\mathcal{L}_\infty = \chi \mathcal{L}_\omega \chi^{-1}, \quad \chi := \Phi_\infty \circ \Upsilon \circ \Phi \circ \Psi \circ \Phi_{\mathbb{M}} \circ \Lambda \circ \Phi_{\mathbb{B}}^{-1}. \quad (12.11)$$

Moreover, by (8.3), (9.6), (9.13), (10.9), (11.6) and (12.6) we have the following estimates

$$\|\chi^{\pm 1} h\|_s^{\gamma, \mathcal{O}_\infty} \lesssim_s \|h\|_s^{\gamma, \mathcal{O}_\infty} + \varepsilon \gamma^{-6} \|\mathfrak{J}_\delta\|_{s+\mu}^{\gamma, \mathcal{O}_0} \|h\|_{s_0}^{\gamma, \mathcal{O}_\infty}. \quad (12.12)$$

We have

$$\mathcal{L}_\infty^{-1} G = \left( \begin{array}{c} \sum_{j \in S^c, \ell \in \mathbb{Z}^\nu} \frac{g_{\ell j}}{i(\omega \cdot \ell + d_j^\infty(\omega))} e^{i(\ell \cdot \varphi + jx)} \\ \sum_{j \in S^c, \ell \in \mathbb{Z}^\nu} \frac{\bar{g}_{\ell j}}{i(\omega \cdot \ell - d_j^\infty(\omega))} e^{-i(\ell \cdot \varphi + jx)} \end{array} \right) \stackrel{(12.9)}{\Rightarrow} \|\mathcal{L}_\infty^{-1} G\|_s^{\gamma, \mathcal{O}_\infty} \lesssim_s \gamma^{-1} \|G\|_{s+2\tau+1}^{\gamma, \mathcal{O}_\infty}. \quad (12.13)$$

Thus, by (12.12), we get the estimate (12.10). Recalling (2.55) we have that if  $g \in S_\nu$  then  $g_{\ell j} \neq 0$  implies  $\mathbf{v} \cdot \ell + j = 0$ . Hence  $\mathcal{L}_\infty^{-1} G$  belongs to  $S_\nu$ . Finally we deduce that  $h$  is in  $S_\nu$  since the map  $\chi$  in (12.12) is  $x$ -translation invariant (see (2.51)).  $\square$

### 13. THE NASH-MOSER NONLINEAR ITERATION

In this section we prove Theorem 5.6. It will be a consequence of the Nash-Moser theorem 13.2. Consider the finite-dimensional subspaces

$$E_n := \{\mathfrak{J}(\varphi) = (\Theta(\varphi), y(\varphi), W(\varphi, x)) \in H^{s_0}(\mathbb{T}^\nu; \mathbb{R}^\nu) \times H^{s_0}(\mathbb{T}^\nu; \mathbb{R}^\nu) \times H_{S^\perp}^{s_0} : \Theta = \Pi_n \Theta, y = \Pi_n y, W = \Pi_n W\}$$

where

$$N_n := N_0^{X^n}, \quad n = 0, 1, 2, \dots, \quad \chi := 3/2, \quad N_0 > 0$$

and  $\Pi_n$  are the projectors

$$\Pi_n \Theta(\varphi) := \sum_{|\ell| < N_n} \Theta_\ell e^{i\ell \cdot \varphi}, \quad \Pi_n y(\varphi) := \sum_{|\ell| < N_n} y_\ell e^{i\ell \cdot \varphi}, \quad \Pi_n W(\varphi, x) := \sum_{|(\ell, j)| < N_n} W_{\ell j} e^{i(\ell \cdot \varphi + jx)},$$

where

$$\Theta(\varphi) = \sum_{\ell \in \mathbb{Z}^\nu} \Theta_\ell e^{i\ell \cdot \varphi}, \quad y(\varphi) = \sum_{\ell \in \mathbb{Z}^\nu} y_\ell e^{i\ell \cdot \varphi}, \quad W(\varphi, x) = \sum_{\ell \in \mathbb{Z}^\nu, j \in S^c} W_{\ell j} e^{i(\ell \cdot \varphi + jx)}.$$

**Remark 13.1.** *We point out that if the embedding  $\mathfrak{J}(\varphi) := (\Theta(\varphi), y(\varphi), W(\varphi))$  is traveling, i.e. it satisfies (5.31), then  $\Pi_n \mathfrak{J} := (\Pi_n \Theta, \Pi_n y, \Pi_n W)$  is traveling too. Indeed the (5.31) is equivalent to say that each component of  $\mathfrak{J}$  satisfies a linear PDE with constant coefficients (both in space and time). The action of these linear operators is diagonal on the monomials. Hence the projection on some Fourier modes of a solution of such PDEs is still a solution of the same.*

We define  $\Pi_n^\perp := I - \Pi_n$ . The classical smoothing properties hold, namely, for all  $\alpha, s \geq 0$ ,

$$\|\Pi_n \mathfrak{J}\|_{s+\alpha}^{\gamma, \mathcal{O}} \leq N_n^\alpha \|\mathfrak{J}_\delta\|_s^{\gamma, \mathcal{O}}, \quad \forall \mathfrak{J}(\omega) \in H^s, \quad \|\Pi_n^\perp \mathfrak{J}\|_s^{\gamma, \mathcal{O}} \leq N_n^{-\alpha} \|\mathfrak{J}\|_{s+\alpha}^{\gamma, \mathcal{O}}, \quad \forall \mathfrak{J}(\omega) \in H^{s+\alpha}.$$

Recall (5.21), (5.6) for the definition of  $b$  we set  $a := 2b - 2$ . We define the following constants

$$\begin{aligned} \alpha_0 &:= 3\mu_1 + 3, & \alpha &:= 3\alpha_0 + 1, & \alpha_1 &:= (\alpha - 3\mu_1)/2, \\ k &:= 3(\alpha_0 + \rho^{-1}) + 1, & \beta_1 &:= 6\alpha_0 + 3\rho^{-1} + 3, & \frac{1}{2} \left( \frac{1 - 8a}{C_1(1+a)} \right) &< \rho < \frac{1 - 8a}{C_1(1+a)} \end{aligned} \quad (13.1)$$

where  $\mu_1 := \mu_1(\nu) > 0$  is the ‘‘loss of regularity’’ given by the Theorem 6.8 and  $C_1$  is fixed below.

**Theorem 13.2. (Nash-Moser).** *Let  $\tau := 3\nu + 7$  and recall the functional in (5.37). Then there exist  $C_1 > \max\{\alpha_0 + \alpha, C_0\}$  (where  $C_0 := C_0(\nu)$  is the one in Theorem 12.1),  $\delta_0 := \delta_0(\nu) > 0$  such that, if*

$$N_0^{C_1} \varepsilon^{b_*+1} \gamma^{-7} < \delta_0, \quad \gamma := \varepsilon^{2+a} = \varepsilon^{2b}, \quad N_0 := (\varepsilon \gamma^{-1})^\rho, \quad b_* = 16 - 2b, \quad (13.2)$$

then there exists  $C_* = C_*(S) > 0$  such that for all  $n \geq 0$  the following holds:

(P1)<sub>n</sub> there exists a function  $(\mathfrak{J}_n, \Xi_n): \mathcal{G}_n \subseteq \Omega_\varepsilon \rightarrow E_{n-1} \times \mathbb{R}^\nu, \omega \mapsto (\mathfrak{J}_n(\omega), \Xi_n(\omega)), (\mathfrak{J}_0, \Xi_0) := (0, 0), E_{-1} := \{0\}$ , where the set  $\mathcal{G}_0$  is defined in (5.25) (with  $C_1, C_2$  given in Proposition 11.1) and the sets  $\mathcal{G}_n$  for  $n \geq 1$  are defined inductively by:

$$\begin{aligned} \mathcal{G}_{n+1} &:= \Lambda_{n+1}^{(0)} \cap \Lambda_{n+1}^{(1)} \cap \Lambda_{n+1}^{(2,+)} \cap \Lambda_{n+1}^{(2,-)}, \\ \Lambda_{n+1}^{(0)} &:= \left\{ \omega \in \mathcal{G}_n : |\omega \cdot \ell + \mathbf{m}_1(i_n)j| \geq \frac{2\gamma_n}{\langle \ell \rangle^\tau}, \quad \mathbf{v} \cdot \ell + j = 0, \quad \forall j \in S^c, \ell \in \mathbb{Z}^\nu \right\}, \\ \Lambda_{n+1}^{(1)} &:= \left\{ \omega \in \mathcal{G}_n : |\omega \cdot \ell + \sigma d_{\sigma j}^\infty(i_n)| \geq \frac{2\gamma_n}{\langle \ell \rangle^\tau}, \quad \forall j \in S^c, \ell \in \mathbb{Z}^\nu, \mathbf{v} \cdot \ell + j = 0, \sigma = \pm \right\}, \\ \Lambda_{n+1}^{(2,+)} &:= \left\{ \omega \in \mathcal{G}_n : |\omega \cdot \ell + \sigma(d_{\sigma j}^\infty(i_n) + d_{-\sigma k}^\infty(i_n))| \geq \frac{2\gamma_n}{\langle \ell \rangle^\tau}, \quad \mathbf{v} \cdot \ell + j - k = 0, \forall j, k \in S^c, \ell \in \mathbb{Z}^\nu, \sigma = \pm \right\}, \\ \Lambda_{n+1}^{(2,-)} &:= \left\{ \omega \in \mathcal{G}_n : |\omega \cdot \ell + \sigma(d_{\sigma j}^\infty(i_n) - d_{\sigma k}^\infty(i_n))| \geq \frac{2\gamma_n^*}{\langle \ell \rangle^\tau}, \quad \forall j, k \in S^c, j \neq k, \ell \in \mathbb{Z}^\nu, \mathbf{v} \cdot \ell + j - k = 0, \sigma = \pm \right\}, \end{aligned} \quad (13.3)$$

where  $\gamma_n := \gamma(1 + 2^{-n})$ ,  $\gamma_n^* := \gamma^3(1 + 2^{-n})$  and  $d_j^\infty(\omega) := d_j^\infty(\omega, i_n(\omega))$  are defined in (12.2). Moreover  $|\zeta_n|^{\gamma, \mathcal{G}_n} \lesssim \|\mathcal{F}(U_n)\|_{s_0}^{\gamma, \mathcal{G}_n}$  and

$$\|\mathfrak{J}_n\|_{s_0+\mu_1}^{\gamma, \mathcal{G}_n} \leq C_* \varepsilon^{b_*} \gamma^{-1}, \quad \|\mathcal{F}(U_n)\|_{s_0+\mu_1+3}^{\gamma, \mathcal{G}_n} \leq C_* \varepsilon^{b_*}, \quad (13.4)$$

where  $U_n := (i_n, \Xi_n)$  with  $i_n(\varphi) = (\varphi, 0, 0) + \mathfrak{J}_n(\varphi)$ . The embedding  $\mathfrak{J}_n$  is traveling, namely it satisfies (5.31).

The differences  $\widehat{\mathfrak{J}}_n := \mathfrak{J}_n - \mathfrak{J}_{n-1}$  (where we set  $\widehat{\mathfrak{J}}_0 := 0$ ) is defined on  $\mathcal{G}_n$ , and satisfy

$$\|\widehat{\mathfrak{J}}_1\|_{s_0+\mu_1}^{\gamma, \mathcal{G}_1} \leq C_* \varepsilon^{b_*} \gamma^{-1}, \quad \|\widehat{\mathfrak{J}}_n\|_{s_0+\mu_1}^{\gamma, \mathcal{G}_n} \leq C_* \varepsilon^{b_*} \gamma^{-1} N_{n-1}^{-\alpha}, \quad \forall n \geq 2. \quad (13.5)$$

(P2)<sub>n</sub>  $\|\mathcal{F}(U_n)\|_{s_0}^{\gamma, \mathcal{G}_n} \leq C_* \varepsilon^{b_*} N_{n-1}^{-\alpha}$  where we set  $N_{-1} := 1$ .

(P3)<sub>n</sub> (High Norms).  $\|\mathfrak{J}_n\|_{s_0+\beta_1}^{\gamma, \mathcal{G}_n} \leq C_* \varepsilon^{b_*} \gamma^{-1} N_{n-1}^k$  and  $\|\mathcal{F}(U_n)\|_{s_0+\beta_1}^{\gamma, \mathcal{G}_n} \leq C_* \varepsilon^{b_*} N_{n-1}^k$ .

(P4)<sub>n</sub> (Measure). The measure of the ‘‘Cantor-like’’ sets  $\mathcal{G}_n$  satisfies

$$|\Omega_\varepsilon \setminus \mathcal{G}_0| \leq C_* \varepsilon^{2(\nu-1)} \gamma, \quad |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq C_* \varepsilon^{2(\nu-1)} \gamma N_{n-1}^{-1}. \quad (13.6)$$

*Proof.* Recalling (5.37) and Lemma 5.7 we have

$$\|\mathcal{F}(U_0)\|_s = \|\mathcal{F}((\varphi, 0, 0), 0)\|_s = \|X_P(\varphi, 0, 0)\|_s \lesssim_s \varepsilon^{16-2b}.$$

Hence the smallness conditions in  $(\mathcal{P}_1)_0, (\mathcal{P}_2)_0, (\mathcal{P}_3)_0$  hold taking  $C_*$  large enough. Assume that  $(\mathcal{P}_1)_n, (\mathcal{P}_2)_n, (\mathcal{P}_3)_n$  hold for some  $n \geq 0$ . By (13.1) and (13.2) we have that  $N_0^{C_1} \varepsilon^{b_*+1} \gamma^{-7}$  is arbitrarily small with  $\varepsilon$ .

If we take  $C_1$  bigger than  $C_0$  in Theorem 12.1 then (12.1) holds for  $\varepsilon$  small enough. In (6.3) we consider  $\mathbf{p}_0 = \mathbf{p}_1 + \mathfrak{d}$ , where  $\mathbf{p}_1 := \mu + 2\tau + 1$  (where  $\mu$  is given in Proposition 11.1) and  $\mathfrak{d}$  appears in Lemma 6.2. Since  $\mu_1 \gg \mathbf{p}_0$  (13.4) implies (6.3) and so (7.1), and Proposition 12.2 applies. Hence the operator  $\mathcal{L}_\omega := \mathcal{L}_\omega(\omega, i_n(\omega))$  in (7.4) is defined on  $\mathcal{O}_0 = \mathcal{G}_n$  and is invertible for all  $\omega \in \mathcal{G}_{n+1}$  since  $\mathcal{G}_{n+1} \subseteq \mathcal{O}_\infty$  and the (12.10) holds. This means that the assumption (6.15) of Theorem 6.8 is verified with  $\Omega_\infty = \mathcal{G}_{n+1}$ . By Theorem 6.8 there exists an approximate inverse  $\mathbf{T}_n(\omega) := \mathbf{T}_0(\omega, i_n(\omega))$  of the linearized operator  $L_n(\omega) := d\mathcal{F}(\omega, i_n(\omega), \Xi_n) \equiv d\mathcal{F}(\omega, i_n(\omega), 0)$ .

Now, for all  $\omega \in \mathcal{G}_{n+1}$ , we can define, for  $n \geq 0$ ,

$$U_{n+1} := U_n + H_{n+1}, \quad H_{n+1} := (\widehat{\mathcal{J}}_{n+1}, \widehat{\Xi}_{n+1}) := -\tilde{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) \in E_n \times \mathbb{R}^\nu, \quad (13.7)$$

where  $\tilde{\Pi}_n(\mathcal{J}, \Xi) := (\Pi_n \mathcal{J}, \Xi)$ . By the inductive hypothesis  $\mathcal{J}_n$  is traveling, namely it satisfies (5.31). We claim that  $\widehat{\mathcal{J}}_{n+1}$  satisfies (5.31). Indeed by Lemma 5.5  $\mathcal{F}(U_n)$  is traveling, by Remark 13.1 the projection  $\Pi_n$  of a traveling embedding is a traveling embedding, by Theorem 6.8 the approximate inverse  $\mathbf{T}_n$  maps traveling embeddings into traveling embeddings. By construction we have

$$\begin{aligned} \mathcal{F}(U_{n+1}) &= \mathcal{F}(U_n) + L_n H_{n+1} + Q_n, \\ Q_n &:= Q(U_n, H_{n+1}), \quad Q(U_n, H) := \mathcal{F}(U_n + H) - \mathcal{F}(U_n) - L_n H, \quad H \in E_n \times \mathbb{R}^\nu. \end{aligned}$$

Then, by the definition of  $H_{n+1}$  in (13.7), using  $[L_n, \Pi_n]$  and writing  $\tilde{\Pi}_n^\perp(\mathcal{J}, \Xi) := (\Pi_n^\perp \mathcal{J}, 0)$  we have

$$\begin{aligned} \mathcal{F}(U_{n+1}) &= \mathcal{F}(U_n) - L_n \tilde{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n = \mathcal{F}(U_n) - L_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + L_n \tilde{\Pi}_n^\perp \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n \\ &= \mathcal{F}(U_n) - \Pi_n L_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + (L_n \tilde{\Pi}_n^\perp - \Pi_n^\perp L_n) \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n = \Pi_n^\perp \mathcal{F}(U_n) + R_n + Q_n + Q'_n \end{aligned}$$

where

$$R_n := (L_n \tilde{\Pi}_n^\perp - \Pi_n^\perp L_n) \mathbf{T}_n \Pi_n \mathcal{F}(U_n), \quad Q'_n := -\Pi_n (L_n \mathbf{T}_n - \text{I}) \Pi_n \mathcal{F}(U_n). \quad (13.8)$$

The estimates in  $(\mathcal{P}1)_n, (\mathcal{P}2)_n, (\mathcal{P}3)_n$  follow word by word as in section 8 in [32]. The measure estimates  $(\mathcal{P}4)_n$ , which are different respect to [32], are proved in the next section.  $\square$

**13.1. Measure estimates.** In this section we give the proof of the bound (13.6). Recall (12.2), (11.3). Let us define for  $0 < \eta \lesssim \sqrt{\varepsilon}$ ,  $\tau_0 \geq 1$  and  $n \in \mathbb{N}$

$$\begin{aligned} R_{\ell j k}^{(-)}(\eta, \tau_0) &:= R_{\ell j k}^{(-)}(i_n, \eta, \tau_0) := \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + \sigma(d_{\sigma j}^\infty - d_{\sigma k}^\infty)| \leq 2\eta \langle \ell \rangle^{-\tau_0}, \mathbf{v} \cdot \ell + j - k = 0, \sigma = \pm\} \\ R_{\ell j k}^{(+)}(\eta, \tau_0) &:= R_{\ell j k}^{(+)}(i_n, \eta, \tau_0) := \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + \sigma(d_{\sigma j}^\infty + d_{\sigma k}^\infty)| \leq 2\eta \langle \ell \rangle^{-\tau_0}, \mathbf{v} \cdot \ell + j - k = 0, \sigma = \pm\}, \\ Q_{\ell j}(\eta, \tau_0) &:= Q_{\ell j}(i_n, \eta, \tau_0) := \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + \mathbf{m}_1(\omega) j| \leq 2\eta \langle \ell \rangle^{-\tau_0}, \mathbf{v} \cdot \ell + j = 0\}, \\ P_{\ell j}(\eta, \tau_0) &:= P_{\ell j}(i_n, \eta, \tau_0) := \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + \sigma d_{\sigma j}^\infty| \leq 2\eta \langle \ell \rangle^{-\tau_0}, \mathbf{v} \cdot \ell + j = 0, \sigma = \pm\}. \end{aligned}$$

Recalling (13.3) we can write, setting  $\eta \rightsquigarrow \gamma_n$  for the sets  $Q_{\ell j}(\eta, \tau_0), P_{\ell j}(\eta, \tau_0)$  and  $R_{\ell j k}^{(+)}(\eta, \tau_0), \eta \rightsquigarrow \gamma_n^*$  for the set  $R_{\ell j k}^{(-)}(\eta, \tau_0)$ , and  $\tau_0 \rightsquigarrow \tau$ ,

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} = \bigcup_{\ell \in \mathbb{Z}^\nu, j, k \in S^c} \left( R_{\ell j k}^+(i_n, \gamma_n, \tau) \cup R_{\ell j k}^-(i_n, \gamma_n^*, \tau) \cup Q_{\ell j}(i_n, \gamma_n, \tau) \cup P_{\ell j}(i_n, \gamma_n, \tau) \right). \quad (13.9)$$

We notice that

- by (5.25) and  $\gamma > \gamma^3$  (see (13.2)),  $R_{\ell j k}^{(-)}(i_n) = \emptyset$  for  $j = k, \ell \neq 0$ .
- If  $\ell = 0$ , by the momentum conservation, we have that  $j = k$ . Hence  $R_{0 j j}^{(+)} = \emptyset$ .

We start with a preliminary lemma, which gives a first relation between  $\ell, j, k$  which must be satisfied in order to have non empty resonant sets.

**Lemma 13.3.** *Let  $n \geq 0$ . There is a constant  $C > 0$  dependent of the tangential set and independent of  $\ell, j, k, n, i_n, \omega$  such that the following holds:*

- if  $R_{\ell j k}^{(+)}(i_n, \eta, \tau_0) \neq \emptyset$  then  $|\ell| \geq C(\sqrt{|j|} + \sqrt{|k|})$ ;
- if  $Q_{\ell j}(i_n, \eta, \tau_0) \neq \emptyset$  then  $|\ell| \geq C|j|$ ;
- if  $P_{\ell j}(i_n, \eta, \tau_0) \neq \emptyset$  then  $|\ell| \geq C|j|$ .

*Proof.* If  $R_{\ell j k}^{(+)}(i_n, \eta, \tau_0) \neq \emptyset$ , then there exists  $\omega$  such that

$$|\sigma(d_{\sigma j}^\infty(\omega, i_n(\omega)) + d_{\sigma k}^\infty(\omega, i_n(\omega))) - \mathbf{m}_1 \mathbf{v} \cdot \ell| < 2\eta \langle \ell \rangle^{-\tau_0} + |(\omega - \mathbf{m}_1 \mathbf{v}) \cdot \ell|.$$

Moreover, using (12.2), (12.3), we get  $|\sigma(d_{\sigma j}^\infty(\omega, i_n(\omega)) + d_{\sigma k}^\infty(\omega, i_n(\omega))) - \mathbf{m}_1 \mathbf{v} \cdot \ell| \geq \frac{1}{3}(\sqrt{|j|} + \sqrt{|k|})$ . Thus, for  $\varepsilon$  small enough

$$|\ell| \gtrsim_S |(\omega - \mathbf{m}_1 \mathbf{v}) \cdot \ell| \gtrsim_S \left( \frac{1}{3} - \frac{2\eta}{\langle \ell \rangle^{\tau_0} (\sqrt{|j|} + \sqrt{|k|})} \right) (\sqrt{|j|} + \sqrt{|k|}) \gtrsim_S \frac{1}{4} (\sqrt{|j|} + \sqrt{|k|})$$

and this proves the first claim on  $R_{\ell j k}^{(+)}(i_n, \eta, \tau_0)$ . The others are similar.  $\square$

**Remark 13.4.** *The above lemma implies that for any  $\ell \in \mathbb{Z}^\nu$  there are only finitely many indexes  $j, k \in S^c$  such that the sets  $R_{\ell, j, k}^{(+)}(\eta, \tau_0)$ ,  $Q_{\ell, j}(\eta, \tau_0)$ ,  $P_{\ell, j}(\eta, \tau_0)$  are not empty. Unfortunately this is not true for the set  $R_{\ell, j, k}^{(-)}(\eta, \tau_0)$ .*

13.1.1. *Measure of a resonant set.* The aim of this subsection is to prove the following lemma.

**Lemma 13.5.** *We have that*

$$|R_{\ell j k}^{(-)}(\eta, \sigma)| \leq K \varepsilon^{2(\nu-1)} \eta \langle \ell \rangle^{-\tau_0}, \quad (13.10)$$

for some  $K = K(S)$ . The same holds for  $R_{\ell j k}^{(+)}(\eta, \tau_0)$ ,  $Q_{\ell j}(\eta, \tau_0)$  and  $P_{\ell j}(\eta, \tau_0)$ .

**Lemma 13.6.** *We have*

$$|\nabla_\omega(\omega - \mathbf{m}_1(\omega) \mathbf{v}) \cdot \ell| \gtrsim_S |\ell| \quad (13.11)$$

*Proof.* By (10.2), estimate (10.4), (5.14) and (5.27) we have that

$$\partial_{\omega_i} m_1(\omega) \mathbf{v} \cdot \ell = \varepsilon^{-2} \sum_j (\mathbb{V} \mathbb{A}^{-1})_i^j \ell_j + O(\varepsilon^2) \quad \Rightarrow \quad \partial_{\omega_i} [(\omega - \mathbf{m}_1(\omega) \mathbf{v}) \cdot \ell] = [(\text{Id} - \mathbb{V} \mathbb{A}^{-1}) \ell]_i + O(\varepsilon^2). \quad (13.12)$$

By Lemma 5.3 we have that  $\det(\mathbb{A} - \mathbb{V}) \neq 0$ , hence the (13.12) implies the (13.11).  $\square$

**Proof of Lemma 13.5.** We prove the lemma for the set  $R_{\ell j k}^{(-)}(\eta, \tau_0)$  that is the most difficult case. The estimates for the other sets follow in the same way. We have that  $|\text{sign}(j) - \text{sign}(k)| \leq 2$  and

$$|\sqrt{|j|} - \sqrt{|k|}| \leq \frac{||j| - |k||}{\sqrt{|j|} + \sqrt{|k|}} \leq \frac{|\mathbf{v}| |\ell|}{\sqrt{|j|} + \sqrt{|k|}}. \quad (13.13)$$

Then by Lemma 13.6, (10.4), (11.4), (12.3) we have that

$$\begin{aligned} |\nabla_\omega(\omega \cdot \ell + \sigma(d_{\sigma j}^\infty - d_{\sigma k}^\infty))| &\geq C(S) |\ell| - \varepsilon^2 |\nabla_\omega \mathbf{m}_{1/2}^{(\geq 4)}(\omega)| |\mathbf{v}| |\ell| - 2\varepsilon^2 |\nabla_\omega \mathbf{m}_0(\omega)| \\ &\quad - |\nabla_\omega r_j| - |\nabla_\omega r_j^\infty| - |\nabla_\omega r_k| - |\nabla_\omega r_k^\infty| \gtrsim_S (1 - \varepsilon^2) |\ell|. \end{aligned}$$

Then the (13.10) follows by Fubini theorem.  $\square$

13.1.2. *Summability.*

**Lemma 13.7.** *For  $n \geq 1$ ,  $|\ell| \leq N_{n-1}$ , one has  $R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau)$ ,  $R_{\ell j k}^{(+)}(i_n, \gamma_n, \tau)$ ,  $Q_{\ell j}(i_n, \gamma_n, \tau)$ ,  $P_{\ell j}(i_n, \gamma_n, \tau) = \emptyset$ .*

*Proof.* Consider the set  $R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau)$ . By the momentum condition  $\mathbf{v} \cdot \ell + j - k = 0$  we have to consider the sets  $R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau)$  only for

$$|j - k| \leq |\mathbf{v}| |\ell| \lesssim_S N_{n-1}.$$

By (10.4), (12.8) (with  $i_\delta^{(1)} \rightsquigarrow i_n$  and  $i_\delta^{(2)} \rightsquigarrow i_{n-1}$ ,  $N \rightsquigarrow N_{n-1}$ ) and (13.5) we have for all  $j, k \in S^c$

$$|\sigma(d_{\sigma j}^\infty - d_{\sigma k}^\infty)(i_n) - \sigma(d_{\sigma j}^\infty - d_{\sigma k}^\infty)(i_{n-1})| \leq \varepsilon \gamma^{-3} N_{n-1}^{-\mathbf{a}} \quad \forall \omega \in \mathcal{G}_n, \quad (13.14)$$

where  $\mathbf{a} := \min\{\kappa_*, \alpha\}$  (recall  $\alpha$  in (13.1) and  $\kappa_*$  in (12.8)). Now for all  $j \neq k$ ,  $|\ell| \leq N_{n-1}$ ,  $\omega \in \mathcal{G}_n$  by (13.14)

$$\begin{aligned} |\omega \cdot \ell + \sigma(d_{\sigma j}^\infty(i_n) - d_{\sigma k}^\infty(i_n))| &\geq |\omega \cdot \ell + \sigma(d_{\sigma j}^\infty(i_{n-1}) - d_{\sigma k}^\infty(i_{n-1}))| - |\sigma(d_{\sigma j}^\infty - d_{\sigma k}^\infty)(i_n) - \sigma(d_{\sigma j}^\infty - d_{\sigma k}^\infty)(i_{n-1})| \\ &\geq 2\gamma_{n-1}^* \langle \ell \rangle^{-\tau} - \varepsilon \gamma^{-3} N_{n-1}^{-\mathbf{a}} \geq 2\gamma_n^* \langle \ell \rangle^{-\tau} \end{aligned} \quad (13.15)$$

since  $\varepsilon \gamma^{-6} N_n^{\tau - (2/3)\mathbf{a}} 2^{n+1} \leq 1$ . Since  $R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau) \subseteq \mathcal{G}_n$  then  $R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau) = \emptyset$ . The others are similar.  $\square$

We have proved that, recall (13.9),

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} \subseteq \bigcup_{\substack{j, k \in S^c \\ |\ell| > N_{n-1}}} \left( R_{\ell j k}^{(+)}(i_n, \gamma_n, \tau) \cup R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau) \cup Q_{\ell j}(i_n, \gamma_n, \tau) \cup P_{\ell j}(i_n, \gamma_n, \tau) \right), \quad \forall n \geq 1. \quad (13.16)$$

By Lemmata 13.3, 13.7, 13.5 we deduce that

$$\left| \bigcup_{\ell \in \mathbb{Z}^\nu, j, k \in S^c} R_{\ell j k}^{(+)}(\gamma_n, \tau) \cup Q_{\ell j}(\gamma_n, \tau) \cup P_{\ell j}(\gamma_n, \tau) \right| \leq \sum_{\substack{|\ell| > N_{n-1}, \\ |j|, |k| \leq C|\ell|^2}} \left( |R_{\ell j k}^{(+)}(\gamma_n, \tau)| + |Q_{\ell j}(\gamma_n, \tau)| + |P_{\ell j}(\gamma_n, \tau)| \right) \\ \lesssim_S \varepsilon^{2(\nu-1)} \sum_{|\ell| > N_{n-1}} \frac{\gamma |\ell|^2}{\langle \ell \rangle^\tau} \lesssim_S \varepsilon^{2(\nu-1)} \gamma N_{n-1}^{-1}. \quad (13.17)$$

Consider now indexes  $\ell, j, k$  such that  $\nu \cdot \ell + j - k = 0$  and  $\text{sign}(j) = -\text{sign}(k)$ . Therefore we deduce that  $|j - k| = |j| + |k| \leq |\nu| |\ell|$ . Then by Lemmata 13.7, 13.5 we get

$$\left| \bigcup_{\substack{\ell \in \mathbb{Z}^\nu, j, k \in S^c \\ \text{sign}(j) = -\text{sign}(k)}} R_{\ell j k}^{(-)}(\gamma_n^*, \tau) \right| \leq \sum_{\substack{|\ell| > N_{n-1}, \\ |j| + |k| \leq |\nu| |\ell|}} |R_{\ell j k}^{(-)}(\gamma_n^*, \tau)| \lesssim_S \varepsilon^{2(\nu-1)} \sum_{|\ell| > N_{n-1}} \frac{\gamma |\ell|}{\langle \ell \rangle^\tau} \lesssim_S \varepsilon^{2(\nu-1)} \gamma^3 N_{n-1}^{-1}. \quad (13.18)$$

It remains to study the measure of the union of the sets  $R_{\ell j k}^{(-)}(i_n)$  over for indexes such that  $\text{sign}(j) = \text{sign}(k)$ ,  $|\ell| > N_{n-1}$ . This is the most difficult case, since by Remark 13.4, for any fixed  $\ell \in \mathbb{Z}^\nu$  there are infinite many indexes  $j, k$  such that  $R_{\ell j k}^{(-)}(i_n) \neq \emptyset$ . So we cannot reason as in (13.17), (13.18) and we need a more refined argument.

**Lemma 13.8.** *There exists  $C > 0$  such that if  $\text{sign}(j) = \text{sign}(k)$  and  $|j|, |k| \geq C \langle \ell \rangle^{2\nu+6} \gamma^{-2}$  then*

$$R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau) \subseteq Q_{\ell, j-k}(\gamma, \nu+2). \quad (13.19)$$

*Proof.* Recalling (12.2), (11.3), (10.2), using that  $\text{sign}(j) = \text{sign}(k)$  we write  $\psi_{\ell, j, k} = \omega \cdot \ell + \sigma(d_{\sigma j}^\infty - d_{\sigma k}^\infty)$  with

$$\psi_{\ell, j, k} = \omega \cdot \ell + \mathbf{m}_1(j - k) + \sigma \left( (1 + \mathbf{m}_{1/2}^{(\geq 4)}) (\sqrt{|j|} - \sqrt{|k|}) + r_{\sigma j} - r_{\sigma k} + r_{\sigma j}^\infty - r_{\sigma k}^\infty \right).$$

We show that if  $\omega \in (Q_{\ell, j-k}(\gamma, \nu+2))^c$  then  $\omega \in (R_{\ell j k}^{(-)}(\gamma_n^*, \tau))^c$ . If  $\omega \in (Q_{\ell, j-k}(\gamma, \nu+2))^c$  one has

$$|\omega \cdot \ell + \mathbf{m}_1(j - k)| \geq \gamma \langle \ell \rangle^{-(\nu+2)}.$$

Moreover since  $|j|, |k| \geq C \langle \ell \rangle^{2(\nu+3)} \gamma^{-2}$  then we have

$$|\psi_{\ell, j, k}| \gtrsim \frac{\gamma}{\langle \ell \rangle^{\nu+2}} - \frac{\langle \ell \rangle}{\sqrt{|k|} + \sqrt{|j|}} - \frac{\varepsilon^3 \gamma}{\sqrt{|j|}} + \frac{\varepsilon^3}{\sqrt{|k|}} \gtrsim \frac{\gamma}{\langle \ell \rangle^{\nu+2}} \left( 1 - \frac{1 + \varepsilon}{C} \right) \gtrsim \frac{\gamma}{2 \langle \ell \rangle^{\nu+2}},$$

for  $C > 0$  large enough. This also implies  $|\psi_{\ell, j, k}| \gtrsim \gamma^3 \langle \ell \rangle^{-\tau}$ , since  $\tau > \nu + 2$  and  $\gamma \geq \gamma^3$ . This means  $\omega \in (R_{\ell j k}^{(-)}(i_n))^c$  and hence we get the thesis.  $\square$

As a consequence of Lemma 13.8 we have

$$\left| \bigcup_{\substack{\ell \in \mathbb{Z}^\nu, j, k \in S^c \\ \text{sign}(j) = \text{sign}(k)}} R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau) \right| \leq \sum_{\substack{|\ell| > N_{n-1}, \\ |j|, |k| \geq C \langle \ell \rangle^{2(\nu+3)} \gamma^{-2}, \\ \text{sign}(j) = \text{sign}(k)}} |R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau)| + \sum_{\substack{|\ell| > N_{n-1}, \\ |j|, |k| \leq 2C \langle \ell \rangle^{2(\nu+3)} \gamma^{-2}, \\ \text{sign}(j) = \text{sign}(k)}} |R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau)|.$$

On one hand, using Lemmata 13.5 and 13.8, we have that

$$\sum_{\substack{|\ell| > N_{n-1}, \\ |j|, |k| \geq C \langle \ell \rangle^{2(\nu+3)} \gamma^{-2}, \\ \text{sign}(j) = \text{sign}(k)}} |R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau)| \stackrel{(13.19)}{\lesssim} \sum_{\substack{|\ell| > N_{n-1}, j-k=h, \\ |h| \leq C|\ell|}} |Q_{\ell h}(i_n, \gamma, \nu+2)| \lesssim_S \sum_{\substack{|\ell| > N_{n-1}, j-k=h, \\ |h| \leq C|\ell|}} \varepsilon^{2(\nu-1)} \gamma \langle \ell \rangle^{-\nu-2} \\ \lesssim_S \varepsilon^{2(\nu-1)} \gamma \sum_{|\ell| \geq N_{n-1}} \langle \ell \rangle^{-(\nu+1)} \lesssim_S K \varepsilon^{2(\nu-1)} \gamma N_{n-1}^{-1}.$$



On the other hand

$$\begin{aligned} \sum_{\substack{|\ell| > N_{n-1}, |j-k| \leq C|\ell| \\ |j|, |k| \leq 2C\langle \ell \rangle^{2(\nu+3)} \gamma^{-2} \\ \text{sign}(j) = \text{sign}(k)}} |R_{\ell j k}^{(-)}(i_n, \gamma_n^*, \tau)| &\stackrel{(13.10)}{\lesssim_S} \varepsilon^{2(\nu-1)} \sum_{|\ell| \geq N_{n-1}} \frac{\gamma^3 \langle \ell \rangle^{2\nu+6}}{\gamma^2 \langle \ell \rangle^\tau} \\ &\lesssim_S \gamma \varepsilon^{2(\nu-1)} \sum_{|\ell| \geq N_{n-1}} \langle \ell \rangle^{-(\tau-2\nu-6)} \lesssim_S \gamma \varepsilon^{2(\nu-1)} N_{n-1}^{-1}. \end{aligned}$$

By the discussion above and (13.17), (13.18) we obtain the estimates (13.6).

**13.2. Conclusion of the Proof of Theorem 5.6.** Theorem 13.2 implies that the sequence  $(\mathcal{J}_n, \Xi_n)$  is well defined for  $\omega \in \mathcal{G}_\infty := \bigcap_{n \geq 0} \mathcal{G}_n$ ,  $\mathcal{J}_n$  is a Cauchy sequence in  $\|\cdot\|_{s_0+\mu_1}^{\gamma, \mathcal{G}_\infty}$  (see (13.5)) and  $|\Xi_n|^\gamma \rightarrow 0$ . Therefore  $\mathcal{J}_n$  converges to a limit  $\mathcal{J}_\infty$  in norm  $\|\cdot\|_{s_0+\mu_1}^{\gamma, \mathcal{G}_\infty}$  and, by  $(\mathcal{P}2)_n$ , for all  $\omega \in \mathcal{G}_\infty$ ,  $i_\infty(\varphi) := (\varphi, 0, 0) + \mathcal{J}_\infty(\varphi)$  is a solution of

$$\mathcal{F}(i_\infty, 0) = 0 \quad \text{with} \quad \|\mathcal{J}_\infty\|_{s_0+\mu_1}^{\gamma, \mathcal{G}_\infty} \lesssim \varepsilon^{16-2b} \gamma^{-1}$$

by (13.4), (13.2). Therefore  $\varphi \mapsto i_\infty(\varphi)$  is an invariant torus for the Hamiltonian vector field  $X_{H_\varepsilon}$  (recall (5.19)). We point out that  $\mathcal{J}_\infty$  is traveling, i.e. satisfies (5.31), since it is limit (in the  $H^{s_0+\mu_1}$  topology) of solutions of the linear constant coefficients PDE  $\mathbf{v} \cdot \partial_\varphi + (0, 0, \partial_x) = 0$ . By (13.6),

$$|\Omega_\varepsilon \setminus \mathcal{G}_\infty| \leq |\Omega_\varepsilon \setminus \mathcal{G}_0| + \sum_{n \geq 0} |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq 2C_* \varepsilon^{2(\nu-1)} \gamma + C_* \varepsilon^{2(\nu-1)} \gamma \sum_{n \geq 1} N_{n-1}^{-1} \lesssim C_* \varepsilon^{2(\nu-1)} \gamma.$$

The set  $\Omega_\varepsilon$  in (5.20) has measure  $|\Omega_\varepsilon| = O(\varepsilon^{2\nu})$ . Hence  $|\Omega_\varepsilon \setminus \mathcal{G}_\infty|/|\Omega_\varepsilon| \rightarrow 0$  as  $\varepsilon \rightarrow 0$  because  $\gamma = o(\varepsilon^2)$ , and therefore the measure of  $\mathcal{C}_\varepsilon := \mathcal{G}_\infty$  satisfies (5.39).

It remains to show the linear stability of the embedding  $i_\infty(\varphi)$ . Recall that in the original coordinates the solution reads as in (5.32). By the discussion of section 6 (see also [8] for further details) and section 12, since  $i_\infty(\varphi)$  is isotropic and solves the equation  $\mathcal{F}(i_\infty, 0) = 0$ , it is possible to find a change of coordinates  $G_\infty$  (of the form (6.4)), so that in the linearized system at  $i_\infty$  of the Hamiltonian  $H_\varepsilon \circ G_\infty(\psi, Y, U)$  the equation for the actions is given by  $\dot{Y} = 0$ . Moreover, by section 12 the linear equation for the normal variables  $U$  is conjugated, by setting (see (12.11), (3.4))  $Z = (\chi \circ \Lambda)(U)$ , to the diagonal system  $\dot{Z}_j + \text{id}_j^\infty(\omega) Z_j = f_j(\omega t)$ ,  $j \in S^c$ , where  $f(\omega t)$  is a forcing term. Since  $d_j^\infty \in \mathbb{R}$  a standard argument shows that the Sobolev norms of  $Z$  do not increase in time.

## APPENDIX A. FLOWS AND CONJUGATIONS

**A.1. Flows of pseudo differential PDEs.** Here we state some preliminary results which are the counterpart of sections 3.1, 3.2 in [31]. In this paper we shall consider the flow of the equation

$$\begin{cases} \partial_\theta \Psi^\theta u = i\Pi_S^\perp [Op^W(f(\tau, \varphi, x, \xi))[\Pi_S^\perp \Psi^\theta u]], \\ \Psi^0 u = u, \end{cases} \quad (\text{A.1})$$

where  $f$  is a symbol assumed to have one of the following forms:

$$f(\tau, \varphi, x, \xi) := b(\tau, \varphi, x) \xi, \quad b(\tau, \varphi, x) := \frac{\beta(\varphi, x)}{1 + \tau \beta_x(\varphi, x)}, \quad \beta(\varphi, x) \in H^s(\mathbb{T}^{\nu+1}; \mathbb{R}) \quad (\text{A.2})$$

$$f(\tau, \varphi, x, \xi) := \beta(\varphi, x) |\xi|^{\frac{1}{2}}, \quad \beta(\varphi, x) \in H^s(\mathbb{T}^{\nu+1}; \mathbb{R}), \quad (\text{A.3})$$

$$f(\tau, \varphi, x, \xi) \in S^m, \quad m \leq 0, \quad f \text{ real valued.} \quad (\text{A.4})$$

Notice that the map  $\Psi^\tau$  in (A.1) (if well posed) is symplectic. Indeed it is the flow associated to the Hamiltonian

$$S(\tau, \varphi, z) = \frac{1}{2} \int_{\mathbb{T}} Op^W(f(\tau; \varphi, x, \xi)) z \cdot \bar{z} dx, \quad z \in H_S^\perp.$$

In the following we study the properties of the flow in (A.1). In particular we shall compare the map  $\Psi^\tau$ ,  $\tau \in [0, 1]$  with the flow  $\Phi^\tau$  of the equation

$$\partial_\theta \Phi^\theta = iOp^W(f(\varphi, x, \xi)) \Phi^\theta, \quad \Phi^0 = \text{Id}, \quad (\text{A.5})$$

with  $f(\tau, \varphi, x, \xi)$  as in (A.2)-(A.4). We start by considering the case  $f$  as in (A.2).

For  $\tau \in [0, 1]$  we consider the map

$$\begin{aligned} \mathcal{A}^\tau h(\varphi, x) &:= \sqrt{1 + \tau\beta_x(\varphi, x)} h(\varphi, x + \tau\beta(\varphi, x)), & \varphi \in \mathbb{T}^\nu, x \in \mathbb{T}, \\ (\mathcal{A}^\tau)^{-1} h(\varphi, y) &:= \sqrt{1 + \tilde{\beta}_y(\tau, \varphi, y)} h(\varphi, y + \tilde{\beta}(\tau, \varphi, y)), & \varphi \in \mathbb{T}^\nu, y \in \mathbb{T}, \end{aligned} \quad (\text{A.6})$$

where  $\beta$  is some smooth function and  $\tilde{\beta}(\tau; x, \xi)$  is such that

$$x \mapsto y = x + \tau\beta(\varphi, x) \Leftrightarrow y \mapsto x = y + \tilde{\beta}(\tau, \varphi, x), \quad \tau \in [0, 1], \quad (\text{A.7})$$

It is easy to check that the map in (A.6) is the flow of the equation

$$\begin{cases} \partial_\tau \mathcal{A}^\tau(u) = Op^W(ib(\tau; \varphi, x)\xi)\mathcal{A}^\tau(u) \\ \mathcal{A}^0 u = u, \end{cases} \quad b(\tau, \varphi, x) := \frac{\beta(\varphi, x)}{1 + \tau\beta_x(\varphi, x)}. \quad (\text{A.8})$$

**Remark A.1.** By an explicit computation it is easy to check that

$$\partial_\tau \mathcal{A}^\tau u = b\partial_x \mathcal{A}^\tau u + \frac{b_x}{2} \mathcal{A}^\tau u =: Op(a(\tau, \varphi, x, \xi))\mathcal{A}^\tau u, \quad a(\tau, \varphi, x, \xi) := ib(\tau, \varphi, x)\xi + \frac{1}{2}b_x(\tau, \varphi, x),$$

where  $Op(a)$  is defined in (2.20). By using (2.21) one deduces the (A.8).

We first need to show that  $\mathcal{A}^\tau$  is well defined as map on  $H^s$ .

**Lemma A.2.** Fix  $n \in \mathbb{N}$ . There exists  $\sigma = \sigma(\rho, \nu)$  such that, if  $\|\beta\|_{s_0+\sigma}^{\gamma, \mathcal{O}} < 1$ , then the flow  $\mathcal{A}^\tau(\varphi)$  of (A.8) satisfies for  $s \in [s_0, S]$ ,

$$\sup_{\tau \in [0, 1]} \|\mathcal{A}^\tau u\|_s^{\gamma, \mathcal{O}} + \sup_{\tau \in [0, 1]} \|(\mathcal{A}^\tau)^* u\|_s^{\gamma, \mathcal{O}} \lesssim_s \left( \|u\|_s^{\gamma, \mathcal{O}} + \|b\|_{s+\sigma}^{\gamma, \mathcal{O}} \|u\|_{s_0}^{\gamma, \mathcal{O}} \right), \quad (\text{A.9})$$

$$\sup_{\tau \in [0, 1]} \|(\mathcal{A}^\tau - \text{Id})u\|_s^{\gamma, \mathcal{O}} + \sup_{\tau \in [0, 1]} \|((\mathcal{A}^\tau)^* - \text{Id})u\|_s^{\gamma, \mathcal{O}} \lesssim_s \left( \|\beta\|_{s_0+\sigma}^{\gamma, \mathcal{O}} \|u\|_{s+1}^{\gamma, \mathcal{O}} + \|\beta\|_{s+\sigma}^{\gamma, \mathcal{O}} \|u\|_{s_0+1}^{\gamma, \mathcal{O}} \right).$$

For any  $|\alpha| \leq n$ ,  $m_1, m_2 \in \mathbb{R}$  such that  $m_1 + m_2 = |\alpha|$ , for any  $s \geq s_0$  there exist  $\mu_*$ ,  $\sigma_*$ , depending on  $|\alpha|$ ,  $m_1$ ,  $m_2$ , and  $\delta = \delta(m_1, s)$  such that if  $\|\beta\|_{s_0+\mu_*}^{\gamma, \mathcal{O}} \leq \delta$ , and  $\|\Delta_{12}\beta\|_{p+\sigma_*}^{\gamma, \mathcal{O}} \leq 1$  for  $p + \sigma_* \leq s_0 + \mu_*$ , then one has

$$\sup_{\tau \in [0, 1]} \|\langle D_x \rangle^{-m_1} \partial_\varphi^\alpha \mathcal{A}^\tau(\varphi) \langle D_x \rangle^{-m_2} u\|_s^{\gamma, \mathcal{O}} \lesssim_{s, b, m_1, m_2} \|u\|_s^{\gamma, \mathcal{O}} + \|\beta\|_{s+\mu_*}^{\gamma, \mathcal{O}} \|u\|_{s_0}^{\gamma, \mathcal{O}}, \quad (\text{A.10})$$

$$\sup_{\tau \in [0, 1]} \|\langle D_x \rangle^{-m_1} \partial_\varphi^\alpha \Delta_{12} \mathcal{A}^\tau(\varphi) \langle D_x \rangle^{-m_2} u\|_p \lesssim_{p, b, m_1, m_2} \|u\|_p \|\Delta_{12}\beta\|_{p+\mu_*}, \quad m_1 + m_2 = |\alpha| + 1. \quad (\text{A.11})$$

*Proof.* It follows by Lemmata B.7, B.8, B.9 in [31].  $\square$

We have the following.

**Lemma A.3.** Fix  $n \in \mathbb{N}$ . There exists  $\sigma = \sigma(\rho)$  such that, if  $|f|_{m, s_0+\sigma, \alpha}^{\gamma, \mathcal{O}} \leq 1$ , with  $f$  in (A.3) or (A.4), then the flow  $\Phi^\tau(\varphi)$  of (A.5) satisfies for  $s \in [s_0, S]$ , bounds like (A.9)-(A.11).

*Proof.* The result for  $f(\varphi, x, \xi)$  as in (A.3) follows by the results in the appendix of [18]. For the case (A.4) one can follow almost word by word the proof of Proposition 3.1 in [31] and use Lemmata 2.14, 2.15 and Remark 2.16 to pass from the standard quantization (2.20) to the Weyl quantization in (2.15).  $\square$

**Lemma A.4.** Fix  $\rho \geq 3$  and  $p \geq s_0$  then the following holds.

(i) There exist  $\delta \ll 1$ ,  $\sigma_1 := \sigma_1(\rho, p)$  such that if

$$|f|_{m, s_0+\sigma_1, \alpha}^{\gamma, \mathcal{O}} \leq \delta \quad (\text{A.12})$$

then the following holds. Let  $\Phi^\tau$  be the flow of the system (A.5), then the flow of

$$\partial_\tau \Psi^\tau u = \Pi_S^\perp [Op^W(if(\tau; \varphi, x, \xi))\Pi_S^\perp [\Psi^\tau u]], \quad \Psi^0 u = u, \quad (\text{A.13})$$

is well defined for  $|\tau| \leq 1$  and one has  $\Psi^1 = \Pi_S^\perp \Phi^1 \Pi_S^\perp \circ (\text{Id} + \mathcal{R})$  where  $\mathcal{R}$  is an operator with the form

$$\mathcal{R}(\varphi)w = \sum_{|j| \leq C} \int_0^1 (w, g_j(\tau, \varphi))_{L^2(\mathbb{T})} \chi_j(\tau, \varphi) d\tau, \quad (\text{A.14})$$

with  $\|g_j^{(i)}\|_{s,\mathcal{O}}^{\gamma,\mathcal{O}} + \|\chi_j^{(i)}\|_{s,\mathcal{O}}^{\gamma,\mathcal{O}} \lesssim_s 1$ . Moreover  $\mathcal{R}$  belongs to  $\mathfrak{L}_{\rho,p}(\mathcal{O})$  and satisfies

$$\mathbb{M}_{\mathcal{R}}^{\gamma}(s, \mathbf{b}) \lesssim_s |f|_{m,s+\sigma_1,\alpha}^{\gamma,\mathcal{O}}, \quad \mathbb{M}_{\Delta_{12}\mathcal{R}}(p, \mathbf{b}) \lesssim_p |\Delta_{12}f|_{m,p+\sigma_1,\alpha}. \quad (\text{A.15})$$

(ii) Assume now that  $f \in S_k^m$  (recall Def. 7.3) and (7.1) holds. Then, for  $\varepsilon > 0$  small enough, the result of item (i) holds with an operator  $\mathcal{R}$  in  $\mathfrak{L}_{\rho,p}^k$ .

*Proof.* (i) By using Lemmata 2.14, 2.15 and Remark 2.16 we can follow the same strategy of the proof of Lemma C.1 in [32]. For  $\delta, \sigma_1$  such that the smallness condition (A.12) is satisfied, we have that the flow  $\Phi^\tau$  is well-defined for  $|\tau| \leq 1$ . This follows by Lemmata A.2, A.3 (see also Lemmata B.7, B.8, B.9 in [31] for more details).

Let us define  $\Upsilon^\tau$  as the flow of the following Cauchy problem

$$\partial_\tau \Upsilon^\tau u = -(\Psi_*^\tau \mathbf{Z}) \Upsilon^\tau u, \quad \Upsilon^0 u = u \quad (\text{A.16})$$

with

$$\mathbf{Z}u := Op^W(\text{if}(\tau; \varphi, x, \xi)) \Pi_S[u] + \Pi_S Op^W(\text{if}(\tau; \varphi, x, \xi)) \Pi_S^\perp[u] = \sum_{j \in S} (g_j(\tau), u)_{L^2} \chi_j(\tau) + \sum_{j \in S} (\tilde{g}_j(\tau), u)_{L^2} \tilde{\chi}_j(\tau),$$

$$g_j = \tilde{\chi}_j := e^{ijx}, \quad \chi_j = Op^W(\text{if}(\tau; \varphi, x, \xi)) e^{ijx}, \quad \tilde{g}_j := \Pi_S^\perp Op^W(-\text{if}(\tau; \varphi, x, -\xi)) e^{ijx}. \quad (\text{A.17})$$

Equation (A.16) is well posed on  $H^s$  since its vector field is finite rank. By an explicit computation, using (A.5) and (A.16), we deduce that  $\Psi^\tau = \Pi_S^\perp \circ \Phi^\tau \circ \Upsilon^\tau$  is well-defined on  $H^s$  and solves (A.13). In order to show that  $\Upsilon^\tau - \text{I}$  is of the form (A.14) one can follow almost word by word the proof of Lemma C.1 in [32] using the estimates (A.9)-(A.11) on the flow  $\Phi^\tau$ .

(ii) By the estimates on the symbol  $f \in S_k^m$  we deduce that (A.12) holds. The operator  $\mathbf{Z}$  in (A.17) is in  $\mathfrak{L}_{\rho,p}^k$  (see also Lemma C.7 in [32]) by using its explicit expression. By Taylor expanding the flow  $\Psi^\tau$  in (A.13) one can also deduce that the generator of the flow in (A.16) is in  $\mathfrak{L}_{\rho,p}^k$ . Then the result follows.  $\square$

As a consequence of Lemma A.4 we have the following

**Lemma A.5.** (i) Consider the operator

$$A := \Pi_S^\perp (\omega \cdot \partial_\varphi + Op^W(d(\varphi, x, \xi)) + \mathcal{Q}), \quad d \in S^{m'}, \quad \mathcal{Q} \in \mathfrak{L}_{\rho,p},$$

and let  $\Psi^\tau$  be the flow of (A.13). There exist  $\delta \ll 1$ ,  $\sigma_1 := \sigma_1(\rho, p)$  such that if

$$|d|_{m',s_0+\sigma_1,\alpha}^{\gamma,\mathcal{O}} + \mathbb{M}_{\mathcal{Q}}^{\gamma}(s_0 + \sigma_1, \mathbf{b}) \leq \delta$$

then the following holds. One has

$$\Psi^1 A (\Psi^1)^{-1} = \Pi_S^\perp \Phi^1 A (\Phi^1)^{-1} \Pi_S^\perp + \tilde{\mathcal{R}}(\varphi), \quad (\text{A.18})$$

where  $\tilde{\mathcal{R}}$  is a finite rank operator of the form (A.14) satisfying

$$\begin{aligned} \mathbb{M}_{\tilde{\mathcal{R}}}^{\gamma}(s, \mathbf{b}) &\lesssim_s |f|_{m,s+\sigma_1,\alpha}^{\gamma,\mathcal{O}} + |f|_{m,s_0+\sigma_1,\alpha}^{\gamma,\mathcal{O}} (|d|_{m',s+\sigma_1,\alpha}^{\gamma,\mathcal{O}} + \mathbb{M}_{\mathcal{Q}}^{\gamma}(s, \mathbf{b})), \\ \mathbb{M}_{\Delta_{12}\tilde{\mathcal{R}}}(p, \mathbf{b}) &\lesssim_p |\Delta_{12}f|_{m,p+\sigma_1,\alpha} (1 + C_p), \quad C_p := |\Delta_{12}f|_{m,p+\sigma_1,\alpha} + \mathbb{M}_{\Delta_{12}\mathcal{Q}}(p, \mathbf{b}). \end{aligned}$$

(ii) Assume that the generator  $f$  of the flow  $\Psi^\tau$  of (A.13) is in  $S_k^m$ , and that  $d \in S_{k'}^{m'}$ ,  $\mathcal{Q} \in \mathfrak{L}_{\rho,p}^{k'}$ . Assume also that (7.1) holds. Then, for  $\varepsilon > 0$  small enough, the result of item (i) holds true with a remainder  $\tilde{\mathcal{R}}$  in the class  $\mathfrak{L}_{\rho,p}^{\max\{k,k'\}}$ .

*Proof.* (i) Formula (A.18) follows by Lemma A.4 to write  $\Psi^1 = \Pi_S^\perp \Phi^1 \Pi_S^\perp \circ (\text{I} + \mathcal{R})$ . The estimates on the remainder  $\tilde{\mathcal{R}}$  follow by reasoning as in the proof of Lemma C.2 in [32] and by using Lemmata 2.14, 2.15 and Remark 2.16 to pass from the standard quantization (2.20) to the Weyl quantization in (2.15).

(ii) This item follows by using item (ii) in Lemma A.4, reasoning as in the proof of Lemma C.2 in [32] and using Lemma 7.4 (instead of Lemma 2.15) to pass to the Weyl quantization.  $\square$

**Remark A.6.** The same result of Lemmata A.4, A.5 holds also in the case  $d(\varphi, x, \xi)$  is a matrix of symbols in  $S^{m'} \otimes \mathcal{M}_2(\mathbb{C})$ ,  $m' \leq 0$  and  $\mathcal{Q}$  is a matrix of operators in  $\mathfrak{L}_{\rho,p} \otimes \mathcal{M}_2(\mathbb{C})$ .

**A.2. Egorov Theory and conjugation rules.** In this section we study how pseudo differential operators conjugates under the flows in (A.8) and (A.5).

**Notation.** Consider an integer  $n \in \mathbb{N}$ . To simplify the notation from now on we shall write,  $\Sigma_n^*$  the sum over indexes  $k_1, k_2, k_3 \in \mathbb{N}$  such that  $k_1 < n$ ,  $k_1 + k_2 + k_3 = n$  and  $k_1 + k_2 \geq 1$ .

**Theorem A.7. (Egorov).** Fix  $\rho \geq 3$ ,  $p \geq s_0$ ,  $m \in \mathbb{R}$  with  $\rho + m > 0$ . Let  $w(\varphi, x, \xi) \in S^m$  with  $w = w(\omega, \mathfrak{I}(\omega))$ , Lipschitz in  $\omega \in \mathcal{O} \subseteq \mathbb{R}^\nu$  and in the variable  $\mathfrak{I}$ . Let  $\mathcal{A}^\tau$  be the flow of the system (A.8). (i) There exist  $\sigma_1 := \sigma_1(m, \rho)$  and  $\delta := \delta(m, \rho)$  such that, if

$$\|\beta\|_{s_0+\sigma_1}^{\gamma, \mathcal{O}} < \delta, \quad (\text{A.19})$$

then  $\mathcal{A}^\tau Op^W(w)(\mathcal{A}^\tau)^{-1} = Op^W(q(\varphi, x, \xi)) + R$  where  $q \in S^m$  and  $R \in \mathfrak{L}_{\rho, p}(\mathcal{O})$ . In particular

$$q(\varphi, x, \xi) = q_0(\varphi, x, \xi) + q_1(\varphi, x, \xi), \quad (\text{A.20})$$

where  $q_1 \in S^{m-2}$  and (recall (A.7))

$$q_0(\varphi, x, \xi) = w(x + \beta(\varphi, x), \xi(1 + \tilde{\beta}_y(1, \varphi, y))|_{y=x+\beta(\varphi, x)}). \quad (\text{A.21})$$

The symbol  $q_0$  has the form  $q_0(\varphi, x, \xi) = p_0(\tau, \varphi, x, \xi)|_{\tau=1}$  where  $p_0$  solves the equation

$$\frac{d}{d\tau} p_0(\tau) = \{b(\tau, x)\xi, \varphi, p_0(\tau)\}, \quad p_0(0) = w(\varphi, x, \xi). \quad (\text{A.22})$$

Moreover, one has that the following estimates hold:

$$\begin{aligned} |q|_{m, s, \alpha}^{\gamma, \mathcal{O}} &\leq_{m, s, \alpha, \rho} |w|_{m, s, \alpha + \sigma_1}^{\gamma, \mathcal{O}} + \sum_s^* |w|_{m, k_1, \alpha + k_2 + \sigma_1}^{\gamma, \mathcal{O}} \|\beta\|_{k_3 + \sigma_1}^{\gamma, \mathcal{O}}, \\ |\Delta_{12} q|_{m, p, \alpha} &\leq_{m, p, \alpha, \rho} |w|_{m, p+1, \alpha + \sigma_1} \|\Delta_{12} \beta\|_{p+1} + |\Delta_{12} w|_{m, p, \alpha + \sigma_1} \\ &\quad + \sum_{p+1}^* |w|_{m, k_1, \alpha + k_2 + \sigma_1} \|\beta\|_{k_3 + \sigma_1} \|\Delta_{12} \beta\|_{s_0+1} + \sum_p^* |\Delta_{12} w|_{m, k_1, k_2 + \alpha + \sigma_1} \|\beta\|_{k_3 + \sigma_1}. \end{aligned} \quad (\text{A.23})$$

Furthermore for any  $s_0 \leq s \leq \mathcal{S}$

$$\begin{aligned} \mathbb{M}_R^\gamma(s, \mathbf{b}) &\leq_{s, m, \rho} |w|_{m, s+\rho, \sigma_1}^{\gamma, \mathcal{O}} + \sum_{s+\rho}^* |w|_{m, k_1, k_2 + \sigma_1}^{\gamma, \mathcal{O}} \|\beta\|_{k_3 + \sigma_1}^{\gamma, \mathcal{O}}, \quad 0 \leq \mathbf{b} \leq \rho - 2, \\ \mathbb{M}_{\Delta_{12} R}(p, \mathbf{b}) &\leq_{m, p, \rho} |w|_{m, p+\rho, \sigma_1} \|\Delta_{12} \beta\|_{p+\sigma_1} + |\Delta_{12} w|_{m, s+\rho, \sigma_1} \\ &\quad + \sum_{p+\rho}^* |w|_{m, k_1, k_2 + \sigma_1} \|\beta\|_{k_3 + \sigma_1} \|\Delta_{12} \beta\|_{s_0 + \sigma_1} + \sum_{p+\rho}^* |\Delta_{12} w|_{m, k_1, k_2 + \sigma_1} \|\beta\|_{k_3 + \sigma_1}, \quad 0 \leq \mathbf{b} \leq \rho - 3. \end{aligned} \quad (\text{A.24})$$

(ii) Assume that  $\beta \in S_k^0$ ,  $w \in S_{k'}^m$  and that (7.1) holds. Then, for  $\varepsilon > 0$  small enough, the result of item (i) holds for  $q \in S_{\max\{k, k'\}}^m$  and  $R \in \mathfrak{L}_{\rho, p}^{\max\{k, k'\}}$ .

*Proof.* (i) First of all we write  $\mathcal{A}^\tau := \mathcal{A}_2^\tau \circ \mathcal{A}_1^\tau$  (recall (A.6)) where

$$\begin{aligned} \mathcal{A}_1^\tau h(\varphi, x) &:= h(\varphi, x + \tau\beta(\varphi, x)), \quad \varphi \in \mathbb{T}^\nu, x \in \mathbb{T}, \\ \mathcal{A}_2^\tau h(\varphi, x) &:= \sqrt{1 + \tau\beta_x(\varphi, x)} h(\varphi, x) \quad \varphi \in \mathbb{T}^\nu, x \in \mathbb{T}. \end{aligned}$$

The operator  $\mathcal{A}_1^\tau Op^W(w)(\mathcal{A}_1^\tau)^{-1}$  satisfies (A.20)-(A.24) by Theorem 3.4 in [31]. The same bounds are satisfied by  $\mathcal{A}_2^\tau Op^W(w)(\mathcal{A}_2^\tau)^{-1}$  by explicit computations since  $\mathcal{A}_2$  is a multiplication operator. Equation (A.22) follows by the first step in the proof of Theorem 3.4 in [31].

(ii) The symbol  $q_0$  can be expanded in  $\varepsilon$  by using the equation (A.22) and the expansion on  $w$  and  $\beta$ . By the proof of Theorem 3.4 in [31] one can see that the symbol  $q_1$  is constructed iteratively by solving transport equations similar to (A.22) and hence one can conclude  $q_1 \in S_{\max\{k, k'\}}^{m-2}$ . Again by Theorem 3.4 in [31] one sees that the remainder  $R$  has the form

$$R = R(\tau) = \int_0^\tau \mathcal{A}_1^\tau (\mathcal{A}_1^s)^{-1} \mathcal{M} \mathcal{A}_1^s (\mathcal{A}_1^\tau)^{-1} ds, \quad (\text{A.25})$$

for some  $\mathcal{M} \in \mathfrak{L}_{\rho, p}^{\max\{k, k'\}}$ . Notice that the flow  $\mathcal{A}_1^\tau$  satisfies  $\partial_\tau \mathcal{A}_1^\tau = \partial_x b(\tau; x) \mathcal{A}_1^\tau$ . Then we have  $R \in \mathfrak{L}_{\rho, p}^{\min\{k, k'\}}$  by Taylor expanding the flow  $\mathcal{A}_1^\tau$  and reasoning as in the proof of Lemma B.10 in [31].  $\square$

**Lemma A.8. (Conjugation of  $\omega \cdot \partial_\varphi$ ).** *Let  $\mathcal{A}^\tau$  be the flow of the system (A.8). (i) There exist  $\sigma_1 := \sigma_1(m, \rho)$  and  $\delta := \delta(m, \rho)$  such that, if (A.19) holds, then*

$$\mathcal{A}^1 \circ \omega \cdot \partial_\varphi \circ (\mathcal{A}^1)^{-1} = \omega \cdot \partial_\varphi - iOp^W(g(\varphi, x)\xi)$$

where  $g \in S^0$ , independent of  $\xi$ . In particular one has  $g(\varphi, x)\xi = p_0(\tau, \varphi, x, \xi)|_{\tau=1}$  where  $p_0$  solves the equation

$$\frac{d}{d\tau} p_0(\tau, \varphi, x, \xi) = \{b(\tau, \varphi, x)\xi, p_0(\tau, \varphi, x, \xi)\} + i\omega \cdot \partial_\varphi b(\tau, \varphi, x)\xi, \quad p_0(0) = 0. \quad (\text{A.26})$$

Finally one has

$$|g|_{0,s,\alpha}^{\gamma,\mathcal{O}} \leq_{s,\rho,\alpha} |\beta|_{0,s+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}} \quad (\text{A.27})$$

for all  $s_0 \leq s \leq \mathcal{S}$ . Moreover one has

$$|\Delta_{12}g|_{0,p,\alpha} \leq_{p,\alpha,\rho} |\Delta_{12}\beta|_{0,p+\sigma_1,\alpha+\sigma_1} (1 + |\beta|_{0,p+\sigma_1,\alpha+\sigma_1}), \quad (\text{A.28})$$

where  $p$  is the constant given in Definition 2.10.

(ii) Assume that  $\beta \in S_k^0$  and that (7.1) holds. Then, for  $\varepsilon > 0$  small enough, the result of item (i) holds with  $g \in S_k^0$ .

*Proof.* (i) One can reason essentially as in Lemma A.5 in [14]. (ii) One gets that  $g \in S_k^0$  by using the equation (A.26).  $\square$

**Lemma A.9. (Conjugation of a pseudo differential operator).** *Let  $\Phi^\theta(\varphi)$  be the flow of (A.5) with symbol  $f(\varphi, x, \xi)$  in  $S^m$  with  $m \leq 1/2$ , of the form (A.3) or (A.4). Let  $a = a(\omega, \mathfrak{J}(\omega))$  in  $S^{m'}$  depending on  $\omega \in \mathcal{O} \Subset \mathbb{R}^\nu$  and on  $\mathfrak{J}$  in a Lipschitz way for some  $m' \in \mathbb{R}$ . (i) There exist  $\sigma_1 := \sigma_1(m, m', \rho)$  and  $\delta := \delta(m, m', \rho)$  such that, if*

$$|f|_{m,s_0+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}} < \delta, \quad (\text{A.29})$$

then

$$\Phi^1(\varphi)Op^W(a(\varphi, x, \xi))(\Phi^1(\varphi))^{-1} = Op^W(c(\varphi, x, \xi)) + R(\varphi), \quad (\text{A.30})$$

where

$$\begin{aligned} |c|_{m',s,\alpha}^{\gamma,\mathcal{O}} &\leq_{s,\rho,\alpha,m,m'} |a|_{m',s+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}} (1 + |f|_{m,s_0+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}}) + |a|_{m',s_0+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}} |f|_{m',s+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}}, \\ \mathbb{M}_R^\gamma(s, \mathbf{b}) &\leq_{s,\rho,m,m'} |a|_{m',s+\sigma_1,\sigma_1}^{\gamma,\mathcal{O}} |f|_{m',s_0+\sigma_1,\sigma_1}^{\gamma,\mathcal{O}} + |a|_{m',s_0,\sigma_1}^{\gamma,\mathcal{O}} |b|_{m',s+\sigma_1,\sigma_1}^{\gamma,\mathcal{O}}, \end{aligned} \quad (\text{A.31})$$

for all  $0 \leq \mathbf{b} \leq \rho - 2$  and  $s_0 \leq s \leq \mathcal{S}$ . Moreover one has

$$\begin{aligned} |\Delta_{12}c|_{m',p,\alpha} &\leq_{p,\alpha,\rho,m,m'} |\Delta_{12}a|_{m,p+\sigma_1,\alpha+\sigma_1} (1 + |f|_{m,p+\sigma_1,\alpha+\sigma_1}) \\ &\quad + |a|_{m',p+\sigma_1,\alpha+\sigma_1} |\Delta_{12}f|_{m',p+\sigma_1,\alpha+\sigma_1}, \\ \mathbb{M}_{\Delta_{12}R_p}(p, \mathbf{b}) &\leq_{p,\rho,m,m'} |\Delta_{12}a|_{m+1,p+\sigma_1,\sigma_1} |b|_{m',p+\sigma_1,\sigma_1} \\ &\quad + |a|_{m,p+\sigma_1,\sigma_1} |\Delta_{12}b|_{m'+1,p+\sigma_1,\sigma_1}, \end{aligned} \quad (\text{A.32})$$

for all  $0 \leq \mathbf{b} \leq \rho - 3$  and where  $p$  is the constant given in Definition 2.10. In particular the symbol  $c(\varphi, x, \xi)$  admits the expansion

$$c = a + \{f, a\} + \frac{1}{2}\{f, \{f, a\}\} + r, \quad (\text{A.33})$$

where  $r \in S^q$  where  $q = m' + \max\{m - 3, 2m - 4, 3m - 3\}$ . (ii) Assume that  $a \in S_{k'}^{m'}$ ,  $f \in S_k^m$  and that (7.1) holds. Then, for  $\varepsilon > 0$  small enough, the result of item (i) holds with  $c \in S_{\max\{k,k'\}}^{m'}$  and  $R_\rho \in \mathfrak{L}_{\rho,p}^{\max\{k,k'\}}$ .

*Proof.* (i) To study the conjugated operator  $M := \Phi^1(\varphi)Op^W(a(\varphi, x, \xi))(\Phi^1(\varphi))^{-1}$  we apply the usual Lie expansion up to order  $L \geq 1$ , i.e.

$$\begin{aligned} M &= Op^W(a(\varphi, x, \xi)) + \sum_{q=1}^L \frac{1}{q!} \text{ad}_{iOp^W(f(\varphi,x,\xi))}^q [Op^W(a(\varphi, x, \xi))] \\ &\quad + \frac{1}{L!} \int_0^1 (1-\theta)^L \Phi^\theta \text{ad}_{iOp^W(f(\varphi,x,\xi))}^{L+1} [Op^W(a(\varphi, x, \xi))](\Phi^\theta)^{-1} d\theta, \end{aligned} \quad (\text{A.34})$$

where, given two linear operators  $M$  and  $B$ , we defined

$$\text{ad}_M[B] := [M, B], \quad \text{ad}_M^q[B] := [M, \text{ad}_M^{q-1}[B]], \quad q \geq 1.$$

By applying Lemma 2.19 we get (recall (2.44))

$$\text{ad}_{Op^W(i_f)}[Op^W(a)] = [Op^W(i_f), Op^W(a)] = Op^W(\{f, a\} + r_1), \quad r_1 \in S^{m+m'-3},$$

up to a smoothing operator in  $\mathfrak{L}_{\rho,p}$ . Moreover (choosing  $\rho$  possibly larger)

$$\text{ad}_{Op^W(i_f)}^2[Op^W(a)] = Op^W(\{f, \{f, a\}\} + r_2), \quad r_2 \in S^{2m+m'-4},$$

up to a smoothing operator in  $\mathfrak{L}_{\rho,p}$ . By induction, for  $k \geq 3$  we have

$$\text{ad}_{Op^W(i_f)}^k[Op^W(a)] = Op^W(b_k), \quad b_k \in S^{k(m-1)+m'},$$

up to a smoothing operator in  $\mathfrak{L}_{\rho,p}$ . We choose  $L$  in such a way that  $(L+1)(1-m) - m' \geq \rho$  and  $L+1 \geq 3$ , so that the operator  $Op^W(b_{L+1})$  belongs to  $\mathfrak{L}_{\rho,p}$ . The integral Taylor remainder in (A.34) belongs to  $\mathfrak{L}_{\rho,p}$  as well by Lemma B.2 in [31]. The estimates (A.31)-(A.32) follows by using (2.34)-(2.37). (ii) The second item follows by reason exactly as in item (i) but using the composition Lemma 7.5 instead of Lemma 2.18, 2.19.  $\square$

**Lemma A.10. (Conjugation of  $\omega \cdot \partial_\varphi$ ).** *Let  $\Phi^\theta(\varphi)$  be the flow of (A.5) with symbol  $f(\varphi, x, \xi) \in S^m$  with  $m \leq 1/2$  of the form (A.3) or (A.4). Assume also that  $f = f(\omega, \mathfrak{J}(\omega))$  depends  $\omega \in \mathcal{O} \Subset \mathbb{R}^\nu$  and on  $\mathfrak{J}$  in a Lipschitz way. (i) There exist  $\sigma_1 := \sigma_1(m, \rho)$  and  $\delta := \delta(m, \rho)$  such that, if (A.29) holds then*

$$\Phi^1(\varphi) \circ \omega \cdot \partial_\varphi \circ \Phi^{-1}(\varphi) = \omega \cdot \partial_\varphi - iOp^W(c) + R_\rho \quad (\text{A.35})$$

where  $c \in S^m$  and  $R_\rho \in \mathfrak{L}_{\rho,p}$ . In particular one has

$$\begin{aligned} |c|_{m,s,\alpha}^{\gamma,\mathcal{O}} &\leq_{s,\rho,\alpha,m} |f|_{m,s+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}} + |f|_{m,s_0+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}} |f|_{m,s+\sigma_1,\alpha+\sigma_1}^{\gamma,\mathcal{O}}, \\ \mathbb{M}_{R_\rho}^\gamma(s, \mathbf{b}) &\leq_{s,\rho,m} |f|_{m,s+\sigma_1,\sigma_1}^{\gamma,\mathcal{O}} + |f|_{m,s_0,\sigma_1}^{\gamma,\mathcal{O}} |f|_{m,s+\sigma_1,\sigma_1}^{\gamma,\mathcal{O}}, \end{aligned} \quad (\text{A.36})$$

for all  $0 \leq \mathbf{b} \leq \rho - 2$  and  $s_0 \leq s \leq \mathcal{S}$ . Moreover one has

$$|\Delta_{12}c|_{m,p,\alpha} \leq_{p,\alpha,\rho,m} |\Delta_{12}f|_{m,p+\sigma_1,\alpha+\sigma_1} + |f|_{m,p+\sigma_1,\alpha+\sigma_1} |\Delta_{12}f|_{m,p+\sigma_1,\alpha+\sigma_1}, \quad (\text{A.37})$$

$$\mathbb{M}_{\Delta_{12}R_\rho}(p, \mathbf{b}) \leq_{p,\rho,m,m'} |\Delta_{12}f|_{m+1,p+\sigma_1,\sigma_1} |f|_{m,p+\sigma_1,\sigma_1}, \quad (\text{A.38})$$

for all  $0 \leq \mathbf{b} \leq \rho - 3$  and where  $p$  is the constant given in Definition 2.10. In particular the symbol  $c(\varphi, x, \xi)$  admits the expansion

$$c = \omega \cdot \partial_\varphi f + \frac{1}{2} \{f, \omega \cdot \partial_\varphi f\} + r, \quad (\text{A.39})$$

where  $r \in S^{2m-3}$  satisfies bounds as (A.36) and (A.37). (ii) Assume that  $f \in S_k^m$  and that (7.1) holds. Then, for  $\varepsilon > 0$  small enough, the result of item (i) holds with  $c \in S_k^m$  and  $R_\rho \in \mathfrak{L}_{\rho,p}^k$ .

*Proof.* By using the Lie expansion (A.34) we have

$$\begin{aligned} \Phi^1 \circ \omega \cdot \partial_\varphi \circ \Phi^{-1} &= \omega \cdot \partial_\varphi - Op^W(i\omega \cdot \partial_\varphi f) - \sum_{k=2}^L \frac{1}{k!} \text{ad}_{Op^W(i_f)}^{k-1}[Op^W(i\omega \cdot \partial_\varphi f)] \\ &\quad + \frac{1}{L!} \int_0^1 (1-\theta)^L \Phi^\theta \left( \text{ad}_{Op^W(i_f)}^L [Op^W(i\omega \cdot \partial_\varphi f)] \right) (\Phi^\theta)^{-1} d\theta \end{aligned}$$

and the lemma follows (both items (i) and (ii)) using Lemma 2.19, noting that  $f \#_\rho^W \omega \cdot \partial_\varphi - \omega \cdot \partial_\varphi \#_\rho f = \frac{1}{i} \{f, f_t\}$  plus a symbol of order  $2m - 3$  and reasoning as in the proof of Lemma A.9.  $\square$

**Lemma A.11.** *Fix  $\rho \geq 3$ , consider a compact subset  $\mathcal{O} \Subset \mathbb{R}^\nu$  and let  $R \in \mathfrak{L}_{\rho,p}(\mathcal{O})$  (see Def. 2.10). Consider a function  $\beta$  such that  $\beta := \beta(\omega, i(\omega)) \in H^s(\mathbb{T}^{\nu+1})$  for some  $s \geq s_0$ , assume that it is Lipschitz in  $\omega \in \mathcal{O}$  and  $\mathfrak{J}$ . Let  $\mathcal{A}^\tau$  be the operator defined in (A.6).*

(i) *There exists  $\mu = \mu(\rho) \gg 1$ ,  $\sigma = \sigma(\rho)$  and  $\delta > 0$  small such that if  $\|\beta\|_{s_0+\mu}^{\gamma,\mathcal{O}} \leq \delta$  and  $\|\Delta_{12}\beta\|_{p+\sigma}^{\gamma,\mathcal{O}} \leq 1$ , then the operator  $M^\tau := \mathcal{A}^\tau R (\mathcal{A}^\tau)^{-1}$  belongs to the class  $\mathfrak{L}_{\rho,p}$ . In particular one has, for  $s_0 \leq s \leq \mathcal{S}$ ,*

$$\begin{aligned} \mathbb{M}_{M^\tau}^\gamma(s, \mathbf{b}) &\leq \mathbb{M}_R^\gamma(s, \mathbf{b}) + \|\beta\|_{s+\mu}^{\gamma,\mathcal{O}} \mathbb{M}_R^\gamma(s_0, \mathbf{b}), \quad \mathbf{b} \leq \rho - 2, \\ \mathbb{M}_{\Delta_{12}M^\tau}(p, \mathbf{b}) &\leq \mathbb{M}_{\Delta_{12}R^\tau}(p, \mathbf{b}) + \|\Delta_{12}\beta\|_{p+\mu} \mathbb{M}_{R^\tau}^\gamma(p, \mathbf{b}), \quad \mathbf{b} \leq \rho - 3. \end{aligned} \quad (\text{A.40})$$

(ii) *Assume that  $\beta \in S_k^0$ ,  $R \in \mathfrak{L}_{\rho,p}^{k'}$  and that (7.1) holds. Then, for  $\varepsilon > 0$  small enough and  $\rho$  large enough, the result of item (i) follows with  $M^\tau \in \mathfrak{L}_{\rho',p}^{\max\{k,k'\}}$  with  $\rho' = \rho - 28 + 4 \max\{k, k'\}$ .*



*Proof.* (i) It follows following word by word the proof of Lemma B.10 in [31] and using Lemma A.2. (ii) We Taylor expand the flow  $\mathcal{A}^\tau$  by using (A.8). Notice that

$$\partial_{\tau\tau}\mathcal{A}^\tau = Op^W(b(\tau, \varphi, x)\xi)^2\mathcal{A}^\tau + Op^W(\partial_\tau b(\tau, \varphi, x)\xi)\mathcal{A}^\tau.$$

Recalling  $\beta \in S_k^0$ , the (A.8) and using Lemma 7.5 we deduce

$$\partial_{\tau\tau}\mathcal{A}^\tau = Op^W(B^{(2)}(\tau, \varphi, x, \xi))\mathcal{A}^\tau + R^{(2)} \circ \mathcal{A}^\tau$$

where  $B^{(2)} \in S_k^2$  and  $R^{(2)} \in \mathfrak{L}_{\rho,p}^k$ . Using (A.8) one can prove inductively that

$$\partial_\tau^q \mathcal{A}^\tau = Op^W(B^{(q)}(\tau, \varphi, x, \xi))\mathcal{A}^\tau + R^{(q)} \circ \mathcal{A}^\tau$$

where  $B^{(q)} \in S_k^2$  and  $R^{(q)} \in \mathfrak{L}_{\rho,p}^k$ , for  $2 \leq q \leq 15 - 2k$ . In conclusion we get that

$$\mathcal{A}^\tau = \sum_{i=1}^{14-2k} \varepsilon^i Op^W(A_i(\tau; \varphi, x, \xi)) + Op^W(\tilde{A}(\tau; \varphi, x, \xi)) + \sum_{i=1}^{14-2k} \varepsilon^i \mathcal{R}_i + \tilde{\mathcal{R}} + \int_0^\tau \frac{(\tau - \sigma)^{14-2k}}{(14-2k)!} \partial_\tau^{15-2k} \mathcal{A}^\sigma d\sigma \quad (\text{A.41})$$

where  $A_i \in S^i$ ,  $\mathcal{R}_i \in \mathfrak{L}_{\rho,p}$  have the form respectively (7.14), (7.17),  $\tilde{A} \in S^{15-2k}$ ,  $\tilde{\mathcal{R}} \in \mathfrak{L}_{\rho,p}$  satisfy estimates (7.15), (7.18). Moreover, by assumption, we have  $R = \sum_{q=1}^{14-2k'} \varepsilon^q R_q + \tilde{R}$  with  $R_i$  as in (7.17) and  $\tilde{R}$  satisfying (7.18). Reasoning as in Lemma B.10 in [31] one can check that the operators

$$Op^W(A_i(\tau; \varphi, x, \xi))R_q, \quad \mathcal{R}_i R_q$$

are  $(i+q)$ -homogeneous remainder as in (7.17) belonging to  $\mathfrak{L}_{\rho',p}$ . Similarly one has that

$$\left( Op^W(A_i(\tau; \varphi, x, \xi)) + Op^W(\tilde{A}(\tau; \varphi, x, \xi)) + \mathcal{R}_i + \tilde{\mathcal{R}} \right) \tilde{R}, \quad \left( Op^W(\tilde{A}(\tau; \varphi, x, \xi)) + \tilde{\mathcal{R}} \right) R_q$$

belong to  $\mathfrak{L}_{\rho',p}$  and satisfy (7.18) with  $k \rightsquigarrow \max\{k, k'\}$ . Reasoning in the same way one gets that  $(\partial_\tau^{15-2k} \mathcal{A}^\sigma)R$  is in  $\mathfrak{L}_{\rho',p}^{\max\{k,k'\}}$ . The inverse flow  $(\mathcal{A}^\tau)^{-1}$  admits an expansion similar to (A.41). Then the Lemma follows.  $\square$

**Lemma A.12.** Fix  $\rho \geq 3$ , consider a compact subset  $\mathcal{O} \subset \mathbb{R}^\nu$  and let  $R \in \mathfrak{L}_{\rho,p}(\mathcal{O})$ . (i) There exist  $\mu = \mu(\rho)$ ,  $\sigma = \sigma(\rho)$  and  $\delta > 0$  such that, if  $|f|_{m,s_0+\sigma,\alpha}^{\gamma,\mathcal{O}} \leq \delta$  and  $|\Delta_{12}f|_{p+\sigma}^{\gamma,\mathcal{O}} \leq 1$ , with  $f$  in (A.3) or (A.4), then the operator  $M^\tau := \Phi^\tau R (\Phi^\tau)^{-1}$ , where  $\Phi^\tau$  is the flow of (A.5), belongs to the class  $\mathfrak{L}_{\rho,p}$ . In particular, for  $s_0 \leq s \leq \mathcal{S}$ , the bounds (A.40) hold with  $\|\beta\|_{s+\mu}^{\gamma,\mathcal{O}}$  and  $\|\Delta_{12}\beta\|_{p+\sigma}^{\gamma,\mathcal{O}}$  replaced by  $|f|_{m,s+\sigma,\alpha}^{\gamma,\mathcal{O}}$  and  $|\Delta f|_{m,p+\sigma,\alpha}^{\gamma,\mathcal{O}}$ . (ii) Assume that  $f \in S_k^m$ ,  $R \in \mathfrak{L}_{\rho,p}^{k'}$  and that (7.1) holds. Then, for  $\varepsilon > 0$  small enough, the result of item (i) follows with  $M^\tau \in \mathfrak{L}_{\rho,p}^{\max\{k,k'\}}$ .

*Proof.* By Lemma A.3 we have that the flow  $\Phi^\tau$  in (A.5) satisfies the same bounds given in Lemma A.2. Then one can reason as in Lemma A.11.  $\square$

## APPENDIX B. TECHNICAL LEMMATA

**B.1. Dirichlet-Neumann operator.** Here we collect some results about the Dirichlet-Neumann operator. For more details we refer, for instance, to [4], [57] and reference therein. We also remark that, in our context, we shall also consider  $C^\infty$  profile  $\eta(x)$ . This is due to the fact that at each step of the Nash-Moser we perform a  $C^\infty$ -regularization.

Let  $\eta \in C^\infty$ . It is known that the Dirichlet-Neumann operator is (in the infinite depth case) a *pseudo differential* operator of the form

$$G(\eta) := |D| + R_G(\eta), \quad (\text{B.1})$$

where  $G(0) = |D|$  and the remainder  $R_G(\eta) \in OPS^{-\infty}$  (see Def. 2.12). The key result of this section is the following.

**Proposition B.1. (Dirichlet-Neumann).** Assume that  $\eta(\omega, \varphi, x)$ , for  $\omega \in \mathcal{O} \Subset \mathbb{R}^\nu$ , is  $C^\infty$  in the variables  $\varphi, x$  and Lipschitz in  $\omega$ . Then there exist constants  $\sigma = \sigma(s_0)$ ,  $\delta = \delta(s_0)$ ,  $s_0 \geq (d+1/2)$ , such that, if

$$\|\eta\|_{s_0+\sigma}^{\gamma,\mathcal{O}} \leq \delta,$$

then the Dirichlet-Neumann operator in (1.4) has the form (B.1) where  $R_G(\eta)$  is a pseudo differential operator satisfying the following. For any  $m, s, \alpha \in \mathbb{N}$  there exists  $\sigma_1 = \sigma_1(m, s, \alpha)$  such that

$$|R_G(\eta)|_{-m,s,\alpha}^{\gamma,\mathcal{O}} \lesssim_{m,s,\alpha} \|\eta\|_{s+\sigma_1}^{\gamma,\mathcal{O}}. \quad (\text{B.2})$$

Let  $p \geq s_0 + \sigma$ , consider  $\eta_1, \eta_2 \in C^\infty$  and set  $\Delta_{12}R_G := R_G(\eta_1) - R_G(\eta_2)$ . There exists  $\delta(p) > 0$  such that, if

$$\|\eta_1\|_p + \|\eta_2\|_p \leq \delta_1(p),$$

then, for any  $m, \alpha \in \mathbb{N}$ ,

$$|\Delta_{12}R_G|_{-m,p,\alpha} \lesssim_p \|\eta_1 - \eta_2\|_p.$$

*Proof.* One can deduce the result by following almost word by word the proof Proposition 2.37 in [18].  $\square$

We also have the following.

**Proposition B.2. (Tame estimates on the Dirichlet-Neumann).** *There exist  $\sigma = \sigma(s_0) > 0$ ,  $\delta = \delta(s_0) > 0$  such that, if  $\|\eta\|_{s_0+\sigma}^{\gamma,\mathcal{O}} \leq \delta$ , then, for all  $s \geq s_0$*

$$\begin{aligned} \|(G(\eta) - |D|)\psi\|_s^{\gamma,\mathcal{O}} &\leq \|\eta\|_{s+\sigma}^{\gamma,\mathcal{O}} \|\psi\|_{s_0}^{\gamma,\mathcal{O}} + \|\eta\|_{s_0+\sigma}^{\gamma,\mathcal{O}} \|\psi\|_s^{\gamma,\mathcal{O}}, \\ \|(G'(\eta)[\widehat{\eta}]\psi)\|_s^{\gamma,\mathcal{O}} &\leq \|\psi\|_{s+2}^{\gamma,\mathcal{O}} \|\widehat{\eta}\|_{s_0+1}^{\gamma,\mathcal{O}} + \|\psi\|_{s_0+2}^{\gamma,\mathcal{O}} \|\widehat{\eta}\|_{s+1}^{\gamma,\mathcal{O}} + \|\eta\|_{s+\sigma}^{\gamma,\mathcal{O}} \|\psi\|_{s_0+2}^{\gamma,\mathcal{O}} \|\widehat{\eta}\|_{s_0+1}^{\gamma,\mathcal{O}}, \\ \|(G''(\eta)[\widehat{\eta}, \widehat{\eta}]\psi)\|_s^{\gamma,\mathcal{O}} &\leq \|\psi\|_{s+3}^{\gamma,\mathcal{O}} (\|\widehat{\eta}\|_{s_0+2}^{\gamma,\mathcal{O}})^2 + \|\psi\|_{s_0+3}^{\gamma,\mathcal{O}} \|\widehat{\eta}\|_{s+2}^{\gamma,\mathcal{O}} \|\widehat{\eta}\|_{s_0+2}^{\gamma,\mathcal{O}} \\ &\quad + \|\eta\|_{s+\sigma}^{\gamma,\mathcal{O}} \|\psi\|_{s_0+3}^{\gamma,\mathcal{O}} (\|\widehat{\eta}\|_{s_0+2}^{\gamma,\mathcal{O}})^2. \end{aligned}$$

*Proof.* It follows by Lemma 2.32 in [18].  $\square$

**Algebraic properties.** One has that, setting  $f^\vee(x) := f(-x)$ , the Dirichlet-Neumann operator satisfies

$$G(\eta^\vee)[\psi^\vee](x) = G(\eta)[\psi](-x).$$

This implies that

$$B(\eta^\vee, \psi^\vee)(x) = B(\eta, \psi)(-x), \quad V(\eta^\vee, \psi^\vee)(x) = -V(\eta, \psi)(-x).$$

Moreover, if  $(\eta, \psi)$  satisfy

$$\eta(-t, -x) = \eta(t, x), \quad \psi(-t, -x) = -\psi(t, x),$$

then

$$B(\eta, \psi)(-t, -x) = -B(\eta, \psi)(t, x), \quad V(\eta, \psi)(-t, -x) = V(\eta, \psi)(t, x).$$

The following Lemma is fundamental for our scopes.

**Lemma B.3.** *Let  $\eta, \psi \in S_\nu$  be  $C^\infty$  functions in the variables  $(\varphi, x) \in \mathbb{T}^{\nu+1}$ . Then the following holds. (i)  $G(\eta)\psi \in S_\nu$ . (ii)  $V(\eta, \psi), B(\eta, \psi) \in S_\nu$ . (iii) One has*

$$\left( -(\mathbf{v} \cdot \partial_\varphi G)(\eta) + G(\eta) \circ \partial_x - \partial_x \circ G(\eta) \right) h = 0, \quad (\text{B.3})$$

for any  $h \in H^s(\mathbb{T}^{\nu+1}; \mathbb{R})$ .

*Proof.* Since  $\eta, \psi \in S_\nu$ , Lemma 2.21 implies that  $\mathbf{v} \cdot \partial_\varphi \psi + \partial_x \psi = \mathbf{v} \cdot \partial_\varphi \eta + \partial_x \eta = 0$ . Then, using the ‘‘shape derivative’’ formula in (3.9) we deduce

$$(\mathbf{v} \cdot \partial_\varphi + \partial_x)G(\eta)\psi = G(\eta)[\mathbf{v} \cdot \partial_\varphi \psi + \partial_x \psi] + G'(\eta)[\mathbf{v} \cdot \partial_\varphi \eta + \partial_x \eta]\psi = 0. \quad (\text{B.4})$$

Then Lemma 2.21 implies item (i). Item (ii) follows by item (i) and formulæ (3.10) and (3.11). Reasoning as in (B.4) we have

$$(\mathbf{v} \cdot \partial_\varphi G)(\eta)h = G'(\eta)[\mathbf{v} \cdot \partial_\varphi \eta]h, \quad \partial_x \circ G(\eta)h = G(\eta)h + G'(\eta)[\eta_x]h.$$

Then (B.3) follows using that  $\eta \in S_\nu$ .  $\square$

**Homogeneity expansion of the function  $V$ .** We now compute an expansion in degree of homogeneity of  $V$  by following the strategy used in section 3.2 in [14]. Recalling the good unknown map  $\mathcal{G}$  in (8.1) and the map  $\Lambda$  in (3.4) we define

$$u := \frac{1}{\sqrt{2}}|D|^{-\frac{1}{4}}\eta + \frac{i}{\sqrt{2}}|D|^{\frac{1}{4}}(\psi - B\eta). \quad (\text{B.5})$$

Notice that the variable  $u$  is just the first component of  $\Lambda\mathcal{G}^{-1}\left[\begin{smallmatrix} \eta \\ \psi \end{smallmatrix}\right]$ . We have the following Lemma.

**Lemma B.4. (Expansion of  $V$ ).** *The function  $V$  defined in (3.10) admits the expansion*

$$V = (\psi - B\eta)_x + (|D|(\psi - B\eta)_x)\eta + V_{\geq 3} \quad (\text{B.6})$$

where  $V_{\geq 3} \sim O(u^3)$  is a cubic function in  $(\eta, \psi)$ . In the complex variable  $(u, \bar{u})$  defined in (B.5),  $V$  can be expanded as in (7.29), more precisely we have

$$\tilde{V}_1 = \frac{1}{i\sqrt{2}} \partial_x |D|^{-\frac{1}{4}} (u - \bar{u}), \quad (\text{B.7})$$

$$\tilde{V}_2 = \frac{1}{2i} \left( |D|^{\frac{3}{4}} \partial_x (u - \bar{u}) \right) (|D|^{\frac{1}{4}} (u + \bar{u})). \quad (\text{B.8})$$

*Proof.* By (3.11) and using the expansion (3.13), we deduce  $B = |D|\psi$  up to a quadratic function in  $u$ . As a consequence, by (3.10) and (B.5), we have

$$V = \psi_x - B\eta_x = (\psi - B\eta)_x + \partial_x (|D|\psi)\eta - (|D|\psi)\eta_x$$

up to terms  $O_2(\eta, \psi)$ . Since  $B\eta$  is  $O_2(\eta, \psi)$  we get (B.6). The (B.7), (B.8) follow by passing to the complex variable  $u$  in (B.5).  $\square$

**Lemma B.5. (Coefficients of  $\tilde{V}_1$  and  $\tilde{V}_2$ ).** *The coefficients of the functions  $\tilde{V}_1$  and  $\tilde{V}_2$  in (B.7)-(B.8) are, for all  $n \in \mathbb{Z} \setminus \{0\}$*

$$(\tilde{V}_1)_n^+ = (\tilde{V}_1)_n^- = \frac{1}{\sqrt{2}} n |n|^{-1/4}, \quad (\tilde{V}_2)_{n,n}^{+-} = (\tilde{V}_2)_{n,n}^{-+} = \frac{1}{2} n |n|. \quad (\text{B.9})$$

*Proof.* By (B.7) and expanding in Fourier series the function  $\bar{u}$  as

$$u(x) = \sum_{n \in \mathbb{Z} \setminus \{0\}} u_n \frac{e^{inx}}{\sqrt{2\pi}}, \quad u_n := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{T}} u(x) e^{-inx} dx,$$

we have

$$\tilde{V}_1 = \frac{1}{i\sqrt{2}} \partial_x |D|^{-1/4} (u - \bar{u}) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} \sum_{n \neq 0} n |n|^{-1/4} u_n e^{inx} + n |n|^{-1/4} \bar{u}_n e^{-inx},$$

which implies the expressions for  $(\tilde{V}_1)_n^{\pm}$  in (B.9). By (B.8) we have

$$\begin{aligned} \tilde{V}_2 = \frac{1}{4i\pi} & \left( \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{\frac{3}{4}} (ij) u_j e^{ijx} - \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{\frac{3}{4}} (-ij) \bar{u}_j e^{-ijx} \right) \times \\ & \times \left( \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{\frac{1}{4}} u_j e^{ijx} + \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{\frac{1}{4}} \bar{u}_j e^{-ijx} \right). \end{aligned}$$

By collecting from the formula above the terms  $u_{j_1} \bar{u}_{j_2}$  and  $\bar{u}_{j_1} u_{j_2}$  we obtain

$$\begin{aligned} \frac{1}{4\pi} \sum_{j_1, j_2 \in \mathbb{Z} \setminus \{0\}} u_{j_1} \bar{u}_{j_2} (\tilde{V}_2)_{j_1, j_2}^{+-} e^{i(j_1 - j_2)x}, \quad \frac{1}{4\pi} \sum_{j_1, j_2 \in \mathbb{Z} \setminus \{0\}} \bar{u}_{j_1} u_{j_2} (\tilde{V}_2)_{j_1, j_2}^{-+} e^{-i(j_1 - j_2)x}, \\ (\tilde{V}_2)_{j_1, j_2}^{+-} := \frac{1}{2} |j_1|^{\frac{3}{4}} |j_2|^{\frac{1}{4}} j_1 \quad (\tilde{V}_2)_{j_1, j_2}^{-+} := \frac{1}{2} |j_1|^{\frac{3}{4}} |j_2|^{\frac{1}{4}} j_1. \end{aligned}$$

Taking  $j_1 = j_2 = n$  or  $j_1 = n = -j_2$  we obtain the (B.9) for the coefficients  $(\tilde{V}_2)_{n,n}^{+-}, (\tilde{V}_2)_{n,n}^{-+}$ .  $\square$

**B.2. Technical Lemmata of Section 11.** Here we collect the proofs of some technical Lemmata appearing in the paper.

**Proof of Lemma 11.3.** We use the following notation: given a set of integer indexes  $j_1, \dots, j_n$  we write

$$\max(p) := p\text{-largest value among } |j_1|, \dots, |j_n|.$$

*Proof of (11.9).* Let us call  $j_*$  the only component of  $\ell$  such that  $\ell_{j_*} = 1$ . We first consider the case  $\sigma = -\sigma'$ . We have that

$$|\delta_{\sigma, \sigma', j, k}^{(1)}(\ell)| = |\sqrt{|j_*|} + \sigma(\sqrt{|j|} + \sqrt{|k|})|.$$

If  $\sigma = 1$  then the lower bound is given by 3. If  $\sigma = -1$  then by the momentum conservation we have that  $j + k = j_*$ , and so  $|j| + |k| - |j_*| \geq 0$ . If  $j_* \neq \max(1)$  then

$$|\sqrt{|j|} + \sqrt{|k|} - \sqrt{|j_*|}| \geq \sqrt{|j|} \geq 1.$$

If  $j_* = \max(1)$ , suppose without loss of generality that  $j = \max(2)$ , then by Remark 11.2

$$|\sqrt{|j|} + \sqrt{|k|} - \sqrt{|j_*|}| = \frac{(\sqrt{|j|} + \sqrt{|k|})^2 - |j_*|}{\sqrt{|j|} + \sqrt{|k|} + \sqrt{|j_*|}} \geq \frac{2\sqrt{|j||k|}}{\sqrt{|j|} + \sqrt{|k|} + \sqrt{|j_*|}} = \left( \frac{\sqrt{|j|}}{\sqrt{|j_*|}} \right) \frac{2\sqrt{|k|}}{\left(1 + \frac{\sqrt{|k|} + \sqrt{|j|}}{\sqrt{|j_*|}}\right)} \geq \frac{2}{9}.$$

Now we consider the case  $\sigma = \sigma'$ . The only non-trivial case is when  $j = \max(1)$  and

$$\delta_{\sigma, \sigma', j, k}^{(1)}(\ell) = \sqrt{|j_*|} + \sqrt{|k|} - \sqrt{|j|} = \sqrt{|j_*|} - \frac{|j| - |k|}{\sqrt{|j|} + \sqrt{|k|}}.$$

By the momentum conservation (see (11.8)) we have that  $|j| - |k| \leq |j - k| = |j_*|$ . If  $2\sqrt{|j_*|} \leq \sqrt{|j|} + \sqrt{|k|}$ , then  $|\delta_{\sigma, \sigma', j, k}^{(1)}(\ell)| \geq 1/2$ . If  $\sqrt{|j|} + \sqrt{|k|} < 2\sqrt{|j_*|}$  then

$$|\delta_{\sigma, \sigma', j, k}^{(1)}(\ell)| > \inf_{j, k} \frac{1}{\sqrt{|j|} + \sqrt{|k|}} =: C_1(S)$$

indeed, by using that  $|j_*| \geq ||j| - |k||$  and  $|j| \geq |j_*|$

$$(\sqrt{|j|} + \sqrt{|k|})|\delta_{\sigma, \sigma', j, k}^{(1)}(\ell)| \geq \sqrt{|j j_*|} - |j_*| + \sqrt{|j_* k|} \geq \sqrt{|j_* k|} \geq 1.$$

This implies the (11.9).

*Proof of (11.10).* By (11.7), (11.8) we have to bound from below the function

$$\sqrt{|j_1|} \pm \sqrt{|j_2|} \pm \sqrt{|j_3|} \pm \sqrt{|j_4|} \quad \text{when} \quad j_1 \pm j_2 \pm j_3 \pm j_4 = 0,$$

for all possible combinations of signs. *Case (+, +, +).* This case is trivial since

$$|\sqrt{|j_1|} + \sqrt{|j_2|} + \sqrt{|j_3|} + \sqrt{|j_4|}| \geq \max(3)^{1/2}. \quad (\text{B.10})$$

We have to consider the case when two indexes are in  $S$  and the other two are in  $S^c$ . Assume, w.l.o.g that  $\max(3) = |j_1|$ . Then if  $j_1 \in S$  the (B.10) implies the (11.10). If  $j_1 \in S^c$  then there exists  $i = 2, 3, 4$  such that  $j_i \in S$  and  $|j_i| > |j_1|$ . Then again we obtain the (11.10).

*Case (+, +, -).* We have

$$|\sqrt{|j_1|} + \sqrt{|j_2|} + \sqrt{|j_3|} - \sqrt{|j_4|}| = \frac{||j_1| + |j_2| + |j_3| - |j_4| + 2\sqrt{|j_1 j_2|} + 2\sqrt{|j_2 j_3|} + 2\sqrt{|j_1 j_3|}|}{|\sqrt{|j_1|} + \sqrt{|j_2|} + \sqrt{|j_3|} + \sqrt{|j_4|}|}.$$

The momentum condition implies  $|j_1| + |j_2| + |j_3| - |j_4| \geq 0$ . Therefore the quantity above is bounded from below by  $\max(3)^{1/2}$ . We get the (11.10) as in the previous case.

*Case (-, +, -).* This is the most difficult case. Consider the function

$$\psi(j_1, j_2, j_3, j_4) = \sqrt{|j_1|} - \sqrt{|j_2|} + \sqrt{|j_3|} - \sqrt{|j_4|},$$

on the hyperplane  $j_1 - j_2 + j_3 - j_4 = 0$ . We claim that

$$\begin{aligned} \text{either } \psi &= 0 & \text{or } j_i &= -j_k \text{ for some } i, k \in \{1, 2, 3, 4\}, \\ \text{or } |\psi| &\geq \frac{c}{(\max(3))^9}, \end{aligned} \quad (\text{B.11})$$

for some absolute constant  $0 < c \ll 1$ . The (B.11) implies the (11.10). Indeed if the lower bound holds the one concludes as in the previous cases. If, for instance  $j_1 = -j_2$  then, since  $\psi \neq 0$ ,

$$|\psi| = \frac{||j_3| - |j_4||}{\sqrt{|j_3|} + \sqrt{|j_4|}} \geq \frac{1}{\max\{|j_3|, |j_4|\}^{1/2}}. \quad (\text{B.12})$$

By the choice of  $S$  in (1.23) it is not possible that both  $j_1, j_2$  are in  $S$ . Hence we can assume  $j_1, j_3 \in S$  and  $j_2, j_4 \in S^c$ . By the momentum and  $j_1 = -j_2$  we deduce  $|j_4| \leq C(S)$ . Then (B.12) implies the (11.10).

*Proof of the Claim B.11.* Assume that  $\max(1) \geq 100(\max(3))^3$  and without loss of generality we assume  $\max(1) = |j_1|$ ,  $\max(2) = |j_2|$ . We write

$$\psi = \frac{|j_3| - |j_4|}{\sqrt{|j_3|} + \sqrt{|j_4|}} + \frac{|j_1| - |j_2|}{\sqrt{|j_1|} + \sqrt{|j_2|}}. \quad (\text{B.13})$$

If  $j_1 = j_2$  then  $\psi = 0$ . If  $j_1 = -j_2$  then or  $\psi = 0$  or  $|\psi| \geq 1/(\max(3))^{1/2}$  which implies the (11.10). If  $j_1 \neq j_2$  we have, by (B.13),

$$|\psi| \gtrsim \frac{||j_3| - |j_4||}{\sqrt{|j_3|} + \sqrt{|j_4|}} - \frac{|j_1 - j_2|}{\sqrt{|j_1|} + \sqrt{|j_2|}} \gtrsim \frac{1}{(\max(3))^{1/2}} - \frac{2 \max(3)}{\max(1)^{1/2}} \gtrsim \frac{1}{2(\max(3))^{1/2}},$$

since  $\max(1) \geq 100(\max(3))^3$ . It remains to study the case  $\max(1) < 100(\max(3))^3$ . In this case to get the (11.10) it is sufficient to prove

$$|\psi| \geq \frac{1}{10(\max(1))^{N_0}}, \quad (\text{B.14})$$

for some  $N_0 > 0$ . The (B.14) follows by Proposition 6.3 in [14].

*Proof of (11.11).* We prove the result for  $\sigma = \sigma' = +$  which is the most difficult case. By a generic choice of the set  $S$  one can assume that  $\bar{\omega} \cdot \ell \neq 0$  for  $0 < |\ell| \leq 6$ . Therefore one has that there is a constant  $K(S)$  such that

$$|\bar{\omega} \cdot \ell| \geq 2K(S),$$

since  $|\ell| \leq 6$ . Moreover by the constraint in (11.8) we deduce that  $||j| - |k|| \leq \tilde{C}(S)$ , for some  $\tilde{C}(S)$  depending on the tangential sites in  $S$ . Then, by taking  $|j|, |k| \geq C(S) := \tilde{C}^2(S)/K^2(S)$  and  $\varepsilon$  small enough we get (see (11.7))

$$|\delta_{\sigma, \sigma', j, k}^{(p)}(\ell)| \geq |\bar{\omega} \cdot \ell| - \frac{||j| - |k||}{\sqrt{|j|} + \sqrt{|k|}} \geq 2K(S) - K(S)$$

which implies the (11.11).  $\square$

**Proof of Lemma 11.6.** If the operator  $\mathbf{B}$  is bounded, then  $|(\mathbf{B})_{\sigma, j}^{\sigma', k}(\ell)| \leq c$  for some constant  $c > 0$ . Since  $\mathbf{A}$  is almost-diagonal and it preserves the momentum, there exist two constants  $C_1, C_2$  depending on  $S$  such that the following holds: for all  $j, k$  such that  $(\mathbf{A})_{\sigma, j}^{\sigma', k}(\ell) \neq 0$  we have  $C_1 \langle k \rangle \leq \langle j \rangle \leq C_2 \langle k \rangle$ . Suppose that  $|j| \geq |k|$ , then

$$c \geq |(\mathbf{B})_{\sigma, j}^{\sigma', k}(\ell)| = |(\mathbf{A})_{\sigma, j}^{\sigma', k}(\ell) \langle j \rangle^m \langle k \rangle^m| \geq C_2^{-1} |(\mathbf{A})_{\sigma, j}^{\sigma', k}(\ell) \langle j, k \rangle^{2m}|.$$

This implies the thesis. Assume now that (11.12) holds. Then, since  $\mathbf{A}$  is almost-diagonal it is easy to note that  $|(\mathbf{B})_{\sigma, j}^{\sigma', k}(\ell)| \leq \tilde{C}(S) |j - k|^{-1} |\ell|^{-\nu-1}$ . Therefore, for any  $v \in H_{S^\perp}^s$ , we have (using Cauchy-Schwarz)

$$\|\underline{\mathbf{B}}_{\sigma, j}^{\sigma'} v\|_s^2 \lesssim \sum_{k, p} |v_k(p)|^2 \langle k, p \rangle^{2s} \sum_{j, \ell} \langle j - k, \ell - p \rangle^{2s} |(\mathbf{B})_{\sigma, j}^{\sigma', k}(\ell - p)|^2 \lesssim_{s, S} \|v\|_s,$$

which implies the thesis.  $\square$

**Proof of Lemma 11.7.** Recall that  $\mathbf{q} \in S_4^{-1/2}$  and  $\mathcal{Q} \in \mathfrak{L}_{\rho, p}^4 \otimes \mathcal{M}_2(\mathbb{C})$ . Moreover (see (7.13), (7.16)) we have

$$\mathbf{q} = \sum_{i=1}^6 \varepsilon^i \mathbf{q}_i + \mathbf{q}_{\geq 7}, \quad \mathcal{Q} = \sum_{i=1}^6 \varepsilon^i \mathcal{Q}_i + \mathcal{Q}_{\geq 7}, \quad \mathbf{Q}_i := Op^W \left[ \begin{array}{c} i\mathbf{q}_i(\varphi, x, \xi) \\ 0 \\ -i\mathbf{q}_i(\varphi, x, \xi) \end{array} \right] + \mathcal{Q}_i,$$

where  $\mathbf{q}_i, \mathcal{Q}_i, i = 1, \dots, 6$ , are respectively  $i$ -homogeneous symbols and operators as in (7.14), (7.17). By Lemma A.4 in [31], where it is proved the immersion of smoothing pseudo differential tame operators in the class of modulo tame operators, we have that  $\mathbf{Q}_{\geq 7} := Op^W(\mathbf{q}_{\geq 7}) + \mathcal{Q}_{\geq 7}$  is  $\text{Lip}(-1/2)$ -modulo tame with

$$\begin{aligned} \mathfrak{M}_{\mathbf{Q}_{\geq 7}}^{\sharp, \gamma}(-1/2, s, \mathbf{b}_0) &\lesssim_s |\mathbf{q}_{\geq 7}|_{-\frac{1}{2}, s+\mu, \alpha}^{\gamma, \Omega_\infty^{2\gamma}} + \mathbb{M}_{\mathbf{Q}_{\geq 7}}^\gamma(s, \mathbf{b}), \\ \mathfrak{M}_{\Delta_{12} \mathbf{Q}_{\geq 7}}^{\sharp}(-1/2, p, \mathbf{b}_0) &\lesssim_p |\Delta_{12} \mathbf{q}_{\geq 7}|_{-\frac{1}{2}, p+\mu, \alpha} + \mathbb{M}_{\Delta_{12} \mathbf{Q}_{\geq 7}}(\mathbf{b}) \|i_1 - i_2\|_{p+\mu}, \end{aligned} \quad (\text{B.15})$$

for some  $\mu = \mu(\nu) > 0$  and  $\mathbf{b} = s_0 + \mathbf{b}_0$ . By estimates (7.15), (7.18) with  $k = 4$  we have that the (B.15) imply the (11.15).

By Remark 11.5 the operators  $\mathbf{Q}_i, i = 1, \dots, 6$  are almost-diagonal and  $x$ -translation invariant. Moreover we have that  $\langle D \rangle^{1/4} (\mathbf{Q}_i)_\sigma \langle D \rangle^{1/4}$ , is a bounded operator since  $(\mathbf{Q}_i)_\sigma = Op^W(i\mathbf{q}_i) + \mathcal{Q}_i$  is a pseudo differential operator of

order  $-1/2$  plus a smoothing operator. The operator  $\langle D \rangle^{\rho/2} (\mathbb{Q}_i)_{-\sigma} \langle D \rangle^{\rho/2}$ , is bounded since  $\mathbb{Q}_i \in \mathfrak{L}_{\rho,p}$ . Then (11.14) follows by Lemma 11.6.  $\square$

**Proof of Lemma 11.10.** Note that the operator  $\mathbf{A}$  given in (11.19) is well-defined by (11.11) in Lemma 11.3. Since  $\mathbf{B}$  is  $x$ -translation invariant, almost-diagonal, using Lemma 2.22 and formulæ (11.19), (11.8), it is easy to check that also  $\mathbf{A}$  is  $x$ -translation invariant, almost-diagonal. Moreover we have that  $\mathbf{B}$  is Hamiltonian, i.e.  $\mathbf{B} = iE\tilde{\mathbf{B}}$  with  $\tilde{\mathbf{B}}$  self-adjoint (it satisfies (2.50)). By Lemma 2.23 we have

$$\tilde{\mathbf{B}}_{\sigma,j}^{\sigma,k}(\ell) = \overline{\tilde{\mathbf{B}}_{\sigma,k}^{\sigma,j}(-\ell)}, \quad \tilde{\mathbf{B}}_{\sigma,-j}^{-\sigma,-k}(\ell) = \tilde{\mathbf{B}}_{\sigma,k}^{-\sigma,j}(\ell).$$

Using the explicit expression of  $\delta_{\sigma,\sigma',j,k}^{(p)}(\ell)$  in (11.7) one can check that  $\mathbf{A}$  is Hamiltonian. Passing to the coefficients of the matrices we have that the left hand side of equation (11.20) is given by

$$-i\delta_{\sigma,\sigma',j,k}^{(p)}(\ell)\mathbf{A}_{\sigma,j}^{\sigma',k}(\ell) + \mathbf{B}_{\sigma,j}^{\sigma',k}(\ell).$$

Then by definition (11.19) the equation (11.20) is verified. The (11.21) follows by Lemma 11.3, indeed the condition  $\max\{|j|, |k|\} \geq C$  implies that  $|j|, |k| \geq C(S)$  by the definition of  $C$  and the momentum relation (11.8).  $\square$

**Proof of Lemma 11.11.** We observe that

$$(\langle \partial_\varphi \rangle^{\mathbf{b}_0} \langle D \rangle^{1/2} \mathbf{A} \langle D \rangle^{1/2})_{\sigma,j}^{\sigma',k}(\ell) = \langle \ell \rangle^{\mathbf{b}_0} \langle j \rangle^{\frac{1}{2}} \langle k \rangle^{\frac{1}{2}} \mathbf{A}_{\sigma,j}^{\sigma',k}(\ell).$$

Then it is easy to deduce the (11.23) by using that  $|\ell| \leq i$  and  $\mathbf{v} \cdot \ell = \sigma j - \sigma' k$ . Item (ii) follows by Lemma A.5 in [31] where estimates for composition of modulo-tame operators are provided. Regarding item (iii) we have, by the bounds in item (i), (ii),

$$\begin{aligned} \mathfrak{M}_{\mathbf{F}}^{\sharp,\gamma}(s, \mathbf{b}_0) &\lesssim_s \mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(s, \mathbf{b}_0) \mathfrak{M}_{\mathbf{B}}^{\sharp,\gamma}(s_0, \mathbf{b}_0) \varepsilon^n (\mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(s_0, \mathbf{b}_0))^{n-1} \sum_{k-n \geq 0} \frac{\varepsilon^{k-n}}{k!} (C(\mathbf{A}))^{k-n} \\ &+ \mathfrak{M}_{\mathbf{B}}^{\sharp,\gamma}(s, \mathbf{b}_0) \varepsilon^n (\mathfrak{M}_{\mathbf{A}}^{\sharp,\gamma}(s_0, \mathbf{b}_0))^n \sum_{k-n \geq 0} \frac{\varepsilon^{k-n}}{k!} C(\mathbf{A})^{k-n}. \end{aligned}$$

Hence for  $\varepsilon$  small enough we get the bound (11.24). To prove (11.25) is sufficient to reason as above and using the estimates of item (i) and (ii).  $\square$

**Proof of Lemma 11.12.** (i) We have that

$$\mathbf{C}_{\sigma,j}^{\sigma',k}(\ell) = \sum_{\ell_1, k_1, \sigma''} \mathbf{A}_{\sigma,j}^{\sigma'',k_1}(\ell_1) \mathbf{B}_{\sigma'',k_1}^{\sigma',k}(\ell - \ell_1) - \mathbf{B}_{\sigma,j}^{\sigma'',k_1}(\ell_1) \mathbf{A}_{\sigma'',k_1}^{\sigma',k}(\ell - \ell_1) \quad (\text{B.16})$$

and we observe that when  $\sigma = \sigma'$  then the matrices involved are  $1/2$ -smoothing, if  $\sigma \neq \sigma'$  then there is always an out of diagonal component of  $\mathbf{A}$  or  $\mathbf{B}$  that is  $\rho$ -smoothing. The sum in (B.16) is restricted to indexes (recall  $\mathbf{A}, \mathbf{B}$  are almost-diagonal and  $x$ -translation invariant) satisfying

$$|j - k| = |\mathbf{v} \cdot \ell|, \quad |j - k_1| = |\mathbf{v} \cdot \ell_1|, \quad |k_1 - k| = |\mathbf{v} \cdot (\ell - \ell_1)|. \quad (\text{B.17})$$

First of all we have  $|j| \leq |k|(1 + |\mathbf{v} \cdot \ell|) \leq |k|(1 + C(S)(i_1 + i_2))$ . Therefore

$$\frac{|k|}{(1 + C(S)(i_1 + i_2))} \leq |j| \leq |k|(1 + C(S)(i_1 + i_2)), \quad \Rightarrow \quad \langle j \rangle^{1/2} \langle k \rangle^{1/2} \geq \frac{\langle j, k \rangle}{(1 + C(S)(i_1 + i_2))}.$$

Notice also that, fixed  $j$  and  $k$ , the cardinality of indexes  $\ell_1, k_1$  satisfying the constraints (B.17) is bounded by

$$K(i_1)^{\nu-1} C(S)(i_1 + i_2),$$

with  $K > 0$  some absolute constant. Then, using the bounds (11.22) on  $\mathbf{A}, \mathbf{B}$  and the (B.16) we deduce (11.27) with

$$C(\mathbf{C}) \lesssim C(\mathbf{A})C(\mathbf{B})(i_1)^{\nu-1} C(S)(i_1 + i_2)(1 + C(S)(i_1 + i_2)).$$

(ii) By an explicit computation on can check

$$\mathbf{M}_+^\dagger := [\mathbf{A}_+^\dagger, \mathcal{D}_+^\dagger], \quad \mathbf{M}_-^\dagger := \mathbf{A}_-^\dagger \mathcal{D}_-^\dagger - \mathcal{D}_+^\dagger \mathbf{A}_-^\dagger.$$



The bound on the coefficients of  $\mathbf{M}_+^+$  follows reasoning as in item (i). Using (10.7), (11.26) we have that the coefficients of  $\mathbf{M}_+^-$  are

$$(\mathbf{M}_{+,j}^-)^{k}(\ell) = i \sum_{k=1}^3 \varepsilon^{2k} m_{1/2}^{(2k)} (\sqrt{|k|} + \sqrt{|k|}) \mathbf{A}_{+,j}^{-,k}(\ell).$$

Then the second in (11.28) follows by the estimates (11.22) on the coefficients of  $\mathbf{A}_+^-$ . Item (iii) follows by item (ii) using the estimates in (10.4), and by item (i) of Lemma 11.11.  $\square$

## REFERENCES

- [1] Alazard T., Baldi P., *Gravity capillary standing water waves*, Arch. Rat. Mech. Anal, 217(3):741-830, (2015).
- [2] Alazard T., Burq N., Zuily C., *On the water-wave equations with surface tension*. Duke Math. J., 158, 413-499, 2011.
- [3] Alazard T., Burq N., Zuily C., *On the Cauchy problem for gravity water waves*. Invent. Math., 198, 71–163, 2014.
- [4] Alazard T., Delort J.-M., *Sobolev estimates for two dimensional gravity water waves*, Astérisque, vol. 374, (2015).
- [5] Alazard T., Delort J.-M. *Global solutions and asymptotic behavior for two dimensional gravity water waves*. Ann. Sci. Éc. Norm. Supér., 48(5):1149–1238, (2015).
- [6] Baldi P., Berti M., Haus E., Montalto R., *Time quasi-periodic gravity water waves in finite depth*, Invent. math. 214, 739–911, (2018).
- [7] Baldi P., Berti M., Montalto R., *KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation*, Math. Ann. 359, 471–536 (2014).
- [8] Baldi P., Berti M., Montalto R., *KAM for autonomous quasi-linear perturbations of KdV*, Ann. I. H. Poincaré (C) Anal. Non Linéaire, 33(6):1589–1638, (2015).
- [9] Baldi P., Montalto R., *Quasi-periodic incompressible Euler flows in 3D*, preprint arXiv:2003.14313.
- [10] Berti M., Biasco L., Procesi M., *KAM theory for reversible derivative wave equations*, Arch. Rat. Mech. and Anal., 212(3): 905–955 (2014).
- [11] Berti M., Bolle P., *A Nash-Moser approach to KAM theory*, Fields Institute Communications, volume 75 “Hamiltonian PDEs and Applications”, 255-284, (2014).
- [12] Berti M., Delort J.-M., *Almost Global Solutions of Capillary-gravity Water Waves Equations on the Circle*. UMI Lecture Notes 2018 (awarded UMI book prize 2017), ISBN 978-3-319-99486-4.
- [13] Berti M., Feola R., Franzoi L., *Quadratic life span of periodic gravity-capillary water waves*, Water Waves, 3:85–115 (2021).
- [14] Berti M., Feola R., Pusateri F., *Birkhoff normal form and long time existence for periodic gravity Water Waves*, accepted on CPAM, preprint arXiv:1810.11549.
- [15] Berti M., Feola R., Pusateri F., *Birkhoff Normal form for Gravity Water Waves*, Water Waves, 3:117–126 (2021).
- [16] Berti M., Franzoi L., Maspero A., *Traveling Quasi-Periodic Water Waves with Constant Vorticity*, Arch. Rational Mech. Anal., 240, 99–202 (2021).
- [17] Berti M., Franzoi L., Maspero A., *Pure gravity traveling quasi-periodic water waves with constant vorticity*, preprint arXiv:2101.12006.
- [18] Berti M., Montalto R., *KAM for gravity capillary water waves*, Memoires of AMS, Memo 891, vol. 263 (2020).
- [19] Bourgain J., *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices, 11, 475–497, (1994).
- [20] Coutand D., Shkoller S., *Well-posedness of the free-surface incompressible Euler equations with or without surface tension*. J. Amer. Math. Soc. 20, no. 3, 829-930, 2007.
- [21] Craig W., Nicholls D., *Travelling two and three dimensional capillary gravity water waves*, SIAM J. Math. Anal., 32(2):323-359, (2000).
- [22] Craig W., Sulem C., *Mapping properties of normal forms transformations for water waves*, Boll. Unione Mat. Ital. 9(2):289-318, (2016).
- [23] Craig W., Sulem C., *Numerical simulation of gravity waves*, J. Comput. Phys., 108(1):73–83, (1993).
- [24] Craig W., Wayne C. E., *Newton’s method and periodic solutions of nonlinear wave equations*. Comm. Pure Appl. Math., 46(11):1409–1498 (1993).
- [25] Craig W., Worfolk P., *An integrable normal form for water waves in infinite depth*, Phys. D, 84(3-4):513-531, (1995).
- [26] Degasperis A., Holm D. D., Hone A. N. W., *A new integrable Equation with Peakon Solutions*, Theoretical and Mathematical Physics, 133(2): 1463-1474 (2002).
- [27] Dyachenko A.I., Lvov Y.V., Zakharov V.E., *Five-wave interaction on the surface of deep fluid*, Physica D 87(1-4):233-261, (1995).
- [28] Feola R., Giuliani F., *Time quasi-periodic traveling gravity water waves in infinite depth*, Rend. Lincei Mat. Appl. 31: 901–916, (2020).
- [29] Feola R., Giuliani F., Montalto R., Procesi M., *Reducibility of first order linear operators on tori via Moser’s theorem*, Journal of Functional Analysis 276(3):932-970 (2019).
- [30] Feola R., Giuliani F., Pasquali S., *On the integrability of Degasperis-Procesi equation: control of the Sobolev norms and Birkhoff resonances*, Journal of Differential Equations 266(6):3390–3437, (2019).
- [31] Feola R., Giuliani F., Procesi M., *Reducibility for a class of weakly dispersive linear operators arising from the Degasperis Procesi equation*, Dynamics of Partial Differential Equations, 16(1): 25-94 (2019).
- [32] Feola R., Giuliani F., Procesi M., *Reducibile KAM tori for the Degasperis-Procesi equation*, Commun. Math. Phys. 377, 1681–1759 (2020).
- [33] Feola R., Iandoli F., *Local well-posedness for quasi-linear NLS with large Cauchy data on the circle*, Ann. I. H. Poincaré (C) Anal. Non Linéaire, 36(1):119-164, (2019).
- [34] Feola R., Iandoli F., *Long time existence for fully nonlinear NLS with small Cauchy data on the circle*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 22(1): 109–182 (2021).

- [35] Feola R., Procesi M., *Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations*, J. Differential Equations 259(7):3389–3447, (2015).
- [36] Geng J., Xu X., You J., *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math. 226(6): 5361–5402 (2011).
- [37] Geng J., You J., *A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions*, J. Differential Equations 209(1): 1–56, (2005).
- [38] Geng J., You J., *A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces*, Commun. Math. Phys. 262, 343–372, (2006).
- [39] Germain P., Masmoudi N., Shatah J., *Global solutions for the gravity water waves equation in dimension 3*, Ann. of Math. (2), 175(2):691–754, (2012).
- [40] Giuliani F., *Quasi-periodic solutions for quasi-linear generalized KdV equations*, J. Differential Equations, 262(10):5052–5132 (2017).
- [41] Hörmander L., *The Analysis of Linear Partial Differential Operators III - Pseudo differential operators*, Springer.
- [42] Ifrim M., Tataru D., *Two dimensional water waves in holomorphic coordinates II: global solutions*, Bull. Soc. Math. France 144(2):369–394, (2016).
- [43] Ifrim M., Tataru D., *The lifespan of small data solutions in two dimensional capillary water waves*, Arch. Rat. Mech. Anal. 225(3):1279–1346, (2017).
- [44] Ionescu A., Pusateri F., *Global solutions for the gravity water waves system in 2d*, Invent. Math., 199(3):653–804, (2015).
- [45] Ionescu A., Pusateri F., *Global regularity for 2d water waves with surface tension*, Mem. Amer. Math. Soc., vol. 256(1227), (2018).
- [46] Ionescu A., Pusateri F., *Recent advances on the global regularity for water waves*, Philos. Trans. Roy. Soc. A. 376(2111), (2018).
- [47] Ionescu A., Pusateri F., *Long-time existence for multi-dimensional periodic water waves*, Geom. Funct. Anal. 29(3):811–870, (2019).
- [48] Iooss G., Plotnikov P., *Multimodal standing gravity waves: a completely resonant system*, J. math. fluid mech. 7, S110–S126, (2005).
- [49] Iooss G., Plotnikov P., *Small divisor problem in the theory of three-dimensional water gravity waves*, Mem. Amer. Math. Soc., 200(940), (2009).
- [50] Iooss G., Plotnikov P., *Asymmetrical tridimensional traveling gravity waves*, Arch. Rat. Mech. Anal., 200(3):789–880, (2011).
- [51] 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.
- [52] Kappeler T., Pöschel J., *KAM and KdV*, Springer (2003).
- [53] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional. Anal. i Prilozhen. 21(3):22–37, 95 (1987).
- [54] S. Kuksin. *A KAM theorem for equations of the Korteweg-de Vries type*. Rev. Math. Phys., 10(3):1–64 (1998).
- [55] Kuksin S., Pöschel J., *Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation*. Ann. of Math., 143(1):149–179 (1996).
- [56] Lannes D., *Well-posedness of the water-waves equations*, J. Amer. Math. Soc., 18(3):605–654, (2005).
- [57] Lannes D., *The water waves problem: mathematical analysis and asymptotics*. Mathematical Surveys and Monographs, 188, 2013.
- [58] Levi-Civita T., *Détermination rigoureuse des ondes permanentes d’amplitude finie*. Math. Ann., 93, pp. 264–314, 1925.
- [59] Lindblad H., *Well-posedness for the motion of an incompressible liquid with free surface boundary*. Ann. of Math., 162, no. 1, 109–194, 2005.
- [60] Liu J., Yuan X., *A KAM Theorem for Hamiltonian partial differential equations with unbounded perturbations*, Comm. Math. Phys., 307(3):629–673, (2011).
- [61] Montalto R., *Quasi-periodic solutions of forced Kirchoff equation*, Nonlinear Differ. Equ. Appl. 24(9) (2017).
- [62] Plotnikov P., Toland J., *Nash-Moser theory for standing water waves*. Arch. Ration. Mech. Anal., 159(1):1–83, 2001.
- [63] Pöschel J., *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv., 71(2): 269–296, (1996).
- [64] Procesi C., Procesi M., *A KAM algorithm for the non-linear Schrödinger equation*, Advances in Math. 272, 399–470, (2015).
- [65] Stokes G., *On the theory of oscillatory waves*, Trans. Cambridge Phil. Soc. 8, 441–455, (1847).
- [66] Struik D., *Détermination rigoureuse des ondes irrotationnelles périodiques dans un canal à profondeur finie*, Math. Ann., 95: 595–634, (1926).
- [67] Wu S., *Almost global wellposedness of the 2-D full water wave problem*, Invent. Math., 177(1):45–135, (2009).
- [68] Wu S., *Global well-posedness of the 3-D full water wave problem*, Invent. Math., 184(1):125–220, (2011).
- [69] Zakharov V.E., *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, J Appl Mech Tech Phys 9:190–194, (1968).
- [70] Zakharov V.E., Dyachenko A.I., *Is free-surface hydrodynamics an integrable system?* Physics Letters A, 190(2):144–148, (1994).

UNIVERSITY OF NANTES

Email address: roberto.feola@univ-nantes.fr

UPC, BARCELONA

Email address: filippo.giuliani@upc.edu