Scuola Internazionale Superiore di Studi Avanzati - Trieste

SISSA-ISAS

PhD Course in Mathematical Analysis, Modelling and Applications

KAM for quasi-linear PDE's

Ph.D. Thesis

Supervisor:
Prof. Massimiliano Berti

Candidate:
Filippo Giuliani

Academic Year 2016/2017

SISSA - Via Bonomea 265 - 34136 TRIESTE - ITALY

Declaration

Il presente lavoro costituisce la tesi presentata da Filippo Giuliani, sotto la direzione del Prof. Massimiliano Berti, al fine di ottenere l'attestato di ricerca post-universitaria Doctor Philosophiae presso la SISSA, Curriculum in Analisi Matematica, Modelli e Applicazioni, Area di Matematica. Ai sensi dell'art. 1, comma 4, dello Statuto della Sissa pubblicato sulla G.U. no. 36 del 13.02.2012, il predetto attestato e' equipollente al titolo di Dottore di Ricerca in Matematica.

Trieste, Anno Accademico 2016-2017.

Abstract

In this Thesis we present two new results of existence and stability of Cantor families of small amplitude quasi-periodic in time solutions for quasi-linear Hamiltonian PDE's arising as models for shallow water phenomena.
The considered problems present serious small divisors difficulties and the results are achieved by implementing Nash-Moser algorithms and by exploiting pseudo differential calculus techniques.

The first result concerns a generalized quasi-linear KdV equation

$$u_t + u_{xxx} + \mathcal{N}_2(x, u, u_x, u_{xx}, u_{xxx}) = 0, \quad x \in \mathbb{T},$$

where $\mathcal{N}_2$ is a nonlinearity originating from a cubic Hamiltonian.
The nonlinear part depends upon some parameters and it is intriguing to study how the choice of these parameters affects the bifurcation analysis.
The linearized equation at the origin is resonant, namely the linear solutions are all periodic, hence the existence of the expected quasi-periodic solutions is due only to the presence of the nonlinearities.
The nonlinear terms of these equations are quadratic and contains derivatives of the same order of the linear part, thus they produce strong perturbative effect near the origin.

The second result is the first KAM result for quasi-linear PDE's with asymptotically linear dispersion law and it implies the first existence result for quasi-periodic solutions of the Degasperis-Procesi equation.
We consider Hamiltonian perturbations of the Degasperis-Procesi equation

$$u_t - u_{xxt} + u_{xxx} - 4u_x - uu_{xxx} - 3u_x u_{xx} + 4uu_x + \mathcal{N}_6(u, u_x, u_{xx}, u_{xxx}) = 0, \quad x \in \mathbb{T},$$

where $\mathcal{N}_6$ is a nonlinearity originating from a Hamiltonian density with a zero of order seven at the origin.
We exploit the integrable structure of the unperturbed equation $\mathcal{N}_6 = 0$ to overcome some small divisors problems.
The complicated symplectic structure and the asymptotically linear dispersion law make harder the analysis of the linearized operator in a neighborhood of the origin, which is required by the Nash-Moser scheme, and the measure estimates for the frequencies of the expected quasi-periodic solutions.

To my family

Acknowledgements

I am grateful to my supervisor, Massimiliano Berti, for having introduced me in the reality of the mathematical research, for his precious teaching and the help given me for the realization of this thesis.

I wish to express my deep gratitude to Michela Procesi and Roberto Feola for having supported me during my PhD experience and for the time spended to work together.

I wish also to thank Luca Biasco, Riccardo Montalto, Felice Iandoli, Alberto Maspero for their support and for many useful discussions.

A special thanks to Manuela because she's always been there for me.

Contents

CHAPTER 1

Introduction

In the last years many new results have been achieved in the theory of quasi-periodic motions for infinite dimensional, Hamiltonian, quasi-linear and fully nonlinear partial differential equations (PDE's), namely for nonlinear partial differential equations in which the linear and nonlinear terms contain derivatives of the same order. These progresses have been achieved thanks to the introduction of new ideas and techniques, among which the pseudo differential calculus plays an important role.
In this Thesis we present two new results of existence and stability of quasi-periodic solutions for quasi-linear generalized KdV equations (1.0.1) and for quasi-linear Hamiltonian perturbations of the Degasperis-Procesi equation (1.0.5), which arise as models for shallow water phenomena (see for instance [38], [42]).
More precisely, for both these PDE's we prove existence and linear stability of Cantor families of small amplitude quasi-periodic solutions (see Theorem 1.1.3 and Theorem 1.2.3).

(1) We first deal with the following family of quasi-linear generalized KdV equations

$$u_t + u_{xxx} + \mathcal{N}_2(x, u, u_x, u_{xx}, u_{xxx}) = 0, \tag{1.0.1}$$

under periodic boundary conditions $x \in \mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$, where

$$\mathcal{N}_2(x, u, u_x, u_{xx}, u_{xxx}) := -\partial_x[(\partial_u f)(x, u, u_x) - \partial_x((\partial_{u_x} f)(x, u, u_x))] \tag{1.0.2}$$

and f is the most general quasi-linear Hamiltonian density with 5-jet independent of x

$$\begin{aligned} f(x, u, u_x) := {} & c_1\, u_x^3 + c_2\, u_x^2\, u + c_3\, u^3 + c_4\, u_x^4 \\ & + c_5\, u_x^3\, u + c_6\, u_x^2\, u^2 + c_7\, u^4 + f_{\geq 5}(x, u, u_x), \end{aligned} \tag{1.0.3}$$

where the coefficients $c_i, i = 1, 2, \dots, 7$, are real numbers, and, for some $q > 0$ big enough,

$$f_{\geq 5}(x, u, u_x) := f_5(u, u_x) + f_{\geq 6}(x, u, u_x) \in C^q(\mathbb{T} \times \mathbb{R} \times \mathbb{R}; \mathbb{R}) \tag{1.0.4}$$

is the sum of the homogeneous component of f of degree five and all the higher order terms.

(2) Secondly we deal with quasi-linear Hamiltonian perturbations of the Degasperis-Procesi (DP) equation

$$u_t - u_{xxt} + u_{xxx} - 4u_x - uu_{xxx} - 3u_x u_{xx} + 4uu_x + \mathcal{N}_6(u, u_x, u_{xx}, u_{xxx}) = 0 \tag{1.0.5}$$

under periodic boundary conditions $x \in \mathbb{T}$, where

$$\mathcal{N}_6(u, u_x, u_{xx}, u_{xxx}) := -(4 - \partial_{xx})\partial_x[(\partial_u f)(u)] \tag{1.0.6}$$

and $f \in C^\infty(\mathbb{R};\mathbb{R})$ is a Hamiltonian density

$$f(u) = O(u^7), \tag{1.0.7}$$

where $O(u^7)$ denotes a function with a zero of order at least seven at the origin.

The linearized equation at the origin of (1.0.1) is the well-known Airy equation

$$u_t + u_{xxx} = 0 \tag{1.0.8}$$

and for the Degasperis-Procesi equation (1.0.5) is

$$u_t - u_{xxt} + u_{xxx} - 4u_x = 0. \tag{1.0.9}$$

All the solutions of (1.0.8) and (1.0.9) have the form

$$u(t,x) = \sum_{j\in\mathbb{Z}} u_j\, e^{\mathrm{i}(\omega(j)t + jx)}, \tag{1.0.10}$$

where $\omega(j) = j^3$ in the KdV case and $\omega(j) = j + 3j/(1+j^2)$ in the DP case. The function $j \mapsto \omega(j)$ is called (linear) *dispersion law* (or dispersion relation).
We note that both these problems are *resonant*, in the sense that the dispersion relations are rational, and the existence of quasi-periodic solutions for the equations (1.0.1) and (1.0.5) depends only on the presence of the nonlinearity.
In particular, the KdV case is *completely resonant*, namely all the solutions of the Airy equation (1.0.8) are 2π-periodic in time. Actually the linear situation of the DP problem is more delicate. Indeed all the functions of the form (1.0.10) with compact Fourier support are periodic, but with period depending on the support. This difference arises from the fact that, for any $j \in \mathbb{Z}$, $\omega(j) \in \mathbb{Z}$ in the KdV case, while $\omega(j) \in \mathbb{Q}$ in the DP case.

In the KdV case the dispersion law is superlinear as $j \to \infty$, in the DP case is *asymptotically linear*. This fact makes a significant difference in the study of quasi-periodic motions for these equations. In particular the case with less dispersion, namely the DP case, is much harder and we underline that Theorem 1.2.3, at the best of our knowledge, is the first KAM result for quasi-linear PDE's with this kind of dispersion, which also implies the first existence result for quasi-periodic solutions of the Degasperis-Procesi equation. This is part of a joint work with Roberto Feola and Michela Procesi. The result on the generalized quasi-linear KdV equations is contained in [61].

Hamiltonian PDE's and KAM theory. The Hamiltonian partial differential equations appear naturally in many areas of physics, especially to model the behaviour of idealised vibrating media (as waves on string or on the surface of a fluid), in the absence of friction or other dissipative forces. We briefly describe these equations following [42] and [21]. A Hamiltonian system is given in terms of a function $H \colon \mathcal{P} \to \mathbb{R}$, called *Hamiltonian*, defined on the phase space $\mathcal{P}$. We restrict to the

case of Hilbert phase spaces and we denote by $(\cdot,\cdot)$ the inner product of such spaces. In the infinite dimensional case, namely for PDE's, the Hamiltonian is usually well defined and smooth only on a dense subset of the phase space.
The symplectic structure is provided by a non-degenerate, anti-symmetric 2-form Ω defined by

$$\Omega(u,v) := (J^{-1}u, v), \quad u, v \in \mathcal{P},$$

where $J^{-1}\colon \mathcal{P} \to \mathcal{P}$ is a bounded operator with trivial kernel. By the anti-symmetry of Ω the operator J satisfies the following relation

$$J^{-T} = -J^{-1}.$$

The vector field X_H of the system with Hamiltonian H is defined, at least formally, by the relation

$$d_u H(u)[\cdot] = \Omega(X_H(u), \cdot), \quad u \in \mathcal{P}.$$

Through the inner product we can define also the gradient ∇H of the function H in the usual way

$$d_u H(u)[\cdot] = (\nabla H(u), \cdot), \quad u \in \mathcal{P}.$$

Thus the Hamiltonian vector field can be represented as

$$X_H(u) = J\,\nabla H(u), \quad u \in \mathcal{P}.$$

The Hamiltonian function H is a *constant of motion* for the system $u_t = X_H(u)$, in the sense that H assumes a constant value along its orbits. Indeed, if $\dot{u} = J\nabla H(u)$ then

$$\partial_t H(u(t)) = d_u H(u)[\dot{u}] = \Omega(X_H(u), \dot{u}) = \Omega(X_H(u), X_H(u)) = 0,$$

where we used, in the last relation, the anti-symmetry of the symplectic form.
For Hamiltonian PDE's defined on a compact spatial domain the existence of recurrence phenomena, as periodic or quasi-periodic motions, are expected. We recall that a function $u(t)$ is said *quasi-periodic* with frequency vector $\omega \in \mathbb{R}^\nu$ if there exists a function $\mathcal{U}\colon \mathbb{T}^\nu \to \mathbb{C}$ such that $u(t) = \mathcal{U}(\omega t)$ and ω is *irrational*, in the sense that $\omega \cdot \ell \neq 0$ for all $\ell \in \mathbb{Z}^\nu$, $\ell \neq 0$.
In this Thesis we study the existence and the stability of such motions for one dimensional Hamiltonian quasi-linear PDE's with periodic boundary conditions, namely with spatial variable x belonging to the compact torus manifold $\mathbb{T}$.

In the sequel we refer to some well-known Hamiltonian PDE's. Here we list part of them.

- The Korteweg de Vries (KdV) equation

$$u_t + u_{xxx} + 6uu_x = 0, \tag{1.0.11}$$

  where $u\colon \mathbb{R} \times \mathbb{T} \to \mathbb{R}$ represents the (low) amplitude of a shallow water wave $u(t,x)$.

- The nonlinear Schrödinger (NLS) equations

$$\mathrm{i}u_t + \Delta u + V(x)u = \pm |u|^{p-1}u, \qquad p \geq 2, \tag{1.0.12}$$

  where $u\colon \mathbb{R} \times \mathbb{T}^d \to \mathbb{C}$ is the wave function of a quantum particle and $V(x)$ is a real valued multiplicative potential.

- The nonlinear wave (NLW) equation and the Klein-Gordon equation

$$u_{tt} - \Delta u + V(x)u + \mathcal{N}(u) = 0, \qquad u_{tt} - \Delta u + mu + \mathcal{N}(u) = 0, \tag{1.0.13}$$

  where $u\colon \mathbb{R} \times \mathbb{T}^d \to \mathbb{R}$ represent the amplitude of a wave $u(t, x)$ at a point in spacetime, $\mathcal{N}(u)$ is the nonlinear vector field of a Hamiltonian $\int_{\mathbb{T}^d} F(u)\, dx$ for some Hamiltonian density F, $V(x)$ is a real valued multiplicative potential and m is a mass parameter.

- The water waves equations (we refer to [42] for a detailed presentation of the equations)

$$\begin{cases} \eta_t = G(\eta)\psi, \\ \psi_t + g\eta + \frac{1}{2}\psi_x^2 - \dfrac{1}{2}\dfrac{(G(\eta)\psi + \eta_x\psi_x)^2}{1 + \eta_x^2} = \kappa \dfrac{\eta_{xx}}{(1 + \eta_x^2)^{3/2}}, \end{cases} \tag{1.0.14}$$

  written in the Hamiltonian formulation following Zakahrov [96] and Craig-Sulem [44]. Here $G(\eta)$ is the so-called Dirichlet-Neumann operator, $\eta(t)$ is the profile of the fluid at time t and $\psi(t, x) = \Phi(t, x, \eta(t, x))$ is the restriction of the velocity potential Φ to the free boundary. We note also the presence of two parameters: the capillarity κ and the gravity g.

Near the origin these models behave like an infinite system of harmonic oscillators coupled by the nonlinear terms. We refer to the modes and the frequencies of these oscillators as linear or *normal modes* and *linear frequencies* of oscillations. It is natural, in this context, to approach these problems as infinte dimensional dynamical systems.
For instance if the functional space $\mathcal{P}$ is equipped with a basis in which the linearized operator at the origin is diagonal then the equations above can be written as infinite systems of ordinary differential equations for the coefficients of $u \in \mathcal{P}$ with respect to such basis.
One of the most important ideas carried out from the dynamical systems theory is to look for invariant manifolds (equilibria, periodic solutions, ...) on which the dynamic is simple, in order to deduce the behaviour of the orbits near these sets. A way to do that is by suitable perturbative techniques.
For instance it is natural to look for invariant subsets of the phase space on which the system is *integrable*. There exist different notions of integrability, but all these ones imply that, in some sense, the orbits are explicitly computable. In the finite dimensional Hamiltonian context one of the most important notion is the Liouville-integrability. A system is integrable in this sense if there exists a linearly independent maximal set of Hamiltonians which Poisson commute one each other. As a consequence there exists a foliation of the phase space in invariant submanifolds and a set of canonical coordinates (*action-angle variables*) in which the dynamic on these submanifolds is easily described.
Since Poincaré, many mathematicians have been interested in the study of Hamiltonian systems close to integrable ones (also said *nearly integrable*). These ones can be seen as small perturbations of Hamiltonian integrable systems.
In the 50's-60's Kolmogorov [71] and Arnold [3] provided a fundamental result for the theory of such systems. In the case of analytic perturbations of analytic Hamiltonians, they proved, under suitable *non-degeneracy conditions*, the existence of a positive measure set of initial data from which quasi-periodic motions originate. Later on, Moser extended this result for finite differentiable perturbations and for reversible systems, see [80], [83]. This branch of the dynamical systems theory has been called KAM (Kolmogorov, Arnold and Moser) theory.
The idea of the KAM method developed by these authors is to produce, by an iterative scheme,

a sequence of change of coordinates which puts the perturbed Hamiltonian in a normal form with an invariant torus at the origin. From another point of view this scheme provides a sequence of approximately invariant tori which converges to a final invariant torus.
The main difficulty in implementing this procedure is due to the presence of *small divisors*, which enter at the denominator of the Fourier coefficients of the approximate solution at each step of the iteration. In particular the small divisors are numbers

$$\omega \cdot \ell, \qquad \ell \in \mathbb{Z}^\nu, \tag{1.0.15}$$

where $\omega \in \mathbb{R}^\nu$ is the frequency of oscillation of the torus. The problem of the small divisor persists even if the frequency of the torus is irrational, namely if the quantity (1.0.15) is not zero, since the Fouries series of the approximate solution might not converge if $\omega \cdot \ell$ is too small. Actually it is well known that for almost every $\omega \in \mathbb{R}^\nu$ the set $\{\omega \cdot \ell, \, \ell \in \mathbb{Z}^\nu\}$ accumulates to zero. This problem is overcome by imposing non-resonance *diophantine* conditions of the form

$$|\omega \cdot \ell| \geq \frac{\gamma}{|\ell|^\tau} \qquad \ell \in \mathbb{Z}^\nu \setminus \{0\}, \gamma \in (0,1), \tag{1.0.16}$$

which, combined to a quadratic Newton-type scheme, provide the convergence of the sequence of approximate solutions. The inequality (1.0.16) is also called 0-th Melnikov non-resonance condition.
The divisors (1.0.15) appear typically when one looks for invariant Lagrangian tori, namely tori of maximal dimension, as in the cases considered in the aforementioned papers. Moser [83], Eliasson [51] and Pöschel [85] developed a theory for the search of quasi-periodic solutions supported on lower dimensional tori. In these cases other conditions are required

$$|\omega \cdot \ell \pm d_j| \geq \frac{\gamma}{\langle \ell \rangle^\tau} \qquad \ell \in \mathbb{Z}^\nu, \; j \in \mathbb{Z}, \; \gamma \in (0,1), \tag{1.0.17}$$

$$|\omega \cdot \ell + (d_j \pm d_k)| \geq \frac{\gamma}{\langle \ell \rangle^\tau} \qquad \ell \in \mathbb{Z}^\nu, \; j,k \in \mathbb{Z}, \; \gamma \in (0,1), \tag{1.0.18}$$

where d_j are the normal (to the torus) frequencies. If the sign in (1.0.18) is a minus then this bound should hold for all $(\ell, j, k) \neq (0, j, j)$. The relations (1.0.17), (1.0.18) are called 1-st and 2-nd Melnikov non-resonance conditions.
We note that a condition as (1.0.18) with a minus sign implies that the frequencies d_j and d_k are distinct (case $\ell = 0$). Hence, in a case with multiple eigenvalues, the 2-nd order Melnikov conditions do not hold.

KAM for PDE's. The extension of the classical KAM theory to the infinite dimensional case is called KAM (theory) for PDE's. In the infinite dimensional framework, except for some particular cases ([87], [36], [35]), the search for Lagrangian invariant tori, or *almost periodic* solutions, is out of reach up to now. Hence we will focus on the study of periodic and quasi-periodic motions. This could be seen as the counterpart of the analysis of lower dimensional tori for the finite dimensional case.

In the infinite dimensional context the small divisors problems are much harder (we refer to the monograph [41] and [12] for a recent survey).
Indeed, also for the search of time periodic solutions for an infinite dimensional Hamiltonian system, small divisor issues arise. To overcome such problems, conditions as (1.0.17) or (1.0.18) are required

for an infinite number of normal frequencies. In order to prove that the set of frequencies that satisfy such (infinite) conditions has a large measure (*measure estimates*) one can exploit the presence of external parameters which control the perturbed frequencies, such as the mass m in the Klein-Gordon equation (1.0.13), the capillarity κ and the depth in the water waves equations (1.0.14) and the potential $V(x)$ in NLS (1.0.12) and NLW (1.0.13).
When there are no outer parameters one can extract them from the equation by performing a Birkhoff normal form. This fact has been highlighted by Kuksin-Pöschel in [75] and Pöschel in [86].

In the 90's, at the beginning of the study of periodic and quasi-periodic solutions for PDE's, two main approaches have been developed:

- **normal form KAM methods**,
- **Newton Nash-Moser implicit function iterative scheme**.

The first one is a generalization of the theory developed by Eliasson [51] and Pöschel [85] for lower dimensional tori. This method consists in a Newton-type iteration which brings, up to smaller and smaller remainders, the Hamiltonian into a normal form with an invariant torus at the origin by using canonical transformations. Here the small divisors problem arises in the so-called *homological equations*, which one needs to solve at each step in order to find a suitable symplectic change of coordinates which reduces the size of the remainders. Such equations are constant coefficients linear PDE's and to solve them one needs to impose 2-nd order Melnikov non-resonance conditions. The final KAM invariant torus will be *reducible*, in the sense that the linearized equation at it will be constant coefficient.

The second one is a method proposed first by Craig-Wayne in [45] for the search of periodic solutions for problems with periodic boundary conditions, for which double eigenvalues arise, and extended by Bourgain [30] for the search of quasi-periodic solutions and for PDE's in higher space dimension, see [32]. In all these cases the 2-nd order Melnikov conditions are violated.
After a Lyapunov-Schmidt decomposition, the search of invariant tori is reduced to solve some non-linear functional equation for the embedded torus. By means of a quadratic Newton-type scheme, the solutions are found as limit of a sequence of approximate solutions. This scheme requires to invert the linearized operator at any approximate solution and to do that, a priori, only Melnikov conditions of first order type are needed. As a consequence of having imposed only these conditions, the PDE's which one has to solve at any step are non-constant coefficients.

Actually Berti-Bolle in [21] pointed out that these two approaches have many connections (see also [39]). They highlighted the fact that around an (approximately) invariant torus there always exists a (approximate) Hamiltonian normal form. One can see that the stability in the actions of this normal form is actually a consequence of the Hamiltonian structure. Hence the difference between the two methods above lies in whether or not one diagonalizes the part of the normal form which gives the linear dynamic in the normal (to the torus) directions.
In this perspective the authors in [21] presented a general approach to the problem of finding quasi-periodic solutions for Hamiltonian systems based on a Nash-Moser implicit function theorem. Since we adopt this approach for the results presented in this Thesis we describe more in detail the Nash-Moser scheme following this point of view.

A quasi-periodic solution $u(\omega t)$ with frequency $\omega \in \mathbb{R}^\nu$ of a system with Hamiltonian H can be seen as an embedding $i := i(\varphi)$ into the phase space of the ν-dimensional torus $\mathbb{T}^\nu$ supporting the Kronecker flow $\Psi^t_\omega(\varphi) = \varphi + \omega t$ and such that

$$i \circ \Psi^t_\omega = \Phi^t_H \circ i, \tag{1.0.19}$$

where Φ^t_H is the flow of the Hamiltonian system. From a functional point of view (1.0.19) is equivalent to the equation

$$\mathcal{F}(i) := \omega \cdot \partial_\varphi i - X_H(i) = 0. \tag{1.0.20}$$

In this formulation the search for quasi-periodic solutions is reduced to the search for zeros of the functional $\mathcal{F}$ in (1.0.20). The sequence of approximate solutions i_n defined by the Nash-Moser algorithm is the following (given a sufficiently good approximate solution i_0)

$$i_{n+1} := i_n - (\mathcal{L}(i_n))^{-1}\mathcal{F}(i_n) \qquad \mathcal{L}(i_n) := d_i\mathcal{F}(i_n). \tag{1.0.21}$$

Usually $\mathcal{F}$ is defined on a scale of functional spaces since the inverse of the linearized operators $\mathcal{L}(i_n)$ loses derivatives and the classical implicit function theorems do not apply to solve (1.0.20). Consider for example the scale of Sobolev spaces

$$H^s(\mathbb{T}^{\nu+1}) = \Big\{ u = \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} u_{\ell j} e^{\mathrm{i}(\varphi\cdot\ell + jx)} \,:\, \|u\|_s^2 := \sum_{(\ell,j)\in\mathbb{Z}^{\nu+1}} |u_{\ell j}|^2 \langle \ell, j\rangle^{2s} < +\infty \Big\} \tag{1.0.22}$$

and the linear operator (see (1.0.8))

$$\mathcal{L}_{\mathrm{Airy}} := \omega \cdot \partial_\varphi + \partial_{xxx}.$$

The Fourier representation of its inverse involves the small divisors

$$\mathrm{i}\,\omega \cdot \ell - \mathrm{i}\, j^3, \quad \ell \in \mathbb{Z}^\nu, \quad j \in \mathbb{Z}$$

and by imposing 1-st order Melnikov non-resonance conditions as (1.0.17) with $d_j = -j^3$, the best estimate that one obtains is the following

$$\|\mathcal{L}^{-1}_{\mathrm{Airy}} g\|_s \leq \Big(\sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \frac{|g_{\ell j}|^2}{|\omega\cdot\ell - j^3|^2} \langle \ell, j\rangle^{2s} \Big)^{1/2} \overset{(1.0.17)}{\leq} C\gamma^{-1}\|g\|_{s+\tau}.$$

Note that in this case there is a *loss* of τ derivatives, namely $\mathcal{L}^{-1}_{\mathrm{Airy}} \colon H^{s+\tau}(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$ for any s.

The Nash-Moser scheme also requires to estimate the inverse $\mathcal{L}^{-1}$ in high Sobolev norms at each step. This task may be very hard since the linearized equations are PDE's with non-constant coefficients, represented by differential operators which are small perturbations of a diagonal operator with arbitrarily small eigenvalues. Moreover the tangential and the normal components to the torus of the linearized equations are strongly coupled.

In [21] the authors constructed a symplectic change of variables in which the linearized system at an (approximately) invariant torus is (approximately) triangularized. Then the problem is reduced to the study of the linearized operator in the normal directions only. We refer to Chapter 3 for a more detailed description of this procedure.

A way to obtain good estimates on the inverse of the linearized operator in the normal directions is to diagonalize it. Actually, in dimension one this task is usually doable.
We underline that for this purpose also 2-nd order Melnikov non-resonance conditions are required. Let us describe a general case of *linear reducibility*, namely the diagonalization procedure for a linear operator. Consider the linear operator

$$\mathcal{L} := \omega \cdot \partial_\varphi + D + \varepsilon R, \tag{1.0.23}$$

where $D = \mathrm{diag}_{j\in\mathbb{Z}}(\mathrm{i}d_j)$ is a diagonal matrix such that $d_j = d_k$ only for $j = k$. We call $\omega \cdot \partial_\varphi + D$ the *normal part* or *normal form* of $\mathcal{L}$ and we consider $\varepsilon R = \varepsilon R(\omega t)$ as a small, quasi-periodically time-dependent perturbation. For instance, these matrices can represent, in a Fourier basis, linear operators on Hilbert spaces. It is well known since Moser [83] that, in the finite dimensional framework, if the eigenvalues of the normal part are well separated then there exist, for ε small enough, a change of coordinates Φ such that

$$\Phi^{-1}\mathcal{L}\Phi = \omega \cdot \partial_\varphi + D_\infty, \tag{1.0.24}$$

where D_∞ is a diagonal operator. This task can be carried out also in the infinite dimensional case under suitable assumptions on the spectrum of the operators D, R and the frequency ω.
The function Φ in (1.0.24) is constructed iteratively as limit of a sequence of transformations $(\Upsilon_n)_{n\geq 1}$, which are close to the identity . Let us show one step of this iteration.
Let us consider a transformation of the form $\Upsilon = \mathrm{I} + \varepsilon\Psi$, where $\Psi = \Psi(\omega t)$. Then

$$\mathcal{L}_+ := \Upsilon^{-1}(\omega \cdot \partial_\varphi + D + \varepsilon R)\Upsilon = \omega \cdot \partial_\varphi + D + \varepsilon(\omega \cdot \partial_\varphi \Psi + [D, \Psi] + R) + O(\varepsilon^2)$$

and we look for Ψ that solves the following homological equation

$$\omega \cdot \partial_\varphi \Psi + [D, \Psi] + R = 0. \tag{1.0.25}$$

At the level of the entries of the matrices the equation (1.0.25) reads as follows

$$\mathrm{i}(\omega \cdot \ell + d_j - d_k)\Psi_j^k(\ell) = -R_j^k(\ell). \tag{1.0.26}$$

Clearly if $\ell = 0$ and $j = k$ this equation cannot be solved and the diagonal terms $R_j^j(0)$ will contribute to the new normal form. These terms are called *resonant*. If $\ell \neq 0$ or $j \neq k$ one can choose Ψ such that (1.0.26) holds by imposing 2-nd Melnikov non-resonance conditions

$$|\omega \cdot \ell + d_j - d_k| \geq \frac{\gamma}{\langle \ell \rangle^\tau} \qquad \forall (\ell, j, k) \neq (0, j, j). \tag{1.0.27}$$

The choice of Ψ in (1.0.25) brings the new operator $\mathcal{L}_+$ in the form $\omega \cdot \partial_\varphi + D_+ + \varepsilon^2 R_+$, where $D_+ := D + \varepsilon R_j^j(0)$, $\omega \cdot \partial_\varphi + D_+$ is the new normal part and $\varepsilon^2 R_+$ the new remainder. We underline the quadratic reduction of the size of the latter.
The convergence of this iteration has to be quite fast in order to overcome the loss of derivatives due to the small divisors in (1.0.26).

Note that in general it is not easy to impose conditions like (1.0.27), in particular when the dispersion relation is not so strong to guarantee (enough) separation between the eigenvalues $\mathrm{i}d_j$ and $\mathrm{i}d_k$. We point out two important features which usually allow to impose these conditions:

- the fact that the eigenvalues of the linear operator are simple,
- a good knowledge of the asymptotic expansion of the eigenvalues.

Even when one of these two facts does not hold one can hope to impose 2-nd Melnikov relations. As an example, in the recent work of Berti-Kappeler-Montalto [26] on the defocusing NLS with periodic boundary conditions, a case with double eigenvalues, the authors imposed these non-resonance conditions thanks to the good asymptotic of the eigenvalues.

Historical preface and literature. The KAM theory for PDE's has been developed in the eighties by Kuksin [72] and Wayne [95] for the one dimensional parameter dependent nonlinear wave and Schrödinger equations with Dirichlet boundary conditions. Then Kuksin-Pöeschel [75] and Pöeschel [86] extended these results for the one dimensional parameter independent nonlinear wave and Schrödinger equations with analytic perturbations. We remark that these results were restricted to Dirichlet boundary conditions in order to assure the 2-nd order Melnikov conditions. Indeed in this case the normal frequencies are simple. This is not true, for instance, for periodic boundary conditions.
At the beginning of the 90's Craig-Wayne [45] provided a generalization of the Lyapunov center theorem in a infinite dimensional non-resonant case. These authors proposed a Nash-Moser-Newton method to find periodic solutions of the 1-d nonlinear Klein-Gordon and Schrödinger equations under periodic boundary conditions and, in [46], for analytic perturbations of the defocusing NLS (dNLS). In this approach the reducibility of the linearized operator it is not required and the linear equations to solve are PDE's with variable coefficients, thus 2-nd Melnikov conditions are not needed.
This method has been generalized for completely resonant cases by Berti-Bolle [16], [17], still for periodic solutions. We mention also for completely resonant problems Gentile-Mastropietro-Procesi [60] (see also the monograph [11]). In [37] Chierchia-You discussed the problem of double eigenvalues for the one dimensional NLW equation with periodic boundary conditions. They proposed a version of a KAM method in which the normal form is not diagonal, but only block (2×2) diagonal. We remark that all these works treat one dimensional *semi-linear* cases, namely when the order of the derivatives in the nonlinearity is strictly less than the order of the derivatives in the linear part.

In higher space dimensions, different techniques have been adopted since the 2-nd order Melnikov conditions are violated, for instance, by the multiplicity of the eigenvalues.
Bourgain in [31], [34], [32] extended the Craig-Wayne approach to get estimates of the inverse of the linearized operators in high norms without imposing second Melnikov conditions. In the mentioned papers these techniques are applied by the author for the search of quasi-periodic solutions for analytic NLS and NLW with convolution potential.
These methods have been further generalized by Wang [94] for completely resonant NLS and by Berti-Bolle [20], [19], [22] for NLS, in the forced case, and for forced and autonomous NLW with a multiplicative potential and differentiable nonlinearities. All these works are based on *multiscale analysis* which involve only conditions like the first order Melnikov relations and the linear stability of the solutions is not implied.
The first results on the existence of reducible KAM tori are due to Geng-You [58] and Eliasson-Kuksin [54] for NLS with convolutive potential on $\mathbb{T}^d$. In the latter, the second order Melnikov conditions are verified thanks to the introduction of the notion of Töplitz-Lipschitz Hamiltonians.

In the aforementioned paper the space variable $x \in \mathbb{T}^d$. For problems on a more general spatial domain we cite the papers of Berti-Procesi [29] for periodic solutions of the nonlinear wave and Schrödinger equations on compact Lie groups and homogeneous spaces, and Berti-Corsi-Procesi [24], in which the authors provided an abstract Nash-Moser implicit function theorem with applications to the quasi-periodic case for NLW and NLS on compact Lie groups. In 2011 Geng-Xu-You [59] proved a KAM result for the cubic NLS on $\mathbb{T}^2$. Later on, Procesi-Procesi [89] provided a normal form result for the completely resonant NLS with periodic boundary conditions in any dimension.
Procesi-Xu [91] introduced the notion of quasi-Töplitz functions and exploited it to prove existence and stability of quasi-periodic solutions for NLS on $\mathbb{T}^d$. In [90] Procesi-Procesi provided the existence of quasi-periodic solutions for the completely resonant NLS with periodic boundary conditions, in any dimension, by using the results of [89] and [91]. We cite also the reducibility result of Eliasson-Kuksin [53] and the recent work on the Beam equation on $\mathbb{T}^d$ by Eliasson-Grebert-Kuksin [52]. All the aforementioned papers treat problems with bounded perturbations.

To explain the main issues of the cases with unbounded perturbations we refer to the linear reducibility model proposed above, recall (1.0.23). Suppose that the perturbation R is an unbounded operator. Then a priori the trasformation Ψ defined by (1.0.26) loses derivatives and, along the iteration, the order of the transformed vector field $\mathcal{L}_+$ increases quadratically. Moreover it is not even clear that the transformation Ψ is invertible. In general this is not true.
The first KAM result for systems with unbounded perturbation is due to Kuksin [73] and Kappeler-Pöeschel [70] for Hamiltonian analytic perturbations of the KdV equation

$$u_t + u_{xxx} - 6uu_x + \varepsilon \partial_x f(x, u) = 0$$

with periodic boundary conditions. Note that the nonlinearity contains one derivative. The key ideas were to exploit the strong dispersion of KdV (recall $\omega(j) = j^3$) by imposing stronger non-resonance conditions

$$|\omega \cdot \ell + j^3 - k^3| \geq \frac{\gamma}{\langle \ell \rangle^\tau} |j^3 - k^3|, \quad |j^3 - k^3| \geq \frac{j^2 + k^2}{2}, \qquad \text{for } j \neq k, \tag{1.0.28}$$

which give a gain of regularity, in particular two derivatives, and to insert the remaining diagonal angle-dependent terms (for $j = k$) in the normal form. In this way the homological equations have variable coefficients and one has to be able to solve them. This was the purpose of the so-called *Kuksin Lemma.* Note that such homological equations are scalar and so they are much easier than the variable coefficients functional equations which appear in the Newton-Nash-Moser approach of Craig-Wayne-Bourgain.
Later on Liu-Yuan in [76] extended the result of Kuksin [73] and Kappeler-Pöeschel [70] for the less dispersive case of the NLS with one derivative in the nonlinearity. This result is based on a improvement of the Kuksin Lemma. We mention also the result of Zhang-Gao-Yuan [98] for the derivative NLS.
The above method does not apply in the case of the derivative NLW (DNLW) which contains the first order derivatives ∂_x, ∂_t in the nonlinearity.
In [14]-[15] Berti-Biasco-Procesi provided existence of quasi-periodic solutions for the DNLW in the Hamiltonian case

$$u_{tt} - u_{xx} + mu + g(Du) = 0 \qquad D := \sqrt{-\partial_{xx} + m}, \quad x \in \mathbb{T}$$

and in the reversible case

$$u_{tt} - u_{xx} + mu + g(x, u, u_t, u_x) = 0,$$

when

$$g(x, u, u_t, u_x) = g(x, u, -u_t, u_x), \qquad g(x, u, u_t, u_x) = g(-x, u, u_t, -u_x).$$

We remark that in this case the dispersion law is linear. We quote also a previous result on the DNLW due to Bourgain [33] whose extended the Craig-Wayne approach [45]. The papers quoted above concern cases with semi-linear perturbations.

In 2008 Baldi [4] proved via a Nash-Moser method the existence of periodic forced vibrations of the quasi-linear Kirchoff equations

$$u_{tt} - \Big(1 + \int_\Omega |\nabla u|^2 \, dx\Big)\Delta u = \varepsilon f(\omega t, x), \quad x \in \Omega \tag{1.0.29}$$

with Dirichlet boundary conditions for $\Omega \subset \mathbb{R}^d$ and for periodic boundary conditions $\Omega = \mathbb{T}^d$.
For the water waves equations, which are fully nonlinear, the previous results are not applicable. For water waves problems which do not involve small divisors we mention the work of Craig-Nicholls [43] for the existence of periodic travelling waves with capillarity.
For small divisors problems we cite the pioneering work of Iooss-Plotnikov-Toland [69] about the existence of periodic standing waves for the water waves equations and of Iooss-Plotnikov [67], [68] for 3-d travelling waves.
These results are obtained via Nash-Moser methods combined with a Lyapunov-Schimdt reduction. The authors realized a new technique for the analysis of the linearized operator, which successively inspired the first works on quasi-periodic solutions for quasi-linear and fully nonlinear problems due to Baldi-Berti-Montalto [7], [8]. To better explain this point we refer to the KdV model considered for the search of periodic solutions, since we will discuss shortly the quasi-periodic cases [7], [8] on the KdV equation. We consider the linear operator

$$\mathcal{L}_\omega := \omega\partial_t + a_3(\omega t, x)\partial_{xxx} + a_2(\omega t, x)\partial_{xx} + a_1(\omega t, x)\partial_x + a_0(\omega t, x), \quad x \in \mathbb{T}, \tag{1.0.30}$$

where $a_3 = 1 + O(\varepsilon)$, $a_2, a_1, a_0 = O(\varepsilon)$.
The key idea is to reduce the order (as pseudo differential operator) of $\mathcal{L}_\omega$ instead of reducing the size of the perturbation, as we did in the reducibility example (1.0.23). The operator $\mathcal{L}_\omega$ is conjugated to a new operator $\mathcal{L}_+$ which is of the form

$$\mathcal{L}_+ = \omega\partial_t + D + R(\omega t, x), \quad D := m_3\partial_{xxx} + m_1\partial_x, \tag{1.0.31}$$

where $m_3 = 1 + O(\varepsilon)$ and $m_1 = O(\varepsilon)$ are real constants, hence the part $\omega\partial_t + D$ is diagonal, and R is a bounded remainder. This conjugation is achieved by using changes of coordinates as diffeomorphisms of the torus $\mathbb{T}^2$ and pseudo differential maps. We call this method *regularization procedure.*
One can perform other steps of this algorithm to regularize the remainder R, but it is not possible to iterate it infinitely many times because the convergence of this scheme is not sufficiently fast to overcome the loss of derivatives due to the small divisors. In the applications the success of this procedure depends highly on the PDE in exam. Once R is smoothing enough, namely once

$(\omega\partial_t + D)^{-1}R$ is bounded, one can invert $\mathcal{L}_+$ by Neumann series.
Note that in the periodic case $\omega \cdot \ell = \omega\ell$ and this quantity is approximately ℓ if $\omega \neq 0$. The eigenvalues of $\omega\partial_t + D$ are

$$\mathrm{i}(\omega\ell - m_3 j^3 + m_1 j), \quad \ell \in \mathbb{Z}, \quad j \in \mathbb{Z},$$

then for inverting $\mathcal{L}_+$ by Neumann series one has to deal with these numbers at the denominator of the Fourier representation of $\mathcal{L}_+^{-1}$. These quantities are small if $\ell = O(j^3)$. Thus the loss of derivatives of the small divisor, which is a loss in time, can be compensated by R if this operator is regularizing enough (also only in space), in this example if $R = O(\partial_x^{-d})$, $d \geq 3$.
In the quasi-periodic case $\omega \cdot \ell \neq \omega\ell$ and the above argument fails.

In [5] Baldi, inspired by the work of Iooss-Plotnikov-Toland, proved the existence of periodic solutions for fully nonlinear perturbations of the Benjamin-Ono equation

$$u_t + \mathcal{H}u_{xx} + \partial_x(u^3) + \mathcal{N}_4(u) = 0, \quad x \in \mathbb{T},$$

where $\mathcal{H}e^{\mathrm{i}jx} = -\mathrm{i}\,\mathrm{sign}(j)\, e^{\mathrm{i}jx}$, $j \in \mathbb{Z}$ is the Hilbert transform and $\mathcal{N}_4$ is a quasi-linear or fully nonlinear perturbation $O(u^4)$. In the regularization procedure the author consider changes of coordinates which preserve the dynamical structure of the equation.
For fully nonlinear problems we cite also the paper of Alazard-Baldi [1] for periodic standing waves for gravity-capillary water waves equations.
All these works provide results of existence for time periodic solutions.

The first breakthrough results for quasi-periodic solutions of quasi-linear and fully nonlinear PDE's are due to Baldi-Berti-Montalto for the forced Airy equation in [7]

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

and in [8], [9] for the autonomous KdV equation

$$u_t + u_{xxx} - 6uu_x + \mathcal{N}_4(x, u, u_x, u_{xx}, u_{xxx}) = 0, \quad x \in \mathbb{T} \tag{1.0.32}$$

and the mKdV equation

$$u_t + u_{xxx} \pm \partial_x u^3 + \mathcal{N}_4(x, u, u_x, u_{xx}, u_{xxx}) = 0, \quad x \in \mathbb{T}, \tag{1.0.33}$$

where $\mathcal{N}_4$ is a nonlinearity with a zero of order four at the origin and a 5-jet independent of x.
These results are proved via Nash-Moser methods. The authors in [7] extended the regularization procedure of Iooss-Plotnikov-Toland for the study of the quasi-periodically forced linearized KdV operator

$$\mathcal{L}_\omega := \omega \cdot \partial_\varphi + a_3(\varphi, x)\partial_{xxx} + a_2(\varphi, x)\partial_{xx} + a_1(\varphi, x)\partial_x + a_0(\varphi, x), \quad x \in \mathbb{T},\ \varphi \in \mathbb{T}^\nu. \tag{1.0.34}$$

After many changes of coordinates the operator $\mathcal{L}_\omega$ is conjugated to an operator $\mathcal{L}_+$ of the form

$$\mathcal{L}_+ = \omega \cdot \partial_\varphi + D + R(\varphi, x), \quad D := m_3\partial_{xxx} + m_1\partial_x,$$

where $m_3 = 1 + O(\varepsilon)$ and $m_1 = O(\varepsilon)$ are real constants. We saw that in the quasi-periodic case it is not possible to invert $\mathcal{L}_+$ by Neumann series. In [7] the authors performed a KAM algorithm to completely diagonalize $\mathcal{L}_+$ thanks to a good control on the asymptoic expansion of the eigenvalues of (1.0.34)

$$\mathrm{i}(\omega\cdot\ell+d_j),\quad d_j=-m_3 j^3+m_1 j+r_j,\qquad |r_j|=O(\varepsilon),\quad j\in\mathbb{Z},$$

which allows to impose 2-nd order Menlnikov non-resonance conditions. The KAM scheme is initialized by a smallness assumption on the norm of the remainder R in (1.0.31) and exploits the boundness of this operator.
In contrast with the works of Iooss-Plotnikov-Toland the transformations involved in the regularization procedure are quasi-periodically time dependent maps of the phase space and quasi-periodic reparametrization of time. Hence the dynamical structure of the equation is preserved. Moreover these transformations are tame and bounded on H^s (recall (1.0.22)) for any s in a suitable (large) range. Thus one can easily get tame estimates for the inverse of the diagonal operator and then come back to the original (physical) coordinates obtaining tame estimates on $\mathcal{L}_\omega^{-1}$ (see (1.0.34)). This procedure provides also the *stability* of the final solutions, which is not guaranteed by the results mentioned above.
In [8], [9] the systems considered are autonomous and different problems arise. The frequency of the expected solutions are unknown and a careful bifurcation analysis is needed. The authors implemented the general strategy proposed by Berti-Bolle in [21] in order to reduce the analysis of the linear equations to the inversion of the quasi-periodically forced PDE (1.0.34). In Section 3 we describe more in details this approach by refering to the strategy of the results presented in this Thesis.
Later on Feola-Procesi [56], [55] extended the results of Baldi-Berti-Montalto [7], [8] in order to prove existence and stability of quasi-periodic solutions for fully nonlinear perturbations of the reversible and Hamiltonian Schrödinger equation

$$\mathrm{i}u_t=u_{xx}+\varepsilon f(\omega t,x,u,u_x,u_{xx}),\quad x\in\mathbb{T},$$

where $f\in C^q(\mathbb{T}^{\nu+1}\times\mathbb{C}^3)$ for some large $q>0$. We mention also Montalto [79], [78] for quasi-periodic solutions of the Kirchoff equation (1.0.29) and a reducibility result for a class of linear wave equations with unbounded perturbations on $\mathbb{T}^d$.
Recently Berti-Kappeler-Montalto [26] provided the existence of finite dimensional invariant tori of any size by perturbing the finite-gap solutions of the defocusing NLS

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

The dNLS frequencies are asymptotically double, but, adapting the strategy of [8], the authors obtained an asymptotic expansion that allows to impose 2-nd order Melnikov conditions and prove the linear stability of the solutions. This result extended the work of Kuksin-Pöschel [75] for small amplitude quasi-periodic solutions of dNLS under Dirichlet boundary conditions.
For related results on the dNLS equation, we cite also the paper [57] by Geng-You which exploit the momentum to deal with resonant frequencies.

In 2016 Berti-Montalto [27] (see also [28]) extended the work of Alazard-Baldi [1] by providing the existence of quasi-periodic, instead of periodic, solutions for water waves equations with capillarity (see (1.0.14)). In this work the authors followed the general strategy proposed in [8] and [9]. The analysis of the linearized operator is performed by fully exploiting pseudo differential calculus

techniques, particularly a careful Egorov analysis. The KAM reducibility iteration is implemented on a more sophisticated class of operators, (modulo-tame operators, which are introduced in Section 2), which is closed by the operations involved in the KAM scheme, namely the composition, the projection and the solution of the homological equations. We note that in this case the surface tension κ is used to control the Melnikov non-resonance conditions via the so-called degenerate KAM theory (see [10]). Indeed the frequencies behaves like $\kappa\, j^{3/2}$. We note that the dispersion law is superlinear as $j \to \infty$.
Very recently Baldi-Berti-Haus-Montalto [6] extended the above result providing existence and stability of small amplitude quasi-periodic solutions for the water waves equations with finite depth under the action of pure gravity. In this case the dispersion relation $\omega(j) \sim \sqrt{j}$ is sublinear and the 2-nd Melnikov conditions produces a loss of derivatives both in time and space. This loss is compensated by a stronger regularization procedure.
For new results on the water waves equations we cite also the paper [25] by Berti-Delort on the almost global existence of solutions for the gravity-capillary case with periodic spatial boundary conditions.

Comments on the results of the Thesis. The results of this Thesis follow the strategy of the papers [8], [9], [6], [27] but differ from these problems for various aspects. Here we discuss some of them, but we refer to Section 3 for more details.

- In the DP case we exploit the integrable structure of the Degasperis-Procesi equation, proved by Degasperis-Holm-Hone in [48], to overcome some small divisors problems arising in the bifurcation analysis and along the linear reducibility procedure.
  Actually the Birkhoff normal form of the Degasperis-Procesi equation presents some non trivial resonances at order four which are very hard to compute explicitly. The same phenomenon occurs for the water waves equations, as discussed by Craig-Worfolk [47] and Craig [40] (and references therein). In these papers the authors analyse the Birkhoff normal form of water waves with infinite depth observing the presence of non trivial resonances at order four, called *Benjamin-Feir*, and at order five. They also show that the order four resonances do not contribute to the normal form of the Hamiltonian. Thanks to the integrable structure, we are able to prove the same at order four, five and six for the normal form of the Degasperis-Procesi equation (we refer to Sections 5.1, 5.2 and Proposition 5.1.3).

- The nonlinearities which contain the highest order derivatives of the linear part in (1.0.1) and (1.0.5) are *quadratic*, instead of quartic as in (1.0.32) and (1.0.33). This is actually a natural feature of equations which arise from fluid dynamical models, as for instance the water waves equations (1.0.14). However, the latter involve parameters, as the capillarity κ, unlike the cases (1.0.1) and (1.0.5). In such problems one can extract parameters from the equation by a Birkhoff normal form method. These parameters are typically the amplitudes of the approximate solutions from which the iterative scheme bifurcates. The condition for which these amplitudes are in one-to-one relation with the frequency of these approximate solutions is called *twist condition.* In [8], [9] this condition is trivially satisfied. This is not the case for the problems (1.0.1) and (1.0.5). In the first case the twist condition depends on the interactions of many terms, which are parametrized by the real numbers c_i, $i = 1, \dots, 7$ (see (1.0.3), (4.2.7), (4.2.8)). In the second, the expression of the frequency-amplitude map is quite complicated, see (5.3.8), (5.3.9). In both cases, we need to impose *generic* (see Definitions 1.1.2 and 1.2.2) conditions on the tangential set (see (1.1.5), (1.2.8)) to get the *twist* of the frequency-amplitude map.

- In implementing a Birkhoff normal form procedure the linear tangential and normal frequencies of oscillations are corrected by the nonlinearities. The presence of quadratic and quasi-linear nonlinear terms guarantees that there are corrections at the normal frequencies $\omega(j)$ with the same size of the ones at the linear frequencies $\overline{\omega}$. Both in our cases the computations of these corrections are by no means an easy task (see Sections 4.6.5, 4.6.6 and 5.7.5). Typically we deal with these issues when we impose the 2-nd order Melnikov non-resonance conditions. For instance, consider the following expression (recall for instance (1.0.18))

$$\omega\cdot\ell+d_j-d_k=\overline{\omega}\cdot\ell+\omega(j)-\omega(k)+\big((\omega-\overline{\omega})\cdot\ell+(d_j-\omega(j))-(d_k-\omega(k))\big) \qquad (1.0.35)$$

where $\mathrm{i}d_j$ and $\mathrm{i}d_k$ are the eigenvalues of the linearized operator at some approximate solution. In the resonant case $\overline{\omega}\cdot\ell+\omega(j)-\omega(k)=0$ we have to prove a bound from below for $(\omega-\overline{\omega})\cdot\ell+(d_j-\omega(j))-(d_k-\omega(k))$, but these terms interact, actually they have the same size, and they could become arbitrarily small. In order to avoid these stronger resonant cases we impose some conditions which we prove to be generic (see Definitions 1.1.2 and 1.2.2).

- We point out that we have to require stronger generic conditions for the DP case (1.0.5) respect to the KdV case (1.0.1). This is due to the following facts:

  (1) the linear frequencies of the DP equation are rational numbers (see (1.0.10)). In particular the normal part $\overline{\omega}\cdot\ell+\omega(j)-\omega(k)$ of the small divisor (1.0.35), even if it is not zero, can be arbitrarily small.

  (2) Because of the weak dispersion of the DP equation (1.0.5), some identically zero relations as (recall (1.0.35))

$$\begin{cases}\overline{\omega}\cdot\ell+\omega(j)-\omega(k)=0,\\(\omega-\overline{\omega})\cdot\ell+(d_j-\omega(j))-(d_k-\omega(k))=0\end{cases}$$

  could occur for an infinte set of indices (ℓ,j,k).

  As a consequence, in the DP case we need to require further non-degeneracy conditions, which, a priori, could be not satisfied in the generic sense of the KdV case (1.0.1).

- A substantial difference between the DP case (1.0.5) and the works on KdV equations is the linear dispersion relation. As we said above, in this case it is more difficult to prove good bounds for the measure of the set of frequencies satisfying the 2-nd order Melnikov conditions (see (1.0.18)), since the normal frequencies $d_j-d_k\sim j-k$ are not strongly separated, comparing for instance with (1.0.28). Actually, for these reasons, the study of the measure estimates is close to the one adopted for the wave equation (see [14]).
  The analysis of the linearized operator is quite complex and requires a full exploitation of pseudo differential calculus techniques. The presence of the pseudo differential operator J in the symplectic form (1.2.5) complicates this analysis, since its symbol has an *infinite* asymptotic expansion in (decreasing order) homogeneous symbols. Due to the asymptotically linear

dispersion, in the regularization procedure we have to deal with homological equations which are quasi-periodic *transport equations*. To solve these equations we need to impose further conditions on the frequencies similar to the 1-st Melnikov conditions (recall (1.0.17)).

Plan of the Thesis. In the rest of this Chapter we state the main results of this Thesis, namely Theorem 1.1.3 and Theorem 1.2.3.

In Chapter 2 we introduce the notation and the mathematical tools that we used for the proofs of the main results.

In Chapter 3 we first discuss the general strategy adopted in [8], [9], [27], [6] by refering to the KAM problems presented in Chapters 4 and 5. We show also the *stability* argument for the quasi-periodic solutions that we find for the equations (1.0.1) and (1.0.5).
Then we underline the main differences with the works mentioned above and we discuss the main novelties of the results 1.1.3 and 1.2.3.

In Chapter 4 we present the proof of Theorem 1.1.3 about the existence and the stability of quasi-periodic solutions for the equations (1.0.1).

In Chapter 5 we give the proof of Theorem 1.2.3 about the existence and the stability of quasi-periodic solutions for the equations (1.0.5).

In the Appendix are collected some technical lemmata, the proofs of some propositions omitted in Chapters 4 and 5 and some facts about the integrability structure of the DP equation.

1.1 Main results for quasi-linear generalized KdV equations

The equation (1.0.1) can be formulated as a Hamiltonian PDE $u_t = \partial_x \nabla_{L^2} H(u)$, where $\nabla_{L^2} H$ is the $L^2(\mathbb{T})$ gradient of the Hamiltonian (recall (1.0.3))

$$H(u) = \int_{\mathbb{T}} \frac{u_x^2}{2} + f(x, u, u_x)\, dx \tag{1.1.1}$$

on the real phase space

$$H_0^1(\mathbb{T}_x) := \left\{ u \in H^1(\mathbb{T}, \mathbb{R}) : \int_{\mathbb{T}} u(x)\, dx = 0 \right\} \tag{1.1.2}$$

endowed with the non-degenerate symplectic form

$$\Omega(u, v) := \int_{\mathbb{T}} (\partial_x^{-1} u)\, v\, dx, \quad \forall u, v \in H_0^1(\mathbb{T}_x), \tag{1.1.3}$$

where $\partial_x^{-1} u$ is the periodic primitive of u with zero average defined by

$$\partial_x^{-1} e^{\mathrm{i} j x} = \frac{1}{\mathrm{i} j} e^{\mathrm{i}\, j\, x} \quad \text{if } j \neq 0, \qquad \partial_x^{-1} 1 := 0.$$

The Poisson bracket induced by Ω in (1.1.3) between two functions $F, G \colon H_0^1(\mathbb{T}) \to \mathbb{R}$ is

$$\{F(u), G(u)\} := \Omega(X_F, X_G) = \int_{\mathbb{T}} \nabla F(u)\, \partial_x \nabla G(u)\, dx, \tag{1.1.4}$$

where X_F and X_G are the vector fields associated to the Hamiltonians F and G, respectively.

The solutions that we find are localized in Fourier space close to finitely many *tangential sites*

$$S^+ := \{\overline{\jmath}_1, \dots, \overline{\jmath}_\nu\}, \quad S := S^+ \cup (-S^+) = \{\pm j : j \in S^+\}, \quad \overline{\jmath}_i \in \mathbb{N} \setminus \{0\}, \quad \forall i = 1, \dots, \nu \tag{1.1.5}$$

and the linear frequencies of oscillation on the tangential sites are

$$\overline{\omega} := (\overline{\jmath}_1^3, \dots, \overline{\jmath}_\nu^3) \in \mathbb{N}^\nu. \tag{1.1.6}$$

We assume the following hypotesis on the set S:

(S) $\nexists\, j_1, j_2, j_3, j_4 \in S$ such that

$$j_1 + j_2 + j_3 + j_4 \neq 0, \; j_1^3 + j_2^3 + j_3^3 + j_4^3 - (j_1 + j_2 + j_3 + j_4)^3 = 0. \tag{1.1.7}$$

We shall also assume “non-resonance" and “non-degeneracy" conditions on the nonlinearity (1.0.3). In particular we require for the coefficients $c_1, \dots, c_7$ in (1.0.3) to be non-resonant according to the following definition.

Definition 1.1.1. We say that the coefficients $c_1, \dots, c_7$ are *resonant* if the following holds

$$c_3 = c_7 = 2c_1^2 - c_4 = 7c_2^2 - 6c_6 = 0 \tag{1.1.8}$$

and we say that $c_1, \dots, c_7$ are *non-resonant* if (1.1.8) does not hold.

Moreover, we require the following “non-degeneracy" conditions

(C1) fixed $\nu \in \mathbb{N}$, the coefficients $c_1, \dots, c_7$ satisfy

$$(7 - 16\nu)\, c_2^2 \neq 6\,(1 - 2\nu) c_6, \tag{1.1.9}$$

(C2) fixed $\nu \in \mathbb{N}$, if the coefficients $c_1, \dots, c_7$ are non-resonant the following holds

$$\nu\, \frac{3c_6 - 4c_2^2}{9c_4 - 18c_1^2} \notin \{j^2 + k^2 + jk \,:\, j, k \in \mathbb{Z} \setminus \{0\},\, j \neq k\}. \tag{1.1.10}$$

In Section 4 we prove that the assumptions (S), (C1)-(C2) are satisfied for a large choice of the tangential set S. In particular they are “*generic*" according to the following definition.

Definition 1.1.2. (**Genericity**) Fixed $\nu \in \mathbb{N}$ and given a non-trivial, i.e non identically zero, polynomial $P(z)$, with $z \in \mathbb{C}^\nu$, we say that a vector of integers $z_0 \in \mathbb{N}^\nu$ is *generic* if $P(z_0) \neq 0$. We shall say that “*there is a generic choice of the tangential sites S for which some condition holds*" if this condition is satisfied by every vector of integers $(\overline{\jmath}_1, \dots \overline{\jmath}_\nu)$ (see (1.1.5)) that are not zeros of some non-trivial polynomial.

Now we state the main result of Section 4. This is the Theorem 1.3 of [61].

Theorem 1.1.3. *Given $\nu \in \mathbb{N}$, let $f \in C^q$ (with $q := q(\nu)$ large enough) satisfy* (1.0.3). *If $c_1, \dots, c_7$ in* (1.0.3) *are non-resonant (see Definition 1.1.1) and conditions* (C1)-(C2) *hold, then for a generic choice of tangential sites (see Definition 1.1.2 and* (1.1.5)*), in particular satisfying* (S)*, the equation* (1.0.1) *possesses small amplitude quasi-periodic solutions, with diophantine frequency vector $\omega := \omega(\xi) = (\omega_j)_{j \in S^+} \in \mathbb{R}^\nu$, of the form*

$$u(t,x) = 2 \sum_{j \in S^+} \sqrt{j\,\xi_j}\, \cos(\omega_j t + jx) + o(\sqrt{|\xi|}), \qquad \omega_j = j^3 + O(|\xi|) \tag{1.1.11}$$

for a Cantor-like set of small amplitudes $\xi \in \mathbb{R}^\nu_+$ with density 1 at $\xi = 0$. The term $o(\sqrt{|\xi|})$ is small in some H^s-Sobolev norm (see (1.0.22)*), $s < q$. These quasi-periodic solutions are **linearly stable**.*

Let us make some comments.

- We briefly explain why we consider the function f in (1.1.1) of the form (1.0.3). The coefficients $c_1, \dots, c_7$ in (1.0.3) are not considered as external parameters for the equation (1.0.1). Actually we regard (1.0.1) as a family of equations parametrized by these coefficients and we are intersted in studying how the quadratic quasi-linear nonlinearities modulate the frequency-amplitude map (4.2.18) respect to the tangential sites of the solutions. To do that we perform a Birkhoff normal form. The Hamiltonian terms $O(u^4)$ give the frequency-amplitude relation and we want that the Birkhoff procedure is not affected at this order by the x-dependent part of f (see (1.1.1) and (1.0.3)).

- The non-resonance condition stated in Definition 1.1.1 arises by asking that the frequency-amplitude map (4.2.18) is a diffeomorphism. The invertibility of this map is equivalent to require that $\det \mathbb{M} \neq 0$, where the determinant of $\mathbb{M}$ (see (4.2.8) and (4.2.18)) is a polynomial in the variables $(c_1, \dots, c_7, \bar{\jmath}_1, \dots, \bar{\jmath}_\nu)$. In Theorem 1.1.3 we fix non-resonant coefficients $c_1, \dots, c_7$ and we prove in Lemma 4.2.2 that the condition $\det \mathbb{M} \neq 0$ is satisfied for a generic choice of the tangential sites S. We remark that this explicit condition could be verified also adopting a different perspective. We might fix the integers $\bar{\jmath}_1, \dots, \bar{\jmath}_\nu$, hence the tangential set, and choose the real parameters $c_1, \dots, c_7$, thus the equations to study, outside the zeros of some polynomial, since the determinant of $\mathbb{M}$ (see (4.2.21)) can be written as

  $$\det \mathbb{M} = \sum_{n,m} P_n(\bar{\jmath}_i) Q_m(c_i)$$

  where P_n is a homogenous function of degree n, for some $n > 0$, in the variables $\bar{\jmath}_1, \dots, \bar{\jmath}_\nu$ and Q_m is a homogenous function of degree m, for some $m > 0$, in the variables $c_1, \dots, c_7$.

- For the equations (1.0.1) with resonant coefficients (according to the Definition 1.1.1) one could expect that quasi-periodic solutions do not exist at all. We did not investigate in this direction, but we refer to similar cases discussed by Feola-Procesi in [55] for the autonomous NLS. They provide some examples of resonant equations (according to a similar definition) which admit only periodic solutions.

- For the measure estimates of Section 4.7.1, we shall avoid some lower order resonances by imposing the assumptions (H1) and $(\mathtt{H2})_{j,k}$ (see (4.7.33), (4.7.34)). These ones imply that some polynomials are non zero at $(c_1, \dots, c_7, \bar{\jmath}_1, \dots, \bar{\jmath}_\nu)$. If (C1)-(C2) hold and $c_1, \dots, c_7$ are non-resonant then these polynomials are not trivial in the variables $(\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu)$ (see Lemma 4.7.7 and Lemma 4.7.8) and, for a finite number of $j, k \in S^c$, (H1) and $(\mathtt{H2})_{j,k}$ are verified by fixing non-resonant parameters $c_1, \ldots, c_7$ and by choosing a generic set of integers $\{\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu\}$.

- We assume the Hypotesis (S), because we want to perform three steps of Birkhoff normal form and the homogenous part of degree five of the perturbation f_5, which has not an explicit expression (see (1.0.3), (1.0.4)), contributes to the third normal form step. This fact occurs also in [8], indeed the perturbation considered has a non zero homogenous part of degree five and three steps of Birkhoff normal form are nedeed to satisfy the smallness condition (see (4.7.4)) required in the Nash-Moser Theorem (see Theorem 4.7.1); actually this request depends on the quadraticity of the nonlinearity.
  We remark that the assumption (S) can be reformulate as a condition that is satisfied for a generic choice of the tangential sites (according to Definition 1.1.2).

1.2 Main results for Hamiltonian perturbations of the Degasperis-Procesi equation

In 1999 Degasperis and Procesi [49] applied the method of asymptotic integrability to the family of third-order dispersive PDE conservation laws

$$u_t + c_0 u_x + \gamma u_{xxx} - \alpha^2 u_{xxt} = (c_1 u^2 + c_2 u_x^2 + c_3 u u_{xx})_x, \tag{1.2.1}$$

where the constants α^2 and γ / c_0 are squares of lenght scales, and c_0 is the linear wave speed for undisturbed water at rest at spatial infinity (see [50]).
In this family only three equations result to satisfy the asymptotic integrability condition up to the third order, the KdV equation ($\alpha = c_2 = c_3 = 0$), the Camassa-Holm equation ($c_1 = -3c_3/2\alpha^2, c_2 = c_3/2$) and the Degaperis-Procesi equation

$$u_t + c_0 u_x + \gamma u_{xxx} - \alpha^2 u_{xxt} = \left(-\frac{2c_3}{\alpha^2} u^2 + c_2 (u_x^2 + u u_{xx}) \right)_x . \tag{1.2.2}$$

In [48] the authors showed the integrability of equation (1.2.2) by proving the existence of a Lax pair and they provide a recursive method to generate infinitely many constants of motion (see Section 4 in [48]).
The DP equation can be regarded as a model for nonlinear shallow water dynamics and its asymptotic accuracy is the same as for the Camassa-Holm equation and a degree more than the KdV equation. Since its discovery, many results have been produced on the DP equation, for instance, for local and global well-posedness, existence of wave breaking phenomena (peakons, N-peakons solutions).
For proving existence of wave breaking phenomena it is natural to consider the equation (1.2.2) in its dispersionless form, which can be obtained by translations on the phase space and Galilean boosts. Note that the family of equations (1.2.1) is not Galilean invariant.

For our purpose the presence of the dispersive terms $c_0 u_x$ and γu_{xxx} is fundamental, since we study the existence of quasi-periodic waves in the small amplitude regime, namely in a neighborhood of the origin $u = 0$, where the dispersive effects are much stronger than the nonlinear ones. In particular we choose to set in (1.2.2)

$$\alpha^2 = 1, \quad \gamma = 1, \quad c_0 = -4, \quad c_2 = c_3 = -1. \tag{1.2.3}$$

With this choice of the parameter we obtain the equation (1.0.5) with $\mathcal{N}_6 = 0$. Actually we need only to have $c_0 \neq 0, \gamma \neq 0$, the other choices have been done for simplicity of notations.

The equation (1.0.5) can be formulated as a Hamiltonian PDE $u_t = J\nabla_{L^2}H(u)$, where $\nabla_{L^2}H$ is the $L^2(\mathbb{T})$ gradient of the Hamiltonian

$$H(u) = \int_{\mathbb{T}} \frac{u^2}{2} - \frac{u^3}{6} + f(u)\,dx \tag{1.2.4}$$

with $f \in C^\infty(\mathbb{R};\mathbb{R})$, $f(u) = O(u^7)$. The Hamiltonian (1.2.4) is defined on the real phase space

$$H_0^1(\mathbb{T}_x) := \Big\{ u \in H^1(\mathbb{T}, \mathbb{R}) : \int_{\mathbb{T}} u\,dx = 0 \Big\}$$

endowed with the non-degenerate symplectic form

$$\Omega(u,v) := \int_{\mathbb{T}} (J^{-1}u)\,v\,dx, \quad \forall u, v \in H_0^1(\mathbb{T}_x), \qquad J := (1-\partial_{xx})^{-1}(4-\partial_{xx})\partial_x. \tag{1.2.5}$$

The Poisson bracket induced by Ω between two functions $F, G\colon H_0^1(\mathbb{T}) \to \mathbb{R}$ is

$$\{F(u), G(u)\} := \Omega(X_F, X_G) = \int_{\mathbb{T}} \nabla F(u)\, J\nabla G(u)\,dx, \tag{1.2.6}$$

where X_F and X_G are the vector fields associated to the Hamiltonians F and G, respectively. The *dispersion law* of the DP equation is given by

$$\omega(j) := j\,\frac{4+j^2}{1+j^2} = j + \frac{3j}{1+j^2}, \qquad j \in \mathbb{Z}. \tag{1.2.7}$$

We note the Hamiltonian (1.2.4) preserves the momentum, since f does not depend on x.

The solutions that we find are localized in Fourier space close to finitely many *tangential sites*

$$S^+ := \{\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu\}, \quad S := S^+ \cup (-S^+) = \{\pm j : j \in S^+\}, \quad \bar{\jmath}_i \in \mathbb{N}\setminus\{0\}, \quad \forall i = 1, \ldots, \nu \tag{1.2.8}$$

and the linear frequencies of oscillation on the tangential sites are

$$\overline{\omega} := \left(\frac{\bar{\jmath}_1(4+\bar{\jmath}_1^2)}{1+\bar{\jmath}_1^2}, \ldots, \frac{\bar{\jmath}_\nu(4+\bar{\jmath}_\nu^2)}{1+\bar{\jmath}_\nu^2} \right) \in \mathbb{Q}^\nu. \tag{1.2.9}$$

We note that the linear frequencies (1.2.9) are rational numbers such that

$$\frac{\bar{\jmath}_i(4+\bar{\jmath}_i^2)}{1+\bar{\jmath}_i^2} - \bar{\jmath}_i \to 0 \qquad \text{as } \bar{\jmath}_i \to \infty. \tag{1.2.10}$$

For this reason we shall deal with stronger degeneracy relations which we are able to satisfy with further restrictions on the choice of the tangential set respect to the KdV case (see (1.1.6), Definition 1.1.2). In particular we need to formulate a slightly different notion of genericity respect to the one given in Definition 1.1.2.

Definition 1.2.1. Let $1 > \mathtt{c} > 0$ be a constant and fix $\nu \in \mathbb{N}$, $\nu \geq 1$. We define the set $\mathcal{V}(\mathtt{c}, \nu)$ as the set of the ν-ples $(j_1, \dots, j_\nu) \in (\mathbb{N} \setminus \{0\})^\nu$ such that

$$\max_{i=1,\dots,\nu} j_i \leq \frac{1}{\mathtt{c}} \tag{1.2.11}$$

or

$$\max_{i=1,\dots,\nu} j_i > \frac{1}{\mathtt{c}} \qquad \text{and} \qquad \left| \frac{j_i}{\max_{i=1,\dots,\nu}\{j_i\}} - 1 \right| \leq \mathtt{c}. \tag{1.2.12}$$

The set $\mathcal{V}(\nu, \mathtt{c}) \subset \mathbb{N}^\nu$ is the union of a ball centered at the origin of radius $1/\mathtt{c}$ with a cone whose vertex is the origin and with axis $\{(j_1, \dots, j_\nu) \in \mathbb{N}^\nu : j_1 = \dots = j_\nu\}$.

Definition 1.2.2. (**Genericity**) Let $1 > \mathtt{c} > 0$ be a constant and fix $\nu \in \mathbb{N}$, $\nu \geq 1$. We say that a set $\{j_1, \dots, j_\nu\}$, $j_i \in \mathbb{N} \setminus \{0\}$, is generic in $\mathcal{V}(\mathtt{c}) := \mathcal{V}(\nu, \mathtt{c})$ if $(j_1, \dots, j_\nu) \in \mathcal{V}(\mathtt{c})$ and it is generic according to Definition 1.1.2. We shall say that "*a choice of the tangential sites S is generic*" if S^+ defined in (1.2.8) is generic in $\mathcal{V}(\mathtt{c})$ for some constant $\mathtt{c}$.

We shall assume the following non-degeneracy conditions which are proved to be generic in $\mathcal{V}(c)$ for some fixed constant $\mathtt{c}$ (see Appendix B.1):

(H0) For $\ell \in \mathbb{Z}^\nu$

$$\sum_{i=1}^{\nu} \bar{\jmath}_i \ell_i = 0 \quad \Longleftrightarrow \ell = 0 \qquad \text{for } |\ell| = 7, 8, \tag{1.2.13}$$

(H1) for $\ell \in \mathbb{Z}^\nu$

$$\sum_{i=1}^{\nu} \frac{\bar{\jmath}_i}{1 + \bar{\jmath}_i^2} \ell_i \neq 0 \qquad \text{for } |\ell| = 3, 4, 5, \tag{1.2.14}$$

(H2) for some constant $\mathtt{c}_*$ independent of the set S^+

$$|\det \mathbb{A}| \geq \mathtt{c}_* \max_{i=1,\dots,\nu} \bar{\jmath}_i, \qquad |1 - \mathbb{A}^{-T} \vec{v} \cdot \overline{\omega}| \geq \mathtt{c}_* \tag{1.2.15}$$

where $\mathbb{A}$ is defined in (5.3.9), $\vec{v}$ in (5.9.46),

(H3) For all $\ell \in \mathbb{Z}^\nu$, $j, k \in S^c$

$$\Big(\mathrm{I} - \mathbb{A}^{-T} \vec{v}\, \overline{\omega}^T\Big)\ell \neq \mathbb{A}^{-T}(w_j - w_k) \tag{1.2.16}$$

where (recall (1.2.7))

$$w_j := \frac{2\omega(j)}{3} \left(\frac{(1 + \bar{\jmath}_k^2)(7 + 5\bar{\jmath}_k^2 + \bar{\jmath}_k^4 + 3j^2)}{(3 + j^2)^2 + (6 + j^2)\bar{\jmath}_k^2 + \bar{\jmath}_k^4} \right)_{k=1}^{\nu} \in \mathbb{R}^\nu.$$

The main result of Section 5 is the following. This is part of a joint work with Roberto Feola and Michela Procesi.

Theorem 1.2.3. *Given $\nu \in \mathbb{N}$, let $f \in C^\infty$ satisfy* (1.0.7). *There exist a constant $\mathtt{c} > 0$ such that for a generic choice of S^+ in $\mathcal{V}(\mathtt{c}, \nu)$ (see Definitions 1.1.2, 1.2.2 and* (1.2.8)*) the equation* (1.0.5) *possesses small amplitude quasi-periodic solutions, with diophantine frequency vector* $\omega := \omega(\xi) = (\omega_j)_{j\in S^+} \in \mathbb{R}^\nu$*, of the form*

$$u(t,x) = 2\sum_{j\in S^+} \sqrt{\xi_j}\,\cos(\omega_j t + jx) + o(\sqrt{|\xi|}), \quad \omega_j = j\frac{(4+j^2)}{1+j^2} + O(|\xi|) \tag{1.2.17}$$

for a Cantor-like set of small amplitudes $\xi \in \mathbb{R}^\nu_+$ *with density* 1 *at* $\xi = 0$*. The term* $o(\sqrt{|\xi|})$ *is small in some* H^s*-Sobolev norm (see* (1.0.22)*) for* s *large. These quasi-periodic solutions are* ***linearly stable****.*

Let us make some comments.

- As commented under Theorem 1.1.3, we require that the analysis of the twist condition is not affected by the perturbation f (see (1.2.4)). Moreover we want to fully exploit the integrability of the Degasperis-Procesi equation (see [48]) to prove that there are only trivial resonances involved in the steps of weak Birkhoff normal form (see Proposition 5.1.3).
- We could deal also with perturbations of the form

$$f(x,u) = f_{\le 4}(u) + f_{\ge 5}(x,u)$$

  where $f_{\le 4}$ denotes the homogeneous part of degree less than four of f near the origin and $f_{\ge 5}$ the part of degree greater or equal than five. For simplicity we avoid some technical issues by considering a Hamiltonian density f independent of x. In particular we use this fact to exploit the conservation of momentum in implementing the normal form steps in Sections 5.7.4 and 5.7.5.
- We consider $f \in C^\infty$ in order to work in the usual framework of the C^∞ pseudo differential operators. In particular with this choice the coefficients of the linearized operator (5.6.31) are C^∞ in (φ, x) since the vector field $J\nabla H(u)$ is C^∞ (see (1.2.4), (1.2.5)) and the approximate solutions at which we linearize are trigonometric polynomial.
- The generic assumption (H0) (see (1.2.13)) is needed to perform the 5-th and 6-th step of (weak) Birkhoff normal form. We could avoid this condition by computing other two suitable constants of motion and by reasoning as in the first four steps (see Section 5.2).
  The generic assumption (H1) (see (1.2.14)) is used in Section 5.7.4.
  The assumption (H2) is imposed to get the invertibility of the frequency-amplitude map (5.3.8). Note that we cannot simply require that the determinant of the twist matrix $\mathbb{A}$ (see (4.2.8)) is not zero, since it is an analytic function of the tangential variables $\bar{\jmath}_1, \dots, \bar{\jmath}_\nu$ that could accumulate to zero as $\max \bar{\jmath}_i$ increases. In the set (1.2.12) the asymptotic behaviour of this function is under control, this is the reason for defining the notion of genericity (1.2.2).
  The hypothesis (H3) is assumed in order to prove Lemma 5.9.9 for the measure estimates.

CHAPTER 2

Functional setting

In this Chapter we introduce some notations, definitions and technical tools which will be used along the proofs of Chapters 4 and 5.

Notations. We list some of the notations used along the Thesis.

- We denote by I the identity operator on some space X. When there is possible confusion about the domain of definition we denote it by I_X.

- We write $a \leq_s b$ to mean that $a \leq C(s) b$ with $C(s)$ a constant depending on some index s. In Chapters 4 and 5 we use this notation for the index s of a Sobolev space H^s. In Chapter 5 we use this notation also for constants depending on parameters $\mathtt{b}$ or ρ (see Sections 5.7, 5.8) and for the constants depending on the index α in (2.2.10) and (2.2.11) .
  When we deal with inequalities which involve constants depending on the dimension ν, the index s_0, the number τ of the diophantine inequalities or other quantities which come from the nonlinear terms, we simply call all these constants C. When it is necessary to distinguish different constants we give different names.

- The notation $R(v^{k-q} z^q)$ indicates a homogeneous polynomial of degree k in the variables (v, z) of the form

$$R(v^{k-q} z^q) = M[\underbrace{v, \ldots, v}_{(k-q) \ times}, \underbrace{z, \ldots, z}_{q \ times}], \quad M = k - \text{linear form}.$$

  In particular, for homogenous Hamiltonian $H = H(v, z)$ we denote by $H^{(n)}$ the homogenous part of degree n of H, by $H^{(n, \leq m)}$ the terms $R(v^{n-q} z^q)$ for $q \leq m$.

- Sometimes we will denote the operator $\omega \cdot \partial_\varphi$ with $\mathcal{D}_\omega$. We denote by $[r]$ the integer part of a real number r.

- We denote by U and $\vec{1}$ respectively the matrix with all entries equal to 1 and the vector with all components equal to 1.

2.1 Sobolev functions and Lipschitz norms.

Sobolev functions. We consider a function $u(\varphi,x)\in L^2(\mathbb{T}^\nu\times\mathbb{T},\mathbb{C})$ as a φ-dependent family of functions $u(\varphi,\cdot)\in L^2(\mathbb{T}_x,\mathbb{C})$ with the Fourier series expansion

$$u(\varphi,x)=\sum_{j\in\mathbb{Z}}u_j(\varphi)\,e^{\mathrm{i}jx}=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}u_{\ell j}\,e^{\mathrm{i}(\ell\cdot\varphi+jx)}.$$

We denote the Fourier coefficients of u as u_j and $u_{\ell j}$, or $u_{\ell,j}$, with respect to the variables x and (φ,x). Sometimes, in order to distinguish the Fourier taransform only in one of the variables x and φ we will use the notation $\hat{\cdot}$.
We shall consider *real valued* functions. For the Fourier coefficients this means that

$$\overline{u}_j(\varphi)=u_{-j}(\varphi),\quad \overline{u}_{\ell j}=u_{-\ell,-j}. \tag{2.1.1}$$

We use the simplified notation L^2 to denote $L^2(\mathbb{T}^\nu\times\mathbb{T})$ and $L^2_x:=L^2(\mathbb{T}_x)$. We define the Sobolev space

$$H^s(\mathbb{T}^{\nu+1};\mathbb{C}):=\left\{u(\varphi,x)\in L^2:\|u\|_s^2:=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}|u_{\ell j}|^2\langle\ell,j\rangle^{2s}<\infty\right\} \tag{2.1.2}$$

where $\langle\ell,j\rangle:=\max\{1,|\ell|,|j|\}$, $|\ell|:=\sum_{i=1}^\nu|\ell_i|$ and it is equivalent to $1+|\ell|+|j|$. The norm defined in (2.1.2) is equivalent to the norm

$$\|\cdot\|_{H^s(\mathbb{T}^\nu,L^2_x)}+\|\cdot\|_{L^2(\mathbb{T}^\nu,H^s_x)}. \tag{2.1.3}$$

If $s_0\geq(\nu+1)/2$ then for $s\geq s_0$ the spaces $H^s(\mathbb{T}^{\nu+1})\hookrightarrow L^\infty(\mathbb{T}^{\nu+1})$ and they have the algebra structure. Moreover they satisfies the interpolation inequalities (see for instance [23], [69], Appendix G [62], Appendix [18])

$$\|u\,v\|_s\leq C(s_0)\,\|u\|_s\|v\|_{s_0}+C(s)\|u\|_{s_0}\|v\|_s,\quad\forall u,v\in H^s(\mathbb{T}^{\nu+1}). \tag{2.1.4}$$

Linear operators. Let X,Y be two Banach spaces and let us denote by $\mathcal{L}(X,Y)$ the set of linear operators from X to Y.
Let $A\colon\mathbb{T}^\nu\to\mathcal{L}(L^2(\mathbb{T}_x))$, $\varphi\mapsto A(\varphi)$, be a φ-dependent family of linear operators acting on $L^2(\mathbb{T}_x)$. We consider A as an operator acting on the functions $u(\varphi,x)$ in the following way

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

This action is represented in Fourier coordinates as

$$Au(\varphi,x)=\sum_{j,j'\in\mathbb{Z}}A_j^{j'}(\varphi)\,u_{j'}(\varphi)\,e^{\mathrm{i}jx}=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}A_j^{j'}(\ell-\ell')\,u_{\ell'j'}\,e^{\mathrm{i}(\ell\cdot\varphi+jx)}. \tag{2.1.5}$$

We define for $m=1,\dots,\nu$ the operator $\partial_{\varphi_m}A(\varphi)$ as

$$(\partial_{\varphi_m}A(\varphi))u(\varphi,x)=\sum_{\ell\in\mathbb{Z}^\nu,j\in\mathbb{Z}}\sum_{\ell'\in\mathbb{Z}^\nu,j'\in\mathbb{Z}}\mathrm{i}(\ell-\ell')\,A_j^{j'}(\ell-\ell')\,u_{\ell'j'}\,e^{\mathrm{i}(\ell\cdot\varphi+jx)}. \tag{2.1.6}$$

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

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

Lipschitz norm. Fix $\nu \geq 1$ and let $\mathcal{O}$ be a compact subset of $\mathbb{R}^\nu$. We call parameters the elements of this set.
For a function $u\colon \mathcal{O} \to E$, where $(E, \|\cdot\|_E)$ is a Banach space, we define the sup-norm and the lip-seminorm of u as

$$\begin{aligned} \|u\|_E^{\sup} &:= \|u\|_E^{\sup,\mathcal{O}} := \sup_{\omega\in\mathcal{O}} \|u(\omega)\|_E, \\ \|u\|_E^{lip} &:= \|u\|_E^{lip,\mathcal{O}} := \sup_{\substack{\omega_1,\omega_2\in\mathcal{O},\\ \omega_1\neq\omega_2}} \frac{\|u(\omega_1)-u(\omega_2)\|_E}{|\omega_1-\omega_2|}. \end{aligned} \tag{2.1.7}$$

Fix $\gamma > 0$. In Chapter 4 we will use the following Lipschitz norms

$$\begin{aligned} \|u\|_s^{Lip(\gamma)} &:= \|u\|_s^{sup} + \gamma \|u\|_s^{lip}, \quad \forall s \geq (\nu+2)/2, \\ |m|^{Lip(\gamma)} &:= |m|^{sup} + \gamma |m|^{lip}, \quad m \in \mathbb{R}. \end{aligned} \tag{2.1.8}$$

In Chapter 5 we will use the following Lipschitz norms

$$\|u\|_s^{\gamma,\mathcal{O}} := \|u\|_s^{sup,\mathcal{O}} + \gamma \|u\|_{s-1}^{lip,\mathcal{O}}, \quad \forall s \geq [\nu/2]+3 \tag{2.1.9}$$

$$|m|^{\gamma,\mathcal{O}} := |m|^{sup,\mathcal{O}} + \gamma |m|^{lip,\mathcal{O}}, \quad m \in \mathbb{R}. \tag{2.1.10}$$

For convenience we use two slightly different notations for the Lip norm in the KdV and DP cases. In the DP case we need to be more careful about the domains of parameters, thus we prefer to recall the set $\mathcal{O}$ in the notation used for the Lip norms (2.1.9) and (2.1.10). Note that in (2.1.9) we require a weaker lip-seminorm respect to (2.1.8).

We refer to the Appendix A for technical lemmata on the tameness properties of the Lipschitz and Sobolev norms.

2.1.1 Symplectic Sobolev scales

In Chapters 4 and 5 we work on a scale of symplectic Sobolev spaces. We consider $H_0^1(\mathbb{T}_x;\mathbb{R})$ (see (1.1.2)) as phase space and the symplectic structure is given by

$$\Omega(u,v) := \langle J^{-1}u, v\rangle_{L^2(\mathbb{T}_x)} = \sum_{j\in\mathbb{Z}\setminus\{0\}} \frac{1}{\mathrm{i}\lambda(j)} u_j\, \overline{v}_j, \qquad u,v \in H_0^1(\mathbb{T}_x), \tag{2.1.11}$$

where $J := \partial_x$, $\lambda(j) := j$ in the KdV case and $J := (1-\partial_{xx})^{-1}(4-\partial_{xx})\partial_x$, $\lambda(j) := (1+j^2)/(j(4+j^2))$ in the DP case. Both these operators are unbounded and the vector field $X_H := J\nabla H$ is defined on the scale

$$H^s(\mathbb{T}_x;\mathbb{R}) := \left\{ u(x) \in L^2 : \|u\|_s^2 := \sum_{j\in\mathbb{Z}\setminus\{0\}} |u_j|^2 \langle j\rangle^{2s} < \infty \right\}$$

and in Fourier coordinates reads as

$$[X_H(u)]_j = \mathrm{i}\lambda(j)(\partial_{\overline{u}_j} H)(u), \quad j \in \mathbb{Z}\setminus\{0\}.$$

Consider two real functions $F, G\colon H_0^1(\mathbb{T}_x) \to \mathbb{R}$ then the Poisson bracket associated to (2.1.11) is

$$\{F,G\}(u) := \langle \nabla F(u), J\nabla G(u)\rangle_{L^2(\mathbb{T}_x)} = -\sum_{j\in\mathbb{Z}} \mathrm{i}\lambda(j)\partial_{u_{-j}} F(u)\, \partial_{u_j} G(u). \tag{2.1.12}$$

Definition 2.1.1. We say that

(1) a map is symplectic if it preserves the 2-form Ω in (2.1.11);

(2) an operator $(Ah)(\varphi,x) := A(\varphi)h(\varphi,x)$ is symplectic if each $A(\varphi), \varphi \in \mathbb{T}^\nu$, is a symplectic map of the phase space (or of a symplectic subspace like $H_S^\perp$);

(3) the operator $\omega \cdot \partial_\varphi - JG(\varphi)$ is Hamiltonian if each $G(\varphi), \varphi \in \mathbb{T}^\nu$, is symmetric.

Conservation of momentum. In the KdV case, Chapter 4, the *momentum* is the quadratic Hamiltonian

$$M_{KdV}(u) = (1/2)\int_{\mathbb{T}} u^2\,dx.$$

In the DP case, Chapter 5, the momentum is

$$M_{DP}(u) = \int_{\mathbb{T}} J^{-1}u_x\,u\,dx.$$

In both cases, the vector field generated by M_{KdV}, M_{DP} is

$$\partial_x \nabla M_{KdV}(u) = J\nabla M_{DP}(u) = u_x.$$

It is easy to see that a homogeneous Hamiltonian of degree m

$$H(u) = \sum_{j_1,\dots,j_m \in \mathbb{Z}\setminus\{0\}} H_{j_1,\dots,j_m} u_{j_1}\dots u_{j_m} \tag{2.1.13}$$

preserves the momentum, namely Poisson commutes with M_{KdV} or M_{DP} with the respective Poisson structures (see (2.1.12)), if it is supported on the set $\{(j_1,\dots,j_m) \in \mathbb{Z}^m \setminus \{\mathbf{0}\} : j_1 + \dots + j_m = 0\}$.

2.2 Pseudo differential calculus

In Chapter 5 we exploit some pseudo differential calculus techniques (see for instance Section 5.7). Here we recall some basic definitions and properties of pseudo differential calculus. For more details we refer to Section 2 of [27] or [63], [77], [93], [92].

The pseudo differential operators on the torus $\mathrm{Op}(a(x,j))$ may be seen as a particular case of pseudo differential operators $\mathrm{Op}(a(x,\xi))$ on $\mathbb{R}^n$ (we refer for instance to [63]).
In the following we give the definitions of both these class of operators, but, along the Chapter 5 we will always use the continuos notation $a(x,\xi)$ for the symbols, also if we think $\mathrm{Op}(a)$ as an operator acting on 2π-periodic functions $u(x)$.

Given a function $\beta \in \mathbb{N}$ we denote by $\Delta_j^\beta := \Delta_j \circ \dots \circ \Delta_j$ the composition of β-discrete derivatives. Now we give the definition of a pseudo differential operator on the torus.

Definition 2.2.1. Let $u = \sum_{j\in\mathbb{Z}} u_j\, e^{\mathrm{i}jx}$. A linear operator A defined by

$$Au(x) := \sum_{j\in\mathbb{Z}} a(x,j)\,u_j\,e^{\mathrm{i}jx} \tag{2.2.1}$$

is called pseudo differential of order $\le m$ if its symbol $a(x,j)$ is 2π-periodic and C^∞ smooth in x, and it satisfies

$$|\partial_x^\alpha \Delta_j^\beta a(x,j)| \le C_{\alpha,\beta}\langle j\rangle^{m-\beta}, \ \ \forall\, \alpha,\beta \in \mathbb{N}. \tag{2.2.2}$$

We give the definition of a pseudo differential operator on $\mathbb{T}$ derived from the corresponding one on $\mathbb{R}$.

Definition 2.2.2. A linear operator A is called pseudo differential of order $\le m$ if its symbol $a(x,j)$ is the restriction to $\mathbb{R}\times\mathbb{Z}$ of a complex valued function $a(x,\xi)$ which is C^∞ smooth on $\mathbb{R}\times\mathbb{R}$, 2π-periodic in x and satisfies

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \le C_{\alpha,\beta}\langle\xi\rangle^{m-\beta}, \ \ \forall\, \alpha,\beta\in\mathbb{N}. \tag{2.2.3}$$

The Definitions 2.2.1 and 2.2.2 are equivalent and we refer to [92] for the proof of this fact. We denote by

$$A = \mathrm{Op}(a)$$

the pseudo operator with symbol $a := a(x,\xi)$. 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 define the class S^m as the set of symbols which satisfy (2.2.3).

Composition of pseudo differential operators. One of the fundamental properties of pseudo differential operators is the following: given two pseudo differential operators $\mathrm{Op}(a)\in OPS^m$ and $\mathrm{Op}(b)\in OPS^{m'}$, for some $m,m'\in\mathbb{R}$, the composition $\mathrm{Op}(a)\circ\mathrm{Op}(b)$ is a pseudo differential operator of order $m+m'$. In particular

$$\mathrm{Op}(a)\circ\mathrm{Op}(b) = \mathrm{Op}(a\# b), \tag{2.2.4}$$

where the symbol of the composition is given by

$$(a\# b)(x,\xi) = \sum_{j\in\mathbb{Z}} a(x,\xi+j)\hat{b}_j(\xi)e^{\mathrm{i}jx} = \sum_{k,j\in\mathbb{Z}} \hat{a}_{k-j}(\xi+j)\hat{b}_j(\xi)e^{\mathrm{i}kx}. \tag{2.2.5}$$

Here the $\hat{\ }$ denotes the Fourier transform of the symbols $a(x,\xi)$ and $b(x,\xi)$ in the variable x. The symbol $a\# b$ has the following asymptotic expansion: for any $N\ge 1$ one can write

$$\begin{aligned}
(a\# b)(x,\xi) &= \sum_{n=0}^{N-1}\frac{1}{n!\mathrm{i}^n}\partial_\xi^n a(x,\xi)\partial_x^n b(x,\xi) + r_N(x,\xi), \qquad r_N\in S^{m+m'-N},\\
r_N(x,\xi) &= \frac{1}{(N-1)!\mathrm{i}^N}\int_0^1 (1-\tau)^N \sum_{j\in\mathbb{Z}}(\partial_\xi^N a)(x,\xi+\tau j)\widehat{\partial_x^N b}(j,\xi)e^{\mathrm{i}jx}d\tau.
\end{aligned} \tag{2.2.6}$$

Definition 2.2.3. Let $N\in\mathbb{N}$, $0\le k\le N$, $a\in S^m$ and $b\in S^{m'}$, we define

$$a\#_k b := \frac{1}{k!\mathrm{i}^k}(\partial_\xi^k a)(\partial_x^k b), \qquad a\#_{<N} b := \sum_{k=0}^{N-1} a\#_k b, \qquad a\#_{\ge N} b := r_N. \tag{2.2.7}$$

Adjoint operator. Let $A := \mathrm{Op}(a)\in OPS^m$. Then its L^2-adjoint A^* is a pseudo differential operator such that

$$A^* = \mathrm{Op}(a^*), \qquad a^*(x,\xi) = \overline{\sum_{j\in\mathbb{Z}}\hat{a}_j(\xi-j)e^{\mathrm{i}jx}}. \tag{2.2.8}$$

Parameter family of pseudo differential operators. We shall deal also with pseudo differential operators depending on parameters $\varphi \in \mathbb{T}^\nu$:

$$(Au)(\varphi,x) = \sum_{j\in\mathbb{Z}} a(\varphi,x,j) u_j e^{\mathrm{i}jx}, \quad a(\varphi,x,j) \in S^m.$$

The symbol $a(\varphi,x,\xi)$ is C^∞ smooth also in the variable φ. We still denote

$$A := A(\varphi) = \mathrm{Op}(a(\varphi,\cdot)) = \mathrm{Op}(a).$$

For the symbols of the composition operator with $\mathrm{Op}(b(\varphi,x,\xi))$ and the L^2-adjoint we have the following formulas

$$\begin{aligned}(a\#b)(\varphi,x,\xi) &= \sum_{j\in\mathbb{Z}} a(\varphi,x,\xi+j)\hat{b}(\varphi,j,\xi)\, e^{\mathrm{i}jx} = \sum_{\substack{j,j'\in\mathbb{Z},\\ \ell,\ell'\in\mathbb{Z}^\nu}} \hat{a}(\ell-\ell',j'-j,\xi+j)\hat{b}(\ell',j,\xi)\, e^{\mathrm{i}(\ell\cdot\varphi+jx)},\\ a^*(\varphi,x,\xi) &= \overline{\sum_{j\in\mathbb{Z}} \hat{a}(\varphi,j,\xi-j)\, e^{\mathrm{i}jx}} = \overline{\sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \hat{a}(\ell,j,\xi-j)\, e^{\mathrm{i}(\ell\cdot\varphi+jx)}}.\end{aligned} \tag{2.2.9}$$

Following [27] and [77], on such operators we define the following norm.

Definition 2.2.4. Let $a(\varphi,x,\xi)\in S^m$ and set $A = \mathrm{Op}(a) \in OPS^m$,

$$|A|_{m,s,\alpha} := \max_{0\le\beta\le\alpha} \sup_{\xi\in\mathbb{R}} \|\partial_\xi^\beta a(\cdot,\cdot,\xi)\|_s \langle\xi\rangle^{-m+\beta}. \tag{2.2.10}$$

We will use also the notation

$$|a|_{m,s,\alpha} := |A|_{m,s,\alpha}.$$

Note that the norm $|\cdot|_{m,s,\alpha}$ in non-decreasing in s and α. Moreover given a symbol $a(\varphi,x)$ independent of ξ, the norm of the associated multiplication operator $\mathrm{Op}(a)$ is just the H^s norm of the function a. If on the contrary the symbol $a(\xi)$ depends only on ξ, then the norm of the corresponding Fourier multipliers $\mathrm{Op}(a(\xi))$ is just controlled by a constant.

Weighted pseudo differential norm. Let $A = \mathrm{Op}(a(\lambda,\varphi,x,\xi)) \in OPS^m$ be a family of pseudo differential operators with symbol $a(\lambda,\varphi,x,\xi)\in S^m$ depending in a Lipschitz way on some parameter $\lambda \in \mathcal{O} \subset \mathbb{R}^\nu$. We introduce the following weighted norm

$$|A|^{\gamma,\mathcal{O}}_{m,s,\alpha} := \sup_{\lambda\in\mathcal{O}} |A|_{m,s,\alpha} + \gamma \sup_{\lambda_1,\lambda_2\in\mathcal{O}} \frac{|\mathrm{Op}\big(a(\lambda_1,\varphi,x,\xi) - a(\lambda_2,\varphi,x,\xi)\big)|_{m,s-1,\alpha}}{|\lambda_1-\lambda_2|}. \tag{2.2.11}$$

Note that the norm (2.2.11) satisfies

$$\begin{aligned}&\forall\, s\le s',\ \alpha\le\alpha' \ \Rightarrow |\cdot|^{\gamma,\mathcal{O}}_{m,s,\alpha} \le |\cdot|^{\gamma,\mathcal{O}}_{m,s',\alpha}, \quad |\cdot|^{\gamma,\mathcal{O}}_{m,s,\alpha} \le |\cdot|^{\gamma,\mathcal{O}}_{m,s,\alpha'}\\ &m\le m' \quad \Rightarrow |\cdot|^{\gamma,\mathcal{O}}_{m',s,\alpha} \le |\cdot|^{\gamma,\mathcal{O}}_{m',s,\alpha}.\end{aligned} \tag{2.2.12}$$

In the following lemma we collect properties of pseudo differential operators which will be used in the sequel.
We remark that along the Nash-Moser iteration we shall control the Lipschitz variation respect to the torus embedding $i := i(\varphi)$ of the terms of the linearized operator at i. Hence we consider pseudo differential operators which depend on this variable.

Lemma 2.2.5. *Fix $m, m', m'' \in \mathbb{R}$. Let i be a torus embedding. Consider symbols*

$$a(i, \lambda, \varphi, x, \xi) \in S^m, \ b(i, \lambda, \varphi, x, \xi) \in S^{m'}, \ c(\lambda, \varphi, x, \xi) \in S^{m''}, \ d(\lambda, \varphi, x, \xi) \in S^0$$

which depend on $\lambda \in \mathcal{O}$ and $i \in H^s$ in a Lipschitz way. Set

$$\begin{aligned} A &:= \mathrm{Op}(a(\lambda, \varphi, x, \xi)), \quad B := \mathrm{Op}(b(\lambda, \varphi, x, \xi)), \\ C &:= \mathrm{Op}(c(\lambda, \varphi, x, \xi)), \quad D := \mathrm{Op}(d(\lambda, \varphi, x, \xi)). \end{aligned}$$

Then one has
(i) for any $\alpha \in \mathbb{N}$, $s \geq s_0$,

$$|A \circ B|^{\gamma,\mathcal{O}}_{m+m',s,\alpha} \leq_{m,\alpha} C(s)|A|^{\gamma,\mathcal{O}}_{m,s,\alpha}|B|^{\gamma,\mathcal{O}}_{m',s_0+\alpha+|m|,\alpha} + C(s_0)|A|^{\gamma,\mathcal{O}}_{m,s_0,\alpha}|B|^{\gamma,\mathcal{O}}_{m',s+\alpha+|m|,\alpha}. \tag{2.2.13}$$

One has also that, for any $N \geq 1$, the operator $R_N := \mathrm{Op}(r_N)$ with r_N defined in (2.2.6) *satisfies*

$$\begin{aligned} |R_N|^{\gamma,\mathcal{O}}_{m+m'-N,s,\alpha} \leq_{m,N,\alpha} \frac{1}{N!}\Big(&C(s)|A|^{\gamma,\mathcal{O}}_{m,s,\alpha+N}|B|^{\gamma,\mathcal{O}}_{m',s_0+2N+\alpha+|m|,\alpha} + \\ &C(s_0)|A|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+N}|B|^{\gamma,\mathcal{O}}_{m',s+2N+\alpha+|m|,\alpha}\Big); \end{aligned} \tag{2.2.14}$$

$$\begin{aligned} |\partial_i R_N[\hat{\imath}]|^{\gamma,\mathcal{O}}_{m+m'-N,s,\alpha} \leq_{m,N,\alpha} \frac{1}{N!}\Big(&C(s)|\partial_i A[\hat{\imath}]|^{\gamma,\mathcal{O}}_{m,s,\alpha+N}|B|^{\gamma,\mathcal{O}}_{m',s_0+2N+\alpha+|m|,\alpha} + \\ &C(s_0)|\partial_i A[\hat{\imath}]|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+N}|B|^{\gamma,\mathcal{O}}_{m',s+2N+\alpha+|m|,\alpha}\Big) \\ &+ \frac{1}{N!}\Big(C(s)|A|^{\gamma,\mathcal{O}}_{m,s,\alpha+N}|\partial_i B[\hat{\imath}]|^{\gamma,\mathcal{O}}_{m',s_0+2N+\alpha+|m|,\alpha} \\ &+ C(s_0)|A|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+N}|\partial_i B[\hat{\imath}]|^{\gamma,\mathcal{O}}_{m',s+2N+\alpha+|m|,\alpha}\Big); \end{aligned} \tag{2.2.15}$$

(ii) the adjoint operator $C^* := \mathrm{Op}(c^*(\lambda, \varphi, x, \xi))$ in (2.2.8) *satisfies*

$$|C^*|^{\gamma,\mathcal{O}}_{m'',s,0} \leq_m |C|^{\gamma,\mathcal{O}}_{m'',s+s_0+|m''|,0}; \tag{2.2.16}$$

(iii) consider the map $\Phi := \mathrm{I} + D$, then there are constants $C(s_0, \alpha), C(s, \alpha) \geq 1$ such that if

$$C(s_0, \alpha)|D|^{\gamma,\mathcal{O}}_{0,s_0+\alpha,\alpha} \leq \frac{1}{2}, \tag{2.2.17}$$

then, for all λ, the map Φ is invertible and $\Phi^{-1} \in OPS^0$ and for any $s \geq s_0$ one has

$$|\Phi^{-1} - \mathrm{I}|^{\gamma,\mathcal{O}}_{0,s,\alpha} \leq C(s, \alpha)|D|^{\gamma,\mathcal{O}}_{0,s+\alpha,\alpha}. \tag{2.2.18}$$

Proof. Item (i) and (iii) are proved respectively in Lemmata 2.13 and 2.17 of [27]. The estimates (2.2.13) and (2.2.14) are proved in Lemma 2.16 of [27]. The bound (2.2.15) is obtained following the proof of Lemma 2.16 of [27] and exploiting the Leibniz rule. □

Remark 2.2.6. When the domain of parameters $\mathcal{O}$ depends on the variable i then we are interested in estimating the variation $\Delta_{12}A := A(i_1) - A(i_2)$ on $\mathcal{O}(i_1) \cap \mathcal{O}(i_2)$ instead of the derivative ∂_i. The bound (2.2.15) holds also for Δ_{12} by replacing $i_1 - i_2 \rightsquigarrow \hat{\imath}$.

Commutators. By formula (2.2.6) the commutator between two pseudo differential operators $A := \mathrm{Op}(a(\lambda,\varphi,x,\xi))$, $B := \mathrm{Op}(b(\lambda,\varphi,x,\xi))$ with $a \in S^m$ and $b \in S^{m'}$, is a pseudo differential operator such that

$$[A,B] := \mathrm{Op}(a\star b), \quad a\star b(\lambda,\varphi,x,\xi) := \big(a\#b - b\#a\big)(\lambda,\varphi,x,\xi). \tag{2.2.19}$$

The symbols $a\star b$ (called the Moyal parenthesis of a and b) admits the expansion

$$a\star b = -\mathrm{i}\{a,b\} + \mathtt{r}_2(a,b), \quad \{a,b\} = \partial_\xi a\partial_x b - \partial_x a\partial_\xi b \in S^{m+m'-1}, \tag{2.2.20}$$

where

$$\mathtt{r}_2(a,b) = \Big[(a\#b) - \frac{1}{\mathrm{i}}\partial_\xi a\partial_x b\Big] - \Big[(b\#a) - \frac{1}{\mathrm{i}}\partial_\xi b\partial_x a\Big] \in S^{m+m'-2}. \tag{2.2.21}$$

Following Definition 2.2.3 we also set

$$a\star_k b := a\#_k b - b\#_k a, \qquad a\star_{<N} b := \sum_{k=0}^{N-1} a\star_k b, \qquad a\star_{\geq N} b := a\#_{\geq N} b - b\#_{\geq N} a. \tag{2.2.22}$$

As a consequence, using bounds (2.2.13) and (2.2.14), one has

$$\begin{aligned}|[A,B]|^{\gamma,\mathcal{O}}_{m+m'-1,s,\alpha} \leq_{m,m'} &\, C(s)|A|^{\gamma,\mathcal{O}}_{m,s+2+|m'|+\alpha,\alpha+1}|B|^{\gamma,\mathcal{O}}_{m',s_0+2+\alpha+|m|,\alpha+1}\\ &+ C(s_0)|A|^{\gamma,\mathcal{O}}_{m,s_0+2+|m'|+\alpha+1,\alpha+1}|B|^{\gamma,\mathcal{O}}_{m',s+2+\alpha+|m|,\alpha+1}.\end{aligned} \tag{2.2.23}$$

The last inequality is proved in Lemma 2.15 of [27].

2.3 Setting for KAM reducibility

In Chapters 4 and 5 we implement two different KAM reducibility schemes, see Section s 4.6.8 and 5.8. The first one refers to a KAM iteration performed in [7], [8] where the non diagonal bounded remainders resulted from the regularization procedure of Section 4.6 have finite *decay norm* , see Definition 2.3.1 below.
In the DP case, Chapter 5, the remainders of the regularization procedure of Section 5.7 do not have this property. Following [27] and [28] we perform the diagonalization procedure of Section 5.8 with the class of *modulo-tame operators*. Our case involves tame-operator of negative order as in the work of Baldi-Berti-Haus-Montalto [28] and we thank the authors for useful discussions about this key point of the proof.

2.3.1 Matrices with off-diagonal decay

We recall the definition of the s-decay norm (introduced in [20]) of an infinite dimensional matrix. This norm is used in [7] for the KAM reducibility scheme of the linearized operators and we refer to Section 2 of [7] for further details.

Definition 2.3.1. (**Decay norm**) The s-decay norm of an infinite dimensional matrix $A := (A^{i_2}_{i_1})_{i_1,i_2\in\mathbb{Z}^b}$, $b \geq 1$ is

$$|A|^2_s := \sum_{i\in\mathbb{Z}^b} \langle i\rangle^{2s}\Big(\sup_{i_1-i_2=i}|A^{i_2}_{i_1}|\Big)^2. \tag{2.3.1}$$

For parameter dependent matrices $A := A(\omega), \omega \in \mathcal{O} \subseteq \mathbb{R}^\nu$, the definitions (2.1.7) and (2.1.8) become

$$\begin{aligned} & |A|_s^{\sup} := \sup_{\omega \in \mathcal{O}} |A(\omega)|_s, \ \ |A|_s^{lip} := \sup_{\substack{\omega_1,\omega_2 \in \mathcal{O}, \\ \omega_1 \neq \omega_2}} \frac{|A(\omega_1) - A(\omega_2)|_s}{|\omega_1 - \omega_2|}, \\ & |A|_s^{Lip(\gamma)} := |A|_s^{\sup} + \gamma |A|_s^{lip}. \end{aligned} \tag{2.3.2}$$

Such a norm is modelled on the behavior of matrices representing the multiplication operator by a function. Actually, given a function $p \in H^s(\mathbb{T}^b)$, the multiplication operator $h \to p\,h$ is represented by the Töplitz matrix $T_i^j = p_{i-j}$ and $|T|_s = \|p\|_s$. If $p = p(\omega)$ is a Lipschitz family of functions, then

$$|T|_s^{Lip(\gamma)} = \|p\|_s^{Lip(\gamma)}.$$

The s-norm satisfies classical algebra and interpolation inequalities proved in [20].

Lemma 2.3.2. *Let $A = A(\omega), B = B(\omega)$ be matrices depending in a Lipschitz way on the parameter $\omega \in \mathcal{O} \subseteq \mathbb{R}^\nu$. Then for all $s \geq s_0 > b/2$ there are $C(s) \geq C(s_0) \geq 1$ such that*

$$\begin{aligned} & |A\,B|_s^{Lip(\gamma)} \leq C(s) |A|_s^{Lip(\gamma)} |B|_s^{Lip(\gamma)}, \\ & |A\,B|_s^{Lip(\gamma)} \leq C(s) |A|_s^{Lip(\gamma)} |B|_{s_0}^{Lip(\gamma)} + C(s_0) |A|_{s_0}^{Lip(\gamma)} |B|_s^{Lip(\gamma)}. \end{aligned}$$

The s-decay norm controls the Sobolev norm, namely

$$\|Ah\|_s^{Lip(\gamma)} \leq C(s) \left(|A|_{s_0}^{Lip(\gamma)} \|h\|_s^{Lip(\gamma)} + |A|_s^{Lip(\gamma)} \|h\|_{s_0}^{Lip(\gamma)} \right). \tag{2.3.3}$$

An important sub-algebra is formed by the *Töplitz in time matrices* defined by

$$A_{(\ell_1,j_1)}^{(\ell_2,j_2)} := A_{j_1}^{j_2}(\ell_1 - \ell_2), \tag{2.3.4}$$

whose decay norm (2.3.1) is

$$|A|_s^2 = \sum_{j \in \mathbb{Z}, \ell \in \mathbb{Z}^\nu} \left(\sup_{j_1 - j_2 = j} |A_{j_1}^{j_2}(\ell)| \right)^2 \langle \ell, j \rangle^{2\,s}. \tag{2.3.5}$$

These matrices are identified with the φ-dependent family of operators

$$A(\varphi) := (A_{j_1}^{j_2}(\varphi))_{j_1,j_2 \in \mathbb{Z}}, \quad A_{j_1}^{j_2}(\varphi) := \sum_{\ell \in \mathbb{Z}^\nu} A_{j_1}^{j_2}(\ell)\, e^{\mathrm{i}\,\ell \cdot \varphi}$$

which act on functions of the x-variables as

$$A(\varphi) : h(x) = \sum_{j \in \mathbb{Z}} h_j\, e^{\mathrm{i}\,j\,x} \mapsto A(\varphi)h(x) = \sum_{j_1,j_2 \in \mathbb{Z}} A_{j_1}^{j_2}(\varphi) h_{j_2}\, e^{\mathrm{i}\,j_1\,x}.$$

Lemma 2.3.3. *(Lemma 2.4 in [7]) Let A be a Töpliz matrix as in (2.3.4) and $s_0 \geq (\nu+2)/2$. Then*

$$|A(\varphi)|_s \leq C(s_0) |A|_{s+s_0} \quad \forall \varphi \in \mathbb{T}^\nu.$$

2.3.2 Tame operators

In [27] the authors deal with linear operators depending on a parameter λ and they need to control, in a quantitative way, a certain number of derivatives in λ in order to develop a degenerate KAM theory. For this reason, in Section 2.2 of [27], they introduced the notion of $\mathcal{D}^{k_0}$-tame and $\mathcal{D}^{k_0}$-modulo-tame operators, where $\mathcal{D}^{k_0}$ denotes the regularity on the parameter λ.
In the DP case, Chapter 5, we need only to control the Lipschitz variation on the parameters, hence we use slightly different (weaker) notions of the operators involved in our scheme. The proofs of the results presented here derive from the results proved in Section 2.2 of [27] with some small technical variations.

Definition 2.3.4 (**σ-Tame operators).** For $\sigma \geq 0$ a linear operator A is *σ-tame* if, for any $s_0 \leq s \leq S_{max}$ with possibly $S_{max} = +\infty$, the following estimate holds:

$$\|Au\|_s \leq \mathfrak{M}_A(\sigma, s)\|u\|_{s_0+\sigma} + \mathfrak{M}_A(\sigma, s_0)\|u\|_{s+\sigma} \qquad u \in H^s, \tag{2.3.6}$$

where the functions $s \mapsto \mathfrak{M}_A(\sigma, s)$ are non-decreasing in s. We call $\mathfrak{M}_A(\sigma, s)$ the *tame constant* of the operator A. When the index σ is not relevant we write $\mathfrak{M}_A(\sigma, s) = \mathfrak{M}_A(s)$.

Definition 2.3.5 (**Lip-σ-Tame operators).** Let $A = A(\omega)$ be a linear operator defined for $\omega \in \mathcal{O} \subset \mathbb{R}^\nu$. Let us define

$$\Delta_{\omega,\omega'} A := \frac{A(\omega) - A(\omega')}{\omega - \omega'}, \quad \omega, \omega' \in \mathcal{O}. \tag{2.3.7}$$

Then A is σ-tame with $\sigma \geq 0$ if, for any $s_0 \leq s \leq S_{max}$, with possibly $S_{max} = +\infty$, the following estimate holds

$$\sup_{\omega\in\mathcal{O}} \|Au\|_s, \gamma \sup_{\omega\neq\omega'} \|(\Delta_{\omega,\omega'} A)\|_{s-1} \leq_s \mathfrak{M}^\gamma_A(\sigma, s)\|u\|_{s_0+\sigma} + \mathfrak{M}^\gamma_A(\sigma, s)\|u\|_{s+\sigma}, \quad u \in H^s, \tag{2.3.8}$$

where the functions $s \mapsto \mathfrak{M}^\gamma_A(\sigma, s)$ are non-decreasing in s. We call $\mathfrak{M}^\gamma_A(\sigma, s)$ the *Lip-tame constant* of the operator A. When the index σ is not relevant we write $\mathfrak{M}^\gamma_A(\sigma, s) = \mathfrak{M}^\gamma_A(s)$.

Lemma 2.3.6. *Let A and B be respectively Lip-σ_A-tame and Lip-σ_b-tame operators with tame constants respectively $\mathfrak{M}^\gamma_A(s)$ and $\mathfrak{M}^\gamma_B(s)$. Then the composition $A \circ B$ is a Lip-$(\sigma_A + \sigma_B)$-operator with*

$$\mathfrak{M}^\gamma_{A\circ B}(s) \leq \mathfrak{M}^\gamma_A(s)\mathfrak{M}^\gamma_B(s_0 + \sigma_A) + \mathfrak{M}^\gamma_A(s_0)\mathfrak{M}^\gamma_B(s + \sigma_A). \tag{2.3.9}$$

The same holds for σ-tame operators.

Lemma 2.3.7. *Let A be a Lip-σ-tame operator. Let $u(\omega)$, $\omega \in \mathcal{O} \subset \mathbb{R}^\nu$ be a ω-parameter family of Sobolev functions H^s, for $s \geq s_0$. Then*

$$\|Au\|^{\gamma,\mathcal{O}}_s \leq_s \mathfrak{M}^\gamma_A(\sigma, s)\|u\|^{\gamma,\mathcal{O}}_{s_0} + \mathfrak{M}^\gamma_A(\sigma, s_0)\|u\|^{\gamma,\mathcal{O}}_s. \tag{2.3.10}$$

Proof. By definition (2.3.8) we have $\mathfrak{M}_A(\sigma, s) \leq \mathfrak{M}^\gamma_A(\sigma, s)$ and $\|u\|_s \leq \|u\|^{\gamma,\mathcal{O}}_s$. Then the thesis follows by the triangle inequalities

$$\frac{1}{|\omega - \omega'|}\|A(\omega)u(\omega) - A(\omega')u(\omega')\|_s \leq \|(\Delta_{\omega,\omega'} A)u(\omega)\|_s + \|A(\omega')\Delta_{\omega,\omega'} u\|_s.$$

□

Lemma 2.3.8. *Let $A = \mathrm{Op}(a(\varphi, x, D)) \in OPS^{\sigma}$ with $\sigma \geq 0$ be a family of pseudo differential operators which are Lipschitz in a parameter $\omega \in \mathcal{O} \subset \mathbb{R}^{\nu}$. If $|A|^{\gamma,\mathcal{O}}_{\sigma,s,0} < +\infty$ then A is a σ-tame operator with*

$$\mathfrak{M}^{\gamma}_{A}(\sigma, s) \leq C(s)|A|^{\gamma,\mathcal{O}}_{\sigma,s,0}. \tag{2.3.11}$$

Proof. We refer to the proof of Lemma 2.21 of [27]. □

2.3.3 Modulo-tame operators and majorant norms

The modulo-tame operators are introduced in Section 2.2 of [27]. As we said before, our definitions are slightly different since we are interested only in the Lipschitz variation of the operators respect to the parameters of the problem. The main difference with the work [27] is that in the KAM reducibility procedure we involve modulo-tame operators which regularize in space. This fact holds also in [28].

Definition 2.3.9. Let $u \in H^s(\mathbb{T}^{\nu+1})$, we define the majorant function

$$\underline{u}(\varphi, x) := \sum_{\ell\in\mathbb{Z}^{\nu}, j\in\mathbb{Z}} |u_{\ell j}| e^{\mathrm{i}(\ell\cdot\varphi + jx)}.$$

Note that $\|u\|_s = \|\underline{u}\|_s$.

Definition 2.3.10. Let A be a bounded linear operator from $H^s(\mathbb{T}^{\nu+1})$ into $H^s(\mathbb{T}^{\nu+1})$ and recall its matrix representation (2.1.5). We define the madjorant matrix $\underline{A}$ as the matrix with entries

$$(\underline{A})^{j'}_{j}(\ell) := |(A)^{j'}_{j}(\ell)| \qquad j, j' \in \mathbb{Z},\ \ell \in \mathbb{Z}^{\nu}.$$

We consider the majorant operator norms

$$\|\underline{M}\|_{\mathcal{L}(H^s)} := \sup_{\|u\|_s \leq 1} \|\underline{M}u\|_s. \tag{2.3.12}$$

We have a partial ordering relation, i.e. if

$$M \preceq N \Leftrightarrow |M^{j'}_{j}(\ell)| \leq |N^{j'}_{j}(\ell)| \ \ \forall j, j', \ell \ \Rightarrow \|M\|_{\mathcal{L}(H^s)} \leq \|N\|_{\mathcal{L}(H^s)}, \quad \|Mu\|_s \leq \|\underline{M}\,\underline{u}\|_s \leq \|\underline{N}\,\underline{u}\|_s.$$

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

$$\Pi_K M := \begin{cases} M^{j'}_{j}(\ell) & \text{if } |\ell| \leq K, \\ 0 & \text{otherwise} \end{cases} \qquad \Pi^{\perp}_{K} := \mathrm{I} - \Pi_K.$$

Finally we define for $\mathtt{b}_0 \in \mathbb{N}$

$$(\langle\partial_{\varphi}\rangle^{\mathtt{b}_0} M)^{j'}_{j}(\ell) = \langle\ell\rangle^{\mathtt{b}_0} M^{j'}_{j}(\ell).$$

If $A = A(\omega)$ is an operator depending on a parameter ω, we control the Lipschitz variation, see formula 2.3.7. In the sequel let $1 > \gamma > \gamma_* > 0$ be fixed constants.

Definition 2.3.11 (Lip-σ-modulo tame)**.** A linear operator $A := A(\omega)$, $\omega \in \mathcal{O} \subset \mathbb{R}^\nu$, is Lip-$\sigma$-modulo-tame w.r.t. an increasing sequence $\{\mathfrak{M}^\sharp_A, \gamma_*(s)\}_{s=s_0}^{S_{max}}$ if the majorant operators $\underline{A}, \underline{\Delta_{\omega,\omega'} A}$ are Lip-σ-tame w.r.t. these constants, i.e. they satisfy the following weighted tame estimates: For $\sigma \geq 0$, for all $s \geq s_0$ and for any $u \in H^s$,

$$\sup_{\omega\in\mathcal{O}} \|\underline{A}u\|_s, \sup_{\omega\neq\omega'\in\mathcal{O}} \gamma_* \|\underline{\Delta_{\omega,\omega'} A}u\|_s \leq \mathfrak{M}^{\sharp,\gamma_*}_A(\sigma, s_0)\|u\|_{s+\sigma} + \mathfrak{M}^{\sharp,\gamma_*}_A(\sigma, s)\|u\|_{s_0+\sigma} \tag{2.3.13}$$

where the functions $s \mapsto \mathfrak{M}^{\sharp,\gamma_*}_A(s) \geq 0$ are non-decreasing in s. The constant $\mathfrak{M}^{\sharp,\gamma_*}_A(\sigma, s)$ is called the MODULO-TAME CONSTANT of the operator A.

Definition 2.3.12. We say that A is Lip--1-modulo tame if $\langle D_x\rangle^{1/2} A \langle D_x\rangle^{1/2}$ is Lip-0-modulo tame. We denote

$$\begin{aligned} \mathfrak{M}^{\sharp,\gamma_*}_A(s) &:= \mathfrak{M}^{\sharp,\gamma_*}_{D^{1/2}AD^{1/2}}(0, s) \\ \mathfrak{M}^{\sharp,\gamma_*}_A(s, a) &:= \mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^a A}(s), \quad a \geq 0. \end{aligned} \tag{2.3.14}$$

In the following we shall systematically use -1 modulo-tame operators. Here we list and prove some properties.

Lemma 2.3.13. *Let $s_0 \geq [(\nu + 2)/2 + 1] \in \mathbb{N}$. Let A be a -1-modulo tame operator and $\mathtt{b}_0 \in \mathbb{N}$. Then*

$$\mathfrak{M}^{\sharp,\gamma_*}_A(s) \leq \max_{m=1,\dots,\nu} \mathfrak{M}^{\gamma_*}_{\partial^{s_0}_{\varphi_m}[A,\partial_x]}(-1, s), \tag{2.3.15}$$

$$\mathfrak{M}^{\sharp,\gamma_*}_A(s, \mathtt{b}_0) \leq \max_{m=1,\dots,\nu} \mathfrak{M}^{\gamma_*}_{\partial^{s_0+\mathtt{b}_0}_{\varphi_m}[A,\partial_x]}(-1, s). \tag{2.3.16}$$

Proof. We have

$$\begin{aligned} &\|\langle D_x\rangle^{1/2}\underline{A}\langle D_x\rangle^{1/2}u\|_s^2 \leq \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \langle \ell, j\rangle^{2s}\Big(\sum_{\ell'\in\mathbb{Z}^\nu, j'\in\mathbb{Z}} \langle j'\rangle^{1/2}|(A)^{j'}_j(\ell-\ell')\langle j\rangle^{1/2}|u_{\ell'j'}|\Big)^2 \\ &\leq \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \langle \ell, j\rangle^{2s}\Big(\sum_{\ell'\in\mathbb{Z}^\nu, j'\in\mathbb{Z}} \frac{\langle \ell-\ell'\rangle^{s_0}|j-j'|}{\langle \ell-\ell'\rangle^{s_0}|j-j'|}\langle j'\rangle^{1/2}|(A)^{j'}_j(\ell-\ell')\langle j\rangle^{1/2}|u_{\ell'j'}|\Big)^2 \\ &\leq C\sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \langle \ell, j\rangle^{2s} \sum_{\ell'\in\mathbb{Z}^\nu, j'\in\mathbb{Z}} \langle j\rangle\langle j'\rangle|j-j'|^2\langle \ell-\ell'\rangle^{2s_0}|(A)^{j'}_j(\ell-\ell')|^2|u_{\ell'j'}|^2 \\ &= C\sum_{\ell'\in\mathbb{Z}^\nu, j'\in\mathbb{Z}} |u_{\ell'j'}|^2 \sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \langle j\rangle\langle j'\rangle|j-j'|^2\langle \ell-\ell'\rangle^{2s_0}|(A)^{j'}_j(\ell-\ell')|^2 \end{aligned}$$

since

$$C := \sum_{\ell'\in\mathbb{Z}^\nu, j'\in\mathbb{Z}} \frac{1}{\langle \ell-\ell'\rangle^{2s_0}|j-j'|^2} < +\infty.$$

By the fact that for any $1 \leq m \leq \nu$

$$\begin{aligned} &\sum_{\ell\in\mathbb{Z}^\nu, j\in\mathbb{Z}} \langle \ell, j\rangle^{2s}\langle j\rangle\langle j'\rangle|j-j'|^2\langle \ell_m-\ell'_m\rangle^{2s_0}|(A)^{j'}_j(\ell-\ell')|^2 \\ &\leq 2(\mathfrak{M}^{\gamma_*}_{\partial^{s_0}_{\varphi_m}[A,\partial_x]}(-1, s))^2\langle \ell', j'\rangle^{2s_0} + 2(\mathfrak{M}^{\gamma_*}_{\partial^{s_0}_{\varphi_m}[A,\partial_x]}(-1, s_0))^2\langle \ell', j'\rangle^{2s} \end{aligned} \tag{2.3.17}$$

and

$$\langle \ell-\ell'\rangle \leq \max_{m=1,\dots,\nu} \langle \ell_m-\ell'_m\rangle \tag{2.3.18}$$

we obtain

$$\|\langle D_x\rangle^{1/2}\underline{A}\langle D_x\rangle^{1/2}u\|_s^2 \leq 2\max_{m=1,\dots,\nu}(\mathfrak{M}^{\gamma_*}_{\partial^{s_0}_{\varphi_m}[A,\partial_x]}(-1,s_0))^2\|u\|_s^2 + 2\max_{m=1,\dots,\nu}(\mathfrak{M}^{\gamma_*}_{\partial^{s_0}_{\varphi_m}[A,\partial_x]}(-1,s))^2\|u\|_{s_0}^2.$$

Following the same computations above we conclude the same bound for $\|\langle D_x\rangle^{1/2}\underline{\Delta_{\omega,\omega'}A}\langle D_x\rangle^{1/2}u\|_s^2$. By the fact that $\gamma_* < 1$ we deduce (2.3.15). The proof of (2.3.16) is analogous. $\square$

Lemma 2.3.14. *(i) If $A \preceq B$ and $\Delta_{\omega,\omega'}A \preceq \Delta_{\omega,\omega'}B$ for all $\omega \neq \omega' \in \mathcal{O}$, we may choose the tame constants of A so that*

$$\mathfrak{M}^{\sharp,\gamma_*}_A(s) \leq \mathfrak{M}^{\sharp,\gamma_*}_B(s)\,.$$

(ii) Let A be a -1 modulo-tame operator with modulo-tame constant $\mathfrak{M}^{\sharp,\gamma_*}_A(s)$. Then the operator $\langle D_x\rangle^{1/2}A\langle D_x\rangle^{1/2}$ is majorant bounded $H^s \to H^s$

$$\|\langle D_x\rangle^{1/2}\underline{A}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^s)} \leq 2\mathfrak{M}^{\sharp,\gamma_*}_A(s)\,, \qquad |A^j_j(0)|^{\gamma_*} \leq \frac{\mathfrak{M}^{\sharp,\gamma_*}_A(s_0)}{\langle j\rangle}.$$

(iii) Suppose that $\langle\partial_\varphi\rangle^{\mathtt{b}_0}A$, $\mathtt{b}_0 \geq 0$, is -1 modulo-tame. Then the operator $\Pi^\perp_N A$ is -1 modulo-tame with tame constant

$$\mathfrak{M}^{\sharp,\gamma_*}_{\Pi^\perp_N A}(s) \leq \min(N^{-\mathtt{b}_0}\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}A}(s), \mathfrak{M}^{\sharp,\gamma_*}_A(s))\,. \tag{2.3.19}$$

(iv) Let B be a -1 modulo-tame operator with modulo-tame constant $\mathfrak{M}^{\sharp,\gamma_*}_B(s)$. Then $A+B$ is -1 modulo-tame with modulo-tame constant

$$\mathfrak{M}^{\sharp,\gamma_*}_{A+B}(s) \leq \mathfrak{M}^{\sharp,\gamma_*}_A(s) + \mathfrak{M}^{\sharp,\gamma_*}_B(s)\,. \tag{2.3.20}$$

The composed operator $A\circ B$ is -1 modulo-tame with modulo-tame constant

$$\mathfrak{M}^{\sharp,\gamma_*}_{AB}(s) \leq C(s)\big(\mathfrak{M}^{\sharp,\gamma_*}_A(s)\mathfrak{M}^{\sharp,\gamma_*}_B(s_0) + \mathfrak{M}^{\sharp,\gamma_*}_A(s_0)\mathfrak{M}^{\sharp,\gamma_*}_B(s)\big)\,. \tag{2.3.21}$$

Assume in addition that $\langle\partial_\varphi\rangle^{\mathtt{b}_0}A$, $\langle\partial_\varphi\rangle^{\mathtt{b}_0}B$ are -1 modulo-tame with constant respectively $\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}A}(s)$ and $\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}B}(s)$, then $\langle\partial_\varphi\rangle^{\mathtt{b}_0}(AB)$ is -1 modulo-tame with modulo-tame constant satisfsying

$$\begin{aligned}\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}(AB)}(s) \leq C(s,\mathtt{b}_0)\Big(&\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}A}(s)\mathfrak{M}^{\sharp,\gamma_*}_B(s_0) + \mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}A}(s_0)\mathfrak{M}^{\sharp,\gamma_*}_B(s)\\ &+ \mathfrak{M}^{\sharp,\gamma_*}_A(s)\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}B}(s_0) + \mathfrak{M}^{\sharp,\gamma_*}_A(s_0)\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}B}(s)\Big)\,.\end{aligned} \tag{2.3.22}$$

Finally, for any $k \geq 1$ we have, setting $L = \mathrm{ad}^k(A)B$, $\mathrm{ad}(A)B := AB - BA$:

$$\begin{aligned}&\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}L}(s) \leq C(s,\mathtt{b}_0)^k\Big[(\mathfrak{M}^{\sharp,\gamma_*}_A(s_0))^k\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}B}(s)\\ &+ k(\mathfrak{M}^{\sharp,\gamma_*}_A(s_0))^{k-1}\Big(\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}A}(s)\mathfrak{M}^{\sharp,\gamma_*}_B(s_0) + \mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}A}(s_0)\mathfrak{M}^{\sharp,\gamma_*}_B(s)\Big)\\ &+ k(k-1)(\mathfrak{M}^{\sharp,\gamma_*}_A(s_0))^{k-2}\mathfrak{M}^{\sharp,\gamma_*}_A(s)\mathfrak{M}^{\sharp,\gamma_*}_{\langle\partial_\varphi\rangle^{\mathtt{b}_0}A}(s_0)\mathfrak{M}^{\sharp,\gamma_*}_B(s_0)\Big]\,.\end{aligned} \tag{2.3.23}$$

The same bound holds if we set $L = A^kB$.

(v) Let $\Phi := \mathrm{I} + A$ *and assume, for some* $\mathtt{b}_0 \geq 0$*, that* $\langle \partial_\varphi \rangle^{\mathtt{b}_0} A$ *is Lip–1-modulo tame and the smallness condition*

$$8C(S_{max}, \mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_A(s_0) < 1\,, \qquad C(S_{max}, \mathtt{b}_0) = \max_{s_0 \leq s \leq S_{max}} C(s, \mathtt{b}_0) \tag{2.3.24}$$

holds. Then the operator Φ *is invertible,* $\breve{A} := \Phi^{-1} - \mathrm{Id}$ *is* -1 *modulo-tame together with* $\langle \partial_\varphi \rangle^{\mathtt{b}_0} A$ *with modulo-tame constants*

$$\mathfrak{M}^{\sharp,\gamma_*}_{\breve{A}}(s) \leq 2\mathfrak{M}^{\sharp,\gamma_*}_A(s)\,, \tag{2.3.25}$$

$$\mathfrak{M}^{\sharp,\gamma_*}_{\langle \partial_\varphi \rangle^{\mathtt{b}_0} \breve{A}}(s) \leq 2\mathfrak{M}^{\sharp,\gamma_*}_{\langle \partial_\varphi \rangle^{\mathtt{b}_0} A}(s) + 8C(S_{max}, \mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_{\langle \partial_\varphi \rangle^{\mathtt{b}_0} A}(s_0)\, \mathfrak{M}^{\sharp,\gamma_*}_A(s)\,. \tag{2.3.26}$$

Proof. In the following we shall sistematically use the fact that if B is an operator with matrix coefficients ≥ 1, then $A \preceq \underline{A \circ B} = \underline{A} \circ \underline{B} = \underline{A} \circ B$. Note that $\langle D_x \rangle^{1/2}$ is a diagonal operator with positive eignevalues.

(i) Assume that $A \preceq B$ i.e. $|A_j^{j'}(\ell)| \leq |B_j^{j'}(\ell)|$ for all j, j', ℓ. Then

$$\|\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} u\|_s \leq \|\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \underline{u}\|_s \leq \|\langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2} \underline{u}\|_s.$$

The same reasoning holds for $\langle D_x \rangle^{1/2} \underline{\Delta_{\omega,\omega'} A} \langle D_x \rangle^{1/2}$, so that the result follows.

(ii) The first bound is just a reformulation of the definition, indeed

$$\sup_{\|u\|_s \leq 1} \|\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} u\|_s \leq \sup_{\|u\|_s \leq 1} (\mathfrak{M}^{\sharp,\gamma_*}_A(s_0)\|u\|_s + \mathfrak{M}^{\sharp,\gamma_*}_A(s)\|u\|_{s_0}) \leq 2\mathfrak{M}^{\sharp,\gamma_*}_A(s).$$

In order to prove the second bound we notice that setting

$$B_j^{j'}(\ell) = \begin{cases} \langle j \rangle A_j^j(0) & \ell = 0 \text{ and } j = j', \\ 0 & \text{otherwise}, \end{cases}$$

we have $B \preceq \langle D_x \rangle^{1/2} A \langle D_x \rangle^{1/2}$, same for $\Delta_{\omega,\omega'} B$. Fix any j_0 and consider the unit vector $u^{(j_0)}$ in $H^{s_0}(\mathbb{T}^{\nu+1})$ defined by $u_{j,\ell} = 0$ if $(j,\ell) \neq (j_0, 0)$ and $u_{j_0,0} = \langle j_0 \rangle^{-s_0}$. We have

$$\langle j_0 \rangle |A_{j_0}^{j_0}(0)| = \|\underline{B} u^{(j_0)}\|_{s_0} \leq \|\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \underline{u}^{(j_0)}\|_{s_0} \leq \mathfrak{M}^{\sharp,\gamma_*}_A(s_0).$$

The same holds for $\gamma_* \langle j_0 \rangle |\Delta_{\omega,\omega'} A_{j_0}^{j_0}(0)|$.

(iii) We remark that $|A_j^{j'}(\ell)| \leq N^{-\mathtt{b}_0} \langle \ell \rangle^{\mathtt{b}_0} |A_j^{j'}(\ell)|$ if $|\ell| \geq N$ and the same holds for $|\Delta_{\omega,\omega'} A_j^{j'}(\ell)|$. Therefore we have

$$\Pi_N^\perp A \preceq N^{-\mathtt{b}_0} \langle \partial_\varphi \rangle^{\mathtt{b}_0} \Pi_N^\perp A \preceq N^{-\mathtt{b}_0} \langle \partial_\varphi \rangle^{\mathtt{b}_0} A$$

and the result follows, see also Lemma 2.27 of [27].

(iv) For the first bound we just remark that

$$\langle D_x \rangle^{1/2} (\underline{A + B}) \langle D_x \rangle^{1/2} \preceq \langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} + \langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2},$$

and the same for the Lipschitz variation, so that the result follows from Lemma 2.25 of [27]. Regarding the second we note that

$$\begin{aligned} &\langle D_x \rangle^{1/2} \underline{A \circ B} \langle D_x \rangle^{1/2} \preceq \langle D_x \rangle^{1/2} \underline{A} \circ \underline{B} \langle D_x \rangle^{1/2} \preceq \langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \circ \langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2}\,, \\ &\langle D_x \rangle^{1/2} \underline{\Delta_{\omega,\omega'} A \circ B} \langle D_x \rangle^{1/2} \preceq \langle D_x \rangle^{1/2} \underline{\Delta_{\omega,\omega'} A} \langle D_x \rangle^{1/2} \circ \langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2} \\ &+ \langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \circ \langle D_x \rangle^{1/2} \underline{\Delta_{\omega,\omega'} B} \langle D_x \rangle^{1/2}, \end{aligned}$$

so that the result follows from Lemma 2.25 of [27]. For the third bound we note that

$$\langle \ell \rangle^{\mathtt{b}_0} \sum_{j_1, \ell_1+\ell_2=\ell} A_j^{j_1}(\ell_1) B_{j_1}^{j'}(\ell_2) \leq C(\mathtt{b}_0) \sum_{j_1, \ell_1+\ell_2=\ell} (\langle \ell_1 \rangle^{\mathtt{b}_0} + \langle \ell_2 \rangle^{\mathtt{b}_0}) A_j^{j_1}(\ell_1) B_{j_1}^{j'}(\ell_2)$$

and the same holds for $\Delta_{\omega,\omega'} A \circ B$ and $A \circ \Delta_{\omega,\omega'} B$. Hence

$$\begin{aligned}\langle D_x \rangle^{1/2} \underline{\langle \partial_\varphi \rangle^{\mathtt{b}_0} (A \circ B)} \langle D_x \rangle^{1/2} \preceq C(\mathtt{b}_0) \Big(& \langle D_x \rangle^{1/2} \underline{\langle \partial_\varphi \rangle^{\mathtt{b}_0} A} \langle D_x \rangle^{1/2} \circ \langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2} \\ & + \langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \circ \langle D_x \rangle^{1/2} \underline{\langle \partial_\varphi \rangle^{\mathtt{b}_0} B} \langle D_x \rangle^{1/2} \Big) ,\end{aligned}$$

same for the Lipshitz variations. The result follows from the estimate on the composition.

In order to prove (2.3.23) we note that

$$\langle D_x \rangle^{1/2} \mathrm{ad}^k(A) B \langle D_x \rangle^{1/2} \preceq \underline{\mathrm{ad}}^k \left(\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \right) \langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2} ,$$

where $\underline{\mathrm{ad}}(A)B := AB + BA$ and similarly

$$\begin{aligned}\langle \partial_\varphi^{\mathtt{b}_0} \rangle \langle D_x \rangle^{1/2} \mathrm{ad}^k(A) B \langle D_x \rangle^{1/2} \preceq & \ \underline{\mathrm{ad}}^k \left(\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \right) \langle D_x \rangle^{1/2} \langle \partial_\varphi^{\mathtt{b}_0} \rangle \underline{B} \langle D_x \rangle^{1/2} \\ & + \sum_{k_1+k_2=k-1} \underline{\mathrm{ad}}^{k_1} \left(\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \right) \underline{\mathrm{ad}} (\langle D_x \rangle^{1/2} \langle \partial_\varphi^{\mathtt{b}_0} \rangle \underline{A} \langle D_x \rangle^{1/2}) \\ & \qquad \underline{\mathrm{ad}}^{k_2} \left(\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \right) \langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2} .\end{aligned}$$

Completely analogous bounds can be proved for the Lipschitz variations, by recalling that

$$\Delta_{\omega,\omega'} \mathrm{ad}(A) B = \mathrm{ad}(\Delta_{\omega,\omega'} A) B + \mathrm{ad}(A) \Delta_{\omega,\omega'} B.$$

The result follows, by induction, from the estimate on the composition. The estimate (2.3.23) when $C = A^k \circ B$ follows in the same way using

$$\begin{aligned}\langle \partial_\varphi^{\mathtt{b}_0} \rangle \langle D_x \rangle^{1/2} (A)^k \circ B \langle D_x \rangle^{1/2} \preceq & \ (\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2})^k \circ \langle D_x \rangle^{1/2} \langle \partial_\varphi^{\mathtt{b}_0} \rangle \underline{B} \langle D_x \rangle^{1/2} \\ & + \sum_{k_1+k_2=k-1} \left(\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \right)^{k_1} \langle D_x \rangle^{1/2} \langle \partial_\varphi^{\mathtt{b}_0} \rangle \underline{A} \langle D_x \rangle^{1/2}) \left(\langle D_x \rangle^{1/2} \underline{A} \langle D_x \rangle^{1/2} \right)^{k_2} \\ & \qquad \langle D_x \rangle^{1/2} \underline{B} \langle D_x \rangle^{1/2}.\end{aligned}$$

Item *(v)* follows trivially from (2.3.23) with $C = A^k \circ B$. □

CHAPTER 3

Strategy of the proofs and main novelties

The general strategy used for the proofs of the Theorems 1.1.3 and 1.2.3 follows the one of the papers [8], [9] and for the analysis of the linearized operators we use some tools developed in [27], [6]. In turn these papers are in the general framework developed in [21]. In this Chapter we first describe the common strategy for the proof of Theorems 1.1.3 and 1.2.3. Then we discuss more in details the features of these problems.

3.1 General strategy

We describe the key ingredients of the strategy adopted in Chapters 4 and 5. We show also how we deduce the linear stability for the quasi-periodic solutions of the equations (1.0.1) and (1.0.5).

We look for small amplitude, quasi-periodic solutions for Hamiltonian, autonomous and parameter independent systems

$$u_t = X_H(u) := J\nabla H(u), \tag{3.1.1}$$

where H is given in (1.1.1) for the KdV case and in (1.2.4) for the DP case, with periodic boundary conditions, $x \in \mathbb{T}$. We consider the Hamiltonian H defined on the phase space $H_0^1(\mathbb{T}_x)$ (see (1.1.2)), since this space is left invariant by the flow of (1.0.1) and (1.0.5). The symplectic structure is given by (1.1.3) in the KdV case and (1.2.5) in the DP case.

Tangential sites. We look for solutions that are mainly Fourier supported on a finite set of modes. These ones are obtainied by perturbative arguments starting from approximately invariant finite dimensional tori. The set S of these modes is called *tangential set* (see (1.1.5)) and its elements are called *tangential sites*. S is a *symmetric* subset of $\mathbb{Z}$, since we look for real solutions (see (2.1.1)).

Autonomous and parameter independent system. Since the system (3.1.1) is autonomous, the frequency of the expected solutions is a priori unknown.
Since we deal with resonant problems the existence of quasi-periodic motions is due to the nonlinearity and the main modulation of the frequency vector of the solution respect to its amplitude is produced by the quadratic and cubic nonlinear terms of the vector field (or equivantely by the third and fourth degree terms in the Hamiltonian H). Moreover, the problems (1.0.1) and (1.0.5) have not external parameters (see the first comment after Theorem 1.1.3) which may influence the frequency. In order to deduce the approximate relation between amplitudes and frequencies we perform a Birkhoff normal form.

Weak Birkhoff normal form. The goal of this procedure, introduced in [8], is to find an approximately invariant finite dimensional manifold of the phase space on which the dynamic is *integrable* and *non-isochronous*, namely a set foliated by approximately invariant tori whose frequencies *twist* with their amplitudes ξ. The vectors ξ are used as *parameters* to control the non-resonance conditions that we shall impose on the frequencies.
Let us explain in detail this procedure, which is performed in Sections 4.1 and 5.2 respectively for the KdV and DP case. We decompose the phase space as

$$H_0^1(\mathbb{T}_x) := H_S \oplus H_S^{\perp}, \quad H_S := \mathrm{span}\{e^{\mathrm{i}\, j\, x} : j \in S\}, \quad H_S^{\perp} := \{u = \sum_{j \in S^c} u_j \, e^{\mathrm{i}\, j\, x} \in H_0^1(\mathbb{T}_x)\},$$

and we denote by $\Pi_S, \Pi_S^{\perp}$ the corresponding orthogonal projectors. The subspaces H_S and $H_S^{\perp}$ are symplectic respect to the 2-form Ω (see (1.1.3)). We write

$$u = v + z, \quad v := \Pi_S u := \sum_{j \in S} u_j \, e^{\mathrm{i}\, j\, x}, \quad z = \Pi_S^{\perp} u := \sum_{j \in S^c} u_j \, e^{\mathrm{i}\, j\, x}, \tag{3.1.2}$$

where v is called the *tangent* variable and z the *normal* one. In the following, we will identify $v = (v_j)_{j \in S}$ and $z = (z_j)_{j \in S^c}$. The dynamics on the tangential and normal part is quite different, hence it is useful to distinguish these components of the space.
We look for an analytic, invertible, symplectic change of coordinates Φ_B which transforms the Hamiltonian H into another Hamiltonian $\mathcal{H}$, whose homogeneous monomials of degree $\leq N$, for some number N which depends on the specifical problem (in the KdV case $N = 5$, in the DP case $N = 8$), do not include terms independent or linear in z (see Propositions 4.1.1 and 5.2.3). In this way, the set $\{z = 0\}$ is a submanifold of the phase space which is invariant for the Hamiltonian $\mathcal{H}$ truncated at degree N.
We verify that the dynamic on this set is integrable. Most of the equations in (1.0.1) parametrized by the coefficients c_i are not integrable, but the integrability of the truncated system on $\{z = 0\}$ is guaranteed by the particular form of the dispersion relation (see Lemma 4.1.3), which is the same of the KdV equation, and by imposing the assumption (1.1.7) on the tangential sites.
For the equation (1.0.5) we exploit the integrability of the Degasperis-Procesi equation (see Proposition 5.1.3) and we impose the assumption (1.2.13). In both cases we exploit the conservation of momentum (see Section 2.1.1).
Thus the submanifold $\{z = 0\}$ is foliated by finite dimensional approximately invariant tori with amplitudes ξ and frequency vectors $\omega(\xi)$. Then we require that the truncated system at $\{z = 0\}$ is non-isochronous, namely that the map $\xi \mapsto \omega(\xi)$ is a diffeomorphism, see Lemma 4.2.2 and Remark 5.3.1.

We note that the map Φ_B is of the form identity map plus a *finite rank* operator (see (4.1.6) and (5.2.6)) and for this reason the linearized operator is mildly modified respect to the one in the original coordinates. Another advantage of this procedure is that Φ_B is obtained as the time-one flow of an auxiliary Hamiltonian which is compactly Fourier supported thanks to the conservation of momentum (see Section 2.1.1 and Remark 4.1.2) and the fact that we eliminate only terms independent or linear in z. Hence there are no problems with the possible ill-posedness of any auxiliary system.
The disvantage is that the weak Birkhoff normal form does not normalize the terms $O(z^2)$.
This could be done for instance in [8], but the changes of coordinates would be of the form $\mathrm{I} + O(\partial_x^{-1})$

and such transformations produce terms ∂_{xx} and ∂_x in the transformed vector field $\mathcal{N}_4$ (see (1.0.32)). Actually, for the equations (1.0.1) and (1.0.5) we are not able to apply this stronger normal form method, called *Partial Birkhoff normal form* by Pöschel in [88]. Indeed we have problems with the definition of the transformations.
We remark that two steps of the Birkhoff procedure are sufficient to extract the frequency-amplitude modulation. The other steps are needed to have sufficiently good approximate solutions for the convergence of the Nash-Moser scheme. We note that the map (5.3.8) is a better approximation of the frequency-amplitude relation of the solutions respect to (4.2.7), since in the DP case we perform more steps of Birkhoff normal.

Action-angle variables We put action-angle variables (θ, y) (see (4.2.9) and (5.3.10)) on the finite dimensional submanifold $\{z = 0\}$ and we rescale the amplitudes ξ, the actions y and the normal variable z in order to work in a neighborhood of the torus $\{y = 0, z = 0\}$ (see (4.2.13) and (5.3.14)). This scaling is given in powers of a small parameter $\varepsilon > 0$. In this way, after these transformations the Hamiltonian $\mathcal{H}$ has the form (see (4.2.17) and (5.3.18))

$$H_\varepsilon = \mathcal{N} + P, \qquad \mathcal{N} := \alpha(\xi) \cdot y + \frac{1}{2}(N(\theta)z, z)_{L^2(\mathbb{T})} \tag{3.1.3}$$

where $\alpha(\xi)$ is the frequency-amplitude map and $N(\theta) := (\partial_z \nabla H_\varepsilon)(\theta, 0, 0)$. The Hamiltonian $\mathcal{N}$ is called *normal part* and it collects all the linear effects. Note that the coefficient of the normal form $\mathcal{N}$ depends on the angles θ, since the weak normal form procedure did not normalize the terms $O(z^2)$. The Hamiltonian P is regarded as a small perturbation of the normal part $\mathcal{N}$ and its size decreases when the number of steps of weak Birkhoff normal form increases (see Lemmata 4.3.3 and 5.4.2).

Nonlinear functional setting. In order to use the frequencies of the solutions as parameters, we embedd the system (3.1.3) in a ω-parameter family of Hamiltonians by setting $\xi = \alpha^{-1}(\omega)$ in (3.1.3). Note that these Hamiltonians, for $P = 0$, possess a invariant torus at the origin $(\varphi, 0, 0)$ with frequency ω. The parameters ω belong to a compact subset Ω_ε of $\mathbb{R}^\nu$, which is the image through $\alpha(\xi)$ of a ν-dimensional real cube (see (4.3.2)).
We look for zeros of the nonlinear functional equation (see (4.3.7) and (5.4.8))

$$\mathcal{F}(\omega, \varepsilon, i) = \big(\omega \cdot \partial_\varphi - X_{H_{\varepsilon,\omega}}\big)i = 0, \tag{3.1.4}$$

where the variable $i = i(\varphi)$, with $\varphi \in \mathbb{T}^\nu$, is an embedding of the torus $\mathbb{T}^\nu$ into the phase space supporting a quasi-periodic motion of frequency ω. The solutions of the equation (3.1.4) are constructed by a Nash-Moser iteration.
In this perspective we state the Theorems 4.3.2 and 5.4.1, which imply respectively Theorem 1.1.3 and Theorem 1.2.3.
Any solution of the problem (3.1.4) corresponds to a quasi-periodic solution for the Hamiltonian system (3.1.3) originating from an approximately invariant torus of amplitude $\xi = \alpha^{-1}(\omega)$. We require that ω satisfies diophantine conditions as (4.3.3) and (5.4.5). Note that in the DP case we require an additional condition (see (5.4.4)) in order to overcome the small divisor problem in the third step of Section 5.7.5.
We underline that the diophantine constant γ is very small (see (4.3.4) and (5.4.6)), actually it has

size $O(|\xi|)$, which is small with ε, since the frequencies ω are $O(|\xi|)$-close to vectors with rational components (see (1.1.6) and (1.2.9)).

The inversion of the linearized operator. The main issue in implementing a Nash-Moser scheme is the inversion of the linearized operator $d_i\mathcal{F}(i_0)$ at each approximate solution i_0. Actually, Zehnder [97] noted that it is sufficient only to approximately invert the linearized operator, in the sense that it is enough to construct an operator $\mathbf{T}_0$ such that

$$d_i\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathrm{I} = O(\mathcal{F}(i_0)\gamma^{-1}).$$

Note that the operator $\mathbf{T}_0$ is an exact right inverse of $d_i\mathcal{F}(i_0)$ if i_0 is a solution of (3.1.4).
The major difficulty is that the linear tangential and normal dynamcs are strongly coupled around an approximately invariant torus. To overcome this problem, in Sections 4.4 and 5.5, we use the abstract procedure developed by Berti-Bolle in [21]. This method reduces the search of an approximate inverse of the linearized operator to the invertibility of a quasi-periodically forced PDE restricted to the normal directions. More precisely, by introducing a suitable set of symplectic coordinates (ψ, η, w) the linearized system at an approximately invariant torus i_δ, which is close to i_0 and isotropic, is approximately triangularized. The canonicity of the change of variables $G_\delta(\theta, y, z) = (\psi, \eta, w)$ (see (4.4.12) and (5.5.12)) is proved thanks to the isotropy of the torus i_δ. In this way, the analysis of the linearized operator is reduced to the solvability of the linear equations in the normal variables. Once this problem is solved we deduce tame estimates on $\mathbf{T}_0$ thanks to the regularity of the functional $\mathcal{F}$ and the vicinity of i_δ to i_0.
For the analysis of the linearized equations we mention also a parallel method presented in [39] which does not exploit the Hamiltonian structure.

The linearized operator in the normal directions. Let us denote with $\mathcal{L}_\omega$ the linearized operator in the normal directions (see (4.5.33) and (5.6.31)). In order to get tame estimates on the inverse of $\mathcal{L}_\omega$ we conjugate it to a diagonal operator $\mathcal{L}_\infty$. This is done in two main steps: first, in Sections 4.6 and 5.7, we apply a regularization procedure, namely we conjugate $\mathcal{L}_\omega$, via changes of variables which are close to the identity and satisfy tame estimates on the Sobolev spaces H^s, to an operator that is diagonal up to a smoothing remainder R (see (4.6.128) and Theorem 5.7.3). Then, in Sections 4.6.8 and 5.8, we apply a KAM reducibility scheme to complete the diagonalization.
These two steps, in particular the regularization procedure, depends on the structure of the PDE which we consider. For this reason we refer to the next section for more details about these steps.
We point out that, in both cases (1.0.1) and (1.0.5), we have to deal with the terms $O(z^2)$ that the weak Birkhoff procedure did not touch. Indeed, some of these terms are non-perturbative for the KAM reducibility scheme, in the sense that they do not satisfy a smallness condition required by this procedure (see (4.6.130) and (5.8.16)). The normalization of the non-perturbative terms is done through a *linear Birkhoff normal form procedure* performed in Sections 4.6.5, 4.6.6 for the KdV case and 5.7.5 for the DP case. We remark that in the latter the steps of this procedure are 3 instead of 2. This is due to the fact that we are able to impose the 2-nd order Melnikov conditions (5.8.5) with a diophantine constant γ^* smaller than γ (actually $\gamma^* = \gamma^{3/2}$). This implies that the smallness condition (5.8.16) requires a smaller remainder coming from the regularization procedure.

The Nash-Moser nonlinear iteration. The Nash-Moser iteration (see Sections 4.7 and 5.9) requires a smallness condition like $\mathcal{F}(\varphi,0,0)\gamma^{-2} \ll 1$ in the KdV case and $\mathcal{F}(\varphi,0,0)\gamma^{-1}\gamma_*^{-1} \ll 1$ in the DP case. This is verified thanks to the weak Birkhoff steps. Then the algorithm provides a sequence of approximate solutions which converges to a final torus i_∞ such that $\mathcal{F}(\omega,\varepsilon,i_\infty)=0$ for ε small enough and for a frequency ω which satisfies the infinitely many Melnikov conditions imposed along the iteration. The set $\mathcal{C}_\varepsilon$ (see Theorems 4.3.2 and 5.4.1) of such frequencies is proved to have asymptotically full measure, in the sense that

$$\lim_{\varepsilon\to 0}\frac{|\mathcal{C}_\varepsilon|}{|\Omega_\varepsilon|}=1.$$

For the measure estimates we refer to Sections 4.7.1 and 5.9.1. We underline that in the DP case the proof of these estimates is quite more complicated.

3.1.1 The linear stability

At an exact solution $i_\infty(\omega t)$, the change of variables G_δ (see (4.4.12) and (5.5.12)), which depends on i_∞, puts the Hamiltonian $H_{\varepsilon,\omega}$ (see (3.1.3)) in the normal form

$$K := H_{\varepsilon,\omega}\circ G_\delta = \frac{1}{2}K_{20}(\psi)\eta\cdot\eta + (K_{11}(\psi)\eta, w)_{L^2(\mathbb{T}_x)} + \frac{1}{2}(K_{02}(\psi)w,w)_{L^2(\mathbb{T}_x)} + K_{\geq 3}(\psi,\eta,w),$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (η,w). Hence the linearized equations at i_∞ are

$$\begin{cases}\dot\psi = K_{20}(\omega t)\eta + K_{11}^T(\omega t)w,\\ \dot\eta = 0,\\ \dot w - JK_{02}(\omega t)w = JK_{11}(\omega t)\eta,\end{cases}$$

where $J=\partial_x$ in the KdV case and $J=(1-\partial_{xx})^{-1}(4-\partial_{xx})\partial_x$ in the DP case. Thus the actions $\eta(t)=\eta(0)$ do not evolve in time and the third equation reduces to

$$\dot w - JK_{02}(\omega t)w = JK_{11}(\omega t)\eta(0). \tag{3.1.5}$$

The right hand side of (3.1.5) is the linearized operator in the normal directions (see (4.5.33) and (5.6.31)), the left hand side is a forcing term. In Section 4.7 and 5.8 we, respectively, semi-conjugate and conjugate (3.1.5) to the diagonal system

$$h_t = -\mathcal{D}_\infty h + f(\omega t), \tag{3.1.6}$$

where $\mathcal{D}_\infty = \mathrm{diag}_j(\mathrm{i}d_j^\infty)$ and the eigenvalues $\mathrm{i}d_j^\infty\in\mathrm{i}\mathbb{R}$ are the *Floquet exponents* of the quasi-periodic solution i_∞. In the KdV case (see Theorem 4.6.19) we have

$$d_j^\infty = -m_3 j^3 + m_1 j + r_j^\infty, \quad j\in S^c$$

where (recall (2.1.7))

$$|m_3-1|\leq C\varepsilon^2, \quad |m_1|\leq C\varepsilon^2, \quad |r_j^\infty|^{sup}\leq C\varepsilon^{3-\delta}$$

for a small constant $\delta > 0$.
In the DP case (see Theorem 5.8.1) we have

$$d_j^\infty = m\frac{j(4+j^2)}{1+j^2} + \varepsilon^2\kappa_j + r_j^\infty, \quad j \in S^c$$

where κ_j is defined in (5.7.408) (see also (5.7.407)),

$$|m-1| \leq C\varepsilon^2, \quad |r_j^\infty|^{sup} \leq C\varepsilon^{4-\delta'}$$

for a small constant $\delta' > 0$.
For the KdV case, let us denote by Φ_ρ and Φ the two maps of the phase space such that $\Phi_\rho \mathcal{L}_\omega \Phi^{-1}$ is diagonal. These transformations have the form $\Phi_\rho := \Psi_1\rho\Psi_2$, $\Phi := \Psi_1\Psi_2$, where ρ denotes the multiplication by a function $\rho(\varphi)$ (see the proof of Theorem 4.6.21). Then the forcing term in (3.1.6) is

$$f(\omega t) := \Psi_1\rho\Psi_2\Big(JK_{11}(\omega t)\eta(0)\Big).$$

In the DP case, let us call Φ the map which conjugates $\mathcal{L}_\omega$ to a diagonal operator (see Theorem 5.8.5). Then the forcing term in (3.1.6) is

$$f(\omega t) := \Phi\Big(JK_{11}(\omega t)\eta(0)\Big).$$

The solutions of the non-homogeneous scalar equation

$$\dot h_j = -\mathrm{i}d_j^\infty\, h_j + f_j(\omega t), \quad j \in S^c$$

are

$$h_j(t) = c_j\, e^{-\mathrm{i}d_j^\infty t} + \tilde v_j(t), \quad \tilde v_j := \sum_{\ell\in\mathbb{Z}^\nu} \frac{f_{j\ell}\, e^{\mathrm{i}(\omega\cdot\ell)t}}{\mathrm{i}(\omega\cdot\ell + d_j^\infty)}, \quad j \in S^c.$$

Note that the first order Melnikov conditions (4.6.138) and (5.8.55) hold at a solution, so that $\tilde v_j$ is well defined. Since the changes of coordinates applied for the diagonalization are bounded on $H^s(\mathbb{T}_x)$, for $s \geq s_0$, then $\|f(\omega t)\|_{H^s(\mathbb{T}_x)} \leq C(s)|\eta(0)|$ for all t. As a consequence, the Sobolev norm of the solutions of (3.1.6) with initial condition $h(0) \in H^r(\mathbb{T}_x)$, for some $r \in (s_0, s)$ (in a suitable range of values), does not increase in time. In particular they satisfy

$$\|h(t)\|_{H_x^r} \leq C(s)\big(|\eta(0)| + \|h(0)\|_{H_x^r}\big), \quad \forall t \in \mathbb{R}.$$

Thus the linear stability of the solution i_∞ is proved.

3.2 The generalized KdV case

The main difference of the result presented in Section 1.1 with respect to the papers [8] , [9] is that we consider in (1.0.1) also quasi-linear quadratic and cubic terms in the nonlinear part of the equation. Now we explain the main consequences of this fact.

Twist condition. The presence of the quasi-linear monomials of degree three and four in the Hamiltonian (1.1.1) makes significantly harder the computations of the new Hamiltonian after two steps of Birkhoff normal form with respect to the case examined in [8] and [9] for the Hamiltonians

$$H_{\mathrm{KdV}}(u) = \frac{1}{2}\int_{\mathbb{T}} u_x^2\, dx + \frac{1}{6}\int_{\mathbb{T}} u^3\, dx, \quad H_{\mathrm{mKdV}}(u) = \frac{1}{2}\int_{\mathbb{T}} u_x^2\, dx \pm \frac{1}{4}\int_{\mathbb{T}} u^4\, dx.$$

Because of the integrability of the KdV system, in [8] and [9] the *twist condition*, namely, the invertibility of the frequency-amplitude map, is obtained for every choice of the tangential set S (see (1.1.5)). On the contrary, for the general case (1.1.1) the twist condition depends on the choice of the coefficients $c_1, \dots, c_7$ and the tangential sites $\bar{\jmath}_1, \dots, \bar{\jmath}_\nu$.
In Lemma 4.2.2 we provide the invertibility of the frequency-amplitude map for a large choice of the tangential sites and of the coefficients.

Accurate bounds on the small divisors. The diagonalization of the linearized operator in the normal directions $\mathcal{L}_\omega$ (see (4.5.33)) is obtained by conjugation with the same transformations defined in [8]. The main perturbative effect to the spectrum of $\mathcal{L}_\omega$ is due to the term $a_1(\omega t)\partial_{xxx}$ (see (4.5.33)) and the presence of u_x in the cubic part of the Hamiltonian density (1.0.3) affects this coefficient. In particular, $a_1 - 1 = O(\varepsilon)$, instead of $O(\varepsilon^3)$ as in [8]. In general, the corrections of the coefficients of $\mathcal{L}_\omega$ are bigger in size and this fact implies some difficulties in providing the smallness condition (4.6.132) required in Theorem 4.6.19. In particular, in Sections 4.6.3 and 4.6.4, the definition of the transformations (4.6.53) (a quasi-periodically reparametrization of time) and (4.6.75) (a quasi-periodically time dependent translation of the space variable) involve some small divisors, see (4.6.57) and (4.6.78). In order to prove that these changes of variables are close to the identity, one has to carefully estimate the inverse of the operator $\mathcal{D}_\omega^{-1}$, see Remark 4.6.8.

Non-perturbative terms. The normal frequencies are corrected by terms which do not satisfy the smallness condition required by Theorem 4.6.19, terms of size ε and ε^2. In order to normalize such terms one has to know their explicit expression. Indeed the ε^2-corrections are crucial for the measure estimate of the Cantor set $\mathcal{C}_\varepsilon$ (see Theorem 4.3.2).
Due to the quadraticity of the nonlinearities the ε^2-corrections change along the weak Birkhoff normal form and the regularization procedure. Thus we need to keep track of these terms after any change of coordinates. In the DP case the computations are harder and so we adopt a different strategy (see Proposition 5.7.34), which explain in the next section.

Measure estimates. The correction at the linear frequencies given by (4.2.7) has the same size of the correction at the normal frequencies j^3, $j \in S^c$, see (4.6.134). This fact brings to consider degenerate cases which do not appear in the study of the second Melnikov conditions for the equations considered in [8] and [9]. As a consequence, for the measure estimates we need to require conditions like (4.7.33) and (4.7.34), which are generic under suitable choices of the coefficients $c_1, \dots, c_7$.

3.3 The Degasperis-Procesi case

We discuss the main issues of Chapter 5.

Integrability and normal form. In [48] the authors prove the integrability of the Degasperis-Procesi equation by explicitly constructing its Laix pair. Moreover they provide a method for computing infinte conserved quantities for this system through the expansion of a spectral parameter. We are intersted in exploiting the integrable structure of this equation in order to overcome some small divisors problems.
For instance, at the N-th step of the weak Birkhoff normal form we deal with small denominators of the form $\overline{\omega} \cdot \ell$ (recall (1.2.9)) for $\ell \in \mathbb{Z}^\nu$ such that $|\ell| \leq N+2$. By the fact that the linear frequencies

of oscillations are rational numbers (see (1.2.9)), these denominators are rational functions of the tangental sites (1.2.8). In particular, for $N = 2$, the functions $\overline{\omega} \cdot \ell$ with $|\ell| = 4$ have the usual *trivial resonances*, i.e. 4-ples of the form $(i, -i, j, -j)$ (and its permutations), and other non trivial 4-ples coming from the zeros of a polynomial P (see (5.7.381)). Thus the normalized terms of the Hamiltonian at the second step of the weak Birkhoff normal form are determined by these 4-ples of integers and might not be integrable. In principle one might hope that P has no integer zeros, but we found a integer 4-ple of solutions with the assistance of Wolphram Mathematica.
Thanks to the constants of motion (5.1.1) we are able to prove that the coefficients of the Hamiltonian (after the first step of weak Birkhoff normal form) corresponding to the zeros of P are naught. In other words, there are only trivial resonances at order four.
The proof of this fact is based on the following aspects (verified along Section 5.2, see also Lemma C.0.1):

- all the Hamiltonians commuting with H are put in the same normal form simultaneously, namely by the same transformation,
- the normalized Hamiltonians commute with the quadratic part of any other commuting Hamiltonians.

In Proposition 5.1.3 we prove that for $n \leq 6$ there are no n-resonances (see Definition 5.1.2) unless the trivial ones. This fact will be fundamental also for the second step of the linear Birkhoff procedure of Section 5.7.5 (see Proposition 5.7.34).
We underline that the presence of non trivial (or generic) resonaces is a phenomenon which occurs also in the water waves equations (see [47], [40]). In this case no integrable structures are known and resonances at order four and five appear.
We remark that we are able to perform four steps of weak Birkhoff normal form without requiring any assumption on the tangential set S. However, we need to require generic conditions for the invertibility of the twist map. This is an interesting point, indeed in the absence of the perturbation f one should expect to be able to prove existence of quasi-periodic solutions of the Degasperis-Procesi equation for any choice of the tangential sites.

Analysis of the linearized operator in the normal directions. We describe the strategy of the reducibility for the linearized operator in the normal directions performed in Sections 5.7 and 5.8.

Reduction at the highest order. At any step of the Nash-Moser iteration the linearized operator in the normal directions has the form (recall that J in (1.2.5) is an operator of order 1)

$$\mathcal{L}_\omega = \Pi_S^\perp \Big(\mathcal{D}_\omega - J \circ (1 + a_0(\varphi, x)) + R \Big) \tag{3.3.1}$$

where $a_0(\varphi, x) = O(\varepsilon)$ (see (5.7.7), (5.7.8)) and R is a finite rank operator (see Proposition 5.6.5). The aim of Section 5.7 is to make constant the coefficient $a_0(\varphi, x)$. To do that, we conjugate $\mathcal{L}_\omega$ by the flow at time one $\Phi := \Phi^\tau_{|\tau=1}$ of a quasi-periodically parameter dependent Hamiltonian $S = (1/2) \int_{\mathbb{T}} b(\varphi)\, z^2\, dx$, where $b := \beta/(1 + \tau\beta_x)$ for some function $\beta(\varphi, x)$ to be determined (see (5.7.143)).
Actually this transformation is used also in the KdV case (see Lemma 4.6.3) in order to make

constant the leading term of the linearized operator, but here we have different issues. First, the symplectic structure (1.2.5) is more complicated than (1.1.3). Secondly, due to the asimptotically linear dispersion the time and space derivatives interact and the equation for determining β is not trivial, as we see in a moment.
In Proposition 5.7.21 we prove that

$$\Phi\mathcal{L}_\omega\Phi^{-1} = \Pi_S^\perp\Big(\mathcal{D}_\omega - J\circ(1+a_+(\varphi,x)) + \mathcal{Q}\Big) \tag{3.3.2}$$

where $\mathcal{Q}$ is a smoothing remainder and the new coefficient is (see (5.7.185))

$$1+a_+(\varphi,x) = -(\mathcal{D}_\omega\tilde{\beta})(\varphi,x+\beta(\varphi,x)) + (1+a_0(\varphi,x+\beta(\varphi,x)))(1+\tilde{\beta}_x(\varphi,x+\beta(\varphi,x))), \tag{3.3.3}$$

where $\tilde{\beta} := \tilde{\beta}(\tau,\varphi,x)$ is such that $x\mapsto x+\tau\beta(\varphi,x)$ is the inverse of the diffeomorphism of the torus $x\mapsto x+\tilde{\beta}(\tau,\varphi,x)$.
Then our goal is to find $\tilde{\beta}$, or equivalently β, such that the right hand side of (3.3.3) is constant. This problem is tantamount to find a diffeomorphism of the torus $(\varphi,x)\mapsto(\varphi,x+\beta(\varphi,x))$ which "straightens" the following vector field on the torus $\mathbb{T}^{\nu+1}$

$$\omega\cdot\frac{\partial}{\partial\varphi} - (1+a_0(\varphi,x))\frac{\partial}{\partial_x}.$$

This is the content of Proposition 5.7.22. This result is proved via a quadratic KAM iteration which requires the fullfillment of a smallness condition (see (5.7.224)). This condition is not satisfied by the initial coefficient a_0 (see (5.7.7), (5.7.8)), since $O(\varepsilon)\gamma^{-1}$ is big (recall (5.4.6)). Hence, in Section 5.7.4, we perform some preliminary steps in order to reduce the size of the coefficient a_0 by using transformations like Φ. The equations for the new coefficients are solved thanks to the generic assumption (H1) in (1.2.14). The transformed vector field by the diffeomorphism of the torus obtained by the KAM scheme of the Proposition 5.7.22 is constant coefficient on a restricted domain of parameters $\mathcal{O}_\infty^{2\gamma}$ (see (5.7.226)), since we need to impose first order Melnikov conditions to solve the (transport) homological equations at each step of this iteration.

Class of remainders. After the regularization procedure of Section 5.7 the linearized operator has a form like

$$\mathcal{L}_+ = \Pi_S^\perp\Big(\mathcal{D}_\omega - mJ - \varepsilon^2\mathfrak{D} + \mathcal{R}\Big),$$

where m is a real constant, $\mathfrak{D}$ is a diagonal operator 1-smoothing in space and $\mathcal{R}$ is a bounded remainder (smoothing in space) (see Theorem 5.7.3). In Section 5.8 we perform a KAM reducibility scheme (as in [27], [6]) which completely diagonalizes $\mathcal{L}_+$. This scheme works for remainders $\mathcal{R}$ which are -1-modulo-tame operators (see Definition 2.3.12) and fullfill the smallness condition (5.8.16). In order to obtain such a remainder after the regularization procedure, we define two classes of operators $\mathfrak{L}_\rho$ (see Definition 5.7.4) and $\mathfrak{C}_{1,\mathtt{b}}$ (see Definition 5.7.1), depending on two parameters ρ and $\mathtt{b}$ which are fixed respectively in (5.7.273) and (5.7.341), and we require that the remainders of the conjugations applied in Sections 5.7 and 5.7.5 belong to these sets.
The class $\mathfrak{L}_\rho$ is a set of operators ρ-smoothing in space, closed for the operations involved in the reduction at the highest order of the linearized operator, as for instance the composition and the conjugation by $\mathcal{A}^\tau$ (see (5.7.60)). All the properties of these operators are stated and proved in Section 5.7.1.

After the linear Birkhoff normal form, see Section 5.7.5, the transformed elements of this class belong to $\mathfrak{C}_{1,\mathtt{b}}$ for opportune values of the parameters ρ and $\mathtt{b}$, see Lemma 5.7.15. This set is constructed to be closed under the changes of coordinates defined by the Birkhoff maps (5.7.349), (5.7.374), (5.7.420). Moreover its elements have modulo-tame constant finite (see Lemma 5.8.2).

Flow of pseudo hyperbolic PDEs and Egorov analysis. The transformation Φ in (3.3.2) is the flow at time one of an hyperbolic PDE whose non autonomous φ-dependent vector field is $J \circ b(\varphi, \tau)$ (to shorten the notation we omit the projection $\Pi_S^{\perp}$). In Section 5.7.2 we investigate the structure of such flow, in particular in Proposition 5.7.16 we prove that Φ^{τ} is the composition of the flow $\mathcal{A}^{\tau}$ (see (5.7.60) for the definition and Appendix A.1 for the proof of some properties and estimates), which has been studied in [8], with an operator which is the sum of a pseudo differential operator of order -1 and an element of the class $\mathfrak{L}_{\rho}$ (see Definition (5.7.4)). This allows us to get tame estimates for Φ and to study the structure of an operator of the form (3.3.1) conjugated by Φ, see Proposition 5.7.21.
When it is possible we exploit the explicit expression (5.7.59) of the flow $\mathcal{A}^{\tau}$ to write the conjugated operators by Φ. In the other cases, we use the result given in Theorem 5.7.20, which is a *quantitative version* of an Egorov-type theorem. This asserts that a conjugated pseudo differential operator with symbol $w \in S^m$ (see Section 2.2) by $\mathcal{A}^{\tau}$ is the sum of a pseudo differential operator, whose weighted norm (see (2.2.10), (2.2.11)) is controlled by the weighted norm of w and the Lipschitz norm (see (2.1.9)) of the function β in (5.7.59), and an element of the class $\mathfrak{L}_{\rho}$.

Linear Birkhoff normal form. To implement the diagonalization procedure of Section 5.8 we impose second order Melnikov conditions (5.8.5) with a diophantine constant γ^* (see Theorem 5.8.1) smaller than γ (see (5.4.6)). As a consequence the remainder $\mathcal{R}$ (see Theorem 5.7.3) which comes out from the regularization procedure has to be very small in size. This is a bifurcation issue and we solve this problem by implementing three steps of linear Birkhoff normal form, which eliminate (or normalize) terms of order ε, ε^2 and ε^3.
For the second step we point out the result proved in Proposition 5.7.34. It asserts that the ε^2-corrections at the normal frequencies are the same which one would obtain by performing a *partial* Birkhoff normal form, namely by normalizing also the terms $O(z^2)$ in the Hamiltonian. This allows to easily compute these resonant terms also in a case which present complicated expressions of the eigenvalues (see for instance the expression of the ε^2-correction $\mathrm{i}\kappa_j$ defined in (5.7.407) and (5.7.408)).
For the third step we have to require some diophantine conditions (see (5.4.4) and (5.4.5)) in order to define the transformation Υ_3 in (5.7.420), (5.7.423). Actually this small divisor problem arises by the fact that the linear frequencies of oscillations satisfies (1.2.10).

Measure estimates. In Section 5.9.1 we prove the bounds (5.9.8). As we said above, the presence of quasi-linear terms which has the same size of the corrections at the linear frequencies might present degenerate cases in the study of the second Melnikov conditions. In the KdV case (1.0.1) the set of indices ℓ, j, k for which there are identically zero relations between the eigenvalues and the elements of the lattice $\omega \cdot \ell$, $\ell \in \mathbb{Z}^{\nu}$ is finite, thanks to the strong dispersion relation, in particular by the fact that there are only a finite number of $j, k \in \mathbb{Z} \setminus \{0\}$ such that $|\omega(j) - \omega(k)| = |j^3 - k^3| \leq C$. In the DP case this clearly is not true.
By the fact that the linear frequencies of oscillations are rational we have to control the asymptotic

behaviour $(\min_i |\bar{\jmath}_i| \to \infty)$ of some functions of the tangential sites $\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu$, as for instance the determinant of the twist matrix (see (1.2.15)), which could enter at some denominator and assume very small values.
This task might be not easy due to the presence of several variables $\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu$. Hence, we restrict the choice of the tangential sites, out of a (possibly large) ball (see (1.2.11)), to a cone (see (1.2.12)), where the functions above behave like functions of one variable only.

CHAPTER 4

Quasi-periodic solutions for quasi-linear generalized KdV equations

In this Chapter we prove Theorem 1.1.3. In Section 4.1 we perform three steps of weak Birkhoff normal form in order to extract parameters, which modulate the frequency-amplitude relation (4.2.7), and to provide a good starting point for the Nash-Moser iteration.
In Sections 4.2 and 4.3 we introduce action-angle variables (4.2.9) and we reformulate the problem of finding quasi-periodic solutions as the search for the zeros of the nonlinear functional $\mathcal{F}$ defined in (4.3.7). Adopting this new point of view, we devote the rest of the Chapter to the proof of Theorem 4.3.2, which implies Theorem 1.1.3.
In Section 4.4 we describe the construction of the approximate inverse for the linearized operator (4.4.2) following the abstract procedure developed in [21]. Thus the main issue is the approximate inversion of the linearized equations restricted at the normal directions, or equivalently the approximate inversion of the operator $\mathcal{L}_\omega$ in (4.5.33), which acts on the normal variables space $H_S^\perp$.
In Section 4.5 we prove that $\mathcal{L}_\omega$ has the form (4.5.33). In Section 4.6 we semi-conjugate $\mathcal{L}_\omega$ to a a diagonal operator $\mathcal{L}_\infty$ (see Theorem 4.6.19) and we provide tame estimates for the inverse of $\mathcal{L}_\omega$ (see Section 4.6.8).
In Section 4.7 we implement the Nash-Moser scheme of Theorem 4.7.1 to the functional $\mathcal{F}$ (recall (4.3.7)). In Section 4.7.1 we prove the measure estimates (4.7.8). This concludes the proof of Theorem 1.1.3.

4.1 Weak Birkhoff Normal form

The Hamiltonian (1.1.1) is $H = H^{(2)} + H^{(3)} + H^{(4)} + H^{(\geq 5)}$, where

$$
\begin{aligned}
&H^{(2)}(u) := \frac{1}{2}\int_{\mathbb{T}} u_x^2\,dx, \quad H^{(3)}(u) := \int_{\mathbb{T}} c_1\,u_x^3 + c_2\,u_x^2\,u + c_3\,u^3\,dx, \\
&H^{(4)}(u) := \int_{\mathbb{T}} c_4\,u_x^4 + c_5\,u_x^3\,u + c_6\,u_x^2\,u^2 + c_7\,u^4\,dx, \quad H^{(\geq 5)}(u) := \int_{\mathbb{T}} f_{\geq 5}(x, u, u_x)\,dx.
\end{aligned} \tag{4.1.1}
$$

For a finite dimensional space

$$
E := E_C := \mathrm{span}\left\{e^{\mathrm{i}\,j\,x} : 0 < |j| \leq C\right\}, \quad C > 0, \tag{4.1.2}
$$

let Π_E denote the corresponding L^2-projector on E.
The notation $R(v^{k-q}z^q)$ indicates a homogeneous polynomial of degree k in (v,z) of the form

$$R(v^{k-q}z^q) = M[\underbrace{v,\dots,v}_{(k-q)\ times},\underbrace{z,\dots,z}_{q\ times}], \quad M = k - \text{linear}.$$

We denote with $H^{(n,\geq k)}, H^{(n,k)}, H^{(n,\leq k)}$ the terms of type $R(v^{n-s}\,z^s)$, where, respectively, $s \geq k, s = k, s \leq k$, that appear in the homogeneous polynomial $H^{(n)}$ of degree n in the variables (v,z).
In particular, we have

$$H^{(3,\leq 1)} = \int_{\mathbb{T}} \left\{c_1(v_x^3 + 3\,v_x^2\,z_x) + c_2(v_x^2\,v + 2\,v_x\,v\,z_x + v_x^2\,z) + c_3(v^3 + 3\,v^2\,z)\right\}\,dx, \tag{4.1.3}$$

$$H^{(3,\geq 2)} = \int_{\mathbb{T}} \{c_1(z_x^3 + 3\,z_x^2\,v_x) + c_2(z_x^2\,z + z_x^2\,v + 2\,z_x\,z\,v_x) + c_3(z^3 + 3\,v^2\,z)\}\,dx, \tag{4.1.4}$$

$$H^{(4,0)} = \int_{\mathbb{T}} \{c_4\,v_x^4 + c_5\,v_x^3\,v + c_6\,v_x^2\,v^2 + c_7\,v^4\}\,dx. \tag{4.1.5}$$

Proposition 4.1.1. *(**Weak Birkhoff Normal form**) Assume Hypotesis* (S) *(see* (1.1.7)*). Then there exists an analytic invertible transformation of the phase space* $\Phi_B\colon H_0^1(\mathbb{T}_x) \to H_0^1(\mathbb{T}_x)$ *of the form*

$$\Phi_B(u) = u + \Psi(u), \quad \Psi(u) := \Pi_E\Psi(\Pi_E u), \tag{4.1.6}$$

where E *is a finite dimensional space as in* (4.1.2)*, such that the transformed Hamiltonian is*

$$\mathcal{H} = H \circ \Phi_B = H^{(2)} + \mathcal{H}^{(3)} + \mathcal{H}^{(4)} + \mathcal{H}^{(5)} + \mathcal{H}^{(\geq 6)}, \tag{4.1.7}$$

where $H^{(2)}$ *is defined in* (4.1.1)*,*

$$\begin{aligned}
&\mathcal{H}^{(3)} = c_1\int_{\mathbb{Z}}(z_x^3 + 3\,z_x^2\,v_x)\,dx + c_2\int_{\mathbb{Z}}(z_x^2\,z + z_x^2\,v + 2\,v_x\,z_x\,z)\,dx + c_3\int_{\mathbb{T}}(z^3 + 3v\,z^2)\,dx,\\
&\mathcal{H}^{(4)} = H^{(4)}_{4,0} + \mathcal{H}_{4,2} + \mathcal{H}^{(4,3)} + \mathcal{H}^{(4,4)}, \quad \mathcal{H}^{(4,2)} = R(v^2\,z^2), \quad \mathcal{H}_{4,3} = R(v\,z^3),\\
&\mathcal{H}^{(4,4)} = \int_{\mathbb{T}} c_4\,z_x^4 + c_5\,z_x^3\,z + c_6\,z_x^2\,z^2 + c_7\,z^4\,dx, \quad \mathcal{H}^{(5)} = \sum_{q=2}^{5} R(v^{5-q}\,z^q),
\end{aligned} \tag{4.1.8}$$

$H_4^{(4,0)}$ *is defined in* (4.1.25) *and* $\mathcal{H}^{(\geq 6)}$ *collects all the terms of order at least six in* (v,z).

The rest of this section is devoted to the proof of the Proposition 4.1.1.
We construct a symplectic map Φ_B as the composition of analytic and invertible transformations on the phase space that eliminates the terms linear in z and independent of it from the Hamiltonian (1.1.1). In this way, the Hamiltonian system (1.0.1) tranforms into one that is integrable and non-isocronous on the subspace $\{z = 0\}$.

Remark 4.1.2. We note that if $j_1,\dots,j_N \in \mathbb{Z}\setminus\{0\}, j_1 + \dots + j_N = 0$ and at most one of these integers does not belong to S, then $\max_{i=1,\dots,N}|j_i| \leq (N-1)C_S$, where $C_S := \max_{j\in S}|j|$. Thus, the vector field $X_{F^{(N)}}$, generated by the finitely supported Hamiltonian

$$F^{(N)} = \sum_{j_1+\dots+j_N=0} F^{(N)}_{j_1\dots j_N} u_{j_1}\dots u_{j_N},$$

is finite rank, and, in particular, it vanishes outside the finite dimensional subspace $E := E_{(N-1)C_S}$ (see (4.1.2)) and it has the form

$$X_{F^{(N)}}(u) = \Pi_E X_{F^{(N)}}(\Pi_E u).$$

Hence its flow $\Phi^{(N)}$ is analytic and invertible on the phase space $H_0^1(\mathbb{T}_x)$.

Step one. First we remove the cubic terms independent of z and linear in z from the Hamiltonian $H^{(3)}$ defined in (4.1.1) . We look for a symplectic transformation Φ_3 of the phase space which eliminates the monomials $u_{j_1} u_{j_2} u_{j_3}$ of $H^{(3)}$ with at most one index outside S.
We look for $\Phi_3 := (\Phi^t_{F^{(3)}})_{|t=1}$ as the time-1 flow map generated by the Hamiltonian vector field $X_{F^{(3)}}$, with an auxiliary Hamiltonian of the form

$$F^{(3)}(u) := \sum_{j_1+j_2+j_3=0} F^{(3)}_{j_1\, j_2\, j_3}\, u_{j_1}\, u_{j_2}\, u_{j_3}.$$

The transformed Hamiltonian is

$$\begin{aligned} &H_3 := H \circ \Phi_3 = H^{(2)} + H_3^{(3)} + H_3^{(4)} + H_3^{(\geq 5)}, \\ &H_3^{(3)} = H^{(3)} + \{H^{(2)}, F^{(3)}\}, \quad H_3^{(4)} = \frac{1}{2}\{\{H^{(2)}, F^{(3)}\}, F^{(3)}\} + \{H^{(3)}, F^{(3)}\} + H^{(4)}, \end{aligned} \tag{4.1.9}$$

where $H_3^{(\geq 5)}$ collects all the terms of order at least five in (v, z). In order to find the exact expression of $F^{(3)}$, we have to solve the homological equation

$$H^{(3)} + \{H^{(2)}, F^{(3)}\} = H^{(3,\geq 2)} \tag{4.1.10}$$

or, equivalently, $\{H^{(2)}, F^{(3)}\} = -\Pi_{\mathrm{Rg}(H^{(2)})} H^{(3,\leq 1)}$, see (4.1.3). In the Fourier representation, by (1.1.4) and (4.1.1), the equation (4.1.10) writes

$$\sum_{j_1+j_2+j_3=0} \mathrm{i}\,(j_1^3 + j_2^3 + j_3^3)\, F^{(3)}_{j_1 j_2 j_3}\, u_{j_1}\, u_{j_2}\, u_{j_3} = \sum_{(j_1,j_2,j_3)\in\mathcal{A}_3} (-\mathrm{i}\, c_1\, j_1 j_2 j_3 - c_2\, j_1 j_2 + c_3)\, u_{j_1} u_{j_2} u_{j_3} \tag{4.1.11}$$

where

$$\mathcal{A}_3 := \{(j_1, j_2, j_3) \in \mathbb{Z}^3 \setminus \{\mathbf{0}\} : j_1 + j_2 + j_3 = 0 \text{ and at least 2 indices among } j_1, j_2, j_3 \text{ belong to } S\}.$$

We note that if $(j_1, j_2, j_3) \in \mathcal{A}_3$ then $j_1^3 + j_2^3 + j_3^3 \neq 0$, because

$$j_1 + j_2 + j_3 = 0 \quad \Rightarrow \quad j_1^3 + j_2^3 + j_3^3 = 3\, j_1\, j_2\, j_3 \tag{4.1.12}$$

and $j_1, j_2, j_3 \in \mathbb{Z} \setminus \{0\}$.
Hence, to solve the equation (4.1.10) we choose

$$F^{(3)}_{j_1 j_2 j_3} := \begin{cases} \dfrac{-\mathrm{i}\, c_1\, j_1 j_2 j_3 - c_2\, j_1 j_2 + c_3}{\mathrm{i}(j_1^3 + j_2^3 + j_3^3)} & \text{if } (j_1, j_2, j_3) \in \mathcal{A}_3, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1.13}$$

By construction, all the monomials of $H^{(3)}$ with at least two indices outside S are not modified by the transformation Φ_3. Hence we have

$$H_3^{(3)} = c_1 \int_{\mathbb{Z}} (z_x^3 + 3\, z_x^2\, v_x)\, dx + c_2 \int_{\mathbb{Z}} (z_x^2\, z + z_x^2\, v + 2\, v_x\, z_x\, z)\, dx + c_3 \int_{\mathbb{T}} (z^3 + v\, z^2)\, dx. \tag{4.1.14}$$

Now we compute the fourth order term $H_3^{(4)}$ in (4.1.9). We have, by (4.1.10)

$$H_3^{(4)} = \frac{1}{2}\{\{H^{(2)}, F^{(3)}\}, F^{(3)}\} + \{H^{(3)}, F^{(3)}\} + H^{(4)} = \frac{1}{2}\{H^{(3,\leq 1)}, F^{(3)}\} + \{H_3^{(3)}, F^{(3)}\} + H^{(4)} \tag{4.1.15}$$

and by (4.1.11) and (4.1.13)

$$\begin{aligned} F^{(3)}(u) = & - \frac{c_1}{3} \int_{\mathbb{T}} v^3\, dx - c_1 \int_{\mathbb{T}} v^2\, z\, dx - \frac{c_2}{3} \int_{\mathbb{T}} (\partial_x^{-1} v)\, v^2\, dx - \frac{c_2}{3} \int_{\mathbb{T}} v^2\, (\partial_x^{-1} z)\, dx - \\ & - \frac{2\, c_2}{3} \int_{\mathbb{T}} v\, (\partial_x^{-1} v)\, z\, dx - \frac{c_3}{3} \int_{\mathbb{T}} (\partial_x^{-1} v)^3\, dx - c_3 \int_{\mathbb{T}} (\partial_x^{-1} v)^2\, (\partial_x^{-1} z)\, dx. \end{aligned} \tag{4.1.16}$$

Thus

$$\begin{aligned} \partial_x \nabla F^{(3)}(u) = & - c_1 \partial_x (v^2) - 2\, c_1 \partial_x \Pi_S [v\, z] + \frac{c_2}{3} \pi_0 [v^2] - \frac{c_2}{3} \partial_{xx} [(\partial_x^{-1} v)^2] - \\ & - \frac{2\, c_2}{3} \partial_x \Pi_S [(\partial_x^{-1} v) z + (\partial_x^{-1} z) v] + \frac{2\, c_2}{3} \Pi_S [v\, z] + c_3 \pi_0 [(\partial_x^{-1} v)^2] + \\ & + 2\, c_3\, \Pi_S [(\partial_x^{-1} v)(\partial_x^{-1} z)] \end{aligned} \tag{4.1.17}$$

where π_0 denotes the projection on the space of functions with zero space average, namely

$$\pi_0[u] := u(x) - \frac{1}{2\pi} \int_{\mathbb{T}} u(x)\, dx.$$

By (4.1.3), we get

$$\begin{aligned} \nabla H^{(3,\leq 1)}(u) = & - 3\, c_1 \partial_x (v_x^2) - 6\, c_1 \partial_x \Pi_S [v_x\, z_x] - c_2 \partial_{xx} (v^2) - 2\, c_2 \partial_{xx} \Pi_S [v\, z] + \\ & + c_2 \pi_0 [v_x^2] + 2\, c_2 \Pi_S [v_x\, z_x] + 3\, c_3 \pi_0 [v^2] + 6\, c_3 \Pi_S [v\, z]. \end{aligned} \tag{4.1.18}$$

Hence, by (1.1.4), (4.1.17), (4.1.18), we have

$$\begin{aligned} \frac{1}{2}\{H^{(3,\leq 1)}, F^{(3)}\} = & \frac{3\, c_1^2}{2} \int_{\mathbb{T}} \partial_x (v_x^2)\, \partial_x (v^2)\, dx - \frac{c_1\, c_2}{2} \int_{\mathbb{T}} v^2\, \partial_x (v_x^2)\, dx \\ & + \frac{c_1\, c_2}{2} \int_{\mathbb{T}} \partial_x (v_x^2) \partial_{xx} [(\partial_x^{-1} v)^2]\, dx - \frac{3\, c_1\, c_3}{2} \int_{\mathbb{T}} \partial_x (v_x^2)\, (\partial_x^{-1} v)^2 \\ & + \frac{c_2^2}{6} \int_{\mathbb{T}} (\partial_x (v^2))^2\, dx + \frac{c_2^2}{6} \int_{\mathbb{T}} \partial_{xx} (v^2)\, \partial_{xx} [(\partial_x^{-1} v)^2]\, dx \\ & - c_2\, c_3 \int_{\mathbb{T}} \partial_{xx} (v^2)\, (\partial_x^{-1} v)^2\, dx - \frac{c_1\, c_2}{2} \int_{\mathbb{T}} v_x^2\, \partial_x (v^2)\, dx + \frac{c_2^2}{6} \int_{\mathbb{T}} v_x^2\, \pi_0 [v^2]\, dx \\ & - \frac{c_2^2}{6} \int_{\mathbb{T}} v_x^2\, \partial_{xx} [(\partial_x^{-1} v)^2]\, dx + \frac{c_2\, c_3}{2} \int_{\mathbb{T}} v_x^2\, \pi_0 [(\partial_x^{-1} v)^2] + \frac{c_2\, c_3}{2} \int_{\mathbb{T}} (\pi_0 [v^2])^2\, dx \\ & - \frac{3\, c_3^2}{2} \int_{\mathbb{T}} v^2\, \pi_0 [(\partial_x^{-1} v)^2]\, dx + R(v^3\, z) + R(v^2\, z^2). \end{aligned} \tag{4.1.19}$$

By (4.1.4), we get

$$\begin{aligned}\nabla H_3^{(3)}(u) = &- 3\,c_1\,\partial_x(z_x^2) - 6\,c_1\,\partial_x\Pi_S^{\perp}[v_x\,z_x] - c_2\partial_{xx}(z^2) + c_2\pi_0[z_x^2] - 2\,c_2\partial_{xx}\Pi_S^{\perp}[v\,z]+\\ &+ 2\,c_2\Pi_S^{\perp}[v_x\,z_x] + 3\,c_3\pi_0[z^2] + 2\,c_3\Pi_S^{\perp}[v\,z].\end{aligned} \tag{4.1.20}$$

Thus by (1.1.4), (4.1.17), (4.1.20), we have

$$\begin{aligned}\{H_3^{(3)}, F^{(3)}\} = &\,3\,c_1^2\int_{\mathbb{T}}\partial_x(z_x^2)\,\partial_x(v^2)\,dx - c_1\,c_2\int_{\mathbb{T}} v^2\,\partial_x(z_x^2)\,dx+\\ &+ c_1\,c_2\int_{\mathbb{T}}\partial_x(z_x^2)\,\partial_{xx}[(\partial_x^{-1}v)^2]\,dx - 3\,c_1\,c_3\int_{\mathbb{T}}(\partial_x^{-1}v)^2\,\partial_x(z_x^2)\,dx+\\ &+ c_1\,c_2\int_{\mathbb{T}}(\partial_x^{-1}v)^2\,\partial_x(z_x^2)\,dx - \frac{c_2^2}{3}\int_{\mathbb{T}} v^2\,\partial_{xx}(z^2)\,dx+\\ &+ \frac{c_2^2}{3}\int_{\mathbb{T}}\partial_{xx}(z^2)\,\partial_{xx}[(\partial_x^{-1}v)^2]\,dx - c_2\,c_3\int_{\mathbb{T}}(\partial_x^{-1}v)^2\partial_{xx}(z^2)\,dx-\\ &- c_1\,c_2\int_{\mathbb{T}} z_x^2\,\partial_x(v^2)\,dx + \frac{c_2^2}{3}\int_{\mathbb{T}} z_x^2\,\pi_0[v^2]\,dx-\\ &- \frac{c_2^2}{3}\int_{\mathbb{T}} z_x^2\,\partial_{xx}[(\partial_x^{-1}v)^2]\,dx - c_2\,c_3\int_{\mathbb{T}} z_x^2\,\pi_0[(\partial_x^{-1}v)^2]\,dx-\\ &- 3\,c_1\,c_3\int_{\mathbb{T}} z^2\,\partial_x(v^2)\,dx + c_2\,c_3\int_{\mathbb{T}} v^2\,\pi_0[z^2]\,dx-\\ &- c_2\,c_3\int_{\mathbb{T}} z^2\,\partial_{xx}[(\partial_x^{-1}v)^2]\,dx + 3\,c_3^2\int_{\mathbb{T}}(\partial_x^{-1}v)^2\,\pi_0[z^2]\,dx+\\ &+ R(v^3\,z) + R(v\,z^3).\end{aligned} \tag{4.1.21}$$

Step two. We now construct a symplectic map Φ_4 to eliminate the term $H_3^{(4,1)}$ (which is linear in z) and to normalize $H_3^{(4,0)}$ (which is independent of z). We need the following elementary lemma.

Lemma 4.1.3. *(Lemma 13.4 in [70]) Let $j_1, j_2, j_3, j_4 \in \mathbb{Z}$ such that $j_1 + j_2 + j_3 + j_4 = 0$. Then*

$$j_1^3 + j_2^3 + j_3^3 + j_4^3 = -3(j_1 + j_2)(j_1 + j_3)(j_2 + j_3).$$

We look for a map $\Phi_4 := (\Phi^t_{F^{(4)}})_{|t=1}$ which is the time-1 flow map of an auxiliary Hamiltonian

$$F^{(4)}(u) := \sum_{\substack{j_1+j_2+j_3+j_4=0,\\ \text{at least 3 indices belong to } S}} F^{(4)}_{j_1j_2j_3j_4}\,u_{j_1}u_{j_2}u_{j_3}u_{j_4},$$

which has the same form of the Hamiltonian $H_3^{(4,0)} + H_3^{(4,1)}$. The transformed Hamiltonian is

$$H_4 := H_3\circ\Phi_4 = H^{(2)} + H_3^{(3)} + H_4^{(4)} + H_4^{(\geq 5)}, \quad H_4^{(4)} := \{H^{(2)}, F^{(4)}\} + H_3^{(4)} \tag{4.1.22}$$

and $H_4^{(\geq 5)}$ collects all the terms of order at least five in (v, z). We write

$$H_3^{(4)}(u) = \sum_{j_1+j_2+j_3+j_4=0} H^{(4)}_{3,\,j_1j_2j_3j_4}u_{j_1j_2j_3j_4}. \tag{4.1.23}$$

This makes sense since $H^{(3,\leq 1)}, H_3^{(3)}$ and $F^{(3)}$ preserve the momentum, hence also $H_3^{(4)}$ does it. We choose the coefficients

$$F^{(4)}_{j_1j_2j_3j_4} := \begin{cases} \dfrac{H^{(4)}_{3,j_1j_2j_3j_4}}{\mathrm{i}(j_1^3+j_2^3+j_3^3+j_4^3)} & \text{if } (j_1,j_2,j_3,j_4)\in\mathcal{A}_4, \\ 0 & \text{otherwise,} \end{cases} \tag{4.1.24}$$

where

$$\mathcal{A}_4 := \{(j_1,j_2,j_3,j_4)\in\mathbb{Z}^4\setminus\{\mathbf{0}\} : j_1+j_2+j_3+j_4=0, j_1^3+j_2^3+j_3^3+j_4^3\neq 0, \text{ and at most one among } j_1,j_2,j_3,j_4 \text{ outside } S\}.$$

By this definition, the symmetry of S and the Lemma 4.1.3, we have $H_4^{(4,1)}=0$, because there no exist $j_1,j_2,j_3\in S$ and $j_4\in S^c$ such that $j_1+j_2+j_3+j_4=0$, $j_1^3+j_2^3+j_3^3+j_4^3=0$. By construction, the terms $H_4^{(4,i)}=H_3^{(4,i)}, i=2,3,4$ are not changed by Φ_4.
It remains to compute the resonant part of $H_3^{(4,0)}$, i.e. the terms of $H_3^{(4)}$ of type $R(v^4)$ supported on the modes (j_1,j_2,j_3,j_4) that do not belong to $\mathcal{A}_4$.
If we call

$$\mathcal{B} := \{(j_1,j_2,j_3,j_4)\in S^4 \ : \ j_1+j_2+j_3+j_4=0,\ j_1^3+j_2^3+j_3^3+j_4^3=0,\ j_1+j_2\neq 0\}$$

then by (4.1.5), (4.1.19) we have

$$\begin{aligned} H_4^{(4,0)} &= -\frac{3\,c_1^2}{2}\sum_{\mathcal{B}}(j_1+j_2)^2 j_3j_4\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} + \frac{c_2^2}{6}\sum_{\mathcal{B}}(j_3+j_4)^2\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} \\ &- \frac{c_2^2}{6}\sum_{\mathcal{B}} j_3j_4\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} - \frac{c_2^2}{6}\sum_{\mathcal{B}}(j_1+j_2)^2(j_3+j_4)^2\frac{1}{j_1j_2}\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} \\ &+ \frac{c_2^2}{6}\sum_{\mathcal{B}}\frac{(j_1+j_2)^2\,j_3j_4}{j_1j_2}\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} + \frac{3}{2}c_3^2\sum_{\mathcal{B}}\frac{1}{\mathrm{i}j_1\,\mathrm{i}j_2}\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} \\ &+ \frac{c_2\,c_3}{2}\sum_{\mathcal{B}}u_{j_1}u_{j_2}u_{j_3}u_{j_4} - \frac{c_2\,c_3}{2}\sum_{\mathcal{B}}\frac{(j_1+j_2)^2}{j_1j_2}u_{j_1}u_{j_2}u_{j_3}u_{j_4} \\ &- \frac{c_2c_3}{2}\sum_{\mathcal{B}}\frac{(j_3+j_4)^2}{j_1\,j_2}\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} + \frac{c_2c_3}{2}\sum_{\mathcal{B}}\frac{j_3j_4}{j_1j_2}\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} \\ &+ c_4\sum_{\mathcal{B}\cup\{j_1+j_2=0\}} j_1\,j_2\,j_3\,j_4\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} - c_6\sum_{\mathcal{B}\cup\{j_1+j_2=0\}} j_1\,j_2\,u_{j_1}u_{j_2}u_{j_3}u_{j_4} \\ &+ c_7\sum_{\mathcal{B}\cup\{j_1+j_2=0\}} u_{j_1}u_{j_2}u_{j_3}u_{j_4}. \end{aligned} \tag{4.1.25}$$

By Lemma 4.1.3, if $j_1+j_2+j_3+j_4=0, j_1^3+j_2^3+j_3^3+j_4^3=0$ then $(j_1+j_2)(j_1+j_3)(j_2+j_3)=0$. We develop all the sums in (4.1.25) with respect to the first index j_1. The possible cases are:

(i) $\{j_2\neq -j_1, j_3=-j_1, j_4=-j_2\}$ $\quad$ (ii) $\{j_2\neq -j_1, j_3\neq -j_1, j_3=-j_2, j_4=-j_1\}$

(iii) $\{j_1+j_2=0\}$.

If $I := (I_{\bar{\jmath}_1}, \ldots, I_{\bar{\jmath}_\nu}) \in \mathbb{R}^\nu_+$ with $I_j := |u_j|^2, j \in S$, we get

$$\begin{aligned} H_4^{(4,0)}(I) = & - 12\, c_1^2 \sum_{j \in S^+} j^4\, I_j^2 - 24 c_1^2 \sum_{\substack{j,j' \in S^+, \\ j \neq j'}} j^2\, j'^2\, I_j\, I_{j'} - \frac{7c_2^2}{3} \sum_{j \in S^+} j^2\, I_j^2 \\ & - \frac{8c_2^2}{3} \sum_{\substack{j,j' \in S^+, \\ j \neq j'}} (j^2 + j'^2)\, I_j\, I_{j'} - 3\, c_3^2 \sum_{j \in S^+} \frac{1}{j^2} I_j^2 - 2\, c_2\, c_3 \sum_{j \in S^+} I_j^2 \\ & - 8 c_2 c_3 \sum_{\substack{j,j' \in S^+, \\ j \neq j'}} I_j\, I_{j'} + 6\, c_4 \sum_{j \in S^+} j^4\, I_j^2 + 12 c_4 \sum_{\substack{j,j' \in S, \\ j \neq j'}} j^2\, j'^2\, I_j\, i_{j'} \\ & + 2 c_6 \sum_{j \in S^+} j^2\, I_j^2 + 2 c_6 \sum_{\substack{j,j' \in S^+, \\ j \neq j'}} (j^2 + j'^2)\, I_j\, I_{j'} \\ & + 6\, c_7 \sum_{j \in S^+} I_j^2 + 12 c_7 \sum_{\substack{j,j' \in S^+, \\ j \neq j'}} I_j\, I_{j'}. \end{aligned} \tag{4.1.26}$$

The Hamiltonian system $H^{(2)} + H_3^{(3)} + H_4^{(4)}$, obtained by truncation at order 4 of the transformed Hamiltonian $H \circ \Phi_3 \circ \Phi_4$, possesses the invariant submanifold $\{z = 0\}$, and, restricted to this subspace, it is integrable. Indeed, if we introduce on H_S the action-angle variables $u \mapsto (\theta, I)$ by defining

$$u_j := v_j = \sqrt{I_j}\, e^{\mathrm{i}\theta_j}, \qquad I_j = I_{-j}, \quad \theta_{-j} = -\theta_j \quad j \in S, \tag{4.1.27}$$

the restriction of the Hamiltonian $H^{(2)} + H_3^{(3)} + H_4^{(4)}$ to $\{z = 0\}$, namely $\frac{1}{2} \int v_x^2\, dx + H_{4,0}^{(4)}$, depends only on the actions $I_{\bar{\jmath}_1}, \ldots, I_{\bar{\jmath}_\nu}$. We will prove later that, for a generic choice of the tangential sites, this system is also non-isochronous (actually it is formed by ν decoupled oscillators).
Due to the presence of a quadratic nonlinearity in the equation (1.0.1), we have to eliminate further monomials of $H^{(4)}$ in (4.1.22) in order to enter in a perturbative regime. Indeed, the minimal requirement for the convergence of the nonlinear Nash-Moser iteration is to eliminate the monomials $R(v^5)$ and $R(v^4\, z)$. Here we need the choice of the sites of Hypotesis $(\mathtt{S})$.

Step three. The homogeneous component of degree five of H_4 has the form

$$H_4^{(5)}(u) = \sum_{j_1 + \cdots + j_5 = 0} H^{(5)}_{4, j_1, \ldots, j_5}\, u_{j_1} u_{j_2} u_{j_3} u_{j_4} u_{j_5},$$

indeed, the Hamiltonian $H_4^{(5)}$ preserves the momentum, because $f_5(u, u_x)$ does not depend on x (see (1.0.4)). We want to remove from $H_4^{(5)}$ the terms with at most one index among $j_1, \ldots, j_5$ outside S. We consider the auxiliary Hamiltonian

$$F^{(5)} = \sum_{\substack{j_1 + \cdots + j_5 = 0, \\ \text{at most one index outside } S}} F^{(5)}_{j_1, \ldots, j_5}\, u_{j_1} \ldots u_{j_5}, \quad F^{(5)}_{j_1, \ldots, j_5} := \frac{H^{(5)}_{4, j_1, \ldots, j_5}}{\mathrm{i}(j_1^3 + \cdots + j_5^3)}. \tag{4.1.28}$$

Hypotesis $(\mathtt{S})$ implies that

($\mathtt{S}_0$) there is no choice of 5 integers $j_1, \dots, j_5 \in S$ such that

$$j_1 + \cdots + j_5 = 0, \quad j_1^3 + \cdots + j_5^3 = 0, \tag{4.1.29}$$

($\mathtt{S}_1$) there is no choice of 4 integers $j_1, \dots, j_4 \in S$ and $j_5 \in S^c$ such that (4.1.29) holds.

Hence $F^{(5)}$ in (4.1.28) is well defined. Let Φ_5 be the time-1 flow generated by $X_{F^{(5)}}$. The new Hamiltonian is

$$H_5 := H_4 \circ \Phi_5 = H^{(2)} + H_3^{(3)} + H_4^{(4)} + H_5^{(5)} + H_5^{(\geq 6)}, \quad H_5^{(5)} = \{H^{(2)}, F^{(5)}\} + H_4^{(5)}, \tag{4.1.30}$$

where $H_5^{(\geq 6)}$ collects all the terms of degree greater or equal than six, and, by the definition of $F^{(5)}$,

$$H_5^{(5)} = \sum_{q=2}^{5} R(v^{5-q} z^q). \tag{4.1.31}$$

Setting $\Phi_B := \Phi_3 \circ \Phi_4 \circ \Phi_5$ and renaming $\mathcal{H} := H^{(5)} = H \circ \Phi_B, \mathcal{H}^{(n)} = H_n^{(n)}$, by Remark (4.1.2), we conclude the proof of Proposition 4.1.1.

4.2 Action-angle variables

Consider the change of variable $v \mapsto (\theta, I)$ in (4.1.27), where the actions I are defined in the positive half space $\{v \in \mathbb{R}^\nu : v_i \geq 0, \forall i = 1, \dots, \nu\}$ and $\theta \in \mathbb{T}^\nu$. The symplectic form in (1.1.3) restricted to the subspace H_S transforms into the 2-form

$$\tilde{\Omega}_S = \sum_{j \in S^+} d\theta_j \wedge \frac{1}{j}\, dI_j. \tag{4.2.1}$$

Hence the Hamiltonian system $\mathcal{H}^{(\leq 5)} := H^{(2)} + H_3^{(3)} + H_4^{(4)} + H_5^{(5)}$ restricted to $\{z = 0\}$ writes

$$\begin{cases} \dot\theta_j = j\, \dfrac{\partial}{\partial I_j} \mathcal{H}^{(\leq 5)}(\theta, I, 0), & j \in S^+, \\ \dot I_j = -\dfrac{\partial}{\partial \theta_j} \mathcal{H}^{(\leq 5)}(\theta, I, 0), & j \in S^+. \end{cases} \tag{4.2.2}$$

We have that

$$\tilde{h}(I) := \mathcal{H}^{(\leq 5)}(\theta, I, 0) := \sum_{j \in S^+} j^2\, I_j + H_4^{(4,0)}(I) \tag{4.2.3}$$

depends only by the actions I, and, if we call $\omega_j(I) := j\, \partial_{I_j} \tilde{h}(I)$, we have

$$\begin{cases} \dot\theta_j = \omega_j(I), & j \in S^+, \\ \dot I_j = 0, & j \in S^+. \end{cases} \tag{4.2.4}$$

By (4.1.26)

$$
\begin{aligned}
\omega_j(I) = & j^3 - 24\, c_1^2\, j^5\, I_j - 48 c_1^2\, j^3 \sum_{k\in S^+, k\neq j} k^2 I_k - c_2^2 \frac{14}{3} j^3\, I_j - \frac{16}{3} c_2^2\, j^3 \sum_{k\in S^+, k\neq j} I_k \\
& - \frac{16}{3} c_2^2\, j \sum_{k\in S^+, k\neq j} k^2 I_k - 6\, c_3^2\, \frac{1}{j}\, I_j - 2\, c_2\, c_3 j\, I_j - 8 c_2 c_3\, j \sum_{k\in S^+, k\neq j} I_k + 12\, c_4 j^5\, I_j \\
& + 24 c_4\, j^3 \sum_{k\in S^+, k\neq j} k^2\, I_k + 4 c_6 j^3\, I_j + 4 c_6 j^3 \sum_{k\in S^+, k\neq j} I_k + 4 c_6\, j \sum_{k\in S^+, k\neq j} k^2 I_k \\
& + 12\, c_7\, j\, I_j + 24 c_7\, j \sum_{k\in S^+, k\neq j} I_k .
\end{aligned}
\tag{4.2.5}
$$

Hence, in a small neighbourhood of the origin of the phase space $H_0^1(\mathbb{T}_x)$, the submanifold $\{z = 0\}$ is foliated by invariant tori of amplitude ξ and frequency vector $\omega(\xi) := (\omega_j(\xi))_{j\in S^+}$ as in (4.2.5). We shall select from this set of tori the approximately invariant quasi-periodic solutions to be continued and we will use their *unperturbed actions* ξ as parameters. Moreover, we shall require that the frequencies of these tori vary in a one-to-one way with the actions ξ. Thanks to this fact, we could control the conditions that we shall impose on the frequencies ω through the amplitudes, and viceversa.

If we call $\vec{1}$ the vector in $\mathbb{R}^\nu$ with all components equal to 1 and

$$
D_S := \mathrm{diag}_{i=1,\dots,\nu}\{\jmath_i\}, \qquad v_k := D_S^k\, \vec{1}, \qquad U := \vec{1}^T\, \vec{1}, \tag{4.2.6}
$$

where the notaion $\vec{1}^T$ denotes the row vector with all components equal to 1, then we can write, in a compact form, the vector with components $\omega_j(I)$, with $j \in S^+$, in (4.2.5), as

$$
\omega(\xi) = \overline{\omega} + \mathbb{A}\, \xi, \tag{4.2.7}
$$

where $\overline{\omega}$ is the vector of the linear frequencies (see (1.1.6)) and

$$
\begin{aligned}
\mathbb{A} : = & (24 c_1^2 - 12 c_4) D_S^5 \{\mathrm{I} - 2 D_S^{-2} U D_S^2\} + (\frac{14}{3} c_2^2 - 4 c_6) D_S^3 \\
& + (4 c_6 - \frac{16}{3} c_2^2)\{D_S^3 U + D_S U D_S^2\} + 12 (c_2 c_3 - c_7) D_S + (24 c_7 - 16 c_2 c_3) D_S U - 6 c_3^2 D_S^{-1}.
\end{aligned}
\tag{4.2.8}
$$

The function of ξ in (4.2.7) is the *frequency-amplitude map*, which describes, at the main order, how the tangential frequencies are shifted by the amplitudes ξ.

In order to work in a neighbourhood of the unperturbed torus $\{I \equiv D_1 \xi\}$ it is advantageous to introduce a set of coordinates $(\theta, y, z) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$ adapted to it, defined by

$$
\begin{cases}
u_j := \sqrt{I_j}\, e^{\mathrm{i}\theta_j}\, e^{\mathrm{i}\, j\, x}, \qquad I_j := |j|(\xi_j + y_j), & j \in S, \\
u_j := z_j, & j \in S^c,
\end{cases}
\tag{4.2.9}
$$

where (recall $\overline{u}_j = u_{-j}$)

$$
\xi_{-j} = \xi_j, \quad \xi_j > 0, \quad y_{-j} = y_j, \quad \theta_{-j} = -\theta_j, \quad \theta_j \in \mathbb{T},\ y_j \in \mathbb{R}, \quad \forall j \in S. \tag{4.2.10}
$$

For the tangential sites $S^+ := \{\bar{\jmath}_1, \dots, \bar{\jmath}_\nu\}$ we will also denote

$$\theta_{\bar{\jmath}_i} := \theta_i, \quad y_{\bar{\jmath}_i} := y_i, \quad \xi_{\bar{\jmath}_i} := \xi_i, \quad \omega_{\bar{\jmath}_i} = \omega_i, \quad i = 1, \dots, \nu.$$

The symplectic 2-form Ω in (1.1.3) becomes

$$\mathcal{W} := \sum_{i=1}^{\nu} d\theta_i \wedge dy_i + \frac{1}{2} \sum_{j \in S^c} \frac{1}{\mathrm{i} j} dz_j \wedge dz_{-j} = \Big(\sum_{i=1}^{\nu} d\theta_i \wedge dy_i \Big) \oplus \Omega_{S^\perp} = d\Lambda, \tag{4.2.11}$$

where $\Omega_{S^\perp}$ denotes the restriction of Ω to $H_S^\perp$ and Λ is the Liouville 1-form on $\mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$ defined by $\Lambda_{(\theta,y,z)} \colon \mathbb{R}^\nu \times \mathbb{R}^\nu \times H_S^\perp \to \mathbb{R}$,

$$\Lambda_{(\theta,y,z)}[\hat{\theta}, \hat{y}, \hat{z}] := -y \cdot \hat{\theta} + \frac{1}{2} (\partial_x^{-1} z, \hat{z})_{L^2(\mathbb{T})}. \tag{4.2.12}$$

Working in a neighbourhood of the origin of the phase space, it is convenient to rescale the unperturbed actions ξ and the variables θ, y, z as

$$\xi \mapsto \varepsilon^2 \xi, \quad y \mapsto \varepsilon^{2b} y, \quad z \mapsto \varepsilon^b z. \tag{4.2.13}$$

The symplectic form in (4.2.11) transforms into $\varepsilon^{2b} \mathcal{W}$. Hence the Hamiltonian system generated by $\mathcal{H}$ in (4.1.7) transforms into the new Hamiltonian system

$$\begin{cases} \dot{\theta} = \partial_y H_\varepsilon(\theta, y, z), \\ \dot{y} = -\partial_\theta H_\varepsilon(\theta, y, z), \\ \dot{z} = \partial_x \nabla_z H_\varepsilon(\theta, y, z), \end{cases} \qquad H_\varepsilon := \varepsilon^{-2b} \mathcal{H} \circ A_\varepsilon, \tag{4.2.14}$$

where

$$A_\varepsilon(\theta, y, z) := \varepsilon\, v_\varepsilon(\theta, y) + \varepsilon^b z, \quad v_\varepsilon(\theta, y) := \sum_{j \in S} \sqrt{|j|} \sqrt{\xi_j + \varepsilon^{2(b-1)} y_j}\, e^{\mathrm{i}\theta_j} e^{\mathrm{i} j x}. \tag{4.2.15}$$

We still denote by

$$X_{H_\varepsilon} = (\partial_y H_\varepsilon, -\partial_\theta H_\varepsilon, \partial_x \nabla_z H_\varepsilon)$$

the Hamiltonian vector field in the variables $(\theta, y, z) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$. We now write explicitly the Hamiltonian defined in (4.2.14). The quadratic Hamiltonian $H^{(2)}$ in (4.1.1) becomes

$$\varepsilon^{-2b} H^{(2)} \circ A_\varepsilon = const + \sum_{j \in S^+} j^3 y_j + \frac{1}{2} \int_{\mathbb{T}} z_x^2 \, dx, \tag{4.2.16}$$

and by (4.1.1), (4.1.21) and (4.1.25) we have (writing $v_\varepsilon := v_\varepsilon(\theta, y)$)

$$\begin{aligned} H_\varepsilon(\theta, y, z) &= e(\xi) + \alpha(\xi) \cdot y + \frac{1}{2} \int_{\mathbb{T}} z_x^2 \, dx + \varepsilon \int_{\mathbb{T}} (3\, c_1 z_x^2\, (v_\varepsilon)_x + 3\, c_2 z_x^2\, v_\varepsilon + 2\, c_2 (v_\varepsilon)_x z_x z) \, dx \\ &\quad + \varepsilon^b \int_{\mathbb{T}} \big(c_1\, z_x^3 + c_2\, z_x^2\, z + c_3\, z^3 \, dx \big) \, dx + \frac{\varepsilon^{2\,b}}{2} \mathbb{M}\, y \cdot y + \varepsilon^{2b} \int_{\mathbb{T}} (c_4\, z_x^4 + c_5\, z_x^3\, z \\ &\quad + c_6\, z_x^2\, z^2 + c_7\, z^4) \, dx + \varepsilon^2 R((v_\varepsilon(\theta, y))^2 z^2) + \varepsilon^{1+b} R(v_\varepsilon(\theta, y)\, z^3) \\ &\quad + \varepsilon^3 R((v_\varepsilon(\theta, y))^3 z^2) + \varepsilon^{2+b} \sum_{q=3}^{5} \varepsilon^{(q-3)(b-1)} R((v_\varepsilon(\theta, y))^{5-q} z^q) \\ &\quad + \varepsilon^{-2b} \mathcal{H}^{(\geq 6)}(\varepsilon v_\varepsilon(\theta, y) + \varepsilon^b z) \end{aligned} \tag{4.2.17}$$

where the function $e(\xi)$ is a constant and

$$\alpha(\xi) = \overline{\omega} + \varepsilon^2 \mathbb{M}\xi\,, \qquad \mathbb{M} := \mathbb{A}\, D_S \tag{4.2.18}$$

is the frequency amplitude-map after the change of coordinates in (4.2.9) and the rescaling in (4.2.13). Usually $\mathbb{M}$ is called the *twist matrix* and we note that is symmetric.
We write the Hamiltonian in (4.2.17), eliminating the constant $e(\xi)$ which is irrelevant for the dynamics, as

$$\begin{aligned} &H_\varepsilon = \mathcal{N} + P, \qquad \mathcal{N}(\theta, y, z) = \alpha(\xi)\cdot y + \frac{1}{2}(N(\theta)z, z)_{L^2(\mathbb{T})},\\ &\frac{1}{2}(N(\theta)z, z)_{L^2(\mathbb{T})} := \frac{1}{2}((\partial_z \nabla H_\varepsilon)(\theta,0,0)[z], z)_{L^2(\mathbb{T})} = \frac{1}{2}\int_{\mathbb{T}} z_x^2\, dx + \\ &+\varepsilon \int_{\mathbb{T}} c_1 z_x^2\,(v_\varepsilon)_x(\theta,0)\,dx + \varepsilon\int_{\mathbb{T}} c_2\, z_x^2\, v_\varepsilon(\theta,0)\,dx + 2\,\varepsilon\, c_2 \int_{\mathbb{T}} z\, z_x\,(v_\varepsilon)_x(\theta,0)\,dx + \dots \end{aligned} \tag{4.2.19}$$

where $\mathcal{N}$ describes the linear dynamics, and $P := H_\varepsilon - \mathcal{N}$ collects the nonlinear perturbative effects.

As we said before, we require that the map (4.2.18) is a diffeomorphism. This function is affine, thus its invertibility is equivalent to the nondegenerancy (or *twist*) condition

$$\det \mathbb{M} := \det(D_S)\, \det\left(\frac{\partial^2}{\partial I_j\, I_k}\tilde{h}(I)\right)_{j,k\in\{1,\dots,\nu\}} \det(D_S) \neq 0. \tag{4.2.20}$$

Remark 4.2.1. The inequality (4.2.20) is equivalent to the classical Kolmogorov condition that requires the invertibility of the Hessian of the Hamiltonian $\tilde{h}$ in (4.2.3). The presence of the diagonal matrix D_S in (4.2.20) is due to the symplectic form (1.1.3) and the choice of the action-angle variables (4.2.9).

In the following lemma we prove that the condition (4.2.20) is satisfied for non-resonant coefficients and a generic choice of the tangential sites (see Definition 1.1.2).

Lemma 4.2.2. *If the coefficients $c_1, \dots, c_7$ are non-resonant, for a generic choice of the tangential sites $\overline{\jmath}_1, \dots, \overline{\jmath}_\nu$ (see Definition 1.1.2) the condition* (4.2.20) *is satisfied.*

Proof. We write $\mathbb{M} = D_S^{-1}\,\mathbb{B}\, D_S$, with

$$\begin{aligned} \mathbb{B} :&= (24c_1^2 - 12c_4)D_S^6\{\mathrm{I} - 2D_{-2}UD_S^2\} + (\frac{14}{3}c_2^2 - 4c_6)D_S^4 \\ &+ (4c_6 - \frac{16}{3}c_2^2)\{D_S^4U + D_S^2UD_S^2\} - 6c_3^2\mathrm{I} + 12(c_2c_3 - c_7)D_S^2 \\ &+ (24c_7 - 16c_2c_3)D_S^2U, \end{aligned} \tag{4.2.21}$$

where I is the identity $\nu\times\nu$ matrix. The determinant of $\mathbb{B}$ is a polynomial in the variables $(\overline{\jmath}_1, \dots, \overline{\jmath}_\nu)$ and, if $c_3 \neq 0$, it is not trivial, namely it is not identically zero. Indeed, the monomial of minimal degree of this polynomial originates from the matrix $6\,c_3^2\,\mathrm{I}$, that is invertible, and so it cannot be naught.
Similarly, if $c_3 = 0$ and $2\,c_1^2 - c_4 \neq 0$ then the monomial of maximal degree, i.e. six, is not zero, beacuse $(24\,c_1^2 - 12\,c_4)\,D_S^6\{\mathrm{I} - 2\,D_S^{-2}\,U\,D_S^2\}$ is invertible.

If $c_3 = 2\,c_1^2 - c_4 = 0$ and $c_7 \neq 0$ then the monomial of minimal degree, i.e. two, is $12\,c_7\,D_S^2\,(2\,U - \mathrm{I})$, that is invertible, indeed

$$(2\,U - \mathrm{I})^{-1} = \mathrm{I} - \frac{2}{2\,\nu + 1}\,U,$$

where $2\,\nu + 1 \neq 0$, because $\nu \in \mathbb{N}$. If $c_3 = 2\,c_1^2 - c_4 = c_7 = 0$ then

$$\mathbb{B} = D_S^4\left\{(\frac{14}{3}c_2^2 - 4c_6)\mathrm{I} + (4c_6 - \frac{16}{3}c_2^2)\{U + D_{-2}UD_S^2\}\right\}$$

The matrix $U + D_S^{-2}UD_S^2$ has rank 2 and its image is spanned by the vectors $\vec{1} := (1, \dots, 1)$ and v_{-2} The eigenvalues of this matrix, different from zero, are

$$\lambda_1 := \nu + \sqrt{\left(\sum_{i=1}^{\nu} \jmath_i^2\right)\left(\sum_{i=1}^{\nu} \jmath_i^{-2}\right)}, \quad \lambda_2 := \nu - \sqrt{\left(\sum_{i=1}^{\nu} \jmath_i^2\right)\left(\sum_{i=1}^{\nu} \jmath_i^{-2}\right)}. \tag{4.2.22}$$

Then, if $7c_2^2 - 6\,c_6 \neq 0$ and $\alpha := (8\,c_2^2 - 6\,c_6)/(7\,c_2^2 - 6\,c_6)$, we require that

$$\begin{cases} 1 - \alpha\,\lambda_1 \neq 0, \\ 1 - \alpha\,\lambda_2 \neq 0. \end{cases} \tag{4.2.23}$$

The conditions (4.2.23) are satisfied for every choice of the tangential sites if $4\,c_2^2 = 3\,c_6$; otherwise, it is satisfied by generic integer vectors $(\jmath_i)_{i=1}^{\nu}$. □

4.3 The nonliner functional setting

We look for an embedded invariant torus

$$i\colon \mathbb{T}^\nu \to \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp, \quad \varphi \mapsto i(\varphi) := (\theta(\varphi), y(\varphi), z(\varphi)) \tag{4.3.1}$$

of the Hamiltonian vector field X_{H_ε} filled by quasi-periodic solutions with diophantine frequency $\omega \in \mathbb{R}^\nu$, that we consider as independent parameters. We require that ω belongs to the set

$$\Omega_\varepsilon := \{\alpha(\xi) : \xi \in [1,2]^\nu\}, \tag{4.3.2}$$

where α is the function defined in (4.2.18) and, by Lemma 4.2.20, it is a diffeomorphism for a generic choice of the tangential sites.

Remark 4.3.1. We could consider any compact subset of $\{v \in \mathbb{R}^\nu : v_i > 0, \forall i = 1, \dots, \nu\}$ instead of the set $[1,2]^\nu$ in the definition (4.3.2).

Since any $\omega \in \Omega_\varepsilon$ is ε^2-close to the integer vector $\overline{\omega} := (\jmath_1^3, \dots, \jmath_\nu^3) \in \mathbb{N}^\nu$, we require that the constant γ in the diophantine inequality

$$|\omega \cdot \ell| \geq \gamma\,\langle \ell \rangle^{-\tau}, \quad \forall \ell \in \mathbb{Z}^\nu \setminus \{0\} \tag{4.3.3}$$

satisfies

$$\gamma = \varepsilon^{2+a}, \quad \text{for some } a > 0. \tag{4.3.4}$$

Note that the definition of γ in (4.3.4) is slightly stronger than the minimal condition, namely $\gamma \leq c\varepsilon^2$, with $c > 0$ small enough. In addition to (4.3.3) we shall also require that ω satisfies the first and the second order Melnikov non-resonance conditions. We fix the amplitude ξ as a function of ω and ε, as

$$\xi := \varepsilon^{-2}\, \mathbb{M}^{-1}[\omega - \overline{\omega}], \tag{4.3.5}$$

so that $\alpha(\xi) = \omega$ (see (4.2.18)). Consequently, H_ε in (4.2.19) becomes a (ω, ε)-parameter family of Hamiltonians which possess an invariant torus at the origin with frequency vector close to ω.
Now we look for an embedded invariant torus of the modified Hamiltonian vector field $X_{H_{\varepsilon,\zeta}} = X_{H_\varepsilon} + (0, \zeta, 0), \zeta \in \mathbb{R}^\nu$, which is generated by the Hamiltonian

$$H_{\varepsilon,\zeta}(\theta, y, z) := H_\varepsilon(\theta, y, z) + \zeta \cdot \theta, \quad \zeta \in \mathbb{R}^\nu. \tag{4.3.6}$$

We introduce ζ in order to control the average in the y-component of the linearized equations (4.4.23) (see (4.4.26)). However, the vector ζ has no dynamical consequences. Indeed it turns out that an invariant torus for the Hamiltonian vector field $X_{H_{\varepsilon,\zeta}}$ is actually invariant for X_{H_ε} itself.
Thus, we look for zeros of the nonlinear operator

$$\mathcal{F}(i, \zeta) := \mathcal{F}(i, \zeta, \omega, \varepsilon) := \mathcal{D}_\omega i(\varphi) - X_{\mathcal{N}}(i(\varphi)) - X_P(i(\varphi)) + (0, \zeta, 0) \tag{4.3.7}$$

$$:= \begin{pmatrix} \mathcal{D}_\omega \theta(\varphi) - \partial_y H_\varepsilon(i(\varphi)) \\ \mathcal{D}_\omega y(\varphi) + \partial_\theta H_\varepsilon(i(\varphi)) + \zeta \\ \mathcal{D}_\omega z(\varphi) - \partial_x \nabla_z H_\varepsilon(i(\varphi)) \end{pmatrix} = \begin{pmatrix} \mathcal{D}_\omega \Theta(\varphi) - \partial_y P(i(\varphi)) \\ \mathcal{D}_\omega y(\varphi) + \frac{1}{2}\partial_\theta (N(\theta(\varphi)) z(\varphi))_{L^2(\mathbb{T})} + \partial_\theta P(i(\varphi)) + \zeta \\ \mathcal{D}_\omega z(\varphi) - \partial_x N(\theta(\varphi))\, z(\varphi) - \partial_x \nabla_z P(i(\varphi)) \end{pmatrix}$$

where $\Theta(\varphi) := \theta(\varphi) - \varphi$ is $(2\pi)^\nu$-periodic and we use the short notation

$$\mathcal{D}_\omega := \omega \cdot \partial_\varphi. \tag{4.3.8}$$

The Sobolev norm of the periodic component of the embedded torus

$$\mathfrak{I}(\varphi) := i(\varphi) - (\varphi, 0, 0) := (\Theta(\varphi), y(\varphi), z(\varphi)), \tag{4.3.9}$$

is

$$\|\mathfrak{I}\|_s := \|\Theta\|_{H^s_\varphi} + \|y\|_{H^s_\varphi} + \|z\|_s \tag{4.3.10}$$

where $\|z\|_s := \|z\|_{H^s_{\varphi,x}}$ is defined in (2.1.2).
We link the rescaling of the domain of the variables (4.2.13) with the diophantine constant $\gamma = \varepsilon^{2+a}$ by choosing

$$\gamma = \varepsilon^{2+a} = \varepsilon^{2b}, \quad b := 1 + (a/2). \tag{4.3.11}$$

Other choices are possible (see Remark 5.2 in [9]).

Theorem 4.3.2. *If $c_1, \ldots, c_7$ are non-resonant and conditions* (C1)-(C2) *(see* (1.1.9), (1.1.10)*) hold, then for a generic choice of the tangential sites S (see* (1.1.7)*), satisfying the assumption* (S)*, there exists $\varepsilon_0 > 0$ small enough such that for all $\varepsilon \in (0, \varepsilon_0)$ there exist a constant $C > 0$ and a Cantor-like set $\mathcal{C}_\varepsilon \subseteq \Omega_\varepsilon$ (see* (4.3.2)*), with asymptotically full measure as $\varepsilon \to 0$, namely*

$$\lim_{\varepsilon \to 0} \frac{|\mathcal{C}_\varepsilon|}{|\Omega_\varepsilon|} = 1, \tag{4.3.12}$$

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_ε}*, and it is filled by quasi-periodic solutions with frequency* ω*. The torus* i_∞ *satisfies*

$$\|i_\infty(\varphi) - (\varphi, 0, 0)\|_{s_0+\mu}^{Lip(\gamma)} \leq C\,\varepsilon^{6-2b}\,\gamma^{-1} \tag{4.3.13}$$

for some $\mu := \mu(\nu) > 0$*. Moreover the torus* i_∞ *is linearly stable.*

Theorem 4.3.2 is proved in Sections 4.4-4.7. It implies Theorem 1.1.3 where the ξ_j in (1.1.11) are the components of the vector $\mathbb{M}^{-1}[\omega - \overline{\omega}]$.

Now we give tame estimates for the composition operator induced by the Hamiltonian vector fields $X_{\mathcal{N}}$ and X_P in (4.3.7).
Since the functions $y \to \sqrt{\xi + \varepsilon^{2(b-1)} y}, \theta \to e^{\mathrm{i}\theta}$ are analytic for ε small enough and $|y| \leq C$, the composition lemma A.0.3 implies that, for all $\Theta, y \in H^s(\mathbb{T}^\nu, \mathbb{R}^\nu)$ with $\|\Theta\|_{s_0}, \|y\|_{s_0} \leq 1$, one has the tame estimate

$$\|v_\varepsilon(\theta(\varphi), y(\varphi))\|_s \leq_s 1 + \|\Theta\|_s + \|y\|_s. \tag{4.3.14}$$

Hence the map A_ε in (4.2.15) satisfies, for all $\|\mathfrak{I}\|_{s_0}^{Lip(\gamma)} \leq 1$

$$\|A_\varepsilon(\theta(\varphi), y(\varphi), z(\varphi))\|_s^{Lip(\gamma)} \leq_s \varepsilon(1 + \|\mathfrak{I}\|_s^{Lip(\gamma)}). \tag{4.3.15}$$

In the following lemma we collect tame estimates for the Hamiltonian vector fields $X_{\mathcal{N}}$, X_P and X_{H_ε}, see (4.2.19).

Lemma 4.3.3. *Let* $\mathfrak{I}(\varphi)$ *in* (4.3.9) *satisfy* $\|\mathfrak{I}\|_{s_0+3}^{Lip(\gamma)} \leq C\,\varepsilon^{6-2b}\gamma^{-1}$*. Then*

$$\|\partial_y P(i)\|_s^{Lip(\gamma)} \leq_s \varepsilon^4 + \varepsilon^{2b}\|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}, \qquad \|\partial_\theta P(i)\|_s^{Lip(\gamma)} \leq_s \varepsilon^{6-2b}(1 + \|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}), \tag{4.3.16}$$

$$\|\nabla_z P(i)\|_s^{Lip(\gamma)} \leq_s \varepsilon^{5-b} + \varepsilon^{6-b}\gamma^{-1}\|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}, \qquad \|X_P(i)\|_s^{Lip(\gamma)} \leq_s \varepsilon^{6-2b} + \varepsilon^{2b}\|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}, \tag{4.3.17}$$

$$\|\partial_\theta\partial_y P(i)\|_s^{Lip(\gamma)} \leq_s \varepsilon^4 + \varepsilon^5\gamma^{-1}\|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}, \qquad \|\partial_y\nabla_z P(i)\|_s^{Lip(\gamma)} \leq_s \varepsilon^{b+3} + \varepsilon^{2b-1}\|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}, \tag{4.3.18}$$

$$\|\partial_{yy} P(i) - \frac{\varepsilon^{2b}}{2}\mathbb{M}\|_s^{Lip(\gamma)} \leq_s \varepsilon^{2+2b} + \varepsilon^{2b+3}\gamma^{-1}\|\mathfrak{I}\|_{s+2}^{Lip(\gamma)} \tag{4.3.19}$$

and for all $\hat{\imath} := (\hat{\Theta}, \hat{y}, \hat{z})$*,*

$$\|\partial_y d_i X_P(i)[\hat{\imath}]\|_s^{Lip(\gamma)} \leq_s \varepsilon^{2b-1}(\|\hat{\imath}\|_{s+3}^{Lip(\gamma)} + \|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}\|\hat{\imath}\|_{s_0+3}^{Lip(\gamma)}), \tag{4.3.20}$$

$$\|d_i X_{H_\varepsilon}(i)[\hat{\imath}] + (0, 0, \partial_{xxx}\hat{z})\|_s^{Lip(\gamma)} \leq_s \varepsilon(\|\hat{\imath}\|_{s+3}^{Lip(\gamma)} + \|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}\|\hat{\imath}\|_{s_0+3}^{Lip(\gamma)}), \tag{4.3.21}$$

$$\|d_i^2 X_{H_\varepsilon}(i)[\hat{\imath}, \hat{\imath}]\|_s^{Lip(\gamma)} \leq_s \varepsilon(\|\hat{\imath}\|_{s+3}^{Lip(\gamma)}\|\hat{\imath}\|_{s_0+3}^{Lip(\gamma)} + \|\mathfrak{I}\|_{s+3}^{Lip(\gamma)}(\|\hat{\imath}\|_{s_0+3}^{Lip(\gamma)})^2). \tag{4.3.22}$$

In the sequel we will use that, by the diophantine condition (4.3.3), the operator $\mathcal{D}_\omega^{-1}$ (see (4.3.8)) is defined for all functions u with zero φ-average, and satisfies

$$\|\mathcal{D}_\omega^{-1} u\|_s \leq_s \gamma^{-1}\,\|u\|_{s+\tau}, \quad \|\mathcal{D}_\omega^{-1} u\|_s^{Lip(\gamma)} \leq_s \gamma^{-1}\|u\|_{s+2\tau+1}^{Lip(\gamma)}. \tag{4.3.23}$$

4.4 Approximate inverse

We will apply a Nash-Moser iterative scheme in order to find a zero of the functional $\mathcal{F}(i,\zeta)$ defined in (4.3.7). In particular, we shall construct a sequence of approximate solutions of

$$F(i,\zeta)=0 \tag{4.4.1}$$

that converges to a solution in some Sobolev norm. In order to define this sequence we need to solve some linearized equations and this is the main difficulty for implementing the Nash-Moser algorithm. Zehnder noted in [97] that it is sufficient to invert these equations only approximately to get a scheme with still quadratic speed of convergence. We refer to [97] for the precise notion of *approximate right inverse*, whose main feature is to be an *exact right inverse* when the equation is linearized at an exact solution. Hence, our aim is to construct an approximate right inverse of the linearized operator

$$d_{i,\zeta}\mathcal{F}(i_0,\zeta_0)[\hat\imath,\hat\zeta]=\mathcal{D}_\omega\hat\imath-d_iX_{H_\varepsilon}(i_0(\varphi))[\hat\imath]+(0,\hat\zeta,0) \tag{4.4.2}$$

at any approximate solution i_0 of the equation (4.4.1), and to verify that satisfies some tame estimates.
Note that $d_{i,\zeta}\mathcal{F}(i_0,\zeta_0)=d_{i,\zeta}\mathcal{F}(i_0)$ is independent of ζ_0 (see (4.3.7)).
We will implement the general strategy in [21], [22] which reduces the search of an approximate right inverse of (4.4.2) to the search of an approximate inverse on the normal directions only.

It is well known that an invariant torus i_0 with diophantine flow is isotropic (see e.g.[21]), namely the pull-back 1-form $i_0^*\Lambda$ is closed, where Λ is the Liouville 1-form in (4.2.12). This is tantamount to say that the 2-form $\mathcal{W}$ in (4.2.11) vanishes on the torus $i_0(\mathbb{T}^\nu)$, because $i_0^*\mathcal{W}=i_0^*d\Lambda=d\,i_0^*\Lambda$. For an "approximately invariant" embedded torus i_0 the 1-form $i_0^*\Lambda$ is only "approximately closed". In order to make this statement quantitative we consider

$$i_0^*\Lambda=\sum_{k=1}^{\nu}a_k(\varphi)\,d\varphi_k,\quad a_k(\varphi):=-([\partial_\varphi\theta_0(\varphi)]^Ty_0(\varphi))_k+\frac{1}{2}(\partial_{\varphi_k}z_0(\varphi),\partial_x^{-1}z_0(\varphi))_{L^2(\mathbb{T})} \tag{4.4.3}$$

and we quantify how small is

$$i_0^*\mathcal{W}=d\,i_0^*\Lambda=\sum_{1\le k<j\le\nu}A_{kj}(\varphi)\,d\varphi_k\wedge d\varphi_j,\quad A_{kj}(\varphi):=\partial_{\varphi_k}a_j(\varphi)-\partial_{\varphi_j}a_k(\varphi). \tag{4.4.4}$$

In order to get estimates for an approximate inverse we need to take in account the size of the "error" function

$$Z(\varphi):=(Z_1,Z_2,Z_3)(\varphi):=\mathcal{F}(i_0,\zeta_0)(\varphi)=\omega\cdot\partial_\varphi i_0(\varphi)-X_{H_{\varepsilon,\zeta_0}}(i_0(\varphi)), \tag{4.4.5}$$

which gives a measure of how i_0 is near to be an exact solution.
Along this section we will always assume the following hypotesis (which will be proved at each step of the Nash-Moser iteration):

- **Assumption**. The map $\omega\mapsto i_0(\omega)$ is a Lipschitz function defined on some subset $\mathcal{O}_0\subseteq\Omega_\varepsilon$, where Ω_ε is defined in (4.3.2), and, for some $\mu:=\mu(\tau,\nu)>0$,

$$\|\mathfrak{I}_0\|_{s_0+\mu}^{Lip(\gamma)}\le\varepsilon^{6-2b}\gamma^{-1},\quad \|Z\|_{s_0+\mu}^{Lip(\gamma)}\le\varepsilon^{6-2b},\quad\gamma=\varepsilon^{2+a},\quad a\in(0,1/6), \tag{4.4.6}$$

  where $\mathfrak{I}_0(\varphi):=i_0(\varphi)-(\varphi,0,0)$.

The next lemma proves that if i_0 is a solution of the equation (4.4.1), then the parameter ζ has to be naught, hence the embedded torus i_0 supports a quasi-periodic solution of the "original" system with Hamiltonian H_ε.

Lemma 4.4.1. *(Lemma 6.1 in [8]) We have*

$$|\zeta_0|^{Lip(\gamma)} \leq C\|Z\|_{s_0}^{Lip(\gamma)}.$$

In particular, if $\mathcal{F}(i_0,\zeta_0)=0$ then $\zeta_0=0$ and the torus $i_0(\varphi)$ is invariant for the vector field X_{H_ε}.

Now we estimate the size of $i_0^*\mathcal{W}$ in terms of the error function Z.
By (4.4.3), (4.4.4) we get

$$\|A_{k\,j}\|_s^{Lip(\gamma)} \leq_s \|\mathfrak{I}_0\|_{s+2}^{Lip(\gamma)}.$$

Moreover, we have the following bound.

Lemma 4.4.2. *(Lemma 6.2 in [8]) The coefficients $A_{k\,j}(\varphi)$ in (4.4.4) satisfy*

$$\|A_{k\,j}\|_s^{Lip(\gamma)} \leq_s \gamma^{-1}(\|Z\|_{s+2\tau+2}^{Lip(\gamma)} + \|Z\|_{s_0+1}^{Lip(\gamma)}\|\mathfrak{I}_0\|_{s+2\tau+2}^{Lip(\gamma)}). \tag{4.4.7}$$

As in [21], the idea is to analyze the operator linearized at an isotropic embedded torus i_δ, because the isotropy of the torus allows to construct a symplectic set of coordinates around it for which the linear tangential dynamic and the normal one are decoupled. Thus, the linear system becomes "triangular" and the hard part is to solve the equation in the normal directions (see Section 7).
Now we see that we can slightly modify i_0 (indeed, it is sufficient to move the y-component only) to obtain an isotropic torus i_δ, that is an approximate solution as well as i_0. At the end of this section, we will prove that we are able to construct an approximate right inverse of (4.4.2) starting from an approximate inverse of $d_{i,\zeta}\mathcal{F}(i_\delta,\zeta_0)[\hat{\imath},\hat{\zeta}]$.

In the paper we denote equivalently the differential ∂_i or d_i. We use the notation $\Delta_\varphi := \sum_{k=1}^\nu \partial_{\varphi_k}^2$ and we denote by $\sigma := \sigma(\nu,\tau)$ possibly different (larger) "loss of derivatives" constants.

Lemma 4.4.3. ***(Isotropic torus)****(Lemma 6.3 in [8]) The torus $i_\delta = (\theta_0(\varphi), y_\delta(\varphi), z_0(\varphi))$ defined by*

$$y_\delta := y_0 + [\partial_\varphi\theta_0(\varphi)]^{-T}\rho(\varphi), \quad \rho_j(\varphi) := \Delta_\varphi^{-1}\sum_{k=1}^\nu \partial_{\varphi_j}A_{k\,j}(\varphi), \tag{4.4.8}$$

is isotropic. If (4.4.6) holds, then, for some $\sigma := \sigma(\nu,\tau)$,

$$\|y_\delta - y_0\|_s^{Lip(\gamma)} \leq_s \gamma^{-1}(\|Z\|_{s+\sigma}^{Lip(\gamma)}\|\mathfrak{I}_0\|_{s_0+\sigma}^{Lip(\gamma)} + \|Z\|_{s_0+\sigma}^{Lip(\gamma)}\|\mathfrak{I}_0\|_{s+\sigma}^{Lip(\gamma)}), \tag{4.4.9}$$

$$\|\mathcal{F}(i_\delta,\zeta_0)\|_s^{Lip(\gamma)} \leq_s \|Z\|_{s+\sigma}^{Lip(\gamma)} + \|Z\|_{s_0+\sigma}^{Lip(\gamma)}\|\mathfrak{I}_0\|_{s+\sigma}^{Lip(\gamma)}, \tag{4.4.10}$$

$$\|\partial_i i_\delta[\hat{\imath}]\|_s \leq_s \|\hat{\imath}\|_s + \|\mathfrak{I}_0\|_{s+\sigma}\|\hat{\imath}\|_s. \tag{4.4.11}$$

We introduce a set of symplectic coordinates adapted to the isotropic torus i_δ. We consider the map $G_\delta\colon (\psi,\eta,w)\to(\theta,y,z)$ of the phase space $\mathbb{T}^\nu\times\mathbb{R}^\nu\times H_S^\perp$ defined by

$$\begin{pmatrix}\theta\\ y\\ z\end{pmatrix} := G_\delta\begin{pmatrix}\psi\\ \eta\\ w\end{pmatrix} := \begin{pmatrix}\theta_0(\psi)\\ y_\delta(\psi) + [\partial_\psi\theta_0(\psi)]^{-T}\eta + [(\partial_\theta\tilde{z}_0)(\theta_0(\psi))]^T\partial_x^{-1}w\\ z_0(\psi)+w\end{pmatrix} \tag{4.4.12}$$

where $\tilde{z}_0 := z_0(\theta_0^{-1}(\theta))$ (indeed $\theta_0 \colon \mathbb{T}^\nu \to \mathbb{T}^\nu$ is a diffeomorphism, because $\theta_0(\varphi) - \varphi$ is small). It is proved in [21] (Lemma 6.3) that G_δ in (4.4.12) is symplectic, using that the torus i_δ is isotropic. In the new coordinates, i_δ is at the origin, i.e. $(\psi, \eta, w) = (\psi, 0, 0)$. The transformed Hamiltonian $K := K(\psi, \eta, w, \zeta_0)$ is (recall (4.3.6))

$$\begin{aligned} K := H_{\varepsilon,\zeta_0} \circ G_\delta = \theta_0(\psi) \cdot \zeta_0 + K_{00}(\psi) + K_{10}(\psi) \cdot \eta + (K_{01}(\psi), w)_{L^2(\mathbb{T})} + \frac{1}{2} K_{20}(\psi)\eta \cdot \eta + \\ + (K_{11}(\psi)\eta, w)_{L^2(\mathbb{T})} + \frac{1}{2}(K_{02}(\psi)w, w)_{L^2(\mathbb{T})} + K_{\geq 3}(\psi, \eta, w) \end{aligned} \tag{4.4.13}$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (η, w). At any fixed ψ, the Taylor coefficient $K_{00}(\psi) \in \mathbb{R}, K_{10}(\psi) \in \mathbb{R}^\nu, K_{01}(\psi) \in H_S^\perp, K_{20}(\psi)$ is a $\nu \times \nu$ real matrix, $K_{02}(\psi)$ is a linear self-adjoint operator of $H_S^\perp$ and $K_{11}(\psi) \colon \mathbb{R}^\nu \to H_S^\perp$.
Note that the above Taylor coefficients do not depend on the parameter ζ_0.
The Hamilton equations associated to (4.4.13) are

$$\begin{cases} \dot\psi = K_{10}(\psi) + K_{20}(\psi)\eta + K_{11}^T(\psi)w + \partial_\eta K_{\geq 3}(\psi, \eta, w) \\ \dot\eta = -[\partial_\psi \theta_0(\psi)]^T \zeta_0 - \partial_\psi K_{00}(\psi) - [\partial_\psi K_{10}(\psi)]^T \eta - [\partial_\psi K_{01}(\psi)]^T w - \\ \qquad - \partial_\psi \left(\frac{1}{2} K_{20}(\psi)\eta \cdot \eta + (K_{11}(\psi)\eta, w)_{L^2(\mathbb{T})} + \frac{1}{2}(K_{02}(\psi)w, w)_{L^2(\mathbb{T})} + K_{\geq 3}(\psi, \eta, w) \right) \\ \dot w = \partial_x (K_{01}(\psi) + K_{11}(\psi)\eta + K_{02}(\psi)w + \nabla_w K_{\geq 3}(\psi, \eta, w)) \end{cases} \tag{4.4.14}$$

where $[\partial_\psi K_{10}(\psi)]^T$ is the $\nu \times \nu$ transposed matrix and $[\partial_\psi K_{01}(\psi)]^T, K_{11}^T(\psi) \colon H_S^\perp \to \mathbb{R}^\nu$ are defined by the duality relation

$$(\partial_\psi K_{01}(\psi)[\hat\psi], w)_{L^2(\mathbb{T})} = \hat\psi \cdot [\partial_\psi K_{01}(\psi)]^T w, \quad \forall \hat\psi \in \mathbb{R}^\nu, w \in H_S^\perp,$$

and similarly for K_{11}. Explicitly, for all $w \in H_S^\perp$, and denoting $\underline{e}_k$ the k-th versor of $\mathbb{R}^\nu$,

$$K_{11}^T(\psi)w = \sum_{k=1}^{\nu} (K_{11}^T(\psi)w \cdot \underline{e}_k)\, \underline{e}_k = \sum_{k=1}^{\nu} (w, K_{11}(\psi)\underline{e}_k)_{L^2(\mathbb{T})} \underline{e}_k \in \mathbb{R}^\nu. \tag{4.4.15}$$

In the next lemma we estimate the coefficients K_{00}, K_{10}, K_{01} in the Taylor expansion (4.4.13). The term K_{10} describes how the tangential frequencies vary with respect to ω. Note that on an exact solution (i_0, ζ_0) we have $K_{00}(\psi) = const, K_{10} = \omega$ and $K_{01} = 0$.

Lemma 4.4.4. *(Lemma 6.4 in [8]) Assume (4.4.6). Then there is $\sigma := \sigma(\tau, \nu)$ such that*

$$\|\partial_\psi K_{00}\|_s^{Lip(\gamma)} + \|K_{10} - \omega\|_s^{Lip(\gamma)} + \|K_{01}\|_s^{Lip(\gamma)} \leq_s \|Z\|_{s+\sigma}^{Lip(\gamma)} + \|Z\|_{s_0+\sigma}^{Lip(\gamma)} \|\mathfrak{I}_0\|_{s+\sigma}^{Lip(\gamma)}.$$

Remark 4.4.5. By Lemma 4.4.1 if $\mathcal{F}(i_0, \zeta_0) = 0$ and, by Lemma 4.4.4, the Hamiltonian (4.4.13) simplifies to

$$K = const + \omega \cdot \eta + \frac{1}{2} K_{20}(\psi)\eta \cdot \eta + (K_{11}(\psi)\eta, w)_{L^2(\mathbb{T})} + \frac{1}{2}(K_{02}(\psi)w, w)_{L^2(\mathbb{T})} + K_{\geq 3}. \tag{4.4.16}$$

In general, the normal form (4.4.16) provides a control of the linearized equations in the normal bundle of the torus.

We now estimate K_{20}, K_{11} in (4.4.13). The norm of K_{20} is the sum of the norms of its matrix entries.

Lemma 4.4.6. *(Lemma* 6.6 *in [8]) Assume* (4.4.6)*. Then for some* $\sigma := \sigma(\nu,\tau)$ *we have*

$$\|K_{20} - \frac{\varepsilon^{2b}}{2}\mathbb{M}\|_s^{Lip(\gamma)} \leq_s \varepsilon^{2b+2} + \varepsilon^{2b}\|\mathfrak{I}_0\|_{s+\sigma}^{Lip(\gamma)} + \varepsilon^3\gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{Lip(\gamma)}\|Z\|_{s+\sigma}^{Lip(\gamma)}, \tag{4.4.17}$$

$$\|K_{11}\eta\|_s^{Lip(\gamma)} \leq_s \varepsilon^5\gamma^{-1}\|\eta\|_s^{Lip(\gamma)} + \varepsilon^{2b-1}(\|\mathfrak{I}_0\|_{s+\sigma}^{Lip(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{Lip(\gamma)}\|Z\|_{s+\sigma}^{Lip(\gamma)})\|\eta\|_{s_0}^{Lip(\gamma)}, \tag{4.4.18}$$

$$\|K_{11}^T w\|_s^{Lip(\gamma)} \leq_s \varepsilon^5\gamma^{-1}\|w\|_{s+2}^{Lip(\gamma)} + \varepsilon^{2b-1}(\|\mathfrak{I}_0\|_{s+\sigma}^{Lip(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{Lip(\gamma)}\|Z\|_{s+\sigma}^{Lip(\gamma)})\|w\|_{s_0+2}^{Lip(\gamma)}. \tag{4.4.19}$$

In particular

$$\|K_{20} - \frac{\varepsilon^{2b}}{2}\mathbb{M}\|_{s_0}^{Lip(\gamma)} \leq \varepsilon^6\gamma^{-1}, \quad \|K_{11}\eta\|_{s_0}^{Lip(\gamma)} \leq \varepsilon^5\gamma^{-1}\|\eta\|_{s_0}^{Lip(\gamma)},$$
$$\|K_{11}^T w\|_{s_0}^{Lip(\gamma)} \leq \varepsilon^5\gamma^{-1}\|w\|_{s_0}^{Lip(\gamma)}.$$

We apply the linear change of variables

$$DG_\delta(\varphi,0,0)\begin{pmatrix}\hat\psi\\ \hat\eta\\ \hat w\end{pmatrix} := \begin{pmatrix}\partial_\psi\theta_0(\varphi) & 0 & 0\\ \partial_\psi y_\delta(\varphi) & [\partial_\psi\theta_0(\varphi)]^{-T} & -[(\partial_\theta\tilde z_0)(\theta_0(\varphi))]^T\partial_x^{-1}\\ \partial_\psi z_0(\varphi) & 0 & \mathrm{I}\end{pmatrix}\begin{pmatrix}\hat\psi\\ \hat\eta\\ \hat w\end{pmatrix}. \tag{4.4.20}$$

In these new coordinates the linearized operator $d_{i,\zeta}\mathcal{F}(i_\delta,\zeta_0)$ is "approximately" the operator obtained linearizing (4.4.14) at $(\psi,\eta,w,\zeta) = (\varphi,0,0,\zeta_0)$ with $\mathcal{D}_\omega$ instead of ∂_t, namely

$$\begin{pmatrix}\mathcal{D}_\omega\hat\psi - \partial_\psi K_{10}(\varphi)[\hat\psi] - K_{20}(\varphi)\hat\eta - K_{11}^T(\varphi)\hat w\\ \mathcal{D}_\omega\hat\eta + [\partial_\psi\theta_0(\varphi)]^T\hat\zeta + \partial_\psi[\partial_\psi\theta_0(\varphi)]^T[\hat\psi,\zeta_0] + \partial_{\psi\psi}K_{00}(\varphi)[\hat\psi] + [\partial_\psi K_{10}(\varphi)]^T\hat\eta + [\partial_\psi K_{01}(\varphi)]^T\hat w\\ \mathcal{D}_\omega\hat w - \partial_x\{\partial_\psi K_{01}(\varphi)[\hat\psi] + K_{11}(\varphi)\hat\eta + K_{02}(\varphi)\hat w\}\end{pmatrix}. \tag{4.4.21}$$

We give estimate on the composition operator induced by the transformation (4.4.20).

Lemma 4.4.7. *(Lemma* 6.7 *in [8]) Assume* (4.4.6) *and let* $\hat\imath := (\hat\psi,\hat\eta,\hat w)$*. Then, for some* $\sigma := \sigma(\tau,\nu)$*, we have*

$$\begin{aligned}&\|DG_\delta(\varphi,0,0)[\hat\imath]\|_s + \|DG_\delta(\varphi,0,0)^{-1}[\hat\imath]\|_s \leq_s \|\hat\imath\|_s + (\|\mathfrak{I}_0\|_{s+\sigma} + \gamma^{-1}\|\mathfrak{I}_0\|_{s+\sigma}\|Z\|_{s+\sigma})\|\hat\imath\|_{s_0}\\ &\|D^2G_\delta(\varphi,0,0)[\hat\imath_1,\hat\imath_2]\|_s \leq_s \|\hat\imath_1\|_s\|\hat\imath_2\|_{s_0} + \|\hat\imath_1\|_{s_0}\|\hat\imath\|_s\\ &\qquad + (\|\mathfrak{I}_0\|_{s+\sigma} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}\|Z\|_{s+\sigma})\|\hat\imath\|_{s_0}\|\hat\imath_2\|_{s_0}.\end{aligned} \tag{4.4.22}$$

Moreover the same estimates hold if we replace $\|\cdot\|_s$ *with* $\|\cdot\|_s^{Lip(\gamma)}$*.*

In order to construct an approximate inverse of (4.4.21) it is sufficient to solve the system of equations

$$\mathbb{D}[\hat\psi,\hat\eta,\hat w,\hat\zeta] := \begin{pmatrix}\mathcal{D}_\omega\hat\psi - K_{20}(\varphi)\hat\eta - K_{11}^T(\varphi)\hat w\\ \mathcal{D}_\omega\hat\eta + [\partial_\psi\theta_0(\varphi)]^T\hat\zeta\\ \mathcal{D}_\omega\hat w - \partial_x K_{11}(\varphi)\hat\eta - \partial_x K_{02}(\varphi)\hat w\end{pmatrix} = \begin{pmatrix}g_1\\ g_2\\ g_3\end{pmatrix} \tag{4.4.23}$$

which is obtained by (4.4.21) neglecting the terms that are naught at a solution, namely, by Lemmata (4.4.1) and (4.4.4), $\partial_\psi K_{10}, \partial_{\psi\psi}K_{00}, \partial_\psi K_{00}, \partial_\psi K_{01}$ and $\partial_\psi[\partial_\psi\theta_0(\varphi)]^T[\cdot,\zeta_0]$.

Remark 4.4.8. We will use the following notations for the averages of a function $v(\varphi, x)$

$$M_x[v] := \frac{1}{2\pi}\int_{\mathbb{T}} v(\varphi, x)\, dx, \quad M_\varphi[v] := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} v(\varphi, x)\, d\varphi \tag{4.4.24}$$

and $M_{\varphi,x}[v] := M_x[M_\varphi[v]] = M_\varphi[M_x[v]]$.

First, we solve the second equation, namely

$$\mathcal{D}_\omega \hat{\eta} = g_2 - [\partial_\psi \theta_0(\varphi)]\hat{\zeta}. \tag{4.4.25}$$

We choose $\hat{\zeta}$ so that the φ-average of the right hand side of (4.4.25) is zero, namely

$$\hat{\zeta} = M_\varphi[g_2]. \tag{4.4.26}$$

Note that the φ-averaged matrix $M_\varphi[(\partial_\psi \theta_0)^T] = M_\varphi[\mathrm{I} + (\partial_\psi \Theta_0)^T] = \mathrm{I}$ since we have that $\theta_0(\varphi) = \varphi + \Theta_0(\varphi)$ and $\Theta_0(\varphi)$ is periodic. Therefore

$$\hat{\eta} = \mathcal{D}_\omega^{-1}(g_2 - [\partial_\psi \theta_0(\varphi)]^T M_\varphi[g_2]) + M_\varphi[\hat{\eta}], \qquad M_\varphi[\hat{\eta}] \in \mathbb{R}^\nu, \tag{4.4.27}$$

where the average $M_\varphi[\hat{\eta}]$ will be fix when we deal with the first equation.
We now analyze the third equation, namely

$$\mathcal{L}_\omega \hat{w} = g_3 + \partial_x K_{11}(\varphi)\hat{\eta}, \quad \mathcal{L}_\omega := \omega \cdot \partial_\varphi - \partial_x K_{02}(\varphi). \tag{4.4.28}$$

If we fix $\hat{\eta}$, then solving the equation (4.4.28) is tantamount to invert the operator $\mathcal{L}_\omega$. For the moment we assume the following hypotesis (that will be proved in Section 8)

- **Inversion Assumption**. There exists a set $\Omega_\infty \subseteq \Omega_\varepsilon$ such that for all $\omega \in \Omega_\infty$, for every function $g \in H^{s+\mu}_{S^\perp}(\mathbb{T}^{\nu+1})$ there exists a solution $h := \mathcal{L}_\omega^{-1} g$ of the linear equation $\mathcal{L}_\omega h = g$ which satisfies

$$\|\mathcal{L}_\omega^{-1} g\|_s^{Lip(\gamma)} \leq_s \gamma^{-1}(\|g\|_{s+\mu}^{Lip(\gamma)} + \varepsilon\gamma^{-1}\{\|\mathfrak{I}_0\|_{s+\mu}^{Lip(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s+\mu}^{Lip(\gamma)}\|Z\|_{s+\mu}^{Lip(\gamma)}\}\|g\|_{s_0}^{Lip(\gamma)}) \tag{4.4.29}$$

for some $\mu := \mu(\tau, \nu)$.

Remark 4.4.9. The term $\varepsilon\gamma^{-1}\|\mathfrak{I}_0\|_{s+\mu}^{Lip(\gamma)}$ arises because the remainder $\mathcal{R}_6$ in Section 8 contains the term $\varepsilon(\|\Theta_0\|_{s+\mu}^{Lip(\gamma)} + \|y_\delta\|_{s+\mu}^{Lip(\gamma)}) \leq_s \varepsilon\|\mathfrak{I}_0\|_{s+\mu}^{Lip(\gamma)}$, see Lemma 4.6.18.
These big constants coming from the tame estimates for the inverse of the linearized operators at any approximate solution will be dominated by the quadraticity of the Nash-Moser scheme.

By the above assumption, there exists a solution of (4.4.28)

$$\hat{w} = \mathcal{L}_\omega^{-1}[g_3 + \partial_x K_{11}(\varphi)\hat{\eta}]. \tag{4.4.30}$$

Now consider the first equation

$$\mathcal{D}_\omega \hat{\psi} = g_1 + K_{20}\hat{\eta} - K_{11}^T(\varphi)\hat{w}. \tag{4.4.31}$$

Substituting (4.4.27), (4.4.30) in the equation (4.4.31), we get

$$\mathcal{D}_\omega \hat{\psi} = g_1 + M_1(\varphi)M_\varphi[\hat{\eta}] + M_2(\varphi)g_2 + M_3(\varphi)g_3 - M_2(\varphi)[\partial_\psi \theta_0]^T M_\varphi[g_2], \tag{4.4.32}$$

where

$$M_1(\varphi) := K_{20}(\varphi) + K_{11}^T(\varphi)\mathcal{L}_\omega^{-1}\partial_x K_{11}(\varphi),\ M_2(\varphi) := M_1(\varphi)\mathcal{D}_\omega^{-1},\ M_3(\varphi) := K_{11}^T(\varphi)\mathcal{L}_\omega^{-1}.$$

In order to solve the equation (4.4.32) we have to choose $M_\varphi[\hat\eta]$ such that the right hand side in (4.4.32) has zero φ-average.
By Lemma 4.4.6 and (4.4.6), the φ-averaged matrix $M_\varphi[M_1] = \varepsilon^{2b}M + O(\varepsilon^{10}\gamma^{-3})$. Therefore, for ε small, $M_\varphi[M_1]$ is invertible and $M_\varphi[M_1]^{-1} = O(\varepsilon^{-2b}) = O(\gamma^{-1})$. Thus we define

$$M_\varphi[\hat\eta] := -(M_\varphi[M_1])^{-1}\{M_\varphi[g_1] + M_\varphi[M_2 g_2] + M_\varphi[M_3 g_3] - M_\varphi[M_2(\partial_\psi\theta_0)^T]\, M_\varphi[g_2]\}. \tag{4.4.33}$$

With this choice of $M_\varphi[\hat\eta]$ the equation (4.4.32) has the solution

$$\hat\psi := \mathcal{D}_\omega^{-1}\{g_1 + M_1(\varphi)M_\varphi[\hat\eta] + M_2(\varphi)g_2 + M_3(\varphi)g_3 - M_2(\varphi)[\partial_\psi\theta_0]^T M_\varphi[g_2]\}. \tag{4.4.34}$$

In conclusion, we have constructed a solution $(\hat\psi, \hat\eta, \hat w, \hat\zeta)$ of the linear system (4.4.23). We resume this in the following proposition, giving also estimates on the inverse of the operator $\mathbb{D}$ defined in (4.4.23).

Proposition 4.4.10. *(Proposition 6.9 in [8]) Assume (4.4.6) and (4.4.29). Then, for all $\omega \in \Omega_\infty$, for all $g := (g_1, g_2, g_3)$, the system (4.4.23) has a solution $\mathbb{D}^{-1}g := (\hat\psi, \hat\eta, \hat w, \hat\zeta)$ where $(\hat\psi, \hat\eta, \hat w, \hat\zeta)$ are defined in (4.4.34), (4.4.27), (4.4.30), (4.4.26). Moreover, we have*

$$\|\mathbb{D}^{-1}g\|_s^{Lip(\gamma)} \le_s \gamma^{-1}(\|g\|_{s+\mu}^{Lip(\gamma)} + \varepsilon\gamma^{-1}\{\|\mathfrak{I}_0\|_{s+\mu}^{Lip(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\mu}^{Lip(\gamma)}\|\mathcal{F}(i_0,\zeta_0)\|_{s+\mu}^{Lip(\gamma)}\}\|g\|_{s_0+\mu}^{Lip(\gamma)}). \tag{4.4.35}$$

Eventually we prove that the operator

$$\mathbf{T}_0 := (D\tilde G_\delta)(\varphi,0,0) \circ \mathbb{D}^{-1} \circ (DG_\delta(\varphi,0,0))^{-1} \tag{4.4.36}$$

is an approximate right inverse of $d_{i,\zeta}\mathcal{F}(i_0)$ where $\tilde G_\delta((\psi,\eta,w),\zeta)$ is the identity on the ζ-component. We denote the norm $\|(\psi,\eta,w,\zeta)\|_s^{Lip(\gamma)} := \max\{\|(\psi,\eta,w)\|_s^{Lip(\gamma)}, |\zeta|^{Lip(\gamma)}\}$.

Theorem 4.4.11. *(Theorem 6.10 in [8]) Assume (4.4.6) and the inversion assumption (4.4.29). Then there exists $\mu := \mu(\tau,\nu)$ such that, for all $\omega \in \Omega_\infty$, for all $g := (g_1, g_2, g_3)$, the operator $\mathbf{T}_0$ defined in (4.4.36) satisfies*

$$\|\mathbf{T}_0 g\|_s^{Lip(\gamma)} \le_s \gamma^{-1}(\|g\|_{s+\mu}^{Lip(\gamma)} + \varepsilon\gamma^{-1}\{\|\mathfrak{I}_0\|_{s+\mu}^{Lip(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\mu}^{Lip(\gamma)}\|\mathcal{F}(i_0,\zeta_0)\|_{s+\mu}^{Lip(\gamma)}\}\|g\|_{s_0+\mu}^{Lip(\gamma)}). \tag{4.4.37}$$

It is an approximate inverse of $d_{i,\zeta}\mathcal{F}(i_0)$, namely

$$\begin{aligned}\|(d_{i,\zeta}\mathcal{F}(i_0) \circ \mathbf{T}_0 - \mathrm{I})g\|_s^{Lip(\gamma)} \le_s \gamma^{-1}\Big(&\|\mathcal{F}(i_0,\zeta_0)\|_{s_0+\mu}^{Lip(\gamma)}\|g\|_{s+\mu}^{Lip(\gamma)} \\ &+ \{\|\mathcal{F}(i_0,\zeta_0)\|_{s+\mu}^{Lip(\gamma)} + \varepsilon\gamma^{-1}\|\mathcal{F}(i_0,\zeta_0)\|_{s_0+\mu}^{Lip(\gamma)}\|\mathfrak{I}_0\|_{s+\mu}^{Lip(\gamma)}\}\|g\|_{s_0+\mu}^{Lip(\gamma)}\Big).\end{aligned} \tag{4.4.38}$$

4.5 The linearized operator in the normal directions

In this section we give an explicit expression of the linearized operator

$$\mathcal{L}_\omega := \omega\cdot\partial_\varphi - \partial_x K_{02}(\varphi). \tag{4.5.1}$$

To this aim we compute $\frac{1}{2}(K_{02}(\psi)w, w)_{L^2(\mathbb{T})}, w \in H_S^\perp$, which collects all the terms of $(H_\varepsilon \circ G_\delta)(\psi, 0, w)$ that are quadratic in w.
First we recall some preliminary lemmata.

Lemma 4.5.1. *(Lemma 7.1 in [8]) Let H be a Hamiltonian function of class $C^2(H_0^1(\mathbb{T}_x),\mathbb{R})$ and consider a map $\Phi(u) := u + \Psi(u)$ satisfying $\Psi(u) = \Pi_E\Psi(\Pi_E u)$, for all u, where E is a finite dimensional subspace as in* (4.1.2)*. Then*

$$\partial_u[\nabla(H\circ\Phi)](u)[h] = (\partial_u\nabla H)(\Phi(u))[h] + \mathcal{R}(u)[h], \tag{4.5.2}$$

where $\mathcal{R}(u)$ has the "finite dimensional" form

$$\mathcal{R}(u)[h] = \sum_{|j|\le C}(h, g_j(u))_{L^2(\mathbb{T})}\chi_j(u) \tag{4.5.3}$$

with $\chi_j(u) = e^{\mathrm{i}jx}$ or $g_j(u) = e^{\mathrm{i}jx}$. The remainder in (4.5.3) *is*

$$\mathcal{R}(u) = \mathcal{R}_0(u) + \mathcal{R}_1(u) + \mathcal{R}_2(u)$$

with

$$\begin{aligned}&\mathcal{R}_0(u) := (\partial_u\nabla H)(\Phi(u))\partial_u\Psi(u), \quad \mathcal{R}_1(u) := [\partial_u\{\Psi'(u)^T\}][\cdot,\nabla H(\Phi(u))],\\ &\mathcal{R}_2(u) := [\partial_u\Psi(u)]^T(\partial_u\nabla H)(\Phi(u))\partial_u\Phi(u).\end{aligned} \tag{4.5.4}$$

Lemma 4.5.2. *(Lemma 7.3 in [8]) Let $\mathcal{R}$ be an operator of the form*

$$\mathcal{R}h = \sum_{|j|\le C}\int_0^1 (h, g_j(\tau))_{L^2(\mathbb{T})}\chi_j(\tau)\,d\tau, \tag{4.5.5}$$

where the functions $g_j(\tau),\chi_j(\tau)\in H^s, \tau\in[0,1]$ depend in a Lipschitz way on the parameter ω. Then its matrix s-decay norm (see (2.3.1)-(2.3.2)*) satisfies*

$$|\mathcal{R}|_s^{Lip(\gamma)} \le_s \sum_{|j|\le C}\sup_{\tau\in[0,1]}(\|\chi_j(\tau)\|_s^{Lip(\gamma)}\|g_j\|_{s_0}^{Lip(\gamma)} + \|\chi_j(\tau)\|_{s_0}^{Lip(\gamma)}\|g_j(\tau)\|_s^{Lip(\gamma)}). \tag{4.5.6}$$

4.5.1 Composition with the map G_δ

In the sequel we use the fact that $\mathfrak{I}_\delta := \mathfrak{I}_\delta(\varphi;\omega) = i_\delta(\varphi;\,\omega) - (\varphi,\,0,\,0)$ satisfies

$$\|\mathfrak{I}_\delta\|_{s_0+\mu}^{Lip(\gamma)} \le C\,\varepsilon^{6-2b}\gamma^{-1}. \tag{4.5.7}$$

We now study the Hamiltonian $K := H_\varepsilon\circ G_\delta = \varepsilon^{-2b}\mathcal{H}\circ A_\varepsilon\circ G_\delta$ (see (4.2.19)). Recalling (4.2.15), $A_\varepsilon\circ G_\delta$ has the form

$$A_\varepsilon(G_\delta(\psi,\eta,w)) = \varepsilon v_\varepsilon(\theta_0(\psi), y_\delta(\psi) + L_1(\psi)\eta + L_2(\psi)w) + \varepsilon^b(z_0(\psi)+w) \tag{4.5.8}$$

where

$$L_1(\Psi) := [\partial_\psi\theta_0(\psi)]^{-T}, \quad L_2(\psi) := [(\partial_\theta\tilde{z}_0)(\theta_0(\psi))]^T\partial_x^{-1}. \tag{4.5.9}$$

By Taylor formula, we develop (4.5.8) in w at $(\eta,w)=(0,0)$, and we get

$$(A_\varepsilon\circ G_\delta)(\psi,0,w) = T_\delta(\psi) + T_1(\psi)w + T_2(\psi)[w,w] + T_{\ge 3}(\psi,w),$$

where

$$T_\delta(\psi) := A_\varepsilon(G_\delta(\psi,0,0)) = \varepsilon v_\delta(\psi) + \varepsilon^b z_0(\psi), \quad v_\delta(\psi) := v_\varepsilon(\theta_0(\psi), y_\delta(\psi)) \tag{4.5.10}$$

is the approximate isotropic torus in the phase space $H^1_0(\mathbb{T})$ (it corresponds to i_δ),

$$T_1(\psi)w := \varepsilon^{2b-1}U_1(\psi)w + \varepsilon^b w; \quad T_2(\psi)[w,w] := \varepsilon^{4b-3}U_2(\psi)[w,w] \tag{4.5.11}$$

$$U_1(\psi)w := \varepsilon \sum_{j\in S} \frac{|j|\,[L_2(\psi)w]_j\, e^{\mathrm{i}[\theta_0(\psi)]_j}}{2\sqrt{|j|}\sqrt{\xi_j + \varepsilon^{2(b-1)}[y_\delta(\psi)]_j}}, \tag{4.5.12}$$

$$U_2(\psi)[w,w] := -\varepsilon \sum_{j\in S} \frac{j^2\,[L_2(\psi)w]_j^2\, e^{\mathrm{i}[\theta_0(\psi)]_j}}{8|j|^{\frac32}\{\xi_j + \varepsilon^{2(b-1)}[y_\delta(\psi)]_j\}^{\frac32}}, \tag{4.5.13}$$

and $T_{\geq 3}(\psi,w)$ collects all the terms of order at least cubic in w. In the notation of (4.2.15), the function $v_\delta(\Psi)$ in (4.5.10) is $v_\delta(\psi) = v_\varepsilon(\theta_0(\psi), y_\delta(\psi))$. The terms U_1, U_2 in (4.5.12), (4.5.13) are $O(1)$ in ε. Moreover, using that $L_2(\psi)$ in (4.5.9) vanishes at $z_0 = 0$, they satisfy

$$\begin{aligned} &\|U_1 w\|_s \leq_s \|\mathfrak{I}_\delta\|_s \|w\|_{s_0} + \|\mathfrak{I}_\delta\|_{s_0}\|w\|_s, \\ &\|U_2[w,w]\|_s \leq_s \|\mathfrak{I}_\delta\|_s \|\mathfrak{I}_\delta\|_{s_0}\|w\|_{s_0}^2 + \|\mathfrak{I}_\delta\|_{s_0}^2 \|w\|_{s_0}\|w\|_s \end{aligned} \tag{4.5.14}$$

and also in the norm $\|\cdot\|_s^{Lip(\gamma)}$. We expand $\mathcal{H}$ by Taylor formula

$$\mathcal{H}(u+h) = \mathcal{H}(u) + ((\nabla\mathcal{H})(u), h)_{L^2(\mathbb{T})} + \frac12((\partial_u\nabla\mathcal{H})(u)[h], h)_{L^2(\mathbb{T})} + O(h^3). \tag{4.5.15}$$

Specifying at $u = T_\delta(\psi)$ and $h = T_1(\psi)w + T_2(\psi)[w,w] + T_{\geq 3}(\psi,w)$, we obtain that the sum of all components of $K = \varepsilon^{-2b}(\mathcal{H}\circ A_\varepsilon \circ G_\delta)(\psi,0,w)$ that are quadratic in w is

$$\frac12 (K_{02}w,w)_{L^2(\mathbb{T})} = \varepsilon^{-2b}((\nabla\mathcal{H})(T_\delta), T_2[w,w])_{L^2(\mathbb{T})} + \frac{\varepsilon^{-2b}}{2}((\partial_u\nabla\mathcal{H})(T_\delta)[T_1 w], T_1 w)_{L^2(\mathbb{T})}. \tag{4.5.16}$$

Inserting the expressions (4.5.12), (4.5.13) in the equality (4.5.16), we get

$$\begin{aligned} K_{02}(\psi)w =& (\partial_u\nabla\mathcal{H})(T_\delta)[w] + 2\varepsilon^{b-1}(\partial_u\nabla\mathcal{H})(T_\delta)[U_1 w] + \\ &+ \varepsilon^{2(b-1)}U_1^T(\partial_u\nabla\mathcal{H})(T_\delta)[U_1 w] + 2\,\varepsilon^{2b-3}U_2[w,\cdot]^T(\nabla\mathcal{H})(T_\delta). \end{aligned} \tag{4.5.17}$$

Lemma 4.5.3. *The operator* K_{02} *reads*

$$(K_{02}w,w)_{L^2(\mathbb{T})} = ((\partial_u\nabla\mathcal{H})(T_\delta)[w], w)_{L^2(\mathbb{T})} + (R(\psi)w,w)_{L^2(\mathbb{T})} \tag{4.5.18}$$

where $R(\psi)$ *has the "finite dimensional" form*

$$R(\psi)w = \sum_{|j|\leq C} (w, g_j(\psi))_{L^2(\mathbb{T})}\,\chi_j(\psi). \tag{4.5.19}$$

The functions g_j, χ_j *satisfy, for some* $\sigma := \sigma(\nu,\tau) > 0$,

$$\|g_j\|_s^{Lip(\gamma)}\|\chi_j\|_{s_0}^{Lip(\gamma)} + \|g_j\|_{s_0}^{Lip(\gamma)} \leq_s \varepsilon^{1+b}\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \tag{4.5.20}$$

$$\begin{aligned} &\|\partial_i g_j[\hat\imath]\|_s\|\chi_j\|_{s_0} + \|\partial_i g_j[\hat\imath]\|_{s_0}\|\chi_j\|_s + \|g_j\|_s\|\partial_i\chi_j[\hat\imath]\|_{s_0} + \|g_j\|_{s_0}\|\partial_i\chi_j[\hat\imath]\|_s \\ &\leq_s \varepsilon^{1+b}\|\hat\imath\|_{s+\sigma} + \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s+\sigma} \end{aligned} \tag{4.5.21}$$

In conclusion, the linearized operator to analyze after the composition with the action-angle variables, the rescaling and the transformation G_δ is

$$w \mapsto (\partial_u\nabla\mathcal{H})(T_\delta)[w], \quad w \in H_S^\perp$$

up to finite dimensional operators which have form (4.5.19) and size (4.5.20).

4.5.2 The linearized operator in the normal directions

In this section we compute $((\partial_u \nabla \mathcal{H})(T_\delta)[w], w)_{L^2(\mathbb{T})}, w \in H_S^\perp$, recalling that $\mathcal{H} = H \circ \Phi_B$ and Φ_B is the Birkhoff map of Proposition 4.1.1. It is convenient to write separately the terms in

$$\mathcal{H} = H \circ \Phi_B = (H^{(2)} + H^{(3)}) \circ \Phi_B + H^{(4)} \circ \Phi_B + H^{(\geq 5)} \circ \Phi_B, \tag{4.5.22}$$

where $H^{(2)}, H^{(3)}, H^{(4)}, H^{(\geq 5)}$ are defined in (4.1.1). First we consider $H^{(\geq 5)} \circ \Phi_B$. By (4.1.1) we get

$$\nabla H^{(\geq 5)}(u) = \pi_0[(\partial_u f)(x, u, u_x)] - \partial_x\{(\partial_{u_x} f)(x, u, u_x)\}.$$

Since the Birkhoff transformation Φ_B has the form (4.1.6), Lemma 4.5.1 (at $u = T_\delta$) implies that

$$\begin{aligned} \partial_u \nabla (H^{(\geq 5)} \circ \Phi_B)(T_\delta)[h] &= (\partial_u \nabla H^{(\geq 5)})(\Phi_B(T_\delta))[h] + \mathcal{R}_{H^{(\geq 5)}}(T_\delta)[h] = \\ &= \partial_x (r_1(T_\delta)\, \partial_x h) + r_0(T_\delta) h + \mathcal{R}_{H^{(\geq 5)}}(T_\delta)[h] \end{aligned} \tag{4.5.23}$$

where the multiplicative functions $r_0(T_\delta), r_1(T_\delta)$ are

$$r_0(T_\delta) := \sigma_0(\Phi_B(T_\delta)), \quad \sigma_0(u) := (\partial_{uu} f)(x, u, u_x) - \partial_x\{(\partial_{u u_x} f)(x, u, u_x)\}, \tag{4.5.24}$$

$$r_1(T_\delta) := \sigma_1(\Phi_B(T_\delta)), \quad \sigma_1(u) := -(\partial_{u_x u_x} f)(x, u, u_x), \tag{4.5.25}$$

the remainder $\mathcal{R}_{H^{(\geq 5)}}(u)$ has the form (4.5.3) with $\chi_j = e^{\mathrm{i} j x}$ or $g_j = e^{\mathrm{i} j x}$ and it satisfies, for some $\sigma := \sigma(\nu, \tau) > 0$,

$$\|g_j\|_s^{Lip(\gamma)} \|\chi_j\|_{s_0}^{Lip(\gamma)} + \|g_j\|_{s_0}^{Lip(\gamma)} \leq_s \varepsilon^4 (1 + \|\mathfrak{I}_\delta\|_{s+2}^{Lip(\gamma)}),$$

$$\begin{aligned} &\|\partial_i g_j[\hat\imath]\|_s \|\chi_j\|_{s_0} + \|\partial_i g_j[\hat\imath]\|_{s_0} \|\chi_j\|_s + \|g_j\|_s \|\partial_i \chi_j[\hat\imath]\|_{s_0} + \|g_j\|_{s_0} \|\partial_i \chi_j[\hat\imath]\|_s \\ &\leq_s \varepsilon^4 (\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+2} \|\hat\imath\|_{s_0+2}). \end{aligned}$$

Now consider the contribution of $(H^{(2)} + H^{(3)} + H^{(4)}) \circ \Phi_B$. By Lemma 4.5.1 and (4.1.1) we have

$$\begin{aligned} &\partial_u \nabla ((H^{(2)} + H^{(3)} + H^{(4)}) \circ \Phi_B)(T_\delta)[h] = -h_{xx} - 6\, c_1\, \partial_x[\Phi_B(T_\delta)_x\, h_x] - 2\, c_2\, \partial_{xx}(\Phi_B(T_\delta) h) \\ &+ 2\, c_2\, \Phi_B(T_\delta)_x\, h_x + 6\, c_3\, \Phi_B(T_\delta)\, h - 12\, c_4\, \partial_x[(\Phi_B(T_\delta))_x^2\, h_x] - 3\, c_5\, \partial_x[(\Phi_B(T_\delta))_x^2\, h] \\ &+ 3\, c_5\, (\Phi_B(T_\delta))_x^2\, h_x - 2\, c_6\, \partial_x[\Phi_B(T_\delta)^2\, h_x] - 2\, c_6\, \partial_{xx}(\Phi_B(T_\delta)^2)\, h + 2\, c_6\, \Phi_B(T_\delta)_x^2\, h \\ &+ 12\, c_7\, \Phi_B(T_\delta)^2\, h + \mathcal{R}_{H^{(2)}}(T_\delta) + \mathcal{R}_{H^{(3)}}(T_\delta) + \mathcal{R}_{H^{(4)}}(T_\delta)[h], \end{aligned} \tag{4.5.26}$$

where $\Phi_B(T_\delta)$ is a zero space average function, indeed Φ_B maps $H_0^1(\mathbb{T}_x)$ in itself by Proposition (4.1.1). The remainder $\mathcal{R}_{H^{(2)}}, \mathcal{R}_{H^{(3)}}, \mathcal{R}_{H^{(4)}}$ have the form (4.5.3) and, by (4.5.4), the size $(\mathcal{R}_{H^{(2)}} + \mathcal{R}_{H^{(3)}} + \mathcal{R}_{H^{(4)}})(T_\delta) = O(\varepsilon)$. We develop this sum as

$$(\mathcal{R}_{H^{(2)}} + \mathcal{R}_{H^{(3)}} + \mathcal{R}_{H^{(4)}})(T_\delta) = \varepsilon \mathcal{R}_1 + \varepsilon^2 \mathcal{R}_2 + \tilde{\mathcal{R}}_{>2}, \tag{4.5.27}$$

where $\tilde{\mathcal{R}}_{>2}$ has size $o(\varepsilon^2)$. Thus we get, for all $h \in H_S^\perp$,

$$\begin{aligned} &\Pi_S^\perp \partial_u \nabla ((H^{(2)} + H^{(3)} + H^{(4)}) \circ \Phi_B)(T_\delta)[h] = -h_{xx} + \Pi_S^\perp \{-6\, c_1\, \partial_x[\Phi_B(T_\delta)_x\, h_x] \\ &- 2\, c_2\, \partial_{xx}(\Phi_B(T_\delta) h) + 2\, c_2\, \Phi_B(T_\delta)_x\, h_x + 6\, c_3\, \Phi_B(T_\delta)\, h - 12\, c_4\, \partial_x[(\Phi_B(T_\delta))_x^2\, h_x] \\ &- 3\, c_5\, \partial_x[(\Phi_B(T_\delta))_x^2\, h] + 3\, c_5\, (\Phi_B(T_\delta))_x^2\, h_x - 2\, c_6\, \partial_x[\Phi_B(T_\delta)^2\, h_x] - 2\, c_6\, \partial_{xx}(\Phi_B(T_\delta)^2)\, h \\ &+ 2\, c_6\, \Phi_B(T_\delta)_x^2\, h + 12\, c_7\, \Phi_B(T_\delta)^2\, h\} + \Pi_s^\perp (\varepsilon \mathcal{R}_1 + \varepsilon^2 \mathcal{R}_2 + \tilde{\mathcal{R}}_{>2})[h]. \end{aligned} \tag{4.5.28}$$

Now we expand $\Phi_B(u) = u+\Psi_2(u)+\Psi_{\geq 3}(u)$, where $\Psi_2(u)$ is a quadratic function of u, $\Psi_{\geq 3} = O(u^3)$ and both map $H^1_0(\mathbb{T}_x)$ in itself. At $u = T_\delta = \varepsilon v_\delta + \varepsilon^b z_0$ we get

$$\Phi_B(T_\delta) = T_\delta + \Psi_2(T_\delta) + \Psi_{\geq 3}(T_\delta) = \varepsilon v_\delta + \varepsilon^2 \Psi_2(v_\delta) + \tilde{q}, \tag{4.5.29}$$

where $\tilde{q} = \varepsilon^b z_0 + \Psi_2(T_\delta) - \Psi_2(v_\delta) + \Psi_{\geq 3}(T_\delta)$ and it satisfies

$$\|\tilde{q}\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 + \varepsilon^b \|\mathfrak{I}_\delta\|_s^{Lip(\gamma)}, \quad \|\partial_i \tilde{q}[\hat{\imath}]\|_s \leq_s \varepsilon^b(\|\hat{\imath}\|_s + \|\mathfrak{I}_\delta\|_s \|\hat{\imath}\|_{s_0}). \tag{4.5.30}$$

Note that also $\tilde{q}$ has zero space average, indeed $\tilde{q} = \Phi_B(T_\delta) - \varepsilon v_\delta - \varepsilon^2 \Psi_2(v_\delta)$ and the functions $\Phi_B(T_\delta), v_\delta, \Psi_2(v_\delta)$ belong to $H^1_0(\mathbb{T}_x)$.
We observe that the terms $O(\varepsilon)$ come from the monomials $R(v\, z^2)$ of $\mathcal{H}^{(3)}$ and the ones of size $O(\varepsilon^2)$ from $H^{(2)}+\mathcal{H}^{(4,2)}$ (see (4.1.8)). Thus, we compare (4.5.28) with $\Pi_S^\perp(\partial_u \nabla(H^{(2)}+\mathcal{H}^{(3)}+\mathcal{H}^{(4,2)}))(T_\delta)[h]$, using (4.1.8), and, by (4.5.29), we obtain $\mathcal{R}_1 = 0$,

$$\Psi_2(v_\delta) = -c_1\, \partial_x(v_\delta^2) - \frac{c_2}{3}\partial_{xx}[(\partial_x^{-1} v_\delta)^2] + \frac{c_2}{3}\pi_0[v_\delta^2] + c_3 \pi_0[(\partial_x^{-1} v_\delta)^2] \tag{4.5.31}$$

and

$$\begin{aligned}
\mathcal{R}_2[h] = & - 6c_1^2\{v_\delta \partial_{xx}(\Pi_S[(v_\delta)_x h_x]) - \partial_x((v_\delta)_x \partial_{xx}\Pi_S[v_\delta h])\} \\
& + 2c_1c_2\, v_\delta\, \partial_x(\Pi_S[(v_\delta)_x\, h_x]) + 2c_1c_2\, \partial_x((v_\delta)_x\, \partial_x \Pi_S[v_\delta\, h]) \\
& - 2c_1c_2\, (\partial_x^{-1} v_\delta)\, \partial_{xx}\Pi_S[(v_\delta)_x\, h_x] + 2c_1c_2\, \partial_x\{(v_\delta)_x\, \partial_{xx}\Pi_S[(\partial_x^{-1} v_\delta)\, h]\} \\
& - \frac{2\,c_2^2}{3}\, (\partial_x^{-1} v_\delta)\, \partial_{xxx}\Pi_S[v_\delta\, h] + \frac{2\,c_2^2}{3}\, (\partial_x^{-1} v_\delta)\, \partial_x \Pi_S[(v_\delta)_x\, h_x] \\
& + \frac{2\,c_2^2}{3}\, \partial_x\{(v_\delta)_x\, \partial_x \Pi_S[(\partial_x^{-1} v_\delta)\, h] + 2c_2c_3\, (\partial_x^{-1} v_\delta)\, \partial_x \Pi_S[v_\delta\, h] \\
& - 2c_2c_3\, v_\delta\, \partial_x \Pi_S[(\partial_x^{-1} v_\delta) h] + 2c_1c_2\, \partial_x^{-1}\{v_\delta\, \partial_{xx}\Pi_S[(v_\delta)_x\, h_x]\} \\
& + 2c_1c_2\, \partial_x\{(v_\delta)_x\, \partial_{xx}\Pi_S[v_\delta\, (\partial_x^{-1} h)]\} + \frac{2\,c_2^2}{3}\, \partial_x^{-1}\{v_\delta\, \partial_{xxx}\Pi_S[v_\delta h]\} \\
& + \frac{2\,c_2^2}{3}\, v_\delta\, \partial_{xxx}\Pi_S[v_\delta\, (\partial_x^{-1} h)] - \frac{2\,c_2^2}{3}\, (\partial_x^{-1}\{v_\delta)\, \partial_x \Pi_S[(v_\delta)_x h_x]\} \\
& + \frac{2\,c_2^2}{3}\, \partial_x\{(v_\delta)_x\, \partial_x \Pi_S[v_\delta (\partial_x^{-1} h)] - 2c_2c_3\, \partial_x^{-1}\{v_\delta\, \partial_x \Pi_S[v_\delta\, h]\} \\
& - 2c_2c_3 v_\delta\, \partial_x \Pi_S[v_\delta\, (\partial_x^{-1} h)] - 2c_1c_2\, v_\delta\, \partial_x \Pi_S[(v_\delta)_x\, h_x] \\
& - 2c_1c_2\, \partial_x\{(v_\delta)_x\, \partial_x \Pi_S[v_\delta\, h]\} - \frac{4\,c_2^2}{3}\, v_\delta\, \partial_{xx}\Pi_S[v_\delta\, h] \\
& + \frac{2\,c_2^2}{3}\, v_\delta\, \Pi_S[(v_\delta)_x\, h_x] - \frac{2\,c_2^2}{3}\, \partial_x\{(v_\delta)_x\, \Pi_S[v_\delta\, h]\} \\
& + 4c_2c_3\, v_\delta\, \Pi_S[v_\delta\, h] + 6c_1c_3\, \partial_x^{-1}\{(\partial_x^{-1} v_\delta)\, \partial_x \Pi_S[(v_\delta)_x\, h_x]\} \\
& - 6c_1c_3\, \partial_x\{(v_\delta)_x\, \partial_x \Pi_S[(\partial_x^{-1} v_\delta)(\partial_x^{-1} h)]\} + 2c_2c_3 \partial_x^{-1}\{(\partial_x^{-1} v_\delta)\, \partial_{xx}\Pi_S[v_\delta h]\} \\
& - 2c_2c_3\, v_\delta\, \partial_{xx}\Pi_S[(\partial_x^{-1} v_\delta)(\partial_x^{-1} h)] - 2c_2c_3\, \partial_x^{-1}\{(\partial_x^{-1} v_\delta)\, \Pi_S[(v_\delta)_x h_x]\} \\
& - 2c_2c_3\, \partial_x\{(v_\delta)_x\, \Pi_S[(\partial_x^{-1} v_\delta)(\partial_x^{-1} h)]\} - 6\, c_3^2\, \partial_x^{-1}\{(\partial_x^{-1} v_\delta)\, \Pi_S[v_\delta\, h]\} \\
& + 6c_3^2\, v_\delta\, \Pi_S[(\partial_x^{-1} v_\delta)(\partial_x^{-1} h)] + \frac{2}{3} c_2^2\, v_\delta\, \partial_{xxx}\Pi_S[(\partial_x^{-1} v_\delta)\, h].
\end{aligned} \tag{4.5.32}$$

In conclusion, we have the following proposition.

Proposition 4.5.4. *Assume* (4.5.7). *Then the Hamiltonian operator* $\mathcal{L}_\omega$, $\forall h \in H^s_{S^\perp}(\mathbb{T}^{\nu+1})$, *has the form*

$$\mathcal{L}_\omega h := \omega \cdot \partial_\varphi h - \partial_x K_{02} h = \Pi_S^\perp(\omega \cdot \partial_\varphi h + \partial_{xx}(a_1 h_x) + \partial_x(a_0 h) - \varepsilon^2 \partial_x \mathcal{R}_2 h - \partial_x \mathcal{R}_* h) \tag{4.5.33}$$

where $\mathcal{R}_2$ *is defined in* (4.5.32),

$$R_* := \tilde{\mathcal{R}}_{>2} + R_{H^{(\geq 5)}}(T_\delta) + R(\psi), \tag{4.5.34}$$

with $R(\psi)$ *defined in Lemma 4.5.3, the functions*

$$\begin{aligned} a_1 :=& 1 + 6c_1 (\Phi_B(T_\delta))_x + 2 c_2 \Phi_B(T_\delta) + 12 c_4 (\Phi_B(T_\delta))_x^2 + 3c_5 \partial_x[\Phi_B(T_\delta)^2] + \\ &+ 2c_6 \Phi_B(T_\delta)^2 - r_1(T_\delta), \end{aligned} \tag{4.5.35}$$

$$\begin{aligned} a_0 :=& 2c_2 (\Phi_B(T_\delta))_{xx} - 6c_3 \Phi_B(T_\delta) + 3c_5 \partial_x[(\Phi_B(T_\delta))_x^2] + 2c_6 \{\Phi_B(T_\delta)_x^2 + \\ &+ 2\Phi_B(T_\delta) (\Phi_B(T_\delta))_{xx}\} - 12 c_7 \Phi_B(T_\delta)^2 - r_0(T_\delta) \end{aligned} \tag{4.5.36}$$

the function r_1 *is defined in* (4.5.25), r_0 *in* (4.5.24), T_δ *and* v_δ *in* (4.5.10).
Furthermore, we have, for some $\sigma := \sigma(\nu, \tau) > 0$,

$$\|a_1 - 1\|_s^{Lip(\gamma)} \leq_s \varepsilon (1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}), \qquad \|\partial_i a_1[\hat{\imath}]\|_s \leq_s \varepsilon(\|i\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|i\|_{s_0+\sigma}), \tag{4.5.37}$$

$$\|a_0\|_s^{Lip(\gamma)} \leq_s \varepsilon (1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}), \qquad \|\partial_i a_0[\hat{\imath}]\|_s \leq_s \varepsilon(\|i\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|i\|_{s_0+\sigma}), \tag{4.5.38}$$

where $\mathfrak{I}_\delta(\varphi) := (\theta_0(\varphi) - \varphi, y_\delta(\varphi), z_0(\varphi))$ *corresponds to* T_δ. *The remainder* $\mathcal{R}_2$ *has the form* (4.5.3) *with*

$$\begin{aligned} &\|g_j\|_s^{Lip(\gamma)} + \|\chi_j\|_s^{Lip(\gamma)} \leq_s 1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \\ &\|\partial_i g_j[\hat{\imath}]\|_s + \|\partial_i \chi_j[\hat{\imath}]\|_s \leq_s \|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma} \end{aligned} \tag{4.5.39}$$

and also $\mathcal{R}_*$ *has the form* (4.5.3) *with*

$$\|g_j^*\|_s^{Lip(\gamma)}\|\chi_j^*\|_{s_0}^{Lip(\gamma)} + \|g_j^*\|_{s_0}^{Lip(\gamma)}\|\chi_j^*\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 + \varepsilon^{1+b}\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \tag{4.5.40}$$

$$\begin{aligned} &\|\partial_i g_j^*[\hat{\imath}]\|_s\|\chi_j^*\|_{s_0} + \|\partial_i g_j^*[\hat{\imath}]\|_{s_0}\|\chi_j^*\|_s + \|g_j^*\|_{s_0}\|\partial_i \chi_j^*\|_s + \|g_j^*\|_s\|\partial_i \chi_j^*\|_{s_0} \\ &\leq_s \varepsilon^{1+b}\|\hat{\imath}\|_{s+\sigma} + \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}. \end{aligned} \tag{4.5.41}$$

The bounds (4.5.39) and (4.5.40) imply, by Lemma 4.5.2, estimates for the s-decay norms of $\mathcal{R}_2$ and $\mathcal{R}_*$.
The linearized operator $\mathcal{L}_\omega := \mathcal{L}_\omega(\omega, i_\delta(\omega))$ depends on the parameter ω both directly and also through the dependence on the embedded torus $i_\delta(\omega)$. The estimates on the partial derivative respect to i (see (4.3.1)) allow us to control, along the Nash-Moser iteration, the Lipschitz variation of the eigenvalues of $\mathcal{L}_\omega$ with respect to ω and the approximate solution i_δ.

4.6 Reduction of the linearized operator in the normal directions

The goal of this section is to conjugate the Hamiltonian linear operator $\mathcal{L}_\omega$ in (4.5.33) to a constant coefficients linear operator $\mathcal{L}_\infty$. For this purpose, we shall apply the same kind of symplectic transformations used in [8], whose aim is to diagonalize the operator $\mathcal{L}_\omega$ up to a bounded remainder

$\mathcal{R}_6$ (see (4.6.128)). This one has to satisfy the smallness condition (4.6.132) in order to initialize the KAM reducibility scheme of Theorem 4.6.19, that completes the diagonalization procedure.
The size of all these transformations will be greater than the ones used in [8] (see Section 8 in [8]) and, as a consequence, some non perturbative terms will be modified by them. Thus, in order to prove (4.6.132) we will have to overcome two main difficulties: (a) computing the terms of order ε and ε^2 after each transformation, since we need to normalize them through the Birkhoff steps of Section 8.5 and 8.6, (b) providing optimal estimates for the transformations and, consequently, for the remainder $\mathcal{R}_6$ (see (4.6.128)).

Consider

$$\overline{v}(\varphi, x) := \sum_{j\in S} \sqrt{|j|\xi_j}\, e^{\mathrm{i}\mathtt{l}(j)\cdot\varphi}\, e^{\mathrm{i}jx} \tag{4.6.1}$$

and $\mathtt{l}\colon S \to \mathbb{Z}^\nu$ is the odd injective map

$$\mathtt{l}\colon S \to \mathbb{Z}^\nu, \quad \mathtt{l}(\overline{\jmath}_i) := \mathtt{e}_i, \quad \mathtt{l}(-\overline{\jmath}_i) = -\mathtt{l}(\overline{\jmath}_i) = -\mathtt{e}_i, \quad i = 1, \ldots, \nu, \tag{4.6.2}$$

denoting by $\mathtt{e}_i = (0, \ldots, 1, \ldots, 0)$ the i-th vector of the canonical basis of $\mathbb{R}^\nu$. We observe that

$$\|v_\delta - \overline{v}\|_s^{Lip(\gamma)} \le_s \|\mathfrak{I}_\delta\|_s^{Lip(\gamma)}, \qquad \|\partial_i(v_\delta - \overline{v})[\hat{\imath}]\|_s \le_s \|\hat{\imath}\|_s + \|\mathfrak{I}_\delta\|_s\|\hat{\imath}\|_{s_0}. \tag{4.6.3}$$

Remark 4.6.1. The function $\overline{v}(\varphi, x)$ in (4.6.1) corresponds to the torus $(\varphi, 0, 0)$ after the transformation A_ε defined in (4.2.15). In particular, this torus is invariant under the flow of the integrable Hamiltonian $\varepsilon^{-2b}\tilde{h} \circ A_\varepsilon$ (recalling (4.2.3)), which preserves the momentum. Hence, the square of the L^2 norm of $\overline{v}$ is independent of the time φ, as we can deduce by the properties of the map $\mathtt{l}$ defined in (4.6.2).
We shall expand the coefficients of the linearized operator at $y = z = 0$ to get the bounds on the transformations defined along this section, thus we will frequently use the inequalities (4.6.3) and the assumption (4.5.7). Moreover, we will use the fact that $\overline{v}$ satisfies the equation $L_{\overline{\omega}} = 0$, where $\overline{\omega}$ is the vector of the linear frequencies (see (1.1.6)) and $L_\omega := \omega \cdot \partial_\varphi + \partial_{xxx}$.

Remark 4.6.2. We recall that $\omega = \overline{\omega} + O(\varepsilon^2)$, see for instance (4.2.18). Moreover, note that $\mathcal{D}_\omega\overline{v} = \mathcal{D}_{\overline{\omega}}\overline{v} + \mathcal{D}_{\omega-\overline{\omega}}\overline{v}$ and

$$\mathcal{D}_{\omega-\overline{\omega}}\overline{v} = \sum_{j\in S} \mathrm{i}(\omega - \overline{\omega}) \cdot \mathtt{l}(j)\, \sqrt{|j|\xi_j}\, e^{\mathrm{i}\mathtt{l}(j)\cdot\varphi}\, e^{\mathrm{i}jx}.$$

Then $\|\mathcal{D}_{\omega-\overline{\omega}}\overline{v}\|_s^{Lip(\gamma)} \le C\varepsilon^2$ and $\mathcal{D}_{\omega-\overline{\omega}}\overline{v}$ has zero spatial average.

We expand in powers of ε the coefficients a_0 and a_1 in (4.5.36) and (4.5.35) as

$$a_0 = \varepsilon a_{0,1} + \varepsilon^2 a_{0,2} + \mathtt{R}_{a_0}, \qquad a_1 - 1 = \varepsilon a_{1,1} + \varepsilon^2 a_{1,2} + \mathtt{R}_{a_1}, \tag{4.6.4}$$

where

$$\begin{aligned}
a_{0,1} &:= 2c_2\, \overline{v}_{xx} - 6c_3\, \overline{v}, \qquad a_{1,1} := 6c_1\overline{v}_x + 2c_2\overline{v}_{xx},\\
a_{0,2} &:= 2c_2\, (\Psi_2(\overline{v}))_{xx} - 6c_3\, \Psi_2(\overline{v}) + 3c_5\partial_x(\overline{v}_x^2) + 2c_6\{\overline{v}_x^2 + 2\overline{v}\overline{v}_{xx}\} - 12c_7\overline{v}^2,\\
a_{1,2} &:= 6c_1(\Psi_2(\overline{v}))_x + 2c_2\Psi_2(\overline{v}) + 12c_4\overline{v}_x^2 + 3c_5\partial_x(\overline{v}^2) + 2c_6\overline{v}^2
\end{aligned}$$

and, by (4.6.3), $\|\mathtt{R}_{a_k}\|_s^{Lip(\gamma)} \le \varepsilon^3 + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}$, for some $\sigma > 0$.

4.6.1 Space reduction at the order ∂_{xxx}

First we conjugate $\mathcal{L}_\omega$ in (4.5.33) to an operator $\mathcal{L}_1$ whose coefficient in front of ∂_{xxx} is independent on the space variable x. Because of the Hamiltonian structure, the terms $O(\partial_{xx})$ will be simultaneously eliminated.
We look for a φ-dependent family of symplectic diffeomorphisms $\Phi(\varphi)$ of $H_S^\perp$ which differ from

$$\mathcal{A}_\perp := \Pi_S^\perp \mathcal{A} \Pi_S^\perp, \quad (\mathcal{A}h)(\varphi, x) := (1 + \beta_x(\varphi, x))\, h(\varphi, x + \beta(\varphi, x)), \tag{4.6.5}$$

up to a small "finite dimensional" remainder, see (4.6.9).
If $\|\beta\|_{W^{1,\infty}} < \frac{1}{2}$ then $\mathcal{A}$ is invertible and its inverse and adjoin map are

$$(\mathcal{A}^{-1}h)(\varphi, y) := (1 + \tilde{\beta}_y(\varphi, y))\, h(\varphi, y + \tilde{\beta}(\varphi, y)), \quad (\mathcal{A}^T h)(\varphi, y) = h(\varphi, y + \tilde{\beta}(\varphi, y)) \tag{4.6.6}$$

For each $\varphi \in \mathbb{T}^\nu$, $\mathcal{A}(\varphi)$ is a symplectic transformation of the phase space, see Remark 3.3 in [7], but the restricted map $\mathcal{A}_\perp(\varphi)$ is not.
In order to find a symplectic diffeomorphism near $\mathcal{A}_\perp$ first we observe that $\mathcal{A}_\perp$ is the time-1 flow map of the linear Hamiltonian PDE

$$\partial_\tau u = \partial_x(b(\varphi, \tau, x)u), \quad b(\varphi, \tau, x) := \frac{\beta(\varphi, x)}{1 + \tau\beta_x(\varphi, x)}. \tag{4.6.7}$$

The equation (4.6.7) is a linear transport equation, whose characteristic curves are the solutions of the ODE

$$\frac{d}{d\tau} x = -b(\varphi, \tau, x).$$

As in [8], we define a symplectic map Φ of $H_S^\perp$ as the time-1 flow of the Hamiltonian PDE

$$\partial_\tau u = \Pi_S^\perp \partial_x(b(\tau, x)u) = \partial_x(b(\tau, x)u) - \Pi_S \partial_x(b(\tau, x)u), \quad u \in H_S^\perp \tag{4.6.8}$$

generated by the quadratic Hamiltonian $\frac{1}{2}\int_{\mathbb{T}} b(\tau, x)u^2\, dx$ restricted to $H_S^\perp$. The flow of (4.6.8) is well defined in the Sobolev spaces $H^s_{S^\perp}(\mathbb{T}_x)$ for $b(\tau, x)$ smooth enough, by standard theory of linear hyperbolic PDE's. We obtained a symplectic diffeomorphism Φ that differs from $\mathcal{A}_\perp$ by a "finite dimensional" remainder of small size, more precisely, of size $O(\beta)$.

Lemma 4.6.3. *(Lemma 8.2 in [8]) For $\|\beta\|_{W^{s_0+1,\infty}}$ small, there exists an invertible symplectic transformation $\Phi = \mathcal{A}_\perp + \mathcal{R}_\Phi$ of $H^s_{S^\perp}$, where $\mathcal{A}_\perp$ is defined in (4.6.5) and $\mathcal{R}_\Phi$ is a "finite dimensional" remainder*

$$\mathcal{R}_\Phi h = \sum_{j \in S} \int_0^1 (h, g_j(\tau))_{L^2(\mathbb{T})} \chi_j(\tau)\, d\tau + \sum_{j \in S} (h, \psi_j)_{L^2(\mathbb{T})} e^{\mathrm{i}jx} \tag{4.6.9}$$

for some functions $\chi_j(\tau), g_j(\tau), \psi_j(\tau) \in H^s$ satisfying for all $\tau \in [0, 1]$

$$\|\psi_j\|_s + \|g_j(\tau)\|_s \leq_s \|\beta\|_{W^{s+2,\infty}}, \quad \|\chi_j(\tau)\|_s \leq_s 1 + \|\beta\|_{W^{s+1,\infty}}. \tag{4.6.10}$$

Moreover

$$\|\Phi h\|_s + \|\Phi^{-1} h\|_s \leq_s \|h\|_s + \|\beta\|_{W^{s+2,\infty}} \|h\|_{s_0} \quad \forall h \in H^s_{S^\perp}. \tag{4.6.11}$$

We conjugate $\mathcal{L}_\omega$ in (4.5.33) via the symplectic map $\Phi = \mathcal{A}_\perp + \mathcal{R}_\Phi$ of Lemma (4.6.3). Using the splitting $\Pi_S^\perp = \mathrm{I} - \Pi_S$, we compute

$$\mathcal{L}_\omega \Phi = \Phi \mathcal{D}_\omega + \Pi_S^\perp \mathcal{A}(b_3 \partial_{yyy} + b_2 \partial_{yy} + b_1 \partial_y + b_0)\Pi_S^\perp + \mathcal{R}_I, \tag{4.6.12}$$

where the coefficients are

$$b_3(\varphi, y) := \mathcal{A}^T[a_1 (1+\beta_x)^3] \qquad b_2(\varphi, y) := \mathcal{A}^T[2(a_1)_x(1+\beta_x)^2 + 6\, a_1\, \beta_{xx}(1+\beta_x)] \tag{4.6.13}$$

$$b_1(\varphi, y) := \mathcal{A}^T \left[(\mathcal{D}_\omega \beta) + 3\, a_1 \frac{\beta_{xx}^2}{1+\beta_x} + 4\, a_1\, \beta_{xxx} + 6\,(a_1)_x \beta_{xx} + (a_1)_{xx}(1+\beta_x) + a_0(1+\beta_x) \right] \tag{4.6.14}$$

$$b_0(\varphi, y) := \mathcal{A}^T \left[\frac{(\mathcal{D}_\omega \beta_x)}{1+\beta_x} + a_1 \frac{\beta_{xxxx}}{1+\beta_x} + 2(a_1)_x \frac{\beta_{xxx}}{1+\beta_x} + (a_1)_{xx} \frac{\beta_{xx}}{1+\beta_x} + a_0 \frac{\beta_{xx}}{1+\beta_x} + (a_0)_x \right] \tag{4.6.15}$$

and the remainder

$$\begin{aligned} \mathcal{R}_I := & - \Pi_S^\perp \partial_x (\varepsilon^2 \mathcal{R}_2 + \mathcal{R}_*)\, \mathcal{A}_\perp - \Pi_S^\perp (a_1 \partial_{xxx} + 2(a_1)_x \partial_{xx} + ((a_1)_{xx} + a_0)\partial_x + (a_0)_x) \Pi_S \mathcal{A} \Pi_S^\perp + \\ & + [\mathcal{D}_\omega, \mathcal{R}_\Phi] + (\mathcal{L}_\omega - \mathcal{D}_\omega)\mathcal{R}_\Phi. \end{aligned} \tag{4.6.16}$$

The commutator $[\mathcal{D}_\omega, \mathcal{R}_\Phi]$ has the form (4.6.9) with $\mathcal{D}_\omega g_j$ or $\mathcal{D}_\omega \chi_j, \mathcal{D}_\omega \psi_j$ instead of χ_j, g_j, ψ_j respectively. Also the last term $(\mathcal{L}_\omega - \mathcal{D}_\omega)\mathcal{R}_\Phi$ in (4.6.16) has the form (4.6.9) (note that $\mathcal{L}_\omega - \mathcal{D}_\omega$ does not contain derivatives with respect to φ). By (4.6.12), and decomposing $\mathrm{I} = \Pi_S + \Pi_S^\perp$, we get

$$\mathcal{L}_\omega \Phi = \Phi(\mathcal{D}_\omega + b_3 \partial_{yyy} + b_2 \partial_{yy} + b_1 \partial_y + b_0)\Pi_S^\perp + \mathcal{R}_{II}, \tag{4.6.17}$$

$$\mathcal{R}_{II} := \{\Pi_S^\perp (\mathcal{A} - \mathrm{I})\Pi_S - \mathcal{R}_\Phi\}(b_3 \partial_{yyy} + b_2 \partial_{yy} + b_1 \partial_y + b_0)\Pi_S^\perp + \mathcal{R}_I. \tag{4.6.18}$$

In order to solve the equation

$$b_3(\varphi, y) = b_3(\varphi)$$

for some function $b_3(\varphi)$, so that the coefficient in front of ∂_{xxx} depends only on φ, we choose the function $\beta = \beta(\varphi, x)$ such that

$$a_1(\varphi, x)(1 + \beta_x(\varphi, x))^3 = b_3(\varphi), \tag{4.6.19}$$

where we used that $\mathcal{A}^T[b_3(\varphi)] = b_3(\varphi)$. The only solution of (4.6.19) with zero space average is

$$\beta := \partial_x^{-1} \rho_0, \quad \rho_0 := b_3(\varphi)^{\frac{1}{3}} (a_1(\varphi, x))^{-\frac{1}{3}} - 1, \quad b_3(\varphi) := \left(\frac{1}{2\pi} \int_{\mathbb{T}} (a_1(\varphi, x))^{-\frac{1}{3}}\, dx \right)^{-3}. \tag{4.6.20}$$

Applying the symplectic map Φ^{-1} in (4.6.17) we obtain the Hamiltonian operator

$$\mathcal{L}_1 := \Phi^{-1} \mathcal{L}_\omega \Phi = \Pi_S^\perp (\omega \cdot \partial_\varphi + b_3(\varphi)\partial_{yyy} + b_1 \partial_y + b_0)\Pi_S^\perp + \mathfrak{R}_1 \tag{4.6.21}$$

where $\mathfrak{R}_1 := \Phi^{-1} \mathcal{R}_{II}$. We used that, by the Hamiltonian nature of $\mathcal{L}_1$, the coefficient $b_2 = 2\,(b_3)_y$ and so, by the choice (4.6.20), we have $b_2 = 2\,(b_3)_y = 0$.

Lemma 4.6.4. *(Lemma* 8.3 *in [8]) The operator* $\mathfrak{R}_1$ *in* (4.6.21) *has the form* (4.5.5).

In the proofs of the estimates for the transformations and the coefficients, we will always use the index σ to denote a certain loss of derivatives, since we do not need to know exactly the total amount of this loss. This, in fact, involves only the regularity required for the Hamiltonian nonlinearity $f(x,u,u_x)$ in (1.1.1).

Lemma 4.6.5. *There is $\sigma := \sigma(\tau,\nu) > 0$ such that, for $k = 0,1$,*

$$\|\beta\|_s^{Lip(\gamma)} \le_s \varepsilon\,(1+\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}) \qquad \|\partial_i\beta[\hat\imath]\|_s \le_s \varepsilon(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}) \tag{4.6.22}$$

$$\|b_3-1\|_s^{Lip(\gamma)} \le_s \varepsilon^2\,(1+\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}) \qquad \|\partial_i b_3[\hat\imath]\| \le_s \varepsilon^2(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}) \tag{4.6.23}$$

$$\|b_k\|_s^{Lip(\gamma)} \le_s \varepsilon(1+\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}) \qquad \|\partial_i b_k[\hat\imath]\|_s \le_s \varepsilon(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}). \tag{4.6.24}$$

The transformations Φ,Φ^{-1} satisfy

$$\|\Phi^{\pm1}h\|_s^{Lip(\gamma)} \le_s \|h\|_{s+1}^{Lip(\gamma)} + \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}\|h\|_{s_0+1}^{Lip(\gamma)} \tag{4.6.25}$$

$$\|\partial_i(\Phi^{\pm1}h)[\hat\imath]\|_s \le_s \|h\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma} + \|h\|_{s_0+\sigma}\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|h\|_{s_0+\sigma}\|\hat\imath\|_{s_0+\sigma}. \tag{4.6.26}$$

Moreover the remainder $\mathcal{R}_*$ has the form (4.5.5) *where the functions $\chi_j(\tau), g_j(\tau)$ satisfy the estimates* (4.5.40) *uniformly in $\tau\in[0,1]$.*

Proof. To shorten the notation we write $\|\cdot\|_s := \|\cdot\|_s^{Lip(\gamma)}$.
Estimate (4.6.23): Consider the functions $g(t) = (1+t)^{-\frac{1}{3}}$ and $\Upsilon(t) = (1+t)^{-3}$, analytic in a small neighbourhood of the origin. Then we have

$$b_3 - 1 = \Upsilon(M_x[g(a_1-1)-g(0)]) - \Upsilon(0). \tag{4.6.27}$$

By the mean value theorem, $\|b_3-1\|_s \le_s \|M_x[g(a_1-1)-g(0)]\|_s$. By Taylor expansion, we get

$$M_x[g(a_1-1)-g(0)] = g'(0)M_x[a_1-1] + \int_{\mathbb{T}}\int_0^1 (1-s)\,g''(s(a_1-1))\,(a_1-1)^2\,ds\,dx \tag{4.6.28}$$

and we note that, by Remark 4.6.4,

$$M_x[a_1-1] = \varepsilon^2 M_x[a_{1,2}] + M_x[\mathtt{R}_{a_1}].$$

Moreover, $\|M_x[\mathtt{R}_{a_1}]\|_s \le_s \varepsilon^3 + \varepsilon^{2b}\|\mathfrak{I}_\delta\|_{s+\sigma}$, because $M_x[v_\delta - \overline{v}] = M_x[\tilde q] = 0$ and $\mathtt{R}_{a_1}$ contains terms like $\varepsilon^2(v_\delta^2 - \overline{v}^2)$ and cubic in the x-derivatives of v_δ.
The second addend in the right hand side of (4.6.28) can be estimated by $\varepsilon^2(1+\|\mathfrak{I}_\delta\|_{s+\sigma})$. Hence

$$\|b_3-1\|_s \le_s \varepsilon^2(1+\|\mathfrak{I}_\delta\|_{s+\sigma}). \tag{4.6.29}$$

Now we consider the partial derivative respect to the variable i (see (4.3.1)) of b_3, namely

$$\partial_i b_3[\hat\imath] = \Upsilon'(M_x[g(a_1-1)-g(0)])\,M_x[g'(a_1-1)\,\partial_i a_1[\hat\imath]].$$

The derivatives of the functions g and Υ, for ε small enough, are approximately 1. Therefore, the estimate

$$\|\partial_i b_3[\hat\imath]\|_s \le_s \varepsilon^2(\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}) \tag{4.6.30}$$

derived from the estimate on $M_x[\partial_i a_1[\hat{\imath}]]$ and the fact that $M_x[\partial_i \overline{v}[\hat{\imath}]] = 0$. By (4.6.29) and (4.6.30) we conclude.

Estimate (4.6.22): Consider the functions $\phi(t) := (1+t)^{-1}$ and $g(t) := (1+t)^{-\frac{1}{3}}$. Recalling that $\beta_x = (b_3^{-1} a_1)^{\frac{1}{3}} - 1$, we have

$$\beta_x = g^{-1}(b_3^{-1}\, a_1 - 1) - g^{-1}(0) \quad \text{and} \quad b_3^{-1}\, a_1 - 1 = a_1\,(\phi(b_3 - 1) - \phi(0)) + (a_1 - 1).$$

Then, by (2.1.4),

$$\begin{aligned}\|\beta_x\|_s &\leq_s \|\phi(b_3 - 1) - \phi(0)\|_s \|a_1\|_{s_0} + \|\phi(b_3 - 1) - \phi(0)\|_{s_0} \|a_1\|_s + \|a_1 - 1\|_s \\ &\leq_s \|b_3 - 1\|_{s+\sigma} + \|b_3 - 1\|_{s_0+\sigma} \|a_1\|_s + \|a_1 - 1\|_s \leq_s \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\sigma}).\end{aligned}$$

Estimate (4.6.24): By (4.5.38), (4.5.37), (4.6.22) we get the estimates (4.6.24).
For the estimates (4.6.25), (4.6.26) on Φ, Φ^{-1} we apply Lemma 4.6.3 and the estimate (4.6.22) for β. We estimate the remainder $\mathcal{R}_*$ using (4.6.16), (4.6.18) and (4.5.40). □

4.6.2 Terms of order ε and ε^2

The diffeomorphism of the torus $\Phi = \mathcal{A}_\perp + \mathcal{R}_\Phi$ defined in Lemma 4.6.3 is, by (4.6.10) and (4.6.22), of the form $\mathrm{I} + O(\varepsilon)$, hence, the terms $O(\varepsilon^2)$ of $\mathcal{L}_\omega$ are modified by it.
From now on, the transformations we shall apply to reduce the linearized operator $\mathcal{L}_\omega$ to a constant coefficient operator will be $\mathrm{I} + O(\varepsilon^d)$ with $d > 1$, hence the terms of order $\varepsilon, \varepsilon^2$ will not be changed anymore.
In this section, our goal is to identify them in view of the linear Birkhoff steps of Section 8.5 and 8.6.

We have to put in evidence the terms $O(\varepsilon), O(\varepsilon^2)$ of b_0, b_1, b_3 in (4.6.21) and the ones in the remainder $\mathfrak{R}_1$ defined in (4.6.51).

Coefficients b_k

First, we note that $b_k = \mathcal{A}^T \alpha_k = \alpha_k + (\mathcal{A}^T - \mathrm{I})\alpha_k$, $k = 0, 1$, where

$$\alpha_1 := (\mathcal{D}_\omega \beta) + 3\, a_1 \frac{\beta_{xx}^2}{1+\beta_x} + 4\, a_1\, \beta_{xxx} + 6\,(a_1)_x \beta_{xx} + (a_1)_{xx}(1+\beta_x) + a_0(1+\beta_x), \tag{4.6.31}$$

$$\alpha_0 := \frac{(\mathcal{D}_\omega \beta_x)}{1+\beta_x} + a_1 \frac{\beta_{xxxx}}{1+\beta_x} + 2(a_1)_x \frac{\beta_{xxx}}{1+\beta_x} + (a_1)_{xx} \frac{\beta_{xx}}{1+\beta_x} + a_0 \frac{\beta_{xx}}{1+\beta_x} + (a_0)_x. \tag{4.6.32}$$

By (4.5.35), (4.6.20), we have

$$\begin{aligned}\beta = &-2\, c_1 \Phi_B(T_\delta) - \frac{2}{3}\, c_2 \partial_x^{-1}[\Phi_B(T_\delta)] - 4\, c_4 \partial_x^{-1}[\Phi_B(T_\delta)_x^2] - c_5\, \pi_0[\Phi_B(T_\delta)^2] \\ &- \frac{2}{3}\, c_6 \partial_x^{-1}[\Phi_B(T_\delta)^2] + 8\, c_1^2 \partial_x^{-1}[\Phi_B(T_\delta)_x^2] + \frac{8}{9}\, c_2^2 \partial_x^{-1}[\Phi_B(T_\delta)^2] + \frac{8}{3}\, c_1 c_2 \pi_0[\Phi_B(T_\delta)^2] + \mathtt{R}\end{aligned} \tag{4.6.33}$$

where, by (4.5.30), $\|\mathtt{R}\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 + \varepsilon^b \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}$. Then we write $\beta = \varepsilon\, \beta_1 + \varepsilon^2\, \beta_2 + \mathtt{R}_\beta$, where

$$\begin{aligned}\beta_1 :&= -2c_1 \overline{v} - \frac{2}{3} c_2 \partial_x^{-1}(\overline{v}), \\ \beta_2 :&= -2c_1 \Psi_2(\overline{v}) - \frac{2}{3} c_2 \partial_x^{-1}(\Psi_2(\overline{v})) - 4c_4 \partial_x^{-1}(\overline{v}_x^2) - c_5 \pi_0[\overline{v}^2] \\ &\quad - \frac{2}{3} c_6 \partial_x^{-1}[\overline{v}^2] + 8 c_1^2 \partial_x^{-1}[\overline{v}_x^2] + \frac{8}{9} c_2^2 \partial_x^{-1}[\overline{v}^2] + \frac{8}{3} c_1 c_2 \pi_0[\overline{v}^2]\end{aligned} \tag{4.6.34}$$

and $\mathtt{R}_\beta$ is defined by difference and satisfies, by (4.6.3),

$$\|\mathtt{R}_\beta\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad \|\partial_i \mathtt{R}_\beta[\hat{\imath}]\|_s \leq_s \varepsilon(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}).$$

Now we can develop α_0 and α_1 in powers of ε. By (4.5.35), (4.5.36), (4.6.33) and by Remark 4.6.1 we obtain $\alpha_1 := \varepsilon\alpha_{1,1} + \varepsilon^2\alpha_{1,2} + \mathtt{R}_1$ and $\alpha_0 = \varepsilon\alpha_{0,1} + \varepsilon^2\alpha_{0,2} + \mathtt{R}_0$, where

$$\begin{aligned} \alpha_{1,1} &= 2\,c_2\overline{v}_{xx} - 6\,c_3\overline{v}, \\ \alpha_{1,2} &= L_{\overline{\omega}}[\beta_2] + \frac{8}{3}(\beta_2)_{xxx} - \frac{41}{3}\,\partial_x[(\beta_1)_x\,(\beta_1)_{xx}] + a_{0,2} + a_{0,1}\,(\beta_1)_x, \end{aligned} \tag{4.6.35}$$

and

$$\begin{aligned} \alpha_{0,1} &= 2\,c_2\overline{v}_{xxx} - 6\,c_3\overline{v}_x, \\ \alpha_{0,2} &= \partial_x L_{\overline{\omega}}[\beta_2] - 3\partial_x[(\beta_1)_x\,(\beta_1)_{xxx}] - 3\partial_x[(\beta_1)_{xx}^2] + a_{0,1}(\beta_1)_{xx} + (a_{0,2})_x. \end{aligned} \tag{4.6.36}$$

The functions $\mathtt{R}_0$ and $\mathtt{R}_1$ are defined by difference and satisfy the following estimates

$$\|\mathtt{R}_k\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad \|\partial_i \mathtt{R}_k[\hat{\imath}]\|_s \leq_s \varepsilon(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}), \quad k = 0, 1. \tag{4.6.37}$$

Remark 4.6.6. We note that the terms $O(\varepsilon)$ generated by the Hamiltonian $\int_{\mathbb{T}}(3c_1 v_x + c_2 v)\, z_x^2\, dx$ (see (4.1.8)) are cancelled by the diffeomorphism of the torus Φ.

Remark 4.6.7. The averages of $\alpha_{j,k}, j = 0, 1$ for $k = 1$ are zero and, for $k = 2$, we have

$$M_x[\alpha_{1,2}] = M_x[a_{0,2}] + M_x[a_{0,1}\,(\beta_1)_x] = -2c_6\,M_x[\overline{v}_x^2] - 12c_7 M_x[\overline{v}^2] + \frac{4}{3}c_2^2 M_x[\overline{v}_x^2] + 4c_2c_3 M_x[\overline{v}^2],$$

$$M_x[\alpha_{0,2}] = M_x[a_{0,1}\,(\beta_1)_{xx}] = -4c_1c_2 M_x[\overline{v}_{xx}^2] - 12c_1c_3 M_x[\overline{v}_x^2].$$

We used the fact that $\partial_\varphi M_x[\overline{v}^2] = 0$, see Remark 4.6.1. Moreover, we note that, for a similar argument, $M_{\varphi,x}[\alpha_{k,2}] = M_x[\alpha_{k,2}]$, for $k = 0, 1$.

The transformation $\mathcal{A}^T - \mathrm{I}$ (see Section 8.1) is of order $O(\varepsilon)$, hence it generates new terms of order $O(\varepsilon^2)$ when it is applied to ones of order ε. In particular, by the regularity of the function $\overline{v}(\varphi, x)$, that is at least C^2, we have, for $k = 0, 1$, by Taylor expansion

$$\varepsilon(\mathcal{A}^T - \mathrm{I})\alpha_{k,1}(\varphi, y) = \varepsilon(\alpha_{k,1}(\varphi, y + \tilde{\beta}(\varphi, y)) - \alpha_{k,1}(\varphi, y)) = \varepsilon\partial_y(\alpha_{k,1})(\varphi, y)\,\tilde{\beta}(\varphi, y) + \mathtt{R}_{\tilde{\beta}},$$

where $\|\mathtt{R}_{\tilde{\beta}}\|_s \leq_s \varepsilon^3(1 + \|\mathfrak{I}_\delta\|_{s+\sigma})$ for some $\sigma > 0$.
We observe that $\tilde{\beta}(\varphi, y) = -(\mathcal{A}^T\,\beta)(\varphi, y)$ and by (4.6.34) we get, for $k = 0, 1$,

$$\varepsilon(\mathcal{A}^T - \mathrm{I})\alpha_{k,1}(\varphi, y) = -\varepsilon^2\,\partial_y(\alpha_{k,1})(\varphi, y)\,\beta_1(\varphi, y) + \mathtt{R}_{\tilde{\beta}}, \tag{4.6.38}$$

where we have renamed $\mathtt{R}_{\tilde{\beta}}$ the terms of order $o(\varepsilon^2)$.

Remainder $\mathfrak{R}_1$

The remaining terms of order ε^2 generated by the diffeomorphism of the torus Φ have the form (4.5.5) and originate from $\mathcal{R}_{II} = \Phi\mathfrak{R}_1$ (see (4.6.18)). Thus we analyze the expression

$$\begin{aligned} \mathcal{R}_{II} :&= \Pi_S^\perp(\mathcal{A} - \mathrm{I})\Pi_S[b_3\partial_{yyy} + b_1\partial_y + b_0] - R_\Phi(b_3\partial_{yyy} + b_1\partial_y + b_0) \\ &\quad - \Pi_S^\perp\partial_x(\varepsilon^2\mathcal{R}_2 + \mathcal{R}_*)\mathcal{A}_\perp - \Pi_S^\perp[\partial_{xx}(a_1\partial_x) + \partial_x(a_0\cdot)]\Pi_S\mathcal{A}\Pi_S^\perp + [\mathcal{D}_\omega, \mathcal{R}_\Phi] \\ &\quad + (\mathcal{L}_\omega - \mathcal{D}_\omega)\mathcal{R}_\Phi. \end{aligned} \tag{4.6.39}$$

We start from the first term in (4.6.39). As we said above, the transformation $\mathcal{A} - \mathrm{I}$ has size $O(\varepsilon)$. Hence, we look for the terms $O(\varepsilon)$ of $b_3\partial_{yyy} + b_1\partial_y + b_0$. We have, by (4.6.23), $b_3 = 1 + O(\varepsilon^2)$ and $b_k = \alpha_k + (\mathcal{A}^T - \mathrm{I})\alpha_k$ for $k = 0, 1$. Thus

$$b_3\partial_{yyy} + b_1\partial_y + b_0 = \partial_{yyy} + \varepsilon\partial_y(\alpha_{1,1}\,\cdot) + O(\varepsilon^2).$$

By Taylor expansion at the point $\beta = 0$, we get, for a function $u(\varphi, x)$

$$\begin{aligned}(\mathcal{A} - \mathrm{I})u(\varphi, x) &= (1+\beta_x)u(\varphi, x+\beta) - u(\varphi, x) = u(\varphi, x+\beta) - u(\varphi, x) + \beta_x u(\varphi, x+\beta)\\ &= u_x(\varphi, x)\beta(\varphi, x) + \beta_x(\varphi, x)u(\varphi, x) + O(\beta^2)\\ &= \varepsilon\partial_x(\beta_1(\varphi, x)\,u(\varphi, x)) + O(\varepsilon^2).\end{aligned} \tag{4.6.40}$$

Therefore we have

$$\Pi_S^{\perp}(\mathcal{A} - \mathrm{I})\Pi_S[b_3\partial_{yyy} + b_2\partial_{yy} + b_1\partial_y + b_0] = \varepsilon^2\Pi_S^{\perp}[\partial_x(\beta_1\,\partial_x(\alpha_{1,1}\,\cdot))] + o(\varepsilon^2) \tag{4.6.41}$$

Now we extract the homogeneous terms of order ε from $\mathcal{R}_\Phi$ (see (4.6.9)). We recall the exact expressions of g_k and χ_k in (4.6.9) refering to the proof of Lemma 8.2 in [8]. We have

$$g_k(\tau, x) := -(\Phi^\tau)^T[b(\tau)\partial_x e^{\mathrm{i}kx}], \tag{4.6.42}$$

where $(\Phi^\tau)^T$ is the flow of the adjoint PDE

$$\partial_\tau z = \Pi_S^{\perp}\{b(\tau, x)\partial_x z\}, \qquad b(\tau, x) = \frac{\beta(x)}{1 + \tau\beta_x(x)} = \varepsilon\beta_1 + O(\varepsilon^2). \tag{4.6.43}$$

This equation is well defined on $H^s_{S^\perp}(\mathbb{T}_x)$, because the function b is smooth enough. By (4.6.42) we have

$$g_k(\tau, x) = -b(\tau)\partial_x e^{\mathrm{i}kx} + (\mathrm{I}_{H_S^\perp} - (\Phi^\tau)^T)[b(\tau)\partial_x e^{\mathrm{i}kx}]$$

and, for $z \in H^s_{S^\perp}(\mathbb{T}_x)$, by (4.6.22) and (4.6.43), $\|(\Phi^\tau)^T z - z\|_s \leq_s \varepsilon C(\|z\|_{s+1} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|z\|_{s+1})$, where C is the Lipschitz constant, in time, on the interval $[0, 1]$ of the flow $(\Phi^\tau)^T$. Hence, by (4.6.42),

$$g_k = -\varepsilon\beta_1\,\partial_x e^{\mathrm{i}kx} + O(\varepsilon^2). \tag{4.6.44}$$

Now consider

$$\chi_k := -\frac{1+\beta_x}{1+\tau\beta_x}\ \exp(\mathrm{i}k\gamma^\tau(x + \beta(x))),$$

where γ^τ is the flow of the characteristic ODE

$$\frac{d}{d\tau}x = -b(\tau, x). \tag{4.6.45}$$

By (4.6.43), the vector field of (4.6.45) has size $O(\varepsilon)$ and, by similar arguments used above for the flow of (4.6.43), we have $\gamma^\tau(x) - x = O(\varepsilon)$. By Taylor expansion of the function $\exp(\mathrm{i}k\gamma^\tau(x+\beta(x)))$ at $\beta = 0$ we have

$$\chi_k = e^{\mathrm{i}kx} + O(\varepsilon). \tag{4.6.46}$$

Recalling (4.6.40) we have

$$\psi_k = (\mathcal{A}^T - \mathrm{I})e^{\mathrm{i}kx} = \varepsilon\partial_x(\beta_1 e^{\mathrm{i}kx}) + O(\varepsilon^2) = \varepsilon(\beta_1)_x e^{\mathrm{i}kx} + \varepsilon\beta_1\,\partial_x e^{\mathrm{i}kx} + O(\varepsilon^2). \tag{4.6.47}$$

Eventually, by (4.6.44), (4.6.46) and (4.6.47), we have $\mathcal{R}_\Phi = \varepsilon \mathrm{R}_\Phi + O(\varepsilon^2)$, where

$$\begin{aligned}\mathrm{R}_\Phi(h) : &= -\sum_{k\in S}(h, \beta_1\partial_x e^{\mathrm{i}kx})_{L^2(\mathbb{T})} e^{\mathrm{i}kx} + \sum_{k\in S}(h, (\beta_1)_x e^{\mathrm{i}kx})_{L^2(\mathbb{T})} e^{\mathrm{i}kx} + \sum_{k\in S}(h, \beta_1\partial_x e^{\mathrm{i}kx})_{L^2(\mathbb{T})} e^{\mathrm{i}kx} \\ &= \Pi_S[(\beta_1)_x\, h].\end{aligned} \tag{4.6.48}$$

By (4.6.48) the range of R_Φ is orthogonal to the subspace $H_S^\perp$, hence the term $\Phi^{-1} R_\Phi(b_3\partial_{yyy} + b_1\partial_y + b_0)$ will have size at least $O(\varepsilon^3)$, indeed $\Phi = \mathrm{I}_{H_S^\perp} + O(\varepsilon)$.
We ignore the terms $\varepsilon^2\mathcal{R}_2$ and $\mathcal{R}_*$ because are too small. Then, we can consider

$$(\mathcal{L}_\omega - \mathcal{D}_\omega)\mathrm{R}_\Phi = \Pi_S^\perp[\partial_{xx}(a_1\partial_x) + \partial_x(a_0\cdot)]\Pi_S^\perp \mathrm{R}_\Phi = 0.$$

By (4.6.40) we have

$$\Pi_S^\perp[\partial_{xx}(a_1\partial_x) + \partial_x(a_0\cdot)]\Pi_S(\mathcal{A} - \mathrm{I})\Pi_S^\perp = \varepsilon^2\Pi_S^\perp[\partial_{xx}(a_{1,1}\partial_{xx}\Pi_S[\beta_1\,\cdot]) + \partial_x(a_{0,1}\,\partial_x\Pi_S[\beta_1\,\cdot])] + o(\varepsilon^2). \tag{4.6.49}$$

It remains to study the commutator $[\mathcal{D}_\omega, \mathcal{R}_\Phi] = [D_{\overline{\omega}}, \mathcal{R}_\Phi^\varepsilon] + O(\varepsilon^3)$. We have

$$[D_{\overline{\omega}}, \mathcal{R}_\Phi^\varepsilon]h = \varepsilon D_{\overline{\omega}}\Pi_S[(\beta_1)_x h] - \varepsilon\Pi_S[(\beta_1)_x D_{\overline{\omega}} h] = \varepsilon\Pi_S[(D_{\overline{\omega}}(\beta_1)_x)h]$$

and so $\Phi^{-1}[\mathcal{D}_\omega, \mathcal{R}_\Phi] = o(\varepsilon^2)$.
Finally, by (4.6.41), (4.6.49), we obtained $\mathcal{R}_{II} = \varepsilon^2\mathrm{R}_2 + o(\varepsilon^2)$, where, for $h \in H_S^\perp$,

$$\begin{aligned}\mathrm{R}_2[h] &= \Pi_S^\perp\{\partial_x(\beta_1\,\Pi_S[\partial_x(\alpha_{1,1}h)]) - \partial_{xx}(a_{1,1}\partial_x\Pi_S[\partial_x(\beta_1 h)]) - \partial_x(\alpha_{1,1}\Pi_S[\partial_x(\beta_1 h)])\} \\ &= 4\,c_1c_2\,\Pi_S^\perp\{-\partial_x(v_\delta\,\partial_x\Pi_S[(v_\delta)_{xx}\,h]) + \partial_{xx}((v_\delta)_x\partial_{xx}\Pi_S[(\partial_x^{-1}v_\delta)h]) \\ &+ \partial_{xx}(v_\delta\,\partial_{xx}\Pi_S[v_\delta h]) + \partial_x((v_\delta)_{xx}\,\partial_x\Pi_S[v_\delta h])\} \\ &+ \frac{4}{3}\,c_2^2\,\Pi_S^\perp\{-\partial_x((\partial_x^{-1}v_\delta)\,\partial_x\Pi_S[(v_\delta)_{xx}\,h]) + \partial_{xx}(v_\delta\partial_{xx}\,\Pi_S[(\partial_x^{-1}v_\delta)h]) \\ &+ \partial_x((v_\delta)_{xx}\,\partial_x\Pi_S[(\partial_x^{-1}v_\delta)h])\} \\ &+ 12\,c_1c_3\,\Pi_S^\perp\{\partial_x(v_\delta\,\partial_x\Pi_S[v_\delta\,h]) - \partial_x(v_\delta\,\partial_x\Pi_S[v_\delta h])\} \\ &+ 4\,c_2c_3\,\Pi_S^\perp\{\partial_x((\partial_x^{-1}v_\delta)\,\partial_x\Pi_S[v_\delta\,h]) - \partial_x(v_\delta\,\partial_x\Pi_S[(\partial_x^{-1}v_\delta)h])\} \\ &+ 12c_1^2\Pi_S^\perp\{\partial_{xx}((v_\delta)_x\,\partial_{xx}\Pi_S[v_\delta h])\}\end{aligned} \tag{4.6.50}$$

Using (4.6.16), (4.6.18) we get

$$\mathfrak{R}_1 := \Phi^{-1}\mathcal{R}_{\mathrm{II}} = -\varepsilon^2\Pi_S^\perp\partial_x\mathcal{R}_2 + \mathcal{R}_* \tag{4.6.51}$$

where $\mathcal{R}_2$, defined in (4.5.32), has been renamed as

$$\mathcal{R}_2 := \mathcal{R}_2 - \partial_x^{-1}\mathrm{R}_2 \tag{4.6.52}$$

and we have renamed $\mathcal{R}_*$ the term $o(\varepsilon^2)$. Note that $\mathcal{R}_{II}^{\varepsilon^2}[h]$ has zero spatial average for every h belonging to $H_{S^\perp}^s(\mathbb{T}^{\nu+1})$ and the remainder $\mathcal{R}_*$ has the form (4.5.5).

4.6.3 Time reduction at the order ∂_{xxx}

The goal of this section is to make constant the coefficient of the highest order spatial derivative operator ∂_{yyy} by a quasi-periodic reparametrization of time. We consider the change of variable

$$(Bw)(\varphi, y) := w(\varphi + \omega\alpha(\varphi), y), \quad (B^{-1}h)(\vartheta, y) := h(\vartheta + \omega\tilde{\alpha}(\vartheta), y), \tag{4.6.53}$$

where $\varphi = \vartheta + \omega\tilde{\alpha}(\vartheta)$ is the inverse diffeomorphism of $\vartheta = \varphi + \omega\alpha(\varphi)$ in $\mathbb{T}^\nu$. By conjugation, the differential operators transform into

$$B^{-1}\omega \cdot \partial_\varphi B = \rho(\vartheta)\omega \cdot \partial_\vartheta, \quad B^{-1}\partial_y B = \partial_y, \quad \rho := B^{-1}(1 + \omega \cdot \partial_\varphi \alpha). \tag{4.6.54}$$

By (4.6.21), using also that B and B^{-1} commute with $\Pi_S^\perp$, we get

$$B^{-1}\mathcal{L}_1 B = \Pi_S^\perp[\rho\,\omega \cdot \partial_\vartheta + (B^{-1}b_3)\partial_{yyy} + (B^{-1}b_1)\partial_y + (B^{-1}b_0)]\Pi_S^\perp + B^{-1}\mathfrak{R}_1 B. \tag{4.6.55}$$

We choose α such that the new coefficient at order ∂_{yyy} is proportional to the function $\rho(\vartheta)$, namely

$$(B^{-1}b_3)(\vartheta) = m_3\,\rho(\vartheta), \quad m_3 \in \mathbb{R} \quad \Longrightarrow \quad b_3(\varphi) = m_3(1 + \omega \cdot \partial_\varphi \alpha(\varphi)). \tag{4.6.56}$$

The unique solution with zero average of (4.6.56) is

$$\alpha(\varphi) := \frac{1}{m_3}(\omega \cdot \partial_\varphi)^{-1}(b_3 - m_3)(\varphi), \quad m_3 := \frac{1}{(2\pi)^\nu}\int_{\mathbb{T}^\nu} b_3(\varphi)\, d\varphi. \tag{4.6.57}$$

Hence, by (4.6.55) we have

$$B^{-1}\mathcal{L}_1 B = \rho\,\mathcal{L}_2, \quad \mathcal{L}_2 := \Pi_S^\perp(\omega \cdot \partial_\vartheta + m_3\partial_{yyy} + c_1\partial_y + c_0)\Pi_S^\perp + \mathfrak{R}_2, \tag{4.6.58}$$

$$c_1 := \rho^{-1}(B^{-1}b_1), \quad c_0 := \rho^{-1}(B^{-1}b_0), \quad \mathfrak{R}_2 := \rho^{-1}B^{-1}\mathfrak{R}_1 B. \tag{4.6.59}$$

In order to control the corrections to the normal frequencies also at lower orders of size, we expand the constant coefficient m_3, defined in (4.6.57), in powers of ε. We have

$$m_3 = 1 + \varepsilon^2 d(\xi) + \mathtt{r}_{m_3} \tag{4.6.60}$$

where

$$\begin{aligned} d(\xi) :&= (12c_4 - 24\,c_1^2)\,M_{\varphi,x}[\overline{v}_x^2] + \varepsilon^2(2c_6 - \frac{8}{3}c_2^2)M_{\varphi,x}[\overline{v}^2] \\ &= (24c_4 - 48c_1^2)v_3 \cdot \xi + (4c_6 - \frac{16}{3}c_2^2)v_1 \cdot \xi \end{aligned} \tag{4.6.61}$$

and $|\mathtt{r}_{m_3}|^{Lip(\gamma)} \le \varepsilon^3$. The transformed operator $\mathcal{L}_2$ in (4.6.58) is still Hamiltonian, since the reparametrization of time preserves the Hamiltonian structure (see Section 2.2 and Remark 3.7 in [7]).
We note that, by (4.6.59), for $k = 0, 1$, we have

$$c_k = b_k + (B^{-1} - \mathrm{I})b_k + (\rho^{-1} - 1)\,B^{-1}\,b_k$$

and $b_k = O(\varepsilon)$ is the biggest term in the expression above. We define, for $k = 0, 1$,

$$\tilde{c}_k := c_k - b_k = (B^{-1} - \mathrm{I})b_k + (\rho^{-1} - 1)B^{-1}b_k \tag{4.6.62}$$

and we estimate them in Lemma 4.6.9. The remainder $\mathfrak{R}_2$ in (4.6.59) has still the form (4.5.5) and, by (4.6.51),

$$\mathfrak{R}_2 := -\rho^{-1}B^{-1}\mathfrak{R}_1 B = -\varepsilon^2\Pi_S^\perp\partial_x\mathcal{R}_2 + \mathcal{R}_* \tag{4.6.63}$$

where $\mathcal{R}_2$ is defined in (4.6.52) and we have renamed $\mathcal{R}_*$ the term of order $o(\varepsilon^2)$ in $\mathfrak{R}_2$.

Remark 4.6.8. In the proof of the estimates for the transformations B and $\mathcal{T}$, respectively defined in (4.6.53) and (4.6.75), we have to give a bound to the inverse of the operator $\mathcal{D}_\omega$ applied to the difference of a spatial and total (in space and time) average of some function in $H^s_{S^\perp}(\mathbb{T}^{\nu+1})$.
The main problem is that the estimate (4.3.23) is too rough to deal with functions $h(\varphi,x)$ of size greater or equal than ε^3, indeed, the terms $O(\varepsilon^3\gamma^{-1})$ are just not perturbative.
In the proofs of Lemma 4.6.9 and 4.6.10, we exploit the fact that if $h(\varphi,x)$ is a function supported on few harmonics, then we do not need to use the diophantine inequality (4.3.3) to give a bound to the divisors appearing in the Fourier coefficients of $\mathcal{D}_\omega^{-1}h$.
In this way, we overcome the problem discussed in Remark 8.11 in [8] and we can drop the hypotesis

$$j_1+j_2+j_3\neq 0 \quad \text{for all } j_1,j_2,j_3\in S$$

on the tangential sites assumed in [8].

Lemma 4.6.9. *There is $\sigma=\sigma(\nu,\tau)>0$ (possibly larger than the one in Lemma 4.6.5) such that*

$$|m_3-1|^{Lip(\gamma)}\le C\,\varepsilon^2, \quad |\partial_i m_3[\hat\imath]|\le \varepsilon^2\|\hat\imath\|_{s_0+\sigma}, \tag{4.6.64}$$

$$\|\alpha\|_s^{Lip(\gamma)}\le_s \varepsilon^4\gamma^{-1}+\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad \|\partial_i\alpha[\hat\imath]\|_s\le_s \|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}, \tag{4.6.65}$$

$$\|\rho-1\|_s^{Lip(\gamma)}\le_s \varepsilon^3+\varepsilon^{2b}\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad \|\partial_i\rho[\hat\imath]\|_s\le_s \varepsilon^{2b}(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}), \tag{4.6.66}$$

$$\|\tilde c_k\|_s^{Lip(\gamma)}\le_s \varepsilon^{3-2a}+\varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad \|\partial_i\tilde c_k[\hat\imath]\|_s\le_s \varepsilon(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}). \tag{4.6.67}$$

Proof. Estimate (4.6.64): To shorten the notation we write $\|\cdot\|_s:=\|\cdot\|_s^{Lip(\gamma)}$. We have $m_3-1=\int_{\mathbb{T}^\nu}(b_3-1)\,d\varphi$, then, by (4.6.23),

$$|m_3-1|\le\int_{\mathbb{T}^\nu}|b_3-1|d\varphi\le\|b_3-1\|_{s_0}\le C\varepsilon^2,$$
$$|\partial_i m_3[\hat\imath]|\le\int_{\mathbb{T}^\nu}\partial_i b_3[\hat\imath]d\varphi\le\|\partial_i b_3[\hat\imath]\|_{s_0}\le\varepsilon^2\|\hat\imath\|_{s_0+2}.$$

Estimate (4.6.65): To shorten the notation we write $\|\cdot\|_s:=\|\cdot\|_s^{Lip(\gamma)}$. By (4.6.57) and the fact that m_3 is a constant near to 1, it is sufficient to give a bound to b_3-m_3.
Consider the functions $g(t)=(1+t)^{-\frac{1}{3}}, \Upsilon(t)=(1+t)^{-3}$, defined in a small neighbourhood of the origin.
We have

$$\begin{aligned} b_3-m_3&=(b_3-1)-M_\varphi[b_3-1]\\ &\overset{(4.6.27)}{=}\Upsilon[M_x[g(a_1-1)-g(0)]]-M_\varphi[\Upsilon[M_x[g(a_1-1)-g(0)]]]. \end{aligned} \tag{4.6.68}$$

By the analiticity of Υ

$$\Upsilon(t)-\Upsilon(0)=\Upsilon'(0)\,t+\Upsilon_{\ge2}[t], \qquad \Upsilon_{\ge2}[t]:=\sum_{k\ge2}\frac{\Upsilon^{(k)}(0)}{k!}\,t^k,$$

for $|t|$ small enough. Hence, by (4.6.68),

$$\begin{aligned} b_3-m_3&=\Upsilon'(0)\{M_x[g(a_1-1)-g(0)]-M_{\varphi,x}[g(a_1-1)-g(0)]\}\\ &\quad+\Upsilon_{\ge2}[M_x[g(a_1-1)-g(0)]]-M_\varphi[\Upsilon_{\ge2}[M_x[g(a_1-1)-g(0)]]]. \end{aligned} \tag{4.6.69}$$

The difference of the last two terms in the right hand side of (4.6.69) can be estimated by

$$\begin{aligned}&\|\Upsilon_{\geq 2}[M_x[g(a_1-1)-g(0)]]-M_\varphi[\Upsilon_{\geq 2}[M_x[g(a_1-1)-g(0)]]]\|_s\\&\leq_s \|M_x[g(a_1-1)-g(0)]\|_{s_0}\|M_x[g(a_1-1)-g(0)]\|_s \overset{(4.6.23)}{\leq_s} \varepsilon^4(1+\|\mathfrak{I}_\delta\|_s).\end{aligned}$$

Now we prove a bound for the difference $M_x[g(a_1-1)-g(0)]-M_{\varphi,x}[g(a_1-1)-g(0)]$.
By Taylor expansion

$$\begin{aligned}g(a_1-1)-g(0)&=g'(0)(a_1-1)+\frac{g''(0)}{2}(a_1-1)^2+\frac{g'''(0)}{3!}(a_1-1)^3\\&\quad+\frac{(a_1-1)^4}{6}\int_0^1(1-s)^3\,g^{(4)}(s(a_1-1))\,ds\end{aligned}$$

and the last term of the right hand side can be estimated by $\varepsilon^4(1+\|\mathfrak{I}_\delta\|_{s+\sigma})$.
The function a_1 in (4.5.37) is a linear combination of $\Phi_B(T_\delta),\Phi_B(T_\delta)^2$ (and their derivatives in the x-variable) and $r_1(T_\delta)$, whose coefficients depend on $c_1,\ldots,c_7$ and other real constants. Without loss of generality, to simplify the notations, we can write $a_1=1+\Phi_B(T_\delta)+\Phi_B(T_\delta)^2+r_1(T_\delta)$ (recall (4.5.10), (4.5.29) and (4.5.25)). Thus, we have

$$\begin{aligned}M_x[a_1-1]&=M_x[\Phi_B(T_\delta)^2]+M_x[r_1(T_\delta)],\\M_x[(a_1-1)^2]&=M_x[\Phi_B(T_\delta)^2]+2M_x[\Phi_B(T_\delta)^3]+\mathtt{Q}_2(T_\delta),\\M_x[(a_1-1)^3]&=4M_x[\Phi_B(T_\delta)^3]+\mathtt{Q}_3(T_\delta),\end{aligned}$$

where $\|\mathtt{Q}_i(T_\delta)\|_s\leq_s\varepsilon^4+\varepsilon^{2+b}\|\mathfrak{I}_\delta\|_{s+\sigma}$ for $i=2,3$. By (4.5.25) and the fact that $\Phi_B(T_\delta)$ has size $O(\varepsilon)$, $r_1(T_\delta)$ is a polynomial of degree three in the variables $(\Phi_B(T_\delta),\Phi_B(T_\delta)_x)$, up to a remainder that is bounded in H^s norm by $\varepsilon^4(1+\|\mathfrak{I}_\delta\|_{s+\sigma})$. Thus, we reduced to study the differences

$$M_x[\Phi_B(T_\delta)^2]-M_{\varphi,x}[\Phi_B(T_\delta)^2],\qquad M_x[\Phi_B(T_\delta)^3]-M_{\varphi,x}[\Phi_B(T_\delta)^3].$$

We have, up to constants,

$$\Phi_B(T_\delta)^2=\varepsilon^2v_\delta^2+\varepsilon v_\delta\tilde q+\varepsilon^3v_\delta\Psi_2(v_\delta)+\tilde{\mathtt{Q}}_2(T_\delta),\qquad \Phi_B(T_\delta)^3=\varepsilon^3v_\delta^3+\tilde{\mathtt{Q}}_3(T_\delta),$$

where $\|\tilde{\mathtt{Q}}_i(T_\delta)\|_s\leq_s\varepsilon^4+\varepsilon^{2+b}\|\mathfrak{I}_\delta\|_{s+\sigma}$ for $i=2,3$. By the definition of $\tilde q$ and the fact that v_δ and z_0 are orthogonal in $L^2(\mathbb{T})$, we have

$$\varepsilon M_x[v_\delta\tilde q]=\varepsilon^{2+b}M_x[\Psi_2'(v_\delta)v_\delta z_0]+\varepsilon M_x[v_\delta\,\Psi_3(T_\delta)],\tag{4.6.70}$$

thus $\|M_x[\varepsilon v_\delta\tilde q]-M_{\varphi,x}[\varepsilon v_\delta\tilde q]\|_s\leq_s\varepsilon^4+\varepsilon^{2+b}\|\mathfrak{I}_\delta\|_{s+\sigma}$. It remains to estimate the differences of the averages of polynomial of degree two and three in the variables v_δ and its derivatives. These functions are of order ε^2 and ε^3, respectively, and supported on not many harmonics, because v_δ is not.
By (4.5.10) we get

$$M_x[v_\delta^2]-M_{\varphi,x}[v_\delta^2]=\varepsilon^{2(b-1)}\sum_{j\in S}|j|((y_\delta)_j-M_\varphi[(y_\delta)_j]).$$

We gain an extra smallness factor $\varepsilon^{2(b-1)}$ by the fact that $M_x[\overline{v}^2]$ is independent of φ (see Remark 4.6.1). Thus, we obtain $\varepsilon^2\|M_x[v_\delta^2]-M_{\varphi,x}[v_\delta^2]\|_s\leq_s\varepsilon^{2b}\|\mathfrak{I}_\delta\|_s$.
For the cubic terms in v_δ we use the following equality

$$M_x[v_\delta^3]-M_{\varphi,x}[v_\delta^3]=(M_x[\overline{v}^3]-M_{\varphi,x}[\overline{v}^3])+M_x[v_\delta^3-\overline{v}^3]-M_{\varphi,x}[v_\delta^3-\overline{v}^3],\tag{4.6.71}$$

where $\|M_x[v_\delta^3 - \overline{v}^3] - M_{\varphi,x}[v_\delta^3 - \overline{v}^3]\|_s \leq_s \varepsilon^3 \|\mathfrak{I}_\delta\|_s$.
We now analyze the first difference in the right hand side of (4.6.71). We cannot roughly bound it by ε^3 (see Remark 4.6.8). But we have

$$\mathcal{D}_\omega^{-1}\left(M_x[\overline{v}^3] - M_{\varphi,x}[\overline{v}^3]\right) = \sum_{\substack{j_1,j_2,j_3 \in S, \\ j_1+j_2+j_3=0 \\ \mathtt{l}(j_1)+\mathtt{l}(j_2)+\mathtt{l}(j_3)\neq 0}} \frac{\sqrt{\xi_{j_1}\xi_{j_2}\xi_{j_3}}}{\mathrm{i}\,\omega\cdot(\mathtt{l}(j_1)+\mathtt{l}(j_2)+\mathtt{l}(j_3))} e^{\mathrm{i}(\mathtt{l}(j_1)+\mathtt{l}(j_2)+\mathtt{l}(j_3))\cdot\varphi}. \tag{4.6.72}$$

We recall that $\omega = \overline{\omega} + O(\varepsilon^2)$, hence the denominator in (4.6.72) can be written as

$$\begin{aligned}\omega\cdot(\mathtt{l}(j_1)+\mathtt{l}(j_2)+\mathtt{l}(j_3)) &= \overline{\omega}\cdot(\mathtt{l}(j_1)+\mathtt{l}(j_2)+\mathtt{l}(j_3)) + (\omega-\overline{\omega})\cdot(\mathtt{l}(j_1)+\mathtt{l}(j_2)+\mathtt{l}(j_3)) \\ &= j_1^3+j_2^3+j_3^3+O(\varepsilon^2)\end{aligned}$$

and it is greater or equal than 1, indeed, if $j_1+j_2+j_3=0$, then $|j_1^3+j_2^3+j_3^3| = 3|j_1\,j_2\,j_3| \geq 3$. Thus, actually,

$$\|\mathcal{D}_\omega^{-1}\left(M_x[\overline{v}^3] - M_{\varphi,x}[\overline{v}^3]\right)\|_s \leq \varepsilon^3.$$

Finally, we get

$$\|b_3 - m_3\|_s \leq_s \varepsilon^3 + \varepsilon^{2b}\|\mathfrak{I}_\delta\|_{s+\sigma} \quad \text{and} \quad \|\mathcal{D}_\omega^{-1}(b_3-m_3)\|_s \leq_s \varepsilon^4\gamma^{-1} + \|\mathfrak{I}_\delta\|_{s+\sigma}, \tag{4.6.73}$$

so $\|\alpha\|_s \leq_s \varepsilon^4\gamma^{-1} + \|\mathfrak{I}_\delta\|_{s+\sigma}$.
Now we look to the partial derivative

$$\partial_i\left(\frac{b_3-m_3}{m_3}\right)[\hat{\imath}] = \frac{1}{m_3^2}\left[m_3\partial_i(b_3-m_3)[\hat{\imath}] - (b_3-m_3)\partial_i m_3[\hat{\imath}]\right]. \tag{4.6.74}$$

By (4.6.64) m_3-1 and $\partial_i m_3[\hat{\imath}]$ are of order ε^2, hence the estimate for $\partial_i\alpha[\hat{\imath}]$ comes from $\mathcal{D}_\omega^{-1}(\partial_i(b_3-m_3)[\hat{\imath}])$. By (4.6.69) we have

$$\begin{aligned}\partial_i(b_3-m_3)[\hat{\imath}] &= \Upsilon'(0)\{M_x[\partial_i(g(a_1-1)-g(0))[\hat{\imath}]] - M_{\varphi,x}[\partial_i(g(a_1-1)-g(0))[\hat{\imath}]]\} \\ &\quad + \partial_i\{\Upsilon_{\geq 2}[M_x[g(a_1-1)-g(0)]] - M_\varphi[\Upsilon_{\geq 2}[M_x[g(a_1-1)-g(0)]]]\}[\hat{\imath}]\end{aligned}$$

As before, the bigger terms are the partial derivatives of $M_x[g(a_1-1)-g(0)] - M_{\varphi,x}[g(a_1-1)-g(0)]$. We have

$$\partial_i(g(a_1-1)-g(0))[\hat{\imath}] = g'(0)\partial_i a_1[\hat{\imath}] + g''(0)(a_1-1)\,\partial_i a_1[\hat{\imath}] + \frac{g'''(0)}{2}(a_1-1)^2\,\partial_i a_1[\hat{\imath}] + \mathtt{T}(i_\delta,\hat{\imath})$$

where $\|\mathtt{T}(i_\delta,\hat{\imath})\|_s \leq_s \varepsilon^4(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma})$ and

$$\partial_i a_1[\hat{\imath}] = \partial_i\Phi_B(T_\delta)[\hat{\imath}] + 2\Phi_B(T_\delta)\partial_i\Phi_B(T_\delta)[\hat{\imath}] + \partial_i\tilde{q}[\hat{\imath}].$$

We note that $M_x[\partial_i\Phi_B(T_\delta)[\hat{\imath}]] = M_x[\partial_i\tilde{q}[\hat{\imath}]] = 0$. Thus, we focus on the terms

$$\Phi_B(T_\delta)\partial_i\Phi_B(T_\delta), \quad \Phi_B(T_\delta)^2\partial_i\Phi_B(T_\delta), \quad \Phi_B(T_\delta)\,\partial_i\tilde{q}[\hat{\imath}].$$

Further terms have Sobolev norm bounded by $\varepsilon^{2+b}(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma})$. We have

$$\begin{aligned}\Phi_B(T_\delta)\partial_i\Phi_B(T_\delta) &= \varepsilon^2 v_\delta\,\partial_i v_\delta[\hat{\imath}] + \varepsilon^3(\Psi_2(v_\delta) + \Psi_2'(v_\delta)v_\delta)\partial_i v_\delta[\hat{\imath}] + \varepsilon\partial_i(\tilde{q}\,v_\delta)[\hat{\imath}] + \tilde{\mathtt{T}}(i_\delta,\hat{\imath}), \\ \Phi_B(T_\delta)^2\partial_i\Phi_B(T_\delta)[\hat{\imath}] &= \varepsilon^3 v_\delta^2\,\partial_i v_\delta[\hat{\imath}] + \tilde{\mathtt{T}}(i_\delta,\hat{\imath}),\end{aligned}$$

where $\|\mathtt{T}(i_\delta,\hat\imath)\|_s \le_s \varepsilon^{2+b}(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma})$. We start from the average of the partial derivative of $v_\delta\tilde q$. By (4.6.70) we get $\varepsilon\|\partial_i M_x[v_\delta\tilde q][\hat\imath]\|_s \le_s \varepsilon^{2+b}(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma})$. Then, we reduce to study

$$M_x[v_\delta\partial_i v_\delta[\hat\imath]]-M_{\varphi,x}[v_\delta\partial_i v_\delta[\hat\imath]],\qquad M_x[v_\delta^2\partial_i v_\delta[\hat\imath]]-M_{\varphi,x}[v_\delta^2\partial_i v_\delta[\hat\imath]].$$

If we call $G(i_0(\varphi)):=y_\delta-y_0$, then we have

$$\partial_i v_\delta[\hat\imath]=\sum_{j\in S}\sqrt{|j|}\sqrt{\xi_j+\varepsilon^{2(b-1)}(y_\delta)_j}\,e^{\mathrm{i}(\theta_0)_j}\left(\mathrm{i}\hat\Theta_j+\varepsilon^{2(b-1)}\frac{\hat y_j+(\partial_i G(i_0(\varphi))[\hat\imath])_j}{2\,|j|\,(\xi_j+\varepsilon^{2(b-1)}(y_\delta)_j)}\right)e^{\mathrm{i}jx}$$

and

$$\begin{aligned}M_x[v_\delta\partial_i v_\delta[\hat\imath]]-M_{\varphi,x}[v_\delta\partial_i v_\delta[\hat\imath]]&=\varepsilon^{2(b-1)}\sum_{j\in S}\mathrm{i}\hat\Theta_j((y_\delta)_j-M_\varphi[(y_\delta)_j])\\&\quad+\frac{\varepsilon^{2(b-1)}}{2}\sum_{j\in S}\{(\partial_i G(i_0(\varphi))[\hat\imath])_j-M_\varphi[(\partial_i G(i_0(\varphi))[\hat\imath])_j]\}.\end{aligned}$$

Therefore, $\varepsilon^2\|M_x[v_\delta\partial_i v_\delta[\hat\imath]]-M_{\varphi,x}[v_\delta\partial_i v_\delta[\hat\imath]]\|_s\le_s\varepsilon^{2b}(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma})$. Moreover, we have $\|\varepsilon^3 M_x[v_\delta^2\partial_i v_\delta[\hat\imath]]\|_s\le_s\varepsilon^3(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma})$. Hence, we get

$$\|\partial_i(b_3-m_3)[\hat\imath]\|_s\le_s\varepsilon^{2b}(\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma})$$

and $\|\partial_i\alpha[\hat\imath]\|\le_s\|\hat\imath\|_{s+\sigma}+\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}$.By Lemma A.0.4 we deduce the inequality (4.6.65).

Estimate (4.6.66): Note that $\rho-1=B^{-1}\ ((b_3-m_3)/m_3)$. Thus, by Lemma A.0.5, (4.6.65), (4.6.73) we get

$$\begin{aligned}\|B^{-1}\ ((b_3-m_3)/m_3)\|_s^{Lip(\gamma)}&\le_s\|b_3-m_3\|_{s+1}^{Lip(\gamma)}+\|\alpha\|_{s+s_0}^{Lip(\gamma)}\|b_3-m_3\|_2^{Lip(\gamma)}\\&\le_s\varepsilon^3+\varepsilon^{2b}\|\mathfrak{I}_\delta\|_{s+s_0+\sigma}^{Lip(\gamma)}.\end{aligned}$$

Estimate (4.6.67): Note that $\|\rho^{-1}-1\|_s\le_s\|\rho-1\|_s$. By Lemma A.0.5 and (2.1.4), (4.6.24), we get, for $k=0,1$,

$$\|(B^{-1}-\mathrm{I})\,b_k\|_s^{Lip(\gamma)}\le_s\varepsilon^7\gamma^{-2}+\varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)},\quad\|(\rho^{-1}-1)b_k\|_s^{Lip(\gamma)}\le_s\varepsilon^4+\varepsilon^{1+2b}\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}.$$

□

4.6.4 Translation of the space variable

The goal of this section is to remove the space average from the coefficient in front of ∂_y. This is a preliminary step for the descent method that we apply at Section 8.7.
Consider the change of variable

$$(\mathcal{T}w)(\vartheta,y)=w(\vartheta,y+p(\vartheta)),\quad(\mathcal{T}^{-1}h)(\vartheta,z)=h(\vartheta,z-p(\vartheta)).\tag{4.6.75}$$

The differential operators in $\mathcal{L}_2$ (see (4.6.58)) transform into

$$\mathcal{T}^{-1}\omega\cdot\partial_\vartheta\mathcal{T}=\omega\cdot\partial_\vartheta+\{\omega\cdot\partial_\vartheta p(\vartheta)\}\partial_z,\quad\mathcal{T}^{-1}\partial_y\mathcal{T}=\partial_z.$$

Since $\mathcal{T}, \mathcal{T}^{-1}$ commute with $\Pi_S^\perp$, we get

$$\mathcal{L}_3 := \mathcal{T}^{-1}\mathcal{L}_2\mathcal{T} = \Pi_S^\perp(\omega\cdot\partial_\vartheta + m_3\,\partial_{zzz} + D_S\,\partial_z + d_0)\Pi_S^\perp + \mathfrak{R}_3, \tag{4.6.76}$$

$$d_1 := (\mathcal{T}^{-1}c_1) + \omega\cdot\partial_\vartheta p, \quad d_0 := \mathcal{T}^{-1}c_0, \quad \mathfrak{R}_3 := \mathcal{T}^{-1}\mathfrak{R}_2\mathcal{T} \tag{4.6.77}$$

and we choose

$$m_1 := \frac{1}{(2\pi)^{\nu+1}}\int_{\mathbb{T}^{\nu+1}} c_1\,d\vartheta\,dy, \quad p := (\omega\cdot\partial_\vartheta)^{-1}\Big(m_1 - \frac{1}{2\pi}\int_{\mathbb{T}} c_1\,dy\Big) \tag{4.6.78}$$

so that

$$\frac{1}{2\pi}\int_{\mathbb{T}} d_1(\vartheta,z)\,dz = m_1 \quad \forall\vartheta\in\mathbb{T}^\nu. \tag{4.6.79}$$

We define

$$\tilde d_k := d_k - \varepsilon\,\alpha_{k,1} - \varepsilon^2(\alpha_{k,2} - \alpha_{k,1}\,(\beta_1)_x), \qquad k = 0,1 \tag{4.6.80}$$

and we split $\mathfrak{R}_3 = -\varepsilon^2\partial_x\overline{\mathcal{R}}_2 + \tilde{\mathcal{R}}_*$, where $\overline{\mathcal{R}}_2$ is obtained replacing v_δ with $\overline{v}$ in $\mathcal{R}_2$ and

$$\tilde{\mathcal{R}}_* := \mathcal{T}^{-1}\mathcal{R}_*\mathcal{T} + \varepsilon^2\Pi_S^\perp\partial_x(\mathcal{R}_2 - \mathcal{T}^{-1}\mathcal{R}_2\mathcal{T}) + \varepsilon^2\Pi_S^\perp\partial_x(\overline{\mathcal{R}}_2 - \mathcal{R}_2), \tag{4.6.81}$$

where $\mathcal{R}_*$ has been defined in (4.5.34) and modified along this section by adding terms $o(\varepsilon^2)$. We used that $\mathcal{T}^{-1}$ commutes with ∂_x and $\Pi_S^\perp$.
We define

$$c(\xi) := M_{\varphi,x}[\alpha_{1,2} + \alpha_{1,1}\,(\beta_1)_x]. \tag{4.6.82}$$

This quantity is a correction at order ε^2 to the eigenvalues of the linear operator $\mathcal{L}_\omega$, see (4.5.33). In particular, we have

$$m_1 = \varepsilon^2 c(\xi) + \mathtt{r}_{m_1}, \quad \text{with} \quad |\mathtt{r}_{m_1}|^{Lip(\gamma)} \le \varepsilon^{3-2a}.$$

Lemma 4.6.10. *There is $\sigma := \sigma(\tau,\nu)$ (possibly larger than in Lemma 4.6.9) such that*

$$|m_1 - \varepsilon^2 c(\xi)|^{Lip(\gamma)} \le \varepsilon^7\,\gamma^{-2}, \quad |\partial_i(m_1 - \varepsilon^2 c(\xi))[\hat\imath]| \le \varepsilon^7\,\gamma^{-2}\|\hat\imath\|_{s_0+\sigma}, \tag{4.6.83}$$

$$\|p\|_s^{Lip(\gamma)} \le_s \varepsilon^4\gamma^{-1} + \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad \|\partial_i p[\hat\imath]\|_s \le_s \|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}, \tag{4.6.84}$$

$$\|\tilde d_k\|_s^{Lip(\gamma)} \le_s \varepsilon^{3-2a} + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad \|\partial_i\tilde d_k[\hat\imath]\|_s \le_s \varepsilon(\|\hat\imath\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}) \tag{4.6.85}$$

for $k = 0,1$. Moreover the matrix s-decay norm (see (2.3.1))

$$|\tilde{\mathcal{R}}_*|_s^{Lip(\gamma)} \le_s \varepsilon^3 + \varepsilon^2\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad |\partial_i\tilde{\mathcal{R}}_*[\hat\imath]|_s \le_s \varepsilon^2\|\hat\imath\|_{s+\sigma} + \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat\imath\|_{s_0+\sigma}. \tag{4.6.86}$$

The transformations $\mathcal{T},\mathcal{T}^{-1}$ satisfy (4.6.25), (4.6.26).

Proof. Estimate (4.6.84): To shorten the notation we write $\|\cdot\|_s := \|\cdot\|_s^{Lip(\gamma)}$. By (4.6.59) and (4.6.78) we have

$$\begin{aligned} m_1 - M_x[c_1] &= (M_{\varphi,x}[b_1] - M_x[b_1]) + (M_{\varphi,x}[(\rho^{-1}-1)b_1] - M_x[(\rho^{-1}-1)b_1]) \\ &\quad + (M_{\varphi,x}[(B^{-1}-\mathrm{I})b_1] - M_x[(B^{-1}-\mathrm{I})b_1]) \\ &\quad + (M_{\varphi,x}[(\rho^{-1}-1)(B^{-1}-\mathrm{I})b_1] - M_x[(\rho^{-1}-1)(B^{-1}-\mathrm{I})b_1]). \end{aligned} \tag{4.6.87}$$

By (4.6.65), (4.6.66) and Lemma 2.1.4, we get

$$\|(\rho^{-1}-1)(B^{-1}-\mathrm{I})b_1\|_s \leq_s \varepsilon^9\gamma^{-2}+\varepsilon^6\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma}.$$

Thus, by (4.3.23)

$$\|\mathcal{D}_\omega^{-1}\{M_{\varphi,x}[(\rho^{-1}-1)(B^{-1}-\mathrm{I})b_1]-M_x[(\rho^{-1}-1)(B^{-1}-\mathrm{I})b_1]\}\|_s \leq_s \varepsilon^9\gamma^{-3}+\varepsilon^6\gamma^{-2}\|\mathfrak{I}_\delta\|_{s+\sigma}. \quad (4.6.88)$$

We note that $\rho^{-1}-1$ is independent of x, hence $M_x[(\rho^{-1}-1)b_1]=(\rho^{-1}-1)M_x[b_1]$ and we can estimate the difference between the averages of $(\rho^{-1}-1)b_1$ with

$$\|(\rho^{-1}-1)M_x[b_1]\|_s \leq_s \varepsilon^5+\varepsilon^{2(b+1)}\|\mathfrak{I}_\delta\|_{s+\sigma} \quad (4.6.89)$$

and use again (4.3.23) for $\|\mathcal{D}_\omega^{-1}(M_{\varphi,x}[(\rho^{-1}-1)b_1]-M_x[(\rho^{-1}-1)b_1])\|_s \leq_s \varepsilon^5\gamma^{-1}+\varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}$. By Taylor expansion and the fact that $\tilde{\alpha}=-\alpha+(B-\mathrm{I})\alpha$ (see (4.6.53)), we have

$$b_1(\vartheta+\omega\tilde{\alpha}(\vartheta),x)=b_1(\vartheta,x)-\omega\cdot\partial_\vartheta b_1(\vartheta,x)\,\alpha(\vartheta)+\mathtt{R}_{\tilde{\alpha}}(\vartheta,x)$$

where $\|\mathtt{R}_{\tilde{\alpha}}\|_s \leq_s \varepsilon^8\gamma^{-2}+\varepsilon^4\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma}$. Moreover, by a change of variable

$$\int_{\mathbb{T}^{\nu+1}}(B^{-1}-\mathrm{I})b_1\,d\vartheta\,dx=\int_{\mathbb{T}^{\nu+1}}\omega\cdot\partial_\varphi\alpha(\varphi)\,b_1(\varphi,x)\,d\varphi\,dx. \quad (4.6.90)$$

From these facts and an integration by parts, we obtain

$$M_x[(B^{-1}-\mathrm{I})b_1]-M_{\varphi,x}[(B^{-1}-\mathrm{I})b_1]=\mathcal{D}_\omega\alpha\,M_x[b_1]-M_{\varphi,x}[(\mathcal{D}_\omega\alpha)\,b_1]+M_x[\mathtt{R}_{\tilde{\alpha}}]-M_{\varphi,x}[\mathtt{R}_{\tilde{\alpha}}]$$

and, by the estimate above for $\mathtt{R}_{\tilde{\alpha}}$ and the bound given by (4.6.73) for $\mathcal{D}_\omega\alpha$, we have

$$\|M_x[(B^{-1}-\mathrm{I})b_1]-M_{\varphi,x}[(B^{-1}-\mathrm{I})b_1]\|_s \leq_s \varepsilon^5+\varepsilon^{2(b+1)}\|\mathfrak{I}_\delta\|_{s+\sigma}. \quad (4.6.91)$$

As before, we can use (4.3.23). We remark that

$$\int_{\mathbb{T}}b_1(\varphi,y)\,dy=\int_{\mathbb{T}}(\mathcal{A}^T\alpha_1)(\varphi,y)\,dy=\int_{\mathbb{T}}\alpha_1(\varphi,y+\tilde{\beta}(\varphi,y))\,dy=\int_{\mathbb{T}}\alpha_1(\varphi,x)(1+\beta_x(\varphi,x))\,dx,$$

hence, it remains to estimate

$$M_{\varphi,x}[b_1]-M_x[b_1]=(M_{\varphi,x}[\alpha_1]-M_x[\alpha_1])+(M_{\varphi,x}[\alpha_1\beta_x]-M_x[\alpha_1\beta_x]). \quad (4.6.92)$$

The functions α_1 and $\alpha_1\beta_x$ are linear combinations of powers of $\Phi_B(T_\delta)$ (and its derivatives in the x-variable), $r_1(T_\delta)$, $r_0(T_\delta)$, whose coefficients depend on $c_1,\dots,c_7$ and other real constants. Hence, using the same reasoning adopted in the proof of the estimates (4.6.65), we get

$$\|\mathcal{D}_\omega^{-1}\{M_{\varphi,x}[\alpha_1]-M_x[\alpha_1]\}\|_s \leq_s \varepsilon^4\gamma^{-1}+\|\mathfrak{I}_\delta\|_{s+\sigma} \quad (4.6.93)$$

and the same estimate holds for $\mathcal{D}_\omega^{-1}\{M_{\varphi,x}[\beta_x\alpha_1]-M_x[\beta_x\alpha_1]\}$. By following analogous arguments used in the proof of the estimate (4.6.65) we conclude.

Estimate (4.6.83): By (4.6.59) and (4.6.78)

$$m_1=\int_{\mathbb{T}^{\nu+1}}b_1\,dx\,d\varphi+\int_{\mathbb{T}^{\nu+1}}\tilde{c}_1\,dx\,d\varphi.$$

Moreover,

$$\int_{\mathbb{T}^{\nu+1}} b_1\, dx\, d\varphi = \int_{\mathbb{T}^{\nu+1}} (\varepsilon^2 \alpha_{1,2} + \mathtt{R}_{a_1})\, dx\, d\varphi + \int_{\mathbb{T}^{\nu+1}} (\mathcal{A}^T - \mathrm{I})\alpha_1\, dx\, d\varphi.$$

Thus, the bound (4.6.83) comes from taking the maximum between

$$\left| \int_{\mathbb{T}^{\nu+1}} b_1\, d\varphi\, dy - \varepsilon^2 \int_{\mathbb{T}^{\nu+1}} (\alpha_{1,2} + \alpha_{1,1}\,(\beta_1)_x)\, d\varphi\, dx \right| \leq \|\mathtt{R}_{a_1}\|_{s_0} + \|(\mathcal{A}^T - \mathrm{I})(\alpha_1 - \varepsilon \alpha_{1,1})\|_{s_0} \leq \varepsilon^3$$

and $\|\tilde{c}_1\|_{s_0} \leq \varepsilon^7 \gamma^{-2} = \varepsilon^{3-2a}$.

Estimate (4.6.85): We observe that, by (4.6.38),

$$\tilde{d}_0 := \varepsilon^2 (\mathcal{A}^T - I)\alpha_{0,2} + \mathcal{A}^T \mathtt{R}_0 + \mathcal{R}_{\tilde{\beta}} + (\mathcal{T}^{-1} - \mathrm{I}) b_0 + \mathcal{T}^{-1} \tilde{c}_0.$$

By Lemma A.0.3, A.0.5 we have the following bounds

$$\|\varepsilon^2 (\mathcal{A}^T - I)\alpha_{0,2}\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 (1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}), \qquad \|\mathcal{A}^T \mathtt{R}_0\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 + \varepsilon \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)},$$
$$\|\mathcal{R}_{\tilde{\beta}}\|_s^{Lip(\gamma)} \leq_s \varepsilon^3 + \varepsilon^{1+b} \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \qquad \|\mathcal{T}^{-1} \tilde{c}_0\|_s^{Lip(\gamma)} \leq_s \varepsilon^7 \gamma^{-2} + \varepsilon \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)},$$
$$\|(\mathcal{T}^{-1} - \mathrm{I}) b_0\|_s^{Lip(\gamma)} \leq_s \varepsilon^7 \gamma^{-2} + \varepsilon \|\mathfrak{I}\|_{s+\sigma}^{Lip(\gamma)}.$$

From these estimates we get (4.6.85) for $k = 0$. The estimate for $k = 1$ can be obtained in the same way, considering that $\omega \cdot \partial_\vartheta p = O(\varepsilon^6 \gamma^{-1})$ in low norm by (4.6.84). □

4.6.5 Linear Birkhoff Normal Form (Step one)

Let us collect all the terms of order ε and ε^2 of $\mathcal{L}_3$ (see (4.6.76)) in the operators

$$\begin{aligned} \mathfrak{B}_1[h] &:= \alpha_{1,1}\, \partial_x h + \alpha_{0,1}\, h = \partial_x \{(2 c_2 v_{xx} - 6 c_3 v)\, h\}, \\ \mathfrak{B}_2[h] &:= \{\alpha_{1,2} - (\alpha_{1,1})_x\, \beta_1\}\, \partial_x h + \{\alpha_{0,2} - (\alpha_{0,1})_x\, \beta_1\}\, h - \partial_x \overline{\mathcal{R}}_2[h]. \end{aligned} \tag{4.6.94}$$

Note that $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are not the linear Hamiltonian vector fields of $H_S^\perp$ generated, respectively, by the Hamiltonians $R(v^2 z)$ and $R(v^2 z^2)$ in (4.1.1) at $v = \overline{v}$, as expected. Indeed, as we said in Remark 4.6.6, some Hamiltonians of type $R(v^2 z)$ have been eliminated by the diffeomorphism of the torus Φ defined in Section 8.1, and also the Hamiltonians $R(v^2 z^2)$ have been modified by that. Renaming $\vartheta = \varphi, z = x$ we have

$$\mathcal{L}_3 = \Pi_S^\perp (\omega \cdot \partial_\varphi + m_3 \partial_{xxx} + \varepsilon \mathfrak{B}_1 + \varepsilon^2 \mathfrak{B}_2 + \tilde{d}_1 \partial_x + \tilde{d}_0) \Pi_S^\perp + \tilde{\mathcal{R}}_* \tag{4.6.95}$$

where $\tilde{d}_1, \tilde{d}_0, \tilde{\mathcal{R}}_*$ are defined in (4.6.80) and (4.6.81).
The aim of this section is to eliminate $\mathfrak{B}_1$ from (4.6.95). In the next section we shall normalize the term $\mathfrak{B}_2$.
We conjugate $\mathcal{L}_3$ with a symplectic operator $\Phi_1 \colon H^s_{S^\perp}(\mathbb{T}^{\nu+1}) \to H^s_{S^\perp}(\mathbb{T}^{\nu+1})$ of the form

$$\Phi_1 := \exp(\varepsilon A_1) = \mathrm{I}_{H_S^\perp} + \varepsilon A_1 + \varepsilon^2 \frac{A_1^2}{2} + \varepsilon^3 \hat{A}_1, \quad \hat{A}_1 := \sum_{k \geq 3} \frac{\varepsilon^{k-3}}{k!} A_1^k, \tag{4.6.96}$$

where $A_1(\varphi)h = \sum_{j,j'\in S^c}(A_1)_j^{j'}(\varphi)\, h_{j'}\, e^{\mathrm{i}jx}$ is a Hamiltonian vector field. The map Φ_1 is symplectic, because it is the time-1 flow of a Hamiltonian vector field. Therefore

$$\begin{aligned}
&\mathcal{L}_3\Phi_1 - \Phi_1\Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx})\Pi_S^\perp = \\
&= \Pi_S^\perp(\varepsilon\{\mathcal{D}_\omega A_1 + m_3[\partial_{xxx}, A_1] + \mathfrak{B}_1\} + \varepsilon^2\{\mathfrak{B}_1 A_1 + \mathfrak{B}_2 + \frac{1}{2}m_3[\partial_{xxx}, A_1^2] + \frac{1}{2}(\mathcal{D}_\omega A_1^2)\} \\
&+ \tilde{d}_1\partial_x + R_3)\Pi_S^\perp
\end{aligned} \tag{4.6.97}$$

where

$$R_3 := \tilde{d}_1\partial_x(\Phi_1 - \mathrm{I}) + \tilde{d}_0\Phi_1 + \tilde{\mathcal{R}}_*\Phi_1 + \varepsilon^2\mathfrak{B}_2(\Phi_1 - \mathrm{I}) + \varepsilon^3\{\mathcal{D}_\omega\hat{A}_1 + m_3[\partial_{xxx}, \hat{A}_1] + \frac{1}{2}\mathfrak{B}_1 A_1^2 + \varepsilon\mathfrak{B}_1\hat{A}_1\}. \tag{4.6.98}$$

Remark 4.6.11. R_3 has no longer the form (4.5.5). However $R_3 = O(\partial_x^0)$ because $A_1 = O(\partial_x^{-1})$ and therefore $\Phi_1 - \mathrm{I}_{H_S^\perp} = O(\partial_x^{-1})$. Moreover the matrix decay norm of R_3 is $o(\varepsilon^2)$.

In order to eliminate the order ε from (4.6.97), we choose

$$(A_1)_j^{j'}(\ell) = \begin{cases} -\dfrac{(\mathfrak{B}_1)_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + m_3(j'^3 - j^3))} & \text{if } \overline{\omega}\cdot\ell + j'^3 - j^3 \neq 0, \qquad j,\, j' \in S^c,\, \ell \in \mathbb{Z}^\nu \\ 0 & \text{otherwise} \end{cases} \tag{4.6.99}$$

This definition is well posed. Indeed, by (4.6.1) and (4.6.94)

$$(\mathfrak{B}_1)_j^{j'}(\ell) := \begin{cases} (-2\mathrm{i}j\, c_2\,(j-j')^2 - 6\mathrm{i}j\, c_3)\sqrt{|j-j'|}\xi_{j-j'} & \text{if } j - j' \in S, \quad \ell = \mathtt{l}(j-j') \\ 0 & \text{otherwise.} \end{cases} \tag{4.6.100}$$

In particular $(\mathfrak{B}_1)_j^{j'}(\ell) = 0$ unless $|\ell| \leq 1$. Thus, for (ℓ, j, j') such that $\overline{\omega}\cdot\ell + j'^3 - j^3 \neq 0$, the denominators in (4.6.99) satisfy

$$\begin{aligned}
|\omega\cdot\ell + m_3(j'^3 - j^3)| &= |m_3(\overline{\omega}\cdot\ell + j'^3 - j^3) + (\omega - m_3\overline{\omega})\cdot\ell| \geq \\
&\geq |m_3|\,|\overline{\omega}\cdot\ell + j'^3 - j^3| - |\omega - m_3\overline{\omega}|\,|\ell| \geq 1/2, \ \ \forall|\ell| \leq 1
\end{aligned} \tag{4.6.101}$$

for ε small enough, since $m_3 - 1$ and $\omega - \overline{\omega}$ are $O(\varepsilon^2)$. A_1 defined in (4.6.99) is a Hamiltonian vector field as $\mathfrak{B}_1$.

Lemma 4.6.12. *(Lemma 8.16 in [8]) If $j, j' \in S^c, j - j' \in S, \ell = \mathtt{l}(j-j')$, then*

$$\overline{\omega}\cdot\ell + j'^3 - j^3 = 3\,j\,j'\,(j'-j) \neq 0.$$

Corollary 4.6.13. *(Corollary 8.17 in [8]) Let $j, j' \in S^c$. If $\overline{\omega}\cdot\ell + j'^3 - j^3 = 0$ then $(\mathfrak{B}_1)_j^{j'} = 0$.*

By (4.6.99) and the previous corollary, the term of order ε in (4.6.97) is

$$\Pi_S^\perp(\mathcal{D}_\omega A_1 + m_3[\partial_{xxx}, A_1] + \mathfrak{B}_1)\Pi_S^\perp = 0. \tag{4.6.102}$$

We now prove that A_1 is a bounded transformation.

Lemma 4.6.14. *(Lemma 8.18 in [8])*

(i) *For all* $\ell \in \mathbb{Z}^\nu, j, j' \in S^c$,

$$|(A_1)_j^{j'}(\ell)| \leq C(|j| + |j'|)^{-1}, \quad |(A_1)_j^{j'}(\ell)|^{lip} \leq \varepsilon^{-2}(|j| + |j'|)^{-1}. \tag{4.6.103}$$

(ii) $(A_1)_j^{j'}(\ell) = 0$ *for all* $\ell \in \mathbb{Z}^\nu, j, j' \in S^c$ *such that* $|j - j'| > C_S$, *where* $C_S := \max\{|j| : j \in S\}$.

The previous lemma means that $A = O(\partial_x^{-1})$. More precisely, we deduce that

Lemma 4.6.15. *(Lemma* 8.19 *in [8])* $|A_1\partial_x|_s^{Lip(\gamma)} + |\partial_x A_1|_s^{Lip(\gamma)} \leq C(s)$.

It follows that the symplectic map Φ_1 in (4.6.96) is invertible for ε small, with inverse

$$\Phi_1^{-1} = \exp(-\varepsilon A_1) = \mathrm{I}_{H_S^\perp} + \varepsilon \check{A}_1, \ \check{A}_1 := \sum_{n\geq 1} \frac{\varepsilon^{n-1}}{n!}(-A_1)^n, \ |\check{A}_1\partial_x|_s^{Lip(\gamma)} + |\partial_x \check{A}_1|_s^{Lip(\gamma)} \leq C(s). \tag{4.6.104}$$

Since A_1 solves the homological equation (4.6.102), the ε-term in (4.6.95) is zero, and, with a straightforward calculation, the ε^2-term simplifies to $\mathfrak{B}_2 + \frac{1}{2}[\mathfrak{B}_1, A_1]$. We obtain the Hamiltonian operator

$$\mathcal{L}_4 := \Phi_1^{-1}\mathcal{L}_3\Phi_1 = \Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx} + \tilde{d}_1\partial_x + \varepsilon^2\{\mathfrak{B}_2 + \frac{1}{2}[\mathfrak{B}_1, A_1]\} + \tilde{R}_4)\Pi_S^\perp, \tag{4.6.105}$$

$$\tilde{R}_4 := (\Phi_1^{-1} - \mathrm{I})\Pi_S^\perp[\varepsilon^2(\mathfrak{B}_2 + \frac{1}{2}[\mathfrak{B}_1, A_1]) + \tilde{d}_1\partial_x] + \Phi_1^{-1}\Pi_S^\perp R_3. \tag{4.6.106}$$

We split A_1 defined in (4.6.99), (4.6.100) into $A_1 = \overline{A}_1 + \tilde{A}_1$ where, for all $j, j' \in S^c, \ell \in \mathbb{Z}^\nu$,

$$(\overline{A}_1)_j^{j'}(\ell) := -\frac{2j\, c_2\,(j-j')^2\sqrt{|j-j'|\xi_{j-j'}} + 6j\, c_3\sqrt{|j-j'|\xi_{j-j'}}}{\overline{\omega}\cdot l + j'^3 - j^3} \tag{4.6.107}$$

if $\overline{\omega}\cdot\ell + j'^3 - j^3 \neq 0$, $j - j' \in S$, $\ell = \mathtt{1}(j-j')$, and $(\overline{A}_1)_j^{j'}(\ell) := 0$ otherwise.
By Lemma 4.6.12, for all $j, j' \in S^c, \ell \in \mathbb{Z}^\nu$,

$$(\overline{A}_1)_j^{j'}(\ell) = \begin{cases} -\dfrac{2}{3}c_2\left(\dfrac{j-j'}{j'}\right)\sqrt{|j-j'|\xi_{j-j'}} - 2\,c_3\,\dfrac{1}{j'(j'-j)}\sqrt{|j-j'|\xi_{j-j'}} & \text{if } j - j' \in S, \\ 0 & \text{otherwise,} \end{cases}$$

namely

$$\overline{A}_1 h = -\frac{2}{3}c_2\Pi_S^\perp[\overline{v}_x\,(\partial_x^{-1}h)] + 2\,c_3\Pi_S^\perp[(\partial_x^{-1}\overline{v})(\partial_x^{-1}h)], \quad \forall h \in H^s_{S^\perp}(\mathbb{T}^{\nu+1}). \tag{4.6.108}$$

The difference is

$$(\tilde{A}_1)_j^{j'}(\ell) := -\frac{(2c_2\,j\,(j-j')^2 + 6\,c_3\,j)\sqrt{|j-j'|\xi_{j-j'}}\{(\omega-\overline{\omega})\cdot\ell + (m_3-1)(j'^3 - j^3)\}}{(\omega\cdot\ell + m_3(j'^3 - j^3))(\overline{\omega}\cdot\ell + j'^3 - j^3)} \tag{4.6.109}$$

for $j, j' \in S^c, j - j' \in S, \ell = \mathtt{1}(j-j')$, and $(\tilde{A}_1)_j^{j'}(\ell) = 0$ otherwise. Then, by (4.6.105),

$$\mathcal{L}_4 = \Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx} + \tilde{d}_1\,\partial_x + \varepsilon^2 T + R_4)\Pi_S^\perp, \tag{4.6.110}$$

where

$$T := \mathfrak{B}_2 + \frac{1}{2}[\mathfrak{B}_1, \overline{A}_1], \quad R_4 := \frac{\varepsilon^2}{2}[\mathfrak{B}_1, \tilde{A}_1] + \tilde{R}_4. \tag{4.6.111}$$

The operator T is Hamiltonian as $\mathfrak{B}_1, \mathfrak{B}_2, \overline{A}_1$, because the commutator of two Hamiltonian vector fields is Hamiltonian.

Lemma 4.6.16. *There is* $\sigma = \sigma(\nu, \tau) > 0$ *(possibly larger than in Lemma 4.6.10) such that*

$$|R_4|_s^{Lip(\gamma)} \leq_s \varepsilon^7 \gamma^{-2} + \varepsilon \|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad |\partial_i R_4[\hat{\imath}]|_s \leq_s \varepsilon(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma} \|\hat{\imath}\|_{s_0+\sigma}). \tag{4.6.112}$$

Proof. The proof follows the one of Lemma 8.20 in [8]. The only difference is the estimate on the coefficient $\tilde{d}_0$ (see (4.6.85)), that gives the term of size $\varepsilon^7\gamma^{-2}$ in (4.6.112), instead of $\varepsilon^5\gamma^{-1}$ in the inequality (8.95) in [8]. □

4.6.6 Linear Birkhoff Normal form (Step two)

The goal of this section is to normalize the term $\varepsilon^2 T$ from the operator $\mathcal{L}_4$ defined in (4.6.105). We cannot eliminate the terms $O(\varepsilon^2)$ at all, because some harmonics of $\varepsilon^2 T$, which correspond to null divisors, are not naught.
We conjugate the Hamiltonian operator $\mathcal{L}_4$ via a symplectic map

$$\Phi_2 := \exp(\varepsilon^2 A_2) = \mathrm{I}_{H_S^\perp} + \varepsilon^2 A_2 + \varepsilon^4 \hat{A}_2, \quad \hat{A}_2 := \sum_{k\geq 2} \frac{\varepsilon^{2(k-2)}}{k!} A_2^k \tag{4.6.113}$$

where $A_2(\varphi) = \sum_{j,j'\in S^c} (A_2)_j^{j'}(\varphi) h_{j'} e^{\mathrm{i}jx}$ is a Hamiltonian vector field. We compute

$$\mathcal{L}_4\Phi_2 - \Phi_2\Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx})\Pi_S^\perp = \Pi_S^\perp(\varepsilon^2\{\mathcal{D}_\omega A_2 + m_3[\partial_{xxx}, A_2] + T\} + \tilde{d}_1\partial_x + \tilde{R}_5)\Pi_S^\perp, \tag{4.6.114}$$

$$\tilde{R}_5 := \Pi_S^\perp\{\varepsilon^4(\mathcal{D}_\omega \hat{A}_2 + m_3[\partial_{xxx}, \hat{A}_2]) + (\tilde{d}_1\partial_x + \varepsilon^2 T)(\Phi_2 - \mathrm{I}) + R_4\Phi_2\}\Pi_S^\perp. \tag{4.6.115}$$

We define

$$(A_2)_j^{j'}(\ell) := \begin{cases} -\dfrac{T_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + m_3(j'^3 - j^3))} & \text{if } \overline{\omega}\cdot\ell + j'^3 - j^3 \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.6.116}$$

The definition is well posed. Indeed the matrix entries $T_j^{j'}(\ell) = 0$ for all $|j - j'| > 2C_S, \ell \in \mathbb{Z}^\nu$, where $C_S := \max\{|j| : j \in S\}$. Also $T_j^{j'}(\ell) = 0$ for all $j, j' \in S^c, |\ell| > 2$. Thus, arguing as in (4.6.101), if $\overline{\omega}\cdot\ell + j'^3 - j^3 \neq 0$, then $|\omega\cdot\ell + m_3(j'^3 - j^3)| \geq 1/2$. The operator A_2 is a Hamiltonian vector field because T is Hamiltonian.

Resonant terms

Now we compute the terms of $\varepsilon^2 T$ that cannot be removed by the Birkhoff map Φ_2.
By (4.6.108), (4.6.111) we get, for $h \in H_{S^\perp}^s$,

$$\begin{aligned} \mathfrak{B}_1\,\overline{A}_1[h] &= -\frac{4}{3}c_2^2\,\partial_x\Pi_S^\perp[\overline{v}_{xx}\,\Pi_S^\perp[\overline{v}_x\,(\partial_x^{-1}h)]] + 4c_2c_3\partial_x\Pi_S^\perp[\overline{v}_{xx}\Pi_S^\perp[(\partial_x^{-1}\overline{v})(\partial_x^{-1}h)]] \\ &\quad + 4c_2c_3\partial_x\Pi_S^\perp[\overline{v}\Pi_S^\perp[\overline{v}_x\,(\partial_x^{-1}h)]] - 12c_3^2\partial_x\Pi_S^\perp[\overline{v}\Pi_S^\perp[(\partial_x^{-1}\overline{v})(\partial_x^{-1}h)]] \\ \overline{A}_1\mathfrak{B}_1[h] &= -\frac{4}{3}c_2^2\,\Pi_S^\perp[\overline{v}_x\,\Pi_S^\perp[\overline{v}_{xx}h]] + 4c_2c_3\Pi_S^\perp[\overline{v}_x\,\Pi_S^\perp[\overline{v}\,h]] \\ &\quad + 4c_2c_3\Pi_S^\perp[(\partial_x^{-1}\overline{v})\Pi_S^\perp[\overline{v}_{xx}\,h]] - 12c_3^2\Pi_S^\perp[(\partial_x^{-1}\overline{v})\Pi_S^\perp[\overline{v}\,h]] \end{aligned}$$

whence, for all $j, j' \in S^c, \ell \in \mathbb{Z}^\nu$,

$$
\begin{aligned}
([\mathfrak{B}_1, \overline{A}_1])_j^{j'}(\ell) = \frac{4}{3} c_2^2 \,\mathrm{i} & \sum_{\substack{j_1, j_2 \in S, j_1 + j_2 = j - j', \\ j' + j_2 \in S^c, \mathtt{l}(j_1) + \mathtt{l}(j_2) = \ell}} \Big(\frac{j\, j_1^2\, j_2 - j_1\, j_2^2\, j'}{j'} \Big) \sqrt{|j_1\, j_2| \xi_{j_1} \xi_{j_2}} \\
& + 4 c_2 c_3 \,\mathrm{i} \sum_{\substack{j_1, j_2 \in S, j_1 + j_2 = j - j', \\ j' + j_2 \in S^c, \mathtt{l}(j_1) + \mathtt{l}(j_2) = \ell}} \Big(\frac{-j\, j_1^3 + j\, j_1\, j_2^2 - j_1^2\, j_2\, j' - j_2^3\, j'}{j' j_1 j_2} \Big) \sqrt{|j_1\, j_2| \xi_{j_1} \xi_{j_2}} \\
& + 12\, c_3^2 \,\mathrm{i} \sum_{\substack{j_1, j_2 \in S, j_1 + j_2 = j - j', \\ j' + j_2 \in S^c, \mathtt{l}(j_1) + \mathtt{l}(j_2) = \ell}} \Big(\frac{j j_1 - j' j_2}{j' j_1 j_2} \Big) \sqrt{|j_1\, j_2| \xi_{j_1} \xi_{j_2}} .
\end{aligned}
\tag{4.6.117}
$$

If $([\mathfrak{B}_1, \overline{A}_1])_j^{j'}(\ell) \neq 0$ there are $j_1, j_2 \in S$ such that $j_1 + j_2 = j - j', j' + j_2 \in S^c, \mathtt{l}(j_1) + \mathtt{l}(j_2) = \ell$. Then

$$
\overline{\omega} \cdot \ell + j'^3 - j^3 = \overline{\omega} \cdot \mathtt{l}(j_1) + \overline{\omega} \cdot \mathtt{l}(j_2) + j'^3 - j^3 = j_1^3 + j_2^3 + j'^3 - j^3 . \tag{4.6.118}
$$

Thus, if $\overline{\omega} \cdot \ell + j'^3 - j^3 = 0$, Lemma (4.1.3) implies that $(j_1 + j_2)(j_1 + j')(j_2 + j') = 0$. Now $j_1 + j', j_2 + j' \neq 0$ because $j_1, j_2 \in S, j' \in S^c$ and S is symmetric. Hence $j_1 + j_2 = 0$, which implies $j = j'$ and $\ell = 0$. In conclusion, if $\overline{\omega} \cdot \ell + j'^3 - j^3 = 0$, the only nonzero matrix entry $([\mathfrak{B}_1, \overline{A}_1])_j^{j'}(\ell)$ is

$$
\begin{aligned}
\frac{1}{2}([\mathfrak{B}_1, \overline{A}_1])_j^j(0) = \frac{4}{3} c_2^2 \,\mathrm{i} & \sum_{j_2 \in S, j_2 + j \in S^c} j_2^3\, |j_2| \xi_{j_2} + 8 c_2 c_3 \,\mathrm{i} \sum_{j_2 \in S, j_2 + j \in S^c} j_2\, |j_2| \xi_{j_2} \\
& + 12\, c_3^2 \mathrm{i} \sum_{j_2 \in S, j_2 + j \in S^c} j_2^{-1}\, |j_2|\, \xi_{j_2} .
\end{aligned}
\tag{4.6.119}
$$

Now consider $\mathfrak{B}_2$ defined in (4.6.94). We split $\mathfrak{B}_2 = B_1 + B_2 + B_3 + B_4 + B_5$, where

$$
\begin{aligned}
& B_1[h] := \alpha_{1,2}\, h_x, \quad B_2[h] := \alpha_{0,2}\, h, \quad B_3[h] := -(\alpha_{1,1})_x\, \beta_1\, h_x, \\
& B_4[h] := -(\alpha_{0,1})_x\, \beta_1, \quad B_5[h] := -\partial_x \overline{\mathcal{R}}_2[h].
\end{aligned}
\tag{4.6.120}
$$

We denote by $(\alpha)_{j,\ell}$ the (j, ℓ)-th Fourier coefficient of $\alpha(\varphi, x)$ as function of time and space. The Fourier representation of $B_i, i = 1, \dots, 4$ in (4.6.120) is

$$
\begin{aligned}
(B_1)_j^{j'}(\ell) &= \mathrm{i}\, j'\, (\alpha_{1,2})_{j-j', \mathtt{l}(j-j')}, \qquad (B_2)_j^{j'}(\ell) = (\alpha_{0,2})_{j-j', \mathtt{l}(j-j')} \\
(B_3)_j^{j'}(\ell) &= 4 c_1 c_2 \mathrm{i}\, j' (\overline{v}_{xxx} \overline{v})_{j-j', \mathtt{l}(j-j')} + \frac{4}{3} c_2^2 \mathrm{i} j' (\overline{v}_{xxx} (\partial_x^{-1} \overline{v}))_{j-j', \mathtt{l}(j-j')} \\
& \quad - 12 c_1 c_3 \mathrm{i}\, j'\, (\overline{v}\, \overline{v}_x)_{j-j', \mathtt{l}(j-j')} - 4 c_2 c_3 \mathrm{i}\, j' (\overline{v}_x (\partial_x^{-1} \overline{v}))_{j-j', \mathtt{l}(j-j')}, \\
(B_4)_j^{j'}(\ell) &= 4 c_1 c_2 (\overline{v}_{xxxx} \overline{v})_{j-j', \mathtt{l}(j-j')} + \frac{4}{3} c_2^2 (\overline{v}_{xxxx} (\partial_x^{-1} \overline{v}))_{j-j', \mathtt{l}(j-j')} \\
& \quad - 12 c_1 c_3 (\overline{v} \overline{v}_{xx})_{j-j', \mathtt{l}(j-j')} - 4 c_2 c_3 (\overline{v}_{xx} (\partial_x^{-1} \overline{v}))_{j-j', \mathtt{l}(j-j')}
\end{aligned}
$$

If $(B_k)_j^{j'}(\ell) \neq 0$, $k = 1, \dots, 4$ there are $j_1, j_2 \in S$ such that $j_1 + j_2 = j - j', \ell = \mathtt{l}(j_1) + \mathtt{l}(j_2)$ and (4.6.118) holds. Thus, if $\overline{\omega} \cdot \ell + j'^3 - j^3 = 0$, Lemma (4.1.3) implies that $(j_1 + j_2)(j_1 + j')(j_2 + j') = 0$, and, since $j' \in S^c$ and S is symmetric, the only possibility is $j_1 + j_2 = 0$. Hence $j = j'$, $\ell = 0$. In conclusion, if $\overline{\omega}\cdot\ell + j'^3 - j^3 = 0$, the only nonzero matrix element $(B_i)_j^{j'}(\ell), i = 1,\dots,4$, by (4.6.7), is

$$\begin{aligned}
&(B_1)_j^j(0) = \mathrm{i}j\sum_{k\in S}(-2c_6k^2 - 12c_7 + \frac{4}{3}c_2^2k^2 + 4c_2c_3)\,|k|\xi_k,\\
&(B_2)_j^j(0) = \sum_{k\in S}(-4c_1c_2\,k^4 - 12c_1c_3\,k^2)\,|k|\xi_k,\\
&(B_3)_j^j(0) = \mathrm{i}j\sum_{k\in S}(\frac{4}{3}c_2^2\,k^2 + 4c_2c_3)|k|\xi_k, \quad (B_4)_j^j(0) = \sum_{k\in S}(4c_1c_2\,k^4 + 12c_1c_3\,k^2)\,|k|\xi_k
\end{aligned} \tag{4.6.121}$$

We note that $c(\xi)$ defined in (4.6.82) is equal to $-\mathrm{i}\sum_{i=1}^4 j^{-1}\,(B_i)_j^j(0)$ (observe that the term $j^{-1}\,(B_i)_j^j(0)$ is independent of j) and we write

$$\begin{aligned}
c(\xi) &= \sum_{k\in S^+}(-4c_6\,k^3 - 24c_7k + \frac{16}{3}c_2^2\,k^3 + 16c_2c_3k)\,\xi_k\\
&= (\frac{16}{3}c_2^2 - 4c_6)v_3\cdot\xi + (16c_2c_3 - 24c_7)v_1\cdot\xi,
\end{aligned} \tag{4.6.122}$$

where $v_3\cdot\xi = \sum_{j\in S^+} j^3\,\xi_j$ and $v_1\cdot\xi = \sum_{j\in S^+} j\,\xi_j$.
As before, the only possibility to get a zero at the denominator of (4.6.116) is $j_1 + j_2 = 0$. Therefore

$$\begin{aligned}
(B_5)_j^j(0) = \frac{4}{3}c_2^2\mathrm{i}&\sum_{j_2\in S, j_2+j\in S} j_2^3\,|j_2|\xi_{j_2} + 8c_2c_3\mathrm{i}\sum_{j_2\in S, j_2+j\in S} j_2\,|j_2|\xi_{j_2}\\
&+ 12c_3^2\mathrm{i}\sum_{j_2\in S, j_2+j\in S} j_2^{-1}\,|j_2|\xi_{j_2}.
\end{aligned} \tag{4.6.123}$$

We note that for every odd function $f\colon S\to\mathbb{Z}$, by the simmetry of S, we have

$$\sum_{j_2\in S} f(j_2)\,\xi_{j_2} = 0.$$

Thus, by (4.6.119) and (4.6.123), we get

$$(B_5)_j^j(0) + \frac{1}{2}([\mathfrak{B}_1, \overline{A}_1])_j^j(0) = \frac{4}{3}c_2^2\mathrm{i}\sum_{j_2\in S} j_2^3\,|j_2|\xi_{j_2} + 8c_2c_3\mathrm{i}\sum_{j_2\in S} j_2\,|j_2|\xi_{j_2} + 12c_3^2\mathrm{i}\sum_{j_2\in S} j_2^{-1}\,|j_2|\xi_{j_2} = 0.$$

Finally, we have

$$\mathcal{L}_5 := \Phi_2^{-1}\mathcal{L}_4\Phi_2 = \Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx} + (\tilde{d}_1 + \varepsilon^2 c(\xi))\,\partial_x + R_5)\Pi_S^\perp, \tag{4.6.124}$$
$$R_5 := (\Phi_2^{-1} - \mathrm{I})\Pi_S^\perp(\tilde{d}_1 + \varepsilon^2 c(\xi))\partial_x + \Phi_2^{-1}\Pi_S^\perp\tilde{R}_5. \tag{4.6.125}$$

Lemma 4.6.17. *R_5 satisfies the same estimates* (4.6.112) *as R_4 (with a possibly larger σ).*

4.6.7 Descent method

The goal of this section is to transform $\mathcal{L}_5$ in (4.6.124) in order to make constant the coefficient in front of ∂_x. We conjugate $\mathcal{L}_5$ via a symplectic map of the form

$$\mathcal{S} := \exp(\Pi_S^\perp(w\partial_x^{-1}))\Pi_S^\perp = \Pi_S^\perp(\mathrm{I} + w\partial_x^{-1})\Pi_S^\perp + \hat{\mathcal{S}}, \quad \hat{\mathcal{S}} := \sum_{k\geq 2}\frac{1}{k!}[\Pi_S^\perp(w\partial_x^{-1})]^k\Pi_S^\perp, \tag{4.6.126}$$

where $w\colon \mathbb{T}^{\nu+1} \to \mathbb{R}$ is a function. Note that $\Pi_S^\perp(w\partial_x^{-1})\Pi_S^\perp$ is the Hamiltonian vector field generated by $-\frac{1}{2}\int_{\mathbb{T}} w(\partial_x^{-1}h)^2\,dx, h \in H_S^\perp$. We calculate

$$\begin{aligned}
&\mathcal{L}_5\mathcal{S} - \mathcal{S}\Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx} + m_1\partial_x)\Pi_S^\perp = \Pi_S^\perp(3m_3w_x + \tilde{d}_1 + \varepsilon^2c(\xi) - m_1)\partial_x\Pi_S^\perp + \tilde{R}_6, \qquad (4.6.127)\\
&\tilde{R}_6 : = \Pi_S^\perp\{(3m_3w_{xx} + (\tilde{d}_1 + \varepsilon^2c(\xi))\Pi_S^\perp w - m_1w)\pi_0 + (\mathcal{D}_\omega w + m_3w_{xxx}\\
&\quad + (\tilde{d}_1 + \varepsilon^2c(\xi))\Pi_S^\perp w_x)\partial_x^{-1} + \mathcal{D}_\omega\hat{S} + m_3[\partial_{xxx}, \hat{S}] + (\tilde{d}_1 + \varepsilon^2c(\xi))\partial_x\hat{S} - m_1\hat{S}\partial_x + R_5\mathcal{S}\}\Pi_S^\perp
\end{aligned}$$

where $\tilde{R}_6$ collects all the bounded terms. By (4.6.80), (4.6.82), we solve

$$3m_3w_x + \tilde{d}_1 + \varepsilon^2c(\xi) - m_1 = 0$$

choosing $w := -(3m_3)^{-1}\partial_x^{-1}(\tilde{d}_1 + \varepsilon^2c(\xi) - m_1)$. For ε sufficiently small, the operator $\mathcal{S}$ is invertible and, by (4.6.127),

$$\mathcal{L}_6 := \mathcal{S}^{-1}\mathcal{L}_4\mathcal{S} = \Pi_S^\perp(\mathcal{D}_\omega + m_3\partial_{xxx} + m_1\partial_x)\Pi_S^\perp + R_6, \quad R_6 := \mathcal{S}^{-1}\tilde{R}_6. \qquad (4.6.128)$$

Since $\mathcal{S}$ is symplectic, $\mathcal{L}_6$ is Hamiltonian.

Lemma 4.6.18. *There is $\sigma := \sigma(\nu,\tau) > 0$ (possibly larger than in Lemma 4.6.16) such that*

$$|\mathcal{S}^{\pm1} - \mathrm{I}|_s^{Lip(\gamma)} \leq_s \varepsilon^7\gamma^{-2} + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}, \quad |\partial_i\mathcal{S}^{\pm1}[\hat{\imath}]|_s \leq_s \varepsilon(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}).$$

The remainder R_6 satisfies the same estimates of R_4 (with a possibly larger σ).

Proof. By (4.6.64), (4.6.83), (4.6.85), $\|w\|_s^{Lip(\gamma)} \leq_s \varepsilon^7\gamma^{-2} + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{Lip(\gamma)}$, and the lemma follows by the definition of $\mathcal{S}$, see (4.6.126). Since $\hat{S} = O(\partial_x^{-2})$ the commutator $[\partial_{xxx}, \hat{S}] = O(\partial_x^0)$ and $|[\partial_{xxx}, \hat{S}]|_s^{Lip(\gamma)} \leq_s \|w\|_{s_0+3}^{Lip(\gamma)}\|w\|_{s+3}^{Lip(\gamma)}$. □

4.6.8 KAM reducibility and inversion of $\mathcal{L}_\omega$

The coefficients m_3, m_1 of the operator $\mathcal{L}_6$ in (4.6.128) are constants, and the remainder R_6 is a bounded operator of order ∂_x^0 with small matrix decay norm. Then we can diagonalize $\mathcal{L}_6$ by applying the iterative KAM reducibility Theorem 4.2 in [7] along the sequence of scales

$$N_n := N_0^{\chi^n}, \quad n = 0,1,2,\dots, \quad \chi := 3/2, \quad N_0 > 0. \qquad (4.6.129)$$

In Section 9, the initial N_0 will (slightly) increase to infinity as $\varepsilon \to 0$, see (4.7.4). The required smallness condition (see (4.14) in [7]) is

$$N_0^{C_0}|R_6|_{s_0+\beta}^{Lip(\gamma)}\gamma^{-1} \leq 1, \qquad (4.6.130)$$

where $\beta = 7\tau + 6$ (see (4.1) in [7]), τ is the diophantine exponent in (4.3.3) and (4.6.135), and the constant $C_0 := C_0(\tau,\nu) > 0$ is fixed in Theorem 4.2 in [7]. By Lemma 4.6.18, the remainder R_6 satisfies the bound (4.6.112), and using (4.5.7) we get

$$|R_6|_{s_0+\beta}^{Lip(\gamma)} \leq C\varepsilon^{7-2b}\gamma^{-1} = C\varepsilon^{3-2a}, \qquad |R_6|_{s_0+\beta}^{Lip(\gamma)}\gamma^{-1} \leq C\varepsilon^{1-3a}. \qquad (4.6.131)$$

We use that μ in (4.5.7) is assumed to satisfy $\mu \geq \sigma + \beta$ where $\sigma := \sigma(\tau,\nu)$ is given in Lemma 4.6.18.

Theorem 4.6.19. ***(Reducibility)*** *Assume that $\omega \mapsto i_\delta(\omega)$ is a Lipschitz function defined on some subset $\mathcal{O}_0 \subseteq \Omega_\varepsilon$ (recall (4.3.2)), satisfying (4.5.7) with $\mu \geq \sigma + \beta$ where $\sigma := \sigma(\tau, \nu)$ is given in Lemma 4.6.18 and $\beta := 7\tau + 6$. Then there exists $\delta_0 \in (0,1)$ such that, if*

$$N_0^{C_0} \varepsilon^{7-2b} \gamma^{-2} = N_0^{C_0} \varepsilon^{1-3a} \leq \delta_0, \quad \gamma := \varepsilon^{2+a}, \quad a \in (0, 1/6), \tag{4.6.132}$$

then

(i) ***(Eigenvalues)****. For all $\omega \in \Omega_\varepsilon$ there exists a sequence*

$$d_j^\infty(\omega) := d_j^\infty(\omega, i_\delta(\omega)) := -\tilde{m}_3(\omega)\, j^3 + \tilde{m}_1(\omega) j + r_j^\infty(\omega), \quad j \in S^c, \tag{4.6.133}$$

where $\tilde{m}_3, \tilde{m}_1$ coincide with the coefficients of $\mathcal{L}_6$ of (4.6.128) for all $\omega \in \mathcal{O}_0$. Furthermore, for all $j \in S^c$

$$|\tilde{m}_3 - 1|^{Lip(\gamma)} \leq C\varepsilon^2, \quad |\tilde{m}_1 - \varepsilon^2 c(\xi)|^{Lip(\gamma)} \leq C\varepsilon^{3-2a}, \quad |r_j^\infty|^{Lip(\gamma)} \leq C\varepsilon^{3-2a}, \tag{4.6.134}$$

for some $C > 0$. All the eigenvalues $\mathrm{i}d_j^\infty$ are purely imaginary. We define, for convenience, $d_0^\infty(\omega) := 0$.

(ii) ***(Conjugacy)****. For all ω in the set*

$$\begin{aligned} \Omega_\infty^{2\gamma} := \Omega_\infty^{2\gamma}(i_\delta) := \Big\{ \omega \in \mathcal{O}_0 : |\omega \cdot \ell + d_j^\infty(\omega) - d_k^\infty(\omega)| \geq \frac{2\gamma\, |j^3 - k^3|}{\langle \ell \rangle^\tau}, \\ \forall \ell \in \mathbb{Z}^\nu, \ \forall j, k \in S^c \cup \{0\} \Big\} \end{aligned} \tag{4.6.135}$$

there is a real, bounded, invertible, linear operator $\Phi_\infty(\omega) \colon H^s_{S^\perp}(\mathbb{T}^{\nu+1}) \to H^s_{S^\perp}(\mathbb{T}^{\nu+1})$, with bounded inverse $\Phi_\infty^{-1}(\omega)$, that conjugates $\mathcal{L}_6$ in (4.6.128) to constant coefficients, namely

$$\begin{aligned} \mathcal{L}_\infty(\omega) &:= \Phi_\infty^{-1}(\omega) \circ \mathcal{L}_6 \circ \Phi_\infty(\omega) = \omega \cdot \partial_\varphi + \mathcal{D}_\infty(\omega), \\ \mathcal{D}_\infty(\omega) &:= \mathrm{diag}_{j \in S^c} \{ \mathrm{i} d_j^\infty(\omega) \}. \end{aligned} \tag{4.6.136}$$

The transformations $\Phi_\infty, \Phi_\infty^{-1}$ are close to the identity in matrix decay norm, with

$$|\Phi_\infty^{\pm 1} - \mathrm{I}|^{Lip(\gamma)}_{s, \Omega_\infty^{2\gamma}} \leq_s \varepsilon^7 \gamma^{-3} + \varepsilon \gamma^{-1} \|\mathfrak{I}_\delta\|^{Lip(\gamma)}_{s+\sigma}. \tag{4.6.137}$$

Moreover $\Phi_\infty, \Phi_\infty^{-1}$ are symplectic, and $\mathcal{L}_\infty$ is a Hamiltonian operator.

Remark 4.6.20. Theorem 4.2 in [7] also provides the Lipschitz dependence of the (approximate) eigenvalues d_j^n with respect to the unknown $i_0(\varphi)$, which is used for the measure estimate in Lemma 4.7.3.

Observe that all the parameters $\omega \in \Omega_\infty^{2\gamma}$ satisfy also the first Melnikov condition, namely

$$|\omega \cdot \ell + d_j^\infty(\omega)| \geq 2\gamma |j|^3 \langle \ell \rangle^{-\tau}, \quad \forall \ell \in \mathbb{Z}^\nu, \ j \in S^c, \tag{4.6.138}$$

because, by definition, $\mu_0^\infty = 0$, and the diagonal operator $\mathcal{L}_\infty$ is invertible.
In the following theorem we verify the inversion assumption (4.4.29) for $\mathcal{L}_\omega$.

Theorem 4.6.21. *Assume the hypotesis of Theorem 4.6.19 and* (4.6.132)*. Then there exists* $\sigma_1 := \sigma_1(\tau,\nu) > 0$ *such that, for all* $\omega \in \Omega_\infty^{2\gamma}(i_\delta)$ *(see* (4.6.135)*), for any function* $g \in H_{S^\perp}^{s+\sigma_1}(\mathbb{T}^{\nu+1})$ *the equation* $\mathcal{L}_\omega h = g$ *has a solution* $h = \mathcal{L}_\omega^{-1} g \in H_{S^\perp}^{s}(\mathbb{T}^{\nu+1})$*, satisfying*

$$\begin{aligned}\|\mathcal{L}_\omega^{-1} g\|_s^{Lip(\gamma)} &\leq_s \gamma^{-1}(\|g\|_{s+\sigma_1}^{Lip(\gamma)} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma_1}^{Lip(\gamma)}\|g\|_{s_0}^{Lip(\gamma)}) \\ &\leq_s \gamma^{-1}(\|g\|_{s+\sigma_1}^{Lip(\gamma)} + \varepsilon\gamma^{-1}\{\|\mathfrak{I}_0\|_{s+\sigma_1+\sigma}^{Lip(\gamma)} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{Lip(\gamma)}\|Z\|_{s+\sigma_1+\sigma}^{Lip(\gamma)}\}\|g\|_{s_0}^{Lip(\gamma)}).\end{aligned} \tag{4.6.139}$$

Proof. We semi-conjugated the operator $\mathcal{L}_\omega$ in (4.5.33) to the diagonal operator $\mathcal{L}_\infty$ in (4.6.136) with the following transformations (recall Lemma 4.6.3, (4.6.53), (4.6.54), (4.6.75), (4.6.96), (4.6.113), (4.6.126), (4.6.136))

$$\mathcal{L}_\omega = \mathcal{M}_1\,\mathcal{L}_\infty\mathcal{M}_2^{-1}, \quad \mathcal{M}_1 := \Phi\,B\,\rho\,\mathcal{T}\,\Phi_1\,\Phi_2\,\mathcal{S}\,\Phi_\infty, \quad \mathcal{M}_2 := \Phi\,B\,\mathcal{T}\,\Phi_1\,\Phi_2\,\mathcal{S}\,\Phi_\infty, \tag{4.6.140}$$

where ρ means the multiplication for the function $\rho(\varphi)$ defined in (4.6.54). By (4.6.138) and Lemma 4.2 of [7] we get the bound

$$\|\mathcal{L}_\infty^{-1} g\|_s^{Lip(\gamma)} \leq_s \gamma^{-1}\|g\|_{s+2\tau+1}.$$

By Lemmata 4.6.5, 4.6.9, 4.6.10, 4.6.18, the bound (4.6.137) and the fact that $|\Phi_1^{\pm 1}|_s^{Lip(\gamma)} \leq C(s)$, $|\Phi_2^{\pm 1}|_s^{Lip(\gamma)} \leq C(s)$ (recall that the decay norm controls the Sobolev norm, see (2.3.3)) we get

$$\|\mathcal{M}_2 h\|_s^{Lip(\gamma)} + \|\mathcal{M}_1^{-1} h\|_s^{Lip(\gamma)} \leq_s \|h\|_{s+3}^{Lip(\gamma)} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma+3}^{Lip(\gamma)}\|h\|_{s_0}^{Lip(\gamma)}.$$

By using the bound above and (4.4.9) we obtain (4.6.139). □

4.7 The Nash-Moser nonlinear iteration

In this section we prove Theorem 4.3.2. It will be a consequence of the Nash-Moser theorem 4.7.1.

Consider the finite-dimensional subspaces

$$E_n := \{\mathfrak{I}(\varphi) = (\Theta, y, z)(\varphi) : \Theta = \Pi_n\Theta, y = \Pi_n y, z = \Pi_n z\}$$

where $N_n := N_0^{\chi^n}$ are introduced in (4.6.129), and Π_n are the projectors (which, with a small abuse of notation, we denote with the same symbol)

$$\begin{aligned}&\Pi_n\Theta(\varphi) := \sum_{|\ell|<N_n}\Theta_\ell\,e^{\mathrm{i}\ell\cdot\varphi},\ \Pi_n y(\varphi) := \sum_{|\ell|<N_n} y_\ell\,e^{\mathrm{i}\ell\cdot\varphi},\ \text{where } \Theta(\varphi) = \sum_{\ell\in\mathbb{Z}^\nu}\Theta_\ell\,e^{\mathrm{i}\ell\cdot\varphi},\ y(\varphi) = \sum_{\ell\in\mathbb{Z}^\nu} y_\ell\,e^{\mathrm{i}\ell\cdot\varphi}, \\ &\Pi_n z(\varphi,x) := \sum_{|(\ell,j)|<N_n} z_{\ell j}\,e^{\mathrm{i}(\ell\cdot\varphi + jx)}, \text{ where } \quad z(\varphi,x) = \sum_{\ell\in\mathbb{Z}^\nu, j\in S^c} z_{\ell j}\,e^{\mathrm{i}(\ell\cdot\varphi+jx)}.\end{aligned} \tag{4.7.1}$$

We define $\Pi_n^\perp = \mathrm{I} - \Pi_n$. The classical smoothing properties hold, namely, for all $\alpha, s \geq 0$,

$$\|\Pi_n\mathfrak{I}\|_{s+\alpha}^{Lip(\gamma)} \leq N_n^\alpha\|\mathfrak{I}_\delta\|_s^{Lip(\gamma)}, \quad \forall\mathfrak{I}(\omega) \in H^s, \quad \|\Pi_n^\perp\mathfrak{I}\|_s^{Lip(\gamma)} \leq N_n^{-\alpha}\|\mathfrak{I}\|_{s+\alpha}^{Lip(\gamma)}, \quad \forall\mathfrak{I}(\omega) \in H^{s+\alpha}. \tag{4.7.2}$$

We define the following constants

$$\begin{aligned} &\mu_1 := 3\mu + 9, && \alpha := 3\mu_1 + 1, && \alpha_1 := (\alpha - 3\mu)/2, \\ &k := 3(\mu_1 + \rho^{-1}) + 1, && \beta_1 := 6\mu_1 + 3\rho^{-1} + 3, && 0 < \rho < \frac{1-3a}{C_1(1+a)}. \end{aligned} \tag{4.7.3}$$

where $\mu := \mu(\tau,\nu) > 0$ is the "loss of regularity" given by the Theorem 4.4.36 and C_1 is fixed below. We note that the constants in (4.7.3) are the same of the ones defined in [8], but with a different (larger) μ.

Theorem 4.7.1. ***(Nash-Moser)*** *Assume that $f \in C^q$ with $q > S := s_0 + \beta_1 + \mu + 3$. Let $\tau \geq \nu + 2$. Then there exist $C_1 > \max\{\mu_1 + \alpha, C_0\}$ (where $C_0 := C_0(\tau,\nu)$ is the one in Theorem 4.6.19), $\delta_0 := \delta_0(\tau,\nu) > 0$ such that, if*

$$N_0^{C_1}\varepsilon^{b_*+1}\gamma^{-2} < \delta_0, \quad \gamma := \varepsilon^{2+a} = \varepsilon^{2b}, \quad N_0 := (\varepsilon\gamma^{-1})^{\rho}, \quad b_* = 6 - 2b, \tag{4.7.4}$$

then, for all $n \geq 0$:

$(\mathcal{P}1)_n$ *there exists a function $(\mathfrak{I}_n,\zeta_n)\colon \mathcal{G}_n \subseteq \Omega_\varepsilon \to E_{n-1} \times \mathbb{R}^\nu, \omega \mapsto (\mathfrak{I}_n(\omega),\zeta_n(\omega)), (\mathfrak{I}_0,\zeta_0) := 0, E_{-1} := \{0\}$, satisfying $|\zeta_n|^{Lip(\gamma)} \leq C\|\mathcal{F}(U_n)\|_{s_0}^{Lip(\gamma)}$,*

$$\|\mathfrak{I}_n\|_{s_0+\mu}^{Lip(\gamma)} \leq C_*\varepsilon^{b_*}\gamma^{-1}, \quad \|\mathcal{F}(U_n)\|_{s_0+\mu+3}^{Lip(\gamma)} \leq C_*\varepsilon^{b_*}, \tag{4.7.5}$$

where $U_n := (i_n,\zeta_n)$ with $i_n(\varphi) = (\varphi,0,0) + \mathfrak{I}_n(\varphi)$. The sets $\mathcal{G}_n$ are defined inductively by:

$$\begin{aligned} &\mathcal{G}_0 := \{\omega \in \Omega_\varepsilon : |\omega\cdot\ell| \geq 2\gamma\langle\ell\rangle^{-\tau}, \ \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\}, \\ &\mathcal{G}_{n+1} := \left\{\omega \in \mathcal{G}_n : |\omega\cdot\ell + d_j^\infty(i_n) - d_k^\infty(i_n)| \geq \frac{2\,\gamma_n\,|j^3-k^3|}{\langle\ell\rangle^\tau}, \ \forall j,k \in S^c\cup\{0\}, \ell \in \mathbb{Z}^\nu\right\}, \end{aligned} \tag{4.7.6}$$

where $\gamma_n := \gamma(1+2^{-n})$ and $d_j^\infty(\omega) := d_j^\infty(\omega, i_n(\omega))$ are defined in (4.6.133) (and $d_0^\infty(\omega) = 0$). The differences $\hat{\mathfrak{I}}_n := \mathfrak{I}_n - \mathfrak{I}_{n-1}$ (where we set $\hat{\mathfrak{I}}_0 := 0$) is defined on $\mathcal{G}_n$, and satisfy

$$\|\hat{\mathfrak{I}}_1\|_{s_0+\mu}^{Lip(\gamma)} \leq C_*\varepsilon^{b_*}\gamma^{-1}, \quad \|\hat{\mathfrak{I}}_n\|_{s_0+\mu}^{Lip(\gamma)} \leq C_*\varepsilon^{b_*}\gamma^{-1}N_{n-1}^{-\alpha}, \quad \forall n > 1. \tag{4.7.7}$$

$(\mathcal{P}2)_n$ $\|\mathcal{F}(U_n)\|_{s_0}^{Lip(\gamma)} \leq C_*\varepsilon^{b_*}N_{n-1}^{-\alpha}$ *where we set $N_{-1} := 1$.*

$(\mathcal{P}3)_n$ *(High Norms). $\|\mathfrak{I}_n\|_{s_0+\beta_1}^{Lip(\gamma)} \leq C_*\varepsilon^{b_*}\gamma^{-1}N_{n-1}^k$ and $\|\mathcal{F}(U_n)\|_{s_0+\beta_1}^{Lip(\gamma)} \leq C_*\varepsilon^{b*}N_{n-1}^k$.*

$(\mathcal{P}4)_n$ *(Measure). The measure of the "Cantor-like" sets $\mathcal{G}_n$ satisfies*

$$|\Omega_\varepsilon \setminus \mathcal{G}_0| \leq C_*\varepsilon^{2(\nu-1)}\gamma, \quad |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq C_*\varepsilon^{2(\nu-1)}\gamma N_{n-1}^{-1}. \tag{4.7.8}$$

All the Lip norms are defined on $\mathcal{G}_n$, namely $\|\cdot\|_s^{Lip(\gamma)} = \|\cdot\|_{s,\mathcal{G}_n}^{Lip(\gamma)}$.

Proof. Proof of $(\mathcal{P}_1)_0, (\mathcal{P}_2)_0, (\mathcal{P}_3)_0$. Recalling (4.3.7), we have, by the second estimate in (4.3.17),

$$\|\mathcal{F}(U_0)\|_s = \|\mathcal{F}((\varphi,0,0),0)\|_s = \|X_P(i_0)\|_s \leq_s \varepsilon^{6-2b}.$$

Hence the smallness conditions in $(\mathcal{P}_1)_0, (\mathcal{P}_2)_0, (\mathcal{P}_3)_0$ hold taking $C_* := C_*(s_0+\beta_1)$ large enough. *Assume that* $(\mathcal{P}_1)_n, (\mathcal{P}_2)_n, (\mathcal{P}_3)_n$ *hold for some* $n \geq 0$, *and prove* $(\mathcal{P}_1)_{n+1}, (\mathcal{P}_2)_{n+1}, (\mathcal{P}_3)_{n+1}$. By (4.7.3) and (4.7.4)

$$N_0^{C_1}\varepsilon^{b_*+1}\gamma^{-2} = N_0^{C_1}\varepsilon^{1-3a} = \varepsilon^{1-3a-\rho C_1(1+a)} < \delta_0$$

for ε small enough. If we take $C_1 \geq C_0$ then (4.6.132) holds. Moreover (4.7.5) imply (4.4.6), and so (4.5.7), and Theorem 4.6.21 applies. Hence the operator $\mathcal{L}_\omega := \mathcal{L}_\omega(\omega, i_n(\omega))$ defined in (4.5.33) is invertible for all $\omega \in \mathcal{G}_{n+1}$ and the last estimate in (4.6.139) holds. This means that the assumption (4.4.29) of Theorem 4.4.11 is verified with $\Omega_\infty = \mathcal{G}_{n+1}$. By Theorem 4.4.11 there exists an approximate inverse $\mathbf{T}_n(\omega) := \mathbf{T}_0(\omega, i_n(\omega))$ of the linearized operator $L_n(\omega) := d_{i,\zeta}\mathcal{F}(\omega, i_n(\omega))$, satisfying (4.4.37). By (4.7.4), (4.7.5)

$$\|\mathbf{T}_n g\|_s \leq_s \gamma^{-1}(\|g\|_{s+\mu} + \varepsilon\gamma^{-1}\{\|\mathfrak{I}_n\|_{s+\mu} + \gamma^{-1}\|\mathfrak{I}_n\|_{s_0+\mu}\|\mathcal{F}(U_n)\|_{s+\mu}\}\|g\|_{s_0+\mu}) \tag{4.7.9}$$

$$\|\mathbf{T}_n g\|_{s_0} \leq_{s_0} \gamma^{-1}\|g\|_{s_0+\mu} \tag{4.7.10}$$

and, by (4.4.38), using also (4.7.4), (4.7.5), (4.7.2),

$$\begin{aligned}\|(L_n\circ\mathbf{T}_n - \mathrm{I})g\|_s \leq_s &\gamma^{-1}(\|\mathcal{F}(U_n)\|_{s_0+\mu}\|g\|_{s+\mu} + \|\mathcal{F}(U_n)\|_{s+\mu}\|g\|_{s_0+\mu} \\ &+ \varepsilon\gamma^{-1}\|\mathfrak{I}_n\|_{s+\mu}\|\mathcal{F}(U_n)\|_{s_0+\mu}\|g\|_{s_0+\mu})\end{aligned} \tag{4.7.11}$$

$$\begin{aligned}\|(L_n\circ\mathbf{T}_n - \mathrm{I})g\|_{s_0} &\leq_{s_0}\gamma^{-1}\|\mathcal{F}(U_n)\|_{s_0+\mu}\|g\|_{s_0+\mu} \\ &\leq_{s_0}\gamma^{-1}(\|\Pi_n\mathcal{F}(U_n)\|_{s_0+\mu} + \|\Pi_n^\perp\mathcal{F}(U_n)\|_{s_0+\mu})\|g\|_{s_0+\mu} \\ &\leq_{s_0} N_n^\mu\gamma^{-1}(\|\mathcal{F}(U_n)\|_{s_0} + N_n^{-\beta_1}\|\mathcal{F}(U_n)\|_{s_0+\beta_1})\|g\|_{s_0+\mu}.\end{aligned} \tag{4.7.12}$$

The index β_1 in (4.7.3) is an ultraviolet cut, and it has to be define in order to obtain the convergence of the iteration scheme.
Now, for all $\omega \in \mathcal{G}_{n+1}$, we can define, for $n \geq 0$,

$$U_{n+1} := U_n + H_{n+1}, \quad H_{n+1} := (\hat{\mathfrak{I}}_{n+1}, \hat{\zeta}_{n+1}) := -\tilde{\Pi}_n\mathbf{T}_n\Pi_n\mathcal{F}(U_n) \in E_n \times \mathbb{R}^\nu, \tag{4.7.13}$$

where $\tilde{\Pi}_n(\mathfrak{I}, \zeta) := (\Pi_n\mathfrak{I}, \zeta)$ with Π_n defined in (4.7.1). Since $L_n := d_{i,\zeta}\mathcal{F}(i_n)$, we write

$$\mathcal{F}(U_{n+1}) = \mathcal{F}(U_n) + L_nH_{n+1} + Q_n,$$

where

$$Q_n := Q(U_n, H_{n+1}), \quad Q(U_n, H) := \mathcal{F}(U_n + H) - \mathcal{F}(U_n) - L_nH, \quad H \in E_n \times \mathbb{R}^\nu. \tag{4.7.14}$$

Then, by the definition of H_{n+1} in (4.7.13), using $[L_n, \Pi_n]$ and writing $\tilde{\Pi}_n^\perp(\mathfrak{I}, \zeta) := (\Pi_n^\perp\mathfrak{I}, 0)$ we have

$$\begin{aligned}\mathcal{F}(U_{n+1}) &= \mathcal{F}(U_n) - L_n\tilde{\Pi}_n\mathbf{T}_n\Pi_n\mathcal{F}(U_n) + Q_n \\ &= \mathcal{F}(U_n) - L_n\mathbf{T}_n\Pi_n\mathcal{F}(U_n) + L_n\tilde{\Pi}_n^\perp\mathbf{T}_n\Pi_n\mathcal{F}(U_n) + Q_n \\ &= \mathcal{F}(U_n) - \Pi_nL_n\mathbf{T}_n\Pi_n\mathcal{F}(U_n) + (L_n\tilde{\Pi}_n^\perp - \Pi_n^\perp L_n)\mathbf{T}_n\Pi_n\mathcal{F}(U_n) + Q_n \\ &= \Pi_n^\perp\mathcal{F}(U_n) + R_n + Q_n + Q_n'\end{aligned} \tag{4.7.15}$$

where

$$R_n := (L_n\tilde{\Pi}_n^\perp - \Pi_n^\perp L_n)\mathbf{T}_n\Pi_n\mathcal{F}(U_n), \quad Q_n' := -\Pi_n(L_n\mathbf{T}_n - \mathrm{I})\Pi_n\mathcal{F}(U_n). \tag{4.7.16}$$

Lemma 4.7.2. *(Lemma 9.2 in [8]) Define*

$$w_n := \varepsilon\gamma^{-2}\|\mathcal{F}(U_n)\|_{s_0}, \qquad B_n := \varepsilon\gamma^{-1}\|\mathfrak{I}_n\|_{s_0+\beta_1} + \varepsilon\gamma^{-2}\|\mathcal{F}(U_n)\|_{s_0+\beta_1}. \tag{4.7.17}$$

Then there exists $K := K(s_0, \beta_1) > 0$ *such that, for all* $n \geq 0$, *setting* $\mu_1 := 3\mu + 9$

$$w_{n+1} \leq K N_n^{\mu_1+\rho^{-1}-\beta_1} B_n + K N_n^{\mu_1} w_n^2, \qquad B_{n+1} \leq K N_n^{\mu_1+\rho^{-1}} B_n. \tag{4.7.18}$$

Proof of $(\mathcal{P}_3)_{n+1}$. By (4.7.18) and $(\mathcal{P}_3)_n$

$$B_{n+1} \leq K N_n^{\mu_1+\rho^{-1}} B_n \leq 2C_* K \varepsilon^{b_*+1}\gamma^{-2} N_n^{\mu_1+\rho^{-1}} N_{n-1}^{k} \leq C_*\varepsilon^{b_*+1}\gamma^{-2} N_n^{k}, \tag{4.7.19}$$

provided $2K N_n^{\mu_1+\rho^{-1}-k} N_{n-1}^{k} \leq 1, \forall n \geq 0$. Choosing k as in (4.7.3) and N_0 large enough, i.e. for ε small enough. By (4.7.17) and the bound (4.7.19) $(\mathcal{P}_3)_{n+1}$ holds.
Proof of $(\mathcal{P}_2)_{n+1}$. Using (4.7.17), (4.7.18) and $(\mathcal{P}_2)_n, (\mathcal{P}_3)_n$, we get

$$\begin{aligned} w_{n+1} &\leq K N_n^{\mu_1+\rho^{-1}-\beta_1} B_n + K N_n^{\mu_1} w_n^2 \leq K N_n^{\mu_1+\rho^{-1}-\beta_1} 2C_*\varepsilon^{b_*+1}\gamma^{-2} N_{n-1}^{k} \\ &\quad + K N_n^{\mu_1} (C_*\varepsilon^{b_*+1}\gamma^{-2} N_{n-1}^{-\alpha})^2 \end{aligned}$$

and $w_{n+1} \leq C_*\varepsilon^{b_*+1}\gamma^{-2} N_n^{-\alpha}$ provided that

$$4K N_n^{\mu_1+\rho^{-1}-\beta_1+\alpha} N_{n-1}^{k} \leq 1, \quad 2KC_*\varepsilon^{b_*+1}\gamma^{-2} N_n^{\mu_1+\alpha} N_{n-1}^{-2\alpha} \leq 1, \ \forall n \geq 0. \tag{4.7.20}$$

The inequalities in (4.7.20) hold by (4.7.4), taking α as in (4.7.3), $C_1 > \mu_1 + \alpha$ and δ_0 in (4.7.4) small enough. By (4.7.17), the inequality $w_{n+1} \leq C_*\varepsilon^{b_*+1}\gamma^{-2} N_n^{-\alpha}$ implies $(\mathcal{P}_2)_{n+1}$.
Proof of $(\mathcal{P}_1)_{n+1}$. The bound (4.7.7) for $\hat{\mathfrak{I}}_1$ follows by (4.7.13), (4.7.9) (for $s = s_0 + \mu$) and $\|\mathcal{F}(U_0)\|_{s_0+2\mu} = \|\mathcal{F}((\varphi, 0, 0), 0)\|_{s_0+2\mu} \leq_{s_0+2\mu} \varepsilon^{b_*}$. The bound (4.7.7) for $\hat{\mathfrak{I}}_{n+1}$ follows by (4.7.1), $(\mathcal{P}_2)_n$ and (4.7.3). It remains to prove that (4.7.5) holds at the step $n+1$. We have

$$\|\mathfrak{I}_{n+1}\|_{s_0+\mu} \leq \sum_{k=1}^{n+1} \|\hat{\mathfrak{I}}_k\|_{s_0+\mu} \leq C_*\varepsilon^{b_*}\gamma^{-1} \sum_{k\geq 1} N_{k-1}^{-\alpha_1} \leq C_*\varepsilon^{b_*}\gamma^{-1} \tag{4.7.21}$$

taking α_1 as in (4.7.3) and N_0 large enough, i.e. ε small enough. Moreover, using (4.7.1), $(\mathcal{P}_2)_{n+1}, (\mathcal{P}_3)_{n+1}$, (4.7.3) we get

$$\begin{aligned} \|\mathcal{F}(U_{n+1})\|_{s_0+\mu+3} &\leq N_n^{\mu+3} \|\mathcal{F}(U_{n+1})\|_{s_0} + N_n^{\mu+3-\beta_1} \|\mathcal{F}(U_{n+1})\|_{s_0+\beta_1} \\ &\leq C_*\varepsilon^{b_*} N_n^{\mu+3-\alpha} + C_*\varepsilon^{b_*} N_n^{\mu+3-\beta_1+k} \leq C_*\varepsilon^{b_*}, \end{aligned}$$

which is the second inequality in (4.7.5) at the step $n+1$. The bound

$$|\zeta_{n+1}|^{Lip(\gamma)} \leq C \|\mathcal{F}(U_{n+1})\|_{s_0}^{Lip(\gamma)}$$

is a consequence of Lemma (4.4.1).

4.7.1 Measure estimates

In this section we prove $(\mathcal{P}_4)_n$ for all $n \geq 0$. Fixed $n \in \mathbb{N}$, we have

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} = \bigcup_{\ell \in \mathbb{Z}^\nu, j,k \in S^c \cup \{0\}} R_{\ell jk}(i_n) \tag{4.7.22}$$

where

$$R_{\ell jk}(i_n) := \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + d_j^\infty(i_n) - d_k^\infty(i_n)| < 2\,\gamma_n\, |j^3 - k^3| \langle \ell \rangle^{-\tau}\}. \tag{4.7.23}$$

Since, by (4.3.3), $R_{\ell jk}(i_n) = \emptyset$ for $j = k$, in the sequel we assume that $j \neq k$.

Lemma 4.7.3. *(Lemma 9.3 in [8]) For $n \geq 1, |\ell| \leq N_{n-1}$, one has the inclusion $R_{\ell jk}(i_n) \subseteq R_{\ell jk}(i_{n-1})$.*

By definition, $R_{\ell jk}(i_n) \subseteq \mathcal{G}_n$ (see (4.7.23)). By Lemma 4.7.3, for $n \geq 1$ and $|\ell| \leq N_{n-1}$ we also have $R_{\ell jk}(i_n) \subseteq R_{\ell jk}(i_{n-1})$. On the other hand, $R_{\ell jk}(i_n) \cap \mathcal{G}_n = \emptyset$ (see (4.7.6)). As a consequence, $R_{\ell jk}(i_n) = \emptyset$ for all $|\ell| \leq N_{n-1}$, and

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} \subseteq \bigcup_{\substack{j,k \in S^c \cup \{0\} \\ |\ell| > N_{n-1}}} R_{\ell jk}(i_n) \quad \forall n \geq 1. \tag{4.7.24}$$

Lemma 4.7.4. *Let $n \geq 0$. If $R_{\ell jk}(i_n) \neq \emptyset$, then $|\ell| \geq C_1 |j^3 - k^3| \geq \frac{C_1}{2}(j^2 + k^2)$ for some constant $C_1 > 0$ (independent of $\ell, j, k, n, i_n, \omega$).*

By Lemma 4.7.4 it is sufficient to study the measure of the resonant sets $R_{\ell jk}(i_n)$ defined in (4.7.23) for $(\ell, j, k) \neq (0, j, j)$. In particular we will prove the following Lemma.

Lemma 4.7.5. *For all $n \geq 0$ and for a generic choice of the tangential sites, the measure*

$$|R_{\ell jk}(i_n)| \leq C \varepsilon^{2(\nu-1)} \gamma \langle \ell \rangle^{-\tau}.$$

By (4.7.23), we have to bound the measure of the sublevels of the function $\omega \mapsto \phi(\omega)$ defined by

$$\phi(\omega) : = \omega \cdot \ell + d_j^\infty(\omega) - d_k^\infty(\omega) = \omega \cdot \ell - m_3(\omega)(j^3 - k^3) + m_1(j - k) + (r_j^\infty - r_k^\infty)(\omega) \tag{4.7.25}$$

Note that ϕ also depends on ℓ, j, k, i_n. We recall that

$$m_3 = 1 + \varepsilon^2 d(\xi) + \mathtt{r}_{m_3}(\omega), \qquad m_1 = \varepsilon^2 c(\xi) + \mathtt{r}_{m_1}(\omega) \tag{4.7.26}$$

where

$$|\mathtt{r}_{m_3}|^{Lip(\gamma)} \leq C \varepsilon^3 \qquad |\mathtt{r}_{m_1}|^{Lip(\gamma)} \leq C \varepsilon^{3-2a} \tag{4.7.27}$$

and $d(\xi)$, $c(\xi)$ are defined in (4.6.61) and (4.6.122) respectively.

It will be useful to consider $\phi(\omega)$ in (4.7.25) as a small perturbation of an affine function in ω. We write it as

$$\phi(\omega) := a_{jk} + b_{\ell jk} \cdot \omega + q_{jk}(\omega), \qquad \ell \in \mathbb{Z}^\nu,\ j, k \in S^c, \tag{4.7.28}$$

where, by (4.2.18), (4.6.61), (4.6.122),

$$a_{jk} := -\{(j^3-k^3)[1-d(\mathbb{M}^{-1}\overline{\omega})]+(j-k)c(\mathbb{M}^{-1}\overline{\omega})\}, \tag{4.7.29}$$

$$\begin{aligned} b_{\ell jk} :=& \{\ell-(j^3-k^3)[(24c_4-48c_1^2)\mathbb{M}^{-T}v_3+(4c_6-\frac{16}{3}c_2^2)\mathbb{M}^{-T}v_1], \\ &+(j-k)[(-4c_6+\frac{16}{3}c_2^2)\mathbb{M}^{-1}v_3-(24c_7-16c_2c_3)\mathbb{M}^{-1}v_1]\} \end{aligned} \tag{4.7.30}$$

$$q_{jk}(\omega) := -\,\mathtt{r}_{m_3}(\omega)\,(j^3-k^3)+\mathtt{r}_{m_1}(\omega)\,(j-k)+r_j^\infty(\omega)-r_k^\infty(\omega) \tag{4.7.31}$$

and by (4.6.83), (4.7.27), (4.7.31),

$$\begin{aligned} |q_{jk}(\omega)|^{sup} &\leq \varepsilon^3|j^3-k^3|+\varepsilon^{3-2a}|j-k|+\varepsilon^{3-2a}, \\ |q_{jk}(\omega)|^{lip} &\leq |\mathtt{r}_{m_3}(\omega)|^{lip}|j^3-k^3|+|\mathtt{r}_{m_1}(\omega)|^{lip}|j-k|+|r_j^\infty-r_k^\infty|^{lip} \\ &\leq \varepsilon^3\gamma^{-1}|j^3-k^3|+\varepsilon^{3-2a}\gamma^{-1}|j-k|+\varepsilon^{1-3a}. \end{aligned} \tag{4.7.32}$$

Remark 4.7.6. The idea of the proof of Lemma 4.7.5 is that *generically* (see Definition 1.1.2) a_{jk} has to be sufficiently far from zero or the modulus of the "derivative" b_{ljk} has to be big enough.

We shall use the following *non-degeneracy* assumptions

$$(\mathtt{H1}) \qquad d(\xi)-1\neq 0 \quad \text{at} \quad \xi=\mathbb{M}^{-1}\overline{\omega}, \tag{4.7.33}$$

$$(\mathtt{H2})_{j,k} \quad \text{Fixed } j,k\in S^c,\ j\neq k, \quad \det(\mathbb{M}+B(j,k))\neq 0, \tag{4.7.34}$$

where

$$\begin{aligned} B(j,k) :=& -(24c_4-48c_1^2+\frac{12c_6-16c_2^2}{3(j^2+k^2+jk)})D_S^3UD_S^3 \\ &+(\frac{16c_2^2}{3}-4c_6+\frac{(16c_2c_3-24c_7)}{j^2+k^2+jk})D_SUD_S^3. \end{aligned} \tag{4.7.35}$$

In the next lemmata we prove that if the coefficients $c_1,\dots,c_7$ are non-resonant and conditions (C1)-(C2) hold, then there exist a generic choice of the tangential sites for which Lemma 4.7.10 and Lemma 4.7.13 hold true.

Lemma 4.7.7. *Fix $\nu\in\mathbb{N}$. If the coefficients $c_1,\dots,c_7$ are non-resonant and*

$$(7-16\nu)\,c_2^2\neq 6\,(1-2\nu)c_6 \tag{4.7.36}$$

then the polynomial $P(\bar{\jmath}_1,\dots,\bar{\jmath}_\nu) := d(\mathbb{M}^{-1}\overline{\omega})-1$ is not identically zero. As a consequence, the assumption (H1) *is verified for a generic choice of the tangential sites.*

Proof. Suppose that $d(\mathbb{M}^{-1}\overline{\omega})=1$, namely

$$P(\bar{\jmath}_1,\dots,\bar{\jmath}_\nu) := \{(24c_4-48\,c_1^2)v_3+(4c_6-\frac{16}{3}c_2^2)v_1\}\cdot\mathbb{M}^{-1}\overline{\omega}-1=0. \tag{4.7.37}$$

We evaluate the polynomial P at the point $(\bar{\jmath}_1,\dots,\bar{\jmath}_\nu)=\lambda(1,\dots,1)=\lambda\vec{1}$, for some λ to be determined, and we claim that this is not a zero. This implies that the polinomial P in (4.7.37) cannot be identically zero. We have

$$P(\lambda\vec{1})=\{\lambda^5\,(24c_4-48\,c_1^2)+\lambda^3\,(4c_6-\frac{16}{3}c_2^2)\}\,(\vec{1}\cdot\mathbb{M}(\lambda\vec{1})^{-1}\vec{1})-1$$

and $\mathbb{M}(\lambda\vec{1}) = a(\lambda)\mathrm{I} + b(\lambda)\,U$, where

$$a(\lambda) := (24c_1^2 - 12c_4)\lambda^6 + (\frac{14}{3}c_2^2 - 4c_6)\lambda^4 + (12c_2c_3 - 12c_7)\,\lambda^2 - 6\,c_3^2, \tag{4.7.38}$$

$$b(\lambda) := (-48c_1^2 + 24c_4)\lambda^6 + (-\frac{32}{3}c_2^2 + 8c_6)\lambda^4 + (-16c_2c_3 + 24c_7)\,\lambda^2. \tag{4.7.39}$$

We note that $a(\lambda) \neq 0$, because the coefficients are non-resonant. Moreover, by assumption (4.7.36) $a(\lambda) + \nu b(\lambda) \neq 0$ and we have

$$(\mathbb{M}(\lambda\vec{1}))^{-1} = \frac{\mathrm{I}}{a(\lambda)} - \frac{b(\lambda)}{a(\lambda)\,(a(\lambda) + b(\lambda)\,\nu)}U$$

and, by $\vec{1}\cdot\vec{1} = \nu$, $\vec{1}\cdot U\vec{1} = \nu^2$, we get

$$\vec{1}\cdot\mathbb{M}(\lambda\vec{1})^{-1}\vec{1} = \frac{\nu}{a(\lambda) + b(\lambda)\nu}. \tag{4.7.40}$$

Then $P(\lambda\vec{1}) = 0$ is equivalent to $p(\lambda) = 0$, where

$$\begin{aligned} p(\lambda) : &= \lambda^6\{24c_1^2 - 12c_4\} + \lambda^4\{(\frac{14}{3} - \frac{16\,\nu}{3})c_2^2 - 4(1-\nu)c_6\} \\ &+ \lambda^2\{(12 - 16\nu)c_2c_3 - 12(1-2\nu)c_7\} - 6c_3^2. \end{aligned}$$

Suppose that $c_3 \neq 0$, then $p(\lambda)$ is not trivial. If $c_3 = 0$ and $2c_1^2 \neq c_4$ then we conclude the same, because the monomial of degree six is not naught. If $c_3 = 0, 2c_1^2 = c_4$ then the monomial of minimum degree, namely three, it is not zero if $c_7 \neq 0$, indeed $\nu \in \mathbb{N}$. Suppose now that $c_3 = c_7 = 0, 2c_1^2 = c_4$. Eventually, by assumption (4.7.36) the monomial of maximum degree, namely four, is not naught and we conclude. □

Lemma 4.7.8. *Fix $\nu \in \mathbb{N}$. If $c_1, \ldots, c_7$ are non-resonant and*

$$\nu\,\frac{3c_6 - 4c_2^2}{9c_4 - 18c_1^2} \notin \{j^2 + k^2 + jk \,:\, j,k \in \mathbb{Z}\setminus\{0\},\, j \neq k\}, \tag{4.7.41}$$

then the polynomials $P_{jk}(\bar{\jmath}_1, \ldots, \bar{\jmath}_\nu) := \det(\mathbb{M} + B(j,k))$ are not identically zero, for all $j,k \in S^c$, $j \neq k$.

Proof. By (4.7.35) we have

$$\begin{aligned} \mathbb{M} + B(j,k) &= (24c_1^2 - 12c_4)D_S^6 + (\frac{14}{3}c_2^2 - 4c_6)D_S^4 \\ &+ (4c_6 - \frac{16}{3}c_2^2)D_S^3 U D_S(\mathrm{I} - \frac{1}{j^2+k^2+jk}D_S^2) \\ &+ 12(c_2c_3 - c_7)D_S^2 - 6c_3^2\mathrm{I} + (16c_2c_3 - 24c_7)D_S U D_S(\mathrm{I} - \frac{1}{j^2+k^2+jk}D_S^2). \end{aligned}$$

If $c_3 \neq 0$ then the lowest order monomial of $\det(\mathbb{M}+B(j,k))$ is not zero and the same holds if $c_3 = 0$ and $c_7 \neq 0$. If $c_3 = c_7 = 0$ then the monomial of maximal degree is

$$D_S^3\left((24c_1^2 - 12c_4)\mathrm{I} + \frac{12c_6 - 16c_2^2}{3(j^2+k^2+jk)}U\right)D_S^3$$

and this is invertible if (4.7.41) holds. □

Remark 4.7.9. By Lemma 4.7.8, if (C2) holds, then the assumptions $(\mathtt{H2})_{j,k}$ are satisfied by a generic choice of the tangential sites when j,k vary in a finite set of integers.

The rest of the section is devoted to the proof of Lemma 4.7.5.

Lemma 4.7.10. *Assume* (H1). *Then, for a generic choice of the tangential sites, there exists* $C_0>0$ *such that for all* $j\neq k$, $j,k\in S^c$, *with* $j^2+k^2>C_0$ *and* $\ell\in\mathbb{Z}^\nu$, *we have* $|R_{\ell jk}|\leq C\varepsilon^{2(\nu-1)}\gamma\langle\ell\rangle^{-\tau}$.

Proof. If $j^2+k^2>C_0$ for some constant C_0, then $|j-k|/|j^3-k^3|\leq 2C_0^{-1}$ and

$$|a_{jk}|\geq|j^3-k^3|\left\{|1-d(\mathbb{M}^{-1}\overline{\omega})|-\frac{2}{C_0}\,|c(\mathbb{M}^{-1}\overline{\omega})|\right\}.$$

If $d(\mathbb{M}^{-1}\overline{\omega})\neq 1$ then, by taking C_0 large enough, we get $|a_{jk}|\geq\delta_0|j^3-k^3|$, for some $\delta_0>0$. This implies that for $\delta:=\delta_0/2$ we have $|b_{\ell jk}\cdot\omega|\geq\delta|j^3-k^3|$. Indeed, by (4.7.23), (4.7.32)

$$|b_{\ell jk}\cdot\omega|\geq|a_{jk}|-|\phi(\omega)|-|q_{jk}(\omega)|\geq(\delta_0-2\gamma_n-|q_{jk}(\omega)|^{sup})|j^3-k^3|\geq\frac{\delta_0}{2}|j^3-k^3|,$$

for ε small enough (recall that $\gamma_n=o(\varepsilon^2)$).
If $b:=b_{\ell jk}$ we have $|b\cdot\omega|\leq 2|b||\overline{\omega}|$, because $|\omega|\leq 2|\overline{\omega}|$. Hence $|b|\geq\delta_1\,|j^3-k^3|$ where $\delta_1:=\delta/(2|\overline{\omega}|)$. Split $\omega=s\hat{b}+v$ where $\hat{b}:=b/|b|$ and $v\cdot b=0$. Let $\Psi(s):=\phi(s\hat{b}+v)$. For ε small enough, by (4.7.32), we get

$$\begin{aligned}|\Psi(s_1)-\Psi(s_2)|&\geq(|b|-|q_{jk}|^{lip})|s_1-s_2|\geq\left(\delta_1-\frac{|q_{jk}|^{lip}}{|j^3-k^3|}\right)|j^3-k^3|\,|s_1-s_2|\\&\geq\frac{\delta_1}{2}|j^3-k^3|\,|s_1-s_2|.\end{aligned}$$

As a consequence, the set $\Delta_{\ell jk}(i_n):=\{s:s\hat{\ell}+v\in R_{\ell jk}(i_n)\}$ has Lebesgue measure

$$|\Delta_{\ell jk}(i_n)|\leq\frac{2}{\delta_1\,|j^3-k^3|}\,\frac{4\,\gamma_n\,|j^3-k^3|}{\langle\ell\rangle^\tau}\leq\frac{C\,\gamma}{\langle\ell\rangle^\tau}$$

for some $C>0$. The Lemma follows by Fubini's theorem. □

Lemma 4.7.11. *There exists* $M>0$ *such that for all* $j\neq k$, $j,k\in S^c$, *with* $j^2+k^2\leq C_0$ *(see Lemma 4.7.10) and* $|\ell|\geq M$, *we have* $|R_{\ell jk}|\leq C\varepsilon^{2(\nu-1)}\gamma\langle\ell\rangle^{-\tau}$.

Proof. For $\ell\neq 0$, we decompose $\omega=s\hat{\ell}+v$, where $\hat{\ell}:=\ell/|\ell|, s\in\mathbb{R}$, and $\ell\cdot v=0$. Let $\psi(s):=\phi(s\hat{\ell}+v)$. We remark that $c(\xi)$ and $d(\xi)$ are affine functions of the unperturbed actions ξ, hence

$$\varepsilon^2|c(\xi)|^{lip},\varepsilon^2|d(\xi)|^{lip}\leq K$$

for some constant K depending only on the tangential sites and on the real coefficients $c_1,\ldots,c_7$. Then

$$\begin{aligned}&|\tilde{m}_3(s_1)-\tilde{m}_3(s_2)|\leq K|s_1-s_2|,\\&|\tilde{m}_1(s_1)-\tilde{m}_1(s_2)|\leq(K+\varepsilon^{3-2a}\gamma^{-1})|s_1-s_2|\leq 2K\,|s_1-s_2|,\\&|r_j^\infty(s_1)-r_j^\infty(s_2)|\leq\varepsilon^{3-2a}\gamma^{-1}|s_1-s_2|.\end{aligned}$$

Then, if we take M large enough and ε small, we have

$$
\begin{aligned}
|\psi(s_1)-\psi(s_2)| &\geq |j^3-k^3|\left(\frac{|\ell|}{|j^3-k^3|}-K-\frac{2\,K}{|j^2+k^2+jk|}-\frac{\varepsilon^{3-2a}\gamma^{-1}}{|j^3-k^3|}\right)|s_1-s_2| \\
&\geq \frac{\delta}{4}|j^3-k^3|\,|s_1-s_2|,
\end{aligned}
$$

where δ is a positive constant. Indeed, C_0 and K are fixed and it is sufficient to choose $|\ell|$ such that

$$
\inf_{j\neq k, j^2+k^2\leq C_0}\frac{|\ell|}{|j^3-k^3|}-K-\frac{2K}{C_0}\geq \delta>0.
$$

As a consequence, the set $\Delta_{\ell jk}(i_n) := \{s : s\hat{\ell}+v \in R_{\ell jk}(i_n)\}$ has Lebesgue measure

$$
|\Delta_{\ell jk}(i_n)| \leq \frac{\delta}{|j^3-k^3|}\frac{\gamma_n\,|j^3-k^3|}{\langle \ell\rangle^{\tau}} \leq \frac{C\,\gamma}{\langle \ell\rangle^{\tau}}
$$

for some $C>0$. The Lemma follows by Fubini's theorem. □

It remains to investigate $R_{\ell jk}$ for a finite set of indeces (ℓ, j, k). We need the following Lemma.

Lemma 4.7.12. *Suppose that $\ell\in\mathbb{Z}^\nu$ and $j,k\in S^c$ are such that $\overline{\omega}\cdot\ell\neq j^3-k^3$ and*

$$
|\ell|\leq M, \qquad |j^2+k^2|\leq C_0 \tag{4.7.42}
$$

for some positive constants M and C_0. Then R_{ljk} is empty.

Proof. We have

$$
\begin{aligned}
|\omega\cdot\ell+m_3(k^3-j^3)| &= |m_3(\overline{\omega}\cdot\ell+k^3-j^3)+(\omega-m_3\overline{\omega})\cdot\ell| \geq |m_3||\overline{\omega}\cdot\ell+k^3-j^3| \\
&\quad -|\omega-m_3\overline{\omega}||\ell| \geq 1-|\omega-\overline{\omega}|M-|m_3-1||\overline{\omega}|M\geq 1/2
\end{aligned}
$$

for ε small enough, because $|\omega-\overline{\omega}|, |m_3-1|\leq C\varepsilon^2$. Thus, by (4.7.25) we have

$$
\begin{aligned}
|\phi(\omega)| &\geq 1/2-\varepsilon^2|c(\xi)||j-k|-|m_1-\varepsilon^2 c(\xi)||j-k|-|r_j-r_k| \\
&\geq 1/2-\varepsilon^2\sup_{\xi\in[1,2]^\nu}(|c(\xi)|)-2C\,\varepsilon^{3-2a}\geq 1/4.
\end{aligned}
$$

□

Lemma 4.7.13. *If $j^2+k^2\leq C_0$, $j,k\in S^c$, $|\ell|\leq M$ (see Lemma 4.7.10 and 4.7.11) and $(\mathtt{H2})_{j,k}$ hold, then, for a generic choice of the tangential sites, $|R_{\ell jk}|\leq C\varepsilon^{2(\nu-1)}\gamma\langle\ell\rangle^{-\tau}$.*

Proof. We can write (4.7.25) as an affine function respect to the parameter ξ as

$$
\begin{aligned}
\phi(\xi) &= \overline{\omega}\cdot\ell-(j^3-k^3)+\varepsilon^2\{\mathbb{M}\xi\cdot\ell-d(\xi)(j^3-k^3)+c(\xi)(j-k)\}+q_{jk}(\alpha(\xi)), \\
q_{jk}(\alpha(\xi)) &= -\mathtt{r}_{m_3}(\alpha(\xi))(j^3-k^3)+\mathtt{r}_{m_1}(\alpha(\xi))(j-k)+r_j^\infty(\alpha(\xi))-r_k^\infty(\alpha(\xi)).
\end{aligned} \tag{4.7.43}
$$

By the relation (4.2.18), we can estimate the Lipschitz constant of $\phi(\omega)$ with the derivative respect to ξ of the expression (4.7.43).
By Lemma 4.7.12, we consider the case $\overline{\omega} \cdot \ell = j^3 - k^3$. Thus

$$\begin{aligned}\phi(\xi) =& \varepsilon^2[\mathbb{M}\xi \cdot \ell - d(\xi)\,\overline{\omega} \cdot \ell + c(\xi)(j-k)] + q_{jk}(\alpha(\xi)) \\ =& \varepsilon^2[\mathbb{M}\xi \cdot \ell - d(\xi)\overline{\omega} \cdot \ell + \frac{c(\xi)}{j^2+k^2+jk}\overline{\omega} \cdot \ell] + q_{jk}(\alpha(\xi)) \\ =& \varepsilon^2[\mathbb{M} + B(j,k)]\,\ell \cdot \xi + q_{jk}(\alpha(\xi)).\end{aligned} \tag{4.7.44}$$

where $B(j,k)$ is defined in (4.7.35). By assumption $(\mathtt{H2})$, if $l \neq 0$, then

$$\delta_{\ell jk} := (\mathbb{M} + B(j,k))\ell \neq 0. \tag{4.7.45}$$

Hence, by (4.7.32), (4.7.44) and (4.7.45), for ε small enough, there exist a constant $C > 0$ such that

$$|\phi|^{lip} \geq \delta_{\ell jk} - |q_{jk}|^{lip} \geq C|j^3 - k^3|.$$

Then we conclude as in Lemma 4.7.11. □

We have that Lemmata 4.7.10, 4.7.11, 4.7.13 implies Lemma 4.7.5. By (4.7.22) and Lemma 4.7.5 we get

$$|\mathcal{G}_0 \setminus \mathcal{G}_1| \leq \sum_{\ell \in \mathbb{Z}^\nu, |j|,|k| \leq C|\ell|^{1/2}} |R_{\ell jk}(i_0)| \leq \sum_{\ell \in \mathbb{Z}^\nu} \frac{C\,\varepsilon^{2(\nu-1)}\gamma}{\langle \ell \rangle^{\tau-1}} \leq C'\varepsilon^{2(\nu-1)}\gamma.$$

For $n \geq 1$, by (4.7.24),

$$|\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq \sum_{\substack{|\ell| > N_{n-1}, \\ |j|,|k| \leq C|\ell|^{1/2}}} |R_{\ell jk}(i_n)| \leq \sum_{|\ell| > N_{n-1}} \frac{C\,\varepsilon^{2(\nu-1)}\gamma}{\langle \ell \rangle^{\tau-1}} \leq C'\varepsilon^{2(\nu-1)}\gamma\, N_{n-1}^{-1}$$

because $\tau \geq \nu + 2$. The estimate $|\Omega_\varepsilon \setminus \mathcal{G}_0| \leq C\,\varepsilon^{2(\nu-1)}\gamma$ is elementary.

□

Conclusion of the Proof of Theorem 4.3.2. Theorem 4.7.1 implies that the sequence $(\mathfrak{I}_n, \zeta_n)$ is well defined for $\omega \in \mathcal{G}_\infty := \cap_{n\geq 0}\mathcal{G}_n$, and $\mathfrak{I}_n$ is a Cauchy sequence in $\|\cdot\|^{Lip(\gamma)}_{s_0+\mu,\mathcal{G}_\infty}$, see (4.7.7), and $|\zeta_n|^{Lip(\gamma)} \to 0$. Therefore $\mathfrak{I}_n$ converges to a limit $\mathfrak{I}_\infty$ in norm $\|\cdot\|^{Lip(\gamma)}_{s_0+\mu,\mathcal{G}_\infty}$ and, by $(\mathcal{P}2)_n$, for all $\omega \in \mathcal{G}_\infty, i_\infty(\varphi) := (\varphi, 0, 0) + \mathfrak{I}_\infty(\varphi)$, is a solution of

$$\mathcal{F}(i_\infty, 0) = 0 \qquad \text{with} \qquad \|\mathfrak{I}_\infty\|^{Lip(\gamma)}_{s_0+\mu,\mathcal{G}_\infty} \leq C\,\varepsilon^{6-2b}\gamma^{-1}$$

by (4.7.5). Therefore $\varphi \mapsto i_\infty(\varphi)$ is an invariant torus for the Hamiltonian vector field X_{H_ε} (recall (4.2.19)). By (4.7.8),

$$|\Omega_\varepsilon \setminus \mathcal{G}_\infty| \leq |\Omega_\varepsilon \setminus \mathcal{G}_0| + \sum_{n\geq 0}|\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \leq 2\,C_*\varepsilon^{2(\nu-1)}\gamma + C_*\varepsilon^{2(\nu-1)}\gamma \sum_{n\geq 1} N_{n-1}^{-1} \leq C\varepsilon^{2(\nu-1)}\gamma.$$

The set Ω_ε in (4.3.2) has measure $|\Omega_\varepsilon| = O(\varepsilon^{2\nu})$. Hence $|\Omega_\varepsilon \setminus \mathcal{G}_\infty|/|\Omega_\varepsilon| \to 0$ as $\varepsilon \to 0$ because $\gamma = o(\varepsilon^2)$, and therefore the measure of $\mathcal{C}_\varepsilon := \mathcal{G}_\infty$ satisfies (4.3.12).

CHAPTER 5

Quasi periodic solutions for Hamiltonian perturbations of the Degasperis-Procesi equation

In this Chapter we prove Theorem 1.2.3. In Section 5.1 we introduce the aspects of the integrable structure of the DP equation which we use to perform the Birkhoff normal form of Section 5.2.
In Section 5.2 we perform six steps of weak Birkhoff normal form in order to extract parameters, which modulate the frequency-amplitude relation (5.3.8), and to provide a good first nonlinear approximation of the solutions.
In Sections 5.3 and 5.4 we introduce action-angle variables (5.3.10) and we reformulate the problem of finding quasi-periodic solutions as the search for the zeros of the nonlinear functional $\mathcal{F}$ defined in (5.4.8). Adopting this new point of view, we devote the rest of the Chapter to the proof of Theorem 5.4.1, which implies Theorem 1.2.3.
In Section 5.5 we describe the construction of the approximate inverse for the linearized operator (5.5.2) following the abstract procedure developed in [21]. Thus the main issue is the approximate inversion of the linearized equations restricted at the normal directions, or equivalently the approximate inversion of the operator $\mathcal{L}_\omega$ in (5.6.31), which acts on the normal variables space $H_S^\perp$.
In Section 5.6 we prove that $\mathcal{L}_\omega$ has the form (5.6.31). In Sections 5.7 and 5.8 we conjugate $\mathcal{L}_\omega$ to a a diagonal operator $\mathcal{L}_\infty$ (see Theorem 5.8.1) and we provide tame estimates for the inverse of $\mathcal{L}_\omega$ (see Section 5.8.2).
In Section 5.9 we implement the Nash-Moser scheme of Theorem 5.9.2 to the functional $\mathcal{F}$ (recall (5.4.8)). In Section 5.9.1 we prove the measure estimates. This concludes the proof of Theorem 1.2.3.

5.1 Integrable structure of the DP equation

In [48] the authors proved the existence of a Lax Pair for the Degasperis-Procesi equation. As a consequence, infinitely many commuting constants of motions are produced by using the power series expansion of a parameter of the spectral problem.
By following the procedure described in the Section 4 of [48] we generate some of these symmetries and we use them in order to prove that, at any step of the Birkhoff normal form, there are not small divisors.
In order to do that we have only to consider the quadratic parts of the constants of motion. We list

here the ones that we used. For the complete expressions we refer to Appendix C.

$$
\begin{aligned}
K_0(u) &:= H(u)\\
K_1(u) &:= \frac{1}{2}\int_{\mathbb{T}} (J^{-1}u_x)\, u\, dx,\\
K_2(u) &:= \frac{1}{9}\int_{\mathbb{T}} w^2\, dx + O(u^3),\\
K_3(u) &:= \frac{1}{27}\left\{2\int_{\mathbb{T}} w^2\, dx + \int_{\mathbb{T}} w_x^2\, dx\right\} + O(u^3),\\
K_4(u) &:= \frac{4}{81}\left\{\int_{\mathbb{T}} w_x^2\, dx + \int_{\mathbb{T}} w_{xx}^2\, dx\right\} + O(u^3),\\
K_5(u) &:= \frac{1}{729}\left\{-2\int_{\mathbb{T}} w^2\, dx + 78\int_{\mathbb{T}} w_x^2\, dx + 114\int_{\mathbb{T}} w_{xx}^2\, dx + 153\int_{\mathbb{T}} w_{xxx}^2\, dx\right\} + O(u^3).
\end{aligned}
\tag{5.1.1}
$$

where we denoted by

$$
w := (\Lambda)^{-1}u := u - u_{xx}, \qquad \Lambda := (1-\partial_{xx})^{-1}. \tag{5.1.2}
$$

We remark that K_1 is the *momentum* Hamiltonian (recall Section 2.1.1).

In the sequel we will frequently use the following fact.

Lemma 5.1.1. *Let*

$$
H^{(2)}(u) := \sum_{j\neq 0} h_j |u_j|^2, \qquad K^{(2)}(u) := \sum_{j\neq 0} k_j |u_j|^2
$$

and consider a homogenous Hamiltonian of degree $n \in \mathbb{N}$

$$
F(u) := \sum_{j_1,\dots,j_n \in \mathbb{Z}\setminus\{0\}} F_{j_1,\dots,j_n} u_{j_1}\dots u_{j_n}
$$

such that

$$
\Pi_{\mathrm{Ker}(H^{(2)})}\Pi_{\mathrm{Rg}(K^{(2)})}F = 0.
$$

Then we have that

$$
h_{j_1}\omega(j_1) + \dots + h_{j_n}\omega(j_n) = 0 \quad \Rightarrow \quad k_{j_1}\omega(j_1) + \dots + k_{j_n}\omega(j_n) = 0 \quad \textit{or} \quad F_{j_1\dots j_n} = 0. \tag{5.1.3}
$$

Moreover, $F \in \mathrm{Ker}(H^{(2)}) \cap \mathrm{Ker}(K^{(2)})$ *contains only monomials with*

$$
h_{j_1}\omega(j_1) + \dots h_{j_n}\omega(j_n) = 0 \qquad k_{j_1}\omega(j_1) + \dots k_{j_n}\omega(j_n) = 0.
$$

Definition 5.1.2. By denoting the quadratic part of K_r, $r = 0, \dots, 5$ in (5.1.1) as

$$
K_r^{(2)}(u) := \sum_i k_{j_i}^{(r)}\, |u_{j_i}|^2,
$$

we say that an n-uple $\{j_1, \dots, j_n\}$, with $n \le N$, is a N-*resonance* of order n for the DP hierachy if

$$
\sum_{i=1}^n k_{j_i}^{(r)}\omega(j_i) = 0 \quad \forall r = 0, \dots, N-1. \tag{5.1.4}
$$

Proposition 5.1.3. *All the* 6*-resonances of the DP equations* (5.1.1) *are trivial, namely there are not resonances of order* 3, 5 *and the ones of order* 4, 6 *are, up to permutations, of the form*

$$(i,-i,j,-j)\,,\quad i,j\in\mathbb{N}\backslash\{0\},$$
$$(i,-i,j,-j,k,-k)\,,\quad i,j,k\in\mathbb{N}\backslash\{0\}.$$

Proof. The dispersion law of the costants of motion (5.1.1) are

$$\lambda^{(0)}(j):=\omega(j)k_j^{(0)}=j(4+j^2)(1+j^2)^{-1},\qquad \lambda^{(1)}(j):=\omega(j)k_j^{(1)}=j,$$
$$\lambda^{(2)}(j):=\omega(j)k_j^{(2)}=j\,(4+j^2)\,(1+j^2),$$
$$\lambda^{(3)}(j):=\omega(j)k_j^{(3)}=j(2+j^2)\,(4+j^2)\,(1+j^2),\qquad \lambda^{(4)}(j):=\omega(j)k_j^{(4)}=j^3(1+j^2)^2(4+j^2),$$
$$\lambda^{(5)}(j):=\omega(j)k_j^{(5)}=j(-2+78j^2+114j^4+153j^6)\,(4+j^2)\,(1+j^2).$$

We wish to prove that there are no solutions to the system

$$\begin{cases}\lambda^{(0)}(x_1)+\lambda^{(0)}(x_2)+\lambda^{(0)}(x_3)+\lambda^{(0)}(x_4)+\lambda^{(0)}(x_5)+\lambda^{(0)}(x_6)=0\\ \lambda^{(1)}(x_1)+\lambda^{(1)}(x_2)+\lambda^{(1)}(x_3)+\lambda^{(1)}(x_4)+\lambda^{(1)}(x_5)+\lambda^{(1)}(x_6)=0\\ \lambda^{(2)}(x_1)+\lambda^{(2)}(x_2)+\lambda^{(2)}(x_3)+\lambda^{(2)}(x_4)+\lambda^{(2)}(x_5)+\lambda^{(2)}(x_6)=0\\ \lambda^{(3)}(x_1)+\lambda^{(3)}(x_2)+\lambda^{(3)}(x_3)+\lambda^{(3)}(x_4)+\lambda^{(3)}(x_5)+\lambda^{(3)}(x_6)=0\\ \lambda^{(4)}(x_1)+\lambda^{(4)}(x_2)+\lambda^{(4)}(x_3)+\lambda^{(4)}(x_4)+\lambda^{(4)}(x_5)+\lambda^{(4)}(x_6)=0\\ \lambda^{(5)}(x_1)+\lambda^{(5)}(x_2)+\lambda^{(5)}(x_3)+\lambda^{(5)}(x_4)+\lambda^{(5)}(x_5)+\lambda^{(5)}(x_6)=0\end{cases}\tag{5.1.5}$$

except the trivial ones given by

$$x_1+x_2=x_3+x_3=x_5+x_6=0$$

and all its permutations. First one may replace in (5.1.5),

$$\lambda^{(0)}\rightsquigarrow\mu^{(0)}:=(\lambda^{(0)}-\lambda^{(1)})/3,\quad \lambda^{(1)}\rightsquigarrow\mu^{(1)}:=\lambda^{(1)}$$
$$\lambda^{(2)}\rightsquigarrow\mu^{(2)}:=\lambda^{(2)},\quad \lambda^{(3)}\rightsquigarrow\mu^{(3)}:=\lambda^{(3)}-2\lambda^{(2)},$$
$$\lambda^{(4)}\rightsquigarrow\mu^{(4)}:=\lambda^{(4)}-\mu^{(3)}\,,\qquad \lambda^{(5)}\rightsquigarrow\mu^{(5)}:=(\lambda^{(5)}+2\mu^{(2)}-78\mu^{(3)}+114\mu^{(4)})/153$$

We then may rephrase the problem as follows, set

$$m^{(0)}(j):=1,\qquad m^{(1)}(j):=j^2,\qquad m^{(2)}(j):=\,(4+j^2)\,(1+j^2)^2,$$
$$m^{(3)}(j):=j^2\,(4{+}j^2)\,(1{+}j^2)^2,\qquad m^{(4)}(j):=j^4(4{+}j^2)(1{+}j^2)^2\qquad m^{(5)}(j):=j^6(4{+}j^2)(1{+}j^2)^2$$

finding a solution of (5.1.5) is equivalent to finding integer values of $\{x_i\}_{i=1}^6$ such that the matrix

$$M_6=\begin{pmatrix}m^{(0)}(x_1)&m^{(0)}(x_2)&m^{(0)}(x_3)&m^{(0)}(x_4)&m^{(0)}(x_5)&m^{(0)}(x_6)\\ m^{(1)}(x_1)&m^{(1)}(x_2)&m^{(1)}(x_3)&m^{(1)}(x_4)&m^{(1)}(x_5)&m^{(1)}(x_6)\\ m^{(2)}(x_1)&m^{(2)}(x_2)&m^{(2)}(x_3)&m^{(2)}(x_4)&m^{(2)}(x_5)&m^{(2)}(x_6)\\ m^{(3)}(x_1)&m^{(3)}(x_2)&m^{(3)}(x_3)&m^{(3)}(x_4)&m^{(3)}(x_5)&m^{(3)}(x_6)\\ m^{(4)}(x_1)&m^{(4)}(x_2)&m^{(4)}(x_3)&m^{(4)}(x_4)&m^{(4)}(x_5)&m^{(4)}(x_6)\\ m^{(5)}(x_1)&m^{(5)}(x_2)&m^{(5)}(x_3)&m^{(5)}(x_4)&m^{(5)}(x_5)&m^{(5)}(x_6)\end{pmatrix}$$

has a non-trivial Kernel containing the vector

$$\left(\frac{x_1}{1+x_1^2}, \frac{x_2}{1+x_2^2}, \frac{x_3}{1+x_3^2}, \frac{x_4}{1+x_4^2}, \frac{x_5}{1+x_5^2}, \frac{x_6}{1+x_6^2}\right).$$

This requires in particular that the determinant of M_6 should be equal to zero. It is evident that the determinant of such a matrix has a factor, the term $\prod_{i\neq j}(x_i^2 - x_j^2)$. Indeed we get that

$$\det(M_6) = \prod_{i\neq j}(x_i^2 - x_j^2)P(x)$$

where $P(x)$ is strictly positive. $P(x)$ was explicitly computed using Wolfram Mathematica, it is symmetric, even in each variable, with non-negative coefficients and with $P(0) \neq 0$. Hence in order to have a non-trivial Kernel we need $x_5^2 = x_6^2$ (and all the permutations). Let us show that this leads to the trivial solutions.

Case 1: $x_5 + x_6 = 0$. We just need to look at the system of four equations

$$\begin{cases} \mu^{(0)}(x_1) + \mu^{(0)}(x_2) + \mu^{(0)}(x_3) + \mu^{(0)}(x_4) = 0 \\ \mu^{(1)}(x_1) + \mu^{(1)}(x_2) + \mu^{(1)}(x_3) + \mu^{(1)}(x_4) = 0 \\ \mu^{(2)}(x_1) + \mu^{(2)}(x_2) + \mu^{(2)}(x_3) + \mu^{(2)}(x_4) = 0 \\ \mu^{(3)}(x_1) + \mu^{(3)}(x_2) + \mu^{(3)}(x_3) + \mu^{(3)}(x_4) = 0 \end{cases} \tag{5.1.6}$$

We apply the same procedure as before and obtain a matrix M_4 whose determinant has the form

$$\prod_{i\neq j}(x_i^2 - x_j^2)P_4.$$

So either $x_3 = -x_4$ or $x_3 = x_4$ (or permutations). In the first case we reduce to the set of two equations

$$\begin{cases} \mu^{(0)}(x_1) + \mu^{(0)}(x_2) = 0 \\ \mu^{(1)}(x_1) + \mu^{(1)}(x_2) = 0 \end{cases} \tag{5.1.7}$$

which only has the trivial solution $x_1 + x_2 = 0$ (this is the second equation). If $x_3 = x_4$ we reduce to the following system of three equations

$$\begin{cases} \mu^{(0)}(x_1) + \mu^{(0)}(x_2) + 2\mu^{(0)}(x_3) = 0 \\ \mu^{(1)}(x_1) + \mu^{(1)}(x_2) + 2\mu^{(1)}(x_3) = 0 \\ \mu^{(2)}(x_1) + \mu^{(2)}(x_2) + 2\mu^{(2)}(x_3) = 0 \end{cases} \tag{5.1.8}$$

we proceed as before, associating to this system the matrix

$$M_3 = \begin{pmatrix} m^{(0)}(x_1) & m^{(0)}(x_2) & 2m^{(0)}(x_3) \\ m^{(1)}(x_1) & m^{(1)}(x_2) & 2m^{(1)}(x_3) \\ m^{(2)}(x_1) & m^{(2)}(x_2) & 2m^{(2)}(x_3) \end{pmatrix}$$

and computing its determinant. We get

$$\det(M_3) = \prod_{i\neq j}(x_i^2 - x_j^2)P_3$$

with P_3 strictly positive. If $x_1 = -x_2$ then $x_3 = 0$ and we are in the trivial solution, if $x_1 = x_2$ then we get $x_3 = -x_2$ and $x_4 = -x_1$ (since $x_3 = x_4$ and $x_2 = x_1$), again we are in the trivial solution.

Case 2: $x_5 - x_6 = 0$**.**

We look at the first five equations.

$$\begin{cases} \mu^{(0)}(x_1) + \mu^{(0)}(x_2) + \mu^{(0)}(x_3) + \mu^{(0)}(x_4) + 2\mu^{(0)}(x_5) = 0 \\ \mu^{(1)}(x_1) + \mu^{(1)}(x_2) + \mu^{(1)}(x_3) + \mu^{(1)}(x_4) + 2\mu^{(1)}(x_5) = 0 \\ \mu^{(2)}(x_1) + \mu^{(2)}(x_2) + \mu^{(2)}(x_3) + \mu^{(2)}(x_4) + 2\mu^{(2)}(x_5) = 0 \\ \mu^{(3)}(x_1) + \mu^{(3)}(x_2) + \mu^{(3)}(x_3) + \mu^{(3)}(x_4) + 2\mu^{(3)}(x_5) = 0 \\ \mu^{(4)}(x_1) + \mu^{(4)}(x_2) + \mu^{(4)}(x_3) + \mu^{(4)}(x_4) + 2\mu^{(4)}(x_5) = 0 \end{cases}$$

and repeat the usual procedure, we obtain a matrix M_5 whose determinant is

$$\det(M_5) = \prod_{i \neq j} (x_i^2 - x_j^2) P_5.$$

So either $x_4 = x_5$ or $x_4 = -x_5$ (or permutations). If $x_4 = -x_5$ we get the system 5.1.6. If $x_4 = x_5$ we get the equations

$$\begin{cases} \mu^{(0)}(x_1) + \mu^{(0)}(x_2) + \mu^{(0)}(x_3) + 3\mu^{(0)}(x_4) = 0 \\ \mu^{(1)}(x_1) + \mu^{(1)}(x_2) + \mu^{(1)}(x_3) + 3\mu^{(1)}(x_4) = 0 \\ \mu^{(2)}(x_1) + \mu^{(2)}(x_2) + \mu^{(2)}(x_3) + 3\mu^{(2)}(x_4) = 0 \\ \mu^{(3)}(x_1) + \mu^{(3)}(x_2) + \mu^{(3)}(x_3) + 3\mu^{(3)}(x_4) = 0 \end{cases}$$

We repeat the same procedure as in (5.1.6).

□

5.2 Weak Birkhoff Normal form

The aim of this section is to construct a family of approximately invariant tori for the equation (1.0.5) and to extract parameters ξ that allow to control the frequencies $\omega(\xi)$ of these tori. In particular we will show that there exist a finite dimensional subspace of the phase space closed for these approximate solutions. In order to do that we apply a normal form procedure to the DP Hamiltonian (recall (1.2.4))

$$\begin{aligned} &H(u) = H^{(2)}(u) + H^{(3)}(u) + H^{(\geq 9)}, \\ &H^{(2)}(u) := \frac{1}{2} \int_{\mathbb{T}} u^2 \, dx, \quad H^{(3)}(u) := -\frac{1}{6} \int_{\mathbb{T}} u^3 \, dx, \\ &H^{(\geq 9)}(u) := \int_{\mathbb{T}} f(u) \, dx. \end{aligned} \tag{5.2.1}$$

We decompose the phase space as

$$H_0^1(\mathbb{T}) := H_S \oplus H_S^{\perp}, \quad H_S := \mathrm{span}\{e^{\mathrm{i}\, j\, x} : j \in S\}, \quad H_S^{\perp} := \{u = \sum_{j \in S^c} u_j \, e^{\mathrm{i}\, j\, x} \in H_0^1(\mathbb{T})\}, \tag{5.2.2}$$

and we denote by $\Pi_S, \Pi_S^{\perp}$ the corresponding orthogonal projectors. The subspaces H_S and $H_S^{\perp}$ are symplectic respect to the 2-form Ω (see (1.2.5)). We write

$$u = v + z, \quad v := \Pi_S u := \sum_{j\in S} u_j \, e^{\mathrm{i}\, j\, x}, \quad z = \Pi_S^{\perp} u := \sum_{j \in S^c} u_j \, e^{\mathrm{i}\, j\, x}. \tag{5.2.3}$$

For a finite dimensional space

$$E := E_C := \mathrm{span} \left\{ e^{\mathrm{i}\, j\, x} : 0 < |j| \le C \right\}, \quad C > 0, \tag{5.2.4}$$

let Π_E denote the corresponding L^2-projector on E.
The notation $R(v^{k-q} z^q)$ indicates a homogeneous polynomial of degree k in (v, z) of the form

$$R(v^{k-q} z^q) = M[\underbrace{v, \ldots, v}_{(k-q)\ times}, \underbrace{z, \ldots, z}_{q\ times}], \quad M = k\text{-linear}.$$

We denote with $H^{(n, \ge k)}, H^{(n,k)}, H^{(n, \le k)}$ the terms of type $R(v^{n-s} \, z^s)$, where, respectively, $s \ge k, s = k, s \le k$, that appear in the homogeneous polynomial H_n of degree n in the variables (v, z).
Given an n-uple $\{j_1, \ldots, j_n\} \subset \mathbb{Z} \setminus \{0\}$ and a set $B \subset \mathbb{Z} \setminus \{0\}$ we define

$$\sharp(\{j_1, \ldots, j_n\}, B) := \text{number of } j_i \text{ belonging to } B. \tag{5.2.5}$$

In this way $H^{(n,k)}$ is supported on the set

$$\Big\{ j_1, \ldots, j_n \in \mathbb{Z} \setminus \{0\} \ : \quad \sum_{i=1}^{n} j_i = 0 \, , \quad \sharp(\{j_1, \ldots, j_n\}, S^c) = k \Big\}.$$

Definition 5.2.1. Given a Hamiltonian H, we define $\Pi_{\mathrm{Ker}(H)}$ as the projection on the kernel of the adjoint action $\mathrm{ad}(H) := \{H, \cdot\}$.

Remark 5.2.2. We note that if $j_1, \ldots, j_N \in \mathbb{Z} \setminus \{0\}$, $j_1 + \cdots + j_N = 0$ and

$$\#(\{j_1, \ldots, j_N\}, S^c) \le 1$$

then $\max_{i=1,\ldots,N} |j_i| \le (N-1) C_S$, where $C_S := \max_{j \in S} |j|$. Thus, the vector field $X_{F^{(N)}}$, generated by the finitely supported Hamiltonian

$$F^{(N)} = \sum_{j_1 + \cdots + j_N = 0} F^{(N)}_{j_1 \ldots j_N} u_{j_1} \ldots u_{j_N},$$

is finite rank, and, in particular, it vanishes outside the finite dimensional subspace $E := E_{(N-1)C_S}$ (see (5.2.4)) and it has the form

$$X_{F^{(N)}}(u) = \Pi_E X_{F^{(N)}}(\Pi_E u).$$

Therefore its flow $\Phi^{(N)}$ is analytic and invertible on the phase space $H^1_0(\mathbb{T}_x)$. We recall that the condition $\sum_{i=1}^{N} j_i = 0$ corresponds to the fact that $F^{(N)}$ Poisson commutes with the quadratic Hamiltonian K_1 (see (5.1.1) and Section 2.1.1).

Proposition 5.2.3. *There exists an analytic and symplectic change of coordinates*

$$\Phi_B \colon H_0^1(\mathbb{T}_x) \to H_0^1(\mathbb{T}_x)$$

of the form

$$\Phi_B(u) = u + \Psi(u)\,, \quad \Psi(u) = \Pi_E \Psi(\Pi_E u), \tag{5.2.6}$$

where E *is a finite dimensional space as in* (5.2.4)*, such that the Hamiltonian* H *in* (5.2.1) *transforms into*

$$\begin{aligned}\mathcal{H} := H \circ \Phi_B = H^{(2)} &+ \mathcal{H}^{(3,\geq 2)} + \Pi_{Ker(H^{(2)})}\mathcal{H}^{(4,0)} + \mathcal{H}^{(\geq 4,\geq 2)} + \mathcal{H}^{(5,\geq 2)} \\ &+ \Pi_{Ker(H^{(2)})}\mathcal{H}^{(6,0)} + \mathcal{H}^{(6,\geq 2)} + \mathcal{H}^{(7,\geq 2)} + \Pi_{Ker(H^{(2)})}\mathcal{H}^{(8,0)} \\ &+ \mathcal{H}^{(8,\geq 2)} + \mathcal{H}^{(\geq 9)},\end{aligned} \tag{5.2.7}$$

where

$$\begin{aligned}\mathcal{H}^{(3,\geq 2)} :&= -\frac{1}{2}\int_{\mathbb{T}} v\, z^2\, dx - \frac{1}{6}\int_{\mathbb{T}} z^3\, dx, \\ \Pi_{Ker(H^{(2)})}\mathcal{H}^{(4,0)} &= \frac{1}{2}\sum_{j\in S^+} \frac{\omega(2j)}{2\omega(j)-\omega(2j)}|u_j|^4 \\ &+ \sum_{\substack{j_1,j_2\in S^+,\\ j_1-j_2\neq 0}} \frac{\omega(j_1+j_2)}{\omega(j_1)+\omega(j_2)-\omega(j_1+j_2)}|u_{j_1}|^2|u_{j_2}|^2 \\ &+ \sum_{\substack{j_1,j_2\in S^+,\\ j_1-j_2\neq 0}} \frac{\omega(j_1-j_2)}{\omega(j_1)-\omega(j_2)-\omega(j_1-j_2)}|u_{j_1}|^2|u_{j_2}|^2\end{aligned} \tag{5.2.8}$$

and $\mathcal{H}^{(\geq 9)}$ *collects all the terms of order at least nine in* (v,z)*.*
The same change of variables Φ_B *puts all the Hamiltonians in* (5.1.1) *in weak Birkhoff normal form up to order six as in* (5.2.7)*. In particular we have*

$$K_1 \circ \Phi_B = K_1.$$

Proof. **Step (1)**. First we remove the cubic terms independent of z and linear in z from the Hamiltonian

$$H^{(3)} = -\frac{1}{6}\int_{\mathbb{T}} u^3\, dx = -\frac{1}{6}\int_{\mathbb{T}} v^3\, dx - \frac{1}{2}\int_{\mathbb{T}} v^2 z\, dx - \frac{1}{2}\int_{\mathbb{T}} v z^2\, dx - \frac{1}{6}\int_{\mathbb{T}} z^3\, dx. \tag{5.2.9}$$

Thus we look for a symplectic transformation Φ_3 of the phase space which eliminates the monomials $u_{j_1}\, u_{j_2}\, u_{j_3}$ of $H^{(3)}$ with at most one index outside S.
Note that any homogenous Hamiltonian, which preserves the momentum and which is linear in z or independent of z, has compact support. Thus by Remark 4.1.2 its flow is well defined on the entire phase space.
We look for $\Phi_3 := (\Phi^t_{F^{(3)}})_{|t=1}$ as the time-1 flow map generated by the Hamiltonian vector field $X_{F^{(3)}}$, with an auxiliary Hamiltonian of the form

$$F^{(3)}(u) := \sum_{j_1+j_2+j_3=0} F^{(3)}_{j_1\,j_2\,j_3}\, u_{j_1}\, u_{j_2}\, u_{j_3}. \tag{5.2.10}$$

The transformed Hamiltonian is

$$H_3 := H \circ \Phi_3 = H^{(2)} + H_3^{(3)} + H_3^{(4)} + H_3^{(\geq 5)}, \tag{5.2.11}$$

$$H_3^{(3)} := \{H^{(2)}, F^{(3)}\} + H^{(3)}, \quad H_3^{(4)} := \frac{1}{2}\{\{H^{(2)}, F^{(3)}\}, F^{(3)}\} + \{H^{(3)}, F^{(3)}\}, \tag{5.2.12}$$

where $H_3^{(\geq 5)}$ collects all the terms of order at least five in (v, z). We choose $F^{(3)}$ in (5.2.10) such that the following homological equation holds

$$\{H^{(2)}, F^{(3)}\} + H^{(3)} = H^{(3,\geq 2)} \quad \Leftrightarrow \quad \{H^{(2)}, F^{(3)}\} = -\Pi_{\mathrm{Rg}(H^{(2)})} H^{(3,\leq 1)}. \tag{5.2.13}$$

In Fourier coefficients, by (1.2.6) and (5.2.9), the equation (5.2.13) reads

$$\sum_{\substack{j_1,j_2,j_3 \in \mathbb{Z}\setminus\{0\},\\ j_1+j_2+j_3=0}} \mathrm{i}\,(\omega(j_1)+\omega(j_2)+\omega(j_3))\, F^{(3)}_{j_1 j_2 j_3}\, u_{j_1} u_{j_2} u_{j_3} = -\frac{1}{6} \sum_{\substack{j_1,j_2,j_3 \in \mathbb{Z}\setminus\{0\},\\ \sharp(\{j_1,j_2,j_3\}, S^c)\leq 1\\ \omega(j_1)+\omega(j_2)+\omega(j_3)\neq 0\\ j_1+j_2+j_3=0}} u_{j_1} u_{j_2} u_{j_3} \tag{5.2.14}$$

and we have to determine $F^{(3)}_{j_1 j_2 j_3}$. Since $j_1 + j_2 + j_3 = 0$ we have

$$\omega(j_1)+\omega(j_2)+\omega(j_3) = -\frac{3 j_1 j_2 j_3}{(1+j_1^2)(1+j_2^2)(1+j_3^2)} \left[3 + \left(j_1 + \frac{j_2}{2}\right)^2 + \frac{3}{4} j_2^2\right] \neq 0, \tag{5.2.15}$$

since $j_1, j_2, j_3 \neq 0$. Therefore, in order to solve (5.2.13), we set

$$F^{(3)}_{j_1 j_2 j_3} := \begin{cases} -\dfrac{1}{6\,\mathrm{i}\,(\omega(j_1)+\omega(j_2)+\omega(j_3))} & \text{if } \sharp(\{j_1,j_2,j_3\}, S^c) \leq 1,\ j_1+j_2+j_3=0, \\ 0 & \text{otherwise.} \end{cases} \tag{5.2.16}$$

By (5.2.15) $\Pi_{\mathrm{Rg}(H^{(2)})} H^{(3,\leq 1)} = H^{(3,\leq 1)}$, and we get (see (5.2.12), (5.2.13))

$$H_3^{(3)} = H^{(3,\geq 2)}, \qquad H_3^{(4)} = \frac{1}{2}\{H^{(3,\leq 1)}, F^{(3)}\} + \{H^{(3,\geq 2)}, F^{(3)}\}. \tag{5.2.17}$$

By direct inspection $\{F^{(3)}, K_1\} = 0$, so $K_1 \circ \Phi_3 = K_1$. We now claim that (recall the definition of K_i in (5.1.1))

$$K_{i,3} := K_i \circ \Phi_3 = K_i^{(2)} + K_i^{(3,\geq 2)} + K_{i,3}^{(4)} + K_{i,3}^{(\geq 5)} \tag{5.2.18}$$

for $i = 2, \dots, 5$; namely that the change of variables Φ_3 puts simultaneously all the Hamiltonians commuting with the DP Hamiltonian into Weak Birkhoff normal form up to order 4. We need to show that

$$\{K_i^{(2)}, F^{(3)}\} + K_i^{(3)} = K_i^{(3,\geq 2)} \quad \Leftrightarrow \quad \{K_i^{(2)}, F^{(3)}\} = -\Pi_{\mathrm{Rg}(K_i^{(2)})} K_i^{(3,\leq 1)} \quad i = 2, \dots, 5. \tag{5.2.19}$$

Since H and K_i Poisson commute, by (C.0.8) we have

$$\{H^{(2)}, K_i^{(3,\leq 1)}\} = \{K_i^{(2)}, H^{(3,\leq 1)}\},$$

which in turn implies that

$$\Pi_{\mathrm{Rg}(K_i^{(2)})}\Pi_{\mathrm{Ker}(H^{(2)})}H^{(3,\le 1)} = 0\,, \quad \Pi_{\mathrm{Rg}(H^{(2)})}\Pi_{\mathrm{Ker}(K_i^{(2)})}K^{(3,\le 1)} = 0\,, \tag{5.2.20}$$

$$\Pi_{\mathrm{Rg}(K_i^{(2)})}H^{(3,\le 1)} = (\mathrm{ad}\, K_i^{(2)})^{-1}\{H^{(2)}, K_i^{(3,\le 1)}\}. \tag{5.2.21}$$

By (5.2.20) we have

$$\Pi_{\mathrm{Rg}(H^{(2)})}H^{(3,\le 1)} = \Pi_{\mathrm{Rg}(H^{(2)})}\Pi_{\mathrm{Rg}(K_i^{(2)})}H^{(3,\le 1)}.$$

Then we substitute (5.2.21) in (5.2.13), which defines $F^{(3)}$ and obtain (5.2.19).

Step (2). We now construct a symplectic map Φ_4 to eliminate the term $H_3^{(4,1)}$ (which is linear in z) and to normalize $H_3^{(4,0)}$ (which is independent of z). We look for a map $\Phi_4 := (\Phi^t_{F^{(4)}})_{|t=1}$ which is the time-1 flow map of an auxiliary Hamiltonian

$$F^{(4)}(u) := \sum_{j_1+j_2+j_3+j_4=0} F^{(4)}_{j_1 j_2 j_3 j_4}\, u_{j_1} u_{j_2} u_{j_3} u_{j_4}.$$

Note that we make the ansatz that the function $F^{(4)}$ preserves the momentum, since the Hamiltonians which we want to eliminate do it. The transformed Hamiltonian is

$$H_4 := H_3 \circ \Phi_4 = H^{(2)} + H^{(3,\ge 2)} + H_4^{(4)} + H_4^{(\ge 5)}, \quad H_4^{(4)} := \{H^{(2)}, F^{(4)}\} + H_3^{(4)}, \tag{5.2.22}$$

where $H_4^{(\ge 5)}$ collects all the terms of order at least five in (v,z). We choose $F^{(4)}$ such that

$$\{H^{(2)}, F^{(4)}\} + H_3^{(4)} = \Pi_{\mathrm{Ker}(H^{(2)})}H_3^{(4,\le 1)} + H_3^{(4,\ge 2)}. \tag{5.2.23}$$

In Fourier representation we have

$$\{H^{(2)}, F^{(4)}\} = -\sum_{j_1+j_2+j_3+j_4=0} \mathrm{i}(\omega(j_1)+\omega(j_2)+\omega(j_3)+\omega(j_4))\, F^{(4)}_{j_1 j_2 j_3 j_4}$$

and then we set

$$F^{(4)}_{j_1 j_2 j_3 j_4} := \begin{cases} -\dfrac{H^{(4)}_{3,j_1\,j_2\,j_3\,j_4}}{\mathrm{i}\sum_{i=1}^4 \omega(j_i)} & \text{if} \quad \sharp(\{j_1,j_2,j_3,j_4\}, S^c) \le 1, \quad \sum_{i=1}^4 \omega(j_i) \ne 0, \\ 0 & \qquad\qquad\qquad \text{otherwise} \end{cases} \tag{5.2.24}$$

where $H^{(4)}_{3,j_1\,j_2\,j_3\,j_4}$ is the Fourier coefficient of $H_3^{(4)}$ corresponding to the harmonic (j_1,j_2,j_3,j_4). We claim that for all $i = 1,\dots,5$

$$\Pi_{\mathrm{Ker}(H^{(2)})}\Pi_{\mathrm{Rg}(K_i^{(2)})}H_3^{(4,\le 1)} = 0\,, \tag{5.2.25}$$

indeed, since $\{H_3, K_{i,3}\} = 0$, by (5.2.11), (5.2.18), we have

$$\{H^{(2)}, K_{i,3}^{(4)}\} - \{K^{(2)}, H_{i,3}^{(4)}\} + \{H_3^{(3,\ge 2)}, K_{i,3}^{(3,\ge 2)}\} = 0\,,$$

which implies

$$\{H^{(2)}, K_{i,3}^{(4,\le 1)}\} = \{K^{(2)}, H_{i,3}^{(4,\le 1)}\}\,.$$

Then (5.2.25) follows by using the same strategy as in Step (1). Formula (5.2.25) implies that $\Pi_{\mathrm{Ker}(H^{(2)})} H_3^{(4,\leq 1)}$ is supported on the set of 6-resonances of order 4. By Proposition 5.1.3, we have that such resonances are only the trivial ones, namely

$$(j_1 + j_2)(j_2 + j_3)(j_1 + j_3) = 0. \tag{5.2.26}$$

We note that, by the symmetry of S, the resonances (5.2.26) cannot occur when one of the integers j_1, j_2, j_3, j_4 does not belong to S. Hence $\Pi_{\mathrm{Ker}(H^{(2)})} H_3^{(4,1)} = 0$.
Now we compute $\Pi_{\mathrm{Ker}(H^{(2)})} H_3^{(4,0)}$. We have

$$\Pi_{\mathrm{Ker}(H^{(2)})} H_3^{(4,0)} = \frac{1}{8} \sum_{\substack{j_1,j_2,j_3,j_4 \in S, \\ j_1+j_2+j_3+j_4=0 \\ j_1+j_2 \neq 0,\, j_3+j_4 \neq 0, \\ \sum_{k=1}^4 \omega(j_k)=0}} \frac{\omega(j_1+j_2)}{\omega(j_1)+\omega(j_2)-\omega(j_1+j_2)} u_{j_1} u_{j_2} u_{j_3} u_{j_4}. \tag{5.2.27}$$

By (5.2.26) we have that the possible cases are

$$(i)\ \{j_2 \neq -j_1, j_3 = -j_1, j_4 = -j_2\} \qquad (ii)\ \{j_2 \neq -j_1, j_3 \neq -j_1, j_3 = -j_2, j_4 = -j_1\}$$

and by the fact that $\omega(-j) = -\omega(j)$ (see (1.2.7)) and the simmetry of A we have

$$\begin{aligned} \Pi_{\mathrm{Ker}(H^{(2)})} H_3^{(4,0)} &= \frac{1}{4} \sum_{j \in S^+} \frac{\omega(2j)}{2\omega(j)-\omega(2j)} |u_j|^4 \\ &+ \frac{1}{2} \sum_{\substack{j_1,j_2 \in S^+, \\ j_1-j_2 \neq 0}} \frac{\omega(j_1+j_2)}{\omega(j_1)+\omega(j_2)-\omega(j_1+j_2)} |u_{j_1}|^2 |u_{j_2}|^2 \\ &+ \frac{1}{2} \sum_{\substack{j_1,j_2 \in S^+, \\ j_1-j_2 \neq 0}} \frac{\omega(j_1-j_2)}{\omega(j_1)-\omega(j_2)-\omega(j_1-j_2)} |u_{j_1}|^2 |u_{j_2}|^2. \end{aligned} \tag{5.2.28}$$

We conclude that

$$H_4 = H^{(2)} + H^{(3,\geq 2)} + H_4^{(4,0)} + H_4^{(4,\geq 2)} + H_4^{(\geq 5)}, \quad H_4^{(4,0)} := \Pi_{\mathrm{Ker}(H^{(2)})} H_3^{(4,0)}. \tag{5.2.29}$$

As in the previous step, we have the same normal form for all the $K_{i,3}$, namely

$$\begin{aligned} &K_{i,4} := K_{i,3} \circ \Phi_4 = K_i^{(2)} + K_i^{(3,\geq 2)} + K_{i,4}^{(4,0)} + K_{i,4}^{(4,\geq 2)} + K_{i,4}^{(\geq 5)}, \\ &K_{i,4}^{(4,0)} := \Pi_{Ker(H^{(2)})} K_{i,3}^{(4,0)} \end{aligned} \tag{5.2.30}$$

for $i = 1, \dots, 5$.

Step (3). We want to remove the terms with at most one index among $j_1, \dots, j_5$ outside S from $H_4^{(5)}$. We claim that for all $i = 1, \dots, 5$

$$\Pi_{\mathrm{Ker}(H^{(2)})} \Pi_{\mathrm{Rg}(K_i^{(2)})} H_4^{(5,\leq 1)} = 0\,, \tag{5.2.31}$$

indeed, since $\{H_4, K_{i,4}\} = 0$, using (5.2.29), (5.2.30), we have

$$\{H^{(2)}, K_{i,4}^{(5,\leq 1)}\} = \{K^{(2)}, H_{i,4}^{(5,\leq 1)}\}\,.$$

Then (5.2.31) follows by using the same strategy as in Step (1)-(2). Formula (5.2.31) implies that $\Pi_{\mathrm{Ker}(H^{(2)})} H_4^{(5,\leq 1)}$ is supported on the set of 6-resonances of order 5, but by Proposition (5.1.3) there are not such resonances. We consider the auxiliary Hamiltonian

$$F^{(5)} = \sum_{j_1+\cdots+j_5=0} F^{(5)}_{j_1,\dots,j_5}\, u_{j_1}\dots u_{j_5}, \tag{5.2.32}$$

$$F^{(5)}_{j_1 j_2 j_3 j_4 j_5} := \begin{cases} -\dfrac{H^{(5)}_{3,j_1 j_2 j_3 j_4 j_5}}{\mathrm{i}\sum_{i=1}^{5}\omega(j_i)} & \text{if} \quad \sharp(\{j_1, j_2, j_3, j_4, j_5\}, S^c) \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.2.33}$$

Let Φ_5 be the time-1 flow generated by $X_{F^{(5)}}$. The new Hamiltonian is

$$\begin{aligned} &H_5 := H_4 \circ \Phi_5 = H^{(2)} + H^{(3,\geq 2)} + H_4^{(4,0)} + H_4^{(4,\geq 2)} + H_5^{(5)} + H_5^{(\geq 6)}, \\ &H_5^{(5)} = \{H^{(2)}, F^{(5)}\} + H_4^{(5)}, \end{aligned} \tag{5.2.34}$$

where $H_5^{(\geq 6)}$ collects all the terms of degree greater or equal than six, and, by the definition of $F^{(5)}$,

$$H_5^{(5)} = \sum_{q=2}^{5} R(v^{5-q} z^q). \tag{5.2.35}$$

Step (4). We now construct a symplectic map Φ_6 to eliminate the term $H_5^{(6,1)}$ (which is linear in z) and to normalize $H_5^{(6,0)}$ (which is independent of z). We follow exactly the same procedure adopted in Step (2). We claim that for all $i = 1, \dots, 5$

$$\Pi_{\mathrm{Ker}(H^{(2)})} \Pi_{\mathrm{Rg}(K_i^{(2)})} H_5^{(6,\leq 1)} = 0\,, \tag{5.2.36}$$

indeed, since $\{H_5, K_{i,5}\} = 0$, by (5.2.11), (5.2.18), we have

$$\{H^{(2)}, K_{i,5}^{(6)}\} - \{K^{(2)}, H_{i,5}^{(6)}\} + \{H_5^{(5,\geq 2)}, K_{i,5}^{(5,\geq 2)}\} = 0\,,$$

which implies

$$\{H^{(2)}, K_{i,5}^{(6,\leq 1)}\} = \{K^{(2)}, H_{i,5}^{(6,\leq 1)}\}\,.$$

Then (5.2.25) follows by using the same strategy as in Step (2). Formula (5.2.25) implies that $\Pi_{\mathrm{Ker}(H^{(2)})} H_5^{(6,\leq 1)}$ is supported on the set of 6-resonances of order 6. By Proposition 5.1.3, we have that such resonances are only the trivial ones. We note that, by the symmetry of S, the resonances (5.2.26) cannot occur when one of the integers $j_1, j_2, j_3, j_4, j_5, j_6$ does not belong to S. Hence $\Pi_{\mathrm{Ker}(H^{(2)})} H_5^{(6,1)} = 0$.

We look for a map $\Phi_6 := (\Phi^t_{F^{(6)}})_{|t=1}$ which is the time-1 flow map of an auxiliary Hamiltonian

$$F^{(6)}(u) := \sum_{j_1+j_2+j_3+j_4+j_5+j_6=0} F^{(4)}_{j_1 j_2 j_3 j_4 j_5 j_6}\, u_{j_1} u_{j_2} u_{j_3} u_{j_4} u_{j_5} u_{j_6}$$

and we note that by Proposition 5.1.3 the 8-resonance of order 6 are only the trivial ones, namely the ones given by

$$j_1 + j_2 = j_3 + j_4 = j_5 + j_6 = 0 \tag{5.2.37}$$

and all its permutations. Hence

$$\Pi_{\mathrm{Ker}(H^{(2)})} H_5^{(6,0)} = \sum_{j_1, j_2, j_3 \in S^+} M_{j_1 j_2 j_3} |u_{j_1}|^2 |u_{j_2}|^2 |u_{j_3}|^2 \tag{5.2.38}$$

where $M_{j_1 j_2 j_3}$ are real numbers that we do not need to compute.
The new Hamiltonian is

$$H_6 := H_5 \circ \Phi_6 = H^{(2)} + H^{(3,\geq 2)} + H_4^{(4,0)} + H_4^{(4,\geq 2)} + H_5^{(5,\geq 2)} + H_6^{(6)} + H_6^{(\geq 7)}, \tag{5.2.39}$$

$$H_6^{(6)} := \Pi_{\mathrm{Ker}(H^{(2)})} H_5^{(6,0)} + H_6^{(6,\geq 2)} \tag{5.2.40}$$

where $H_6^{(\geq 7)}$ collects all the terms of degree greater or equal than seven. Moreover

$$H_6^{(6,\geq 2)} = \sum_{q=2}^{6} R(v^{6-q} z^q). \tag{5.2.41}$$

Step (5). We want to remove the terms with at most one index among $j_1, \dots, j_7$ outside S from $H_6^{(7)}$. We consider the auxiliary Hamiltonian

$$F^{(7)} = \sum_{j_1 + \dots + j_7 = 0} F^{(7)}_{j_1, \dots, j_7} u_{j_1} \dots u_{j_7}, \tag{5.2.42}$$

$$F^{(7)}_{j_1 j_2 j_3 j_4 j_5 j_6 j_7} := \begin{cases} -\dfrac{H^{(7)}_{3, j_1 j_2 j_3 j_4 j_5 j_6 j_7}}{\mathrm{i} \sum_{i=1}^{7} \omega(j_i)} & \text{if } \sharp(\{j_1, j_2, j_3, j_4, j_5, j_6, j_7\}, S^c) \leq 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.2.43}$$

Let Φ_7 be the time-1 flow generated by $X_{F^{(7)}}$. By the assumption (H0), see (1.2.13), this transformation is well-defined. The new Hamiltonian is

$$\begin{aligned} H_7 := H_6 \circ \Phi_7 = H^{(2)} + H^{(3,\geq 2)} + H_4^{(4,0)} + H_4^{(4,\geq 2)} + H_5^{(5,\geq 2)} \\ + H_6^{(6,0)} + H_6^{(6,\geq 2)} + H_7^{(7)} + H_7^{(\geq 8)}, \\ H_7^{(7)} = \{H^{(2)}, F^{(7)}\} + H_6^{(7)}, \end{aligned} \tag{5.2.44}$$

where $H_7^{(\geq 8)}$ collects all the terms of degree greater or equal than eight, and, by the definition of $F^{(7)}$,

$$H_7^{(7)} = \sum_{q=2}^{7} R(v^{7-q} z^q). \tag{5.2.45}$$

Step (6). We now construct a symplectic map Φ_8 to eliminate the term $H_7^{(8,1)}$ (which is linear in z) and to normalize $H_7^{(8,0)}$ (which is independent of z). We follow exactly the same procedure adopted in Step (2)-(4). We note that, by the symmetry of S, the resonances (5.2.26) cannot occur when one of the integers $j_1, j_2, j_3, j_4, j_5, j_6, j_7, j_8$ does not belong to S. Hence $\Pi_{\mathrm{Ker}(H^{(2)})} H_7^{(8,1)} = 0$.

We look for a map $\Phi_8 := (\Phi^t_{F^{(8)}})_{|t=1}$ which is the time-1 flow map of an auxiliary Hamiltonian

$$F^{(8)}(u) := \sum_{j_1+j_2+j_3+j_4+j_5+j_6+j_7+j_8=0} F^{(6)}_{j_1j_2j_3j_4j_5j_6j_7j_8} u_{j_1}u_{j_2}u_{j_3}u_{j_4}u_{j_5}u_{j_6}u_{j_7}u_{j_8}$$

and we note that by assumption (H0), see (1.2.13), the 8-resonance of order 8 are only the trivial ones, namely the ones given by

$$j_1 + j_2 = j_3 + j_4 = j_5 + j_6 = j_7 + j_8 = 0 \tag{5.2.46}$$

and all its permutations. Hence

$$\Pi_{\mathrm{Ker}(H^{(2)})} H_7^{(8,0)} = \sum_{j_1,j_2,j_3,j_4 \in S^+} \tilde{M}_{j_1j_2j_3j_4} |u_{j_1}|^2|u_{j_2}|^2|u_{j_3}|^2|u_{j_4}|^2 \tag{5.2.47}$$

where $\tilde{M}_{j_1j_2j_3j_4}$ are real numbers that we do not need to compute.
The new Hamiltonian is

$$\begin{aligned} H_8 := H_7 \circ \Phi_8 = H^{(2)} + H^{(3,\geq 2)} + H_4^{(4,0)} + H_4^{(4,\geq 2)} + H_5^{(5,\geq 2)} + H_6^{(6,0)} \\ + H_6^{(6,\geq 2)} + H_7^{(7,\geq 2)} + H_8^{(8)} + H_8^{(\geq 9)}, \end{aligned} \tag{5.2.48}$$

$$H_8^{(8)} := \Pi_{\mathrm{Ker}(H^{(2)})} H_7^{(8,0)} + H_8^{(8,\geq 2)} \tag{5.2.49}$$

where $H_8^{(\geq 9)}$ collects all the terms of degree greater or equal than seven. Moreover

$$H_8^{(8,\geq 2)} = \sum_{q=2}^{8} R(v^{8-q} z^q). \tag{5.2.50}$$

Setting $\Phi_B := \Phi_3 \circ \Phi_4 \circ \Phi_5 \circ \Phi_6 \circ \Phi_7 \circ \Phi_8$ and renaming $\mathcal{H} := H_8 = H \circ \Phi_B$, by Remark (5.2.2), we conclude the proof of Proposition 5.2.3. □

5.3 Action-angle variables

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

$$\begin{aligned} \mathbb{T}^\nu \times \mathbb{R}^\nu_+ &\longrightarrow \{z = 0\} \\ (\theta, I) &\longmapsto v = \sum_{j \in S} \sqrt{I_j}\, e^{\mathrm{i}\theta_j}\, e^{\mathrm{i}jx} \end{aligned} \tag{5.3.1}$$

where $\mathbb{R}^\nu_+$ is the set of ν-ples of real and positive numbers. Note that this change of coordinates is real (according to Definition (2.1.1)-(4)) if and only if $I_{-j} = I_j$ and $\theta_{-j} = -\theta_j$.
The symplectic form in (1.2.5) restricted to the subspace H_S transforms into the 2-form

$$\tilde{\Omega}_S = \sum_{j \in S^+} d\theta_j \wedge \frac{1}{\omega(j)}\, dI_j. \tag{5.3.2}$$

Hence the Hamiltonian equations of the system $\mathcal{H}^{(\leq 8)} := H^{(2)}+H_3^{(3)}+H_4^{(4)}+H_5^{(5)}+H_6^{(6)}+H_7^{(7)}+H_8^{(8)}$ (see (5.2.12), (5.2.22), (5.2.34), (5.2.39), (5.2.44), (5.2.49)) restricted to $\{z=0\}$ writes

$$\begin{cases} \dot\theta_j = \omega(j)\ \partial_{I_j}\mathcal{H}^{(\leq 8)}(\theta, I, 0), & j\in S^+, \\ \dot I_j = -\partial_{\theta_j}\mathcal{H}^{(\leq 8)}(\theta, I, 0), & j\in S^+. \end{cases} \tag{5.3.3}$$

We have that

$$\mathcal{H}^{(\leq 8)}(\theta, I, 0) := \sum_{j\in S^+} I_j + \Pi_{\mathrm{Ker}(H^{(2)})}H_{(4)}^{(4,0)}(I) + \Pi_{\mathrm{Ker}(H^{(2)})}H_{(6)}^{(6,0)}(I) + \Pi_{\mathrm{Ker}(H^{(2)})}H_{(8)}^{(8,0)}(I) \tag{5.3.4}$$

depends only by the actions I, thus we have

$$\begin{cases} \dot\theta_j = \omega_j(I), & j\in S^+, \\ \dot I_j = 0, & j\in S^+. \end{cases} \tag{5.3.5}$$

By (5.2.28) and (5.2.38)

$$\omega_j(I) = \omega(j) + \frac{1}{2}\frac{\omega(j)\omega(2j)}{2\omega(j)-\omega(2j)} I_j + \omega(j)\sum_{k\in S^+, k\neq j} \mathtt{b}_{jk}\, I_k + \mathcal{G}(I) \tag{5.3.6}$$

where

$$\begin{aligned} \mathtt{b}_{jk} :=& \frac{\omega(k+j)}{\omega(k)+\omega(j)-\omega(k+j)} + \frac{\omega(k-j)}{\omega(k)-\omega(j)-\omega(k-j)} \\ =& \frac{2}{3}\frac{(1+k^2)(1+j^2)(2+k^2+j^2)}{(3+k^2+j^2+kj)(3+k^2+j^2-kj)} \end{aligned} \tag{5.3.7}$$

and $\mathcal{G}(I)$ is the sum of a bilinear and 3-linear functions of the variable I (see (5.2.38) and (5.2.47)). Hence, in a small neighbourhood of the origin of the phase space $H_0^1(\mathbb{T}_x)$, the submanifold $\{z=0\}$ is foliated by invariant tori of amplitude ξ and frequency vector $\omega(\xi) := (\omega_j(\xi))_{j\in S^+}$ as in (5.3.6). We shall select from this set of tori the approximately invariant quasi-periodic solutions to be continued and we will use their *unperturbed actions* ξ as parameters. Moreover, we shall require that the frequencies of these tori vary in a one-to-one way with the actions ξ. Thanks to this fact, we could control the conditions (*Melnikov conditions*) that we shall impose on the frequencies ω through the amplitudes, and viceversa.

We can write, in a compact form, the vector with components $\omega_j(I)$, $j\in S^+$, in (5.3.6), as

$$\omega(\xi) = \overline{\omega} + \mathbb{A}\,\xi + \mathcal{G}(\xi), \tag{5.3.8}$$

where $\overline{\omega}$ is the vector of the linear frequencies (see (1.2.9)),

$$\mathbb{A} := \frac{1}{2}\Omega\ \mathrm{diag}\left(\frac{\omega(2j)}{2\omega(j)-\omega(2j)}\right)_{j\in S^+} + \Omega\ \mathbb{B}, \qquad \Omega := \mathrm{diag}\big(\omega(j)\big)_{j\in S^+}, \tag{5.3.9}$$

where $\mathbb{B}$ is the $\nu\times\nu$ matrix defined by (recall (5.3.7))

$$\mathbb{B} := \begin{cases} \mathtt{b}_{jk} & \text{if } j\neq k, \\ 0 & \text{if } j = k \end{cases}$$

and $\mathcal{G}(\xi)$ is the sum of a bilinear and 3-linear functions of $\xi \in \mathbb{R}^\nu$ (see (5.3.6), (5.2.38), (5.2.47)). The function of ξ in (5.3.8) is the *frequency-amplitude map*, which describes, at the main order, how the tangential frequencies are shifted by the amplitudes ξ.

In order to work in a neighbourhood of the unperturbed torus $\{I \equiv \xi\}$ it is advantageous to introduce a set of coordinates $(\theta, y, z) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$ adapted to it, defined by

$$u = \Phi_{aa}(\theta, y, z) \qquad \Longleftrightarrow \qquad \begin{cases} u_j := \sqrt{I_j}\, e^{\mathrm{i}\theta_j}\, e^{\mathrm{i}\, j\, x}, \qquad I_j := \xi_j + |\omega(j)| y_j, \qquad j \in S, \\ u_j := z_j, \qquad\qquad\qquad\qquad\qquad\qquad\qquad j \in S^c, \end{cases} \tag{5.3.10}$$

where (recall $\overline{u}_j = u_{-j}$)

$$\xi_{-j} = \xi_j, \quad \xi_j > 0, \quad y_{-j} = y_j, \quad \theta_{-j} = -\theta_j, \quad \theta_j \in \mathbb{T},\ y_j \in \mathbb{R}, \quad \forall j \in S. \tag{5.3.11}$$

For the tangential sites $S^+ := \{\overline{\jmath}_1, \ldots, \overline{\jmath}_\nu\}$ we will also denote

$$\theta_{\overline{\jmath}_i} := \theta_i, \quad y_{\overline{\jmath}_i} := y_i, \quad \xi_{\overline{\jmath}_i} := \xi_i, \quad \omega_{\overline{\jmath}_i} := \omega_i, \quad i = 1, \ldots, \nu.$$

The symplectic 2-form Ω in (1.2.5) becomes

$$\mathcal{W} := \sum_{i=1}^{\nu} d\theta_i \wedge dy_i + \frac{1}{2} \sum_{j \in S^c} \frac{1}{\mathrm{i}\omega(j)} dz_j \wedge dz_{-j} = \Big(\sum_{i=1}^{\nu} d\theta_i \wedge dy_i \Big) \oplus \Omega_{S^\perp} = d\Lambda, \tag{5.3.12}$$

where $\Omega_{S^\perp}$ denotes the restriction of Ω to $H_S^\perp$ and Λ is the Liouville 1-form on $\mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$ defined by $\Lambda_{(\theta,y,z)} \colon \mathbb{R}^\nu \times \mathbb{R}^\nu \times H_S^\perp \to \mathbb{R}$,

$$\Lambda_{(\theta,y,z)}[\hat{\theta}, \hat{y}, \hat{z}] := -y \cdot \hat{\theta} + \frac{1}{2} (J^{-1} z, \hat{z})_{L^2(\mathbb{T})}. \tag{5.3.13}$$

Working in a neighbourhood of the origin of the phase space, it is convenient to rescale the unperturbed actions ξ and the variables θ, y, z as

$$\xi \mapsto \varepsilon^2 \xi, \quad y \mapsto \varepsilon^{2b} y, \quad z \mapsto \varepsilon^b z. \tag{5.3.14}$$

The symplectic form in (5.3.12) transforms into $\varepsilon^{2b}\mathcal{W}$. Hence the Hamiltonian system generated by $\mathcal{H}$ in (5.2.7) transforms into the new Hamiltonian system

$$\begin{cases} \dot{\theta} = \partial_y H_\varepsilon(\theta, y, z), \\ \dot{y} = -\partial_\theta H_\varepsilon(\theta, y, z), \\ \dot{z} = \partial_x \nabla_z H_\varepsilon(\theta, y, z), \end{cases} \qquad H_\varepsilon := \varepsilon^{-2b} \mathcal{H} \circ A_\varepsilon, \tag{5.3.15}$$

where

$$A_\varepsilon(\theta, y, z) := \varepsilon\, v_\varepsilon(\theta, y) + \varepsilon^b z, \quad v_\varepsilon(\theta, y) := \sum_{j \in S} \sqrt{\xi_j + \varepsilon^{2(b-1)} |\omega(j)| y_j}\, e^{\mathrm{i}\theta_j} e^{\mathrm{i} j x}. \tag{5.3.16}$$

We still denote by

$$X_{H_\varepsilon} = (\partial_y H_\varepsilon, -\partial_\theta H_\varepsilon, \partial_x \nabla_z H_\varepsilon)$$

the Hamiltonian vector field in the variables $(\theta, y, z) \in \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp$. We now write explicitly the Hamiltonian defined in (5.3.15). The quadratic Hamiltonian $H^{(2)}$ in (5.2.1) becomes

$$\varepsilon^{-2b} H^{(2)} \circ A_\varepsilon = \text{const} + \sum_{j \in S^+} \omega(j)\, y_j + \frac{1}{2}\int_{\mathbb{T}} z^2\, dx, \tag{5.3.17}$$

and by (5.2.1), (5.2.12), (5.2.22), (5.2.40), (5.2.49) and (5.2.45), (5.2.50) we have (writing $v_\varepsilon := v_\varepsilon(\theta, y)$)

$$\begin{aligned} H_\varepsilon(\theta, y, z) &= e(\xi) + \alpha(\xi)\cdot y + \frac{1}{2}\int_{\mathbb{T}} z^2\, dx - \frac{\varepsilon}{2}\int_{\mathbb{T}} v_\varepsilon(\theta, y)\, z^2\, dx - \frac{\varepsilon^b}{6}\int_{\mathbb{T}} z^3\, dx \\ &\quad + \frac{\varepsilon^{2\,b}}{2}\, \mathbb{A}\Omega\, y\cdot y + \varepsilon^{2(b+1)} O(\xi y^2) + \varepsilon^{2(b+2)} O(\xi^2 y^2) + \varepsilon^{4b} O(y^3) \\ &\quad + \varepsilon^2 R((v_\varepsilon(\theta, y))^2 z^2) + \varepsilon^{1+b} R(v_\varepsilon(\theta, y)\, z^3) + \varepsilon^3 R((v_\varepsilon(\theta, y))^3 z^2) \\ &\quad + \varepsilon^{2+b} \sum_{q=3}^{5} \varepsilon^{(q-3)(b-1)} R((v_\varepsilon(\theta, y))^{5-q} z^q) + \varepsilon^{(b+3)} \sum_{q=3}^{6} \varepsilon^{(b-1)(q-3)} R(v^{6-q} z^q) \\ &\quad + \varepsilon^{4+b} \sum_{q=3}^{7} \varepsilon^{(b-1)(q-3)} R(v^{7-q} z^q) + \varepsilon^{5+b} \sum_{q=3}^{8} \varepsilon^{(b-1)(q-3)} R(v^{8-q} z^q) \\ &\quad + \varepsilon^{-2b} \mathcal{H}_{\geq 9}(\varepsilon v_\varepsilon(\theta, y) + \varepsilon^b z) \end{aligned} \tag{5.3.18}$$

where $e(\xi)$ is a constant and $\alpha(\xi)$ is the frequency-amplitude map (recall (5.3.8))

$$\alpha(\xi) = \overline{\omega} + \varepsilon^2 \mathbb{A}\xi + \varepsilon^4\, \mathcal{G}(\xi). \tag{5.3.19}$$

Note that $\mathbb{A}\Omega$ is symmetric.

Remark 5.3.1. By assumption (H2) in (1.2.15) the function (5.3.19) is a diffeomorphism for ε small enough and the system (5.3.5) is integrable and non-isochronous.

We write the Hamiltonian in (5.3.18), eliminating the constant $e(\xi)$, which is irrelevant for the dynamics, as

$$\begin{aligned} &H_\varepsilon = \mathcal{N} + P, \qquad \mathcal{N}(\theta, y, z) = \alpha(\xi)\cdot y + \frac{1}{2}(N(\theta) z, z)_{L^2(\mathbb{T})}, \\ &\frac{1}{2}(N(\theta) z, z)_{L^2(\mathbb{T})} := \frac{1}{2}((\partial_z \nabla H_\varepsilon)(\theta, 0, 0)[z], z)_{L^2(\mathbb{T})} = \frac{1}{2}\int_{\mathbb{T}} z^2\, dx \\ &- \frac{\varepsilon}{2}\int_{\mathbb{T}} v_\varepsilon(\theta, 0)\, z^2\, dx + \varepsilon^2 R((v_\varepsilon(\theta, 0))^2 z^2) + \varepsilon^3 R((v_\varepsilon(\theta, 0))^3 z^2) + \ldots \end{aligned} \tag{5.3.20}$$

where $\mathcal{N}$ describes the linear dynamics, and $P := H_\varepsilon - \mathcal{N}$ collects the nonlinear perturbative effects.

5.4 The nonliner functional setting

We look for an embedded invariant torus

$$i\colon \mathbb{T}^\nu \to \mathbb{T}^\nu \times \mathbb{R}^\nu \times H_S^\perp, \quad \varphi \mapsto i(\varphi) := (\theta(\varphi), y(\varphi), z(\varphi)) \tag{5.4.1}$$

of the Hamiltonian vector field X_{H_ε} supporting quasi-periodic solutions with diophantine frequency $\omega \in \mathbb{R}^\nu$, that we consider as independent parameters. We require that ω belongs to the set

$$\Omega_\varepsilon := \{\alpha(\xi) : \xi \in [1,2]^\nu\}, \tag{5.4.2}$$

where α is the function defined in (5.3.19) and, by Lemma B.1.1, it is a diffeomorphism for a generic choice of the tangential sites.
We shall require also some diophantine conditions on the frequencies $\omega \in \Omega_\varepsilon$. We define the sets

$$\mathcal{G}_0^{(0)} := \{\omega \in \Omega_\varepsilon : |\omega \cdot \ell| \geq \gamma \langle \ell \rangle^{-\tau}, \ \forall \ell \in \mathbb{Z}^\nu \setminus \{0\}\} \tag{5.4.3}$$

$$\begin{aligned} \mathcal{G}_0^{(1)} := \{\omega \in \Omega_\varepsilon : &|\overline{\omega} \cdot \ell + \varepsilon^2 \mathbb{A}\xi \cdot \ell + \omega(j') - \omega(j) + \varepsilon^2(\omega(j')\lambda_{j'} - \omega(j)\lambda_j)| \geq C\gamma, \\ &\sum_{i=1}^{\nu} \overline{\jmath}_i \ell_i + j' - j = 0, \ \forall |\ell| \leq 3, \ \ell \in \mathbb{Z}^\nu \setminus \{0\} \ j, j' \in S^c\} \end{aligned} \tag{5.4.4}$$

for some constant C depending on S, where $\mathbb{A}$ is defined in (5.3.9) and

$$\lambda_j := \frac{2}{3} \sum_{j_2 \in S^+} \frac{(1+j_2^2)(1+j^2)(2+j_2^2+j^2)}{(3+j_2^2-j_2 j+j^2)(3+j_2^2+j_2 j+j^2)} \xi_{j_2}.$$

We require that

$$\omega \in \mathcal{G}_0 := \mathcal{G}_0^{(0)} \cap \mathcal{G}_0^{(1)}. \tag{5.4.5}$$

Since $\omega \in \Omega_\varepsilon$ are ε^2-close to the rational vector $\overline{\omega} := (\omega(\overline{\jmath}_1), \ldots, \omega(\overline{\jmath}_\nu)) \in \mathbb{Q}^\nu$, we require that the constant γ satisfies

$$\gamma = \varepsilon^{2+a}, \quad \text{for some } a > 0 \tag{5.4.6}$$

in order to prove that the set $\mathcal{G}_0$ has large positive measure. This point will be analyzed in Section 5.9.1. Note that the definition of γ in (5.4.6) is slightly stronger than the minimal condition, namely $\gamma \leq c\,\varepsilon^2$, with $c > 0$ small enough. In addition to (5.4.5) we shall also require that ω satisfies the first and the second order Melnikov non-resonance conditions. We fix the amplitude ξ as a function of ω and ε so that (see (5.3.19))

$$\alpha(\xi) = \omega.$$

Consequently, H_ε in (5.3.20) becomes a (ω, ε)-parameter family of Hamiltonians such that for $P = 0$ possess an invariant torus at the origin with frequency ω.
Now we look for an embedded invariant torus of the modified Hamiltonian vector field $X_{H_{\varepsilon,\zeta}} = X_{H_\varepsilon} + (0, \zeta, 0), \zeta \in \mathbb{R}^\nu$, which is generated by the Hamiltonian

$$H_{\varepsilon,\zeta}(\theta, y, z) := H_\varepsilon(\theta, y, z) + \zeta \cdot \theta, \quad \zeta \in \mathbb{R}^\nu. \tag{5.4.7}$$

We introduce ζ in order to control the average in the y-component of the linearized equations (5.5.23) (see (5.5.25)). However, the vector ζ has no dynamical consequences. Indeed it turns out that an invariant torus for the Hamiltonian vector field $X_{H_{\varepsilon,\zeta}}$ is actually invariant for X_{H_ε} itself.
Thus, we look for zeros of the nonlinear operator

$$\mathcal{F}(i, \zeta) := \mathcal{F}(i, \zeta, \omega, \varepsilon) := \mathcal{D}_\omega i(\varphi) - X_{\mathcal{N}}(i(\varphi)) - X_P(i(\varphi)) + (0, \zeta, 0) \tag{5.4.8}$$

$$:= \begin{pmatrix} \mathcal{D}_\omega \theta(\varphi) - \partial_y H_\varepsilon(i(\varphi)) \\ \mathcal{D}_\omega y(\varphi) + \partial_\theta H_\varepsilon(i(\varphi)) + \zeta \\ \mathcal{D}_\omega z(\varphi) - \partial_x \nabla_z H_\varepsilon(i(\varphi)) \end{pmatrix} = \begin{pmatrix} \mathcal{D}_\omega \Theta(\varphi) - \partial_y P(i(\varphi)) \\ \mathcal{D}_\omega y(\varphi) + \frac{1}{2} \partial_\theta (N(\theta(\varphi)) z(\varphi))_{L^2(\mathbb{T})} + \partial_\theta P(i(\varphi)) + \zeta \\ \mathcal{D}_\omega z(\varphi) - \partial_x N(\theta(\varphi))\, z(\varphi) - \partial_x \nabla_z P(i(\varphi)) \end{pmatrix}$$

where $\Theta(\varphi) := \theta(\varphi) - \varphi$ is $(2\pi)^\nu$-periodic and we use the short notation

$$\mathcal{D}_\omega := \omega \cdot \partial_\varphi. \tag{5.4.9}$$

The Sobolev norm of the periodic component of the embedded torus

$$\mathfrak{I}(\varphi) := i(\varphi) - (\varphi, 0, 0) := (\Theta(\varphi), y(\varphi), z(\varphi)), \tag{5.4.10}$$

is

$$\|\mathfrak{I}\|_s := \|\Theta\|_{H^s_\varphi} + \|y\|_{H^s_\varphi} + \|z\|_s \tag{5.4.11}$$

where $\|z\|_s := \|z\|_{H^s_{\varphi,x}}$ is defined in (2.1.2).
We link the rescaling of the domain of the variables (5.3.14) with the diophantine constant $\gamma = \varepsilon^{2+a}$ by choosing

$$\gamma = \varepsilon^{2+a} = \varepsilon^{2b}, \quad b := 1 + (a/2). \tag{5.4.12}$$

Other choices are possible (see Remark 5.2 in [9]). We fix

$$s_0 := \left[\frac{\nu}{2}\right] + 3. \tag{5.4.13}$$

Theorem 5.4.1. *There exists a small constant $\mathtt{c} > 0$ such that for any choice of the tangential sites (see (1.2.8)) generic in $\mathcal{V}(\mathtt{c})$ (according to Definition 1.2.2) there exists $\varepsilon_0 > 0$, small enough, such that the following holds.*
For all $\varepsilon \in (0, \varepsilon_0)$ there exist a constant $C > 0$ and a Cantor-like set $\mathcal{C}_\varepsilon \subseteq \Omega_\varepsilon$ (see (5.4.2)), with asymptotically full measure as $\varepsilon \to 0$, namely

$$\lim_{\varepsilon \to 0} \frac{|\mathcal{C}_\varepsilon|}{|\Omega_\varepsilon|} = 1, \tag{5.4.14}$$

such that, for all $\omega \in \mathcal{C}_\varepsilon$, there exists a solution $i_\infty(\varphi) := i_\infty(\omega, \varepsilon)(\varphi)$ of the equation $\mathcal{F}(i_\infty, 0, \omega, \varepsilon) = 0$ (see (5.4.8)). Hence the embedded torus $\varphi \mapsto i_\infty(\varphi)$ is invariant for the Hamiltonian vector field X_{H_ε}, and it is filled by quasi-periodic solutions with frequency ω. The torus i_∞ satisfies

$$\|i_\infty(\varphi) - (\varphi, 0, 0)\|^{\gamma, \mathcal{C}_\varepsilon}_{s_0+\mu} \le C\, \varepsilon^{9-2b}\, \gamma^{-1} \tag{5.4.15}$$

for some $\mu := \mu(\nu) > 0$. Moreover the torus i_∞ is linearly stable.

Theorem 5.4.1 is proved in Sections 5.5-5.9. It implies Theorem 1.2.3 where the ξ_j in (1.2.17) are the components of the vector $\mathbb{A}^{-1}[\omega - \overline{\omega}]$.

Now we give tame estimates for the composition operator induced by the Hamiltonian vector fields $X_\mathcal{N}$ and X_P in (5.4.8).
Since the functions $y \to \sqrt{\xi + \varepsilon^{2(b-1)} y}, \theta \to e^{\mathrm{i}\theta}$ are analytic for ε small enough and $|y| \le C$, the composition lemma A.0.3 implies that, for all $\Theta, y \in H^s(\mathbb{T}^\nu, \mathbb{R}^\nu)$ with $\|\Theta\|_{s_0}, \|y\|_{s_0} \le 1$, one has the tame estimate

$$\|v_\varepsilon(\theta(\varphi), y(\varphi))\|_s \le_s 1 + \|\Theta\|_s + \|y\|_s. \tag{5.4.16}$$

Hence the map A_ε in (5.3.16) satisfies, for all $\|\mathfrak{I}\|^{\gamma,\mathcal{O}}_{s_0} \le 1$

$$\|A_\varepsilon(\theta(\varphi), y(\varphi), z(\varphi))\|^{\gamma,\mathcal{O}}_s \le_s \varepsilon(1 + \|\mathfrak{I}\|^{\gamma,\mathcal{O}}_s). \tag{5.4.17}$$

In the following lemma we collect tame estimates for the Hamiltonian vector fields $X_\mathcal{N}, X_P, X_{H_\varepsilon}$, see (5.3.20).

Lemma 5.4.2. *Let $\mathfrak{I}(\varphi)$ in (5.4.10) satisfy $\|\mathfrak{I}\|_{s_0+1}^{\gamma,\mathcal{O}} \leq C\,\varepsilon^{9-2b}\gamma^{-1}$. Then*

$$\|\partial_y P(i)\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^7 + \varepsilon^{2b}\|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}, \qquad \|\partial_\theta P(i)\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^{9-2b}(1+\|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}), \tag{5.4.18}$$

$$\|\nabla_z P(i)\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^{8-b} + \varepsilon^{9-b}\gamma^{-1}\|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}, \qquad \|X_P(i)\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^{9-2b} + \varepsilon^{2b}\|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}, \tag{5.4.19}$$

$$\|\partial_\theta\partial_y P(i)\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^7 + \varepsilon^8\gamma^{-1}\|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}, \qquad \|\partial_y\nabla_z P(i)\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^{6+b} + \varepsilon^{2b-1}\|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}, \tag{5.4.20}$$

$$\|\partial_{yy}P(i) - \frac{\varepsilon^{2b}}{2}\mathbb{A}\Omega\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^{5+2b} + \varepsilon^{6+2b}\gamma^{-1}\|\mathfrak{I}\|_{s+2}^{\gamma,\mathcal{O}} \tag{5.4.21}$$

and for all $\hat{\imath} := (\hat{\Theta}, \hat{y}, \hat{z})$,

$$\|\partial_y d_i X_P(i)[\hat{\imath}]\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon^{2b-1}(\|\hat{\imath}\|_{s+3}^{\gamma,\mathcal{O}} + \|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}\|\hat{\imath}\|_{s_0+3}^{\gamma,\mathcal{O}}), \tag{5.4.22}$$

$$\|d_i X_{H_\varepsilon}(i)[\hat{\imath}] + (0,0,J\,\hat{z})\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon(\|\hat{\imath}\|_{s+3}^{\gamma,\mathcal{O}} + \|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}\|\hat{\imath}\|_{s_0+3}^{\gamma,\mathcal{O}}), \tag{5.4.23}$$

$$\|d_i^2 X_{H_\varepsilon}(i)[\hat{\imath},\hat{\imath}]\|_s^{\gamma,\mathcal{O}} \leq_s \varepsilon(\|\hat{\imath}\|_{s+3}^{\gamma,\mathcal{O}}\|\hat{\imath}\|_{s_0+3}^{\gamma,\mathcal{O}} + \|\mathfrak{I}\|_{s+3}^{\gamma,\mathcal{O}}(\|\hat{\imath}\|_{s_0+3}^{\gamma,\mathcal{O}})^2). \tag{5.4.24}$$

In the sequel we will use that, by the diophantine condition (4.3.3), the operator $\mathcal{D}_\omega^{-1}$ (see (5.4.9)) is defined for all functions u with zero φ-average, and satisfies

$$\|\mathcal{D}_\omega^{-1}u\|_s \leq_s \gamma^{-1}\|u\|_{s+\tau}, \quad \|\mathcal{D}_\omega^{-1}u\|_s^{\gamma,\mathcal{O}} \leq_s \gamma^{-1}\|u\|_{s+2\tau+1}^{\gamma,\mathcal{O}}. \tag{5.4.25}$$

5.5 Approximate inverse

We will apply a Nash-Moser iterative scheme in order to find a zero of the functional $\mathcal{F}(i,\zeta)$ defined in (5.4.8). In particular, we shall construct a sequence of approximate solutions of

$$\mathcal{F}(i,\zeta) = 0 \tag{5.5.1}$$

that converges to a solution in some Sobolev norm. In order to define this sequence we need to solve some linearized equations and this is the main difficulty for implementing the Nash-Moser algorithm. Zehnder noted in [97] that it is sufficient to invert these equations only approximately to get a scheme with still quadratic speed of convergence. We refer to [97] for the precise notion of *approximate right inverse*, whose main feature is to be an *exact right inverse* when the equation is linearized at an exact solution. Hence, our aim is to construct an approximate right inverse of the linearized operator

$$d_{i,\zeta}\mathcal{F}(i_0,\zeta_0)[\hat{\imath},\hat{\zeta}] = \mathcal{D}_\omega\hat{\imath} - d_i X_{H_\varepsilon}(i_0(\varphi))[\hat{\imath}] + (0,\hat{\zeta},0) \tag{5.5.2}$$

at any approximate solution i_0 of the equation (5.5.1), and to verify that satisfies some tame estimates.
Note that $d_{i,\zeta}\mathcal{F}(i_0,\zeta_0) = d_{i,\zeta}\mathcal{F}(i_0)$ is independent of ζ_0 (see (5.4.8)).
We will implement the general strategy in [21], [22] which reduces the search of an approximate right inverse of (5.5.2) to the search of an approximate inverse on the normal directions only.

It is well known that an invariant torus i_0 with diophantine flow is isotropic (see e.g.[21]), namely the pull-back 1-form $i_0^*\Lambda$ is closed, where Λ is the Liouville 1-form in (5.3.13). This is tantamount to say that the 2-form $\mathcal{W}$ in (5.3.12) vanishes on the torus $i_0(\mathbb{T}^\nu)$, because $i_0^*\mathcal{W} = i_0^*d\Lambda = d\,i_0^*\Lambda$.

For an "approximately invariant" embedded torus i_0 the 1-form $i_0^*\Lambda$ is only "approximately closed". In order to make this statement quantitative we consider

$$i_0^*\Lambda = \sum_{k=1}^{\nu} a_k(\varphi)\, d\varphi_k, \quad a_k(\varphi) := -([\partial_\varphi \theta_0(\varphi)]^T y_0(\varphi))_k + \frac{1}{2}(\partial_{\varphi_k} z_0(\varphi), \partial_x^{-1} z_0(\varphi))_{L^2(\mathbb{T})} \tag{5.5.3}$$

and we quantify how small is

$$i_0^*\mathcal{W} = d\, i_0^*\Lambda = \sum_{1\le k<j\le\nu} A_{kj}(\varphi)\, d\varphi_k \wedge d\varphi_j, \quad A_{kj}(\varphi) := \partial_{\varphi_k} a_j(\varphi) - \partial_{\varphi_j} a_k(\varphi). \tag{5.5.4}$$

In order to get estimates for an approximate inverse we need to take in account the size of the "error" function

$$Z(\varphi) := (Z_1, Z_2, Z_3)(\varphi) := \mathcal{F}(i_0, \zeta_0)(\varphi) = \omega \cdot \partial_\varphi i_0(\varphi) - X_{H_{\varepsilon,\zeta_0}}(i_0(\varphi)), \tag{5.5.5}$$

which gives a measure of how i_0 is near to be an exact solution.
Along this section we will always assume the following hypotesis (which will be proved at each step of the Nash-Moser iteration):

- **Assumption**. The map $\omega \mapsto i_0(\omega)$ is a Lipschitz function defined on some subset $\mathcal{O}_0 \subseteq \mathcal{G}_0 \subseteq \Omega_\varepsilon$, where Ω_ε is defined in (5.4.2), and, for some $\tilde\mu := \tilde\mu(\tau,\nu) > 0$,

$$\|\mathfrak{I}_0\|_{s_0+\tilde\mu}^{\gamma,\mathcal{O}_0} \le \varepsilon^{9-2b}\gamma^{-1}, \quad \|Z\|_{s_0+\tilde\mu}^{\gamma,\mathcal{O}_0} \le \varepsilon^{9-2b}, \quad \gamma = \varepsilon^{2+a}, \quad a \ll 1, \tag{5.5.6}$$

  where $\mathfrak{I}_0(\varphi) := i_0(\varphi) - (\varphi, 0, 0)$.

The next lemma proves that if i_0 is a solution of the equation (5.5.1), then the parameter ζ has to be naught, hence the embedded torus i_0 supports a quasi-periodic solution of the "original" system with Hamiltonian H_ε.

Lemma 5.5.1. *(Lemma 6.1 in [8]) We have*

$$|\zeta_0|^{\gamma,\mathcal{O}_0} \le C\|Z\|_{s_0}^{\gamma,\mathcal{O}_0}.$$

In particular, if $\mathcal{F}(i_0,\zeta_0) = 0$ then $\zeta_0 = 0$ and the torus $i_0(\varphi)$ is invariant for the vector field X_{H_ε}.

Now we estimate the size of $i_0^*\mathcal{W}$ in terms of the error function Z.
By (5.5.3), (5.5.4) we get

$$\|A_{kj}\|_s^{\gamma,\mathcal{O}_0} \le_s \|\mathfrak{I}_0\|_{s+2}^{\gamma,\mathcal{O}_0}.$$

Moreover, we have the following bound.

Lemma 5.5.2. *(Lemma 6.2 in [8]) The coefficients $A_{kj}(\varphi)$ in (5.5.4) satisfy*

$$\|A_{kj}\|_s^{\gamma,\mathcal{O}_0} \le_s \gamma^{-1}(\|Z\|_{s+2\tau+2}^{\gamma,\mathcal{O}_0} + \|Z\|_{s_0+1}^{\gamma,\mathcal{O}_0}\|\mathfrak{I}_0\|_{s+2\tau+2}^{\gamma,\mathcal{O}_0}). \tag{5.5.7}$$

As in [21], the idea is to analyze the operator linearized at an isotropic embedded torus i_δ, because the isotropy of the torus allows to construct a symplectic set of coordinates around it for which the linear tangential dynamic and the normal one are decoupled. Thus, the linear system becomes "triangular" and the hard part is to solve the equation in the normal directions (see Section 7).
Now we see that we can slightly modify i_0 (indeed, it is sufficient to move the y-component only) to

obtain an isotropic torus i_δ, that is an approximate solution as well as i_0. At the end of this section, we will prove that we are able to construct an approximate right inverse of (5.5.2) starting from an approximate inverse of $d_{i,\zeta}\mathcal{F}(i_\delta,\zeta_0)[\hat{\imath},\hat{\zeta}]$.

In the paper we denote equivalently the differential ∂_i or d_i. We use the notation $\Delta_\varphi := \sum_{k=1}^{\nu}\partial^2_{\varphi_k}$.

Lemma 5.5.3. ***(Isotropic torus)****(Lemma 6.3 in [8]) The torus $i_\delta = (\theta_0(\varphi), y_\delta(\varphi), z_0(\varphi))$ defined by*

$$y_\delta := y_0 + [\partial_\varphi\theta_0(\varphi)]^{-T}\rho(\varphi), \quad \rho_j(\varphi) := \Delta_\varphi^{-1}\sum_{k=1}^{\nu}\partial_{\varphi_j}A_{kj}(\varphi), \tag{5.5.8}$$

is isotropic. If (5.5.6) holds, then, for some $\mathfrak{d} := \mathfrak{d}(\nu,\tau)$,

$$\|y_\delta - y_0\|_s^{\gamma,\mathcal{O}_0} \leq_s \gamma^{-1}(\|Z\|_{s+\mathfrak{d}}^{\gamma,\mathcal{O}_0}\|\mathfrak{I}_0\|_{s_0+\mathfrak{d}}^{\gamma,\mathcal{O}_0} + \|Z\|_{s_0+\mathfrak{d}}^{\gamma,\mathcal{O}_0}\|\mathfrak{I}_0\|_{s+\mathfrak{d}}^{\gamma,\mathcal{O}_0}), \tag{5.5.9}$$

$$\|\mathcal{F}(i_\delta,\zeta_0)\|_s^{\gamma,\mathcal{O}_0} \leq_s \|Z\|_{s+\mathfrak{d}}^{\gamma,\mathcal{O}_0} + \|Z\|_{s_0+\mathfrak{d}}^{\gamma,\mathcal{O}_0}\|\mathfrak{I}_0\|_{s+\mathfrak{d}}^{\gamma,\mathcal{O}_0}, \tag{5.5.10}$$

$$\|\partial_i i_\delta[\hat{\imath}]\|_s \leq_s \|\hat{\imath}\|_s + \|\mathfrak{I}_0\|_{s+\mathfrak{d}}\|\hat{\imath}\|_s. \tag{5.5.11}$$

We introduce a set of symplectic coordinates adapted to the isotropic torus i_δ. We consider the map $G_\delta\colon(\psi,\eta,w)\to(\theta,y,z)$ of the phase space $\mathbb{T}^\nu\times\mathbb{R}^\nu\times H_S^\perp$ defined by

$$\begin{pmatrix}\theta\\ y\\ z\end{pmatrix} := G_\delta\begin{pmatrix}\psi\\ \eta\\ w\end{pmatrix} := \begin{pmatrix}\theta_0(\psi)\\ y_\delta(\psi) + [\partial_\psi\theta_0(\psi)]^{-T}\eta + [(\partial_\theta\tilde{z}_0)(\theta_0(\psi))]^T J^{-1}w\\ z_0(\psi)+w\end{pmatrix} \tag{5.5.12}$$

where $\tilde{z}_0 := z_0(\theta_0^{-1}(\theta))$ (indeed $\theta_0\colon\mathbb{T}^\nu\to\mathbb{T}^\nu$ is a diffeomorphism, because $\theta_0(\varphi)-\varphi$ is small). It is proved in [21] (Lemma 6.3) that G_δ in (5.5.12) is symplectic, using that the torus i_δ is isotropic. In the new coordinates, i_δ is at the origin, i.e. $(\psi,\eta,w)=(\psi,0,0)$. The transformed Hamiltonian $K := K(\psi,\eta,w,\zeta_0)$ is (recall (5.4.7))

$$\begin{aligned}K := H_{\varepsilon,\zeta_0}\circ G_\delta = \theta_0(\psi)\cdot\zeta_0 + K_{00}(\psi) + K_{10}(\psi)\cdot\eta + (K_{01}(\psi),w)_{L^2(\mathbb{T})} + \frac{1}{2}K_{20}(\psi)\eta\cdot\eta +\\ + (K_{11}(\psi)\eta,w)_{L^2(\mathbb{T})} + \frac{1}{2}(K_{02}(\psi)w,w)_{L^2(\mathbb{T})} + K_{\geq 3}(\psi,\eta,w)\end{aligned} \tag{5.5.13}$$

where $K_{\geq 3}$ collects the terms at least cubic in the variables (η,w). At any fixed ψ, the Taylor coefficient $K_{00}(\psi)\in\mathbb{R}, K_{10}(\psi)\in\mathbb{R}^\nu, K_{01}(\psi)\in H_S^\perp, K_{20}(\psi)$ is a $\nu\times\nu$ real matrix, $K_{02}(\psi)$ is a linear self-adjoint operator of $H_S^\perp$ and $K_{11}(\psi)\colon\mathbb{R}^\nu\to H_S^\perp$.
Note that the above Taylor coefficients do not depend on the parameter ζ_0.
The Hamilton equations associated to (5.5.13) are

$$\begin{cases}\dot\psi = K_{10}(\psi) + K_{20}(\psi)\eta + K_{11}^T(\psi)w + \partial_\eta K_{\geq 3}(\psi,\eta,w)\\ \dot\eta = -\,[\partial_\psi\theta_0(\psi)]^T\zeta_0 - \partial_\psi K_{00}(\psi) - [\partial_\psi K_{10}(\psi)]^T\eta - [\partial_\psi K_{01}(\psi)]^T w -\\ \qquad -\,\partial_\psi\left(\frac{1}{2}K_{20}(\psi)\eta\cdot\eta + (K_{11}(\psi)\eta,w)_{L^2(\mathbb{T})} + \frac{1}{2}(K_{02}(\psi)w,w)_{L^2(\mathbb{T})} + K_{\geq 3}(\psi,\eta,w)\right)\\ \dot w = J(K_{01}(\psi) + K_{11}(\psi)\eta + K_{02}(\psi)w + \nabla_w K_{\geq 3}(\psi,\eta,w))\end{cases} \tag{5.5.14}$$

where $[\partial_\psi K_{10}(\psi)]^T$ is the $\nu\times\nu$ transposed matrix and $[\partial_\psi K_{01}(\psi)]^T, K_{11}^T(\psi)\colon H_S^\perp \to \mathbb{R}^\nu$ are defined by the duality relation

$$(\partial_\psi K_{01}(\psi)[\hat\psi], w)_{L^2(\mathbb{T})} = \hat\psi\cdot[\partial_\psi K_{01}(\psi)]^T w, \quad \forall\hat\psi\in\mathbb{R}^\nu, w\in H_S^\perp,$$

and similarly for K_{11}. Explicitly, for all $w\in H_S^\perp$, and denoting $\underline{e}_k$ the k-th versor of $\mathbb{R}^\nu$,

$$K_{11}^T(\psi)w = \sum_{k=1}^\nu (K_{11}^T(\psi)w\cdot\underline{e}_k)\,\underline{e}_k = \sum_{k=1}^\nu (w, K_{11}(\psi)\underline{e}_k)_{L^2(\mathbb{T})}\underline{e}_k \in\mathbb{R}^\nu. \tag{5.5.15}$$

In the next lemma we estimate the coefficients K_{00}, K_{10}, K_{01} in the Taylor expansion (5.5.13). The term K_{10} describes how the tangential frequencies vary with respect to ω. Note that on an exact solution (i_0,ζ_0) we have $K_{00}(\psi) = const, K_{10}=\omega$ and $K_{01}=0$.

Lemma 5.5.4. *(Lemma 6.4 in [8]) Assume* (5.5.6)*. Then there is $\sigma:=\sigma(\tau,\nu)$ such that*

$$\|\partial_\psi K_{00}\|_s^{\gamma,\mathcal{O}_0} + \|K_{10}-\omega\|_s^{\gamma,\mathcal{O}_0} + \|K_{01}\|_s^{\gamma,\mathcal{O}_0} \le_s \|Z\|_{s+\sigma}^{\gamma,\mathcal{O}_0} + \|Z\|_{s_0+\sigma}^{\gamma,\mathcal{O}_0}\|\mathfrak{I}_0\|_{s+\sigma}^{\gamma,\mathcal{O}_0}.$$

Remark 5.5.5. By Lemma 5.5.1 if $\mathcal{F}(i_0,\zeta_0)=0$ and, by Lemma 5.5.4, the Hamiltonian (5.5.13) simplifies to

$$K = const + \omega\cdot\eta + \frac12 K_{20}(\psi)\eta\cdot\eta + (K_{11}(\psi)\eta, w)_{L^2(\mathbb{T})} + \frac12 (K_{02}(\psi)w,w)_{L^2(\mathbb{T})} + K_{\ge 3}. \tag{5.5.16}$$

In general, the normal form (5.5.16) provides a control of the linearized equations in the normal bundle of the torus.

We now estimate K_{20}, K_{11} in (5.5.13). The norm of K_{20} is the sum of the norms of its matrix entries.

Lemma 5.5.6. *Assume* (5.5.6)*. Then for some $\sigma:=\sigma(\nu,\tau)$ we have*

$$\|K_{20} - \frac{\varepsilon^{2b}}{2}\mathbb{A}\Omega\|_s^{\gamma,\mathcal{O}_0} \le_s \varepsilon^{2b+5} + \varepsilon^{2b}\|\mathfrak{I}_0\|_{s+\sigma}^{\gamma,\mathcal{O}_0} + \varepsilon^6\gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{\gamma,\mathcal{O}_0}\|Z\|_{s+\sigma}^{\gamma,\mathcal{O}_0}, \tag{5.5.17}$$

$$\|K_{11}\eta\|_s^{\gamma,\mathcal{O}_0} \le_s \varepsilon^8\gamma^{-1}\|\eta\|_s^{\gamma,\mathcal{O}_0} + \varepsilon^{2b-1}(\|\mathfrak{I}_0\|_{s+\sigma}^{\gamma,\mathcal{O}_0} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{\gamma,\mathcal{O}_0}\|Z\|_{s+\sigma}^{\gamma,\mathcal{O}_0})\|\eta\|_{s_0}^{\gamma,\mathcal{O}_0}, \tag{5.5.18}$$

$$\|K_{11}^T w\|_s^{\gamma,\mathcal{O}_0} \le_s \varepsilon^8\gamma^{-1}\|w\|_{s+2}^{\gamma,\mathcal{O}_0} + \varepsilon^{2b-1}(\|\mathfrak{I}_0\|_{s+\sigma}^{\gamma,\mathcal{O}_0} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}^{\gamma,\mathcal{O}_0}\|Z\|_{s+\sigma}^{\gamma,\mathcal{O}_0})\|w\|_{s_0+2}^{\gamma,\mathcal{O}_0}. \tag{5.5.19}$$

In particular

$$\|K_{20} - \frac{\varepsilon^{2b}}{2}\mathbb{A}\Omega\|_{s_0}^{\gamma,\mathcal{O}_0} \le \varepsilon^9\gamma^{-1}, \quad \|K_{11}\eta\|_{s_0}^{\gamma,\mathcal{O}_0} \le \varepsilon^8\gamma^{-1}\|\eta\|_{s_0}^{\gamma,\mathcal{O}_0}, \quad \|K_{11}^T w\|_{s_0}^{\gamma,\mathcal{O}_0} \le \varepsilon^8\gamma^{-1}\|w\|_{s_0}^{\gamma,\mathcal{O}_0}.$$

We apply the linear change of variables

$$DG_\delta(\varphi,0,0)\begin{pmatrix}\hat\psi\\ \hat\eta\\ \hat w\end{pmatrix} := \begin{pmatrix}\partial_\psi\theta_0(\varphi) & 0 & 0\\ \partial_\psi y_\delta(\varphi) & [\partial_\psi\theta_0(\varphi)]^{-T} & -[(\partial_\theta\tilde z_0)(\theta_0(\varphi))]^T\partial_x^{-1}\\ \partial_\psi z_0(\varphi) & 0 & \mathrm{I}\end{pmatrix}\begin{pmatrix}\hat\psi\\ \hat\eta\\ \hat w.\end{pmatrix} \tag{5.5.20}$$

In these new coordinates the linearized operator $d_{i,\zeta}\mathcal{F}(i_\delta,\zeta_0)$ is "approximately" the operator obtained linearizing (5.5.14) at $(\psi,\eta,w,\zeta)=(\varphi,0,0,\zeta_0)$ with $\mathcal{D}_\omega$ instead of ∂_t, namely

$$\begin{pmatrix}\mathcal{D}_\omega\hat\psi - \partial_\psi K_{10}(\varphi)[\hat\psi] - K_{20}(\varphi)\hat\eta - K_{11}^T(\varphi)\hat w\\ \mathcal{D}_\omega\hat\eta + [\partial_\psi\theta_0(\varphi)]^T\hat\zeta + \partial_\psi[\partial_\psi\theta_0(\varphi)]^T[\hat\psi,\zeta_0] + \partial_{\psi\psi}K_{00}(\varphi)[\hat\psi] + [\partial_\psi K_{10}(\varphi)]^T\hat\eta + [\partial_\psi K_{01}(\varphi)]^T\hat w\\ \mathcal{D}_\omega\hat w - J\{\partial_\psi K_{01}(\varphi)[\hat\psi] + K_{11}(\varphi)\hat\eta + K_{02}(\varphi)\hat w\}.\end{pmatrix} \tag{5.5.21}$$

We give estimate on the composition operator induced by the transformation (5.5.20).

Lemma 5.5.7. *(Lemma 6.7 in [8]) Assume (5.5.6) and let $\hat{\imath} := (\hat{\psi}, \hat{\eta}, \hat{w})$. Then, for some $\sigma := \sigma(\tau, \nu)$, we have*

$$\|DG_\delta(\varphi,0,0)[\hat{\imath}]\|_s + \|DG_\delta(\varphi,0,0)^{-1}[\hat{\imath}]\|_s \leq_s \|\hat{\imath}\|_s + (\|\mathfrak{I}_0\|_{s+\sigma} + \gamma^{-1}\|\mathfrak{I}_0\|_{s+\sigma}\|Z\|_{s+\sigma})\|\hat{\imath}\|_{s_0}$$
$$\|D^2G_\delta(\varphi,0,0)[\hat{\imath}_1,\hat{\imath}_2]\|_s \leq_s \|\hat{\imath}_1\|_s\|\hat{\imath}_2\|_{s_0} + \|\hat{\imath}_1\|_{s_0}\|\hat{\imath}\|_s + (\|\mathfrak{I}_0\|_{s+\sigma} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma}\|Z\|_{s+\sigma})\|\hat{\imath}\|_{s_0}\|\hat{\imath}_2\|_{s_0}. \tag{5.5.22}$$

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 (5.5.21) it is sufficient to solve the system of equations

$$\mathbb{D}[\hat{\psi},\hat{\eta},\hat{w},\hat{\zeta}] := \begin{pmatrix} \mathcal{D}_\omega\hat{\psi} - K_{20}(\varphi)\hat{\eta} - K_{11}^T(\varphi)\hat{w} \\ \mathcal{D}_\omega\hat{\eta} + [\partial_\psi\theta_0(\varphi)]^T\hat{\zeta} \\ \mathcal{D}_\omega\hat{w} - JK_{11}(\varphi)\hat{\eta} - JK_{02}(\varphi)\hat{w} \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \tag{5.5.23}$$

which is obtained by (5.5.21) neglecting the terms that are naught at a solution, namely, by Lemmata (5.5.1) and (5.5.4), $\partial_\psi K_{10}, \partial_{\psi\psi}K_{00}, \partial_\psi K_{00}, \partial_\psi K_{01}$ and $\partial_\psi[\partial_\psi\theta_0(\varphi)]^T[\cdot,\zeta_0]$.

First, we solve the second equation, namely

$$\mathcal{D}_\omega\hat{\eta} = g_2 - [\partial_\psi\theta_0(\varphi)]\hat{\zeta}. \tag{5.5.24}$$

We choose $\hat{\zeta}$ so that the φ-average of the right hand side of (5.5.24) is zero, namely

$$\hat{\zeta} = \langle g_2\rangle_\varphi. \tag{5.5.25}$$

Note that the φ-averaged matrix $\langle(\partial_\psi\theta_0)^T\rangle_\varphi = \langle \mathrm{I} + (\partial_\psi\Theta_0)^T\rangle_\varphi = \mathrm{I}$ because $\theta_0(\varphi) = \varphi + \Theta_0(\varphi)$ and $\Theta_0(\varphi)$ is periodic. Therefore

$$\hat{\eta} = \mathcal{D}_\omega^{-1}(g_2 - [\partial_\psi\theta_0(\varphi)]^T\langle g_2\rangle_\varphi) + \langle\hat{\eta}\rangle_\varphi, \qquad \langle\hat{\eta}\rangle_\varphi \in \mathbb{R}^\nu, \tag{5.5.26}$$

where the average $\langle\hat{\eta}\rangle_\varphi$ will be fix when we deal with the first equation.
We now analyze the third equation, namely

$$\mathcal{L}_\omega\hat{w} = g_3 + JK_{11}(\varphi)\hat{\eta}, \quad \mathcal{L}_\omega := \omega\cdot\partial_\varphi - JK_{02}(\varphi). \tag{5.5.27}$$

If we fix $\hat{\eta}$, then solving the equation (5.5.27) is tantamount to invert the operator $\mathcal{L}_\omega$. For the moment we assume the following hypotesis (that will be proved in Section 8)

- **Inversion Assumption**. There exists a set $\Omega_\infty \subseteq \Omega_\varepsilon$ such that for all $\omega \in \Omega_\infty$, for every function $g \in H^{s+\mu}_{S^\perp}(\mathbb{T}^{\nu+1})$ there exists a solution $h := \mathcal{L}_\omega^{-1}g$ of the linear equation $\mathcal{L}_\omega h = g$ which satisfies

$$\|\mathcal{L}_\omega^{-1}g\|_s^{\gamma,\Omega_\infty} \leq_s \gamma^{-1}(\|g\|_{s+\sigma'}^{\gamma,\Omega_\infty} + \varepsilon\gamma^{-1}\gamma_*^{-1}\{\|\mathfrak{I}_0\|_{s+\sigma'}^{\gamma,\mathcal{O}_0} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\sigma'}^{\gamma,\mathcal{O}_0}\|Z\|_{s+\sigma'}^{\gamma,\mathcal{O}_0}\}\|g\|_{s_0}^{\gamma,\Omega_\infty}) \tag{5.5.28}$$

  for some $\sigma' := \sigma'(\tau,\nu)$.

By the above assumption, there exists a solution of (5.5.27)

$$\hat{w} = \mathcal{L}_\omega^{-1}[g_3 + J\,K_{11}(\varphi)\hat{\eta}]. \tag{5.5.29}$$

Now consider the first equation

$$\mathcal{D}_\omega \hat{\psi} = g_1 + K_{20}\hat{\eta} - K_{11}^T(\varphi)\hat{w}. \tag{5.5.30}$$

Substituting (5.5.26), (5.5.29) in the equation (5.5.30), we get

$$\mathcal{D}_\omega \hat{\psi} = g_1 + M_1(\varphi)\langle\hat{\eta}\rangle_\varphi + M_2(\varphi)g_2 + M_3(\varphi)g_3 - M_2(\varphi)[\partial_\psi \theta_0]^T \langle g_2\rangle_\varphi, \tag{5.5.31}$$

where

$$M_1(\varphi) := K_{20}(\varphi) + K_{11}^T(\varphi)\mathcal{L}_\omega^{-1}\partial_x K_{11}(\varphi), \quad M_2(\varphi) := M_1(\varphi)\mathcal{D}_\omega^{-1}, \quad M_3(\varphi) := K_{11}^T(\varphi)\mathcal{L}_\omega^{-1}. \tag{5.5.32}$$

In order to solve the equation (5.5.31) we have to choose $\langle\hat{\eta}\rangle_\varphi$ such that the right hand side in (5.5.31) has zero φ-average.
By Lemma 5.5.6 and (5.5.6), the φ-averaged matrix $\langle M_1\rangle_\varphi = \varepsilon^{2b}\mathbb{A} + O(\varepsilon^{16}\gamma^{-3})$. Therefore, for ε small, $M_\varphi[M_1]$ is invertible and $M_\varphi[M_1]^{-1} = O(\varepsilon^{-2b}) = O(\gamma^{-1})$. Thus we define

$$\langle\hat{\eta}\rangle_\varphi = -(\langle M_1\rangle_\varphi)^{-1}\{\langle g_1\rangle_\varphi + \langle M_2 g_2\rangle_\varphi + \langle M_3 g_3\rangle_\varphi - \langle M_2(\partial_\psi\theta_0)^T\rangle_\varphi \langle g_2\rangle_\varphi\}. \tag{5.5.33}$$

With this choice of $\langle\hat{\eta}\rangle_\varphi$ the equation (5.5.31) has the solution

$$\hat{\psi} := \mathcal{D}_\omega^{-1}\Big(g_1 + M_1(\varphi)\langle\hat{\eta}\rangle_\varphi + M_2(\varphi)g_2 + M_3(\varphi)g_3 - M_2(\varphi)[\partial_\psi\theta_0]^T\langle g_2\rangle_\varphi\Big). \tag{5.5.34}$$

In conclusion, we have constructed a solution $(\hat{\psi}, \hat{\eta}, \hat{w}, \hat{\zeta})$ of the linear system (5.5.23). We resume this in the following proposition, giving also estimates on the inverse of the operator $\mathbb{D}$ defined in (5.5.23).

Proposition 5.5.8. *Assume* (5.5.6) *and* (5.5.28). *Then, for all* $\omega \in \Omega_\infty$, *for all* $g := (g_1, g_2, g_3)$, *the system* (5.5.23) *has a solution* $\mathbb{D}^{-1}g := (\hat{\psi}, \hat{\eta}, \hat{w}, \hat{\zeta})$ *where* $(\hat{\psi}, \hat{\eta}, \hat{w}, \hat{\zeta})$ *are defined in* (5.5.34), (5.5.26), (5.5.29), (5.5.25). *Moreover, we have*

$$\|\mathbb{D}^{-1}g\|_s^{\gamma,\Omega_\infty} \leq_s \gamma^{-1}(\|g\|_{s+\mu}^{\gamma,\Omega_\infty} + \varepsilon\gamma^{-1}\gamma_*^{-1}\{\mathfrak{I}_0\|_{s+\mu}^{\gamma,\mathcal{O}_0} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\mu}^{\gamma,\mathcal{O}_0}\|\mathcal{F}(i_0,\zeta_0)\|_{s+\mu}^{\gamma,\mathcal{O}_0}\}\|g\|_{s_0+\mu}^{\gamma,\Omega_\infty}). \tag{5.5.35}$$

Proof. The proof follows exactly the same arguments of the Proposition 6.9 in [8]. □

Eventually we prove that the operator

$$\mathbf{T}_0 := (D\tilde{G}_\delta)(\varphi,0,0) \circ \mathbb{D}^{-1} \circ (DG_\delta(\varphi,0,0))^{-1} \tag{5.5.36}$$

is an approximate right inverse of $d_{i,\zeta}\mathcal{F}(i_0)$ where $\tilde{G}_\delta((\psi,\eta,w),\zeta)$ is the identity on the ζ-component. We denote the norm $\|(\psi,\eta,w,\zeta)\|_s^{\gamma,\mathcal{O}} := \max\{\|(\psi,\eta,w)\|_s^{\gamma,\mathcal{O}}, |\zeta|^{\gamma,\mathcal{O}}\}$.

Theorem 5.5.9. *Assume* (5.5.6) *and the inversion assumption* (5.5.28). *Then there exists* $\mu := \mu(\tau,\nu)$ *such that, for all* $\omega \in \Omega_\infty$, *for all* $g := (g_1,g_2,g_3)$, *the operator* $\mathbf{T}_0$ *defined in* (5.5.36) *satisfies*

$$\|\mathbf{T}_0 g\|_s^{\gamma,\Omega_\infty} \leq_s \gamma^{-1}(\|g\|_{s+\mu}^{\gamma,\Omega_\infty} + \varepsilon\gamma^{-1}\gamma_*^{-1}\{\|\mathfrak{I}_0\|_{s+\mu}^{\gamma,\mathcal{O}_0} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\mu}^{\gamma,\mathcal{O}_0}\|\mathcal{F}(i_0,\zeta_0)\|_{s+\mu}^{\gamma,\mathcal{O}_0}\}\|g\|_{s_0+\mu}^{\gamma,\Omega_\infty}). \tag{5.5.37}$$

It is an approximate inverse of $d_{i,\zeta}\mathcal{F}(i_0)$, *namely*

$$\begin{aligned}\|(d_{i,\zeta}\mathcal{F}(i_0)\circ\mathbf{T}_0 - \mathrm{I})g\|_s^{\gamma,\Omega_\infty} &\leq_s \varepsilon^{2b-1}\gamma^{-2}\Big(\|\mathcal{F}(i_0,\zeta_0)\|_{s_0+\mu}^{\gamma,\mathcal{O}_0}\|g\|_{s+\mu}^{\gamma,\Omega_\infty} \\ &+ \{\|\mathcal{F}(i_0,\zeta_0)\|_{s+\mu}^{\gamma,\mathcal{O}_0} + \gamma^{-1}\|\mathcal{F}(i_0,\zeta_0)\|_{s_0+\mu}^{\gamma,\mathcal{O}_0}\|\mathfrak{I}_0\|_{s+\mu}^{\gamma,\mathcal{O}_0}\}\|g\|_{s_0+\mu}^{\gamma,\Omega_\infty}\Big).\end{aligned} \tag{5.5.38}$$

Proof. The proof follows the same arguments of the Theorem 6.10 in [8]. □

5.6 The linearized operator in the normal directions

In this section we give an explicit expression of the linearized operator

$$\mathcal{L}_\omega := \omega \cdot \partial_\varphi - J\, K_{02}(\varphi). \tag{5.6.1}$$

To this aim we compute $\frac{1}{2}\,(K_{02}(\psi)w, w)_{L^2(\mathbb{T})}, w \in H_S^\perp$, which collects all the terms of $(H_\varepsilon \circ G_\delta)(\psi, 0, w)$ that are quadratic in w.
First we recall some preliminary lemmata.

Lemma 5.6.1. *(Lemma 7.1 in [8]) Let H be a Hamiltonian function of class $C^2(H_0^1(\mathbb{T}_x), \mathbb{R})$ and consider a map $\Phi(u) := u + \Psi(u)$ satisfying $\Psi(u) = \Pi_E \Psi(\Pi_E u)$, for all u, where E is a finite dimensional subspace as in* (5.2.4)*. Then*

$$\partial_u[\nabla(H \circ \Phi)](u)[h] = (\partial_u \nabla H)(\Phi(u))[h] + \mathcal{R}(u)[h], \tag{5.6.2}$$

where $\mathcal{R}(u)$ has the "finite dimensional" form

$$\mathcal{R}(u)[h] = \sum_{|j| \le C} (h, g_j(u))_{L^2(\mathbb{T})} \chi_j(u) \tag{5.6.3}$$

with $\chi_j(u) = e^{\mathrm{i}jx}$ or $g_j(u) = e^{\mathrm{i}jx}$. The remainder in (5.6.3) *is*

$$\mathcal{R}(u) = \mathcal{R}_0(u) + \mathcal{R}_1(u) + \mathcal{R}_2(u)$$

with

$$\begin{aligned} &\mathcal{R}_0(u) := (\partial_u \nabla H)(\Phi(u)) \partial_u \Psi(u), \quad \mathcal{R}_1(u) := [\partial_u \{\Psi'(u)^T\}][\cdot, \nabla H(\Phi(u))], \\ &\mathcal{R}_2(u) := [\partial_u \Psi(u)]^T (\partial_u \nabla H)(\Phi(u)) \partial_u \Phi(u). \end{aligned} \tag{5.6.4}$$

Lemma 5.6.2. *(Lemma 7.3 in [8]) Let $\mathcal{R}$ be an operator of the form*

$$\mathcal{R}h = \sum_{|j| \le C} \int_0^1 (h, g_j(\tau))_{L^2(\mathbb{T})} \chi_j(\tau)\, d\tau, \tag{5.6.5}$$

where the functions $g_j(\tau), \chi_j(\tau) \in H^s, \tau \in [0,1]$ depend in a Lipschitz way on the parameter ω. Then its matrix s-decay norm (see (2.3.1)-(2.3.2)*) satisfies*

$$|\mathcal{R}|_s^{Lip(\gamma)} \le_s \sum_{|j| \le C} \sup_{\tau \in [0,1]} (\|\chi_j(\tau)\|_s^{Lip(\gamma)} \|g_j\|_{s_0}^{Lip(\gamma)} + \|\chi_j(\tau)\|_{s_0}^{Lip(\gamma)} \|g_j(\tau)\|_s^{Lip(\gamma)}). \tag{5.6.6}$$

5.6.1 Composition with the map G_δ

In the sequel we use the fact that $\mathfrak{I}_\delta := \mathfrak{I}_\delta(\varphi; \omega) = i_\delta(\varphi; \omega) - (\varphi, 0, 0)$ satisfies, for some $\mu' > 0$,

$$\|\mathfrak{I}_\delta\|_{s_0 + \mu'}^{\gamma, \mathcal{O}_0} \le C\, \varepsilon^{9-2b} \gamma^{-1}. \tag{5.6.7}$$

We now study the Hamiltonian $K := H_\varepsilon \circ G_\delta = \varepsilon^{-2b} \mathcal{H} \circ A_\varepsilon \circ G_\delta$ (see (5.3.20)). Recalling (5.3.16), $A_\varepsilon \circ G_\delta$ has the form

$$A_\varepsilon(G_\delta(\psi, \eta, w)) = \varepsilon v_\varepsilon(\theta_0(\psi), y_\delta(\psi) + L_1(\psi)\eta + L_2(\psi)w) + \varepsilon^b(z_0(\psi) + w) \tag{5.6.8}$$

where

$$L_1(\Psi) := [\partial_\psi \theta_0(\psi)]^{-T}, \quad L_2(\psi) := [(\partial_\theta \tilde z_0)(\theta_0(\psi))]^T J^{-1}. \tag{5.6.9}$$

By Taylor formula, we develop (5.6.8) in w at $(\eta, w) = (0,0)$, and we get

$$(A_\varepsilon \circ G_\delta)(\psi, 0, w) = T_\delta(\psi) + T_1(\psi)w + T_2(\psi)[w,w] + T_{\geq 3}(\psi, w),$$

where

$$T_\delta(\psi) := A_\varepsilon(G_\delta(\psi,0,0)) = \varepsilon v_\delta(\psi) + \varepsilon^b z_0(\psi), \quad v_\delta(\psi) := v_\varepsilon(\theta_0(\psi), y_\delta(\psi)) \tag{5.6.10}$$

is the approximate isotropic torus in the phase space $H_0^1(\mathbb{T})$ (it corresponds to i_δ),

$$T_1(\psi)w := \varepsilon^{2b-1} U_1(\psi)w + \varepsilon^b w; \quad T_2(\psi)[w,w] := \varepsilon^{4b-3} U_2(\psi)[w,w] \tag{5.6.11}$$

$$U_1(\psi)w := \varepsilon \sum_{j\in S} \frac{|\omega(j)|[L_2(\psi)w]_j\, e^{\mathrm{i}[\theta_0(\psi)]_j}}{2\sqrt{\xi_j + \varepsilon^{2(b-1)}|\omega(j)|[y_\delta(\psi)]_j}}, \tag{5.6.12}$$

$$U_2(\psi)[w,w] := -\varepsilon \sum_{j\in S} \frac{\omega(j)^2\, [L_2(\psi)w]_j^2\, e^{\mathrm{i}[\theta_0(\psi)]_j}}{8\{\xi_j + \varepsilon^{2(b-1)}|\omega(j)|[y_\delta(\psi)]_j\}^{\frac{3}{2}}}, \tag{5.6.13}$$

and $T_{\geq 3}(\psi, w)$ collects all the terms of order at least cubic in w. In the notation of (5.3.16), the function $v_\delta(\Psi)$ in (5.6.10) is $v_\delta(\psi) = v_\varepsilon(\theta_0(\psi), y_\delta(\psi))$. The terms U_1, U_2 in (5.6.12), (5.6.13) are $O(1)$ in ε. Moreover, using that $L_2(\psi)$ in (5.6.9) vanishes at $z_0 = 0$, they satisfy

$$\|U_1 w\|_s \leq_s \|\mathfrak{I}_\delta\|_s \|w\|_{s_0} + \|\mathfrak{I}_\delta\|_{s_0}\|w\|_s, \quad \|U_2[w,w]\|_s \leq_s \|\mathfrak{I}_\delta\|_s \|\mathfrak{I}_\delta\|_{s_0}\|w\|_{s_0}^2 + \|\mathfrak{I}_\delta\|_{s_0}^2 \|w\|_{s_0}\|w\|_s \tag{5.6.14}$$

and also in the norm $\|\cdot\|_s^{\gamma,\mathcal{O}_0}$. We expand $\mathcal{H}$ by Taylor formula

$$\mathcal{H}(u+h) = \mathcal{H}(u) + ((\nabla\mathcal{H})(u), h)_{L^2(\mathbb{T})} + \frac{1}{2}((\partial_u \nabla\mathcal{H})(u)[h], h)_{L^2(\mathbb{T})} + O(h^3). \tag{5.6.15}$$

Specifying at $u = T_\delta(\psi)$ and $h = T_1(\psi)w + T_2(\psi)[w,w] + T_{\geq 3}(\psi, w)$, we obtain that the sum of all components of $K = \varepsilon^{-2b}(\mathcal{H}\circ A_\varepsilon \circ G_\delta)(\psi, 0, w)$ that are quadratic in w is

$$\frac{1}{2}(K_{02}w, w)_{L^2(\mathbb{T})} = \varepsilon^{-2b}((\nabla\mathcal{H})(T_\delta), T_2[w,w])_{L^2(\mathbb{T})} + \frac{\varepsilon^{-2b}}{2}((\partial_u\nabla\mathcal{H})(T_\delta)[T_1 w], T_1 w)_{L^2(\mathbb{T})}. \tag{5.6.16}$$

Inserting the expressions (5.6.12), (5.6.13) in the equality (5.6.16), we get

$$\begin{aligned} K_{02}(\psi)w =& (\partial_u\nabla\mathcal{H})(T_\delta)[w] + 2\varepsilon^{b-1}(\partial_u\nabla\mathcal{H})(T_\delta)[U_1 w] + \\ &+ \varepsilon^{2(b-1)} U_1^T (\partial_u\nabla\mathcal{H})(T_\delta)[U_1 w] + 2\,\varepsilon^{2b-3} U_2[w,\cdot]^T (\nabla\mathcal{H})(T_\delta). \end{aligned} \tag{5.6.17}$$

Lemma 5.6.3. *The operator K_{02} reads*

$$(K_{02}w, w)_{L^2(\mathbb{T})} = ((\partial_u\nabla\mathcal{H})(T_\delta)[w], w)_{L^2(\mathbb{T})} + (R(\psi)w, w)_{L^2(\mathbb{T})} \tag{5.6.18}$$

where $R(\psi)$ has the "finite dimensional" form

$$R(\psi)w = \sum_{|j|\leq C} (w, g_j(\psi))_{L^2(\mathbb{T})}\, \chi_j(\psi). \tag{5.6.19}$$

The functions g_j, χ_j *satisfy, for some* $\sigma := \sigma(\nu, \tau) > 0$,

$$\|g_j\|_s^{\gamma,\mathcal{O}_0}\|\chi_j\|_{s_0}^{\gamma,\mathcal{O}_0} + \|g_j\|_{s_0}^{\gamma,\mathcal{O}_0}\|\chi_j\|_s^{\gamma,\mathcal{O}_0} \leq_s \varepsilon^{1+b}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\gamma,\mathcal{O}_0}, \tag{5.6.20}$$

$$\begin{aligned}&\|\partial_i g_j[\hat{\imath}]\|_s\|\chi_j\|_{s_0} + \|\partial_i g_j[\hat{\imath}]\|_{s_0}\|\chi_j\|_s + \|g_j\|_s\|\partial_i\chi_j[\hat{\imath}]\|_{s_0} + \|g_j\|_{s_0}\|\partial_i\chi_j[\hat{\imath}]\|_s \\ &\leq_s \varepsilon^{1+b}\|\hat{\imath}\|_{s+\sigma} + \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s+\sigma}\end{aligned} \tag{5.6.21}$$

In conclusion, the linearized operator to analyze after the composition with the action-angle variables, the rescaling and the transformation G_δ is

$$w \mapsto (\partial_u \nabla \mathcal{H})(T_\delta)[w], \quad w \in H_S^\perp$$

up to finite dimensional operators which have form (5.6.19) and size (5.6.20).

5.6.2 The linearized operator in the normal directions

In this section we compute $((\partial_u \nabla \mathcal{H})(T_\delta)[w], w)_{L^2(\mathbb{T})}, w \in H_S^\perp$, recalling that $\mathcal{H} = H \circ \Phi_B$ and Φ_B is the Birkhoff map of Proposition 5.2.3. It is convenient to write separately the terms in

$$\mathcal{H} = H \circ \Phi_B = (H^{(2)} + H^{(3)}) \circ \Phi_B + H^{(\geq 9)} \circ \Phi_B, \tag{5.6.22}$$

where

$$H^{(2)}(u) := \frac{1}{2}\int_{\mathbb{T}} u^2\,dx, \quad H^{(3)}(u) := -\frac{1}{6}\int_{\mathbb{T}} u^3\,dx, \quad H^{(\geq 9)}(u) := \int_{\mathbb{T}} f(u)\,dx.$$

First we consider $H^{(\geq 9)} \circ \Phi_B$. By (5.2.1) we get

$$\nabla H^{(\geq 9)}(u) = \pi_0[(\partial_u f)(u)].$$

Since the Birkhoff transformation Φ_B has the form (5.2.6), Lemma 5.6.1 (at $u = T_\delta$) implies that

$$\begin{aligned}\partial_u \nabla(H^{(\geq 9)} \circ \Phi_B)(T_\delta)[h] &= (\partial_u \nabla H^{(\geq 9)})(\Phi_B(T_\delta))[h] + \mathcal{R}_{H^{(\geq 9)}}(T_\delta)[h] = \\ &= r_0(T_\delta)h + \mathcal{R}_{H^{(\geq 9)}}(T_\delta)[h]\end{aligned} \tag{5.6.23}$$

where the multiplicative function $r_0(T_\delta)$ is

$$r_0(T_\delta) := \sigma_0(\Phi_B(T_\delta)), \quad \sigma_0(u) := (\partial_{uu} f)(u), \tag{5.6.24}$$

the remainder $\mathcal{R}_{H^{(\geq 9)}}(u)$ has the form (5.6.3) with $\chi_j = e^{\mathrm{i}jx}$ or $g_j = e^{\mathrm{i}jx}$ and it satisfies, for some $\sigma := \sigma(\nu, \tau) > 0$,

$$\|g_j\|_s^{\gamma,\mathcal{O}_0}\|\chi_j\|_{s_0}^{\gamma,\mathcal{O}_0} + \|g_j\|_{s_0}^{\gamma,\mathcal{O}_0} \leq_s \varepsilon^7(1 + \|\mathfrak{I}_\delta\|_{s+2}^{\gamma,\mathcal{O}_0}),$$

$$\|\partial_i g_j[\hat{\imath}]\|_s\|\chi_j\|_{s_0} + \|\partial_i g_j[\hat{\imath}]\|_{s_0}\|\chi_j\|_s + \|g_j\|_s\|\partial_i\chi_j[\hat{\imath}]\|_{s_0} + \|g_j\|_{s_0}\|\partial_i\chi_j[\hat{\imath}]\|_s \leq_s \varepsilon^7(\|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+2}\|\hat{\imath}\|_{s_0+2}).$$

Now consider the contribution of $(H^{(2)} + H^{(3)}) \circ \Phi_B$. By Lemma 5.6.1 and (4.1.1) we have

$$\partial_u \nabla((H^{(2)} + H^{(3)}) \circ \Phi_B)(T_\delta)[h] = (1 - \Phi_B(T_\delta))h + \mathcal{R}_{H^{(2)}}(T_\delta)[h] + \mathcal{R}_{H^{(3)}}(T_\delta)[h], \tag{5.6.25}$$

where $\Phi_B(T_\delta)$ is a zero space average function, indeed Φ_B maps $H^1_0(\mathbb{T}_x)$ in itself by Proposition (5.2.3). The remainder $\mathcal{R}_{H^{(2)}}, \mathcal{R}_{H^{(3)}}$ have the form (5.6.3) and, by (5.6.4), the size $(\mathcal{R}_{H^{(2)}} + \mathcal{R}_{H^{(3)}})(T_\delta) = O(\varepsilon)$. We develop this sum as

$$(\mathcal{R}_{H^{(2)}} + \mathcal{R}_{H^{(3)}})(T_\delta) = \varepsilon \mathcal{R}_1 + \varepsilon^2 \mathcal{R}_2 + \tilde{\mathcal{R}}_{>2}, \tag{5.6.26}$$

where $\tilde{\mathcal{R}}_{>2}$ has size $o(\varepsilon^2)$. Thus we get, for all $h \in H_S^\perp$,

$$\Pi_S^\perp \partial_u \nabla((H^{(2)} + H^{(3)}) \circ \Phi_B)(T_\delta)[h] = \Pi_S^\perp[(1 - \Phi_B(T_\delta))h] + \Pi_s^\perp(\varepsilon\mathcal{R}_1 + \varepsilon^2\mathcal{R}_2 + \tilde{\mathcal{R}}_{>2})[h]. \tag{5.6.27}$$

Now we expand $\Phi_B(u) = u + \Psi_2(u) + \Psi_{\geq 3}(u)$, where $\Psi_2(u)$ is a quadratic function of u, $\Psi_{\geq 3} = O(u^3)$ and both map $H^1_0(\mathbb{T}_x)$ in itself. At $u = T_\delta = \varepsilon v_\delta + \varepsilon^b z_0$ we get

$$\Phi_B(T_\delta) = T_\delta + \Psi_2(T_\delta) + \Psi_{\geq 3}(T_\delta) = \varepsilon v_\delta + \varepsilon^2 \Psi_2(v_\delta) + \tilde{q}, \tag{5.6.28}$$

where $\tilde{q} = \varepsilon^b z_0 + \Psi_2(T_\delta) - \varepsilon^2 \Psi_2(v_\delta) + \Psi_{\geq 3}(T_\delta)$ and it satisfies

$$\|\tilde{q}\|_s^{\gamma,\mathcal{O}_0} \leq_s \varepsilon^3 + \varepsilon^b \|\mathfrak{I}_\delta\|_s^{\gamma,\mathcal{O}_0}, \quad \|\partial_i \tilde{q}[\hat{\imath}]\|_s \leq_s \varepsilon^b(\|\hat{\imath}\|_s + \|\mathfrak{I}_\delta\|_s \|\hat{\imath}\|_{s_0}). \tag{5.6.29}$$

Note that also $\tilde{q}$ has zero space average, indeed $\tilde{q} = \Phi_B(T_\delta) - \varepsilon v_\delta - \varepsilon^2 \Psi_2(v_\delta)$ and

$$\Psi_2(v) := J\nabla F^{(3)}(v) \in H^1_0(\mathbb{T}_x), \tag{5.6.30}$$

$\Phi_B(T_\delta), v_\delta$ belong to $H^1_0(\mathbb{T}_x)$.

Remark 5.6.4. We observe that the terms $O(\varepsilon)$ come from the monomials $R(v\, z^2)$ of $\mathcal{H}_3$ and the ones of size $O(\varepsilon^2)$ from $H^{(2)} + \mathcal{H}^{(4,2)}$ (see (4.1.8)). Thus, we compare (5.6.27) with $\Pi_S^\perp(\partial_u \nabla(H^{(2)} + \mathcal{H}^{(3)} + \mathcal{H}^{(4,2)}))(T_\delta)[h]$, using (4.1.8), and, by (5.6.28), we obtain $\mathcal{R}_1 = 0$ and $\mathcal{R}_2$ is the L^2-gradient of the Hamiltonian composed by the terms $R(v^2\, z^2)$ of the Poisson bracket $2^{-1}\{H^{(3,\leq 1)}, F^{(3)}\}$.

In conclusion, we have the following proposition.

Proposition 5.6.5. *Assume* (5.6.7). *Then the Hamiltonian operator $\mathcal{L}_\omega$, for all $h \in H^s_{S^\perp}(\mathbb{T}^{\nu+1})$, has the form*

$$\mathcal{L}_\omega h := \mathcal{D}_\omega h - J\, K_{02} h = \Pi_S^\perp\big(\mathcal{D}_\omega h - J\,[(1 - \Phi_B(T_\delta) - r_0(T_\delta))h] - \varepsilon^2 J\,\mathcal{R}_2 h - J\,\mathcal{R}_* h\big) \tag{5.6.31}$$

where $\mathcal{R}_2$ is defined in Remark 5.6.4,

$$\mathcal{R}_* := \tilde{\mathcal{R}}_{>2} + R_{H^{(\geq 9)}}(T_\delta) + R(\psi), \tag{5.6.32}$$

with $R(\psi)$ defined in Lemma 5.6.3, r_0 in (5.6.24), *T_δ in* (5.6.10), *$\tilde{\mathcal{R}}_{>2}$ in* (5.6.26) *and $\Phi_B(T_\delta)$ in* (5.6.28).
Furthermore, we have, for some $\sigma := \sigma(\nu, \tau) > 0$,

$$\|\Phi_B(T_\delta)\|_s^{\gamma,\mathcal{O}_0} \leq_s \varepsilon\,(1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{\gamma,\mathcal{O}_0}), \qquad \|\partial_i \Phi_B(T_\delta)[\hat{\imath}]\|_s \leq_s \varepsilon(\|i\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|i\|_{s_0+\sigma}), \tag{5.6.33}$$

where $\mathfrak{I}_\delta(\varphi) := (\theta_0(\varphi) - \varphi, y_\delta(\varphi), z_0(\varphi))$ corresponds to T_δ. The remainder $\mathcal{R}_2$ has the form (5.6.3) *with*

$$\|g_j\|_s^{\gamma,\mathcal{O}_0} + \|\chi_j\|_s^{\gamma,\mathcal{O}_0} \leq_s 1 + \|\mathfrak{I}_\delta\|_{s+\sigma}^{\gamma,\mathcal{O}_0}, \quad \|\partial_i g_j[\hat{\imath}]\|_s + \|\partial_i \chi_j[\hat{\imath}]\|_s \leq_s \|\hat{\imath}\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma} \tag{5.6.34}$$

and also $\mathcal{R}_*$ *has the form* (5.6.3) *with*

$$\|g_j^*\|_s^{\gamma,\mathcal{O}_0}\|\chi_j^*\|_{s_0}^{\gamma,\mathcal{O}_0} + \|g_j^*\|_{s_0}^{\gamma,\mathcal{O}_0}\|\chi_j^*\|_s^{\gamma,\mathcal{O}_0} \leq_s \varepsilon^3 + \varepsilon^{1+b}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\gamma,\mathcal{O}_0}, \tag{5.6.35}$$

$$\begin{aligned}&\|\partial_i g_j^*[\hat{\imath}]\|_s\|\chi_j^*\|_{s_0} + \|\partial_i g_j^*[\hat{\imath}]\|_{s_0}\|\chi_j^*\|_s + \|g_j^*\|_{s_0}\|\partial_i\chi_j^*\|_s + \|g_j^*\|_s\|\partial_i\chi_j^*\|_{s_0} \\ &\leq_s \varepsilon^{1+b}\|\hat{\imath}\|_{s+\sigma} + \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma}\|\hat{\imath}\|_{s_0+\sigma}.\end{aligned} \tag{5.6.36}$$

The linearized operator $\mathcal{L}_\omega := \mathcal{L}_\omega(\omega, i_\delta(\omega))$ depends on the parameter ω both directly and also through the dependence on the embedded torus $i_\delta(\omega)$. The estimates on the partial derivative respect to i (see (5.4.1)) allow us to control, along the Nash-Moser iteration, the Lipschitz variation of the eigenvalues of $\mathcal{L}_\omega$ with respect to ω and the approximate solution i_δ.

Hamiltonian of the linearized operator

Consider the following symplectic form in the extended phase space $\mathbb{R}^\nu \times \mathbb{R}^\nu \times H_S^\perp$

$$\Omega_e(\varphi, \eta, z) := d\eta \wedge d\varphi + \sum_{j\in S^c\setminus\{0\}} \frac{1}{\mathrm{i}\omega(j)} dz_j \wedge dz_{-j} \tag{5.6.37}$$

with the following Poisson brackets

$$\{F, G\}_e := \partial_\eta F \partial_\varphi G - \partial_\varphi F \partial_\eta G + \{F, G\} \tag{5.6.38}$$

where $\{\cdot,\cdot\}$ is defined in (1.2.6).

We denoted by $\overline{v}$ the function

$$\overline{v}(\varphi, x) := \sum_{j\in S} \sqrt{\xi_j} e^{\mathrm{i}(jx + \mathtt{1}(j)\cdot\varphi)} \tag{5.6.39}$$

where $\mathtt{1}(j)$ is the j-th vector of the canonical basis of $\mathbb{Z}^\nu$.
We observe that

$$\|v_\delta - \overline{v}\|_s^{\gamma,\mathcal{O}_0} \leq \|\mathfrak{I}_\delta\|_s^{\gamma,\mathcal{O}_0}. \tag{5.6.40}$$

In the dynamical variables (v, z) the point $(\overline{v}, 0)$ represents a torus supporting a quasi-periodic motion which is invariant for the system (5.3.20) with $P = 0$. Namely $(\overline{v}, 0)$ is the approximate solution from which we bifurcate. By (5.6.39) and (5.3.14) we see that $\overline{v}$ is rescaled with ε.
Along this section we want to point out the Hamiltonians that generate vector fields of size $\varepsilon, \varepsilon^2$ and ε^3. In order to do that we consider the map $\Phi_B(T_\delta)$ in (5.6.28) as function of two variables $x := \varepsilon v_\delta$ and $y := \varepsilon^b z_0$ (recall the definition (5.6.10)).
We Taylor expand $\Phi_B(T_\delta)(x, y)$ at $x = \varepsilon\overline{v}, y = 0$ with increment $h = \varepsilon(v_\delta - \overline{v}) + \varepsilon^b z_0$ and we get

$$\begin{aligned}\Phi_B(T_\delta) &= \Phi_B(T_\delta)(\varepsilon\overline{v}, 0) + O(h) \\ &= \varepsilon\overline{v} + \varepsilon^2\Psi_2(\overline{v}) + \varepsilon^3\Psi_3(\overline{v}) + \Psi_{\geq 4}(\varepsilon\overline{v}) + O(h)\end{aligned} \tag{5.6.41}$$

where $\Psi_{\geq 4}$ is a function with a zero at the origin of degree four. The terms $\Phi_B(T_\delta)(\varepsilon\overline{v}, 0)$ has size, up to constants, a pure power of ε, whereas the remainder denoted by $O(h)$ has size $\varepsilon\|\mathfrak{I}_\delta\|_s$ for some s, see (5.4.11), (5.6.40). In the low norm $\mathfrak{I}_\delta$ is smaller than ε^3 (recall (5.6.7)), hence, whenever we shall focus on the terms $O(\varepsilon), O(\varepsilon^2), O(\varepsilon^3)$, we will consider the Taylor expansion (5.6.41) truncated

at the remainder $\Psi_{\geq 4}(\varepsilon\overline{v}) + O(h)$.
Recall also that the frequency ω has a expansion in powers of ε (see (5.3.19)).

The Hamiltonian of the operator (5.6.31) respect to the symplectic form (5.6.37) is

$$\mathsf{H} := \mathsf{H}_0 + \varepsilon\mathsf{H}_1 + \varepsilon^2\mathsf{H}_2 + \varepsilon^3\mathsf{H}_3 + \mathsf{H}_> + \varepsilon^2\mathsf{H}_{\mathcal{R}_2} + \mathsf{H}_{\mathcal{R}_*} \tag{5.6.42}$$

with

$$\begin{aligned} &\mathsf{H}_0 = \overline{\omega}\cdot\eta + \frac{1}{2}\int_{\mathbb{T}} z^2\,dx, \quad \mathsf{H}_1 = -\frac{1}{2}\int_{\mathbb{T}} \overline{v}\,z^2\,dx, \\ &\mathsf{H}_2 = \mathbb{A}\xi\cdot\eta - \frac{1}{2}\int_{\mathbb{T}} \Psi_2(\overline{v})\,z^2\,dx, \quad \mathsf{H}_3 = -\frac{1}{2}\int_{\mathbb{T}} \Psi_3(\overline{v})\,dx \end{aligned} \tag{5.6.43}$$

and by (5.6.41), (5.6.24)

$$\begin{aligned} &\mathsf{H}_> = -\frac{1}{2}\int_{\mathbb{T}} (\Phi_B(T_\delta) - \varepsilon\overline{v} - \varepsilon^2\Psi_2(\overline{v}) - \varepsilon^3\Psi_3(\overline{v}) + r_0(T_\delta))\,z^2\,dx, \\ &\|\Phi_B(T_\delta) - \varepsilon v - \varepsilon^2\Psi_2(\overline{v}) - \varepsilon^3\Psi_3(\overline{v}) + r_0(T_\delta)\|_s^{\gamma,\mathcal{O}_0} \leq \varepsilon^4 + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma}^{\gamma,\mathcal{O}_0} \end{aligned} \tag{5.6.44}$$

for some $\sigma > 0$ and H_{R_2}, $\mathsf{H}_{\mathcal{R}_*}$ are such that $\nabla_{L^2}\mathsf{H}_{\mathcal{R}_2} = \mathcal{R}_2$, $\nabla_{L^2}\mathsf{H}_{\mathcal{R}_*} = \mathcal{R}_*$.

In the following we adopt the notation $|\cdot|^\gamma := |\cdot|^{\gamma,\Omega_\varepsilon}$.

5.7 Reduction of the linearized operator in the normal directions

In this section we conjugate, by symplectic, tame and bounded changes of coordinates, the linearized operator (5.6.31) to a diagonal one, up to smoothing remainders.
We shall require that these remainders belong to the following class of operators.

Definition 5.7.1. Fix $\mathtt{b} \in \mathbb{N}$ and consider $\mathcal{O} \subseteq \mathbb{R}^\nu$. We denote by $\mathfrak{C}_{1,\mathtt{b}} = \mathfrak{C}_{1,\mathtt{b}}(\mathcal{O})$ the set of the linear operators $A = A(\omega)\colon H^s(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$, $\omega \in \mathcal{O}$ which satisfy the following for any $s_0 \leq s \leq S_{max}$ (with possibly $S_{max} = +\infty$):

- $\langle D_x\rangle^{1/2}A\langle D_x\rangle^{1/2}$, $\langle D_x\rangle^{1/2}\partial_{\varphi_m}^{\mathtt{b}}A\langle D_x\rangle^{1/2}$, $\langle D_x\rangle^{1/2}[\partial_{\varphi_m}^{\mathtt{b}}A, \partial_x]\langle D_x\rangle^{1/2}$, for $m = 1,\dots,\nu$, $0 \leq \mathtt{b}_1 \leq \mathtt{b}$ are Lip-0-tame operators (see Definition 2.3.5) with

$$\begin{aligned} &\mathfrak{M}^\gamma_{\partial_{\varphi_m}^{\mathtt{b}_1}A}(-1,s) := \mathfrak{M}^\gamma_{\langle D_x\rangle^{1/2}\partial_{\varphi_m}^{\mathtt{b}_1}A\langle D_x\rangle^{1/2}}(0,s), \\ &\mathfrak{M}^\gamma_{\partial_{\varphi_m}^{\mathtt{b}_1}[A,\partial_x]}(-1,s) := \mathfrak{M}^\gamma_{\langle D_x\rangle^{1/2}\partial_{\varphi_m}^{\mathtt{b}_1}[A,\partial_x]\langle D_x\rangle^{1/2}}(0,s). \end{aligned} \tag{5.7.1}$$

We define

$$\mathbb{B}^\gamma_A(s,\mathtt{b}) := \max_{\substack{0\leq\mathtt{b}_1\leq\mathtt{b}\\ m=1,\dots,\nu}} \max\Big(\mathfrak{M}^\gamma_{\partial_{\varphi_m}^{\mathtt{b}_1}A}(-1,s), \mathfrak{M}^\gamma_{\partial_{\varphi_m}^{\mathtt{b}_1}[A,\partial_x]}(-1,s)\Big). \tag{5.7.2}$$

Assume now that the set $\mathcal{O}$ and the operator A depend on $i = i(\omega)$, and are well defined for $\omega \in \mathcal{O} \subseteq \mathbb{R}^\nu$ for all i satisfying (5.6.7). We consider $i_1 = i_1(\omega)$, $i_2 = i_2(\omega)$ and for $\omega \in \mathcal{O}(i_1)\cap\mathcal{O}(i_2)$ we define

$$\Delta_{12}A := A(i_1) - A(i_2). \tag{5.7.3}$$

We require the following:

- $\langle D_x\rangle^{1/2}\partial_{\varphi_m}^{\mathtt{b}}\Delta_{12}A\langle D_x\rangle^{1/2}$, $\langle D_x\rangle^{1/2}[\partial_{\varphi_m}^{\mathtt{b}}\Delta_{12}A,\partial_x]\langle D_x\rangle^{1/2}$ for $m=1,\dots,\nu$, $0\le \mathtt{b}_1\le\mathtt{b}$ are 0-tame operators (see Definition 2.3.4) with

$$\begin{aligned}\mathfrak{M}^{\gamma}_{\partial_{\varphi_m}^{\mathtt{b}_1}\Delta_{12}A}(-1,s) &:= \mathfrak{M}^{\gamma}_{\langle D_x\rangle^{1/2}\partial_{\varphi_m}^{\mathtt{b}_1}\Delta_{12}A\langle D_x\rangle^{1/2}}(0,s),\\ \mathfrak{M}^{\gamma}_{\partial_{\varphi_m}^{\mathtt{b}_1}[\Delta_{12}A,\partial_x]}(-1,s) &:= \mathfrak{M}^{\gamma}_{\langle D_x\rangle^{1/2}\partial_{\varphi_m}^{\mathtt{b}_1}[\Delta_{12}A,\partial_x]\langle D_x\rangle^{1/2}}(0,s).\end{aligned} \tag{5.7.4}$$

We define

$$\mathbb{B}_{\Delta_{12}A}(s,\mathtt{b}) := \max_{\substack{0\le\mathtt{b}_1\le\mathtt{b}\\ m=1,\dots,\nu}} \max\Big(\mathfrak{M}_{\partial_{\varphi_m}^{\mathtt{b}_1}\Delta_{12}A}(-1,s), \mathfrak{M}^{\gamma}_{\partial_{\varphi_m}^{\mathtt{b}_1}[\Delta_{12}A,\partial_x]}(-1,s)\Big). \tag{5.7.5}$$

Remark 5.7.2. We note that in Proposition 5.6.5 we have proved C^1 dependence on the embedding i, while in the class (5.7.1) we are just requiring a "Lipschitz" regularity. This is due to the fact that the set $\mathcal{O}$ on which operators are defined, after the reduction procedure, depends on the embedding i; hence we require weaker conditions (see 5.7.4).

We will discuss the choice of the number $\mathtt{b}$ in Section 5.8, see (5.8.1).

For notational convenience we write the linearized operator (5.6.31), defined for $\omega\in\mathcal{O}_0\subset\Omega_\varepsilon$, as

$$\mathcal{L}_\omega = \Pi_S^{\perp}\big(\mathcal{D}_\omega - J\circ(1+a_0(\varphi,x)) + \mathcal{Q}_0\big) \tag{5.7.6}$$

where we have, for some $\sigma_0>0$ (recall (5.6.28) and (5.6.24)),

$$a_0(\varphi,x) := -(\Phi_B(T_\delta)+r_0(T_\delta)), \quad \mathcal{Q}_0 := -J(\varepsilon^2\mathcal{R}_2+\mathcal{R}_*). \tag{5.7.7}$$

$$\|a_0\|_s^{\gamma,\mathcal{O}_0} \le_s \varepsilon(1+\|\mathfrak{I}_\delta\|_{s+\sigma_0}^{\gamma,\mathcal{O}_0}), \qquad \|\partial_i a_0[\hat\imath]\|_s \le_s \varepsilon(\|\hat\imath\|_{s+\sigma_0}+\|\mathfrak{I}_\delta\|_{s+\sigma_0}\|\hat\imath\|_{s_0}) \tag{5.7.8}$$

In particular $\mathcal{Q}_0$ has the finite-dimensional form (5.6.19) and satisfies the following estimate

$$\mathfrak{M}^{\gamma}_{\mathcal{Q}_0}(s) \le_s \varepsilon^2(1+\|\mathfrak{I}_\delta\|_{s+\sigma_0}^{\gamma,\mathcal{O}_0}) \tag{5.7.9}$$

and by (5.6.34) and (5.6.35)

$$\mathfrak{M}_{\partial_i\mathcal{Q}_0[\hat\imath]}(s) \le_s \varepsilon^2\|\hat\imath\|_{s+\sigma_0} + \varepsilon^{2b-1}\|\mathfrak{I}_\delta\|_{s+\sigma_0}\|\hat\imath\|_{s_0+\sigma_0}. \tag{5.7.10}$$

Now we state the precise result we want to prove in this Section.

Theorem 5.7.3. *Consider $\mathcal{L}_\omega=\mathcal{L}_\omega(\mathfrak{I}_\delta)$ in (5.7.6) and fix $\mathtt{b}=s_0+6\tau+6$. There exists $\sigma>0$ such that, if condition (5.6.7) is satisfied with $\mu'=\sigma$, then the following holds. There exists a constant $m(\omega)$ defined for $\omega\in\Omega_\varepsilon$ with*

$$|m-1|^{\gamma,\Omega_\varepsilon}\le C\varepsilon^2, \quad |m|^{lip}\le C \tag{5.7.11}$$

such that for all ω in the set $\mathcal{O}_\infty^{2\gamma}$ (see (5.4.5)), where (recall that $\mathcal{O}_0\subseteq\mathcal{G}_0$, see (5.4.5))

$$\mathcal{O}_\infty^{2\gamma}=\mathcal{O}_\infty^{2\gamma}(i) := \{\omega\in\mathcal{O}_0:\ |\omega\cdot\ell - m(\omega)j| > \frac{2\gamma}{\langle\ell\rangle^{\tau}},\ \forall\ell\in\mathbb{Z}^\nu,\ \forall j\in\mathbb{Z}\setminus\{0\}\}, \tag{5.7.12}$$

there exists a real, bounded linear operator $\Upsilon = \Upsilon(\omega) : H^s_{S^\perp} \to H^s_{S^\perp}$ *such that*

$$\mathcal{L} := \Upsilon \mathcal{L}_\omega \Upsilon^{-1} = \Pi_S^\perp \Big(\mathcal{D}_\omega - mJ - \varepsilon^2 \mathfrak{D}(\xi) + \mathcal{R} \Big) \tag{5.7.13}$$

where $\mathfrak{D}(\xi)$ *is the diagonal operator of order* -1 *defined as*

$$\begin{aligned} \mathfrak{D} &:= \mathfrak{D}(\xi) = \mathrm{diag}(\mathrm{i}\kappa_j)_{j \in S^c}, \\ \kappa_j &= -\frac{2}{3}\omega(j) \sum_{j_2 \in S^+} \frac{(1 + j_2^2)(7 + 5j_2^2 + j_2^4 + 3j^2)}{(3 + j_2^2 - j_2 j + j^2)(3 + j_2^2 + j_2 j + j^2)} \xi_{j_2} \in \mathbb{R}. \end{aligned} \tag{5.7.14}$$

The constant m *depends on* i *and for* $\omega \in \mathcal{O}^{2\gamma}_\infty(i_1) \cap \mathcal{O}^{2\gamma}_\infty(i_2)$ *one has*

$$|\Delta_{12} m| \le \varepsilon \|i_1 - i_2\|_{s_0+\sigma}, \tag{5.7.15}$$

where $\Delta_{12} m := m(i_1) - m(i_2)$*. The* $\mathcal{R}$ *belongs to* $\mathfrak{C}_{1,\mathtt{b}} = \mathfrak{C}_{1,\mathtt{b}}(\mathcal{O}^{2\gamma}_\infty)$ *(see Def. 5.7.1) with*

$$\begin{aligned} \mathbb{B}^\gamma_{\mathcal{R}}(s, \mathtt{b}) &\le_s \varepsilon^{4-3a} + \varepsilon\gamma^{-1} \|\mathfrak{I}_\delta\|^{\gamma,\mathcal{O}_0}_{s+\sigma}, \\ \mathbb{B}_{\Delta_{12}\mathcal{R}}(s, \mathtt{b}) &\le_s \varepsilon\gamma^{-1} (\|i_1 - i_2\|_{s+\sigma} + \|\mathfrak{I}_\delta\|_{s+\sigma} \|i_1 - i_2\|_{s_0+\sigma}). \end{aligned} \tag{5.7.16}$$

Moreover if $u = u(\omega)$ *depends on the parameter* $\omega \in \mathcal{O}^{2\gamma}_\infty$ *in a Lipschitz way then*

$$\|\Upsilon u\|^{\gamma,\mathcal{O}^{2\gamma}_\infty}_s \le_s \|u\|^{\gamma,\mathcal{O}^{2\gamma}_\infty}_s + \varepsilon\gamma^{-1} \|\mathfrak{I}_\delta\|^{\gamma,\mathcal{O}_0}_{s+\sigma} \|u\|^{\gamma,\mathcal{O}^{2\gamma}_\infty}_{s_0}. \tag{5.7.17}$$

The rest of the Section is devoted to the proof of the Theorem 5.7.3. First we exploit the pseudo differential structure of the operator $\mathcal{L}_\omega$ in order to conjugate it to an operator which has constant coefficients up to a smoothing remainder of order -1, but such that the bounds (5.7.16) do not hold.

Then we apply a Linear Birkhoff Normal Form in order to reduce the size of this remainder and achieve (5.7.16). In order to perform the first step we need some abstract results on the flows of pseudo differential hyperbolic PDEs, which we shall use as changes of coordinates for our purposes. In particular we need to work in a smaller class of operators with respect to the class in Definition 5.7.1, because we need some precise information on the pseudo differential structure.

Definition 5.7.4. Fix $\rho \in \mathbb{N}$, with $\rho \ge 3$ and consider any subset $\mathcal{O}$ of $\mathbb{R}^\nu$ (recall (5.4.2)). We denote by $\mathfrak{L}_\rho = \mathfrak{L}_\rho(\mathcal{O})$ the set of the linear operators $A = A(\omega)\colon H^s(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$, $\omega \in \mathcal{O}$ with the following properties:

- A is Lipschitz in ω,
- the operators $\partial_\varphi^{\vec{\mathtt{b}}} A$, $[\partial_\varphi^{\vec{\mathtt{b}}} A, \partial_x]$, for all $\vec{\mathtt{b}} = (\mathtt{b}_1, \ldots, \mathtt{b}_\nu) \in \mathbb{N}^\nu$ with $0 \le |\vec{\mathtt{b}}| \le \rho - 2$ have the following properties, for any $s_0 \le s \le S_{max}$, with possibly $S_{max} = +\infty$:

  (i) for any $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \ge 0$ and $m_1 + m_2 = \rho - |\vec{\mathtt{b}}|$ one has that the operator $\langle D_x \rangle^{m_1} \partial_\varphi^{\vec{\mathtt{b}}} A \langle D_x \rangle^{m_2}$ is Lip-0-tame according to Def. 2.3.5 and we set

$$\mathfrak{M}^\gamma_{\partial_\varphi^{\vec{\mathtt{b}}} A}(-\rho + |\vec{\mathtt{b}}|, s) := \sup_{\substack{m_1 + m_2 = \rho - |\vec{\mathtt{b}}| \\ m_1, m_2 \ge 0}} \mathfrak{M}^\gamma_{\langle D_x \rangle^{m_1} \partial_\varphi^{\vec{\mathtt{b}}} A \langle D_x \rangle^{m_2}}(0, s); \tag{5.7.18}$$

(ii) for any $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho - |\vec{\mathtt{b}}| - 1$ one has that the operator $\langle D_x\rangle^{m_1}[\partial_\varphi^{\vec{\mathtt{b}}} A, \partial_x]\langle D_x\rangle^{m_2}$ is Lip-0-tame according to Def. 2.3.5 and we set

$$\mathfrak{M}^\gamma_{[\partial_\varphi^{\vec{\mathtt{b}}} A,\partial_x]}(-\rho + |\vec{\mathtt{b}}| + 1, s) := \sup_{\substack{m_1+m_2=\rho-|\vec{\mathtt{b}}|-1 \\ m_1,m_2\geq 0}} \mathfrak{M}^\gamma_{\langle D_x\rangle^{m_1}[\partial_\varphi^{\vec{\mathtt{b}}} A,\partial_x]\langle D_x\rangle^{m_2}}(0, s). \tag{5.7.19}$$

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

$$\mathbb{M}^\gamma_A(s, \mathtt{b}) := \max_{0\leq|\vec{\mathtt{b}}|\leq\mathtt{b}} \max\left(\mathfrak{M}^\gamma_{\partial_\varphi^{\vec{\mathtt{b}}} A}(-\rho + |\vec{\mathtt{b}}|, s), \mathfrak{M}^\gamma_{\partial_\varphi^{\vec{\mathtt{b}}}[A,\partial_x]}(-\rho + |\vec{\mathtt{b}}| + 1, s)\right). \tag{5.7.20}$$

Assume now that the set $\mathcal{O}$ and the operator A depend on $i = i(\omega)$, and are well defined for $\omega \in \mathcal{O}_0 \subseteq \Omega_\varepsilon$ for all i satisfying (5.6.7). We consider $i_1 = i_1(\omega)$, $i_2 = i_2(\omega)$ and for $\omega \in \mathcal{O}(i_1) \cap \mathcal{O}(i_2)$ we require the following (see (5.7.3)):

- The operators $\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A$, $[\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A, \partial_x]$, for $0 \leq |\vec{\mathtt{b}}| \leq \rho - 3$, have the following properties, for any $s_0 \leq s \leq S_{max}$, with possibly $S_{max} = +\infty$:

(iii) for any $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho - |\vec{\mathtt{b}}| - 1$ one has that the operator $\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A\langle D_x\rangle^{m_2}$ is 0-tame according to Def. 2.3.4 and we set

$$\mathfrak{M}_{\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A}(-\rho + |\vec{\mathtt{b}}| + 1, s) := \sup_{\substack{m_1+m_2=\rho-|\vec{\mathtt{b}}|-1 \\ m_1,m_2\geq 0}} \mathfrak{M}_{\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A\langle D_x\rangle^{m_2}}(0, s); \tag{5.7.21}$$

(iv) for any $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho - |\vec{\mathtt{b}}| - 2$ one has that the operator $\langle D_x\rangle^{m_1}[\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A, \partial_x]\langle D_x\rangle^{m_2}$ is 0-tame according to Def. 2.3.4 and we set

$$\mathfrak{M}_{[\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A,\partial_x]}(-\rho + |\vec{\mathtt{b}}| + 2, s) := \sup_{\substack{m_1+m_2=\rho-\mathtt{b}-2 \\ m_1,m_2\geq 0}} \mathfrak{M}_{\langle D_x\rangle^{m_1}[\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A,\partial_x]\langle D_x\rangle^{m_2}}(0, s). \tag{5.7.22}$$

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

$$\mathbb{M}_{\Delta_{12}A}(s, \mathtt{b}) := \max_{0\leq|\vec{\mathtt{b}}|\leq\mathtt{b}} \max\left(\mathfrak{M}_{\partial_\varphi^{\vec{\mathtt{b}}}\Delta_{12}A}(-\rho + |\vec{\mathtt{b}}| + 1, s), \mathfrak{M}_{\partial_\varphi^{\vec{\mathtt{b}}}[\Delta_{12}A,\partial_x]}(-\rho + |\vec{\mathtt{b}}| + 2, s)\right). \tag{5.7.23}$$

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

5.7.1 Properties of pseudo differential operators and the class $\mathfrak{L}_\rho$

In the first step of our reduction procedure we work with operators which are pseudo differential up to a remainder in the class $\mathfrak{L}_\rho$. In the following we shall study properties of such operators under composition, inversion etc...

The following Lemma guarantees that the class of operators in 5.7.4 is closed under composition.

Lemma 5.7.5. *If A and B belong to $\mathfrak{L}_\rho$, for $\rho \geq 0$ (see Def. 5.7.4) , then $A \circ B \in \mathfrak{L}_\rho$ and*

$$\mathbb{M}^\gamma_{A\circ B}(s, \mathtt{b}) \leq_{s,\rho} \sum_{\mathtt{b}_1+\mathtt{b}_2=\mathtt{b}} \left(\mathbb{M}^\gamma_A(s_0, \mathtt{b}_1)\mathbb{M}^\gamma_B(s, \mathtt{b}_2) + \mathbb{M}^\gamma_A(s, \mathtt{b}_1)\mathbb{M}^\gamma_B(s_0, \mathtt{b}_2)\right), \tag{5.7.24}$$

$$\begin{aligned}\mathbb{M}_{\Delta_{12}(A\circ B)}(s,\mathtt{b}) \leq_{s,\rho} \sum_{\mathtt{b}_1+\mathtt{b}_2=\mathtt{b}} & (\mathbb{M}_{\Delta_{12}A}(s,\mathtt{b}_1)\mathbb{M}_B(s_0,\mathtt{b}_2)+\mathbb{M}_{\Delta_{12}A}(s_0,\mathtt{b}_1)\mathbb{M}_B(s,\mathtt{b}_2)\\ & +\mathbb{M}_A(s_0,\mathtt{b}_1)\mathbb{M}_{\Delta_{12}B}(s,\mathtt{b}_2)+\mathbb{M}_A(s,\mathtt{b}_1)\mathbb{M}_{\Delta_{12}B}(s_0,\mathtt{b}_2)\ )\ ,\end{aligned} \tag{5.7.25}$$

for $\mathtt{b}\leq\rho-2$ *and* $s_0\leq s\leq S_{max}$.

Proof. We start by noting that $\mathfrak{M}^{\gamma}_{A\circ B}(-\rho,s)$ defined in (5.7.18) with $A\rightsquigarrow A\circ B$ is controlled by the r.h.s. of (5.7.24). Let $m_1,m_2\in\mathbb{R}$, $m_1,m_2\geq 0$ and $m_1+m_2=\rho$. We can write

$$\langle D_x\rangle^{m_1}A\circ B\langle D_x\rangle^{m_2}=\langle D_x\rangle^{m_1}A\langle D_x\rangle^{m_2}\langle D_x\rangle^{-\rho}\langle D_x\rangle^{m_1}B\langle D_x\rangle^{m_2}.$$

By hypothesis we know that A belongs to the class $\mathcal{L}_\rho$, hence by item (i) of Definition 5.7.4 one has that $\langle D_x\rangle^{m_1}A\langle D_x\rangle^{m_2}$ is a $0-$tame operator. For the same reason also $\langle D_x\rangle^{m_1}B\langle D_x\rangle^{m_2}$ is a $0-$tame operator. Note also that, since $\rho\geq 0$, then $\langle D_x\rangle^{-\rho}:H^s(\mathbb{T}^{d+1})\to H^s(\mathbb{T}^{d+1})$ is a $0-$tame operator. Hence, using Lemma 2.3.6 for any $u\in H^s$ one has

$$\begin{aligned}\|\langle D_x\rangle^{m_1}A\circ B\langle D_x\rangle^{m_2}u\|_s \leq_s & (\mathfrak{M}_A(-\rho,s)\mathfrak{M}_B(-\rho,s_0)+\mathfrak{M}_A(-\rho,s_0)\mathfrak{M}_B(-\rho,s))\|u\|_s\\ & +\mathfrak{M}_A(-\rho,s_0)\mathfrak{M}_B(-\rho,s_0)\|u\|_s,\end{aligned} \tag{5.7.26}$$

where $\mathfrak{M}_A(-\rho,s)$, $\mathfrak{M}_B(-\rho,s)$ are defined in (5.7.18). Then we may set

$$\mathfrak{M}_{A\circ B}(-\rho,s)=C(s)\Big(\mathfrak{M}_A(-\rho,s)\mathfrak{M}_B(-\rho,s_0)+\mathfrak{M}_A(-\rho,s_0)\mathfrak{M}_B(-\rho,s)\Big).$$

Reasoning as in (5.7.26) one can check that

$$\mathfrak{M}^{\gamma}_{A\circ B}(-\rho,s)\leq C(s)\Big(\mathfrak{M}^{\gamma}_A(-\rho,s)\mathfrak{M}^{\gamma}_B(-\rho,s_0)+\mathfrak{M}^{\gamma}_A(-\rho,s_0)\mathfrak{M}^{\gamma}_B(-\rho,s)\Big).$$

Let us study the operator $\partial_\varphi^{\vec{\mathtt{b}}}(A\circ B)$ for $\vec{\mathtt{b}}\in\mathbb{N}^\nu$ and $|\vec{\mathtt{b}}|\leq\rho-2$. We have

$$\partial_\varphi^{\vec{\mathtt{b}}}(A\circ B)=\sum_{\vec{\mathtt{b}_1}+\vec{\mathtt{b}_2}=\vec{\mathtt{b}}}(\partial_\varphi^{\vec{\mathtt{b}_1}}A)(\partial_\varphi^{\vec{\mathtt{b}_2}}B). \tag{5.7.27}$$

We show that any summand in (5.7.27) satisfies item (i) of Def. (5.7.4). Let $m_1,m_2\in\mathbb{R}$, $m_1,m_2\geq 0$ and $m_1+m_2=\rho-|\vec{\mathtt{b}}|$. We write

$$\langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}_1}}A)(\partial_\varphi^{\vec{\mathtt{b}_2}}B)\langle D_x\rangle^{m_2}=\langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}_1}}A)\langle D_x\rangle^{y}\langle D_x\rangle^{-y-w}\langle D_x\rangle^{w}(\partial_\varphi^{\vec{\mathtt{b}_2}}B)\langle D_x\rangle^{m_2}$$

with $y:=\rho-|\vec{\mathtt{b}_1}|-m_1$, $w=\rho-|\vec{\mathtt{b}_2}|-m_2$ and note that $-y-w=-\rho\leq 0$. Moreover $m_1+y=\rho-|\vec{\mathtt{b}_1}|$, and $w+m_2=\rho-|\vec{\mathtt{b}_2}|$. Hence the operators $\langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}_1}}A)\langle D_x\rangle^{y}$ and $\langle D_x\rangle^{w}(\partial_\varphi^{\vec{\mathtt{b}_2}}b)\langle D_x\rangle^{m_2}$ are Lip-0-tame operator. Hence, using Lemma 2.3.6 one has

$$\begin{aligned}\|\langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}_1}}A)(\partial_\varphi^{\vec{\mathtt{b}_2}}B)\langle D_x\rangle^{m_2}\|_s^{\gamma,\mathcal{O}} \leq & \mathfrak{M}^{\gamma}_{\partial_\varphi^{\vec{\mathtt{b}_1}}A}(-\rho+|\vec{\mathtt{b}_1}|,s)\mathfrak{M}^{\gamma}_{\partial_\varphi^{\vec{\mathtt{b}_2}}B}(-\rho+|\vec{\mathtt{b}_2}|,s_0)\|u\|_s\\ & +\mathfrak{M}^{\gamma}_A(-\rho+|\vec{\mathtt{b}_1}|,s_0)\mathfrak{M}^{\gamma}_B(-\rho+|\vec{\mathtt{b}_2}|,s)\|u\|_s\\ & +\mathfrak{M}^{\gamma}_A(-\rho+|\vec{\mathtt{b}_1}|,s_0)\mathfrak{M}^{\gamma}_B(-\rho+|\vec{\mathtt{b}_2}|,s_0)\|u\|_s,\end{aligned} \tag{5.7.28}$$

for $u\in H^s$. We can conclude that $\mathfrak{M}^{\gamma}_{\partial_\varphi^{\vec{\mathtt{b}}}(A\circ B)}(-\rho+|\vec{\mathtt{b}}|,s)$ is controlled by the r.h.s. of (5.7.24). Regarding the operator $[A\circ B,\partial_x]$ we reason as follows. We prove that

$$[A\circ B,\partial_x]=A[B,\partial_x]+[A,\partial_x]B. \tag{5.7.29}$$

satisfies item (ii) of Definition (5.7.4). Let $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1+m_2 = \rho-1$. Moreover

$$\langle D_x\rangle^{m_1}[A,\partial_x]B\langle D_x\rangle^{m_2} = \langle D_x\rangle^{m_1}[A,\partial_x]\langle D_x\rangle^{y}\langle D_x\rangle^{-y-z}\langle D_x\rangle^{z}B\langle D_x\rangle^{m_2},$$

with $y = \rho - 1 - m_1$, $z = \rho - m_2$. Hence by definition (see Def. (5.7.4)) we have that the operators $\langle D_x\rangle^{m_1}[A,\partial_x]\langle D_x\rangle^{y}$ and $\langle D_x\rangle^{z}B\langle D_x\rangle^{m_2}$ are Lip-0−tame. Thus one can conclude, as done above, that $\mathfrak{M}_{[A,\partial_x]B}(-\rho+1,s)$ is controlled by the r.h.s. of (5.7.24). One can reason in the same way for the first summand in (5.7.29) and for the operator $[\partial_{\varphi}^{\vec{\mathtt{b}}}(AB),\partial_x]$. This proves (5.7.24).
Let us study the term

$$\Delta_{12}(A\circ B) = \Delta_{12}AB + A\Delta_{12}B. \tag{5.7.30}$$

By definition both $\langle D_x\rangle^{m_1}\Delta_{12}A\langle D_x\rangle^{m_2}, \langle D_x\rangle^{m_1}\Delta_{12}B\langle D_x\rangle^{m_2}$ with $m_1+m_2 = \rho-1$ are 0-tame operators (see (5.7.23) and Def. 2.3.4). In order to prove (5.7.25) one can bound the tameness constant of the two summand in (5.7.30) by following the same procedure used to prove (5.7.24). ☐

The next Lemma shows that $S^{-\rho} \subset \mathfrak{L}_{\rho}$.

Lemma 5.7.6. *Fix $\rho \in \mathbb{N}$ and consider a symbol $a = a(\omega, i(\omega))$ in $S^{-\rho}$ depending on $\omega \in \mathcal{O} \subset \mathbb{R}^{\nu}$ and on i in a Lipschitz way. One has that $A := \mathrm{op}(a(\varphi,x,\xi)) \in \mathfrak{L}_{\rho}$ (see 5.7.4) and*

$$\begin{aligned}\mathbb{M}^{\gamma}_{A}(s,\mathtt{b}) &\leq_{s,\rho} |a|^{\gamma,\mathcal{O}}_{-\rho,s+\rho,0},\\ \mathbb{M}_{\Delta_{12}A}(s,\mathtt{b}) &\leq_{s,\rho} |\Delta_{12}a|_{-\rho+1,s+\rho-1,0}.\end{aligned} \tag{5.7.31}$$

Proof. Let $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1+m_2 = \rho$. We need to show that $\langle D_x\rangle^{m_1}A\langle D_x\rangle^{m_2}$ satisfies item (i) of Definition 5.7.4. By definition it is the composition of three pseudo differential operators hence, by Lemma 2.3.8 and by formula (2.2.13) of Lemma 2.2.5 one has that

$$\mathfrak{M}^{\gamma}_{\langle D_x\rangle^{m_1}A\langle D_x\rangle^{m_2}}(0,s) \leq_s |\langle D_x\rangle^{m_1}A\langle D_x\rangle^{m_2}|^{\gamma,\mathcal{O}}_{0,s,0} \tag{5.7.32}$$

$$\leq_s |\langle D_x\rangle^{m_1}|_{m_1,s,0}|a|^{\gamma,\mathcal{O}}_{-\rho,s+|m_1|,0}|\langle D_x\rangle^{m_2}|_{m_2,s+|m_1|+\rho,0} \leq_s |a|^{\gamma,\mathcal{O}}_{-\rho,s+|m_1|,0} \tag{5.7.33}$$

This means that

$$\mathfrak{M}^{\gamma}_{A}(-\rho,s) \leq_s |a|^{\gamma,\mathcal{O}}_{-\rho,s+\rho,0}.$$

Secondly we consider the operator $(\partial_{\varphi}^{\vec{\mathtt{b}}}\mathrm{op}(a(\varphi,x,\xi))) = \mathrm{op}(\partial_{\varphi}^{\vec{\mathtt{b}}}a(\varphi,x,\xi))$ for $\vec{\mathtt{b}} \in \mathbb{N}^{\nu}$ and $|\vec{\mathtt{b}}| \leq \rho-2$. It is pseudo differential and its symbol $\partial_{\varphi}^{\vec{\mathtt{b}}}a(\varphi,x,\xi)$ is such that

$$|\partial_{\varphi}^{\vec{\mathtt{b}}}a|^{\gamma,\mathcal{O}}_{-\rho,s,\alpha} \leq |a|^{\gamma,\mathcal{O}}_{-\rho,s+|\vec{\mathtt{b}}|,\alpha}.$$

Following the same reasoning used in (5.7.32) one obtains

$$\mathfrak{M}^{\gamma}_{\partial_{\varphi}^{\vec{\mathtt{b}}}A}(-\rho+|\vec{\mathtt{b}}|,s) \leq_s |a|^{\gamma,\mathcal{O}}_{-\rho,s+|\vec{\mathtt{b}}|+\rho-|\vec{\mathtt{b}}|,0}.$$

The operator $[A,\partial_x] = A\partial_x - \partial_x A$ can be treated in the same way, discussing each of the two summands separately, (we are not taking advantage of the pseudo dfferential structure in order to control the order of the commutator).

$$\mathfrak{M}^{\gamma}_{\langle D_x\rangle^{m_1}\partial_x A\langle D_x\rangle^{m_2}}(0,s) \leq_s |\langle D_x\rangle^{m_1}\partial_x A\langle D_x\rangle^{m_2}|^{\gamma,\mathcal{O}}_{0,s,0} \leq_s |a|^{\gamma,\mathcal{O}}_{-\rho,s+\rho,0}$$

The same strategy holds for $[\partial_{\varphi}^{\vec{\mathtt{b}}}A,\partial_x]$ Hence one gets the first of (5.7.31). The second bound in (5.7.31) can be obtained by noting that $\Delta_{12}A = \mathrm{op}(\Delta_{12}a)[\cdot]$ and then following almost word by word the discussion above. ☐

The next Lemma shows that $\mathfrak{L}_\rho$ is closed under left and right multiplication by operators in S^0.

Lemma 5.7.7. *Let $a \in S^0$ and $B \in \mathfrak{L}_\rho$, then $\mathrm{Op}(a) \circ B, B \circ \mathrm{Op}(a) \in \mathfrak{L}_\rho$ and satisfy the bounds*

$$\begin{aligned} \mathbb{M}^\gamma_{\mathrm{Op}(a)\circ B}(s,\mathtt{b}) &\le_{s,\rho} |a|^{\gamma,\mathcal{O}}_{0,s+\rho,0}\mathbb{M}^\gamma_B(s_0,\mathtt{b}) + |a|^{\gamma,\mathcal{O}}_{0,s_0+\rho,0}\mathbb{M}^\gamma_B(s,\mathtt{b}) \\ \mathbb{M}_{\Delta_{12}(\mathrm{Op}(a)\circ B)}(s,\mathtt{b}) &\le_{s,\rho} (|\Delta_{12}a|_{1,s+\rho,0}\mathbb{M}_B(s_0,\mathtt{b}) + |\Delta_{12}a|_{1,s_0+\rho,0}\mathbb{M}_B(s,\mathtt{b}) \\ &\quad + |a|_{0,s_0+\rho,0}\mathbb{M}_{\Delta_{12}B}(s,\mathtt{b}) + |a|_{0,s+\rho,0}\mathbb{M}_{\Delta_{12}B}(s_0,\mathtt{b})\ )\ , \end{aligned} \tag{5.7.34}$$

for all $s_0 \le s \le S_{max}$. Moreover if $B \in \mathfrak{L}_{\rho+1}$ then

$$\partial_{\varphi_m}B, [\partial_x, B] \in \mathfrak{L}_\rho\,, \quad m = 1,\dots,\nu$$

and satisfy the bounds

$$\begin{aligned} &\mathbb{M}^\gamma_{\partial_{\varphi_m}B}(s,\mathtt{b}), \mathbb{M}^\gamma_{[\partial_x,B]}(s,\mathtt{b}) \le \mathbb{M}^\gamma_B(s,\mathtt{b}+1)\,, \quad \mathtt{b} \le \rho-2 \\ &\mathbb{M}_{\partial_{\varphi_m}\Delta_{12}B}(s,\mathtt{b}), \mathbb{M}_{[\partial_x,\Delta_{12}B]}(s,\mathtt{b}) \le \mathbb{M}_{\Delta_{12}B}(s,\mathtt{b}+1)\,, \quad \mathtt{b} \le \rho-3 \end{aligned} \tag{5.7.35}$$

for all $s_0 \le s \le S_{max}$. Note that in (5.7.35) the constants in the right hand side control the tameness constants of B as an element of $\mathfrak{L}_{\rho+1}$.

Proof. We start by studying the Lip-0-tame norm of

$$\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}_1}\mathrm{Op}(a)\circ\partial_\varphi^{\vec{\mathtt{b}}_2}B\langle D_x\rangle^{m_2} = \langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}_1}\mathrm{Op}(a)\langle D_x\rangle^{-m_1}\circ\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}_2}B\langle D_x\rangle^{m_2}.$$

By Lemma 2.3.8 and formula (2.2.13)

$$\mathfrak{M}^\gamma_{\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}_1}\mathrm{Op}(a)\langle D_x\rangle^{-m_1}}(0,s) \le |a|^{\gamma,\mathcal{O}}_{0,s+|\vec{\mathtt{b}}_1|+m_1,0} \le |a|^{\gamma,\mathcal{O}}_{0,s+\rho,0}$$

hence by Lemma 2.3.6 we have

$$\mathfrak{M}^\gamma_{\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}}(\mathrm{Op}(a)B)\langle D_x\rangle^{m_2}}(-\rho+|\vec{\mathtt{b}}|,s) \le |a|^{\gamma,\mathcal{O}}_{0,s+\rho,0}\mathbb{M}^\gamma_B(s_0,\mathtt{b}) + |a|^{\gamma,\mathcal{O}}_{0,s_0+\rho,0}\mathbb{M}^\gamma_B(s,\mathtt{b})\,.$$

Regarding

$$\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}}[\partial_x,\mathrm{Op}(a)B]\langle D_x\rangle^{m_2} = \langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}}([\partial_x,\mathrm{Op}(a)]B)\langle D_x\rangle^{m_2} + \langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}}(\mathrm{Op}(a)[\partial_x,B])\langle D_x\rangle^{m_2}$$

we only need to consider the first summand as the second can be discussed exactly as above. Recalling that by definition $m_1 + m_2 = \rho - |\vec{\mathtt{b}}| - 1$ we write

$$\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}_1}[\partial_x,\mathrm{Op}(a)]\partial_\varphi^{\vec{\mathtt{b}}_2}B\langle D_x\rangle^{m_2} = \langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}_1}[\partial_x,\mathrm{Op}(a)]\langle D_x\rangle^{-m_1-1}\langle D_x\rangle^{m_1+1}\partial_\varphi^{\vec{\mathtt{b}}_2}B\langle D_x\rangle^{m_2}$$

and the result follows by recalling that

$$\mathfrak{M}^\gamma_{\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}_1}[\partial_x,\mathrm{Op}(a)]\langle D_x\rangle^{-m_1-1}}(0,s) \le |a|^{\gamma,\mathcal{O}}_{0,s+|\vec{\mathtt{b}}_1|+m_1,0} \le |a|^{\gamma,\mathcal{O}}_{0,s+\rho,0}\,.$$

□

The next Lemma shows that the finite rank operators of the form (5.6.5) are in $\mathfrak{L}_\rho$.

Lemma 5.7.8. *Fix $\rho \geq 3$. Let $\mathcal{R}$ be an operator of the form (5.6.5), where the functions $g_j(\tau), \chi_j(\tau)$ belong to H^s for $\tau \in [0,1]$ and depend in a Lipschitz way on the parameter $\omega \in \mathcal{O} \subset \mathbb{R}^\nu$. Then there is $\sigma_1 = \sigma_1(\rho) > 0$ such that $\mathcal{R}$ belongs to $\mathfrak{L}_\rho$ and*

$$\mathbb{M}^\gamma_{\mathcal{R}}(s, \mathtt{b}) \leq_s \sum_{|j|\leq C} \sup_{\tau\in[0,1]} (\|\chi_j(\tau)\|^{\gamma,\mathcal{O}}_{s+\sigma_1}\|g_j\|^{\gamma,\mathcal{O}}_{s_0+\sigma_1} + \|\chi_j(\tau)\|^{\gamma,\mathcal{O}}_{s_0+\sigma_1}\|g_j(\tau)\|^{\gamma,\mathcal{O}}_{s+\sigma_1}), \tag{5.7.36}$$

$$\begin{aligned}\mathbb{M}_{\Delta_{12}\mathcal{R}}(s, \mathtt{b}) \leq_s \sum_{|j|\leq C} \sup_{\tau\in[0,1]} \Big(&\|\Delta_{12}\chi_j(\tau)\|_{s+\sigma_1}\|g_j\|_{s_0+\sigma_1} + \|\Delta_{12}\chi_j(\tau)\|_{s_0+\sigma_1}\|g_j(\tau)\|_{s+\sigma_1} \\ &+ \|\chi_j(\tau)\|_{s+\sigma_1}\|\Delta_{12}g_j\|_{s_0+\sigma_1} + \|\Delta_{12}\chi_j(\tau)\|_{s_0+\sigma_1}\|\Delta_{12}g_j(\tau)\|_{s+\sigma_1}\Big).\end{aligned} \tag{5.7.37}$$

Proof. The Lemma follows by using the reasoning in proof of Lemma 5.7.6 and using the explicit formula (5.6.5). □

The next Lemma gives a canonical way to write the composition of two pseudo differential operators as a pseudo differential operator plus a remainder in $\mathfrak{L}_\rho$. Of course Lemma 2.2.5 says that such a composition is itself a pseudo differential operator, so in principle one could take the remainder to be zero. The purpose of this Lemma is to get better bounds with respect to (2.2.13), the price to pay is that we do not control the symbol of the composition but only an approximation up to a smoothing remainder of order $-\rho$.

Lemma 5.7.9. *Let $a = a(\omega) \in S^m$, $b = b(\omega) \in S^{m'}$ be defined on some subset $\mathcal{O} \subset \mathbb{R}^\nu$ with $m, m' \in \mathbb{R}$ and consider any $\rho \geq \max\{-(m + m' + 1), 1\}$. Assume also that a and b depend in a Lipschitz way on the parameter i. There exist an operator $R_\rho \in \mathfrak{L}_\rho$ such that (recall Definition (2.2.3)) setting $N = m + m' + \rho \geq 1$*

$$\mathrm{Op}(a\#b) = \mathrm{Op}(c) + R_\rho, \qquad c := a\#_{<N}b \in S^{m+m'} \tag{5.7.38}$$

$$\begin{aligned}|c|^{\gamma,\mathcal{O}}_{m+m',s,\alpha} \leq_{s,\rho,\alpha,m,m'} &|a|^{\gamma,\mathcal{O}}_{m,s,N-1+\alpha}|b|^{\gamma,\mathcal{O}}_{m',s_0+N-1,\alpha} \\ &+ |a|^{\gamma,\mathcal{O}}_{m,s_0,N-1+\alpha}|b|^{\gamma,\mathcal{O}}_{m',s+N-1,\alpha},\end{aligned} \tag{5.7.39}$$

where

$$\begin{aligned}\mathbb{M}^\gamma_{R_\rho}(s, \mathtt{b}) \leq_{s,\rho,m,m'} &|a|^{\gamma,\mathcal{O}}_{m,s+\rho,N}|b|^{\gamma,\mathcal{O}}_{m',s_0+2N+|m|,0} \\ &+ |a|^{\gamma,\mathcal{O}}_{m,s_0,N}|b|^{\gamma,\mathcal{O}}_{m',s+\rho+2N+|m|,0}.\end{aligned} \tag{5.7.40}$$

for all $0 \leq \mathtt{b} \leq \rho - 2$ and $s_0 \leq s \leq S_{max}$. Moreover one has

$$\begin{aligned}|\Delta_{12}c|_{m+m',s,\alpha} \leq_{s,\alpha,\rho,m,m'} &|\Delta_{12}a|_{m,s,N-1+\alpha}|b|_{m',s_0+N-1,\alpha} \\ &+ |\Delta_{12}a|_{m,s_0,N-1+\alpha}|b|_{m',s+N-1,\alpha} \\ &+ |a|_{m,s,N-1+\alpha}|\Delta_{12}b|_{m',s_0+N-1,\alpha} \\ &+ |a|_{m,s_0,N-1+\alpha}|\Delta_{12}b|_{m',s+N-1,\alpha},\end{aligned} \tag{5.7.41}$$

$$\begin{aligned}\mathbb{M}_{\Delta_{12}R_\rho}(s, \mathtt{b}) \leq_{s,\rho,m,m'} &|\Delta_{12}a|_{m+1,s+\rho,N}|b|_{m',s_0+2N+|m|,0} \\ &+ |\Delta_{12}a|_{m+1,s_0,N}|b|_{m',s+\rho+2N+|m|,0} \\ &+ |a|_{m,s+\rho,N}|\Delta_{12}b|_{m'+1,s_0+2N+|m|,0} \\ &+ |a|_{m,s_0,N}|\Delta_{12}b|_{m'+1,s+\rho+2N+|m|,0}.\end{aligned} \tag{5.7.42}$$

for all $0 \leq \mathtt{b} \leq \rho - 3$ and $s_0 \leq s \leq S_{max}$.

Proof. To shorten the notation we write $\|\cdot\|_s := \|\cdot\|_s^{\gamma,\mathcal{O}}$. For $\beta \in \mathbb{R}$ we have

$$\partial_\xi^\beta c = \sum_{k=0}^{N-1} \frac{1}{k! \mathrm{i}^k} \sum_{\beta_1+\beta_2=\beta} C_{\beta_1\beta_2} \partial_\xi^{\beta_1+k} a \, \partial_\xi^{\beta_2} \partial_x^k b, \tag{5.7.43}$$

then, by the tameness of the product, we get

$$\|\partial_\xi^\beta c\|_s \leq_s \sum_{k=0}^{N-1} \frac{1}{k!} \sum_{\beta_1+\beta_2=\beta} C_{\beta_1\beta_2} \big(\|\partial_\xi^{\beta_1+k} a\|_s \, \|\partial_\xi^{\beta_2} \partial_x^k b\|_{s_0} + \|\partial_\xi^{\beta_1+k} a\|_{s_0} \, \|\partial_\xi^{\beta_2} \partial_x^k b\|_s\big). \tag{5.7.44}$$

Thus

$$\begin{aligned}
|c|_{m+m',s,\alpha} \leq_s & \sum_{k=0}^{N-1} \frac{1}{k!} \max_{0\leq\beta\leq\alpha} \sup_{\xi\in\mathbb{R}} \sum_{\beta_1+\beta_2=\beta} C_{\beta_1\beta_2} \big(\|\partial_\xi^{\beta_1+k} a\|_s \, \|\partial_\xi^{\beta_2} \partial_x^k b\|_{s_0} \\
& + \|\partial_\xi^{\beta_1+k} a\|_{s_0} \, \|\partial_\xi^{\beta_2} \partial_x^k b\|_s\big) \langle\xi\rangle^{-(m+m')+\beta} \\
\leq_s & \sum_{k=0}^{N-1} \frac{1}{k!} \max_{0\leq\beta\leq\alpha} \sum_{\beta_1+\beta_2=\beta} C_{\beta_1\beta_2} \sup_{\xi\in\mathbb{R}} \big(\|\partial_\xi^{\beta_1+k} a\|_s \langle\xi\rangle^{-m+\beta_1} \, \|\partial_\xi^{\beta_2} \partial_x^k b\|_{s_0} \langle\xi\rangle^{-m'+\beta_2} \\
& + \|\partial_\xi^{\beta_1+k} a\|_{s_0} \langle\xi\rangle^{-m+\beta_1} \, \|\partial_\xi^{\beta_2} \partial_x^k b\|_s \langle\xi\rangle^{-m'+\beta_2}\big) \\
\leq_{s,\alpha} & \sum_{k=0}^{N-1} \frac{1}{k!} \sup_{\xi\in\mathbb{R}} \big(\max_{0\leq\beta_1\leq\alpha} \|\partial_\xi^{\beta_1+k} a\|_s \langle\xi\rangle^{-m+\beta_1} \max_{0\leq\beta_2\leq\alpha} \|\partial_\xi^{\beta_2} \partial_x^k b\|_{s_0} \langle\xi\rangle^{-m'+\beta_2} \\
& + \max_{0\leq\beta_1\leq\alpha} \|\partial_\xi^{\beta_1+k} a\|_{s_0} \langle\xi\rangle^{-m+\beta_1} \max_{0\leq\beta_2\leq\alpha} \|\partial_\xi^{\beta_2} \partial_x^k b\|_s \langle\xi\rangle^{-m'+\beta_2}\big).
\end{aligned}$$

Then by definition (2.2.10) we get (5.7.39). In the same way we obtain the bound (5.7.41) by using the following fact

$$\Delta_{12}(\partial_\xi^k a \, \partial_x^k b) = \partial_\xi^k(\Delta_{12} a) \, \partial_x^k b + \partial_\xi^k a \, \partial_x^k(\Delta_{12} b).$$

We remark that R_ρ is the pseudo differential operator R_N considered in Lemma 2.2.5 (recall $N = m + m' + \rho$). By Lemma 5.7.6

$$\mathbb{M}_{R_\rho}^\gamma \overset{(5.7.31)}{\leq_{s,\rho,m,m'}} |R_\rho|_{-\rho,s+\rho,0}$$

then by formula (2.2.14) of Lemma 2.2.5 we get the bounds (5.7.40).

The bounds (5.7.42), follow in the same way.

□

Remark 5.7.10. Note that if $m + m' \leq -\rho$ then by Lemma 5.7.6 $\mathrm{Op}(a) \circ \mathrm{Op}(b) \in \mathfrak{L}_\rho$.

Corollary 5.7.11. *Fix $\rho \geq 3$. Let $a \in S^{-1}$, $\rho > n \in \mathbb{N}$ depending in a Lipschitz way on a parameter i. Then there exist a symbol $c^{(n)} \in S^{-n}$ and a operator $R_\rho^{(n)} \in \mathfrak{L}_\rho$ such that*

$$\mathrm{Op}(a)^n = \mathrm{Op}(c^{(n)}) + R_\rho^{(n)}. \tag{5.7.45}$$

Moreover the following bounds hold

$$|c^{(n)}|_{-n,s,\alpha}^{\gamma,\mathcal{O}} \leq_{n,s,\alpha,\rho} |a|_{-1,s+(n-1)(\rho-3),\alpha+\rho-3}^{\gamma,\mathcal{O}} \big(|a|_{-1,s_0+(n-1)(\rho-3),\alpha+\rho-3}^{\gamma,\mathcal{O}}\big)^{n-1}, \tag{5.7.46}$$

$$|\Delta_{12}c^{(n)}|_{-n,s,\alpha,\rho} \leq |\Delta_{12}a|_{-1,s+(n-1)(\rho-3),\alpha+\rho-3}|a|^{n-1}_{-1,s_0+(n-1)(\rho-3),\alpha+\rho-3} + |a|_{-1,s+(n-1)(\rho-3),\alpha+\rho-3}|\Delta_{12}a|_{-1,s_0+(n-1)(\rho-3),\alpha+\rho-3}|a|^{n-2}_{-1,s_0+(n-1)(\rho-3),\alpha+\rho-3} \tag{5.7.47}$$

and where $R^{(n)}_\rho$ *belongs to* $\mathfrak{L}_\rho$ *with*

$$\mathbb{M}^\gamma_{R^{(n)}_\rho}(s,\mathtt{b}) \leq_{s,\rho,\mathtt{b},n} |a|^{\gamma,\mathcal{O}}_{-1,s+n(\rho-3)+\rho,\rho-2}\big(|a|^{\gamma,\mathcal{O}}_{-1,s_0+n(\rho-3)+\rho,\rho-2}\big)^{n-1} \tag{5.7.48}$$

and

$$\mathbb{M}_{\Delta_{12}R^{(n)}_\rho}(s) \leq_{s,n,\mathtt{b}} |\Delta_{12}a|_{0,s+n(\rho-3)+\rho,\rho-2}\big(|a|_{-1,s_0+n(\rho-3)+\rho,\rho-2}\big)^{n-1} + |\Delta_{12}a|_{0,s_0+n(\rho-3)+\rho,\rho-2}|a|_{-1,s+n(\rho-3)+\rho,\rho-2} \big(|a|_{-1,s_0+n(\rho-3)+\rho,\rho-2}\big)^{n-2} \tag{5.7.49}$$

for all $s_0 \leq s \leq S_{max}$.

Proof. We define for $n \geq 2$

$$c^{(n)} := a\#_{<\rho-2}c^{(n-1)}, \quad \text{where} \quad c^{(1)} := a \in S^{-1},$$

$$R^{(n)}_\rho := \sum_{k=0}^{n-2}[\mathrm{Op}(a)]^k\mathrm{Op}(a\#_{\geq\rho-2}c^{(n-k-1)})$$

By Lemma 5.7.9 we have

$$|c^{(2)}|^{\gamma,\mathcal{O}}_{-2,s,\alpha} \leq_{s,\alpha} |a|^{\gamma,\mathcal{O}}_{-1,s+\rho-3,\alpha+\rho-3}|a|^{\gamma,\mathcal{O}}_{-1,s_0+\rho-3,\alpha+\rho-3}$$

and so (5.7.46) is satisfied for $n=2$. Now given (5.7.46) for n we prove it for $n+1$. For simplicity we write $\leq_{n,s,\alpha}=\leq$. We have

$$\begin{aligned}|a\#_{<\rho-2}c^{(n)}|^{\gamma,\mathcal{O}}_{-n-1,s,\alpha} &\leq |a|^{\gamma,\mathcal{O}}_{-1,s,\alpha+\rho-3}|c^{(n)}|^{\gamma,\mathcal{O}}_{-n,s_0+\rho-3,\alpha} + |a|^{\gamma,\mathcal{O}}_{-1,s_0,\alpha+\rho-3}|a^{(n)}|^{\gamma,\mathcal{O}}_{-n,s+\rho-3,\alpha}\\ &\leq |a|^{\gamma,\mathcal{O}}_{-1,s,\alpha+\rho-3}|a|^{\gamma,\mathcal{O}}_{-1,s_0+n(\rho-3),\alpha+\rho-3}\big(|a|^{\gamma,\mathcal{O}}_{-1,s_0+(n-1)(\rho-3),\alpha+\rho-3}\big)^{n-1}\\ &+ |a|^{\gamma,\mathcal{O}}_{-1,s+n(\rho-3),\alpha+\rho-3}|a|^{\gamma,\mathcal{O}}_{-1,s_0,\alpha+\rho-3}\big(|a|^{\gamma,\mathcal{O}}_{-1,s_0+(n-1)(\rho-3),\alpha+\rho-3}\big)^{n-1}\\ &\leq |a|^{\gamma,\mathcal{O}}_{-1,s+n(\rho-3),\alpha+\rho-3}\big(|a|^{\gamma,\mathcal{O}}_{-n,s_0+n(\rho-3),\alpha+\rho-3}\big)^{n},\end{aligned}$$

hence (5.7.46) is proved. Arguing as above one can prove (5.7.47).

Now fix $2 \leq k \in \mathbb{N}$ and denote by

$$r_k := a\#_{\geq\rho-2}c^{(k-1)}.$$

We apply repeatedly formula (5.7.34) in oder to get

$$\begin{aligned}&\mathbb{M}^\gamma_{(\mathrm{Op}(a)^k\mathrm{Op}(r_{n-k}))}(s,\mathtt{b})\\ &\leq_{s,\rho,\mathtt{b}} (|a|^{\gamma,\mathcal{O}}_{-1,s_0+\rho,0})^{k-1}\left(|a|^{\gamma,\mathcal{O}}_{-1,s+\rho,0}\mathbb{M}^\gamma_{\mathrm{Op}(r_{n-k})}(s_0,\mathtt{b}) + |a|^{\gamma,\mathcal{O}}_{-1,s_0+\rho,0}\mathbb{M}^\gamma_{\mathrm{Op}(r_{n-k})}(s,\mathtt{b})\right).\end{aligned}$$

Now by Lemma 5.7.6 we have that for all $k \geq 2$

$$\mathbb{M}^\gamma_{\mathrm{Op}(r_k)}(s,\mathtt{b}) \leq_{s,\rho,\mathtt{b}} |r_k|^{\gamma,\mathcal{O}}_{-\rho,s+\rho,0}$$

Now by (2.2.14) with $m=-1, m'=-k+1, N=\rho-2$ we have

$$
\begin{aligned}
|r_k|^{\gamma,\mathcal{O}}_{-\rho,s,0} &\le |r_k|^{\gamma,\mathcal{O}}_{-\rho-k+2,s,0} \\
&\le |a|^{\gamma,\mathcal{O}}_{-1,s,\rho-2}|c^{(k-1)}|^{\gamma,\mathcal{O}}_{-k+1,s_0+2(\rho-2)+1,0} + |a|^{\gamma,\mathcal{O}}_{-1,s_0,\rho-2}|c^{(k-1)}|^{\gamma,\mathcal{O}}_{-k+1,s+2(\rho-2)+1,0} \\
&\overset{(5.7.46)}{\le} |a|^{\gamma,\mathcal{O}}_{-1,s+k(\rho-3),\rho-2}\big(|a|^{\gamma,\mathcal{O}}_{-1,s_0+k(\rho-3),\rho-2}\big)^{k-1}
\end{aligned}
$$

Then

$$
\mathbb{M}^{\gamma}_{R^{(n)}_\rho}(s,\mathtt{b}) \le_{s,\rho,\mathtt{b}} |a|^{\gamma,\mathcal{O}}_{-1,s+n(\rho-3)+\rho,\rho-2}\big(|a|^{\gamma,\mathcal{O}}_{-1,s_0+n(\rho-3)+\rho,\rho-2}\big)^{n-1}.
$$

We follow the same strategy in order to study the operator

$$
\Delta_{12}\big(\mathrm{Op}(a)^k R_{\rho^{(n-k)}}\big) = k\mathrm{Op}(a)^{k-1}\mathrm{Op}(\Delta_{12}a)R_{\rho^{(n-k)}} + \mathrm{Op}(a)^k\Delta_{12}R_{\rho^{(n-k)}}
$$

and we get (5.7.49). □

Remark 5.7.12. Note that if $n \ge \rho$ then $\mathrm{Op}(a)^n \in \mathfrak{L}_\rho$, by Lemma 5.7.6.

Corollary 5.7.13. *Let $a \in S^{-1}$ and consider $\mathrm{I}-(\mathrm{Op}(a)+T)$, where $T \in \mathfrak{L}_\rho$. There exist a constant $C(S_{max},\alpha,\rho)$ such that if*

$$
C(S_{max},\alpha,\rho)\Big(|a|^{\gamma,\mathcal{O}}_{-1,s_0+(\rho-1)(\rho-2)+3,\rho-2} + \mathbb{M}^{\gamma}_{T}(s_0,\mathtt{b})\Big) < 1 \tag{5.7.50}
$$

then $\mathrm{I}-(\mathrm{Op}(a)+T)$ is invertible and

$$
(\mathrm{I}-(\mathrm{Op}(a)+T))^{-1} = \mathrm{I} + \mathrm{Op}(c) + R_\rho \tag{5.7.51}
$$

where

$$
|c|^{\gamma,\mathcal{O}}_{-1,s,\alpha} \le_{s,\alpha,\rho} |a|^{\gamma,\mathcal{O}}_{-1,s+(\rho-2)(\rho-3),\alpha+\rho-3}, \tag{5.7.52}
$$

$$
\begin{aligned}
|\Delta_{12}c|_{-1,s,\alpha} \le\; & |\Delta_{12}a|_{-1,s+(\rho-2)(\rho-3),\alpha+\rho-3} \\
& + |a|_{-1,s+(\rho-1)(\rho-3),\alpha+\rho-3}|\Delta_{12}a|_{-1,s_0+(\rho-2)(\rho-3),\alpha+\rho-3}
\end{aligned} \tag{5.7.53}
$$

and $R_\rho \in \mathfrak{L}_\rho$ with

$$
\mathbb{M}^{\gamma}_{R_\rho}(s,\mathtt{b}) \le |a|^{\gamma,\mathcal{O}}_{-1,s+(\rho-1)(\rho-2)+3,\rho-2} + \mathbb{M}^{\gamma}_{T}(s,\mathtt{b}), \tag{5.7.54}
$$

$$
\mathbb{M}_{\Delta_{12}R_\rho}(s,\mathtt{b}) \le |\Delta_{12}a|_{-1,s+(\rho-1)(\rho-2)+3,\rho-2} + \mathbb{M}_{\Delta_{12}T}(s,\mathtt{b}), \tag{5.7.55}
$$

for all $s_0 \le s \le S_{max}$.

Proof. To shorten the notation we write $|\cdot|^{\gamma,\mathcal{O}}_{m,s,\alpha} = |\cdot|_{m,s,\alpha}$. We have by (5.7.50) and Neumann series

$$
\begin{aligned}
(\mathrm{I}-(\mathrm{Op}(a)+T))^{-1} &= \mathrm{I} + \sum_{n\ge1}(\mathrm{Op}(a)+T)^n = \mathrm{I} + \sum_{n=1}^{\rho-1}\Big(\mathrm{Op}(a)^n + \sum_{n=1}^{\infty}\tilde{R}^{(n)}_\rho\Big) + \sum_{n\ge\rho}\mathrm{Op}(a)^n \\
&\overset{\text{Lemma 5.7.11}}{=} \mathrm{I} + \sum_{n=1}^{\rho-1}\Big(\mathrm{Op}(c^{(n)}) + R^{(n)}_\rho + \tilde{R}^{(n)}_\rho\Big) + \sum_{n\ge\rho}\Big(\mathrm{Op}(a)^n + \tilde{R}^{(n)}_\rho\Big)
\end{aligned}
$$

where $\tilde{R}_\rho^{(n)} := (\mathrm{Op}(a)+T)^n - \mathrm{Op}(a)^n$ (and we are setting $R_\rho^{(1)} = 0$). We define $c := \sum_{n=1}^{\rho-1} c^{(n)}$ and by (5.7.46)

$$|c|_{-1,s,\alpha} \leq_{s,\alpha,\rho} \sum_{n=1}^{\rho-1} |a|_{-1,s+(n-1)(\rho-3),\alpha+\rho-3}\big(|a|_{-1,s_0+(n-1)(\rho-3),\alpha+\rho-3}\big)^{n-1}.$$

Formula (5.7.53) s obtained as above by using (5.7.47).

Using (5.7.34), the operator $\tilde{R}_\rho^{(n)} \in \mathfrak{L}_\rho$ with

$$\begin{aligned}\mathbb{M}^\gamma_{\tilde{R}_\rho^{(n)}}(s,\mathtt{b}) \leq_{s,\rho} (C_{s,\rho})^n & \sum_{\substack{n_1+n_2=n,\\ n_1\geq 1, n_2\geq 0}} \binom{n}{n_1} n_1 \Big(\big(|a|_{-1,s_0+\rho,0}\big)^{n_1-1} (\mathbb{M}^\gamma_T(s_0,\mathtt{b}))^{n_2} |a|_{-1,s+\rho,0} \\ & \big(|a|_{-1,s_0+\rho,0}\big)^{n_2} (\mathbb{M}^\gamma_T(s_0,\mathtt{b}))^{n_1-1} \mathbb{M}^\gamma_T(s,\mathtt{b})\Big)\end{aligned} \tag{5.7.56}$$

$$\begin{aligned}&\mathbb{M}_{\Delta_{12}\tilde{R}_\rho^{(n)}}(s,\mathtt{b}) \leq_{s,\rho} \\ &\leq_{s,\rho} (C_{s,\rho})^n \Big(\sum_{\substack{n_1+n_2=n,\\ n_1\geq 1, n_2\geq 0}} \binom{n}{n_1} n_1 |\Delta_{12} a|_{0,s+\rho,0} \mathbb{M}^{\gamma,\mathcal{O}}_T(s_0,\mathtt{b})^{n_2} \big(|a|_{-1,s+\rho,0}\big)^{n_1-1} \\ &+ \sum_{\substack{n_1+n_2=n,\\ n_1\geq 1, n_2\geq 1}} \binom{n}{n_1} n_1 n_2 |a|_{-1,s+\rho,0} \mathbb{M}_{\Delta_{12}T}(s_0,\mathtt{b}) \mathbb{M}^{\gamma,\mathcal{O}}_T(s_0,\mathtt{b})^{n_2-1} \big(|a|_{-1,s_0+\rho,0}\big)^{n_1-1} \\ &+ \sum_{\substack{n_1+n_2=n,\\ n_1\geq 1, n_2\geq 1}} \binom{n}{n_1} n_1 n_2 \mathbb{M}^{\gamma,\mathcal{O}}_T(s,\mathtt{b}) |\Delta_{12} a|_{0,s_0+\rho,0} \big(|a|_{-1,s_0+\rho,0}\big)^{n_1-1} \mathbb{M}^{\gamma,\mathcal{O}}_T(s_0,\mathtt{b})^{n_2-1} \\ &+ \sum_{\substack{n_1+n_2=n,\\ n_1\geq 0, n_2\geq 1}} \binom{n}{n_1} n_2 \big(|a|_{-1,s_0+\rho,0}\big)^{n_1} \mathbb{M}_{\Delta_{12}T}(s,\mathtt{b}) \mathbb{M}^{\gamma,\mathcal{O}}_T(s_0,\mathtt{b})^{n_2-1}\Big)\end{aligned} \tag{5.7.57}$$

We define $Q_\rho := \sum_{n=1}^{\rho-1}(R_\rho^{(n)} + \tilde{R}_\rho^{(n)})$ and then by Lemma 5.7.11 we have (note that if $n \leq \rho$ then all the constants $C^n_{s,\rho}$ can be absorbed in the $\leq_{s,\rho}$ symbol)

$$\begin{aligned}\mathbb{M}_{Q_\rho}(s,\mathtt{b}) \leq_{s,\rho} & \sum_{n=2}^{\rho-1} |a|^{\gamma,\mathcal{O}}_{-1,s+n(\rho-3)+\rho,\rho-2} \big(|a|^{\gamma,\mathcal{O}}_{-1,s_0+n(\rho-3)+\rho,\rho-2}\big)^{n-1} \\ & + \Big(\sum_{n=1}^{\rho-1} \sum_{\substack{n_1+n_2=n,\\ n_1\geq 1, n_2\geq 0}} |a|^{n_1-1}_{-1,s_0+\rho,0} \mathbb{M}^\gamma_T(s_0,\mathtt{b})^{n_2}\Big) |a|_{-1,s+\rho,0} \\ & + \Big(\sum_{n=1}^{\rho-1} \sum_{\substack{n_1+n_2=n,\\ n_2\geq 1, n_1\geq 0}} |a|^{n_1}_{-1,s_0+\rho,0} \mathbb{M}^\gamma_T(s_0,\mathtt{b})^{n_2-1}\Big) \mathbb{M}^\gamma_T(s,\mathtt{b}) \\ & \leq_{s,\rho} |a|_{-1,s+(\rho-1)(\rho-2)+3,\rho-2} + \mathbb{M}^\gamma_T(s,\mathtt{b})\end{aligned} \tag{5.7.58}$$

by our smallness condition.

It remains to bound

$$\sum_{n\geq\rho} \tilde{R}_\rho^{(n)} + \sum_{n\geq\rho} \mathrm{Op}(a)^n.$$

The first summand is controlled by

$$\begin{aligned}
&\sum_{n=\rho}^{\infty}(C_{s,\rho})^n \sum_{\substack{n_1+n_2=n,\\ n_1\geq 1, n_2\geq 0}} \binom{n}{n_1} n_1\Big(\big(|a|_{-1,s_0+\rho,0}\big)^{n_1-1}(\mathbb{M}^{\gamma}_{T}(s_0,\mathtt{b}))^{n_2}|a|_{-1,s+\rho,0}\\
&\big(|a|_{-1,s_0+\rho,0}\big)^{n_2}(\mathbb{M}^{\gamma}_{T}(s_0,\mathtt{b}))^{n_1-1}\mathbb{M}^{\gamma}_{T}(s,\mathtt{b})\Big)\\
&\leq_{s,\rho} |a|_{-1,s+\rho,0}+\mathbb{M}^{\gamma}_{T}(s,\mathtt{b})\,.
\end{aligned}$$

The second summand is controlled by Lemma 5.7.6 and by formula (5.7.34) by

$$|\mathrm{Op}(a)^{\rho}|_{-\rho,s+\rho,0}|a|^{n-\rho}_{-1,s_0+\rho,0}+|\mathrm{Op}(a)^{\rho}|_{-\rho,s_0+\rho,0}|a|^{n-\rho-1}_{-1,s_0+\rho,0}|a|_{-1,s+\rho,0}\,.$$

The bound follows by using repeatedly the bound (2.2.13).
In order to bound the i variation we note

$$\Delta_{12}(1-(\mathrm{Op}(a)+T))^{-1}=-(1-(\mathrm{Op}(a)+T))^{-1}(\mathrm{Op}(\Delta_{12}a)+\Delta_{12}T)(1-(\mathrm{Op}(a)+T))^{-1}\,,$$

and proceed as above. □

We now wish to study conjugation of $\mathfrak{L}_\rho$ with the torus diffeomorphism

$$\begin{aligned}
\mathcal{A}^{\tau}h(\varphi,x)&:=(1+\tau\beta_x(\varphi,x))h(\varphi,x+\tau\beta(\varphi,x)), && \varphi\in\mathbb{T}^{\nu},\ x\in\mathbb{T},\\
(\mathcal{A}^{\tau})^{-1}h(\varphi,y)&:=(1+\tilde{\beta}_y(\tau,\varphi,y))\,h(\varphi,y+\tilde{\beta}(\tau,\varphi,y)), && \varphi\in\mathbb{T}^{\nu},\ y\in\mathbb{T}.
\end{aligned}\tag{5.7.59}$$

We refer to the Appendix A.1 for some properties of the operator $\mathcal{A}^{\tau}$ in (5.7.59). We recall that $\mathcal{A}^{\tau}$ satisfies

$$\begin{cases}\partial_\tau\mathcal{A}^{\tau}=\mathtt{X}\mathcal{A}^{\tau},\\ \mathcal{A}^0=\mathrm{I}.\end{cases},\qquad \mathtt{X}:=\partial_x\circ b,\quad b:=\frac{\beta}{1+\tau\beta_x}.\tag{5.7.60}$$

We have the following Lemma.

Lemma 5.7.14. *Fix $\rho\in\mathbb{N}$ with $\rho\geq 3$, consider $\mathcal{O}\subset\mathbb{R}^{\nu}$ and let $R\in\mathfrak{L}_\rho(\mathcal{O})$ (see Def. 5.7.4). Consider a function β 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 i. Let $\mathcal{A}^{\tau}$ be the operator defined in (5.7.59). There exists $\mu=\mu(\rho)\gg 1$ and $\delta>0$ small such that if $\|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mu}\leq\delta$, then the operator $M^{\tau}:=\mathcal{A}^{\tau}R(\mathcal{A}^{\tau})^{-1}$ belongs to the class $\mathfrak{L}_\rho$. In particular one has, for $s_0\leq s\leq S_{max}$,*

$$\mathbb{M}^{\gamma}_{M^{\tau}}(s,\mathtt{b})\leq\mathbb{M}^{\gamma}_{R}(s,\mathtt{b})+\|\beta\|^{\gamma,\mathcal{O}}_{s+\mu}\mathbb{M}^{\gamma}_{R}(s_0,\mathtt{b}),\tag{5.7.61}$$

for $\mathtt{b}\leq\rho-2$ and

$$\begin{aligned}
\mathbb{M}_{\Delta_{12}M^{\tau}}(s,\mathtt{b})&\leq\mathbb{M}_{\Delta_{12}R^{\tau}}(s,\mathtt{b})+\|\beta\|_{s+\mu}\mathbb{M}_{\Delta_{12}R^{\tau}}(s_0,\mathtt{b})\\
&\quad+\|\Delta_{12}\beta\|_{s+\mu}\mathbb{M}^{\gamma}_{R^{\tau}}(s_0,\mathtt{b})+\|\Delta_{12}\beta\|_{s_0+\mu}\mathbb{M}^{\gamma}_{R^{\tau}}(s,\mathtt{b})
\end{aligned}\tag{5.7.62}$$

for $\mathtt{b}\leq\rho-3$.

Proof. We start by showing that M satisfies item (i) of Definition 5.7.4. Let $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho$. We write

$$\langle D_x \rangle^{m_1} M \langle D_x \rangle^{m_2} = \langle D_x \rangle^{m_1} \mathcal{A}^\tau \langle D_x \rangle^{-m_1} \langle D_x \rangle^{m_1} R \langle D_x \rangle^{m_2} \langle D_x \rangle^{-m_2} (\mathcal{A}^\tau)^{-1} \langle D_x \rangle^{m_2}.$$

Recall that by hypothesis the operator $\langle D_x \rangle^{m_1} R \langle D_x \rangle^{m_2}$ is Lip-0-tame with constants $\mathfrak{M}_R^\gamma(-\rho, s)$ see (5.7.18). Lemma (A.1.2) imply the estimates

$$\|\langle D_x \rangle^{m_1} \mathcal{A}^\tau(\varphi) \langle D_x \rangle^{-m_1} u\|_s^{\gamma,\mathcal{O}}, \|\langle D_x \rangle^{-m_2} (\mathcal{A}^\tau(\varphi))^{-1} \langle D_x \rangle^{m_2} u\|_s^{\gamma,\mathcal{O}} \leq_{s,\rho} \|u\|_s + \|\beta\|_{s+\mu}^{\gamma,\mathcal{O}} \|u\|_{s_0},$$

for $\tau \in [0,1]$, which imply that $\langle D_x \rangle^{m_1} M \langle D_x \rangle^{m_2}$ is Lip-0$-$tame with constant

$$\mathfrak{M}^\gamma_{\langle D_x \rangle^{m_1} M \langle D_x \rangle^{m_2}}(0, s) \leq_{s,\rho} \mathfrak{M}_R^\gamma(-\rho, s) + \|\beta\|_{s+\mu}^{\gamma,\mathcal{O}} \mathfrak{M}_R^\gamma(-\rho, s_0). \tag{5.7.63}$$

Hence M is Lip-$(-\rho)$-tame with constant $\mathfrak{M}_M^\gamma(-\rho, s) = \sup_{\substack{m_1+m_2=\rho \\ m_1,m_2 \geq 0}} \mathfrak{M}^\gamma_{\langle D_x \rangle^{m_1} M \langle D_x \rangle^{m_2}}(0, s)$. Fix $\mathtt{b} \leq \rho - 2$ and let $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho - \mathtt{b}$. We note that for any $\vec{\mathtt{b}} \in \mathbb{N}^\nu$ with $|\vec{\mathtt{b}}| = \mathtt{b}$

$$\partial_\varphi^{\vec{\mathtt{b}}} M = \sum_{\vec{\mathtt{b}_1} + \vec{\mathtt{b}_2} + \vec{\mathtt{b}_3} = \vec{\mathtt{b}}} C(|\vec{\mathtt{b}_1}|, |\vec{\mathtt{b}_2}|, |\vec{\mathtt{b}_3}|) (\partial_\varphi^{\vec{\mathtt{b}_1}} \mathcal{A}^\tau) \partial_\varphi^{\vec{\mathtt{b}_2}} R (\partial_\varphi^{\vec{\mathtt{b}_3}} (\mathcal{A}^\tau)^{-1}), \tag{5.7.64}$$

for some constants $C(|\vec{\mathtt{b}_1}|, |\vec{\mathtt{b}_2}|, |\vec{\mathtt{b}_3}|) > 0$, hence we need to show that each summand in (5.7.64) satisfies item (i) of Definition 5.7.4. We write

$$\begin{aligned} &\langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b}_1}} \mathcal{A}^\tau) \partial_\varphi^{\vec{\mathtt{b}_2}} R (\partial_\varphi^{\vec{\mathtt{b}_3}} (\mathcal{A}^\tau)^{-1}) \langle D_x \rangle^{m_2} = \\ &= \langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b}_1}} \mathcal{A}^\tau) \langle D_x \rangle^{y} \langle D_x \rangle^{-y} (\partial_\varphi^{\vec{\mathtt{b}_2}} R) \langle D_x \rangle^{z} \langle D_x \rangle^{-z} (\partial_\varphi^{\vec{\mathtt{b}_3}} (\mathcal{A}^\tau)^{-1}) \langle D_x \rangle^{m_2}, \end{aligned} \tag{5.7.65}$$

where $y = -|\mathtt{b}_1| - m_1$, $z = \rho - |\mathtt{b}_2| - |\mathtt{b}_1| - m_1$. Since $y + m_1 = -|\mathtt{b}_1|$ and $-z + m_2 = -|\mathtt{b}_3|$, hence by Lemma A.1.2 the operators

$$\langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b}_1}} \mathcal{A}^\tau) \langle D_x \rangle^{y}, \quad \langle D_x \rangle^{-z} (\partial_\varphi^{\vec{\mathtt{b}_3}} (\mathcal{A}^\tau)^{-1}) \langle D_x \rangle^{m_2},$$

satisfy bounds like (A.1.9). Moreover $-y + z = \rho - |\mathtt{b}_2|$ and $-y, z \geq 0$, hence, by the definition of the class $\mathfrak{L}_\rho$, we have that the operator $\langle D_x \rangle^{-y} (\partial_\varphi^{\vec{\mathtt{b}_2}} R) \langle D_x \rangle^{z}$ is Lip-0-tame. Following the reasoning used to prove (5.7.63) one obtains

$$\mathfrak{M}^\gamma_{\langle D_x \rangle^{m_1} \partial_{\varphi_m}^{\mathtt{b}} M \langle D_x \rangle^{m_2}}(0, s) \leq_{s,\rho} \mathfrak{M}_R^\gamma(-\rho + \mathtt{b}, s) + \|\beta\|_{s+\mu}^{\gamma,\mathcal{O}} \mathfrak{M}_R^\gamma(-\rho + \mathtt{b}, s). \tag{5.7.66}$$

Let us consider the operator $[M, \partial_x]$. We write

$$[M, \partial_x] = \mathcal{A}^\tau [R, \partial_x] (\mathcal{A}^\tau)^{-1} + \mathcal{A}^\tau R [(\mathcal{A}^\tau)^{-1}, \partial_x] - [\mathcal{A}^\tau, \partial_x] R (\mathcal{A}^\tau)^{-1}, \tag{5.7.67}$$

for $\tau \in [0,1]$. We need to show that each summand in (5.7.67) satisfies item (ii) in Definition (5.7.4). Let $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho - 1$. We first note that

$$\begin{aligned} &\langle D_x \rangle^{m_1} \mathcal{A}^\tau [R, \partial_x] (\mathcal{A}^\tau)^{-1} \langle D_x \rangle^{m_2} = \\ &= \langle D_x \rangle^{m_1} \mathcal{A}^\tau \langle D_x \rangle^{-m_1} \langle D_x \rangle^{m_1} [R, \partial_x] \langle D_x \rangle^{m_2} \langle D_x \rangle^{-m_2} (\mathcal{A}^\tau)^{-1} \langle D_x \rangle^{m_2}, \end{aligned} \tag{5.7.68}$$

hence, by applying Lemma A.1.2 to estimate the terms

$$\langle D_x \rangle^{-m_2} (\mathcal{A}^\tau)^{-1} \langle D_x \rangle^{m_2}, \quad \langle D_x \rangle^{m_1} (\mathcal{A}^\tau)^{-1} \langle D_x \rangle^{-m_1}$$

and using the tameness of the operator $\langle D_x\rangle^{m_1}[R,\partial_x]\langle D_x\rangle^{m_2}$ one gets

$$\mathfrak{M}^{\gamma}_{\langle D_x\rangle^{m_1}\mathcal{A}^{\tau}[R,\partial_x](\mathcal{A}^{\tau})^{-1}\langle D_x\rangle^{m_2}}(0,s) \leq_{s,\rho} \mathbb{M}^{\gamma}_{R}(s,\mathtt{b}) + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mu}\mathbb{M}^{\gamma}_{R}(s_0,\mathtt{b}). \tag{5.7.69}$$

The term $[\mathcal{A}^{\tau},\partial_x]R(\mathcal{A}^{\tau})^{-1}$ in (5.7.67) is more delicate. Let $m_1,m_2\in\mathbb{R}$, $m_1,m_2\geq 0$ and $m_1+m_2=\rho-1$. We write

$$\langle D_x\rangle^{m_1}[\mathcal{A},\partial_x]\langle D_x\rangle^{-m_1-1}\langle D_x\rangle^{m_1+1}R\langle D_x\rangle^{m_2}\langle D_x\rangle^{-m_2}(\mathcal{A}^{\tau})^{-1}\langle D_x\rangle^{m_2}. \tag{5.7.70}$$

By Lemma A.1.2 we have that $\langle D_x\rangle^{-m_2}(\mathcal{A}^{\tau})^{-1}\langle D_x\rangle^{m_2}$ satisfies a bound like (A.1.9) with $\alpha=0$. The operator $\langle D_x\rangle^{m_1+1}R\langle D_x\rangle^{m_2}\langle D_x\rangle^{m_1+1}R\langle D_x\rangle^{m_2}$ is Lip-0-tame since $R\in\mathfrak{L}_{\rho}$ and $m_1+1+m_2=\rho$. Moreover by an explicit computation (using formula (5.7.59)) we get

$$[\mathcal{A}^{\tau},\partial_x] = \tau\frac{\beta_{xx}}{1+\tau\beta_x}\mathcal{A}^{\tau} + \tau\beta_x\mathcal{A}^{\tau}\partial_x. \tag{5.7.71}$$

We claim that, for $s\geq s_0$ and $u\in H^s$, one has

$$\|\langle D_x\rangle^{m_1}[\mathcal{A}^{\tau},\partial_x]\langle D_x\rangle^{-m_1-1}u\|^{\gamma,\mathcal{O}}_{s} \leq_{s,\rho} \|u\|_s\|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mu} + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mu}\|u\|_{s_0}, \tag{5.7.72}$$

for some $\mu >$ depending only on s,ρ. The first summand in (5.7.71) satisfies the bound (5.7.72) thanks to Lemma 2.2.5 to estimate the composition $\langle D_x\rangle^{m_1}\beta_{xx}(1+\tau\beta_x)^{-1}\langle D_x\rangle^{-m_1}$ and thanks Lemma A.1.2 to estimate $\langle D_x\rangle^{m_1}\mathcal{A}^{\tau}\langle D_x\rangle^{-m_1}$. For the second summand we reason as follow: we write

$$\langle D_x\rangle^{m_1}\tau\beta_x\mathcal{A}^{\tau}\partial_x\langle D_x\rangle^{-m_1-1} = \Big(\langle D_x\rangle^{m_1}\beta_x\langle D_x\rangle^{-m_1}\Big)\Big(\langle D_x\rangle^{m_1}\mathcal{A}^{\tau}\langle D_x\rangle^{-m_1}\Big)\partial_x\langle D_x\rangle^{-1}$$

and we note that the operator $\partial_x\langle D_x\rangle^{-1}$ is bounded on H^s. Hence the bound (5.7.72) follows by applying Lemmata 2.2.5 and A.1.2. By the discussion above one gets

$$\mathfrak{M}^{\gamma}_{\langle D_x\rangle^{m_1}[\mathcal{A}^{\tau},\partial_x]R(\mathcal{A}^{\tau})^{-1}\langle D_x\rangle^{m_2}}(0,s) \leq_{s,\rho} \mathfrak{M}^{\gamma}_{R}(-\rho+1,s) + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mu}\mathfrak{M}^{\gamma}_{R}(-\rho+1,s). \tag{5.7.73}$$

One can study the tameness constant of the operator $\mathcal{A}^{\tau}R[(\mathcal{A}^{\tau})^{-1},\partial_x]$ in (5.7.67) by using the same arguments above.

We check now that M satisfies item (iii) of Def. 5.7.4. Let $m_1,m_2\in\mathbb{R}$, $m_1,m_2\geq 0$ and $m_1+m_2=\rho-\mathtt{b}-1$. We write for $\vec{\mathtt{b}}\in\mathbb{N}^{\nu}$, $|\vec{\mathtt{b}}|=\mathtt{b}$

$$[\partial_{\varphi}^{\vec{\mathtt{b}}}\mathcal{A}^{\tau}R(\mathcal{A}^{\tau})^{-1},\partial_x] = \sum_{\vec{\mathtt{b}}_1+\vec{\mathtt{b}}_2+\vec{\mathtt{b}}_3=\vec{\mathtt{b}}} C(|\vec{\mathtt{b}}_1|,|\vec{\mathtt{b}}_2|,|\vec{\mathtt{b}}_3|)\Big[(\partial_{\varphi}^{\vec{\mathtt{b}}_1}\mathcal{A}^{\tau})(\partial_{\varphi}^{\vec{\mathtt{b}}_2}R)(\partial_{\varphi}^{\vec{\mathtt{b}}_3}(\mathcal{A}^{\tau})^{-1}),\partial_x\Big] \tag{5.7.74}$$

and we note that

$$\begin{aligned}[(\partial_{\varphi}^{\vec{\mathtt{b}}_1}\mathcal{A}^{\tau})(\partial_{\varphi}^{\vec{\mathtt{b}}_2}R)(\partial_{\varphi}^{\vec{\mathtt{b}}_1}(\mathcal{A}^{\tau})^{-1}),\partial_x] &= (\partial_{\varphi}^{\vec{\mathtt{b}}_1}\mathcal{A}^{\tau})\Big[(\partial_{\varphi}^{\vec{\mathtt{b}}_2}R),\partial_x\Big](\partial_{\varphi}^{\vec{\mathtt{b}}_3}\mathcal{A}^{\tau})^{-1}) \\ &\quad + (\partial_{\varphi}^{\vec{\mathtt{b}}_1}\mathcal{A}^{\tau})(\partial_{\varphi}^{\vec{\mathtt{b}}_2}R)\Big[(\partial_{\varphi}^{\vec{\mathtt{b}}_3}(\mathcal{A}^{\tau})^{-1}),\partial_x\Big] \\ &\quad - \Big[(\partial_{\varphi}^{\vec{\mathtt{b}}_1}\mathcal{A}^{\tau}),\partial_x\Big](\partial_{\varphi}^{\vec{\mathtt{b}}_2}R)(\partial_{\varphi}^{\vec{\mathtt{b}}_3}(\mathcal{A}^{\tau})^{-1}).\end{aligned} \tag{5.7.75}$$

The most difficult term to study is the last summand in (5.7.75). We have that

$$\begin{aligned}&\langle D_x\rangle^{m_1}\Big[(\partial_\varphi^{\vec{\mathtt{b}}_1}\mathcal{A}^\tau),\partial_x\Big](\partial_\varphi^{\vec{\mathtt{b}}_2}R)(\partial_\varphi^{\vec{\mathtt{b}}_3}(\mathcal{A}^\tau)^{-1})\langle D_x\rangle^{m_2}=\\&=\langle D_x\rangle^{m_1}\Big[(\partial_\varphi^{\vec{\mathtt{b}}_1}\mathcal{A}^\tau),\partial_x\Big]\langle D_x\rangle^{-y}\langle D_x\rangle^{y}(\partial_\varphi^{\vec{\mathtt{b}}_2}R)\langle D_x\rangle^{z}\langle D_x\rangle^{-z}(\partial_\varphi^{\vec{\mathtt{b}}_3}(\mathcal{A}^\tau)^{-1})\langle D_x\rangle^{m_2},\end{aligned}\tag{5.7.76}$$

with $z=m_2+|\mathtt{b}_3|$ and $y=\rho-|\mathtt{b}_2|-|\mathtt{b}_3|-m_2$. Note the operator $\langle D_x\rangle^{-z}(\partial_\varphi^{\vec{\mathtt{b}}_3}(\mathcal{A}^\tau)^{-1})\langle D_x\rangle^{m_2}$ satisfies bound like (A.1.9) with α_0; moreover the operator $\langle D_x\rangle^{y}(\partial_\varphi^{\vec{\mathtt{b}}_2}R)\langle D_x\rangle^{z}$ is Lip-0-tame since $y+z=\rho-|\mathtt{b}_2|$. Note also that, since $m_1+m_2=\rho-|\mathtt{b}|-1$, one has $y=m_1+|\mathtt{b}_1|+1$. We now study the tameness constant of

$$\langle D_x\rangle^{m_1}\Big[(\partial_\varphi^{\vec{\mathtt{b}}_1}\mathcal{A}^\tau),\partial_x\Big]\langle D_x\rangle^{-m_1-|\mathtt{b}_1|-1}.$$

By differentiating the (5.7.71) we get

$$\partial_\varphi^{\vec{\mathtt{b}}_1}[\mathcal{A}^\tau,\partial_x]=\sum_{\vec{\mathtt{b}}_1'+\vec{\mathtt{b}}_1''=\vec{\mathtt{b}}_1}\tau(\partial_\varphi^{\vec{\mathtt{b}}_1'}g)(\partial_\varphi^{\vec{\mathtt{b}}_1''}\mathcal{A}^\tau)+\tau(\partial_\varphi^{\vec{\mathtt{b}}_1'}\beta_x)(\partial_\varphi^{\vec{\mathtt{b}}_1''}\mathcal{A}^\tau)\partial_x,\tag{5.7.77}$$

where $g=\beta_{xx}/(1+\tau\beta_x)$. We claim that

$$\|\langle D_x\rangle^{m_1}[\partial_\varphi^{\vec{\mathtt{b}}_1}\mathcal{A}^\tau,\partial_x]\langle D_x\rangle^{-m_1-|\mathtt{b}_1|-1}u\|_s^{\gamma,\mathcal{O}}\le_{s,\rho}\|u\|_s\|\beta\|_{s_0+\mu}^{\gamma,\mathcal{O}}+\|\beta\|_{s+\mu}^{\gamma,\mathcal{O}}\|u\|_{s_0},\tag{5.7.78}$$

for some $\mu>0$ depending on s,ρ. We study the most difficult summand in (5.7.77). We have

$$\begin{aligned}\langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}}_1'}\beta_x)(\partial_\varphi^{\vec{\mathtt{b}}_1''}\mathcal{A}^\tau)\partial_x\langle D_x\rangle^{-m_1-|\vec{\mathtt{b}}_1|-1}&=\langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}}_1'}\beta_x)\langle D_x\rangle^{-m_1-|\mathtt{b}_1|+|\mathtt{b}_1''|}\\&\times\langle D_x\rangle^{m_1+|\vec{\mathtt{b}}_1|-|\vec{\mathtt{b}}_1''|}(\partial_\varphi^{\vec{\mathtt{b}}_1''}\mathcal{A}^\tau)\langle D_x\rangle^{-m_1-\vec{\mathtt{b}}_1}\partial_x\langle D_x\rangle^{-1}.\end{aligned}\tag{5.7.79}$$

The (5.7.78) follows for the term in (5.7.79) by using Lemmata 2.2.5, A.1.2 and the fact that $\partial_x\langle D_x\rangle^{-1}$ is bounded on H^s. On the other summand in (5.7.77) one uses similar arguments. By the discussion above one can check that

$$\mathfrak{M}^\gamma_{\langle D_x\rangle^{m_1}[\partial_\varphi^{\vec{\mathtt{b}}}\mathcal{A}^\tau,\partial_x]R(\mathcal{A}^\tau)^{-1}\langle D_x\rangle^{m_2}}(0,s)\le_{s,\rho}\mathbb{M}_R^\gamma(s,\mathtt{b})+\|\beta\|_{s+\mu}^{\gamma,\mathcal{O}}\mathbb{M}_R^\gamma(s_0,\mathtt{b}).\tag{5.7.80}$$

The fact that the operator M satisfies items $(iii)-(iv)$ of Definition (5.7.4) can be proved arguing as done above for items $(i)-(ii)$. □

We conclude this section by showing the connection between the class $\mathfrak{L}_\rho$ and the class $\mathfrak{C}_{1,\mathtt{b}}$ in Definition 5.7.1.

Lemma 5.7.15. *Consider $\mathtt{b}\in\mathbb{N}$ and $\rho\in\mathbb{N}$ with $\rho\ge\mathtt{b}+3$. The following holds.*

(i) If $A\in\mathfrak{L}_\rho$ (see Def. 5.7.4) then $A\in\mathfrak{C}_{1,\mathtt{b}}$ (see Def. 5.7.1) with

$$\mathbb{B}_A^\gamma(s,\mathtt{b})\le_{\rho,s}\mathbb{M}_A^\gamma(s,\rho-2).\tag{5.7.81}$$

$$\mathbb{B}_{\Delta_{12}A}(s,\mathtt{b})\le_{\rho,s}\mathbb{M}_{\Delta_{12}A}(s,\rho-3).\tag{5.7.82}$$

(ii) *Consider a symbol* $a = a(\omega, i(\omega))$ *in* S^m *with* $m \le -1$ *depending on for* $\omega \in \mathcal{O}_0 \subset \mathbb{R}^\nu$ *in a Lipschitz way and on* i *in a Lipschitz way and let* $A := \mathrm{op}(a(x,\xi))$. *Then one has that* $A \in \mathfrak{C}_{1,\mathtt{b}}$ *with*

$$\mathbb{B}^\gamma_A(s,\mathtt{b}) \le_s |a|^{\gamma,\mathcal{O}_0}_{m,s+\mathtt{b}+2,\alpha}. \tag{5.7.83}$$

(iii) *Let* $A, B \in \mathfrak{C}_{1,\mathtt{b}}$. *Then* $A \circ B \in \mathfrak{C}_{1,\mathtt{b}}$ *with*

$$\mathbb{B}^\gamma_{A\circ B}(s,\mathtt{b}) \le_s \mathbb{B}^\gamma_A(s,\mathtt{b})\mathbb{B}^\gamma_A(s_0,\mathtt{b}) + \mathbb{B}^\gamma_A(s_0,\mathtt{b})\mathbb{B}^\gamma_A(s,\mathtt{b}) \tag{5.7.84}$$

$$\begin{aligned}\mathbb{B}_{\Delta_{12}(A\circ B)}(s,\mathtt{b}) \le_{s,\rho} & \,\mathbb{B}_{\Delta_{12}A}(s,\mathtt{b})\mathbb{B}_B(s_0,\mathtt{b}) + \mathbb{B}_{\Delta_{12}A}(s_0,\mathtt{b})\mathbb{B}_B(s,\mathtt{b}) \\ & + \mathbb{B}_{\Delta_{12}B}(s,\mathtt{b})\mathbb{B}_A(s_0,\mathtt{b}) + \mathbb{B}_{\Delta_{12}B}(s_0,\mathtt{b})\mathbb{B}_A(s,\mathtt{b}).\end{aligned} \tag{5.7.85}$$

Proof. Let us check item (i). The fact that $\langle D_x\rangle^{1/2}A\langle D_x\rangle^{1/2}$ is Lip-0-tame follows by (5.7.18) since $-\rho \le -1$. Indeed for any $h \in H^s$

$$\begin{aligned}\|\langle D_x\rangle^{\frac12}A\langle D_x\rangle^{\frac12}h\|^{\gamma,\mathcal{O}_0}_s &\le \|\langle D_x\rangle^{-\rho+1}\langle D_x\rangle^{\rho-\frac12}A\langle D_x\rangle^{\frac12}h\|^{\gamma,\mathcal{O}_0}_s \\ &\le_s \mathfrak{M}^\gamma_A(-\rho,s)\|h\|^{\gamma,\mathcal{O}_0}_{s_0} + \mathfrak{M}^\gamma_A(-\rho,s_0)\|h\|^{\gamma,\mathcal{O}_0}_s.\end{aligned}$$

By studying the tameness constant of the operators

$$\partial^{\vec{\mathtt{b}}}_\varphi A, [A,\partial_x], [\partial^{\vec{\mathtt{b}}}_\varphi A,\partial_x]\Delta_{12}A, \partial^{\vec{\mathtt{b}}}_\varphi\Delta_{12}A, [\Delta_{12}A,\partial_x], [\partial^{\vec{\mathtt{b}}}_\varphi\Delta_{12}A,\partial_x]$$

for $\vec{\mathtt{b}} \in \mathbb{N}^\nu$, $|\vec{\mathtt{b}}| = \mathtt{b}$, following the same reasoning one gets the (5.7.81) and (5.7.82).

In order to prove item (ii) one can follow almost word by word the proof of Lemma 5.7.6. Let us check (5.7.84). Let $\vec{\mathtt{c}} \in \mathbb{N}^\nu$ with $|\vec{\mathtt{c}}| = \mathtt{c}$, $0 \le \mathtt{c} \le \mathtt{b}$. One has

$$\langle D_x\rangle^{\frac12}\partial^{\vec{\mathtt{c}}}_\varphi(A\circ B)\langle D_x\rangle^{\frac12} = \sum_{\vec{\mathtt{b}}_1+\vec{\mathtt{b}}_2=\vec{\mathtt{c}}} C(|\vec{\mathtt{b}}_1|,\vec{\mathtt{b}}_2|)\langle D_x\rangle^{\frac12}(\partial^{\vec{\mathtt{b}}_1}_\varphi A)(\partial^{\vec{\mathtt{b}}_2}_\varphi B)\langle D_x\rangle^{\frac12}. \tag{5.7.86}$$

We show that each summand in (5.7.86) is a Lip-(0)-tame operator. We have for $h \in H^s$

$$\begin{aligned}&\|\langle D_x\rangle^{\frac12}(\partial^{\vec{\mathtt{b}}_1}_\varphi A)(\partial^{\vec{\mathtt{b}}_2}_\varphi B)\langle D_x\rangle^{\frac12}h\|^{\gamma,\mathcal{O}_0}_s \le \\ &\le \|\langle D_x\rangle^{\frac12}(\partial^{\vec{\mathtt{b}}_1}_\varphi A)\langle D_x\rangle^{\frac12}\langle D_x\rangle^{-1}\langle D_x\rangle^{\frac12}(\partial^{\vec{\mathtt{b}}_2}_\varphi B)\langle D_x\rangle^{\frac12}h\|^{\gamma,\mathcal{O}_0}_s \\ &\le_s (\mathbb{B}^\gamma_A(s,\mathtt{b})\mathbb{B}^\gamma_B(s_0,\mathtt{b}) + \mathbb{B}^\gamma_A(s_0,\mathtt{b})\mathbb{B}^\gamma_B(s,\mathtt{b}))\|h\|^{\gamma,\mathcal{O}_0}_{s_0} \\ &+ \mathbb{B}^\gamma_A(s_0,\mathtt{b})\mathbb{B}^\gamma_B(s_0,\mathtt{b})\|h\|^{\gamma,\mathcal{O}_0}_s\end{aligned}, \tag{5.7.87}$$

this bounds holds for any $|\vec{\mathtt{b}}_1|, |\vec{\mathtt{b}}_2| \le \mathtt{b}$. In (5.7.87) we used the fact that $\langle D_x\rangle^{\frac12}(\partial^{\mathtt{b}_1}_{\varphi_m}A)\langle D_x\rangle^{\frac12}$ and $\langle D_x\rangle^{\frac12}(\partial^{\mathtt{b}_2}_{\varphi_m}B)\langle D_x\rangle^{\frac12}$ are 0-tame by hypothesis (see Def. (5.7.1)). This proves (5.7.84) for the operators $A\circ B$ and $\partial^{\vec{\mathtt{c}}}_\varphi(A\circ B)$ for any $|\vec{\mathtt{c}}| \le \mathtt{c}$, $0 \le \mathtt{c} \le \mathtt{b}$. One concludes the proof of (5.7.84) and (5.7.85) followings the same ideas used above. For further details we refer to the proof of Lemma 5.7.5 which is very similar. ∎

5.7.2 Flow of hyperbolic Pseudo PDEs and Egorov Theorem

The goal of this section is to provide the tools we need to construct changes of coordinates that conjugate $\mathcal{L}_\omega$ in (5.6.31) to an operator with constant coefficient at order one (the highest order of differentiation).

We study the flow map Φ^τ of the vector field generated by the Hamiltonian

$$S(\tau,\varphi,z)=\int b(\tau,\varphi,x)z^2dx \qquad b(\tau,\varphi,x):=\frac{\beta(\varphi,x)}{1+\tau\beta_x(\varphi,x)}. \tag{5.7.88}$$

We first need to show that Φ^τ is well defined as map on $H^s_{S^\perp}$ (see Proposition 5.7.16 and Lemma 5.7.18). Then we study the structure of $\Phi^\tau\mathcal{L}_\omega(\Phi^\tau)^{-1}$. This is the core of our analysis and such study is performed in Proposition 5.7.21.

Let us start with two preliminary results. Instead of studying the flow generated by the Hamiltonian (5.7.88) we first consider the time one flow map of the pseudo differential PDE

$$\begin{cases}\partial_\tau\Psi^\tau(u)=(J\circ b)\Psi^\tau(u),\\ \Psi^0u=u,\end{cases} \tag{5.7.89}$$

where $b(\tau,\varphi,x)$ is defined in (5.7.88) with $\beta\in H^s(\mathbb{T}^{\nu+1})$ to be determined.
In the following proposition we prove that the flow of (5.7.89) is the composition of the diffeomorphism of the torus (5.7.59) with a pseudo differential operator of order -1 up to smoothing remainders belonging to the class $\mathfrak{L}_\rho$ for any $\rho\in\mathbb{N}$.

First we note that $J=\partial_x+3\Lambda\partial_x$ (recall (5.1.2) for the definition of Λ) and we define

$$\mathtt{Y}:=3\Lambda\partial_x\circ b,\quad b:=\frac{\beta}{1+\tau\beta_x}. \tag{5.7.90}$$

We note that $\mathcal{A}^\tau$ defined in (5.7.59) is such that (5.7.60) holds.

Proposition 5.7.16. *Fix $\rho\geq 3$, $S_{max}>s_0$ large enough, and consider a function $\beta:=\beta(\omega,i(\omega))\in C^\infty(\mathbb{T}^{\nu+1})$, Lipschitz in $\omega\in\mathcal{O}\subseteq\mathbb{R}^\nu$ and Lipschitz in the variable i. There exist $\sigma_1=\sigma_1(\rho)>0$ and $\delta=\delta(\rho,S_{max})>0$ such that if*

$$\|\beta\|^{\gamma,\mathcal{O}}_{s_0+\sigma_1}\leq\delta, \tag{5.7.91}$$

then, for any $\varphi\in\mathbb{T}^d$, the equation (5.7.89) has a unique solution in the space

$$C^0([0,1];H^s_x)\cap C^1([0,1];H^{s-1}_x),\quad \forall s_0\leq s\leq S_{max}.$$

One has $\Psi^\tau=\mathcal{A}^\tau\circ\mathcal{C}^\tau$ where $\mathcal{A}^\tau$ is defined in (5.7.59) and

$$\mathcal{C}^\tau=\Theta^\tau+R^\tau(\varphi),\qquad \Theta^\tau:=\mathrm{Op}(1+\vartheta(\tau,\varphi,x,\xi)) \tag{5.7.92}$$

with, for any $s\geq s_0$,

$$|\vartheta|^{\gamma,\mathcal{O}}_{-1,s,\alpha}\leq_{s,\alpha,\rho}\|\beta\|^{\gamma,\mathcal{O}}_{s+\sigma_1}, \tag{5.7.93}$$

$$|\Delta_{12}\vartheta|_{-1,s,\alpha}\leq_{s,\alpha,\rho}\|\Delta_{12}\beta\|_{s+\sigma_1}+\|\beta\|_{s+3\rho+2}\|\Delta_{12}\beta\|_{s_0+\sigma_1}. \tag{5.7.94}$$

and $R^\tau(\varphi)\in\mathfrak{L}_\rho$ (see Def. 5.7.4) with, for $s_0\leq s\leq S_{max}$,

$$\mathbb{M}^\gamma_{R^\tau}(s,\mathtt{b})\leq_{s,\alpha,\rho}\|\beta\|^{\gamma,\mathcal{O}}_{s+\sigma_1}, \tag{5.7.95}$$

for all $0 \leq \mathtt{b} \leq \rho - 2$ *and*

$$\mathbb{M}_{\Delta_{12} R^\tau}(s, \mathtt{b}) \leq_{s,\rho} \|\Delta_{12}\beta\|_{s+\sigma_1} + \|\beta\|_{s+\sigma_1}\|\Delta_{12}\beta\|_{s_0+\sigma_1} \tag{5.7.96}$$

for all $0 \leq \mathtt{b} \leq \rho - 3$.

Proof. Let us reformulate the problem (5.7.89) as

$$\Psi^\tau = \mathcal{A}^\tau \circ C^\tau, \qquad C^\tau := (\mathcal{A}^\tau)^{-1} \circ \Psi^\tau \tag{5.7.97}$$

where the operator C^τ satisfies the following system

$$\begin{cases} \partial_\tau C^\tau u = L^\tau C^\tau u, \\ C^0 u = u, \end{cases} \tag{5.7.98}$$

where L^τ is the pseudo differential operator $\mathrm{Op}(l(\tau, \varphi, x, \xi))$ of order -1

$$\begin{aligned} L^\tau &:= \mathcal{A}^\tau \Big(3\Lambda\partial_x \circ b(\tau)\Big)(\mathcal{A}^\tau)^{-1} \\ &= -\Big(\mathrm{I} - \Lambda\mathfrak{R}\Big)^{-1} \circ \Lambda \circ g(\tau, \varphi, x) \circ \partial_x \circ \tilde{\beta}(\varphi, x) \end{aligned} \tag{5.7.99}$$

where (recall (5.1.2))

$$\begin{aligned} & g(\tau, \varphi, x) := 3(1 + \tilde{\beta}_x^2(\varphi, x)), \qquad \mathfrak{R} := \mathrm{Op}(f_0(\varphi, x) + f_1(\varphi)\mathrm{i}\xi), \\ & f_0(\varphi, x) := \tilde{\beta}_x^2 + 2\tilde{\beta}_x - \frac{(1+\tilde{\beta}_x^2)}{2}\partial_{xx}\left(\frac{1}{(1+\tilde{\beta}_x)^2}\right), \\ & f_1(\varphi, x) := -\frac{3}{2}(1+\tilde{\beta}_x)^2\, \partial_x\left(\frac{1}{(1+\tilde{\beta}_x)^2}\right). \end{aligned} \tag{5.7.100}$$

Analysis of L^τ. The following estimates hold

$$\begin{aligned} & \|g\|_s^{\gamma,\mathcal{O}} \leq C(s)(1 + \|\beta\|_{s+1}^{\gamma,\mathcal{O}}\|\beta\|_{s_0+1}^{\gamma,\mathcal{O}}), \quad \|f_0\|_s^{\gamma,\mathcal{O}} \leq C(s)\|\beta\|_{s+3}^{\gamma,\mathcal{O}}, \quad \|f_1\|_s^{\gamma,\mathcal{O}} \leq C(s)\|\beta\|_{s+2}^{\gamma,\mathcal{O}}, \\ & \|\Delta_{12} g\|_s \leq_s \|\Delta_{12}\beta\|_{s+1}\|\beta\|_{s_0+1} + \|\Delta_{12}\beta\|_{s_0+1}\|\beta\|_{s+1}, \\ & \|\Delta_{12} f_0\|_s \leq C(s)\|\Delta_{12}\beta\|_{s+3}, \quad \|\Delta_{12} f_1\|_s \leq C(s)\|\Delta_{12}\beta\|_{s+2}, \\ & |f_0 + f_1\, \mathrm{i}\xi|_{1,s,\alpha}^{\gamma,\mathcal{O}} \leq C(s)\|\beta\|_{s+3}^{\gamma,\mathcal{O}}, \end{aligned} \tag{5.7.101}$$

By the fact that L^τ in (5.7.99) is regularizing, the problem (5.7.98) is locally well-posed. By Lemma 5.7.9 we have that

$$\mathrm{I} - \Lambda\mathfrak{R} = \mathrm{I} - (\mathrm{Op}(r) + R), \tag{5.7.102}$$

such that (see (5.7.101))

$$|r|_{-1,s,\alpha}^{\gamma,\mathcal{O}} \leq C(s,\alpha,\rho)\|\beta\|_{s+\rho+3}^{\gamma,\mathcal{O}}, \quad \mathbb{M}_R^\gamma(s,\mathtt{b}) \leq C(s,\rho)\|\beta\|_{s+2\rho+3}^{\gamma,\mathcal{O}}. \tag{5.7.103}$$

By Lemma 5.7.13 and (5.7.103) we have that

$$\begin{aligned} \Big(\mathrm{I} - \Lambda\mathfrak{R}\Big)^{-1} = \mathrm{I} + \mathrm{Op}(\tilde{r}) + \tilde{R}, \qquad & |\tilde{r}|_{-1,s,\alpha}^{\gamma,\mathcal{O}} \leq C(s,\alpha,\rho)\|\beta\|_{s+\rho^2-3\rho+9}^{\gamma,\mathcal{O}}, \\ & \mathbb{M}_{\tilde{R}}^\gamma(s) \leq C(s,\alpha,\rho)\|\beta\|_{s+\rho^2-3\rho+9}^{\gamma,\mathcal{O}} \end{aligned} \tag{5.7.104}$$

By Lemma 5.7.9 we have $\Lambda \circ g \circ \partial_x \circ \tilde{\beta} = \mathrm{Op}(d) + Q_\rho$ with

$$|d|_{s,\alpha}^{\gamma,\mathcal{O}} \leq \|\beta\|_{s+\rho+1}^{\gamma,\mathcal{O}}, \quad \mathbb{M}_{Q_\rho}^{\gamma}(s) \leq \|\beta\|_{s+2+2\rho}^{\gamma,\mathcal{O}}. \tag{5.7.105}$$

Then

$$\begin{aligned} L^\tau =&(\mathrm{I} + \mathrm{Op}(\tilde{r}) + \tilde{R}) \circ (\mathrm{Op}(d) + Q_\rho) = \mathrm{Op}(d) + \mathrm{Op}(\tilde{r}\#d) + Q_\rho \\ &+ \mathrm{Op}(\tilde{r}) \circ Q_\rho + \tilde{R} \circ \mathrm{Op}(d) + \tilde{R} \circ Q_\rho \overset{\text{Lemma5.7.9}}{=} \mathrm{Op}(d + \tilde{d}) + R_\rho \end{aligned} \tag{5.7.106}$$

where

$$\begin{aligned} &|\tilde{d}|_{s,\alpha}^{\gamma,\mathcal{O}} \leq \|\beta\|_{s+\tilde{\sigma}_1}^{\gamma,\mathcal{O}} \|\beta\|_{s_0+\tilde{\sigma}_1}^{\gamma,\mathcal{O}}, \quad \mathbb{M}_{R_\rho}^{\gamma}(s, \mathtt{b}) \leq_{s,\rho} \|\beta\|_{s+\tilde{\sigma}_1}^{\gamma,\mathcal{O}}, \\ &\mathbb{M}_{\Delta_{12} R_\rho}(s, \mathtt{b}) \leq_{s,\rho} \|\Delta_{12}\beta\|_{s+\tilde{\sigma}_1} + \|\beta\|_{s+\tilde{\sigma}_1} \|\Delta_{12}\beta\|_{s_0+\tilde{\sigma}_1}. \end{aligned} \tag{5.7.107}$$

for some constant $\tilde{\sigma}_1 = \tilde{\sigma}_1(\rho)$. Note that in principle we get a slightly different constant in each inequality, we are just taking the biggest of them for simplicity.
We call $l := d + \tilde{d}$ and we write

$$\begin{aligned} L^\tau = \mathrm{Op}(l) + R_\rho, \qquad &|l|_{-1,s,\alpha}^{\gamma,\mathcal{O}} \leq_{s,\alpha,\rho} \|\beta\|_{s+\tilde{\sigma}_1}^{\gamma,\mathcal{O}}, \\ &|\Delta_{12} l|_{-1,s,\alpha} \leq_{s,\rho} \|\Delta_{12}\beta\|_{s+\tilde{\sigma}_1} + \|\beta\|_{s+\tilde{\sigma}_1} \|\Delta_{12}\beta\|_{s_0+\tilde{\sigma}_1}. \end{aligned} \tag{5.7.108}$$

Approximate solution of (5.7.98). Now we look for an approximate solution

$$\Theta^\tau = \mathrm{Op}(1 + \vartheta(\tau, \varphi, x, \xi))$$

for the system (5.7.98). In order to do that we look for a symbol $\vartheta = \sum_{k=1}^{\rho-1} \vartheta_{-k}(\tau, \varphi, x, \xi)$ such that

$$\begin{cases} \partial_\tau \vartheta = l + l\#\vartheta + S^{-\rho}, \\ \vartheta(0, \varphi, x, \xi) = 0. \end{cases} \tag{5.7.109}$$

We solve recursively as follows

$$\begin{cases} \partial_\tau \vartheta_{-1} = l, \\ \vartheta_{-1}(0, \varphi, x, \xi) = 0, \end{cases} \qquad \begin{cases} \partial_\tau \vartheta_{-k} = \mathtt{r}_{-k}, \qquad 1 < k \leq \rho - 1 \\ \vartheta_{-k}(0, \varphi, x, \xi) = 0, \end{cases} \tag{5.7.110}$$

where

$$\mathtt{r}_{-k} := \sum_{j=1}^{k-1} l \#_{k-1-j} \vartheta_{-j} \in S^{-k} \tag{5.7.111}$$

then

$$\vartheta_{-1}(\tau) = \int_0^\tau l(s)\, ds, \qquad \vartheta_{-k}(\tau) = \int_0^\tau \mathtt{r}_{-k}(s)\, ds. \tag{5.7.112}$$

By recursion we have that

$$|\vartheta_{-k}|_{-k,s,\alpha}^{\gamma,\mathcal{O}} \leq C(s, \alpha, k) \|\beta\|_{s+k+\tilde{\sigma}_1}^{\gamma,\mathcal{O}} (\|\beta\|_{s_0+k+\tilde{\sigma}_1}^{\gamma,\mathcal{O}})^{k-1}, \qquad 1 \leq k \leq \rho - 1, \tag{5.7.113}$$

$$\begin{aligned} |\Delta_{12}\vartheta_{-k}|_{-k,s,\alpha} \leq_{s,\alpha,k} &\|\beta\|_{s+k+\tilde{\sigma}_1} \|\Delta_{12}\beta\|_{s_0+k+\tilde{\sigma}_1} \|\beta\|_{s_0+k+\tilde{\sigma}_1}^{k-2} \\ &+ \|\beta\|_{s_0+k+\tilde{\sigma}_1}^{k-1} \|\Delta_{12}\beta\|_{s+k+\tilde{\sigma}_1}, \end{aligned} \tag{5.7.114}$$

and so we get (5.7.93), (5.7.94). We write $C^\tau = \Theta^\tau + R^\tau$, where R^τ is an operator which satisfies the equation

$$\begin{cases} \partial_\tau R^\tau = L^\tau R^\tau + Q^\tau, \\ R^0 = 0, \end{cases} \tag{5.7.115}$$

where

$$Q^\tau := \mathrm{Op}(\mathtt{q}(\tau)) + R_\rho \Theta^\tau, \quad \mathtt{q}(\tau) := \sum_{k=1}^{\rho-1} l \#_{\geq \rho-1-k} \theta_{-k} \in S^{-\rho} \tag{5.7.116}$$

and by Lemma 5.7.6

$$\mathbb{M}^\gamma_{\mathrm{Op}(\mathtt{q})}(s,\mathtt{b}) \leq_{s,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\tilde\sigma_2} \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\tilde\sigma_2} \tag{5.7.117}$$

with $\tilde\sigma_2 := \tilde\sigma_2(\rho) > \tilde\sigma_1$. By Lemma 5.7.7, the operator Q^τ belongs to $\mathfrak{L}_\rho$ and we have the following bounds

$$\begin{aligned} \mathbb{M}^\gamma_{Q^\tau}(s,\mathtt{b}) &\leq \mathbb{M}^\gamma_{\mathrm{Op}(\mathtt{q})}(s,\mathtt{b}) + \mathbb{M}^\gamma_{R_\rho\Theta^\tau}(s,\mathtt{b}) \leq \mathbb{M}^\gamma_{\mathrm{Op}(\mathtt{q})}(s,\mathtt{b}) + 2\mathbb{M}^\gamma_{R_\rho}(s,\mathtt{b}) \\ &\quad + |\vartheta|^{\gamma,\mathcal{O}}_{-1,s+\rho,0} \mathbb{M}^\gamma_{R_\rho}(s_0,\mathtt{b}) \leq \|\beta\|^{\gamma,\mathcal{O}}_{s+\tilde\sigma_2}, \end{aligned} \tag{5.7.118}$$

$$\mathbb{M}_{\Delta_{12}Q^\tau}(s,\mathtt{b}) \leq \|\Delta_{12}\beta\|_{s+\tilde\sigma_2} + \|\beta\|_{s+\tilde\sigma_2} \|\Delta_{12}\beta\|_{s_0+\tilde\sigma_2}, \tag{5.7.119}$$

note that these bounds hold uniformly for $\tau \in [0,1]$. Now we have to prove that R^τ is belongs to the class $\mathfrak{L}_\rho$ (see Def. 5.7.4). By this fact we will deduce that C^τ and its derivatives are tame on $H^s(\mathbb{T}^{\nu+1})$.

Estimates for the remainder R^τ. We prove the bounds (5.7.95) and (5.7.96) (i.e. we show that R^τ belongs to $\mathfrak{L}_\rho$ in Def. 5.7.4 for $\tau \in [0,1]$). We use the integral formulation for the problem (5.7.115), namely

$$R^\tau = \int_0^\tau (L^t R^t + Q^t)\, dt. \tag{5.7.120}$$

We start by showing that R^τ satisfies item (i) of Definition 5.7.4 with $\vec{\mathtt{b}} = 0$. Let $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho$. We check that the operator $\langle D_x \rangle^{m_1} R^\tau \langle D_x \rangle^{m_2}$ is Lip-0-tame according to Definition 2.3.5. We have

$$\begin{aligned} \langle D_x \rangle^{m_1} R^\tau \langle D_x \rangle^{m_2} &= \int_0^\tau \langle D_x \rangle^{m_1} L^t \langle D_x \rangle^{-m_1} \langle D_x \rangle^{m_1} R^t \langle D_x \rangle^{m_2} dt \\ &\quad + \int_0^\tau \langle D_x \rangle^{m_1} Q^t \langle D_x \rangle^{m_2} dt. \end{aligned} \tag{5.7.121}$$

By (5.7.118) we have, for $s_0 \leq s \leq S_{max}$, that

$$\| \int_0^\tau \langle D_x \rangle^{m_1} Q^t \langle D_x \rangle^{m_2} u\, dt \|^{\gamma,\mathcal{O}}_s \leq \|\beta\|^{\gamma,\mathcal{O}}_{s+\tilde\sigma_2} \|u\|_{s_0} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\tilde\sigma_2} \|u\|_s, \tag{5.7.122}$$

for $\tau \in [0,1]$, $u \in H^s$. Moreover, by recalling the definition of L^t in (5.7.108), by using the fact that R_ρ in (5.7.107) is in the class $\mathfrak{L}_\rho$ and using the estimates (5.7.108) on the symbol l we claim that

$$\| \int_0^\tau \langle D_x \rangle^{m_1} L^t \langle D_x \rangle^{-m_1} u\, dt \|^{\gamma,\mathcal{O}}_s \leq_{s,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\tilde\sigma_1} \|u\|_{s_0} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\tilde\sigma_1} \|u\|_s. \tag{5.7.123}$$

Indeed the bound for $\mathrm{Op}(l)$ are trivial. In order to treat the remainder R_ρ we note that

$$\langle D_x\rangle^{m_1} R_\rho \langle D_x\rangle^{-m_1} = \langle D_x\rangle^{m_1} R_\rho \langle D_x\rangle^{\rho-m_1}\langle D_x\rangle^{-\rho}$$

and $\langle D_x\rangle^{m_1} R_\rho \langle D_x\rangle^{\rho-m_1}$ is Lip-0-tame, since $R_\rho \in \mathfrak{L}_\rho$, moreover $\langle D_x\rangle^{-\rho} \in \mathfrak{L}_\rho$. Then by Lemma 2.3.6 our claim follows.
By using bounds (5.7.122) and (5.7.123) with $s = s_0$ one obtains

$$\sup_{\tau\in[0,1]} \|\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2} u\|_{s_0}^{\gamma,\mathcal{O}} \leq_\rho \|\beta\|_{s_0+\tilde\sigma_1}^{\gamma,\mathcal{O}} \sup_{\tau\in[0,1]} \|\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2} u\|_{s_0}^{\gamma,\mathcal{O}} + \|\beta\|_{s_0+\tilde\sigma_2}\|u\|_{s_0}, \tag{5.7.124}$$

hence, by (5.7.118) and for δ in (5.7.91) small enough, one gets

$$\sup_{\tau\in[0,1]} \|\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2} u\|_{s_0}^{\gamma,\mathcal{O}} \leq_{s,\rho} \|\beta\|_{s_0+\tilde\sigma_2}^{\gamma,\mathcal{O}}\|u\|_{s_0}, \tag{5.7.125}$$

for any $u \in H^s$. Now for any $s_0 \leq s \leq S_{max}$, again by (5.7.122), (5.7.123), we have

$$\begin{aligned}\sup_{\tau\in[0,1]} \|\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2} u\|_{s}^{\gamma,\mathcal{O}} \leq_{s,\rho}\ & \|\beta\|_{s+\tilde\sigma_1}^{\gamma,\mathcal{O}} \sup_{\tau\in[0,1]} \|\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2} u\|_{s_0}^{\gamma,\mathcal{O}} \\ &+ \|\beta\|_{s_0+\tilde\sigma_1}^{\gamma,\mathcal{O}} \sup_{\tau\in[0,1]} \|\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2} u\|_{s}^{\gamma,\mathcal{O}} \\ &+ \|\beta\|_{s_0+\tilde\sigma_2}^{\gamma,\mathcal{O}}\|u\|_s + \|\beta\|_{s+\tilde\sigma_2}^{\gamma,\mathcal{O}}\|u\|_{s_0},\end{aligned} \tag{5.7.126}$$

for any $u \in H^s$, $s_0 \leq s \leq S_{max}$. Thus, by the smallness of β in (5.7.91) and estimate (5.7.125), the bound (5.7.126) implies

$$\sup_{\tau\in[0,1]} \|\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2} u\|_{s}^{\gamma,\mathcal{O}} \leq_{s,\rho} \|\beta\|_{s_0+\tilde\sigma_2}^{\gamma,\mathcal{O}}\|u\|_s + \|\beta\|_{s+\tilde\sigma_2}^{\gamma,\mathcal{O}}\|u\|_{s_0}. \tag{5.7.127}$$

This means that

$$\sup_{\tau\in[0,1]} \mathfrak{M}^\gamma_{R^\tau}(-\rho, s) = \sup_{\tau\in[0,1]} \sup_{\substack{m_1+m_2=\rho\\ m_1,m_2\geq 0}} \mathfrak{M}^\gamma_{\langle D_x\rangle^{m_1} R^\tau \langle D_x\rangle^{m_2}}(0,s) \leq_{s,\rho} \|\beta\|_{s+\tilde s_2}^{\gamma,\mathcal{O}}. \tag{5.7.128}$$

For $\vec{\mathtt{b}} \in \mathbb{N}^\nu$ with $|\vec{\mathtt{b}}| = \mathtt{b} \leq \rho - 2$, we consider the operator $\partial_\varphi^{\vec{\mathtt{b}}} R^\tau$ and we show that the operator $\langle D_x\rangle^{m_1}\partial_{\varphi_m}^{\mathtt{b}} R^\tau \langle D_x\rangle^{m_2}$ is Lip-0-tame for any $m_1, m_2 \in \mathbb{R}$, $m_1, m_2 \geq 0$ and $m_1 + m_2 = \rho - \mathtt{b}$. We prove that

$$\mathfrak{M}^\gamma_{\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x\rangle^{m_2}}(0,s) \leq_{s,\rho} \|\beta\|_{s+\tilde\sigma_3}^{\gamma,\mathcal{O}}, \quad m_1 + m_2 = \rho - \mathtt{b}, \tag{5.7.129}$$

for some $\tilde\sigma_3 := \tilde\sigma_3(\rho) \geq \tilde\sigma_3 > 0$, by induction on $0 \leq \mathtt{b} \leq \rho - 1$. For $\mathtt{b} = 0$ the bound follows by (5.7.128). Assume now that (5.7.129) holds for any $\tilde{\mathtt{b}}$ such that $0 \leq \tilde{\mathtt{b}} < \mathtt{b} \leq \rho - 2$. We show (5.7.129) for $\mathtt{b} = \tilde{\mathtt{b}} + 1$. By (5.7.120) we have

$$\begin{aligned}\langle D_x\rangle^{m_1}\partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x\rangle^{m_2} = &\sum_{\vec{\mathtt{b}_1}+\vec{\mathtt{b}_2}=\vec{\mathtt{b}}} C(|\vec{\mathtt{b}_1}|, |\vec{\mathtt{b}_2}|) \int_0^\tau \langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}_1}} L^t)\partial_\varphi^{\vec{\mathtt{b}_2}}(R^t)\langle D_x\rangle^{m_2}\,dt \\ &+ \int_0^\tau \langle D_x\rangle^{m_1}(\partial_\varphi^{\vec{\mathtt{b}}} Q^t)\langle D_x\rangle^{m_2}\,dt.\end{aligned} \tag{5.7.130}$$

By (5.7.118) we know that, for any $t \in [0,1]$, the operator $\langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b}}} Q^t) \langle D_x \rangle^{m_2}$ is Lip-0-tame. We write

$$\langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b_1}}} L^t) \partial_\varphi^{\vec{\mathtt{b_2}}} (R^t) \langle D_x \rangle^{m_2} = \langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b_1}}} L^t) \langle D_x \rangle^{-m_1 - |\vec{\mathtt{b_1}}|} \langle D_x \rangle^{m_1 + |\vec{\mathtt{b_1}}|} \partial_\varphi^{\vec{\mathtt{b_2}}} (R^t) \langle D_x \rangle^{m_2}. \quad (5.7.131)$$

We study the case $|\vec{\mathtt{b_2}}| \le \mathtt{b} - 1$. By the inductive hypothesis we have that $\langle D_x \rangle^{m_1 + |\vec{\mathtt{b_1}}|} \partial_\varphi^{\vec{\mathtt{b_2}}} (R^t) \langle D_x \rangle^{m_2}$ is Lip-0-tame since $m_1 + |\vec{\mathtt{b_1}}| + m_2 = \rho - |\vec{\mathtt{b_2}}|$, hence the bound (5.7.129) holds for $\mathtt{b} = |\vec{\mathtt{b_2}}|$. By reasoning as for the proof of the bound (5.7.123) we have

$$\| \langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b_1}}} L^t) \langle D_x \rangle^{-m_1 - |\vec{\mathtt{b_1}}|} u \|_s^{\gamma, \mathcal{O}} \le_{s,\rho} \|\beta\|_{s+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \|u\|_{s_0} + \|\beta\|_{s_0+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \|u\|_s, \quad (5.7.132)$$

for $u \in H^s$, $s_0 \le s \le S_{max}$. By (5.7.132), the inductive hypothesis on $\partial_\varphi^{\vec{\mathtt{b_2}}} R^\tau$ and (5.7.118) we get

$$\mathfrak{M}^\gamma_{\langle D_x \rangle^{m_1} (\partial_\varphi^{\vec{\mathtt{b_1}}} L^t) \partial_\varphi^{\vec{\mathtt{b_2}}} (R^t) \langle D_x \rangle^{m_2}} (0, s) \le_{s,\rho} \|\beta\|_{s+\tilde{\sigma}_3}^{\gamma,\mathcal{O}}. \quad (5.7.133)$$

Note also that By Lemma 2.2.5, bounds (5.7.107) and (5.7.108) we have that (5.7.132) holds for $\mathtt{b}_1 = 0$. Hence

$$\begin{aligned}
\sup_{\tau \in [0,1]} \| \langle D_x \rangle^{m_1} \partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x \rangle^{m_2} u \|_s^{\gamma,\mathcal{O}} &\le_{s,\rho} \sup_{\tau \in [0,1]} \| \langle D_x \rangle^{m_1} L^\tau \partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x \rangle^{m_2} u \|_s \\
&\quad + \|\beta\|_{s+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \|u\|_{s_0} + \|\beta\|_{s_0+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \|u\|_s^{\gamma,\mathcal{O}} \\
&\overset{(5.7.132)}{\le_{s,\rho}} \|\beta\|_{s+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \sup_{\tau \in [0,1]} \| \langle D_x \rangle^{m_1} \partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x \rangle^{m_2} u \|_{s_0}^{\gamma,\mathcal{O}} \\
&\quad + \|\beta\|_{s_0+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \sup_{\tau \in [0,1]} \| \langle D_x \rangle^{m_1} \partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x \rangle^{m_2} u \|_s^{\gamma,\mathcal{O}} \\
&\quad + \|\beta\|_{s+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \|u\|_{s_0} + \|\beta\|_{s_0+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \|u\|_s.
\end{aligned} \quad (5.7.134)$$

Hence using (5.7.134) for $s = s_0$ and the smallness of β in (5.7.91) we get

$$\sup_{\tau \in [0,1]} \| \langle D_x \rangle^{m_1} \partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x \rangle^{m_2} u \|_{s_0}^{\gamma,\mathcal{O}} \le_{s,\rho} \|\beta\|_{s_0+\tilde{\sigma}_3}^{\gamma,\mathcal{O}} \|u\|_{s_0}. \quad (5.7.135)$$

Then using again (5.7.135) one obtains the bound for any $s_0 \le s \le S_{max}$

$$\sup_{\tau \in [0,1]} \mathfrak{M}^\gamma_{R^\tau}(-\rho + \mathtt{b}, s) := \sup_{\tau \in [0,1]} \sup_{\substack{m_1 + m_2 = \rho - \mathtt{b} \\ m_1, m_2 \ge 0 \\ |\vec{\mathtt{b}}| \le \mathtt{b}}} \mathfrak{M}^\gamma_{\langle D_x \rangle^{m_1} \partial_\varphi^{\vec{\mathtt{b}}} R^\tau \langle D_x \rangle^{m_2}} (0, s) \le_{s,\rho} \|\beta\|_{s+\tilde{\sigma}_3}^{\gamma,\mathcal{O}}. \quad (5.7.136)$$

The estimates for $\mathfrak{M}_{[R^\tau, \partial_x]}(s)$ and $\mathfrak{M}_{[\partial_\varphi^{\vec{\mathtt{b}}} R^\tau, \partial_x]}(s)$ follow by the same arguments. We have obtained the estimate for $\mathbb{M}^\gamma_{R^\tau}(s, \mathtt{b})$ in (5.7.95). The estimate on the Lipschitz variation with respwct to the variable i (5.7.96) for the variable i follows by by Leibnitz rule and by (5.7.95), (5.7.108), (5.7.119) as in the previous cases.
We proved (5.7.95) with $\sigma_1 = \tilde{\sigma}_3$. □

Corollary 5.7.17. *Fix $\mathtt{b} \in \mathbb{N}$. There exists $\mu = \mu(\rho)$ such that, if $\|\beta\|_{s_0+\mu}^{\gamma,\mathcal{O}} \le 1$, then the flow $\Psi^\tau(\varphi)$ of (5.7.89) satisfies for $s \ge s_0$,*

$$\sup_{\tau \in [0,1]} \|\Psi^\tau u\|_s^{\gamma,\mathcal{O}} \le_s \left(\|u\|_s^{\gamma,\mathcal{O}} + \|b\|_{s+\mu}^{\gamma,\mathcal{O}} \|u\|_{s_0}^{\gamma,\mathcal{O}} \right), \quad (5.7.137)$$

$$\sup_{\tau\in[0,1]} \|(\Psi^\tau - \mathrm{I})u\|_s^{\gamma,\mathcal{O}} \leq_s \left(\|\beta\|_{s_0+\mu}^{\gamma,\mathcal{O}}\|u\|_{s+1}^{\gamma,\mathcal{O}} + \|\beta\|_{s+\mu}^{\gamma,\mathcal{O}}\|u\|_{s_0+1}^{\gamma,\mathcal{O}}\right), \tag{5.7.138}$$

$$\sup_{\tau\in[0,1]} \|(\Psi^\tau)^* u\|_s^{\gamma,\mathcal{O}} \leq_s \left(\|u\|_s^{\gamma,\mathcal{O}} + \|\beta\|_{s+s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s_0}^{\gamma,\mathcal{O}}\right) \tag{5.7.139}$$

$$\sup_{\tau\in[0,1]} \|((\Psi^\tau)^* - \mathrm{I})u\|_s^{\gamma,\mathcal{O}} \leq_s \left(\|\beta\|_{s_0+\mu}^{\gamma,\mathcal{O}}\|u\|_{s+1}^{\gamma,\mathcal{O}} + \|\beta\|_{s+\mu}^{\gamma,\mathcal{O}}\|u\|_{s_0+1}^{\gamma,\mathcal{O}}\right). \tag{5.7.140}$$

For any $|\alpha| \leq \mathtt{b}$, $m_1, m_2 \in \mathbb{R}$ *such that* $m_1 + m_2 = |\alpha|$, *for any* $s \geq s_0$ *there exist* $\mu_* = \mu_*(|\alpha|, m_1, m_2)$ *and* $\delta = \delta(m_1, s)$ *such that if* $\|\beta\|_{s_0+\mu_*}^{\gamma,\mathcal{O}} \leq \delta$, *then one has*

$$\sup_{\tau\in[0,1]} \|\langle D_x\rangle^{-m_1}\partial_\varphi^\alpha \Psi^\tau(\varphi)\langle D_x\rangle^{-m_2}u\|_s^{\gamma,\mathcal{O}} \leq_{s,\mathtt{b},m_1,m_2} \|u\|_s^{\gamma,\mathcal{O}} + \|\beta\|_{s+\mu_*}^{\gamma,\mathcal{O}}\|u\|_{s_0}^{\gamma,\mathcal{O}}. \tag{5.7.141}$$

and for $m_1 + m_2 = |\alpha| + 1$,

$$\begin{aligned}&\sup_{\tau\in[0,1]} \|\langle D_x\rangle^{-m_1}\partial_\varphi^\alpha \Delta_{12}\Psi^\tau(\varphi)\langle D_x\rangle^{-m_2}u\|_s \\ &\qquad \leq_{s,\mathtt{b},m_1,m_2} \|u\|_s\|\Delta_{12}\beta\|_{s_0+\mu} + \|u\|_{s_0+1}(\|\Delta_{12}\beta\|_{s+\mu} + \|\Delta_{12}\beta\|_{s_0+\mu}\|\beta\|_{s+\mu}).\end{aligned} \tag{5.7.142}$$

We omit the proof of the corollary above since it follows by using the same arguments of Lemma A.1.2 and by using the result of Proposition 5.7.16.

In the following we investigate the structure of an operator obtained by conjugating an operator like (5.7.6) by a transformation Φ^τ which is the flow of the system

$$\begin{cases}\partial_\tau \Phi^\tau u = \Pi_S^\perp[(J\circ b)\Pi_S^\perp[\Phi^\tau u]],\\ \Phi^0 u = u.\end{cases} \tag{5.7.143}$$

First we need study how the flow Φ^τ differs from Ψ^τ in (5.7.89). Secondly we study how symbols in the class S^m transform under changes of coordinates in the variables (x,ξ). The proofs of the following Lemmata are in the Appendix B.

Lemma 5.7.18. *Fix* $\rho \geq 0$. *There exist* $\sigma_1 := \sigma_1(\rho)$ *such that if* (5.7.91) *holds with* $\mu = \sigma_1$ *then the following holds. Let* Ψ^τ *be the flow of the system* (5.7.89) *and* Φ^τ *be the flow of* (5.7.143). *The map* Φ^τ *satisfies bounds like* (5.7.137)-(5.7.142) *(with a possible larger* $\mu = \mu(s_0, \rho)$*).*
In particular one has $\Phi^1 = \Pi_S^\perp \Psi^1 \Pi_S^\perp \circ (\mathrm{I} + \mathcal{R})$ *where* $\mathcal{R}$ *is an operator with the form* (5.6.5). *Moreover* $\mathcal{R}$ *belongs to* $\mathfrak{L}_\rho$ *and satisfies*

$$\mathbb{M}_{\mathcal{R}}^\gamma(s,\mathtt{b}) \leq_s \|\beta\|_{s+\sigma_1}^{\gamma,\mathcal{O}}, \tag{5.7.144}$$

$$\mathbb{M}_{\Delta_{12}\mathcal{R}}(s,\mathtt{b}) \leq_s \|\Delta_{12}\beta\|_{s+\sigma_1} + \|\Delta_{12}\beta\|_{s_0+\sigma_1}\|\beta\|_{s+\sigma_1}. \tag{5.7.145}$$

As a consequence $\Phi := \Phi^1$ *satisfies the estimates* (5.7.137)-(5.7.142) *with a possibly larger* $\mu > 0$.

The system (5.7.143) is an Hyperbolic PDE, thus we shall use a version of Egorov Theorem to study how pseudo differential operators change under the flow Φ^τ. This is the content of Theorem 5.7.20 which provides precise estimates for the transformed pseudo differential operators. Before to do that we need the following lemma.

Lemma 5.7.19. *Let $\mathcal{O}$ be a subset of $\mathbb{R}^\nu$. Let A be the operator defined for $w \in S^m$ as*

$$Aw = w(f(x), g(x)\xi), \quad f(x) := x + \beta(x), \quad g(x) = \frac{1}{1+\beta_x} \tag{5.7.146}$$

for some function β such that $\|\beta\|^{\gamma,\mathcal{O}}_{2s_0+2} < 1$. Then A is bounded, namely $Aw \in S^m$ and

$$|Aw|^{\gamma,\mathcal{O}}_{m,s,\alpha} \le |w|^{\gamma,\mathcal{O}}_{m,s,\alpha} + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2}. \tag{5.7.147}$$

for $s \ge 0$. For $s = s_0$ it is convenient to use the rougher estimate

$$|Aw|^{\gamma,\mathcal{O}}_{m,s_0,\alpha} \le |w|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+s_0}. \tag{5.7.148}$$

Lemma 5.7.19 follows directly by Lemma A.0.6 in Appendix A.

Theorem 5.7.20 (Egorov). *Fix $\rho \in \mathbb{N}$, $m \in \mathbb{R}$ with $\rho + m > 0$. Let $w(x,\xi) \in S^m$ with $w = w(\omega, i(\omega))$, Lipschitz in $\omega \in \mathcal{O} \subseteq \mathbb{R}^\nu$ and in the variable i. Let $\mathcal{A}^\tau$ be the flow of the system (5.7.60). There exist $\sigma_1 := \sigma_1(m,\rho)$ and $\delta := \delta(m,\rho)$ such that, if*

$$\|\beta\|^{\gamma,\mathcal{O}}_{s_0+\sigma_1} < \delta, \tag{5.7.149}$$

then $\mathcal{A}^\tau \mathrm{Op}(w)(\mathcal{A}^\tau)^{-1} = \mathrm{Op}(q(x,\xi)) + R$ where $q \in S^m$ and $R \in \mathfrak{L}_\rho$. Moreover the following estimates hold.

$$|q|^{\gamma,\mathcal{O}}_{m,s,\alpha} \le_{m,s,\alpha,\rho} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+\sigma_1} + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+\sigma_1} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+\sigma_1}, \tag{5.7.150}$$

$$\begin{aligned}
|\Delta_{12} q|_{m,s,\alpha} \le_{m,s,\alpha,\rho} & |w|_{m,s+1,\alpha+\sigma_1} \|\Delta_{12}\beta\|_{s_0+1} + |w|_{m,s_0+1,\alpha+\sigma_1} \|\Delta_{12}\beta\|_{s+1} + |\Delta_{12} w|_{m,s,\alpha+\sigma_1} \\
& + \sum_{\substack{s_1+s_2+s_3=s+1,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |w|_{m,s_1,\alpha+s_2+\sigma_1} \|\beta\|_{s_3+\sigma_1} \|\Delta_{12}\beta\|_{s_0+1} \\
& + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |\Delta_{12} w|_{m,s_1,s_2+\alpha+\sigma_1} \|\beta\|_{s_3+\sigma_1}.
\end{aligned} \tag{5.7.151}$$

Furthermore for any $\mathtt{b} \le \rho - 2$ and $s_0 \le s \le S_{max}$

$$\mathbb{M}^\gamma_R(s,\mathtt{b}) \le_{s,m,\rho} |w|^{\gamma,\mathcal{O}}_{m,s+\rho,\sigma_1} + \sum_{\substack{s_1+s_2+s_3=s+\rho,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,s_2+\sigma_1} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+\sigma_1}, \tag{5.7.152}$$

and for any $\mathtt{b} \le \rho - 3$, $s_0 \le s \le S_{max}$

$$\begin{aligned}
\mathbb{M}_{\Delta_{12} R}(s,\mathtt{b}) \le_{m,s,\rho} & |w|_{m,s+\rho,\sigma_1} \|\Delta_{12}\beta\|_{s_0+\sigma_1} + \|\Delta_{12}\beta\|_{s+\sigma_1} |w|_{m,s_0+\rho,\sigma_1} + |\Delta_{12} w|_{m,s+\rho,\sigma_1} \\
& + \sum_{\substack{s_1+s_2+s_3=s+\rho,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |w|_{m,s_1,s_2+\sigma_1} \|\beta\|_{s_3+\sigma_1} \|\Delta_{12}\beta\|_{s_0+\sigma_1} \\
& + \sum_{\substack{s_1+s_2+s_3=s+\rho,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |\Delta_{12} w|_{m,s_1,s_2+\sigma_1} \|\beta\|_{s_3+\sigma_1}.
\end{aligned} \tag{5.7.153}$$

Proof. The operator $P(\tau) := \mathcal{A}^\tau \mathrm{Op}(w)(\mathcal{A}^\tau)^{-1}$ satisfies the equation

$$\begin{cases} \partial_\tau P(\tau) = [\mathtt{X}, P(\tau)], & \mathtt{X} = \partial_x \circ b =: \mathrm{Op}(\chi), \\ P(0) = \mathrm{Op}(w). \end{cases} \tag{5.7.154}$$

We construct an approximate solution of (5.7.154) by considering a pseudo differential operator $\mathrm{Op}(q)$ with

$$q = \sum_{k=0}^{m+\rho-1} q_{m-k}(x,\xi) \tag{5.7.155}$$

such that (see (5.7.154) for the definition of χ)

$$\begin{cases} \partial_\tau q_m = \{b\xi, q_m\}, \\ q_m(0) = w \end{cases} \qquad \begin{cases} \partial_\tau q_{m-k} = \{b\xi, q_{m-k}\} + r_{m-k} \\ q_{m-k}(0) = 0 \end{cases} \qquad k \geq 1 \tag{5.7.156}$$

where for $k \geq 1$ (recall (2.2.20))

$$\begin{aligned} r_{m-k} &:= \frac{1}{\mathrm{i}}\{b_x, q_{m-k+1}\} - \sum_{h=0}^{k-1} q_{m-h} \#_{k+1-h} \chi \\ &= -\frac{1}{\mathrm{i}} \partial_\xi q_{m-k+1}\, b_{xx} - \sum_{h=0}^{k-1} \frac{1}{\mathrm{i}^{k+1-h}(k+1-h)!} (\partial_\xi^{k+1-h} q_{m-h})(\partial_x^{k+1-h}\chi) \in S^{m-k}. \end{aligned}$$

By Lemma 5.7.9, or directly by interpolation, one has

$$|r_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha} \leq \sum_{h=0}^{k-1} |q_{m-h}|^{\gamma,\mathcal{O}}_{m-h,s,\alpha+k+1-h} + \sum_{h=0}^{k-1} |q_{m-h}|^{\gamma,\mathcal{O}}_{m-h,s_0,\alpha+k+1-h} \|\beta\|^{\gamma,\mathcal{O}}_{s+k+3-h}, \tag{5.7.157}$$

$$\begin{aligned} |\Delta_{12} r_{m-k}|_{m-k,s,\alpha} \leq & \sum_{h=0}^{k-1} |\Delta_{12} q_{m-h}|_{m-h,s,\alpha+k+1-h} + \sum_{h=0}^{k-1} |\Delta_{12} q_{m-h}|_{m-h,s_0,\alpha+k+1-h} \|\beta\|_{s+k+3-h} \\ & + \sum_{h=0}^{k-1} |q_{m-h}|_{m-h,s,\alpha+k+1-h} \|\Delta_{12}\beta\|_{s_0+k+3-h} \\ & + \sum_{h=0}^{k-1} |q_{m-h}|_{m-h,s_0,\alpha+k+3-h} \|\Delta_{12}\beta\|_{s+k+3-h}. \end{aligned} \tag{5.7.158}$$

Hence we can solve (5.7.156) iteratively. The first equation has the solution

$$q_m(\tau,x,\xi) = w(\gamma^{\tau,0}(x,\xi)) \tag{5.7.159}$$

where

$$\gamma^{\tau,0}(x,\xi) = \big(f(\tau,x), \xi g(\tau,x)\big), \qquad f(\tau,x) := x + \tau\beta(x), \quad g(\tau,x) := \frac{1}{1+\tau\beta_x(x)}. \tag{5.7.160}$$

Hence by Lemma 5.7.19 we have

$$|q_m|^{\gamma,\mathcal{O}}_{m,s,\alpha} \leq_{s,\alpha} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha} + \sum_{\substack{s_1+s_2+s_3=s, \\ s_1<s, s_1,s_2,s_3\geq 0, \\ s_1+s_2\geq 1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2}. \tag{5.7.161}$$

For any $k \geq 1$, the solution of (5.7.156) is

$$q_{m-k}(\tau, x, \xi) = \int_0^\tau r_{m-k}(\gamma^{0,t}\gamma^{\tau,0}(x,\xi))\,dt. \tag{5.7.162}$$

We observe that

$$\gamma^{0,t}\gamma^{\tau,0}(x,\xi) = (\tilde{f}, \tilde{g}\,\xi) \tag{5.7.163}$$

with

$$\tilde{f}(x) := x + \tau\beta(x) + \tilde{\beta}(t, x + \tau\beta(x)), \qquad \tilde{g}(x) := \frac{1 + t\beta_x(x + \tau\beta(x))}{1 + \tau\beta_x(x)}. \tag{5.7.164}$$

Thus if $\tilde{A}r := r(\tilde{f}, \tilde{g}\,\xi)$ we have (recall that $\tau \in [0,1]$)

$$|q_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha} \leq_{s,\alpha} |\tilde{A}r_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha}, \qquad |q_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s_0,\alpha} \leq_{\alpha} |\tilde{A}r_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s_0,\alpha} \leq |r_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s_0,\alpha+s_0} \tag{5.7.165}$$

and by Lemma 5.7.19 with $A \rightsquigarrow \tilde{A}$

$$|q_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha} \leq_{s,\alpha} |r_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha} + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\geq 0,\\ s_1+s_2\geq 1}} |r_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s_1,\alpha+s_2} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2}. \tag{5.7.166}$$

We want to prove inductively (here we are dropping the constraints $s_1 < s, s_1,\ s_2, s_3 \geq 0$ and $s_1 + s_2 \geq 1$ in the sum over s_1, s_2, s_3 only to shorten the notations)

$$\begin{aligned} |q_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha} &\leq_{s,\alpha,\rho} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+2k} \\ &\quad + \sum_{s_1+s_2+s_3=s} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+k(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2+k}, \quad k = 0, \dots, m+\rho \\ |q_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s_0,\alpha} &\leq_{\alpha,\rho} |w|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+s_0+k(s_0+2)}. \end{aligned} \tag{5.7.167}$$

For $k = 0$ this is proved in (5.7.161). Now assume that (5.7.167) holds, up to some $k-1 \geq 0$. We use (5.7.157) to bound q_{m-k}. First we give a bound for r_{m-k} in terms of the norm of the symbol w

by substituing the inductive hypothesis (5.7.167).

$$
\begin{aligned}
|r_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha} &\leq_{s,\alpha} \sum_{h=0}^{k-1} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+h+k+1} + \sum_{h=0}^{k-1} |w|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+k+1-h+s_0+h(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s+k+3-h} \\
&\quad + \sum_{h=0}^{k-1} \sum_{s_1+s_2+s_3=s} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+h(s_0+2)+k+1-h} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2+h} \\
&\leq_{s,\alpha} \sum_{h=0}^{k-1} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+h+k+1} \\
&\quad + \sum_{h:k\geq 2h-1} |w|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+k+1-h+s_0+h(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s+k+3-h} \\
&\quad + \sum_{\substack{s_1+s_2+s_3=s,\\ h:k\geq 2h-1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+k+1-h+s_2+h(s_0+2)+s_0} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+k-h+3} \\
&\quad + \sum_{h:k<2h-1} |w|^{\gamma,\mathcal{O}}_{m,s_0,\alpha+k+1-h+s_0+h(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s+k+3-h} \\
&\quad + \sum_{\substack{s_1+s_2+s_3=s,\\ h:k<2h-1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+k+1-h+s_2+h(s_0+2)+s_0} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2+h} \\
&\leq_{s,\alpha,\rho} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+2k} + \sum_{s_1+s_2+s_3=s} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+k(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2+k}.
\end{aligned}
\tag{5.7.168}
$$

Then by (5.7.166) and (5.7.168)

$$
\begin{aligned}
|q_{m-k}|^{\gamma,\mathcal{O}}_{m-k,s,\alpha} &\leq_{s,\alpha,k} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+2k} + \sum_{s_1+s_2+s_3=s} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+k(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2+k} \\
&\quad + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\geq 0,\\ s_1+s_2\geq 1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+2k} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2} \\
&\quad + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\geq 0,\\ s_1+s_2\geq 1}} \Big(\sum_{n_1+n_2+n_3=s_1+k} |w|^{\gamma,\mathcal{O}}_{m,n_1,\alpha+n_2+k(s_0+2)+s_2} \|\beta\|^{\gamma,\mathcal{O}}_{n_3+s_0+2+k} \Big) \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2} \\
&\leq_{s,\alpha,k} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+2k} + \sum_{s_1+s_2+s_3=s+k} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+k(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2} \\
&\quad + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\geq 0,\\ s_1+s_2\geq 1}} \sum_{n_1+n_2+n_3=s_1+k} |w|^{\gamma,\mathcal{O}}_{m,n_1,\alpha+n_2+k(s_0+2)+s_2} \|\beta\|^{\gamma,\mathcal{O}}_{n_3+s_3+s_0+2+k} \|\beta\|^{\gamma,\mathcal{O}}_{s_0+2} \\
&\leq_{s,\alpha,k} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha+2k} + \sum_{s_1+s_2+s_3=s} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2+k(s_0+2)} \|\beta\|^{\gamma,\mathcal{O}}_{s_3+s_0+2+k}
\end{aligned}
$$

that is the estimate (5.7.167). By (5.7.162) we have

$$
\Delta_{12} q_{m-k}(\tau,x,\xi) = \int_0^\tau \Delta_{12}\big(r_{m-k}(\gamma^{0,s}\gamma^{\tau,0}(x,\xi))\big)\, ds \tag{5.7.169}
$$

and recalling (5.7.164)

$$\begin{aligned}|\Delta_{12} q_{m-k}|_{m-k,s,\alpha} \leq_{s,\alpha} &|\tilde{A}(\partial_x r_{m-k})\,(\Delta_{12}\tilde{f})|_{m-k,s,\alpha} + |\tilde{A}(\partial_\xi r_{m-k})\,(\Delta_{12}\tilde{g}\,\xi)|_{m-k,s,\alpha} \\ &+ |\tilde{A}(\Delta_{12} r_{m-k})|_{m-k,s,\alpha}.\end{aligned} \tag{5.7.170}$$

The first two terms of the right hand side in (5.7.170) are bounded by (5.7.168) and (2.1.4). For the last summand we proceed by induction as above using (5.7.158). We obtain

$$\begin{aligned}|\Delta_{12} q_{m-k}|_{m-k,s,\alpha} \leq\ & |w|_{m,s+1,\alpha+2k+1}\|\Delta_{12}\beta\|_{s_0+1} \\ &+ \sum_{s_1+s_2+s_3=s+1} |w|_{m,s_1,\alpha+s_2+s_0+1+k(s_0+2)}\|\beta\|_{s_3+2s_0+2+k}\|\Delta_{12}\beta\|_{s_0+1} \\ &+ |w|_{m,s_0+1,\alpha+s_0+1+k(s_0+2)}\|\Delta_{12}\beta\|_{s+1} + |\Delta_{12} w|_{m,s,\alpha+2k} \\ &+ \sum_{s_1+s_2+s_3=s} |\Delta_{12} w|_{m,s_1,s_2+\alpha+k(s_0+2)}\|\beta\|_{s_3+2s_0+2+k}.\end{aligned} \tag{5.7.171}$$

Then we have (5.7.150) and (5.7.151). Now we have (recall (5.7.155))

$$P(\tau) = Q + R, \qquad Q = \mathrm{Op}(q) \in OPS^m \tag{5.7.172}$$

and by the construction of Q we get that

$$\begin{cases}\partial_\tau R(\tau) = [\mathtt{X}, R] + \mathcal{M}, \\ R(0) = 0\end{cases} \tag{5.7.173}$$

where

$$\mathcal{M} = -\mathrm{Op}\Big(\mathrm{i}\{b_x, q_{-\rho+1}\} + \sum_{k=0}^{m+\rho-1} q_{m-k}\#_{\geq m-k+1+\rho}\chi\Big) \in OPS^{-\rho}. \tag{5.7.174}$$

By Lemma 5.7.6 $\mathcal{M} \in \mathfrak{L}_\rho$ and using (2.2.14) (recall also the Definition (2.2.3)) we have for all $s_0 \leq s \leq S_{max}$ with $\mathtt{b} \leq \rho - 2$

$$\mathbb{M}^\gamma_{\mathcal{M}}(s,\mathtt{b}) \leq_{s,\rho,m} |w|^{\gamma,\mathcal{O}}_{m,s+\rho,\sigma_1} + \sum_{s_1+s_2+s_3=s+\rho} |w|^{\gamma,\mathcal{O}}_{m,s_1,s_2+\sigma_1}\|\beta\|^{\gamma,\mathcal{O}}_{s_3+\sigma_1} \tag{5.7.175}$$

and for $\mathtt{b} \leq \rho - 3$

$$\begin{aligned}\mathbb{M}_{\Delta_{12}\mathcal{M}}(s,\mathtt{b}) \leq_s\ & |w|_{m,s+\rho,\sigma_1}\|\Delta_{12}\beta\|_{s_0+\sigma_1} + |w|_{m,s_0+\sigma_1,\sigma_1}\|\Delta_{12}\beta\|_{s+\sigma_1} \\ &+ \|\Delta_{12}\beta\|_{s_0+\sigma_1}\sum_{s_1+s_2+s_3=s+\rho} |w|_{m,s_1,s_2+\sigma_1}\|\beta\|_{s_3+\sigma_1} \\ &+ |\Delta_{12} w|_{m,s+\sigma_1,\sigma_1} + \sum_{s_1+s_2+s_3=s+\rho} |\Delta_{12} w|_{m,s_1,s_2+\sigma_1}\|\beta\|_{s+\sigma_1},\end{aligned} \tag{5.7.176}$$

for some $\sigma_1 > 0$. If $V(\tau) := R(\tau)\mathcal{A}^\tau$ then

$$\partial_\tau V = \mathtt{X} V + \mathcal{M}\mathcal{A}^\tau \tag{5.7.177}$$

and so

$$V^\tau = \int_0^\tau \mathcal{A}^\tau(\mathcal{A}^s)^{-1}\mathcal{M}\mathcal{A}^s\,ds \quad \Rightarrow \quad R(\tau) = \int_0^\tau \mathcal{A}^\tau(\mathcal{A}^s)^{-1}\mathcal{M}\mathcal{A}^s(\mathcal{A}^\tau)^{-1}\,ds. \tag{5.7.178}$$

By Lemma 5.7.14 $R^\tau \in \mathfrak{L}_\rho$ for any $\tau \in [0,1]$. By (5.7.61) we have that, for any $\tau \in [0,1]$, taking σ_1 possibly larger than before in order to fit the assumptions of Lemma 5.7.14

$$\mathbb{M}^\gamma_{R^\tau}(s,\mathtt{b}) \le_s \mathbb{M}^\gamma_{\mathcal{M}}(s) + \|\beta\|^{\gamma,\mathcal{O}}_{s+\sigma_1}\mathbb{M}^\gamma_{\mathcal{M}}(s_0). \tag{5.7.179}$$

Then by Leibniz rule and Lemma A.1.3 we have by (5.7.176)

$$\begin{aligned}\mathbb{M}_{\Delta_{12}R}(s,\mathtt{b}) \le_s\, & \mathbb{M}^\gamma_{\mathcal{M}}(s,\mathtt{b})\|\Delta_{12}\beta\|_{s_0} + \mathbb{M}^\gamma_{\mathcal{M}}(s_0,\mathtt{b})\|\Delta_{12}\beta\|_s + \mathbb{M}^\gamma_{\mathcal{M}}(s_0,\mathtt{b})\|\Delta_{12}\beta\|_{s_0}\|\beta\|_{s+\sigma_1}\\ & + \mathbb{M}_{\Delta_{12}\mathcal{M}}(s,\mathtt{b}) + \mathbb{M}_{\Delta_{12}\mathcal{M}}(s_0,\mathtt{b})\|\beta\|_{s+\sigma_1}.\end{aligned}$$

We obtain (5.7.152) and (5.7.153) by using respectively (5.7.175) and (5.7.176). □

The following proposition describes the structure of an operator like $\mathcal{L}_\omega$ conjugated by a flow of a system like (5.7.143).

Proposition 5.7.21 (Conjugation). *Let $\mathcal{O}$ be a subset of $\mathbb{R}^\nu$. Fix $\rho > 3$, $\alpha \in \mathbb{N}$ and consider a linear operator*

$$\mathcal{L} := \Pi_S^\perp\Big(\mathcal{D}_\omega - J\circ(m + a(\varphi,x)) + \mathcal{Q}\Big) \tag{5.7.180}$$

where $m = m(\omega)$ is a real constant, $a = a(\omega, i(\omega)) \in C^\infty(\mathbb{T}^{\nu+1})$ is real, both are Lipschitz in $\omega \in \mathcal{O}$ and a is Lipschitz in the variable i. Moreover $\mathcal{Q} = \mathrm{Op}(\mathtt{q}(\varphi,x,\xi)) + \widehat{\mathcal{Q}}$ with $\widehat{\mathcal{Q}} \in \mathfrak{L}_\rho$ and $\mathtt{q} = \mathtt{q}(\omega, i(\omega)) \in S^{-1}$ satisfying

$$|\mathtt{q}|^{\gamma,\mathcal{O}}_{-1,s,\alpha} \le_{s,\alpha} \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\sigma_2}, \tag{5.7.181}$$

$$|\Delta_{12}\mathtt{q}|_{-1,s,\alpha} \le_{s,\alpha} \mathtt{k}_3(\|\Delta_{12}p\|_{s+\sigma_2} + \|\Delta_{12}p\|_{s_0+\sigma_2}\|p\|_{s+\sigma_2}). \tag{5.7.182}$$

Here $\mathtt{k}_1, \mathtt{k}_2, \mathtt{k}_3, \sigma_2 > 0$ are constants depending on q while $p = p(\omega, i(\omega)) \in C^\infty(\mathbb{T}^{\nu+1})$, is Lipschitz in ω and in the variable i .
Assume that

$$\|\beta\|^{\gamma,\mathcal{O}}_{s_0+\sigma_3} + \|a\|^{\gamma,\mathcal{O}}_{s_0+\sigma_3} + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s_0+\sigma_3} + \mathtt{k}_1 + \mathbb{M}^\gamma_{\widehat{\mathcal{Q}}}(s_0,\mathtt{b}) \le \delta_*\,, \tag{5.7.183}$$

for some $\sigma_3 := \sigma_3(\rho)$ large enough and $\delta_* := \delta_*(\rho)$ small enough, so that in particular (5.7.91) is satisfied. Consider $\Phi := \Phi^1$ the flow at time one of the system (5.7.143), where b is defined in (5.7.90). Then we have

$$\mathcal{L}_+ := \Phi\mathcal{L}\Phi^{-1} = \Pi_S^\perp\Big(\mathcal{D}_\omega - J\circ(m + a_+(\varphi,x)) + \mathcal{Q}_+\Big) \tag{5.7.184}$$

where

$$m + a_+(\varphi,x) := -(\mathcal{D}_\omega\tilde\beta)(\varphi, x+\beta(\varphi,x)) + (m + a(\varphi, x+\beta(\varphi,x)))(1 + \tilde\beta_x(\varphi, x+\beta(\varphi,x))) \tag{5.7.185}$$

with $\tilde\beta$ the function such that $x + \tilde\beta(\varphi,x)$ is the inverse of the diffeomorphism of the torus $x \mapsto x + \beta(\varphi,x)$. The operator $\mathcal{Q}_+ := \mathrm{Op}(\mathtt{q}_+(\varphi,x,\xi)) + \widehat{\mathcal{Q}}_+$, with

$$\begin{aligned}|\mathtt{q}_+|^{\gamma,\mathcal{O}}_{-1,s,\alpha} &\le_{s,\alpha,\rho} \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\sigma_3} + \|\beta\|^{\gamma,\mathcal{O}}_{s+\sigma_3} + \|a\|^{\gamma,\mathcal{O}}_{s+\sigma_3},\\ |\Delta_{12}\mathtt{q}_+|_{-1,s,\alpha} &\le_{s,\alpha,\rho} \big(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\sigma_3} + \|\beta\|_{s+\sigma_3} + \|a\|_{s+\sigma_3}\big)\|\Delta_{12}\beta\|_{s_0+\sigma_3}\\ &\quad + \mathtt{k}_3(\|\Delta_{12}p\|_{s+\sigma_3} + \|\Delta_{12}p\|_{s_0+\sigma_3}\|p\|_{s+\sigma_3}) + \|\Delta_{12}\beta\|_{s+\sigma_3} + \|\Delta_{12}a\|_{s+\sigma_3}\\ &\quad + \big(\mathtt{k}_3\|\Delta_{12}p\|_{s_0+\sigma_3} + \|\Delta_{12}\beta\|_{s_0+\sigma_3} + \|\Delta_{12}a\|_{s_0+\sigma_3}\big)\|\beta\|_{s+\sigma_3}\end{aligned} \tag{5.7.186}$$

and $\widehat{\mathcal{Q}}_+ \in \mathfrak{L}_\rho$ *with, for* $s_0 \le s \le S_{max}$,

$$\mathbb{M}^{\gamma}_{\widehat{\mathcal{Q}}_+}(s, \mathtt{b}) \le_{s,\rho} \mathbb{M}^{\gamma}_{\widehat{\mathcal{Q}}}(s, \mathtt{b}) + \|\beta\|^{\gamma,\mathcal{O}}_{s+\sigma_3} + \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\sigma_3} + \|a\|^{\gamma,\mathcal{O}}_{s+\sigma_3}, \tag{5.7.187}$$

for any $\mathtt{b} \le \rho - 2$ *and*

$$\begin{aligned} \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}_+}(s, \mathtt{b}) &\le_{s,\rho} \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s, \mathtt{b}) \\ &+ \big(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\sigma_3} + \|\beta\|_{s+\sigma_3} + \|a\|_{s+\sigma_3} + \mathbb{M}^{\gamma}_{\widehat{\mathcal{Q}}}(s, \mathtt{b})\big)\|\Delta_{12}\beta\|_{s_0+\sigma_3} \\ &+ \mathtt{k}_3(\|\Delta_{12}p\|_{s+\sigma_3} + \|\Delta_{12}p\|_{s_0+\sigma_3}\|p\|_{s+\sigma_3}) + \|\Delta_{12}\beta\|_{s+\sigma_3} + \|\Delta_{12}a\|_{s+\sigma_3} \\ &+ \big(\mathtt{k}_3\|\Delta_{12}p\|_{s_0+\sigma_3} + \|\Delta_{12}a\|_{s_0+\sigma_3} + \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s_0, \mathtt{b})\big)\|\beta\|_{s+\sigma_3} \end{aligned} \tag{5.7.188}$$

for any $\mathtt{b} \le \rho - 3$.

Proof. The strategy is the following.

(1) We conjugate
$$\mathcal{L}^0 := \mathcal{D}_\omega - J \circ (m + a(\varphi, x)) + \mathcal{Q} \tag{5.7.189}$$
by the flow Ψ^τ in (5.7.89).We find a transformation W^τ such that $W^\tau\mathcal{L}^0(W^\tau)^{-1}$ differs from $\Psi^\tau\mathcal{L}^0(\Psi^\tau)^{-1}$ by a remainder, which belong to the class $\mathfrak{L}_\rho$. Then we compute explicitly $W^\tau\mathcal{L}^0(W^\tau)^{-1}$.

(2) The operator $\mathcal{L}^0$ differs from $\mathcal{L}$ by a infinitely regularizing operator of the form (5.6.5). By using this fact and Lemma 5.7.18 we estimate the difference between $\Phi^\tau\mathcal{L}(\Phi^\tau)^{-1}$ and $\Psi^\tau\mathcal{L}^0(\Psi^\tau)^{-1}$.

Step (1). Let Ψ^τ be the flow in (5.7.89).
We can write
$$\Psi^\tau := \mathcal{A}^\tau \circ (\Theta^\tau + R^\tau), \tag{5.7.190}$$
where $\mathcal{A}^\tau$ is defined in (5.7.59), and Θ^τ, R^τ given by Prop. 5.7.16 in (5.7.92) with $R^\tau \in \mathfrak{L}_\rho$. We define the map
$$W^\tau := \mathcal{A}^\tau \circ \Theta^\tau. \tag{5.7.191}$$
We claim that setting $\widehat{R}^\tau = (\Theta^\tau)^{-1}\mathcal{R}^\tau$ we have

$$\begin{aligned} S^\tau := W^\tau\mathcal{L}^0(W^\tau)^{-1} - \Psi^\tau\mathcal{L}^0(\Psi^\tau)^{-1} &= W^\tau\big(\mathcal{L}^0 - (I + \widehat{R}^\tau)\mathcal{L}^0(I + \widehat{R}^\tau)^{-1}\big)(W^\tau)^{-1} \\ &= \mathcal{A}^\tau\Theta^\tau[\mathcal{L}^0, \widehat{R}^\tau](I + \widehat{R}^\tau)^{-1}(\Theta^\tau)^{-1}(\mathcal{A}^\tau)^{-1} \in \mathfrak{L}_\rho, \end{aligned} \tag{5.7.192}$$

and $\sup_{\tau\in[0,1]} \mathbb{M}^{\gamma}_{S^\tau}(s, \mathtt{b}), \sup_{\tau\in[0,1]} \mathbb{M}_{\Delta_{12}S^\tau}(s, \mathtt{b})$ satisfy bounds (5.7.187) and (5.7.188). We first study the conjugation of $\mathcal{L}^0$ by W^τ. In order to prove our claim we just have to note that $\widehat{R}^\tau \in \mathfrak{L}_{\rho+1}$ by Lemma 5.7.7, moreover, by formula (5.7.35) , $[\mathcal{D}_\omega, \widehat{R}^\tau] = \omega \cdot \partial_\varphi\widehat{R}^\tau$ and $[\partial_x, \widehat{R}^\tau] \in \mathfrak{L}_\rho$. This means that $[\mathcal{L}^0, \widehat{R}^\tau] \in \mathfrak{L}_\rho$, so that our claim follows by Lemmata 5.7.5, 5.7.7, 5.7.13 and 5.7.14 .

Conjugation by Θ^τ**.** By Lemma 5.7.13

$$(\Theta^\tau)^{-1} := \mathrm{I} - \mathrm{Op}(\tilde{\vartheta}) + \mathtt{R}_\rho,$$

with
$$|\tilde{\vartheta}|^{\gamma,\mathcal{O}}_{-1,s,\alpha} \le_{s,\alpha,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_0} \tag{5.7.193}$$

and for $s_0 \leq s \leq S_{max}$,

$$\begin{aligned} &\mathbb{M}^{\gamma}_{\mathtt{R}_\rho}(s,\mathtt{b}) \leq_{s,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_0}, \quad \mathtt{b} \leq \rho-2, \\ &\mathbb{M}_{\Delta_{12}\mathtt{R}_\rho}(s,\mathtt{b}) \leq_{s,\rho} \|\Delta_{12}\beta\|_{s+\mathfrak{d}_0} + \|\beta\|_{s+\mathfrak{d}_0}\|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_0} \quad \mathtt{b} \leq \rho-3, \end{aligned} \tag{5.7.194}$$

for some $\mathfrak{d}_0 = \mathfrak{d}_0(\rho)$. Throughout the proof we shall denote by $\mathfrak{d}_i$ an increasing sequence of constants, depending on ρ and possibly on σ_1, which keeps track of the loss pf deriviartives in our procedure. Moreover we shall omit writing the constraints $s_0 \leq s \leq S_{max}$, $\mathtt{b} \leq \rho-2$, $\mathtt{b} \leq \rho-3$ in the bounds on the operators belonging to $\mathfrak{L}_\rho$.

We wish to compute

$$\Theta^\tau B(\Theta^\tau)^{-1} = B + [\Theta^\tau, B](\Theta^\tau)^{-1} = B + [\mathrm{Op}(\vartheta), B]\mathrm{Op}(1-\tilde{\vartheta}) + [\mathrm{Op}(\vartheta), B]\mathtt{R}_\rho \tag{5.7.195}$$

for $B = \mathcal{D}_\omega, J\circ(m+a), \mathrm{Op}(\mathtt{q}), \widehat{\mathcal{Q}}$.
Let us start by studying the commutator $[\mathrm{Op}(\vartheta), B]$, our purpose is to write it as a pseudo differential term plus a remainder in $\mathfrak{L}_\rho$. We have (recalling the Definition 2.2.3 and formula (2.2.22))

$$[\mathrm{Op}(\vartheta), \mathcal{D}_\omega] = -\mathrm{Op}(\mathcal{D}_\omega\vartheta) \tag{5.7.196}$$

$$\begin{aligned} [\mathrm{Op}(\vartheta), J\circ(m+a)] &= \mathrm{Op}\big(\vartheta \star_{<\rho+1} (\omega(\xi)\#_{<\rho+1}(m+a))\big) \\ &\quad + \mathrm{Op}\big(\vartheta \star_{\geq\rho+1} (\omega(\xi)\#(m+a)) + \vartheta \star_{<\rho+1} (\omega(\xi)\#_{\geq\rho+1}(m+a))\big) \end{aligned} \tag{5.7.197}$$

$$[\mathrm{Op}(\vartheta), \mathrm{Op}(\mathtt{q})] = \mathrm{Op}\big(\vartheta \star_{<\rho-1} \mathtt{q}\big) + \mathrm{Op}\big(\vartheta \star_{\geq\rho-1} \mathtt{q})\big)\,. \tag{5.7.198}$$

Here $\omega(\xi)$ is the symbol of the Fourier multiplier $J = \partial_x + 3\Lambda\partial_x$

$$\omega(\xi) := \mathrm{i}\xi + 3\frac{\mathrm{i}\xi}{1+\xi^2}.$$

One can directly verify that all the symbols above are in S^{-1}, indeed the commutator of two pseudo differential operators has as order the sum of the orders minus one. By Lemma 5.7.7 we verify that $[\mathrm{Op}(\vartheta), \widehat{\mathcal{Q}}], [\mathrm{Op}(\vartheta), B]\mathtt{R}_\rho \in \mathfrak{L}_\rho$ for all choices of B. By Lemma 5.7.6 and (2.2.6) we have that the second summands in (5.7.197) and (5.7.198) belong to $\mathfrak{L}_\rho$. We have proved that

$$[\mathrm{Op}(\vartheta), B] = \mathrm{Op}(r_B) + R_B\,, \quad r_B \in S^{-1}\,, \quad R_B \in \mathfrak{L}_\rho\,.$$

Using (5.7.93), (5.7.181) and (5.7.183), we have by (5.7.39)

$$|r_B|^{\gamma,\mathcal{O}}_{-1,s,\alpha} \leq_{s,\alpha,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_1} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_1}(\mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_1} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_1}). \tag{5.7.199}$$

Similarly, by (5.7.40) we have

$$\mathbb{M}^{\gamma}_{R_B}(s,\mathtt{b}) \leq_{s,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_1} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_1}(\mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_1} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_1} + \mathbb{M}^{\gamma}_{\widehat{\mathcal{Q}}}(s,\mathtt{b})). \tag{5.7.200}$$

Analogously by (5.7.41) and (5.7.42) we have

$$\begin{aligned} |\Delta_{12} r_B|_{-1,s,\alpha} \leq_{s,\alpha,\rho} &\|\Delta_{12}\beta\|_{s+\mathfrak{d}_1} + \|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_1}(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\mathfrak{d}_1} + \|a\|_{s+\mathfrak{d}_1}) \\ &+ \|\beta\|_{s_0+\mathfrak{d}_1}(\mathtt{k}_3(\|\Delta_{12}p\|_{s+\mathfrak{d}_1} + \|\Delta_{12}p\|_{s_0+\mathfrak{d}_1}\|p\|_{s+\mathfrak{d}_1}) + \|\Delta_{12}a\|_{s+\mathfrak{d}_1}) \\ &+ \|\beta\|_{s+\mathfrak{d}_1}(\mathtt{k}_3\|\Delta_{12}p\|_{s_0+\mathfrak{d}_1} + \|\Delta_{12}a\|_{s_0+\mathfrak{d}_1}). \end{aligned} \tag{5.7.201}$$

Similarly, by (5.7.40) we have

$$\begin{aligned}\mathbb{M}_{\Delta_{12} R_B}(s, \mathtt{b}) \leq_{s,\rho} & \|\Delta_{12}\beta\|_{s+\mathfrak{d}_1} + \|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_1}(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\mathfrak{d}_1} + \|a\|_{s+\mathfrak{d}_1} + \mathbb{M}^{\gamma}_{\widehat{\mathcal{Q}}}(s,\mathtt{b})) \\ & + \|\beta\|_{s+\mathfrak{d}_1}(\mathtt{k}_3\|\Delta_{12}p\|_{s_0+\mathfrak{d}_1} + \|\Delta_{12}a\|_{s_0+\mathfrak{d}_1} + \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s_0,\mathtt{b})) \\ & + \|\beta\|_{s_0+\mathfrak{d}_1}(\mathtt{k}_3(\|\Delta_{12}p\|_{s+\mathfrak{d}_1} + \|\Delta_{12}p\|_{s_0+\mathfrak{d}_1}\|p\|_{s+\mathfrak{d}_1}) + \|\Delta_{12}a\|_{s+\mathfrak{d}_1} + \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s,\mathtt{b})).\end{aligned} \tag{5.7.202}$$

By Lemmata 5.7.9, 5.7.7 and 5.7.5 we have that

$$[\mathrm{Op}(\vartheta), B]\mathrm{Op}(1-\tilde{\theta}) = \mathrm{Op}(\tilde{r}_B) + \tilde{R}_B\,, \quad \tilde{r}_B \in S^{-1}\,, \quad \tilde{R}_B \in \mathfrak{L}_\rho\,,$$

and $\tilde{r}_B, \tilde{R}_B$ satisfy bounds like (5.7.199)-(5.7.202), with possibly a larger $\mathfrak{d}_1$. Analogously, by Lemmata 5.7.7 and 5.7.5, we have that $[\mathrm{Op}(\theta), B]\mathtt{R}_\rho \in \mathfrak{L}_\rho$ satisfies estimates like (5.7.200), (5.7.202). We conclude that

$$\Theta^\tau \mathcal{L}^0 (\Theta^\tau)^{-1} = \mathcal{L}^0 + \mathrm{Op}(r_0) + \mathcal{R}_0$$

where $r_0 \in S^{-1}$, $\mathcal{R}_0 \in \mathfrak{L}_\rho$ and satisfy the bounds (5.7.199)-(5.7.202) with possibly larger $\mathfrak{d}_1$.

Conjugation by $\mathcal{A}^\tau$. We proved that

$$W^\tau \mathcal{L}^0 (W^\tau)^{-1} = \mathcal{A}^\tau \mathcal{L}^0 (\mathcal{A}^\tau)^{-1} + \mathcal{A}^\tau \mathrm{Op}(r_0)(\mathcal{A}^\tau)^{-1} + \mathcal{A}^\tau \mathcal{R}_0 (\mathcal{A}^\tau)^{-1}. \tag{5.7.203}$$

First we note that

$$\mathcal{D}_\omega (\mathcal{A}^\tau)^{-1} = (\mathcal{A}^\tau)^{-1}\mathcal{D}_\omega + (\mathcal{D}_\omega \tilde{\beta}_x)\mathcal{T}_{\tilde{\beta}} + (1+\tilde{\beta}_x)(\mathcal{D}_\omega\tilde{\beta})\mathcal{T}_{\tilde{\beta}}\partial_x \tag{5.7.204}$$

consequently

$$\begin{aligned}\mathcal{A}^\tau \mathcal{D}_\omega (\mathcal{A}^\tau)^{-1} &= \mathcal{D}_\omega + \partial_x \circ (\mathcal{T}_{\tau\beta}\mathcal{D}_\omega\tilde{\beta}) \\ &= \mathcal{D}_\omega + J \circ (\mathcal{T}_{\tau\beta}\mathcal{D}_\omega\tilde{\beta}) + \mathrm{Op}(r_1) + \mathcal{R}_1\end{aligned} \tag{5.7.205}$$

where $r_1 \in S^{-1}$, $\mathcal{R}_1 \in \mathfrak{L}_\rho$ are defined by

$$r_1 := -3(\mathrm{i}\xi/(1+\xi^2))\#_{<\rho-1}\mathcal{T}_{\tau\beta}(\mathcal{D}_\omega\tilde{\beta}), \quad \mathcal{R}_1 := -3\mathrm{Op}((\mathrm{i}\xi/(1+\xi^2))\#_{\geq\rho-1}\mathcal{T}_{\tau\beta}(\mathcal{D}_\omega\tilde{\beta})), \tag{5.7.206}$$

and, by (5.7.39),(5.7.41), (5.7.40), (5.7.42), satisfy the following bounds

$$\begin{aligned}& |r_1|^{\gamma,\mathcal{O}}_{-1,s,\alpha} \leq_{s,\alpha,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_2}, \\ & |\Delta_{12} r_1|_{-1,s,\alpha} \leq_{s,\alpha,\rho} \|\Delta_{12}\beta\|_{s+\mathfrak{d}_2} + \|\beta\|_{s+\mathfrak{d}_2}\|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_2} \\ & \mathbb{M}^\gamma_{\mathcal{R}_1}(s,\mathtt{b}) \leq_{s,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_2}, \\ & \mathbb{M}_{\Delta_{12}\mathcal{R}_1}(s,\mathtt{b}) \leq_{s,\rho} \|\Delta_{12}\beta\|_{s+\mathfrak{d}_2} + \|\beta\|_{s+\mathfrak{d}_2}\|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_2}.\end{aligned} \tag{5.7.207}$$

Moreover

$$\mathcal{A}^\tau (J \circ (m+a))(\mathcal{A}^\tau)^{-1} = J \circ \mathcal{T}_{\tau\beta}\Big((1+\tilde{\beta}_x)(m+a)\Big) + \mathsf{R}^{(2)} \tag{5.7.208}$$

where

$$\begin{aligned}\mathsf{R}^{(2)} := & \Big((1-\Lambda\mathfrak{R})^{-1} - 1\Big) \circ \Lambda \circ g \circ \partial_x \circ \mathcal{T}_{\tau\beta}\Big((1+\tilde{\beta}_x)(m+a)\Big) \\ & + \Big((1-\Lambda\mathfrak{R})^{-1} - 1\Big) \circ \Lambda \circ (g-3) \circ \partial_x \circ \mathcal{T}_{\tau\beta}\Big((1+\tilde{\beta}_x)(m+a)\Big) \\ & + \Big((1-\Lambda\mathfrak{R})^{-1}\Big) \circ \Lambda \circ (g-3) \circ \partial_x \circ \mathcal{T}_{\tau\beta}\Big((1+\tilde{\beta}_x)(m+a)\Big)\end{aligned} \tag{5.7.209}$$

with g and $\mathfrak{R}$ defined in (5.7.100). In particular $\mathsf{R}^{(2)} = \mathrm{Op}(r_2) + \mathcal{R}_2$, $r_2 \in S^{-1}$, $\mathcal{R}_2 \in \mathfrak{L}_\rho$ and satisfy the following bounds

$$
\begin{aligned}
|r_2|^{\gamma,\mathcal{O}}_{-1,s,\alpha} &\leq_{s,\alpha,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_3} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_3}\|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_3}, \\
|\Delta_{12} r_2|_{-1,s,\alpha} &\leq_{s,\alpha,\rho} \|\Delta_{12}\beta\|_{s+\mathfrak{d}_3} + \|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_3}(\|a\|_{s+\mathfrak{d}_3} + \|\beta\|_{s+\mathfrak{d}_3}) \\
&\quad + \|\Delta_{12} a\|_{s_0+\mathfrak{d}_3}\|\beta\|_{s+\mathfrak{d}_3} + \|\Delta_{12} a\|_{s+\mathfrak{d}_3}\|\beta\|_{s_0+\mathfrak{d}_3} \\
\mathbb{M}^\gamma_{\mathcal{R}_2}(s,\mathtt{b}) &\leq_{s,\rho} \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_3} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_3}\|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_3}, \\
\mathbb{M}_{\Delta_{12}\mathcal{R}_2}(s,\mathtt{b}) &\leq_{s,\rho} \|\Delta_{12}\beta\|_{s+\mathfrak{d}_3} + \|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_3}(\|a\|_{s+\mathfrak{d}_3} + \|\beta\|_{s+\mathfrak{d}_3}) \\
&\quad + \|\Delta_{12} a\|_{s_0+\mathfrak{d}_3}\|\beta\|_{s+\mathfrak{d}_3} + \|\Delta_{12} a\|_{s+\mathfrak{d}_3}\|\beta\|_{s_0+\mathfrak{d}_3}.
\end{aligned}
\tag{5.7.210}
$$

Then

$$
W^\tau \mathcal{L}^0 (W^\tau)^{-1} = \mathcal{D}_\omega - J \circ (m + a_+) + \mathcal{Q}_*, \tag{5.7.211}
$$

where by (5.7.203)

$$
\begin{aligned}
\mathcal{Q}_* :=\ & \mathcal{A}^\tau \mathrm{Op}(\mathtt{q} + r_0)(\mathcal{A}^\tau)^{-1} + \mathcal{A}^\tau(\widehat{\mathcal{Q}} + \mathcal{R}_0)(\mathcal{A}^\tau)^{-1} \\
& + \mathrm{Op}(r_1 + r_2) + \mathcal{R}_1 + \mathcal{R}_2.
\end{aligned}
\tag{5.7.212}
$$

By Theorem 5.7.20 and Lemma 5.7.14 we have

$$
\mathcal{A}^\tau \mathrm{Op}(\mathtt{q} + r_0)(\mathcal{A}^\tau)^{-1} = \mathrm{Op}(r_3) + \mathcal{R}_3, \quad \mathcal{A}^\tau(\widehat{\mathcal{Q}} + \mathcal{R}_0)(\mathcal{A}^\tau)^{-1} = \mathcal{R}_4 \tag{5.7.213}
$$

where $r_3 \in S^{-1}$ and $\mathcal{R}_3$, $\mathcal{R}_4 \in \mathfrak{L}_\rho$. In order to bound r_3 we use (5.7.150) with $w = \mathtt{q} + r_0$ so that

$$
|w|^{\gamma,\mathcal{O}}_{-1,s,\alpha} \leq_{s,\alpha,\rho} \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_4} + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_4} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_4}. \tag{5.7.214}
$$

Note that in the formula (5.7.150) (recall the notations used in formula (5.7.150) and the fact that s_1, s_2, $s_3 \geq 0$ and $s_1 + s_2 + s_3 = s$) we have by interpolation

$$
\begin{aligned}
& |w|^{\gamma,\mathcal{O}}_{-1,s_1,\alpha+s_2+\sigma_1}\|\beta\|^{\gamma,\mathcal{O}}_{s_3+\sigma_1} \leq (\mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s_1+\mathfrak{d}_4} + \|\beta\|^{\gamma,\mathcal{O}}_{s_1+\mathfrak{d}_4} + \|a\|^{\gamma,\mathcal{O}}_{s_1+\mathfrak{d}_4})\|\beta\|^{\gamma,\mathcal{O}}_{s_3+\sigma_1} \\
& \leq_s (\mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5})\|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_5} + \|\beta\|_{s+\mathfrak{d}_5}(\mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_5} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_5} + \|a\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_5}).
\end{aligned}
$$

Thus we get by (5.7.183)

$$
\begin{aligned}
|r_3|^{\gamma,\mathcal{O}}_{-1,s,\alpha} &\leq_{s,\alpha,\rho} \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5}, \\
|\Delta_{12} r_3|_{-1,s,\alpha} &\leq_{s,\alpha,\rho} \big(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\mathfrak{d}_5} + \|\beta\|_{s+\mathfrak{d}_5} + \|a\|_{s+\mathfrak{d}_5}\big)\|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_5} \\
&\quad + \mathtt{k}_3(\|\Delta_{12} p\|_{s+\mathfrak{d}_5} + \|\Delta_{12} p\|_{s_0+\mathfrak{d}_5}\|p\|_{s+\mathfrak{d}_5}) + \|\Delta_{12}\beta\|_{s+\mathfrak{d}_5} + \|\Delta_{12} a\|_{s+\mathfrak{d}_5} \\
&\quad + \big(\mathtt{k}_3\|\Delta_{12} p\|_{s_0+\mathfrak{d}_5} + \|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_5} + \|\Delta_{12} a\|_{s_0+\mathfrak{d}_5}\big)\|\beta\|_{s+\mathfrak{d}_5}, \\
\mathbb{M}^\gamma_{\mathcal{R}_3}(s,\mathtt{b}) &\leq_{s,\rho} \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5}, \\
\mathbb{M}_{\Delta_{12}\mathcal{R}_3}(s,\mathtt{b}) &\leq_{s,\rho} \big(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\mathfrak{d}_5} + \|\beta\|_{s+\mathfrak{d}_5} + \|a\|_{s+\mathfrak{d}_5}\big)\|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_5} \\
&\quad + \mathtt{k}_3(\|\Delta_{12} p\|_{s+\mathfrak{d}_5} + \|\Delta_{12} p\|_{s_0+\mathfrak{d}_5}\|p\|_{s+\mathfrak{d}_5}) + \|\Delta_{12}\beta\|_{s+\mathfrak{d}_5} + \|\Delta_{12} a\|_{s+\mathfrak{d}_5} \\
&\quad + \big(\mathtt{k}_3\|\Delta_{12} p\|_{s_0+\mathfrak{d}_5} + \|\Delta_{12} a\|_{s_0+\mathfrak{d}_5}\big)\|\beta\|_{s+\mathfrak{d}_5}.
\end{aligned}
\tag{5.7.215}
$$

Moreover by (5.7.183)

$$
\begin{aligned}
\mathbb{M}^\gamma_{\mathcal{R}_4}(s,\mathtt{b}) &\leq_{s,\rho} \mathbb{M}^\gamma_{\widehat{\mathcal{Q}}}(s,\mathtt{b}) + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_6} + \|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mathfrak{d}_6}(\mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_6} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_6}),\\
\mathbb{M}_{\Delta_{12}\mathcal{R}_4}(s,\mathtt{b}) &\leq_{s,\rho} \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s,\mathtt{b}) + \|\Delta_{12}\beta\|_{s+\mathfrak{d}_6}\\
&\quad + \|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_6}(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\mathfrak{d}_6} + \|a\|_{s+\mathfrak{d}_6} + \|\beta\|_{s+\mathfrak{d}_6} + \mathbb{M}^\gamma_{\widehat{\mathcal{Q}}}(s,\mathtt{b}))\\
&\quad + \|\beta\|_{s_0+\mathfrak{d}_6}(\mathtt{k}_3(\|\Delta_{12}p\|_{s+\mathfrak{d}_6} + \|\Delta_{12}p\|_{s_0+\mathfrak{d}_6}\|p\|_{s+\mathfrak{d}_6}) + \|\Delta_{12}a\|_{s+\mathfrak{d}_6})\\
&\quad + \|\beta\|_{s+\mathfrak{d}_6}(\mathtt{k}_3\|\Delta_{12}p\|_{s_0+\mathfrak{d}_6} + \|\Delta_{12}a\|_{s_0+\mathfrak{d}_6} + \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s_0,\mathtt{b}))
\end{aligned}
\tag{5.7.216}
$$

By (5.7.212) and (5.7.213) $\mathcal{Q}_*$ in (5.7.211) is

$$
\mathcal{Q}_* = \mathrm{Op}(\mathtt{q}_+) + \widehat{\mathcal{Q}}_*, \quad \mathtt{q}_+ := r_1 + r_2 + r_3, \quad \widehat{\mathcal{Q}}_* := \mathcal{R}_1 + \mathcal{R}_2 + \mathcal{R}_3 + \mathcal{R}_4 \tag{5.7.217}
$$

with the following bounds

$$
\begin{aligned}
|\mathtt{q}_+|^{\gamma,\mathcal{O}}_{-1,s,\alpha} &\leq_{s,\alpha,\rho} \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_5},\\
|\Delta_{12}\mathtt{q}_+|_{-1,s,\alpha} &\leq_{s,\alpha,\rho} \big(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\mathfrak{d}_5} + \|\beta\|_{s+\mathfrak{d}_5} + \|a\|_{s+\mathfrak{d}_5}\big)\|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_5}\\
&\quad + \mathtt{k}_3(\|\Delta_{12}p\|_{s+\mathfrak{d}_5} + \|\Delta_{12}p\|_{s_0+\mathfrak{d}_5}\|p\|_{s+\mathfrak{d}_5}) + \|\Delta_{12}\beta\|_{s+\mathfrak{d}_5} + \|\Delta_{12}a\|_{s+\mathfrak{d}_5}\\
&\quad + \big(\mathtt{k}_3\|\Delta_{12}p\|_{s_0+\mathfrak{d}_5} + \|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_5} + \|\Delta_{12}a\|_{s_0+\mathfrak{d}_5}\big)\|\beta\|_{s+\mathfrak{d}_5}
\end{aligned}
\tag{5.7.218}
$$

and

$$
\begin{aligned}
\mathbb{M}^\gamma_{\widehat{\mathcal{Q}}_*}(s,\mathtt{b}) &\leq_{s,\rho} \mathbb{M}^\gamma_{\widehat{\mathcal{Q}}}(s,\mathtt{b}) + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_6} + \mathtt{k}_1 + \mathtt{k}_2\|p\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_6} + \|a\|^{\gamma,\mathcal{O}}_{s+\mathfrak{d}_6},\\
\mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}_*}(s,\mathtt{b}) &\leq_{s,\rho} \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s,\mathtt{b})\\
&\quad + \big(\mathtt{k}_1 + \mathtt{k}_2\|p\|_{s+\mathfrak{d}_6} + \|\beta\|_{s+\mathfrak{d}_6} + \|a\|_{s+\mathfrak{d}_6} + \mathbb{M}^\gamma_{\widehat{\mathcal{Q}}}(s,\mathtt{b})\big)\|\Delta_{12}\beta\|_{s_0+\mathfrak{d}_6}\\
&\quad + \mathtt{k}_3(\|\Delta_{12}p\|_{s+\mathfrak{d}_6} + \|\Delta_{12}p\|_{s_0+\mathfrak{d}_6}\|p\|_{s+\mathfrak{d}_6}) + \|\Delta_{12}\beta\|_{s+\mathfrak{d}_6} + \|\Delta_{12}a\|_{s+\mathfrak{d}_6}\\
&\quad + \big(\mathtt{k}_3\|\Delta_{12}p\|_{s_0+\mathfrak{d}_6} + \|\Delta_{12}a\|_{s_0+\mathfrak{d}_6} + \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}}(s_0,\mathtt{b})\big)\|\beta\|_{s+\mathfrak{d}_6}.
\end{aligned}
\tag{5.7.219}
$$

Step (2). If $y := x + \beta(\varphi, x)$ then we have

$$
\begin{aligned}
\mathcal{L}^0_+ h &:= \Psi\mathcal{L}^0\Psi^{-1} = \mathcal{D}_\omega h - J\{\big(m + a_+(\varphi,x)\big)\, h\} + \mathrm{Op}(\mathtt{q}_+) + \widehat{\mathcal{Q}}_\star h,\\
m + a_+(\varphi,x) &:= -\mathcal{D}_\omega\tilde\beta(\varphi,y) + (m + a(\varphi,y))(1 + \tilde\beta_x(\varphi,y)).
\end{aligned}
$$

By Lemma 5.7.18 we have (recall (B.0.6), (B.0.7))

$$
\begin{aligned}
\widehat{\mathcal{Q}}_{\star\star} := \Phi\mathcal{L}\Phi^{-1} - \Pi_S^\perp\mathcal{L}^0_+\Pi_S^\perp &= \Pi_S^\perp\Psi\mathcal{L}^0\Psi^{-1}\Pi_S + \Pi_S\Psi\mathcal{L}^0\Psi^{-1} + \Psi\mathcal{L}\Gamma\Psi^{-1} + \Psi\mathcal{L}\mathcal{R}\Psi^{-1}\\
&\quad + \Psi\mathcal{R}\mathcal{L}\Gamma\Psi^{-1} - \Psi\mathcal{L}^0\Pi_S\Psi^{-1} - \Psi\Pi_S\mathcal{L}\Pi_S^\perp\Psi^{-1}.
\end{aligned}
\tag{5.7.220}
$$

We define the remainder

$$
\widehat{\mathcal{Q}}_+ := \widehat{\mathcal{Q}}_\star + \widehat{\mathcal{Q}}_{\star\star}.
$$

To conclude the proof we show that $\widehat{\mathcal{Q}}_{\star\star}$ satisfies the bounds (5.7.187) and (5.7.188).

We note that

$$
\begin{aligned}
\Pi_S^\perp \Psi \mathcal{L}^0 \Psi^{-1} \Pi_S h &= \sum_{j\in S} (h, g_j^{(1)})_{L^2} \chi_j^{(1)}, \quad g_j^{(1)} := e^{\mathrm{i}jx}, \quad \chi_j^{(1)} := \Psi \mathcal{L}^0 \Psi^{-1} e^{\mathrm{i}jx}, \\
\Pi_S \Psi \mathcal{L}^0 \Psi^{-1} h &= \sum_{j\in S} (h, g_j^{(2)})_{L^2} \chi_j^{(2)}, \quad g_j^{(2)} := \Psi \mathcal{L}^0 \Psi^{-1} e^{\mathrm{i}jx}, \quad \chi_j^{(2)} = e^{\mathrm{i}jx}, \\
\Psi \mathcal{L}^0 \Pi_S \Psi^{-1} &= \sum_{j\in S} (h, g_j^{(3)})_{L^2} \chi_j^{(3)}, \quad g_j^{(3)} := (\Psi^{-1})^* e^{\mathrm{i}jx}, \quad \chi_j^{(3)} := \Psi \mathcal{L}^0 e^{\mathrm{i}jx}, \\
\Psi \Pi_S \mathcal{L} \Pi_S^\perp \Psi^{-1} h &= \sum_{j\in S} (h, g_j^{(4)})_{L^2} \chi_j^{(4)}, \quad g_j^{(4)} := (\mathcal{L} \Pi_S^\perp \Psi^{-1})^* e^{\mathrm{i}jx}, \quad \chi_j^{(4)} := \Psi e^{\mathrm{i}jx}.
\end{aligned}
\tag{5.7.221}
$$

Thus by Lemma 5.6.2, Corollary 5.7.17 (for the estimates on Ψ) and (5.7.189) we get the bounds (5.7.187) and (5.7.188) for the operators (5.7.221).
The bounds on $\Psi \mathcal{L} \Gamma \Psi^{-1}, \Psi \mathcal{L} \mathcal{R} \Psi^{-1}, \Psi \mathcal{R} \mathcal{L} \Gamma \Psi^{-1}$ follow by Proposition 5.7.16, (5.7.144), (5.7.145).

□

5.7.3 Straightening theorem

By Proposition 5.7.21 the coefficient a_+ of the transformed operator $\mathcal{L}_+ = \Phi \mathcal{L} \Phi^{-1}$ is given by (5.7.185). The aim of this section is to find a function β (see (5.7.88)), or equivalently a flow Φ of (5.7.143), such that a_+ is a constant, namely such that the following equation is solved (recall (5.7.59))

$$
\mathcal{D}_\omega \tilde{\beta} - (m + a)(1 + \tilde{\beta}_x) = \text{constant}. \tag{5.7.222}
$$

This issue is tantamount to finding a change of coordinates that straightens the 1-order vector field

$$
\omega \cdot \frac{\partial}{\partial \varphi} - (m + a(\varphi, x)) \frac{\partial}{\partial x}.
$$

This is the content of the following proposition. Actually this is a classical result on vector fields on a torus (see for instance [2], [82]), adapted to our purposes. Let $\mathcal{O}_0$ be a compact set in Ω_ε (see (5.4.2)). Recall $s_0 = [\nu/2] + 3$ and fix $\tau = \nu + 2$.
We use the notation $\|u\|_s^\gamma := \|u\|_s^{\gamma, \Omega_\varepsilon}$ and $|m|^\gamma := |m|^{\gamma, \Omega_\varepsilon}$, $|m|^{lip} := |m|^{lip, \Omega_\varepsilon}$ (recall (2.1.9), (2.1.10) and (5.4.2)).

Proposition 5.7.22. *Consider for $\omega \in \mathcal{O}_0 \subseteq \Omega_\varepsilon$ a Lipschitz family of vector fields on $\mathbb{T}^{\nu+1}$*

$$
\begin{aligned}
&X_0 := \omega \cdot \frac{\partial}{\partial \varphi} - (m_0 + a_0(x, \varphi; \omega)) \frac{\partial}{\partial x}, \quad \frac{2}{3} < m_0 < \frac{3}{2}, |m_0|^{lip} \le M_0 < 1/2 \\
&a_0 \in H^s(\mathbb{T}^{\nu+1}, \mathbb{R}) \quad \forall s \ge s_0.
\end{aligned}
\tag{5.7.223}
$$

Moreover $a(x, \varphi; \omega) = a(x, \varphi, i(\omega); \omega)$ and it is Lipschitz in the variable i. There exists $\delta_\star = \delta_\star(s_1) > 0$ and $s_1 \ge s_0 + 2\tau + 4$ such that, for any $\gamma > 0$ if

$$
C(s_1) \gamma^{-1} \|a_0\|_{s_1}^{\gamma, \mathcal{O}_0} := \delta \le \delta_\star \tag{5.7.224}
$$

then there exists a Lipschitz function $m_\infty(\omega) = m_\infty(\omega, i(\omega))$ with

$$
\frac{1}{2} < m_\infty < 2, \quad \forall \omega \in \Omega_\varepsilon, \qquad |m_\infty - m_0|^\gamma \le \gamma \delta, \tag{5.7.225}
$$

such that in the set

$$\mathcal{O}^{2\gamma}_{\infty} := \{\omega \in \mathcal{O}_0 : \ |\omega \cdot \ell - m_\infty(\omega) j| > \frac{2\gamma}{\langle \ell \rangle^\tau}\,,\ \forall \ell \in \mathbb{Z}^\nu,\ \forall j \in \mathbb{Z} \setminus \{0\}\} \tag{5.7.226}$$

the following holds. For all $\omega \in \mathcal{O}^{2\gamma}_\infty$

$$|\Delta_{12} m_\infty| \le 2 |\Delta_{12} \langle a_0 \rangle|$$

and there exists a map

$$\beta^{(\infty)} : \mathcal{O}_0 \times \mathbb{T}^{\nu+1} \to \mathbb{R}\,, \quad \|\beta^{(\infty)}\|^{\gamma,\mathcal{O}_0}_s \le_s \gamma^{-1} \|a_0\|^{\gamma,\mathcal{O}_0}_{s+2\tau+4}, \quad \forall s \ge s_0 \tag{5.7.227}$$

so that $\Psi^{(\infty)} : (\varphi, x) \mapsto (\varphi, x + \beta^{(\infty)}(\varphi, x))$ *is a diffeomorphism of* $\mathbb{T}^{\nu+1}$ *and for all* $\omega \in \mathcal{O}^{2\gamma}_\infty$

$$\Psi^{(\infty)}_* X_0 := \omega \cdot \frac{\partial}{\partial \varphi} + (\Psi^{(\infty)})^{-1} \big(\omega \cdot \partial_\varphi \beta^{(\infty)} - (m_0 + a_0)(1 + \beta^{(\infty)}_x)\big) \frac{\partial}{\partial x} = \omega \cdot \frac{\partial}{\partial \varphi} - m_\infty(\omega) \frac{\partial}{\partial x}. \tag{5.7.228}$$

In order to prove Proposition 5.7.22 we apply recursively a KAM step which we now describe.
KAM step. Consider for $\omega \in \mathcal{O} \subseteq \mathcal{O}_0$ a Lipschitz family of vector fields on $\mathbb{T}^{\nu+1}$

$$\begin{aligned} & X := \omega \cdot \frac{\partial}{\partial \varphi} - (m + a(\varphi, x; \omega)) \frac{\partial}{\partial x}\,, \quad \frac{1}{2} < m < 2\,, \quad |m|^{\text{lip}} \le M < 1\,, \\ & a(\cdot, \cdot; \omega) \in H^s(\mathbb{T}^{\nu+1}, \mathbb{R}) \quad \forall s \ge s_0. \end{aligned} \tag{5.7.229}$$

The constant m and the function a in (5.7.229) depend on the variable i.
Fix $s_0 = [\nu/2] + 3$, $\tau = \nu + 2$. Given $K \gg 1$ and $\gamma > 0$ assume that for some domain $\mathcal{O} \subseteq \mathcal{O}_0$ we have

$$C(s_0) \gamma^{-1} K^{2\tau+4} \|a\|^{\gamma,\mathcal{O}}_{s_0+1} < 1. \tag{5.7.230}$$

Let

$$\mathcal{O}_+ \equiv \mathcal{C}_{\gamma,m,K,\mathcal{O}} := \{\omega \in \mathcal{O} : \ |\omega \cdot \ell - m(\omega) j| > \frac{\gamma}{\langle \ell \rangle^\tau}\,,\ \forall \ell \in \mathbb{Z}^\nu : \ |\ell| \le K,\ \forall j \in \mathbb{Z} \setminus \{0\}\}\,, \tag{5.7.231}$$

and for all $\omega \in \mathcal{O}_0$ set $\alpha(\varphi, x; \omega)$ to be

$$\alpha(x, \varphi, \omega) := \sum_{|\ell| \le K} \alpha_{\ell,j} e^{\mathtt{i}(\ell \cdot \varphi + j x)}\,, \quad \alpha_{\ell,j} = \frac{a_{\ell,j}\, \chi_{\ell,\gamma}(\omega \cdot \ell - m\, j)}{\mathtt{i}(\omega \cdot \ell - m\, j)} \tag{5.7.232}$$

where $y \mapsto \chi_{\ell,\gamma}(y)$ is a smooth function defined on $\mathbb{R}$ such that $0 \le \chi_{\ell,\gamma}(y) \le 1$ and

$$\chi_{\ell,\gamma}(y) = \begin{cases} 0 & \text{if} \quad |y| \le \dfrac{\gamma}{2\langle \ell \rangle^\tau} \\ 1 & \text{if} \quad |y| \ge \dfrac{\gamma}{\langle \ell \rangle^\tau}. \end{cases} \tag{5.7.233}$$

Lemma 5.7.23. *We have*

$$\|\alpha\|^{\gamma,\mathcal{O}_0}_s \le \gamma^{-1} \|\Pi_K a\|^{\gamma,\mathcal{O}}_{s+2\tau+1} \le \gamma^{-1} K^{2\tau+1} \|a\|^{\gamma,\mathcal{O}}_s, \tag{5.7.234}$$

moreover, for all $\omega \in \mathcal{O}_+$

$$\begin{aligned} &\|\Delta_{12}\alpha\|_s \leq_s \gamma^{-1}\Big(\|\Pi_K\Delta_{12}a\|_{s+\tau} + \gamma^{-1}|\Delta_{12}m|\|\Pi_K a\|_{s+2\tau+1}\Big), \\ &|\Delta_{12}\langle\alpha\rangle| \leq \gamma^{-1}\Big(|\Delta_{12}\langle a\rangle| + \gamma^{-1}|\Delta_{12}m|\|a\|_{s_0}\Big), \end{aligned} \tag{5.7.235}$$

so that the map

$$\Phi : (\varphi, x) \mapsto (\varphi, x + \alpha(\varphi, x))$$

is a diffeomorphism of $\mathbb{T}^{\nu+1}$*. We may set*

$$\Phi_* X := \omega \cdot \frac{\partial}{\partial\varphi} - (m_+ + a_+(x, \varphi; \omega))\frac{\partial}{\partial x} \tag{5.7.236}$$

with m_+ *defined for* $\omega \in \Omega_\varepsilon$ *and Lipschitz with the bounds*

$$\begin{aligned} &|m| - \|a\|_{s_0}^{\gamma,\mathcal{O}} \leq |m_+| \leq |m| + \|a\|_{s_0}^{\gamma,\mathcal{O}}, \quad \textit{for all } \omega \in \Omega_\varepsilon, \\ &|m_+|^{lip} \leq M + \|a\|_{s_0}^{\gamma,\mathcal{O}}, \\ &|\Delta_{12}m_+| \leq |\Delta_{12}m| + \|\Delta_{12}a\|_{s_0} \quad \textit{for all } \omega \in \mathcal{O}_+. \end{aligned} \tag{5.7.237}$$

The function a_+ *is defined and Lipschitz for all* $\omega \in \mathcal{O}_+$ *(see* (5.7.231)*) and satisfies the bounds*

$$\begin{aligned} &\|a_+\|_{s_0}^{\gamma,\mathcal{O}_+} \leq C_{s_0}(K^{s_0-s_1}\|a\|_{s_1}^{\gamma,\mathcal{O}} + C_{s_0}\gamma^{-1}K^{2\tau+2}(\|a\|_{s_0}^{\gamma,\mathcal{O}})^2)(1 + \gamma^{-1}K^{2\tau+1}\|a\|_{s_0}^{\gamma,\mathcal{O}}) \\ &\|a_+\|_s^{\gamma,\mathcal{O}_+} \leq \|a\|_s^{\gamma,\mathcal{O}} + C_s\gamma^{-1}K^{2\tau+3}\|a\|_{s_0}^{\gamma,\mathcal{O}}\|a\|_s^{\gamma,\mathcal{O}} \end{aligned} \tag{5.7.238}$$

for some constants $C_{s_0}, C_s > 0$*. Moreover for* $p = s_0, s_0 + 1$ *the following estimates hold, for all* $\omega \in \mathcal{O}_+$

$$\begin{aligned} &\|\Delta_{12}a_+\|_p \leq \|\Pi_K^\perp\Delta_{12}a\|_p + \|a\|_{p+1}\gamma^{-1}K^{2\tau+2}\left(\|\Delta_{12}a\|_p + \gamma^{-1}|\Delta_{12}m|\|a\|_p\right), \\ &|\Delta_{12}\langle a_+\rangle| \leq \|a\|_{s_0}\gamma^{-1}\left(|\langle\Delta_{12}a\rangle| + \gamma^{-1}|\Delta_{12}m|\|a\|_{s_0}\right). \end{aligned} \tag{5.7.239}$$

Proof. By definition of $\|\cdot\|_s$, (5.7.232) and (5.7.231) we have $\|\alpha\|_s \leq_s \gamma^{-1}K^\tau\|a\|_s$ for all $\omega \in \mathcal{O}_+$. By construction

$$|\Delta_{\omega,\omega'}\alpha_{\ell,j}| \leq \frac{|\chi_{\ell,\gamma}||\Delta_{\omega,\omega'}a_{\ell,j}|}{|\omega\cdot\ell - m\,j|} + \frac{|a_{\ell,j}||\Delta_{\omega,\omega'}\chi_{\ell,\gamma}|}{|\omega\cdot\ell - m\,j|} + \frac{|a_{\ell,j}|(|\ell| + |\Delta_{\omega,\omega'}m|)}{|\omega\cdot\ell - m\,j|^2} \tag{5.7.240}$$

hence by the fact that $|\chi_{\ell,\gamma}| \leq 1$ for all ℓ, j and by (5.7.231), $K > M$ (see (5.7.229)) we have

$$\|\alpha\|_s^{\gamma,\mathcal{O}} \leq_s \gamma^{-1}K^{2\tau+1}\|a\|_s^{\gamma,\mathcal{O}}. \tag{5.7.241}$$

Similarly one can prove (5.7.235) by using the following expression

$$\Delta_{12}\alpha_{\ell,j} = \frac{(\Delta_{12}a_{\ell,j})\,\chi_{\ell,\gamma}(\omega\cdot\ell - m(i_1)j) + a_{\ell,j}(i_2)(\Delta_{12}\chi_{\ell,\gamma})}{\mathrm{i}(\omega\cdot\ell - m(i_1)j)} + \frac{a_{\ell,j}(i_2)\chi_{\ell,\gamma}(i_2)\mathrm{i}(\Delta_{12}m)j}{(\omega\cdot\ell - m(i_1)j)(\omega\cdot\ell - m(i_2)j)}$$

and the fact that the derivative of $\chi_{\ell,\gamma}$ is bounded. We claim that α satisfies the hypotheses of Lemma A.0.5, hence Φ is a diffeomorphism. Indeed, since $s_0 = [\nu/2] + 3$, by (5.7.241) and (5.7.230) we have

$$|\alpha|_{1,\infty}^{\gamma,\mathcal{O}_0} \leq C_{s_0}\|\alpha\|_{s_0}^{\gamma,\mathcal{O}_0} \leq C_{s_0}\gamma^{-1}K^{2\tau+1}\|a\|_{s_0}^{\gamma,\mathcal{O}} \leq \frac{1}{2}$$

if $C(s_0)$ in (5.7.230) is sufficiently large. By definition of pushforward

$$\Phi_* X := \omega \cdot \frac{\partial}{\partial \varphi} + \Phi^{-1}\big(\mathcal{D}_\omega \alpha - (m+a)(1+\alpha_x)\big)\frac{\partial}{\partial x}\,,$$

and by (5.7.232)

$$(m+a)(1+\alpha_x) - \mathcal{D}_\omega \alpha = m + \langle a(\varphi,x)\rangle_{\mathbb{T}^{\nu+1}} + \Pi_K^{\perp} a(\varphi,x) + a(\varphi,x)\alpha_x(\varphi,x).$$

Now we extend $\langle a(\varphi,x)\rangle_{\mathbb{T}^{\nu+1}}$ from $\mathcal{O}$ to the whole Ω_ε by Kirtzbraun theorem, preserving the Lipschitz norm. We set

$$m_+ := m + \langle a(\varphi,x)\rangle^{\mathrm{Ext}}_{\mathbb{T}^{\nu+1}}\,, \quad a_+(\varphi,x) := \Phi^{-1}(\Pi_K^{\perp} a(\varphi,x) + a(\varphi,x)\alpha_x(\varphi,x)), \tag{5.7.242}$$

and immediately one can check that (5.7.237) hold, and, if $x \mapsto x + \tilde{\alpha}(\varphi,x)$ is the inverse diffeomorphism of $x \mapsto x + \alpha(\varphi,x)$, we have

$$\begin{aligned} \Delta_{12} m_+ &= \Delta_{12} m + \langle \Delta_{12} a\rangle_{\mathbb{T}^{\nu+1}}, \\ \Delta_{12} a_+ &= \Pi_K^{\perp}[\Phi^{-1}(\Delta_{12} a)] + \Pi_K^{\perp}[\Phi^{-1}(a_x)\,(\Delta_{12}\tilde{\alpha})] + \Phi^{-1}(\Delta_{12} a\,\alpha_x) + \Phi^{-1}(a\,\Delta_{12}\alpha_x) \\ &\quad + (\Delta_{12}\tilde{\alpha})\,\Phi^{-1}\big(\partial_x(a\,\alpha_x)\big). \end{aligned} \tag{5.7.243}$$

The bounds follow by repeated use of Lemma A.0.5, indeed setting

$$f := \Pi_K^{\perp} a(\varphi,x) + a(\varphi,x)\alpha_x(\varphi,x) \tag{5.7.244}$$

we have by (A.0.8)

$$\begin{aligned} \|a_+\|_s^{\gamma,\mathcal{O}_+} &\le \|f\|_s^{\gamma,\mathcal{O}_+} + C_s\big(\|f\|_s^{\gamma,\mathcal{O}_+}\|\alpha\|_{s_0+1}^{\gamma,\mathcal{O}} + \|\alpha\|_{s+1}^{\gamma,\mathcal{O}}\|f\|_{s_0}^{\gamma,\mathcal{O}_+}\big), \\ \|f\|_s^{\gamma,\mathcal{O}_+} &\le \|\Pi_K^{\perp} a(\varphi,x)\|_s^{\gamma,\mathcal{O}} + C_s(\|a\|_{s_0}^{\gamma,\mathcal{O}}\|\alpha\|_{s+1}^{\gamma,\mathcal{O}} + \|a\|_s^{\gamma,\mathcal{O}}\|\alpha\|_{s_0+1}^{\gamma,\mathcal{O}}). \end{aligned} \tag{5.7.245}$$

Then if $s = s_0$ by applying the smoothing estimates in the second inequality in (5.7.245) we get

$$\begin{aligned} \|f\|_{s_0}^{\gamma,\mathcal{O}_+} &\le K^{s_0-s_1}\|a\|_{s_1}^{\gamma,\mathcal{O}} + C_{s_0}K\|a\|_{s_0}^{\gamma,\mathcal{O}}\|\alpha\|_{s_0}^{\gamma,\mathcal{O}} \\ &\overset{(5.7.241)}{\le} K^{s_0-s_1}\|a\|_{s_1}^{\gamma,\mathcal{O}} + \gamma^{-1}C_{s_0}K^{2\tau+2}(\|a\|_{s_0}^{\gamma,\mathcal{O}})^2, \\ \|a_+\|_{s_0}^{\gamma,\mathcal{O}_+} &\le (K^{s_0-s_1}\|a\|_{s_1}^{\gamma,\mathcal{O}} + C_{s_0}\gamma^{-1}K^{2\tau+2}(\|a\|_{s_0}^{\gamma,\mathcal{O}})^2)(1 + C_{s_0}\gamma^{-1}K^{2\tau+2}\|a\|_{s_0}^{\gamma,\mathcal{O}}). \end{aligned} \tag{5.7.246}$$

If $s > s_0$ by (5.7.245) and (5.7.241) we just get

$$\|a_+\|_s^{\gamma,\mathcal{O}_+} \le_s \|f\|_s^{\gamma,\mathcal{O}_+}(1 + C_s\gamma^{-1}K^{2\tau+2}\|a\|_{s_0}^{\gamma,\mathcal{O}}) + C_s\gamma^{-1}K^{2\tau+2}\|a\|_s^{\gamma,\mathcal{O}}\|f\|_{s_0}^{\gamma,\mathcal{O}_+} \tag{5.7.247}$$

with

$$\|f\|_s^{\gamma,\mathcal{O}^+} \le \|\Pi_K^{\perp} a\|_s^{\gamma,\mathcal{O}} + 2C_s\gamma^{-1}K^{2\tau+2}\|a\|_{s_0}^{\gamma,\mathcal{O}}\|a\|_s^{\gamma,\mathcal{O}}. \tag{5.7.248}$$

By using (5.7.243), (5.7.235), (A.0.8) and (5.7.230) we get (5.7.238) and (5.7.239).

□

Now we describe the iteration of the KAM step.

Lemma 5.7.24. *Consider the vector field X_0 in (5.7.223). Fix $\gamma > 0$ set (recall $s_0 = [\nu/2] + 3$, $\tau = \nu + 2$)*

$$\chi = \frac{3}{2}, \quad \mu = 6\nu + 22, \quad \varrho = 2\nu + 9, \quad s_1 = 4\nu + 21. \tag{5.7.249}$$

There exists K_0 depending on s_0, ν such that if

$$C(s_1)\, \delta_0(s_1) K_0^{2\tau+\varrho} < 1, \qquad \text{where} \qquad \delta_0(s_1) := \gamma^{-1}\|a_0\|_{s_1}^{\gamma,\mathcal{O}_0} \tag{5.7.250}$$

then, for all $n \geq 0$, the following holds. We set $K_n := K_0^{\chi^n}$, $\chi := 3/2$ and

$$\mathcal{O}_{n+1} = \mathcal{C}_{m_n,K_n,\mathcal{O}_n} := \left\{ \omega \in \mathcal{O}_n : |\omega \cdot \ell - m_n\, j| \geq \frac{\gamma}{\langle \ell \rangle^\tau}, \quad \forall \ell \in \mathbb{Z}^\nu : |\ell| \leq K_n,\ \forall j \in \mathbb{Z} \setminus \{0\} \right\} \tag{5.7.251}$$

and for all $\omega \in \mathcal{O}_0$ we set $\alpha_{n+1}(\varphi, x; \omega)$ to be (recall (5.7.233) for the definition of χ)

$$\alpha_{n+1}(\varphi, x; \omega) := \sum_{|\ell| \leq K_n} \alpha_{\ell,j}^{(n+1)} e^{\mathrm{i}(\ell\cdot\varphi + jx)}, \quad \alpha_{\ell,j}^{(n+1)} := \frac{a_{\ell,j}^{(n)} \chi_{\ell,\gamma}(\omega\cdot\ell - m_n\, j)}{\mathrm{i}(\omega\cdot\ell - m_n\, j)} \tag{5.7.252}$$

and

$$\varepsilon_n := |\langle \Delta_{12} a_n \rangle|, \qquad \delta_n(s) := \gamma^{-1}\|a_n\|_s^{\gamma,\mathcal{O}_n}. \tag{5.7.253}$$

Moreover we set

$$\lambda := 1/(s - s_0 + 2\tau + 2), \quad \mathtt{M}(s) := \max\{\delta_0(s_1), \delta_0(s + 2\tau + 2)\}.$$

Then the following holds.

$(\mathcal{P}_1)_n$. *Set $\alpha_0 = 0$. For all $n \geq 0$ the torus diffeomorphism $\Phi_n : (\varphi, x) \mapsto (\varphi, x + \alpha_n(\varphi, x))$ is well defined from H^s to itself $\forall s \geq s_0$ and setting*

$$X_n := (\Phi_n)_* X_{n-1} := \omega \cdot \frac{\partial}{\partial\varphi} - (m_n + a_n(\varphi, x)) \frac{\partial}{\partial x} \tag{5.7.254}$$

we have the bounds

$$|m_n - m_{n-1}|^\gamma \leq \gamma\delta_0(s_1) K_{n-1}^{-\mu} K_0^\mu, \quad |m_n|^{lip} \leq M_0 + \delta_0(s_1) \sum_{j=1}^n 2^{-j}, \tag{5.7.255}$$

$$|\Delta_{12} m_n| \leq \sum_{j=0}^{n-1} 2^{-j} \varepsilon_0$$

and there exist $C_1(s)$ and $C_2(s)$, positive constants depending on s, such that

$$\delta_n(s_0) \leq \delta_0(s_1)\, K_0^\mu\, K_n^{-\mu}, \quad \delta_n(s) \leq C_1(s)\, \delta_0(s) (1 + \sum_{j=1}^n 2^{-j}), \quad s \geq s_0 \tag{5.7.256}$$

$$\varepsilon_n \leq 2^{-n}\varepsilon_0, \tag{5.7.257}$$

$$\|\alpha_n\|_s^{\gamma,\mathcal{O}_0} \leq \delta_n(s + 2\tau + 1) \leq C_2(s)\, K_n^{-\lambda\mu} K_0^{\lambda\mu}\, \mathtt{M}(s), \quad s \geq s_0. \tag{5.7.258}$$

$(\mathcal{P}_2)_n$. *The torus diffeomorphism defined by*

$$\begin{cases} \Psi_0 = \mathrm{I}, \\ \Psi_n = \Phi_n \circ \Psi_{n-1} \end{cases} \tag{5.7.259}$$

is of the form $\Psi_n : (\varphi, x) \mapsto (\varphi, x + \beta_n(\varphi, x))$ *with (recall* (5.7.258) *for the definition of* $\mathtt{M}(s)$*)*

$$\|\beta_n\|_{s_0}^{\gamma,\mathcal{O}_0} \le C(s_0)\delta_0(s_1)\sum_{j=0}^{n} 2^{-j}\,, \quad \|\beta_n\|_{s}^{\gamma,\mathcal{O}_0} \le C_3(s)\,\mathtt{M}(s+1)\sum_{j=0}^{n} 2^{-j}\,, \quad \forall s \ge s_0, \tag{5.7.260}$$

$$\|\beta_{n-1} - \beta_n\|_{s_0}^{\gamma,\mathcal{O}_0} \le C(s_0)\delta_0(s_1)2^{-n}\,, \quad \|\beta_{n-1} - \beta_n\|_{s}^{\gamma,\mathcal{O}_0} \le C_4(s)\,\mathtt{M}(s+2)2^{-n} \quad \forall s \ge s_0, \tag{5.7.261}$$

Proof. The proof is postponed in Appendix B. □

Now we can prove the Proposition 5.7.22.

Proof of Proposition 5.7.22. We fix s_1 as in (5.7.249) and choose $\delta_\star$ so that (5.7.224) implies (5.7.250). Consider now the sequence β_n defined in Lemma 5.7.24-($\mathcal{P}_2$). By formula (5.7.261) this is a Cauchy sequence in $H^s(\mathbb{T}^{\nu+1})$ for all $s \ge s_0$. Let us denote by $\beta^{(\infty)}$ its limit. We note that $\beta^{(\infty)}$ belongs to $\cap_{s\ge s_0} H^s(\mathbb{T}^{\nu+1})$, hence it is a C^∞ function in the variables φ, x.
In the same way, by (5.7.255) the sequence m_n is a Cauchy sequence and we denote by m_∞ its limit. We claim that

$$(\Psi^{(\infty)})^{-1}\Big(\omega\cdot\partial_\varphi\beta^{(\infty)} - (m_0 + a_0)(1 + (\beta^{(\infty)})_x)\Big) = m_\infty. \tag{5.7.262}$$

First we prove by induction that (recall (5.7.254))

$$(\Psi_n)_* X_0 = X_n. \tag{5.7.263}$$

For $n = 0$ this is trivially true. Now prove the $n+1$-th step. Recalling the definition (5.7.259), by the composition of pushforward

$$(\Psi_{n+1})_* X_0 = (\Phi_{n+1})_*(\Psi_n)_* X_0 = (\Phi_{n+1})_* X_n = X_{n+1}.$$

Now by (5.7.263) we have that

$$(\Psi_n)^{-1}\Big(\omega\cdot\partial_\varphi\beta^{(\infty)} - (m_0 + a_0)(1 + (\beta^{(\infty)})_x)\Big) = m_n + a_n \tag{5.7.264}$$

by (5.7.256) the l. h. s. of (5.7.264) converges in H^{s_0} to m_∞. By the fact that β_n converges to $\beta^{(\infty)}$ in H^s, for every $s \ge s_0$, then

$$(\Psi^{(\infty)})^{-1}\Big(\omega\cdot\partial_\varphi\beta^{(\infty)} - (m_0 + a_0)(1 + (\beta^{(\infty)})_x)\Big) - (\Psi_n)^{-1}\Big(\omega\cdot\partial_\varphi\beta^{(\infty)} - (m_0 + a_0)(1 + (\beta^{(\infty)})_x)\Big)$$

converges to 0 in H^{s_0} by using triangle inequalities and the bounds given in Lemma A.0.5. Then we proved our claim.
By (5.7.262), setting $\Psi^{(\infty)} : (\varphi, x) \mapsto (\varphi, x + \beta^{(\infty)}(\varphi, x))$, we have

$$\Psi^{(\infty)}_* X_0 = \omega\cdot\frac{\partial}{\partial\varphi} - m_\infty(\omega)\frac{\partial}{\partial x}\,, \quad \forall \omega \in \cap_n \mathcal{O}_n.$$

The bounds (5.7.225) follow by (5.7.255). In order to complete the proof we need to show that

$$\mathcal{O}_\infty^{2\gamma} \subset \bigcap_n \mathcal{O}_n.$$

We prove this by induction. By definition $\mathcal{O}_\infty^{2\gamma} \subset \mathcal{O}_0$. Suppose that $\mathcal{O}_\infty^{2\gamma} \subset \mathcal{O}_n$ and we claim that $\mathcal{O}_\infty^{2\gamma}$ is included in $\mathcal{O}_{n+1}$.
Fix $\omega \in \mathcal{O}_\infty^{2\gamma}$ and $|\ell| \le K_n$. Now suppose that $|j| \le K_n^2$, then by (5.7.255), (5.7.250) and recalling that $\mu = 4\nu + 12$ we have

$$|\omega \cdot \ell - m_n j| \ge |\omega \cdot \ell - m_\infty j| - |m_\infty - m_n||j| \ge \frac{\gamma}{\langle \ell \rangle^\tau}\Big(1 - \delta_0(s_1) K_0^\mu K_{n-1}^{-\mu} K_n^2 K_n^\tau\Big) \ge \frac{\gamma}{2\langle \ell \rangle^\tau}.$$

Otherwise, if $|j| > K_n^2$ then

$$|\omega \cdot \ell - m_n j| \ge |j| - |\omega| K_n > K_n^2 - 2|\overline{\omega}| K_n > 1 > \frac{\gamma}{2\langle \ell \rangle^\tau}.$$

This prove the claim and concludes the proof of the proposition. □

Lemma 5.7.25. *Under the assumption of Proposition 5.7.22, the function $\beta^{(\infty)}$ defined in the Proposition 5.7.22 satisfies the following estimate on the variation of the variable i*

$$\|\Delta_{12}\beta^{(\infty)}\|_s \le_s \gamma^{-1}\Big(\|\Delta_{12} a_0\|_{s+2\tau+1} + \|\beta^{(\infty)}\|_{s+\sigma'} \|\Delta_{12} a_0\|_{s_0+2\tau+1}\Big) \tag{5.7.265}$$

for some $\sigma' > 0$.

Proof. By (5.7.228) the function $\beta^{(\infty)}$ satisfies the equation

$$\omega \cdot \partial_\varphi \beta^{(\infty)} - (m_0 + a_0(\varphi, x))(1 + \beta_x^{(\infty)}) = -m_\infty \tag{5.7.266}$$

and differentiating this equation by i we get that $\alpha^{(\infty)}(\varphi, x) := \Delta_{12}\beta^{(\infty)}$ satisfies the following

$$L\alpha^{(\infty)} = f(\varphi, x), \qquad L := \omega \cdot \partial_\varphi - (m_0 + a_0)\partial_x \tag{5.7.267}$$

where

$$f(\varphi, x) := -\Delta_{12} m_\infty + (\Delta_{12} a_0)(1 + \beta_x). \tag{5.7.268}$$

By Proposition 5.7.22 the map $\Phi \colon H^s(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$ defined by

$$\Phi u(\varphi, x) := u(\varphi, x + \beta^{(\infty)}(\varphi, x)) \tag{5.7.269}$$

is such that

$$\Phi L \Phi^{-1}(\tilde{\alpha}^{(\infty)}) = \Big(\omega \cdot \partial_\varphi - m_\infty \partial_x\Big)(\tilde{\alpha}^{(\infty)}), \quad \tilde{\alpha}^{(\infty)} := \Phi(\alpha^{(\infty)}). \tag{5.7.270}$$

Hence the equation (5.7.267) transforms into

$$\Big(\omega \cdot \partial_\varphi - m_\infty \partial_x\Big)\tilde{\alpha}^{(\infty)} = g, \qquad g := \Phi^{-1} f$$

and by Lemma A.0.5, (5.7.227) and (5.7.225) we get

$$\|\tilde{\alpha}^{(\infty)}\|_s \le_s \gamma^{-1}\|g\|_{s+2\tau+1} \le_s \gamma^{-1}\Big(\|\Delta_{12} a_0\|_{s+2\tau+1} + \|\beta^{(\infty)}\|_{s+1}\|\Delta_{12} a_0\|_{s_0}\Big). \tag{5.7.271}$$

By definition

$$\|\alpha^{(\infty)}\|_s \le_s \|\Phi^{-1}\tilde{\alpha}^{(\infty)}\|_s \le_s \gamma^{-1}\Big(\|\Delta_{12} a_0\|_{s+2\tau+1} + \|\beta^{(\infty)}\|_{s+\sigma'}\|\Delta_{12} a_0\|_{s_0+2\tau+1}\Big). \tag{5.7.272}$$

□

5.7.4 Reduction at the highest order

We want to make constant coefficient at the highest order of the operator $\mathcal{L}_\omega$ defined in (5.6.31). In order to do that we shall apply Proposition 5.7.21 which gives an explicit formula for the new coefficient at the highest order (see (5.7.185)). Then Proposition 5.7.22 provides the solution for the equation (5.7.222). This is possible if the smallness assumption (5.7.224) is fullfilled. By (5.7.8) the coefficient a_0 in $\mathcal{L}_\omega$ does not satisfy this condition. Hence we have to perform some preliminary steps in order to enter in a perturbative regime for the scheme described in the proof of Proposition 5.7.22.

Consider

$$\rho \geq s_0 + 6\tau + 9. \tag{5.7.273}$$

We shall use the smallness assumption in (5.6.7) with some μ' such that

$$\mu' > \tilde{\sigma} \geq \sigma_8 + \sigma_0 + (s_1 - s_0) + \sigma_1 + \rho + 1, \tag{5.7.274}$$

where $\tilde{\sigma}$ is the loss of regularity in Lemma 5.7.45, σ_0 has been introduced in Section 5.7, see estimates (5.7.7)- (5.7.9), σ_1 is the index appearing in (5.7.91), while $s_0 + 2\tau + 4 \leq s_1$ (see Theorem 5.7.22). The constant $\sigma_8 \gg 1$ depends only on ρ. Essentially the constant σ_8 will be given by Proposition 5.7.21 (see (5.7.183)).

Step (ε). Consider the Hamiltonian

$$\begin{gathered} S(\tau) := \frac{1}{2}\int_{\mathbb{T}} b_1(\tau,\varphi,x)\, z^2\, dx = \varepsilon S_1(\tau) + \varepsilon^2 S_2(\tau) + \varepsilon^3 S_3(\tau) + S_4(\tau), \\ b_1 := \frac{\varepsilon\beta_1(\varphi,x)}{1+\tau\varepsilon(\beta_1)_x}, \end{gathered} \tag{5.7.275}$$

$$S_1 := \frac{1}{2}\int_{\mathbb{T}} \beta_1\, z^2\, dx, \quad S_2 := -\frac{\tau}{4}\int_{\mathbb{T}} \partial_x(\beta_1^2)\, z^2\, dx, \quad S_3 := \frac{\tau^2}{2}\int_{\mathbb{T}} \beta_1(\beta_1)_x^2\, dx, \tag{5.7.276}$$

for some function β_1 to be determined. The Hamiltonian system associated to εS_1 is

$$u_\tau = \Pi_S^{\perp}[\big(J\circ b_1(\tau)\big)\, u]. \tag{5.7.277}$$

We call Φ_1 the flow at time one of the system (5.7.277), then the Hamiltonian of the conjugated linearized operator $\Phi_1\mathcal{L}_\omega\Phi_1^{-1}$ is (recall (5.6.42), (5.6.43))

$$\begin{aligned} \mathsf{H}\circ\Phi_1^{-1} &= \mathsf{H} + \varepsilon\{S_1,\mathsf{H}\}_e + \varepsilon^2\Big(\frac{1}{2}\{S_1,\{S_1,\mathsf{H}\}_e\}_e + S_2(1)\Big) \\ &+ \frac{\varepsilon^3}{2}\Big(\frac{1}{3}\{S_1,\{S_1,\{S_1,\mathsf{H}\}_e\}_e\}_e + \{S_1,\{S_2,\mathsf{H}\}_e\}_e + \{S_2,\{S_1,\mathsf{H}\}_e\}_e + S_3(1)\Big) + o(\varepsilon^3), \end{aligned} \tag{5.7.278}$$

where $\{\cdot,\cdot\}_e$ are the Poisson brackets defined in (5.6.38) respect to the extended symplectic form (5.6.37).

In particular, we have

$$\begin{aligned} \mathsf{H}\circ\Phi_1^{-1} &= \mathsf{H}_0 + \varepsilon\Big(\mathsf{H}_1 + \{S_1,\mathsf{H}_0\}_e\Big) + \varepsilon^2\Big(\mathsf{H}_2 + S_2(1) + \{S_1,\mathsf{H}_1\}_e + \frac{1}{2}\{\{S_1,\{S_1,\mathsf{H}_0\}_e\}_e + \mathsf{H}_{\mathcal{R}_2}\Big) \\ &+ \varepsilon^3\Big(\mathsf{H}_3 + \frac{1}{2}\{S_2,\{S_1,\mathsf{H}_1\}_e\}_e + \frac{1}{2}\{S_1,\{S_2,\mathsf{H}_1\}_e\}_e \\ &+ \frac{1}{6}\{S_1,\{S_1,\{S_1,\mathsf{H}_0\}_e\}_e\}_e + S_3(1) + \mathsf{H}_{\mathcal{R}_*,3} + \mathbb{A}\xi\cdot\partial_\varphi\mathsf{H}_1\Big) + o(\varepsilon^3), \end{aligned}$$

where $\mathsf{H}_{\mathcal{R}_*,3}$ collects the terms of size $O(\varepsilon^3)$ in $\mathsf{H}_{\mathcal{R}_*}$ (recall that $\mathsf{H}_{\mathcal{R}_*}$ generates a finite rank vector field, see (5.6.32)). We want to solve the following equation

$$\mathsf{H}_1 + \{S_1, \mathsf{H}_0\}_e = \mathsf{H}_1 + \{S_1, \mathsf{H}_0\} + \mathcal{D}_{\overline{\omega}} S_1 = F, \tag{5.7.279}$$

where F is some Hamiltonian of the form $F := (Az, z)_{L^2(\mathbb{T})}$ with A pseudo differential operator of order -1. Recall (5.6.43), then the equation (5.7.279) is equivalent to the following one

$$\mathcal{D}_{\overline{\omega}}\beta_1 - (\beta_1)_x - \overline{v} = 0. \tag{5.7.280}$$

Hence we choose $\beta_1 = \frac{1}{3}(\Lambda\partial_x)^{-1}\overline{v}$ and we note that

$$\|\varepsilon\beta_1\|_s^{\gamma,\mathcal{O}_0} \leq_s \varepsilon, \quad \forall s \geq s_0, \tag{5.7.281}$$

thus the assumption (5.7.91) is satisfied and Proposition 5.7.16 applies.
By (5.7.8), (5.7.9), (5.7.10) and using the assumption (5.6.7) with μ' given in (5.7.274) the condition (5.7.183) holds. In this case $\mathtt{q} \rightsquigarrow 0$ (see (5.7.181)).
Then Proposition 5.7.21 applies and the new linearized operator is

$$\mathcal{L}_1 := \Phi_1 \mathcal{L}_\omega \Phi_1^{-1} = \Pi_S^\perp\Big(\mathcal{D}_\omega - J \circ \big(1 + a_1(\varphi, x)\big) + \mathrm{Op}(\mathtt{q}_1) + \widehat{\mathcal{Q}}_1\Big). \tag{5.7.282}$$

Hence by (5.7.281), (5.7.9), (5.7.187), (5.7.185), (5.7.7), we have that $\widehat{\mathcal{Q}}_1 \in \mathfrak{L}_\rho$ (see Definition 5.7.4) (with ρ as in (5.7.273)) and for $0 \leq \mathtt{b} \leq \rho - 2$

$$\begin{aligned} |\mathtt{q}_1|_{-1,s,\alpha}^{\gamma,\mathcal{O}_0} &\leq_{s,\alpha,\rho} \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_3}^{\gamma,\mathcal{O}_0}), \quad s \geq s_0 \\ \mathbb{M}^\gamma_{\widehat{\mathcal{Q}}_1}(s, \mathtt{b}) &\leq_{s,\rho} \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_3}^{\gamma,\mathcal{O}_0}), \quad s_0 \leq s \leq S_{max}, \end{aligned} \tag{5.7.283}$$

for $0 \leq \mathtt{b} \leq \rho - 3$

$$\begin{aligned} |\Delta_{12}\mathtt{q}_1|_{-1,s,\alpha} &\leq_{s,\alpha,\rho} \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_3} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_3}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_3}), \quad s \geq s_0, \\ \mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}_1}(s, \mathtt{b}) &\leq_{s,\alpha} \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_3} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_3}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_3}), \quad s_0 \leq s \leq S_{max}, \end{aligned} \tag{5.7.284}$$

$$\begin{aligned} \|a_1\|_s^{\gamma,\mathcal{O}_0} &\leq_s \varepsilon^2 + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_3}^{\gamma,\mathcal{O}_0}, \\ \|\Delta_{12}a_1\|_s &\leq_s \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_3} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_3}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_3}), \end{aligned} \tag{5.7.285}$$

with σ_0 in the estimates of (5.7.6) and σ_3 given in Proposition 5.7.21.
The only estimates that are not given by Lemma 5.7.21 are (5.7.285). The coefficient a_1 is given by (5.7.185) with $m \rightsquigarrow 1$, $a \rightsquigarrow a_0$, $a_+ \rightsquigarrow a_1$ and $\tilde{\beta}$ such that $x \mapsto x + \tilde{\beta}$ is the inverse of $x \mapsto x + \beta_1$. By the choice of β_1 in (5.7.280) we have eliminated the ε-terms from a_1. Hence by (5.7.281) and (5.7.8) we get (5.7.285).

Step (ε^2). Now we want to diagonalize the Hamiltonian (5.6.42) at order ε^2. In order to do that, we consider the auxiliary Hamiltonian

$$\tilde{S}(\tau) = \frac{1}{2}\int_{\mathbb{T}} b_2(x,\varphi)\, z^2\, dx = \varepsilon^2\tilde{S}_2 + \tilde{S}_4, \quad b_2 := \frac{\varepsilon^2\beta_2}{1 + \tau\varepsilon^2(\beta_2)_x}, \tag{5.7.286}$$

$$\tilde{S}_2 := \frac{1}{2}\int_{\mathbb{T}} \beta_2\, z^2\, dx, \quad \tilde{S}_4 := \tilde{S} - \varepsilon^2\tilde{S}_2, \tag{5.7.287}$$

where β_2 is some function to be determined. The Hamiltonian system associated to S_2 is

$$u_\tau = \Pi_S^\perp[\big(J \circ b_2(\tau)\big)\, u]. \tag{5.7.288}$$

If Φ_2 is the flow at time one of the system (5.7.288), then the Hamiltonian of the conjugated linearized operator $\mathcal{L}_2 := \Phi_2\mathcal{L}_1\Phi_2^{-1}$ is

$$\begin{aligned}\mathsf{H} \circ \Phi_1^{-1} \circ \Phi_2^{-1} =& \mathsf{H}_0 + \varepsilon \int \mathfrak{B}_1(z)\, z\, dx + \varepsilon^2\Big(\mathsf{H}_2 + S_2(1) + \mathsf{H}_{\mathcal{R}_2} + \{\tilde S_2, \mathsf{H}_0\}_e + \frac{1}{2}\{S_1, \mathsf{H}_1\} \\ &+ \frac{1}{2}\{S_1, \int \mathfrak{B}_1(z)\, z\, dx\}\Big) + \varepsilon^3\Big(\mathsf{H}_3 + \frac{1}{2}\{S_2, \{S_1, \mathsf{H}_1\}_e\}_e + \frac{1}{2}\{S_1, \{S_2, \mathsf{H}_1\}_e\}_e \\ &+ \frac{1}{6}\{S_1, \{S_1, \{S_1, \mathsf{H}_0\}_e\}_e\}_e + S_3(1) + \mathsf{H}_{\mathcal{R}_*,3} + \mathbb{A}\xi \cdot \partial_\varphi \mathsf{H}_1\Big) + o(\varepsilon^3),\end{aligned}$$

with

$$\mathfrak{B}_1 := [3\Lambda\partial_x, \beta_1]. \tag{5.7.289}$$

In particular

$$\frac{1}{2}\{S_1, \mathsf{H}_1\} = \frac{1}{2}\int_{\mathbb{T}} \overline{v} z\, J(\beta_1 z)\, dx - \frac{1}{2}\int_{\mathbb{T}} \overline{v} z\, J\Pi_S[\beta_1 z]\, dx, \tag{5.7.290}$$

$$\frac{1}{2}\{S_1, \int B_1(z)\, z\, dx\} = -\int_{\mathbb{T}} B_1(z)\, J\Pi_S^\perp[\beta_1 z]\, dx. \tag{5.7.291}$$

Since $[3\Lambda\partial_x, \beta_1]$ is a pseudo differential operator of order -2, the Hamiltonian (5.7.291) generates a vector field of order -1. We write

$$\tilde{\mathsf{H}}_{\mathcal{R}_2} := \mathsf{H}_{\mathcal{R}_2} - \frac{1}{2}\int_{\mathbb{T}} \overline{v} z\, J\Pi_S[\beta_1 z]\, dx. \tag{5.7.292}$$

Note that $\frac{1}{2}\int_{\mathbb{T}} vz\, J\Pi_S[\beta_1 z]\, dx$ generates a finite-rank vector field and Φ_2 is $\mathrm{I} + O(\varepsilon^2)$, hence the terms of size $O(\varepsilon)$ in $\Phi_1\mathcal{L}_\omega\Phi_1^{-1}$ are not changed. We want to solve the equation

$$\mathsf{H}_2 + S_2(1) + \tilde{\mathsf{H}}_{\mathcal{R}_2} + \mathcal{D}_{\overline{\omega}}\tilde S_2 + \{\tilde S_2, \mathsf{H}_0\} + \frac{1}{2}\int_{\mathbb{T}} \overline{v} z\, J(\beta_1 z)\, dx - \frac{1}{2}\{\int B_1(z)\, z\, dx, S_1\} = F \tag{5.7.293}$$

where F is some Hamiltonian of the form $F := (Az, z)_{L^2(\mathbb{T})}$ with A a pseudo differential operator of order -1, namely

$$\mathsf{H}_2 + S_2(1) + \mathcal{D}_{\overline{\omega}}\tilde S_2 + \{\tilde S_2, \mathsf{H}_0\} + \frac{1}{2}\int_{\mathbb{T}} \overline{v} z\, \partial_x(\beta_1 z)\, dx = F. \tag{5.7.294}$$

Hence equation (5.7.294) is equivalent to

$$\mathcal{D}_{\overline{\omega}}\beta_2 - (\beta_2)_x - \Psi_2(\overline{v}) + \frac{1}{2}\partial_x(\beta_1^2) - \frac{1}{2}\beta_1\overline{v}_x + \frac{1}{2}\overline{v}(\beta_1)_x = 0. \tag{5.7.295}$$

Note that $\Psi_2(\overline{v}), \beta_1\overline{v}_x, (\beta_1)_x\overline{v}$ and β_1^2 are quadratic functions of $\overline{v}$ (recall (5.6.39)) and by the fact that the index of time $\mathtt{l}(j)$ of $\overline{v}$ is tied with the Fourier index of the space j we have

$$\overline{\omega} \cdot (\mathtt{l}(j_1) + \mathtt{l}(j_2)) = \omega(j_1) + \omega(j_2), \qquad \forall j_1, j_2 \in S. \tag{5.7.296}$$

We have that $\omega(j_1)+\omega(j_2)-(j_1+j_2)=0$ if and only if $j_1+j_2=0$, since $j_1 j_2 \neq -1$. Hence we can choose β_2 such that

$$\mathcal{D}_{\overline{\omega}}\beta_2-(\beta_2)_x-\Psi_2(\overline{v})+\frac{1}{2}\partial_x(\beta_1^2)-\frac{1}{2}\beta_1\overline{v}_x+\frac{1}{2}\overline{v}(\beta_1)_x=\frac{1}{3}\int_{\mathbb{T}^{\nu+1}}(\overline{v}^2+\overline{v}_x^2)\,dx, \tag{5.7.297}$$

since $\Psi_2(\overline{v})$ and $\partial_x(\beta_1^2)$ have spatial zero average. Therefore

$$\mathcal{L}_2:=\Phi_2\mathcal{L}_1\Phi_2^{-1}=\Pi_S^{\perp}\Big(\mathcal{D}_\omega-J\circ(1+\varepsilon^2c(\xi)+a_2(\varphi,x))+\mathcal{Q}_2\Big), \tag{5.7.298}$$

with

$$c(\xi):=\frac{1}{3}\sum_{j\in S}(1+j^2)\,\xi_j=\frac{2}{3}\sum_{j\in S^+}(1+j^2)\,\xi_j, \tag{5.7.299}$$

and, by noting that

$$\|\varepsilon^2\beta_2\|_s^{\gamma,\mathcal{O}}\leq_s\varepsilon^2\quad\forall s\geq s_0, \tag{5.7.300}$$

by (5.7.283)-(5.7.285) and using the assumption (5.6.7) with μ' given in (5.7.274) the condition (5.7.183) holds. In this case $\mathtt{q}\rightsquigarrow\mathtt{q}_1$, hence by (5.7.283), (5.7.284) the bounds (5.7.181), (5.7.182) hold with $\mathtt{k}_1\rightsquigarrow\varepsilon$, $\mathtt{k}_2\rightsquigarrow\varepsilon$, $\mathtt{k}_3\rightsquigarrow\varepsilon$, $p\rightsquigarrow\mathfrak{I}_\delta$. Then Proposition 5.7.21 applies and by (5.7.300), (5.7.283)-(5.7.285), (5.7.188), we have for $0\leq\mathtt{b}\leq\rho-2$

$$\begin{aligned}&|\mathtt{q}_2|_{-1,s,\alpha}^{\gamma,\mathcal{O}_0}\leq_{s,\alpha}\varepsilon(1+\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_4}^{\gamma,\mathcal{O}_0}),\quad s\geq s_0,\\&\mathbb{M}^\gamma_{\widehat{\mathcal{Q}}_2}(s,\mathtt{b})\leq_{s,\rho}\varepsilon(1+\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_4}^{\gamma,\mathcal{O}_0}),\quad s_0\leq s\leq S_{max}\end{aligned} \tag{5.7.301}$$

and for $0\leq\mathtt{b}\leq\rho-3$

$$\begin{aligned}&|\Delta_{12}\mathtt{q}_2|_{-1,s,\alpha}\leq_{s,\alpha}\varepsilon(\|i_1-i_2\|_{s+\sigma_0+\sigma_4}+\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_4}\|i_1-i_2\|_{s_0+\sigma_0+\sigma_4}),\quad s\geq s_0,\\&\mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}_2}(s,\mathtt{b})\leq_{s,\rho}\varepsilon(\|i_1-i_2\|_{s+\sigma_0+\sigma_4}+\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_4}\|i_1-i_2\|_{s_0+\sigma_0+\sigma_4}),\quad s_0\leq s\leq S_{max},\end{aligned} \tag{5.7.302}$$

$$\begin{aligned}&\|a_2\|_s^{\gamma,\mathcal{O}_0}\leq_s\varepsilon^3+\varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_4}^{\gamma,\mathcal{O}_0},\\&\|\Delta_{12}a_2\|_s\leq_s\varepsilon(\|i_1-i_2\|_{s+\sigma_0+\sigma_4}+\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_4}\|i_1-i_2\|_{s_0+\sigma_0+4\sigma_4}),\end{aligned} \tag{5.7.303}$$

for $\sigma_4>0$ given by Proposition 5.7.21.

Step (ε^3). Now we want to diagonalize the Hamiltonian (5.6.42) at order ε^3. In order to do that, we consider the auxiliary Hamiltonian

$$\hat{S}(\tau)=\frac{1}{2}\int_{\mathbb{T}}b_3(x,\varphi)\,z^2\,dx=\varepsilon^3\hat{S}_3+\hat{S}_6,\quad b_3:=\frac{\varepsilon^3\beta_3}{1+\tau\varepsilon^3(\beta_3)_x} \tag{5.7.304}$$

$$\hat{S}_3:=\frac{1}{2}\int_{\mathbb{T}}\beta_3\,z^2\,dx,\quad\hat{S}_6:=\hat{S}-\varepsilon^3\hat{S}_3 \tag{5.7.305}$$

where β_3 is some function to be determined. The Hamiltonian system associated to $\hat{S}$ is

$$u_\tau=\Pi_S^{\perp}[\big(J\circ b_3(\tau)\big)\,u]. \tag{5.7.306}$$

If Φ_3 is the flow at time one of the system (5.7.306), then the Hamiltonian of the conjugated linearized operator $\mathcal{L}_3:=\Phi_3\mathcal{L}_2\Phi_3^{-1}$ is (recall (5.7.299))

$$\mathsf{H}\circ\Phi_1^{-1}\circ\Phi_2^{-1}\circ\Phi_3^{-1}=\left(1+\frac{\varepsilon^2}{2}\,c(\xi)\right)\mathsf{H}_0+\varepsilon\mathcal{K}_1+\varepsilon^2\mathcal{K}_2+\varepsilon^3\mathcal{K}_3+o(\varepsilon^3) \tag{5.7.307}$$

with

$$
\begin{aligned}
&\mathcal{K}_1 := \int \mathfrak{B}_1(z)\, z\, dx, \\
&\mathcal{K}_2 := \mathbb{A}\xi \cdot \eta + \int \mathfrak{B}_2(z) z\, dx + \frac{1}{2}\Big\{S_1, \int \mathfrak{B}_1(z)\, z\, dx\Big\} + \mathcal{K}_{\mathcal{R}_2}, \\
&\mathcal{K}_3 := \{\hat{S}_3, \mathsf{H}_0\}_e + \mathsf{H}_3 + \frac{1}{2}\{S_2, \{S_1, \mathsf{H}_1\}_e\}_e + \frac{1}{2}\{S_1, \{S_2, \mathsf{H}_1\}_e\}_e \\
&+ \frac{1}{6}\{S_1, \{S_1, \{S_1, \mathsf{H}_0\}_e\}_e\}_e + S_3(1) + \mathsf{H}_{\mathcal{R}_*,3} + \mathbb{A}\xi \cdot \partial_\varphi \mathsf{H}_1,
\end{aligned}
\tag{5.7.308}
$$

where $\mathcal{K}_{\mathcal{R}_2}$ is the Hamiltonian $\tilde{\mathsf{H}}_{\mathcal{R}_2}(v, z)$ in (5.7.292) evaluated at $(\overline{v}, 0)$ and

$$
\mathfrak{B}_2 := 3[\Lambda \partial_x, \beta_2]. \tag{5.7.309}
$$

We want to solve the equation

$$
\begin{aligned}
&\mathsf{H}_3 + S_3(1) + \{\hat{S}_3, \mathsf{H}_0\}_e + \frac{1}{2}\{S_2, \{S_1, \mathsf{H}_1\}_e, \}_e + \frac{1}{2}\{S_1, \{S_2, \mathsf{H}_1\}_e\}_e \\
&+ \frac{1}{6}\{S_1, \{S_1, \{S_1, \mathsf{H}_0\}_e\}_e\}_e + \mathcal{D}_{\omega - \overline{\omega}} \mathsf{H}_1 = F
\end{aligned}
\tag{5.7.310}
$$

where F is some Hamiltonian of the form $F := (Az, z)_{L^2(\mathbb{T})}$ with A a pseudo differential operator of order -1. This equation is equivalent to

$$
\mathcal{D}_{\overline{\omega}} \beta_3 - (\beta_3)_x = -\Psi_3(\overline{v}) + f_3(\overline{v}) - \mathbb{A}\xi \cdot \partial_\varphi \beta_1, \tag{5.7.311}
$$

where f_3 and Ψ_3 are cubic functions in $\overline{v}$ (recall (5.6.43), (5.6.28)) and so they are supported on few harmonics in time. We are not intersted in the exact expression of f_3.
We have that

$$
\overline{\omega} \cdot \big(\mathtt{l}(j_1) + \mathtt{l}(j_2) + \mathtt{l}(j_3)\big) - j_1 - j_2 - j_3 \neq 0 \qquad \forall j_1, j_2, j_3 \in S. \tag{5.7.312}
$$

Then we can choose β_3 such that equation (5.7.311) is satisfied. Hence we have

$$
\mathcal{L}_3 := \Phi_3 \mathcal{L}_2 \Phi_3^{-1} = \Pi_S^{\perp}\Big(\mathcal{D}_\omega - J \circ (1 + \varepsilon^2 c(\xi) + a_3(\varphi, x)) + \mathcal{Q}_3\Big), \tag{5.7.313}
$$

and by noting that

$$
\|\varepsilon^3 \beta_3\|_s^{\gamma, \mathcal{O}} \leq_s \varepsilon^3 \quad \forall s \geq s_0, \tag{5.7.314}
$$

the assumption (5.7.91) is satisfied for the system (5.7.306). By (5.7.301)-(5.7.303) and using the assumption (5.6.7) with μ' given in (5.7.274) the condition (5.7.183) holds.
In this case $\mathtt{q} \rightsquigarrow \mathtt{q}_2$, hence by (5.7.301), (5.7.302) the bounds (5.7.181), (5.7.182) hold with $\mathtt{k}_1 \rightsquigarrow \varepsilon$, $\mathtt{k}_2 \rightsquigarrow \varepsilon$, $\mathtt{k}_3 \rightsquigarrow \varepsilon$, $p \rightsquigarrow \mathfrak{I}_\delta$. Hence Proposition 5.7.21 applies and by (5.7.314), (5.7.301)-(5.7.303), (5.7.187)-(5.7.188), we have for $0 \leq \mathtt{b} \leq \rho - 2$

$$
\begin{aligned}
&|\mathtt{q}_3|_{-1,s,\alpha}^{\gamma, \mathcal{O}_0} \leq_{s,\alpha} \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_5}^{\gamma, \mathcal{O}_0}), \quad s \geq s_0, \\
&\mathbb{M}^\gamma_{\widehat{\mathcal{Q}}_3}(s, \mathtt{b}) \leq_{s,\rho} \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_5}^{\gamma, \mathcal{O}_0}), \quad s_0 \leq s \leq S_{max},
\end{aligned}
\tag{5.7.315}
$$

and for $0 \leq \mathtt{b} \leq \rho - 3$

$$
\begin{aligned}
&|\Delta_{12}\mathtt{q}_3|_{-1,s,\alpha} \leq_{s,\alpha} \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_5} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_5}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_5}), \quad s \geq s_0, \\
&\mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}_3}(s, \mathtt{b}) \leq_{s,\rho} \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_5} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_5}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_5}), \quad s_0 \leq s \leq S_{max},
\end{aligned}
\tag{5.7.316}
$$

$$\begin{aligned} &\|a_3\|_s^{\gamma,\mathcal{O}_0} \le_s \varepsilon^4 + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_5}^{\gamma,\mathcal{O}_0}, \\ &\|\Delta_{12}a_3\|_s \le_s \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_5} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_5}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_5}), \end{aligned} \tag{5.7.317}$$

for $\sigma_5 > 0$ given by Proposition 5.7.21.

Step (ε^4)-(ε^5). Consider $i = 4, 5$. We proceed exactly as in the previous steps. We consider a change of coordinates Φ_i as the time-one flow map of

$$u_\tau = \Pi_S^\perp[\big(J \circ b_i(\tau)\big)\, u]. \tag{5.7.318}$$

where

$$b_i := \frac{\varepsilon^i \beta_i}{1 + \varepsilon^i \tau(\beta_i)_x} \tag{5.7.319}$$

for some function $\beta_i \in H^s$ to be determined. We choose β_i in order to solve an equation like the following

$$\mathcal{D}_{\overline{\omega}}\beta_i - (\beta_i)_x = g(\overline{v}), \tag{5.7.320}$$

where g is a function with a zero of order i at the origin.

The condition (H1) in (1.2.14) implies that the equation (5.7.320) for $i = 4$ is solved up to remainders of the form

$$d(\xi) = \sum_{j_1, j_2 \in S} \mathtt{d}(j_1, j_2)\xi_{j_1}\xi_{j_2}. \tag{5.7.321}$$

The condition (H1) in (1.2.14) implies that there are no small divisors for (5.7.320) if $i = 5$.
Thus we have

$$\mathcal{L}_5 := (\Phi_5 \circ \Phi_4)\mathcal{L}_3(\Phi_5 \circ \Phi_4)^{-1} = \Pi_S^\perp\Big(\mathcal{D}_\omega - J \circ (1 + \varepsilon^2 c(\xi) + \varepsilon^4 d(\xi) + a_5(\varphi, x)) + \mathcal{Q}_5\Big) \tag{5.7.322}$$

and by noting that

$$\|\varepsilon^i \beta_i\|_s^{\gamma,\mathcal{O}} \le_s \varepsilon^i \quad i = 4, 5,\ \forall s \ge s_0, \tag{5.7.323}$$

the assumption (5.7.91) is satisfied for the system (5.7.318). By (5.7.315)-(5.7.317) and using the assumption (5.6.7) with μ' given in (5.7.274) the condition (5.7.183) holds.
Hence Proposition 5.7.21 applies and by (5.7.323), (5.7.315)-(5.7.317), (5.7.187), (5.7.188), we have for $0 \le \mathtt{b} \le \rho - 2$

$$\begin{aligned} &|\mathtt{q}_5|_{-1,s,\alpha}^{\gamma,\mathcal{O}_0} \le_{s,\alpha} \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_7}^{\gamma,\mathcal{O}_0}), \quad s \ge s_0, \\ &\mathbb{M}^\gamma_{\widehat{\mathcal{Q}}_5}(s, \mathtt{b}) \le_{s,\rho} \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_7}^{\gamma,\mathcal{O}_0}), \quad s_0 \le s \le S_{max} \end{aligned} \tag{5.7.324}$$

and for $0 \le \mathtt{b} \le \rho - 3$

$$\begin{aligned} &|\Delta_{12}\mathtt{q}_5|_{-1,s,\alpha} \le_{s,\alpha} \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_7} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_7}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_7}), \quad s \ge s_0 \\ &\mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}_5}(s, \mathtt{b}) \le_{s,\rho} \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_7} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_7}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_7}), \quad s_0 \le s \le S_{max}, \end{aligned} \tag{5.7.325}$$

$$\begin{aligned} &\|a_5\|_s^{\gamma,\mathcal{O}_0} \le_s \varepsilon^6 + \varepsilon\|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_7}^{\gamma,\mathcal{O}_0}, \\ &\|\Delta_{12}a_5\|_s \le_s \varepsilon(\|i_1 - i_2\|_{s+\sigma_0+\sigma_7} + \|\mathfrak{I}_\delta\|_{s+\sigma_0+\sigma_7}\|i_1 - i_2\|_{s_0+\sigma_0+\sigma_7}), \end{aligned} \tag{5.7.326}$$

with $\sigma_7 > 0$ given in Proposition 5.7.21. Now we apply Proposition 5.7.22 in order to make constant the coefficient a_5 of the linearized operator $\mathcal{L}_5$, namely we find β such that

$$\mathcal{D}_\omega\tilde{\beta} - (1 + a_5(\varphi, x))(1 + \tilde{\beta}_x) = \text{constant}. \tag{5.7.327}$$

Note that the smallness condition (5.7.224) is satisfied.

Proposition 5.7.26. *Let β be the function such that $x \mapsto x + \beta(\varphi, x)$ is the inverse diffeomorphism of the torus of $x \mapsto x + \beta^{(\infty)}(\varphi, x)$ with $\beta^{(\infty)}$ in* (5.7.227) *and let Φ_6 be the flow of the Hamiltonian PDE*

$$u_\tau = \Pi_S^\perp[\big(J \circ b(\tau)\big) u)], \tag{5.7.328}$$

where

$$b(\tau) := b(\tau, \varphi, x) = \frac{\beta}{1 + \tau\beta_x}, \qquad \|\beta\|_s^{\gamma, \mathcal{O}_0} \leq_s \gamma^{-1}\|a_5\|_{s+2\tau+4}^{\gamma, \mathcal{O}_\infty^{2\gamma}}, \quad \forall s \geq s_0. \tag{5.7.329}$$

Then the conjugated of the operator $\mathcal{L}_5$ in (5.7.313) *is*

$$\mathcal{L}_6 := \Phi_6\, \mathcal{L}_5\, \Phi_6^{-1} = \Pi_S^\perp\Big(\mathcal{D}_\omega - mJ + \mathcal{Q}_6\Big), \tag{5.7.330}$$

where $\mathcal{Q}_6 = \mathrm{Op}(\mathtt{q}_6) + \widehat{\mathcal{Q}}_6$ as in Proposition 5.7.21 and m is a constant such that

$$|m - 1|^\gamma \leq C\varepsilon^2, \quad |m|^{lip} \leq C, \qquad |\Delta_{12} m| \leq \varepsilon\|i_1 - i_2\|_{s_0+2}\,, \quad \forall \omega \in \mathcal{O}_\infty^{2\gamma}. \tag{5.7.331}$$

Moreover, for any $s \geq s_0$

$$|\mathtt{q}_6|_{-1,s,\alpha}^{\gamma, \mathcal{O}_\infty^{2\gamma}} \leq_s \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\hat\sigma}^{\gamma, \mathcal{O}_0}), \tag{5.7.332}$$

$$|\Delta_{12}\mathtt{q}_6|_{-1,s,\alpha} \leq_s \varepsilon\gamma^{-1}(\|i_1 - i_2\|_{s+\hat\sigma} + \|\mathfrak{I}_\delta\|_{s+\hat\sigma}\|i_1 - i_2\|_{s_0+\hat\sigma}) \tag{5.7.333}$$

and $\widehat{\mathcal{Q}}_6 \in \mathfrak{L}_\rho$ with for $0 \leq \mathtt{b} \leq \rho - 2$ and $s_0 \leq s \leq S_{max}$

$$\mathbb{M}^\gamma_{\widehat{\mathcal{Q}}_6}(s, \mathtt{b}) \leq_s \varepsilon(1 + \|\mathfrak{I}_\delta\|_{s+\hat\sigma}^{\gamma, \mathcal{O}_0}) \tag{5.7.334}$$

for $0 \leq \mathtt{b} \leq \rho - 3$ and $s_0 \leq s \leq S_{max}$

$$\mathbb{M}_{\Delta_{12}\widehat{\mathcal{Q}}_6}(s, \mathtt{b}) \leq_s \varepsilon\gamma^{-1}(\|i_1 - i_2\|_{s+\hat\sigma} + \|\mathfrak{I}_\delta\|_{s+\hat\sigma}\|i_1 - i_2\|_{s_0+\hat\sigma}), \tag{5.7.335}$$

with $\hat\sigma = \sigma_0 + \sigma_8 + \rho + s_1 - s_0$ for some σ_8, possibly larger than σ_7.

Proof. The first order linear differential operator (recall (5.7.299), (5.7.321))

$$\mathcal{D}_\omega - \big(1 + \varepsilon^2 c(\xi) + \varepsilon^4 d(\xi) + a_5(\varphi, x)\big)\partial_x \tag{5.7.336}$$

defined on $H^s_{S^\perp}(\mathbb{T}^{\nu+1})$ is associated to the vector field

$$X_0 := \omega \cdot \frac{\partial}{\partial\varphi} - \big(1 + \varepsilon^2 c(\xi) + \varepsilon^4 d(\xi) + a_5(\varphi, x)\big)\frac{\partial}{\partial x}. \tag{5.7.337}$$

By assumption (5.7.274) and by (5.7.326) (recall that s_1 and δ^* are given in Proposition 5.7.22, see (5.7.224))

$$C(s_1)\,\gamma^{-1}\|a_5\|_{s_1}^{\gamma, \mathcal{O}_\infty^{2\gamma}} \leq C(s_1)\varepsilon^{4-3a} = \delta^* \ll 1. \tag{5.7.338}$$

Thus the condition (5.7.224) is satisfied and the Proposition 5.7.22 applies to the vector field (5.7.337). In particular, we have that the operator (5.7.336) conjugated by the transformation

$$\mathcal{T}_{\beta^{(\infty)}} : u(\varphi, x) \mapsto u(\varphi, x + \beta^{(\infty)}(\varphi, x))$$

is associated to the vector field

$$\Psi_*^{(\infty)} X_0 = \omega \cdot \frac{\partial}{\partial \varphi} + (\Psi^{(\infty)})^{-1} \Big(\mathcal{D}_\omega \beta^{(\infty)} - (1 + \varepsilon^2 c(\xi) + \varepsilon^4 d(\xi) + a_5(\varphi, x)(1 + \beta_x^\infty) \Big) \frac{\partial}{\partial x},$$

and by Proposition 5.7.22

$$\mathcal{D}_\omega \beta^{(\infty)} - (1 + a_5(\varphi, x))(1 + \beta_x^{(\infty)}) = -m_\infty. \tag{5.7.339}$$

Hence $\beta^{(\infty)}$ solves the equation (5.7.327). Now we take β in (5.7.329) as the function such that $x \mapsto x + \beta(\varphi, x)$ is the inverse of the torus diffeomorphism $x \mapsto x + \beta^{(\infty)}(\varphi, x)$.
The bound in (5.7.329) follows by (5.7.227).
The constant m in (5.7.330) is m_∞ (recall (5.7.225)).
By (5.7.227), (5.7.317) and (5.7.274), for ε small enough, the function β satisfies (5.7.91), indeed

$$\|\beta\|_{s_0+\sigma_1}^{\gamma, \mathcal{O}_0} \leq C(s_1)\gamma^{-1} \|a_5\|_{s_0+\sigma_1+2\tau+4}^{\gamma, \mathcal{O}_\infty^{2\gamma}} \overset{(5.7.326)}{\leq} C(s_1)\,\gamma^{-1}\big(\varepsilon^6 + \varepsilon \|\mathfrak{I}_\delta\|_{s_0+\mu'}^{\gamma, \mathcal{O}_0}\big) \overset{(5.7.274)}{\leq} C(s_1)\varepsilon^{4-3a}. \tag{5.7.340}$$

Hence Φ_6 is well defined (see Propositions 5.7.16, 5.7.18). By (5.7.326) the bounds (5.7.181), (5.7.182) hold with $\mathtt{k}_1 \rightsquigarrow \varepsilon^6$, $\mathtt{k}_2 \rightsquigarrow \varepsilon$, $\mathtt{k}_3 \rightsquigarrow \varepsilon$ and Proposition 5.7.21 applies.
The estimates (5.7.331) follow by (5.7.225), (5.7.326) and the fact that $m = 1 + O(\varepsilon^2)$.
The estimates on $\mathcal{Q}_5$ satisfies (5.7.324), then by Lemma 5.7.21 the bound (5.7.334) holds on $\widehat{\mathcal{Q}}_6$.

The estimate (5.7.332) follows by (5.7.186) and (5.7.315). By Lemma 5.7.25 we have (5.7.333), (5.7.335). □

5.7.5 Linear Birkhoff Normal Form

Fix

$$\mathtt{b} := s_0 + 6\tau + 6. \tag{5.7.341}$$

We recall that the linearized operator is now

$$\mathcal{L}_6 := \Pi_S^\perp \Big(\mathcal{D}_\omega - mJ + \mathcal{Q}_6 \Big) \tag{5.7.342}$$

where $\mathcal{Q}_6 = \mathrm{Op}(\mathtt{q}_6) + \widehat{\mathcal{Q}}_6$, $\mathtt{q}_6 \in S^{-1}$ and $\widehat{\mathcal{Q}}_6 \in \mathfrak{L}_\rho$ (recall the Definition 5.7.4).
If $\Phi := \Phi_6 \circ \Phi_5 \circ \Phi_4 \circ \Phi_3 \circ \Phi_2 \circ \Phi_1$ then the Hamiltonian of the operator $\mathcal{L}_6$ is (recall (5.6.42) and (5.7.342))

$$\mathcal{K} := \mathsf{H} \circ \Phi^{-1} = \mathsf{H}_0 + \varepsilon\, \mathcal{K}_1 + \varepsilon^2\, \mathcal{K}_2 + \varepsilon^3\, \mathcal{K}_3 + o(\varepsilon^3) \tag{5.7.343}$$

where $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$ collect all the terms of order ε, ε^2 and ε^3 respectively (recall (5.7.308), (5.7.289), (5.7.309)).

The aim of this section is to eliminate $\mathcal{K}_1$, $\mathcal{K}_3$ from (5.7.343) and normalize the Hamiltonian $\mathcal{K}_2$.
We have

$$\mathcal{L}_6 = \Pi_S^\perp \big(\mathcal{D}_\omega - mJ - \varepsilon X_{\mathcal{K}_1} - \varepsilon^2 X_{\mathcal{K}_2} - \varepsilon^3 X_{\mathcal{K}_3} + \mathfrak{R} \big) \tag{5.7.344}$$

where (recall (5.7.330), (5.7.308))

$$\begin{aligned} X_{\mathcal{K}_i} &:= J\nabla \mathcal{K}_i \qquad i = 1, 3, \qquad X_{\mathcal{K}_2} := J\nabla(\mathcal{K}_2 - \mathbb{A}\xi \cdot \eta), \\ \mathfrak{R} &:= \mathcal{Q}_6 + \varepsilon J\nabla \mathcal{K}_1 + \varepsilon^2 J\nabla \mathcal{K}_2 + \varepsilon^3 J\nabla \mathcal{K}_3. \end{aligned} \tag{5.7.345}$$

Remark 5.7.27. Note that $\varepsilon X_{\mathcal{K}_1}$, $\varepsilon^2 X_{\mathcal{K}_2}$ and $\varepsilon^3 X_{\mathcal{K}_3}$ are pseudo differential operators of order -1 (see (5.7.308)). Indeed, $H_{\mathcal{R}_*,3}$ generates a finite rank vector field and S_1, S_2 and $\hat{S}_3$ in (5.7.276), (5.7.305) generate a pseudo differential vector field and clearly also the Poisson brackets of these Hamiltonians.
In particular $\varepsilon X_{\mathcal{K}_1}$, $\varepsilon^2 X_{\mathcal{K}_2}$ and $\varepsilon^3 X_{\mathcal{K}_3}$ are the terms of order ε, ε^2 and ε^3 of the Taylor expansion of $\mathcal{Q}_6$ (see Proposition 5.7.26). Indeed, $\mathsf{H}\circ\Phi^{-1}$ and β in (5.7.329) are functions of $(\varepsilon v_\delta, \varepsilon^b z_0)$. We consider the Taylor expansion at $(\varepsilon\overline{v}, 0)$ and increment $h := \varepsilon(v_\delta - \overline{v}) + \varepsilon^b z_0$

$$\mathsf{H}\circ\Phi^{-1} = \mathsf{H}\circ\Phi^{-1}(\varepsilon\overline{v}, 0) + O(h).$$

Since the coefficient a_5 in (5.7.326) has a zero of order six at $(\varepsilon\overline{v}, 0)$, by (5.7.340) and the fact that $\varepsilon^{10-2b}\gamma^{-2} = \varepsilon^{4-3a}$ with $a \ll 1$, the terms of order $\varepsilon, \varepsilon^2$ and ε^3 in $\mathsf{H}\circ\Phi_1^{-1}\circ\cdots\circ\Phi_5^{-1}$ are not changed by applying the transformation Φ_6 defined in Proposition 5.7.26.
By Lemma 5.7.15-(ii) $\varepsilon X_{\mathcal{K}_1}, \varepsilon^2 X_{\mathcal{K}_2}, \varepsilon^3 X_{\mathcal{K}_3}$ belong to $\mathfrak{C}_{1,\mathtt{b}}$ and satisfy, by (5.7.289), (5.7.281), (5.7.309), (5.7.300), (5.7.314),

$$\begin{aligned} &\mathbb{B}^{\gamma}_{\varepsilon J\nabla\mathcal{K}_1}(s,\mathtt{b}) \le \varepsilon C(s,\mathtt{b}),\\ &\mathbb{B}^{\gamma}_{\varepsilon^2 J\nabla\mathcal{K}_2}(s,\mathtt{b}) \le \varepsilon^2 C(s,\mathtt{b}),\\ &\mathbb{B}^{\gamma}_{\varepsilon^3 J\nabla\mathcal{K}_3}(s,\mathtt{b}) \le \varepsilon^3 C(s,\mathtt{b}). \end{aligned} \tag{5.7.346}$$

Lemma 5.7.28. *We have that* $\mathfrak{R} \in \mathfrak{C}_{1,\mathtt{b}}(\mathcal{O}^{2\gamma}_\infty)$ *with*

$$\begin{aligned} \mathbb{B}^{\gamma}_{\mathfrak{R}}(s,\mathtt{b}) &\le_s \varepsilon^{4-3a} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|^{\gamma,\mathcal{O}_0}_{s+\hat{\sigma}},\\ \mathbb{B}_{\Delta_{12}\mathfrak{R}}(s,\mathtt{b}) &\le_s \varepsilon\gamma^{-1}(\|i_1 - i_2\|_{s+\hat{\sigma}} + \|\mathfrak{I}_\delta\|_{s+\hat{\sigma}}\|i_1 - i_2\|_{s_0+\hat{\sigma}}), \end{aligned} \tag{5.7.347}$$

for $\hat{\sigma} = \sigma_8 + \sigma_0 + (s_1 - s_0) + \rho + 1$ *(recall* (5.7.274)*).*

Proof. By the definition of $\mathfrak{R}$, see (5.7.345), and by the fact that, by Lemma 5.7.15-$(i),(ii)$, the pseudo differential of order -1 belong to $\mathfrak{C}_{1,\mathtt{b}}$ (in particular, for any choice of $\mathtt{b}$) and the elements of the class $\mathfrak{L}_\rho$ with ρ as in (5.7.273) are included in $\mathfrak{C}_{1,\mathtt{b}}$ with $\mathtt{b}$ as in (5.7.341), we have that $\mathfrak{R} \in \mathfrak{C}_{1,\mathtt{b}}$. Note that only $\mathcal{Q}_6$ in (5.7.345) depend on the torus embedding i_δ, then the second bound in (5.7.347) follows by (5.7.333) and (5.7.335). The first bound in (5.7.347) follows by the Remark 5.7.27 and the bounds (5.7.329) and (5.7.326). □

Remark 5.7.29. In the following steps of linear Birkhoff normal form we shall use the relation

$$\sum_{i=1}^{\nu} \overline{\jmath}_i \ell_i + j' - j = 0 \qquad \text{if } |\ell| \le 3, \quad \forall j,\, j' \in S^c,$$

which holds by the conservation of momentum.

Step one (order ε)

We look for a symplectic change of variable $\Upsilon_1 \colon H^s_{S^\perp}(\mathbb{T}^{\nu+1}) \to H^s_{S^\perp}(\mathbb{T}^{\nu+1})$, that is the time-1 flow of a quadratic Hamiltonian

$$H_{A_1}(u) := \varepsilon \sum_{j,j'\in S^c} (A_1)_j^{j'}(\varphi)\, u_{j'}\, \overline{u}_j, \tag{5.7.348}$$

where $A_1(\varphi)$ is a self-adjoint operator $\forall \varphi \in \mathbb{T}^\nu$ and thus

$$\Upsilon_1 := \exp(\varepsilon\, JA_1) = \mathrm{I}_{H_S^\perp} + \varepsilon\, JA_1 + \varepsilon^2 \frac{(JA_1)^2}{2} + \varepsilon^3 R_1, \quad R_1 := \sum_{k\geq 3} \frac{\varepsilon^{k-3}}{k!}\,(JA_1)^k. \tag{5.7.349}$$

Then

$$\begin{aligned} \mathcal{K}^{(1)} &:= \mathcal{K} \circ \Upsilon_1^{-1} = \mathsf{H}_0 + \varepsilon \mathcal{K}_1^{(1)} + \varepsilon^2 \mathcal{K}_2^{(1)} + \varepsilon^3 \mathcal{K}_3^{(1)} + o(\varepsilon^3), \quad \mathcal{K}_1^{(1)} := \mathcal{K}_1 + \mathcal{D}_{\overline{\omega}} H_{A_1} + \{H_{A_1}, \mathsf{H}_0\} \\ \mathcal{K}_2^{(1)} &:= \mathcal{K}_2 + \frac{1}{2}\{H_{A_1}, +\mathcal{D}_{\overline{\omega}} H_{A_1} + \{H_{A_1}, \mathsf{H}_0\} + \mathcal{K}_1\} + \frac{1}{2}\{H_{A_1}, \mathcal{K}_1\} \end{aligned} \tag{5.7.350}$$

We choose A_1 such that

$$\mathcal{D}_{\overline{\omega}} H_{A_1} + \{H_{A_1}, \mathsf{H}_0\} + \mathcal{K}_1 = 0. \tag{5.7.351}$$

We have

$$\begin{aligned} \mathcal{K}_1(u) &:= \sum_{j,j'\in S^c} (\mathfrak{B}_1)_j^{j'}(\varphi)\, u_{j'}\, u_{-j}, \\ (\mathfrak{B}_1)_j^{j'}(\ell) &= \frac{(1+(j-j')^2)(1-jj')\sqrt{\xi_{j-j'}}}{(1+j^2)(1+j'^2)}, \quad j-j' \in S,\ \ell = \mathtt{l}(j-j'), \end{aligned} \tag{5.7.352}$$

then we choose

$$(A_1)_j^{j'}(\ell) = \begin{cases} -\dfrac{(\mathfrak{B}_1)_j^{j'}(\ell)}{\mathrm{i}(\overline{\omega}\cdot \ell + \omega(j') - \omega(j))} & j,j' \in S^c,\ j-j' \in S,\ \ell = \mathtt{l}(j-j'), \\ 0 & \text{otherwise} \end{cases} \tag{5.7.353}$$

(recall (5.6.39) for the definition of $\mathtt{l}(j-j')$). This operator is well defined since

$$\begin{aligned} \overline{\omega}\cdot \mathtt{l}(j-j') + \omega(j') - \omega(j) &= \omega(j-j') - \omega(j) + \omega(j') \\ &= \frac{3jj'(j-j')[3+jj'+(j-j')^2]}{(1+j^2)(1+j'^2)(1+(j-j')^2)} \neq 0 \end{aligned} \tag{5.7.354}$$

and $|\omega - \overline{\omega}| \leq C\varepsilon^2$.

Lemma 5.7.30. *For $j,j' \in S^c$, $j-j' \in S$, $|\ell| \leq 1$ we have that*

$$|(A_1)_j^{j'}(\ell)| \leq \frac{K_1}{|j\, j'|} \tag{5.7.355}$$

for some constant $K_1 = K_1(S) > 0$ depending only by the set S. Otherwise we have that $(A_1)_j^{j'}(\ell) = 0$.

Proof. By (5.7.352), (5.7.353), (5.7.354) we have

$$(A_1)_j^{j'}(\ell) = \frac{1}{3}\, \frac{(1-jj')(1+(j-j')^2)^2}{jj'(j-j')(3+jj'+(j-j')^2)}\, \sqrt{\xi_{j-j'}}.$$

Since $1 < |j-j'| < 2C_S$, where $C_S := max\{|j| : j \in S\}$, $|\omega - \overline{\omega}| \leq C\varepsilon^2$,

$$|j\ j'||j-j'||3+jj'+(j-j')^2| \geq |j\ j'|^2 \quad \forall j \neq j', \qquad |1-j\,j'| \leq 2|j\,j'| \tag{5.7.356}$$

and by (5.7.331) we have (5.7.355).

□

Lemma 5.7.31. *The linear vector field $X_{A_1} := JA_1$ belongs to the class $\mathfrak{C}_{1,\mathtt{b}}$, in particular it satisfies the following*

$$\mathbb{B}^{\gamma}_{\varepsilon X_{A_1}}(s,\mathtt{b}) \leq C(s,\mathtt{b})\varepsilon. \tag{5.7.357}$$

Note that X_{A_1} does not depend on $i(\omega)$.

Proof. We have

$$\begin{aligned}
&\|\langle D_x\rangle^{1/2} JA_1 \langle D_x\rangle^{1/2} h\|_s^2 \\
&\leq \sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\Big(\sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} \langle j\rangle^{1/2}|\omega(j)||(A_1)_j^{j'}(\ell-\ell')|\,\langle j'\rangle^{1/2}|h_{\ell',j'}|\Big)^2\langle \ell,j\rangle^{2s} \\
&+ \sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\Big(\sum_{\substack{l'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle> 2\langle \ell',j'\rangle}} \langle j\rangle^{1/2}|\omega(j)||(A_1)_j^{j'}(\ell-\ell')|\,\langle j'\rangle^{1/2}|h_{\ell',j'}|\Big)^2\langle \ell,j\rangle^{2s}.
\end{aligned} \tag{5.7.358}$$

If $\langle \ell-\ell',j-j'\rangle \leq 2\langle \ell',j'\rangle$ then

$$\langle \ell,j\rangle \leq \langle \ell',j'\rangle + \langle \ell-\ell',j-j'\rangle \leq 3\langle \ell',j'\rangle.$$

Hence

$$\begin{aligned}
&\sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\Big(\sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} \langle j\rangle^{1/2}\langle j'\rangle^{1/2}|\omega(j)||(A_1)_j^{j'}(\ell-\ell')|\,|h_{\ell',j'}|\langle \ell,j\rangle^s\Big)^2 \\
&\leq 3^s \sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\Big(\sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} \langle j\rangle^{1/2}\langle j'\rangle^{1/2}|\omega(j)||(A_1)_j^{j'}(\ell-\ell')|\,\frac{\langle \ell-\ell',j-j'\rangle^{s_0}}{\langle \ell-\ell',j-j'\rangle^{s_0}}|h_{\ell',j'}|\langle \ell',j'\rangle^s\Big)^2
\end{aligned} \tag{5.7.359}$$

and by Cauchy-Schwarz

$$\begin{aligned}
&\sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\Big(\sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} |\omega(j)||(A_1)_j^{j'}(\ell-\ell')|\,\frac{\langle \ell-\ell',j-j'\rangle^{s_0}}{\langle \ell-\ell',j-j'\rangle^{s_0}}|h_{\ell',j'}|\langle \ell',j'\rangle^s\Big)^2 \\
&\leq (4C_S)^{2s_0}\tilde{C}\sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} \langle j\rangle\langle j'\rangle|\omega(j)|^2|(A_1)_j^{j'}(\ell-\ell')|^2|h_{\ell',j'}|^2\langle \ell',j'\rangle^{2s}
\end{aligned} \tag{5.7.360}$$

where

$$\tilde{C} = \sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} \frac{1}{\langle \ell-\ell',j-j'\rangle^{2s_0}} < \infty.$$

Note that fixed ℓ and j, since $|\ell-\ell'|<2$ and $|j-j'|<2C_S$, the sum above is finite and $\tilde{C}$ does not depend on ℓ or j.
By Lemma 5.7.30 and the fact that $|\omega(j)|\leq 3|j|$ we have that

$$\begin{aligned}
&\sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} \langle j\rangle\langle j'\rangle|\omega(j)|^2|(A_1)_j^{j'}(\ell-\ell')|^2|h_{\ell',j'}|^2\langle l',j'\rangle^{2s} \\
&\leq 9K^2\sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\sum_{\substack{\ell\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle \ell-\ell',j-j'\rangle\leq 2\langle \ell',j'\rangle}} |h_{\ell',j'}|^2\langle \ell',j'\rangle^{2s}.
\end{aligned} \tag{5.7.361}$$

Thus we get

$$\sum_{j\in S^c,\ell\in\mathbb{Z}^\nu}\Big(\sum_{\substack{\ell'\in\mathbb{Z}^\nu,j'\in S^c,\\ \langle\ell-\ell',j-j'\rangle\le 2\langle\ell',j'\rangle}}\langle j\rangle^{1/2}\langle j'\rangle^{1/2}|\omega(j)||(A_1)_j^{j'}(\ell-\ell')|\,|h_{\ell',j'}|\langle\ell,j\rangle^s\Big)^2\le 3^s\,C(s_0)\|h\|_s^2. \tag{5.7.362}$$

The case $\langle\ell-\ell',j-j'\rangle>2\langle\ell',j'\rangle$ is analogous. By the definition (5.7.353) of A_1 and (5.7.289) we have that the argument proved above holds also for the norm $\|\cdot\|^{\gamma,\mathcal{O}}$, since β in (5.7.289) is Lipschitz in ω.
The matrix elements of $\partial_{\varphi_m}^{\mathtt{b}}X_{A_1}$, $[X_{A_1},\partial_x]$, $[\partial_{\varphi_m}^{\mathtt{b}}X_{A_1},\partial_x]$ are respectively $\langle\ell_m-\ell'_m\rangle^{\mathtt{b}}\omega(j)(A_1)_j^{j'}(\ell-\ell')$, $(j-j')\omega(j)(A_1)_j^{j'}(\ell-\ell')$, $\langle\ell_m-\ell'_m\rangle^{\mathtt{b}}(j-j')\omega(j)(A_1)_j^{j'}(\ell-\ell')$. Note that by the definition of A_1 in (5.7.353)

$$\langle\ell_m-\ell'_m\rangle^{\mathtt{b}},|j-j'|\le C$$

for some constant C depending on the set S. Thus arguing as above one can easily prove that $\partial_{\varphi_m}^{\mathtt{b}}X_{A_1}$, $[X_{A_1},\partial_x]$, $[\partial_{\varphi_m}^{\mathtt{b}}X_{A_1},\partial_x]$ are -1-Lip-tame operators. This concludes the proof. □

By Lemma 5.7.15 and (5.7.357) we deduce the following result.

Corollary 5.7.32. *The transformation $\Upsilon_1\colon H^s(\mathbb{T}^{\nu+1})\to H^s(\mathbb{T}^{\nu+1})$ defined in (5.7.349) is invertible and satisfies, for any $u=u(\omega)\in H^s$ Lipschitz in $\omega\in\mathcal{O}_\infty^{2\gamma}$,*

$$\|(\Upsilon_1^{\pm1}-\mathrm{I})u\|_s^{\gamma,\mathcal{O}_\infty^{2\gamma}}\le_s C(s_0,\mathtt{b})\varepsilon\|u\|_s^{\gamma,\mathcal{O}_\infty^{2\gamma}}+C(s,\mathtt{b})\varepsilon\|u\|_{s_0}^{\gamma,\mathcal{O}_\infty^{2\gamma}}. \tag{5.7.363}$$

Proof. By (5.7.349) we have

$$(\Upsilon_1-\mathrm{I})u=\sum_{k\ge1}\frac{X_{A_1}^k u}{k!},\qquad(\Upsilon_1^{-1}-\mathrm{I})u=\sum_{k\ge1}(-1)^k\frac{X_{A_1}^k u}{k!}. \tag{5.7.364}$$

By using iteratively the property *(iii)* of Lemma 5.7.15 and Lemma 5.7.31 we have that

$$\begin{aligned}\|X_{A_1}^k u\|^{\gamma,\mathcal{O}_\infty^{2\gamma}}&\le\mathbb{B}_{X_{A_1}}^\gamma(s,\mathtt{b})(\mathbb{B}_{X_{A_1}}^\gamma(s_0,\mathtt{b}))^{k-1}\|u\|_{s_0}^{\gamma,\mathcal{O}_\infty^{2\gamma}}+(\mathbb{B}_{X_{A_1}}^\gamma(s_0,\mathtt{b}))^k\|u\|_s^{\gamma,\mathcal{O}_\infty^{2\gamma}}\\&\le\varepsilon C(s,\mathtt{b})\varepsilon^{k-1}C(s_0,\mathtt{b})^{k-1}\|u\|_{s_0}^{\gamma,\mathcal{O}_\infty^{2\gamma}}+\varepsilon^k C(s_0,\mathtt{b})^k\|u\|_s^{\gamma,\mathcal{O}_\infty^{2\gamma}}.\end{aligned}$$

By using this relation to estimate the Lip-Sobolev norm of (5.7.364) and by noting that $\varepsilon^n C(s_0,\mathtt{b})^n$ is a summable sequence for ε small enough we prove the thesis.

□

To shorten the notation in the following lemma we write

$$\mathrm{ad}_{X_{A_1}}[\cdot]:=[X_{A_1},\cdot].$$

Lemma 5.7.33. *The transformed operator is (recall (5.7.350))*

$$\mathcal{L}_7 h:=\Upsilon_1\mathcal{L}_6\Upsilon_1^{-1}=\Pi_S^\perp\Big(\mathcal{D}_\omega-mJ-\varepsilon^2X_{\mathcal{K}_2^{(1)}}-\varepsilon^3X_{\mathcal{K}_3^{(1)}}+\mathcal{R}_7\Big) \tag{5.7.365}$$

where

$$X_{\mathcal{K}_2^{(1)}}:=J\nabla\mathcal{K}_2^{(1)}=X_{\mathcal{K}_2}+\mathrm{ad}_{X_{A_1}}[X_{\mathcal{K}_1}]+\frac12\mathrm{ad}_{X_{A_1}}^2[\mathcal{D}_\omega-mJ] \tag{5.7.366}$$

$$X_{\mathcal{K}_3^{(1)}} := J\nabla\mathcal{K}_3^{(1)} = X_{\mathcal{K}_3} + \mathrm{ad}_{X_{A_1}}[X_{\mathcal{K}_2}] + \frac{1}{2}\mathrm{ad}^2_{X_{A_1}}[X_{\mathcal{K}_1}] + \frac{1}{6}\mathrm{ad}^3_{X_{A_1}}[\mathcal{D}_\omega - mJ] \tag{5.7.367}$$

$$\begin{aligned}\mathcal{R}_7 :&= \mathfrak{R} + \varepsilon\,\mathrm{ad}_{X_{A_1}}[\varepsilon^3 X_{\mathcal{K}_3} + \mathfrak{R}] \\ &+ \frac{\varepsilon^2}{2}\mathrm{ad}^2_{X_{A_1}}[-\varepsilon^2 X_{\mathcal{K}_2} - \varepsilon^3 X_{\mathcal{K}_3} + \mathfrak{R}] \\ &+ \frac{\varepsilon^3}{6}\mathrm{ad}^3_{X_{A_1}}[-\varepsilon X_{\mathcal{K}_1} - \varepsilon^2 X_{\mathcal{K}_2} - \varepsilon^3 X_{\mathcal{K}_3} + \mathfrak{R}] + \sum_{k\geq 4}\frac{\varepsilon^k}{k!}\mathrm{ad}^k_{X_{A_1}}[\mathcal{L}_6]\end{aligned} \tag{5.7.368}$$

and the following holds : the operator $\mathcal{R}_7 \in \mathfrak{C}_{1,\mathtt{b}}$ *with*

$$\begin{aligned}\mathbb{B}^\gamma_{\mathcal{R}_7}(s,\mathtt{b}) &\leq_s \varepsilon^{4-3a} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|^{\gamma,\mathcal{O}_0}_{s+\tilde\sigma}, \\ \mathbb{B}_{\Delta_{12}\mathcal{R}_7}(s,\mathtt{b}) &\leq_s \varepsilon\gamma^{-1}(\|i_1 - i_2\|_{s+\tilde\sigma} + \|\mathfrak{I}_\delta\|_{s+\tilde\sigma}\|i_1 - i_2\|_{s_0+\tilde\sigma}),\end{aligned} \tag{5.7.369}$$

for some $\tilde\sigma \geq \sigma_0 + \sigma_8 + \mathtt{b} + 1 + s_1 - s_0$ *(recall the loss of regularity in* (5.7.332), (5.7.334), (5.7.335)*).*

Proof. By Lemma 5.7.15 if $Y \in \mathfrak{C}_{1,\mathtt{b}}$ then $\mathrm{ad}^k_{X_{A_1}}[Y] \in \mathfrak{C}_{1,\mathtt{b}}$ for any $k \geq 1$. Moreover

$$\mathbb{B}^\gamma_{\mathrm{ad}_{X_{A_1}}[Y]}(s,\mathtt{b}) \leq_s \mathbb{B}^\gamma_{X_{A_1}}(s,\mathtt{b})\mathbb{B}^\gamma_Y(s_0,\mathtt{b}) + \mathbb{B}^\gamma_{X_{A_1}}(s_0,\mathtt{b})\mathbb{B}^\gamma_Y(s,\mathtt{b}).$$

and by applying iteratively this estimate we get, for any $k \geq 1$,

$$\mathbb{B}^\gamma_{\mathrm{ad}^k_{X_{A_1}}[Y]}(s,\mathtt{b}) \leq_s \mathbb{B}^\gamma_{X_{A_1}}(s,\mathtt{b})\big(\mathbb{B}^\gamma_{X_{A_1}}(s_0,\mathtt{b})\big)^{k-1}\mathbb{B}^\gamma_Y(s_0,\mathtt{b}) + \big(\mathbb{B}^\gamma_{X_{A_1}}(s_0,\mathtt{b})\big)^k\mathbb{B}^\gamma_Y(s,\mathtt{b}). \tag{5.7.370}$$

Hence, if $\mathcal{Z}_n := \sum_{k\geq n}\frac{\varepsilon^k}{k!}\mathrm{ad}^k_{X_{A_1}}[Y]$ for any $n \geq 1$, by (5.7.357) we have

$$\begin{aligned}\mathbb{B}^\gamma_{\mathcal{Z}_n}(s,\mathtt{b}) &\leq_s \mathbb{B}^\gamma_{X_{A_1}}(s,\mathtt{b})\mathbb{B}^\gamma_Y(s_0,\mathtt{b})\sum_{k\geq n}\frac{\varepsilon^k}{k!}\big(\mathbb{B}^\gamma_{X_{A_1}}(s_0,\mathtt{b})\big)^{k-1} + \mathbb{B}^\gamma_Y(s,\mathtt{b})\sum_{k\geq n}\frac{\varepsilon^k}{k!}\big(\mathbb{B}^\gamma_{X_{A_1}}(s_0,\mathtt{b})\big)^k \\ &\leq_s C(s,\mathtt{b},n)\mathbb{B}^\gamma_Y(s_0,\mathtt{b}) + C(s_0,\mathtt{b},n)\mathbb{B}^\gamma_Y(s,\mathtt{b}).\end{aligned} \tag{5.7.371}$$

In (5.7.368) there are terms of the form $\mathrm{ad}^k_{X_{A_1}}[Y]$, for some $k \geq 1$, with $Y = X_{\mathcal{K}_1}, X_{\mathcal{K}_2}, X_{\mathcal{K}_3}, \mathfrak{R}$ which belong to $\mathfrak{C}_{1,\mathtt{b}}$ by Lemmata 5.7.15 and 5.7.28.
We note that by (5.7.351)

$$\mathrm{ad}_{X_{A_1}}[\mathcal{D}_\omega - mJ] = -\varepsilon X_{\mathcal{K}_1} - \mathcal{D}_{\omega-\overline{\omega}}X_{A_1} - (m-1)[X_{A_1}, J] \in \mathfrak{C}_{1,\mathtt{b}}, \tag{5.7.372}$$

since $|\ell - \ell'| \leq C$, $j - j' \in S$, hence $\mathcal{D}_{\omega-\overline{\omega}}X_{A_1}, [X_{A_1}, J] \in \mathfrak{C}_{1,\mathtt{b}}$ (see the proof of Lemma 5.7.31). By (5.7.368), (5.7.370), (5.7.371), (5.7.372), (5.7.346), (5.7.347) and the fact that $|\omega - \overline{\omega}| \leq C\varepsilon^2$ we get the bounds (5.7.369). □

Step two (order ε^2)

The purpose of this section is to normalize the terms of size ε^2. In particular, we look for a symplectic change of cooordinates Υ_2 as the time-1 flow map of an Hamiltonian system

$$H_{A_2}(u) := \sum_{j,j'\in S^c}(A_2)^{j'}_j(\varphi)\,u_{j'}\,\overline{u}_j, \tag{5.7.373}$$

with A_2 a self-adjoint operator and thus it has the form

$$\Upsilon_2 := \exp(\varepsilon^2 JA_2) = \mathrm{I}_{H_S^\perp} + \varepsilon^2 JA_2 + \varepsilon^4 R_2, \quad R_2 := \sum_{k\geq 2} \frac{\varepsilon^{2(k-2)}}{k!} (JA_2)^k. \tag{5.7.374}$$

Then

$$\mathcal{K}^{(2)} := \mathcal{K}^{(1)} \circ \Upsilon_2^{-1} = \mathsf{H}_0 + \varepsilon^2 \mathcal{K}_2^{(2)} + \varepsilon^3 \mathcal{K}_3^{(2)} + o(\varepsilon^3), \qquad \mathcal{K}_2^{(2)} := \{H_{A_2}, \mathsf{H}_0\}_e + \mathcal{K}_2^{(1)}, \tag{5.7.375}$$

where by (5.7.366)

$$\mathcal{K}_2^{(1)}(u) = \mathcal{K}_2(u) + \frac{1}{2}\{H_{A_1}, \mathcal{K}_1\}(u) := \sum_{j,j'\in S^c} (\mathtt{B}_2)_j^{j'}(\varphi)\, u_{j'}\, \overline{u}_j. \tag{5.7.376}$$

Note that, by the definition of $\mathcal{K}_1, \mathcal{K}_2$ in (5.7.308) and of A_1 in (5.7.353), if the matrix element $(\mathtt{B}_2)_j^{j'}(\ell - \ell') \neq 0$ then $|j - j'| \leq 2C_S$, $|\ell - \ell'| \leq 2$.

Therefore we choose A_2 such that

$$(A_2)_j^{j'}(\ell) = \begin{cases} -\dfrac{(\mathtt{B}_2)_j^{j'}(\ell)}{\mathrm{i}(\overline{\omega}\cdot \ell + \omega(j) - \omega(j'))} & \text{if } \overline{\omega}\cdot\ell + \omega(j) - \omega(j') \neq 0, \\ & j, j' \in S^c,\ |j - j'| \leq 2C_S,\ |\ell| \leq 2, \\ 0 & \text{otherwise} \end{cases} \tag{5.7.377}$$

or equivalently

$$\mathcal{D}_{\overline{\omega}} H_{A_2} + \{H_{A_2}, \mathsf{H}_0\} = \mathcal{K}_2^{(1)} - \Pi_{\mathrm{Ker}(\mathsf{H}_0)} \mathcal{K}_2^{(1)}. \tag{5.7.378}$$

We write

$$(\mathtt{B}_2)_j^{j'}(\ell) = \sum_{\substack{j_1, j_2 \in S, \\ j_1 + j_2 = j - j'}} C_{j_1, j_2}^{(j,j')} \sqrt{\xi_{j_1} \xi_{j_2}} \tag{5.7.379}$$

and $(\mathtt{B}_2)_j^{j'}(\ell) \neq 0$ implies that $\ell = \mathtt{l}(j_1) + \mathtt{l}(j_2)$. Moreover, the support of $\Pi_{\mathrm{Ker}(\mathsf{H}_0)} \mathcal{K}_2^{(1)}$ is the set of $j, j' \in S^c$ such that $j_1 + j_2 + j - j' = 0$

$$\begin{aligned} \overline{\omega}\cdot(\mathtt{l}(j_1) + \mathtt{l}(j_2)) + \omega(j) - \omega(j') &= \omega(j_1) + \omega(j_2) + \omega(j) - \omega(j') \\ &= (j_1 + j_2)(j_1 + j)(j_2 + j)\, P(j_1, j_2, j) \end{aligned} \tag{5.7.380}$$

and P is the rational function

$$P(x, y, z) := \frac{3 + x^2 + y^2 + z^2 + xy + xz + yz + xyz(x + y + z)}{(1 + x^2)(1 + y^2)(1 + z^2)(1 + (x + y + z)^2)}. \tag{5.7.381}$$

We claim that $(j_1 + j_2)(j_1 + j)(j_2 + j)\, P(j_1, j_2, j) = 0$ if and only if $j_1 + j_2 = 0$. This claim is a consequence of Proposition 5.7.34. Therefore

$$\Pi_{\mathrm{Ker}(\mathsf{H}_0)} \mathcal{K}_2^{(1)} = \sum_{j\in S^c} \left(\sum_{j_2 \in S} \mathcal{C}(j, j_2)\, \xi_{j_2} \right) |u_j|^2 \tag{5.7.382}$$

where the $\mathcal{C}(j, j_2)$ are constants.

The following Proposition provides also a way to compute explicitly the coefficients $\mathcal{C}(j, j_2)$ in (5.7.382).

Proposition 5.7.34. *If Φ_{aa} is the symplectic change of coordinates* (5.3.10) *that puts a Hamiltonian in action-angle variables and $\Pi^{d_z=k}$ denotes the projection of a homogenous polynomial on the terms with degree k in the z variables, then*

$$\Pi_{Ker(\mathsf{H}_0)}\mathcal{K}_2^{(1)} = \Big[\Pi_{Ker(H^{(2)})}\Pi^{d_z=2}\Big(\frac{1}{2}\{H^{(3)}, F^{(3)}\}\Big)\Big] \circ \Phi_{aa_{|\{y=0\}}}. \tag{5.7.383}$$

Proof. By recalling the notations used in Section 3, we define the projector of a homogenous Hamiltonian of degree n on the monomials with degree less or equal than k in the normal variable z as

$$\Pi^{d_z\le k}H^{(n)} := H^{(n,\le k)}. \tag{5.7.384}$$

We recall the symplectic change of coordinates $\Phi_{aa}(\theta, y, z) = u$ defined in (5.3.10) and we denote by d_ξ, d_y, d_z the degree in the variables ξ, y, z respectively.
By Taylor expansion in $y = 0$ (recall (5.3.10)) we have that

$$u_j = \sqrt{\xi_j}e^{\mathrm{i}\theta_j}\sum_{k\ge0} c_k \left(\frac{\omega_j y_j}{\xi_j}\right)^k, \quad c_k \in \mathbb{R}, \qquad j \in S$$

hence the total degree of a monomial $R(v^{n-k}z^k)$ is given by $d = 2d_\xi + 2d_y + d_z$.

By (5.2.29) we have that, after two steps of weak Birkhoff normal form, the Hamiltonian of degree less or equal than 4 is

$$\Pi^{d\le4}H_4 = H^{(2)} + H^{(3,\ge2)} + H_4^{(4,0)} + H_4^{(4,\ge2)}, \quad H_4^{(4,0)} := \Pi_{\mathrm{Ker}(H^{(2)})}H_3^{(4,0)}. \tag{5.7.385}$$

We look for the correction at order $O(\varepsilon^2)$ of the eigenvalues of the quadratic Hamiltonian in z. By (5.3.14) the monomials of degree greater than 4 are not involved in this computation. We define

$$K := \Pi^{d\le4}H_4 \circ \Phi_{aa}, \tag{5.7.386}$$

in particular

$$\begin{aligned} &K^{(2)} := H^{(2)} \circ \Phi_A := \overline{\omega}\cdot y + \sum_{j\in S^c} z_j z_{-j}, \qquad K_{res}^{(4,0)} := H_4^{(4,0)} \circ \Phi_{aa} = \mathbb{A}(\xi+y)\cdot y, \\ &K^{(3,2)} := H^{(3,2)} \circ \Phi_{aa_{|y=0}}, \qquad K^{(4,2)} := H_4^{(4,2)} \circ \Phi_{aa_{|y=0}}. \end{aligned} \tag{5.7.387}$$

We want to diagonalize $K^{(3,2)} + K^{(4,2)}$, we can ignore the terms $R(v^{n-k}z^k)$ with $k \ge 3$, since we will use changes of coordinates which preserve the degree in z.

The strategy is the following:

(i) we apply *any* transformation generated by an Hamiltonian of the form $S^{(3,2)}(\xi, \theta, z)$ such that the flow is well defined, regular and $d_\xi \ge 1/2$;

(ii) we perform others steps of Birkhoff normal form which diagonalizes the hamiltonian quadratic in z.

We start by applying the flow of a Hamiltonian independent of y

$$S^{(3,2)}(\xi, \theta, z) \tag{5.7.388}$$

with degree $d_\xi = 1/2$ and flow at time 1

$$(\Phi^\tau_{S^{(3,2)}})_{|\tau=1} := \Phi_1.$$

Actually this kind of Hamiltonian is like S_1 defined in (5.7.276) considered in the first preliminary step (Section 5.7.4). We remark that the flow of this Hamiltonian is well defined and smooth.

We evaluate now the terms of the conjugate hamiltonian with $d_\xi \le 1$ and $2d_y + d_z = 2$. We have

$$\begin{aligned}(\Phi_1)_* K = K^{(2)} + K^{(3,2)} + K^{(4,2)} + \{K^{(2)}, S^{(3,2)}\} + \{K^{(3,2)}, S^{(3,2)}\} \\ + \frac{1}{2}\{\{K^{(2)}, S^{(3,2)}\}, S^{(3,2)}\} + K^{(4,0)}_{res} + O(\varepsilon^3).\end{aligned} \tag{5.7.389}$$

We define

$$G^{(3,2)} := -\big[\mathrm{ad}K^{(2)}\big]^{-1}\Big(K^{(3,2)} + \{K^{(2)}, S^{(3,2)}\}\Big) \tag{5.7.390}$$

and we conjugate the Hamiltonian $(\Phi_1)_* K$ in (5.7.389) through the flow at time one of $G^{(3,2)}$. We call

$$\Upsilon_1 := (\Phi^\tau_{G^{(3,2)}})_{|\tau=1}.$$

Actually this transformation can be considered as the map defined in (5.7.349), Section 8.4. We have

$$\begin{aligned}(\Upsilon_1)_*(\Phi_1)_* K = K^{(2)} + K^{(4,0)}_{res} + \{K^{(3,2)}, S^{(3,2)} + G^{(3,2)}\} + K^{(4,2)} \\ + \frac{1}{2}\{\{K^{(2)}, S^{(3,2)} + G^{(3,2)}\}, S^{(3,2)} + G^{(3,2)}\} \\ + \frac{1}{2}\{K^{(2)}, \{S^{(3,2)}, G^{(3,2)}\}\} + O(\varepsilon^3).\end{aligned} \tag{5.7.391}$$

We define

$$\tilde F^{(3)} := S^{(3,2)} + G^{(3,2)} \overset{(5.7.390)}{=} -[\mathrm{ad}K^{(2)}]^{-1}(K^{(3,2)}). \tag{5.7.392}$$

At order ε^2 we have

$$\begin{aligned}\Pi^{d_z=2}\Pi_{\mathrm{Ker}(K^{(2)})}(\Upsilon_1)_*(\Phi_1)_* K = \Pi^{d_z=2}K^{(2)} + K^{4,0}_{res} + \Pi_{\mathrm{Ker}(K^{(2)})}K^{(4,2)} \\ + \frac{1}{2}\Pi_{\mathrm{Ker}(K^{(2)})}\{S^{(3,2)} + G^{(3,2)}, K^{(3,2)}\}.\end{aligned} \tag{5.7.393}$$

Now we want to show that the Partial Birkhoff normal form procedure, that is pure formal, gives the same corrections to the eigenvalues at order ε^2 of the quadratic (in z) Hamiltonian.
In order to do that we consider the first Birkhoff transform that eliminates the monomials $R(vz^2)$ in the Hamiltonian H

$$F^{(3,2)} := [\mathrm{ad}H^{(2)}]^{-1}H^{(3,2)}. \tag{5.7.394}$$

Then by the definition of $\tilde F^{(3)}$ (see (5.7.392))

$$\{H^{(2)} \circ \Phi_{aa}, F^{(3,2)} \circ \Phi_{aa}\}_{|y=0} = (H^{(3,2)} \circ \Phi_{aa})_{|y=0} = K^{(3,2)}, \tag{5.7.395}$$

since Φ_{aa} is symplectic. Moreover we have that

$$\Pi^{d_z=2}\{H^{(3,2)}, F^{(3,2)}\} \circ (\Phi_A)_{|y=0} = \{K^{(3,2)}, \tilde F^{(3,2)}\}, \tag{5.7.396}$$

Hence we rewrite (5.7.393) as

$$\begin{aligned}&\Pi^{d_z=2}K^{(2)}+K^{(4,2)}_{res}+\Pi_{\mathrm{Ker}(K^{(2)})}\frac{1}{2}\{K^{(3,2)},\tilde F^{(3,2)}\}=\\&=\Big(\Pi^{(d_z=2)}\Pi_{\mathrm{Ker}}\Big[H^{(2)}+H_4+\frac12\{H^{(3,2)},F^{(3,2)}\}\Big]\Big)\circ\Phi_{aa|_{y=0}}.\end{aligned}\tag{5.7.397}$$

We claim now that on the diagonal, at order $O(\varepsilon^2)$ we have

$$\Big(\Pi_{\mathrm{Ker}(H^{(2)})}\Pi^{(d_z=2)}\Big[H^{(4)}+\frac12\{H^{(3)},F^{(3)}\}\Big]\Big)\circ\Phi_{aa|_{y=0}},\tag{5.7.398}$$

where

$$F^{(3)}:=-[\mathrm{ad}H^{(2)}]^{-1}H^{(3)}\tag{5.7.399}$$

is the generator of the first transformation of the Full Birkhoff normal form that removes completely $H^{(3)}$.
Since in the original system (5.2.1) the term $H^{(4)}$ is zero, then (5.7.398) implies the thesis.

In order to prove (5.7.398) we evaluate the term H_4 in (5.7.397). We have

$$\begin{aligned}H_4&:=H^{(4)}+\{H^{(3)},F^{(3,\le1)}\}-\frac12\{H^{(3,\le1)},F^{(3,\le1)}\}\\&=H^{(4)}+\frac12\{H^{(3,\le1)},F^{(3,\le1)}\}+\{H^{(3,>1)},F^{(3,\le1)}\}.\end{aligned}\tag{5.7.400}$$

It is important to note that

$$\Pi^{(d_z=2)}\{F^{(3,\le1)},H^{(3,>1)}\}=\Pi^{(d_z=2)}(\{F^{(3,0)},H^{(3,2)}\}+\{F^{(3,1)},H^{(3,3)}\}),\tag{5.7.401}$$

but also

$$\Pi_{\mathrm{Ker}H^{(2)}}\{F^{(3,1)},H^{(3,3)}\}=0.\tag{5.7.402}$$

In the same way

$$\begin{aligned}&\Pi_{\mathrm{Ker}(H^{(2)})}\Pi^{(d_z=2)}\{F^{(3,3)},H^{(3)}\}=\Pi_{\mathrm{Ker}(H^{(2)})}\Pi^{(d_z=2)}\{F^{(3,3)},H^{(3,1)}\}=0,\\&\Pi_{\mathrm{Ker}(H^{(2)})}\{F^{(3,2)},H^{(3,0)}\}=0;\quad \Pi_{\mathrm{Ker}(H^{(2)})}\{H^{(3,2)},F^{(3,0)}\}=0.\end{aligned}\tag{5.7.403}$$

This implies that

$$\begin{aligned}\Pi_{\mathrm{Ker}(H^{(2)})}\Pi^{(d_z=2)}\{F^{(3)},H^{(3)}\}&=\Pi_{\mathrm{Ker}(H^{(2)})}\Pi^{(d_z=2)}\Big(\{H^{(3,\le1)},F^{(3,\le1)}\}+\{H^{(3,>1)},F^{(3,\le1)}\}\Big)\\&+\Pi_{\mathrm{Ker}(H^{(2)})}\Pi^{(d_z=2)}\Big(\{H^{(3,0)},F^{(3,2)}\}+\{H^{(3,2)},F^{(3,2)}\}+\{H^{(3,1)},F^{(3,3)}\}\Big)\\&=\Pi_{\mathrm{Ker}(H^{(2)})}\Pi^{(d_z=2)}\Big(\{H^{(3,\le1)},F^{(3,\le1)}\}+\{H^{(3,2)},F^{(3,2)}\}\Big).\end{aligned}\tag{5.7.404}$$

This implies that equation (5.7.398) is equivalent to (5.7.397). □

By (5.7.383), (5.2.16), (5.2.1) we have

$$\Pi_{\mathrm{Ker}(\mathsf{H}_0)}\mathcal{K}_2^{(1)}=\frac12\sum_{j\in S^c}\Big(\sum_{j_2\in S}\frac{\omega(j_2+j)}{\omega(j_2)+\omega(j)-\omega(j_2+j)}\,\xi_{j_2}\Big)|u_j|^2\tag{5.7.405}$$

and by (5.3.7)

$$\sum_{j_2\in S}\frac{\omega(j_2+j)}{\omega(j_2)+\omega(j)-\omega(j_2+j)}\,\xi_{j_2}=\frac{2}{3}\sum_{j_2\in S^+}\frac{(1+j_2^2)(1+j^2)(2+j_2^2+j^2)}{(3+j_2^2-j_2j+j^2)(3+j_2^2+j_2j+j^2)}\xi_{j_2}. \tag{5.7.406}$$

We define

$$\lambda_j:=\frac{2}{3}\sum_{j_2\in S^+}\frac{(1+j_2^2)(1+j^2)(2+j_2^2+j^2)}{(3+j_2^2-j_2j+j^2)(3+j_2^2+j_2j+j^2)}\xi_{j_2} \tag{5.7.407}$$

and the diagonal operator (recall (5.7.299))

$$\mathfrak{D}:=\mathfrak{D}(\xi)=\mathrm{diag}\,(\mathrm{i}\kappa_j)_{j\in S^c},\qquad \kappa_j=\omega(j)\big(\lambda_j-c(\xi)\big)\in\mathbb{R}. \tag{5.7.408}$$

Lemma 5.7.35. *The operator $\mathfrak{D}(\xi)\in\mathfrak{C}_{1,\mathtt{b}}$ and its eigenvalues satisfies, for some $C>0$,*

$$|j|\,|\kappa_j|\le C\qquad\forall j\in S^c,\quad \kappa_j=-\kappa_{-j}. \tag{5.7.409}$$

Proof. By (5.7.299) and (5.7.407) we have

$$\lambda_j-c(\xi)=-\frac{2}{3}\sum_{j_2\in S^+}\frac{(1+j_2^2)(7+5j_2^2+j_2^4+3j^2)}{(3+j_2^2-j_2j+j^2)(3+j_2^2+j_2j+j^2)}\xi_{j_2}=\frac{P(j)}{Q(j)}$$

where $P(j)$ is a polynomial of degree $4\nu-2$ in j and $Q(j)$ is a polynomial of degree 4ν (recall (1.2.8)).
Then there exist a constant $\tilde{C}>0$ such that $|\lambda_j-c(\xi)|\le\tilde{C}/j^2$. By the definition of κ_j and the fact that $|\omega(j)|\le 3|j|$ we get (5.7.409). By the fact that $\mathfrak{D}(\xi)=\mathrm{diag}_{j\in S^c}(\mathfrak{D}(\xi))_j^j(0)$ we have

$$\|\langle D_x\rangle^{1/2}\mathfrak{D}(\xi)\langle D_x\rangle^{1/2}h\|_s^2\le\sum_{j\in S^c}\langle j\rangle^{2s}\sum_{\ell'\in\mathbb{Z}^\nu}\langle j\rangle|\kappa_j|^2\langle j\rangle|h_{\ell',j}|^2\overset{(5.7.409)}{\le}C\|h\|_s^2.$$

Note that $\mathfrak{D}(\xi)$ does not depend on φ and $[\mathfrak{D}(\xi),\partial_x]=0$ since $\mathfrak{D}(\xi)$ is diagonal. This concludes the proof. □

In the next lemma we provide a bound for the denominators in (5.7.377). Note that if (j_1,j_2,j) are such that $P(j_1,j_2,j)=0$ then by Proposition 5.7.34 the numerator in (5.7.377) is naught.

Lemma 5.7.36. *If $j_1,j_2\in S$, $\ell=\mathtt{l}(j_1)+\mathtt{l}(j_2)$, $j,j'\in S^c$, $|j-j'|\le 2C_S$ and (j,j_1,j_2) are such that $P(j_1,j_2,j)\neq 0$, then*

$$|\overline{\omega}\cdot\ell+\omega(j')-\omega(j)|\ge K_2 \tag{5.7.410}$$

for some constant $K_2=K_2(S)>0$ dependent only by the set S.

Proof. If $|j|>N$, where $N=N(S)$ is a large constant to be fixed and which depends on the set S, then, recalling (5.7.380) and (5.7.381)

$$|(j_1+j_2)(j_1+j)(j_2+j)|\ge C_1\,j^2$$

for some constant $C_1:=C_1(N)>0$ (possibly small). Moreover

$$|(1+j_1j_2)j^2+(j_1^2j_2+j_1j_2^2)j+3+j_1^2+j_2^2|\ge C_2\,j^2,$$

$$|(1+j_1^2)(1+j_2^2)(1+j^2)(1+(j_1+j_2+j)^2| \leq C_3\, j^4$$

for some small constant $C_2 := C_2(N) > 0$ and some large constant $C_3 := C_3(N) > 0$. Thus

$$|\overline{\omega}\cdot\ell + \omega(j') - \omega(j)| \geq \frac{C_1 C_2}{C_3} > 0.$$

Now consider $|j| \leq N$. For (j_1, j_2, j) belonging to the compact set $S \times S \times \{j \in \mathbb{Z} : |j| \leq N\}$ we have

$$|(j_1+j_2)(j_1+j)(j_2+j)||(1+j_1 j_2)j^2 + (j_1^2 j_2 + j_1 j_2^2)j + 3 + j_1^2 + j_2^2| \geq M_1$$
$$|(1+j_1^2)(1+j_2^2)(1+j^2)(1+(j_1+j_2+j)^2| \leq M_2$$

for some constant $M_1, M_2 > 0$ dependent on S. Set $C_4 := M_1/M_2$.
Therefore $|\overline{\omega}\cdot\ell + \omega(j') - \omega(j)| \geq C_4 > 0$. Now take $K := \max\{C_4, C_1C_2/C_3\}$. □

Remark 5.7.37. Let $A = \mathrm{Op}(a)$ be a pseudo differential operator of order $m \in \mathbb{R}$. Then

$$Au = \sum_{j,j' \in S^c, \ell,\ell' \in \mathbb{Z}^\nu} (A)_j^{j'}(\ell-\ell') u_{\ell' j'} e^{\mathrm{i}(jx+\ell\cdot\varphi)} = \sum_{j,j' \in S^c, \ell,\ell' \in \mathbb{Z}^\nu} a(j-j', j', \ell-\ell') u_{\ell' j'} e^{\mathrm{i}(jx+\ell\cdot\varphi)}.$$

We know that there exists a constant $C > 0$ such that

$$\|a(x,\xi)\|_s \langle\xi\rangle^m \leq |a|_{m,s,0} \leq C < +\infty.$$

Hence for any $j, j' \in S^c$ we have

$$|(A)_j^{j'}(\ell)| = |a(j-j', j', \ell)| \leq \frac{C}{\langle j'\rangle^m} \quad \forall \ell \in \mathbb{Z}^\nu.$$

Lemma 5.7.38. *For $j, j' \in S^c$, $|j-j'| \leq 2C_S$, $|\ell| \leq 2$ we have that (recall* (5.7.376), (5.7.377)*)*

$$|(A_2)_j^{j'}(\ell)| \leq \frac{C}{|j'|^2} \tag{5.7.411}$$

for some constant $C > 0$. Otherwise we have that $(A_2)_j^{j'}(\ell) = 0$.

Proof. We have

$$(\mathrm{ad}_{X_{A_1}}[X_{\mathcal{K}_1}])_j^{j'}(\ell) = \sum_{k \in S^c} \omega(j)(A_1)_j^k(\ell)(X_{\mathcal{K}_1})_k^{j'}(\ell) - \omega(j')(X_{\mathcal{K}_1})_j^k(\ell)(A_1)_k^{j'}(\ell). \tag{5.7.412}$$

We have for any $k \in S^c$ that $|j-k|, |j'-k| \leq 2C_S$ by (5.7.377), hence the sum in (5.7.412) is finite and the number of the summands does not depend on j, j'. We know that $X_{\mathcal{K}_1}$ and $X_{\mathcal{K}_2}$ are pseudo differential operators of order -1, hence by Remark 5.7.37 and (5.7.411)

$$\begin{aligned} &|(\mathrm{ad}_{X_{A_1}}[X_{\mathcal{K}_1}])_j^{j'}(\ell)| \leq \tilde{C} \sum_k \frac{|j|}{|j'|^2|j|} \leq \frac{C}{|j'|^2}, \\ &|(X_{\mathcal{K}_2})_j^{j'}(\ell)| \leq \frac{C}{|j'|^2}, \end{aligned} \tag{5.7.413}$$

for some constant $C > 0$. By (5.7.376) and (5.7.410) we conclude. □

Lemma 5.7.39. *The linear vector field $X_{A_2} := JA_2$ belongs to the class $\mathfrak{C}_{1,\mathtt{b}}$, in particular it satisfies the following*

$$\mathbb{B}^{\gamma}_{\varepsilon^2 X_{A_2}}(s,\mathtt{b}) \le C(s,\mathtt{b})\varepsilon^2. \tag{5.7.414}$$

Note that X_{A_2} does not depend on $i(\omega)$.

Proof. The proof follows by the same arguments of the proof of Lemma 5.7.31. The key points are the estimate (5.7.411) on the modulus of the coefficients of the matrix A_2 (defined in (5.7.377)), which is similar to (5.7.355). By the fact that for some constant $C>0$

$$\langle \ell_m - \ell'_m\rangle^{\mathtt{b}}, |j-j'| \le C$$

we obtain an estimate like (5.7.355). ☐

By Lemma 5.7.15 and (5.7.414) we deduce the following result.

Corollary 5.7.40. *The transformation $\Upsilon_2 \colon H^s(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$ defined in (5.7.374) is invertible and satisfies, for any $u = u(\omega) \in H^s$ Lipschitz in $\omega \in \mathcal{O}^{2\gamma}_{\infty}$,*

$$\|(\Upsilon_2^{\pm 1} - \mathrm{I})h\|_s^{\gamma,\mathcal{O}^{2\gamma}_{\infty}} \le_s \varepsilon^2 C(s_0,\mathtt{b})\|u\|_s^{\gamma,\mathcal{O}^{2\gamma}_{\infty}} + \varepsilon^2 C(s,\mathtt{b})\|u\|_{s_0}^{\gamma,\mathcal{O}^{2\gamma}_{\infty}}. \tag{5.7.415}$$

Proof. The proof is analogous of the proof of Corollary 5.7.32. In this case we use Lemma 5.7.39 instead of Lemma 5.7.31. ☐

Lemma 5.7.41. *The transformed operator is (recall (5.7.365))*

$$\mathcal{L}_8 := \Upsilon_2 \mathcal{L}_7 \Upsilon_2^{-1} = \Pi_S^{\perp}\Big(\mathcal{D}_{\omega} - mJ - \varepsilon^2\mathfrak{D}(\xi) - \varepsilon^3 X_{\mathcal{K}_3^{(2)}} + \mathcal{R}_8\Big) \tag{5.7.416}$$

where $\mathcal{K}_3^{(2)} = \mathcal{K}_3^{(1)}$, $\mathfrak{D}(\xi)$ is the diagonal operator of order -1 defined in (5.7.408) and

$$\mathcal{R}_8 := \mathcal{R}_7 + \varepsilon^2 \mathrm{ad}_{X_{A_2}}[-\varepsilon^2\mathfrak{D}(\xi) - \varepsilon^3 X_{\mathcal{K}_3^{(2)}} + \mathcal{R}_8] + \sum_{k\ge 2}\frac{\varepsilon^{2k}}{k!}\mathrm{ad}_{X_{A_2}}[\mathcal{L}_7]. \tag{5.7.417}$$

Moreover the following holds. The operator $\mathcal{R}_8 \in \mathfrak{C}_{1,\mathtt{b}}$ with

$$\begin{aligned} &\mathbb{B}^{\gamma}_{\mathcal{R}_8}(s,\mathtt{b}) \le_s \varepsilon^{4-3a} + \varepsilon\gamma^{-1}\|\mathfrak{I}_{\delta}\|^{\gamma,\mathcal{O}_0}_{s+\tilde{\sigma}}, \\ &\mathbb{B}_{\Delta_{12}\mathcal{R}_8}(s,\mathtt{b}) \le_s \varepsilon\gamma^{-1}(\|i_1 - i_2\|_{s+\tilde{\sigma}} + \|\mathfrak{I}_{\delta}\|_{s+\tilde{\sigma}}\|i_1 - i_2\|_{s_0+\tilde{\sigma}}), \end{aligned} \tag{5.7.418}$$

for some $\tilde{\sigma}$ possibly larger than the one in Lemma 5.7.33.

Proof. The proof follows by using the same arguments of the proof of Lemma 5.7.33, in particular we use the bounds (5.7.370), (5.7.371) and the fact that, by (5.7.378) and Proposition 5.7.34,

$$\mathrm{ad}_{X_{A_2}}[\mathcal{D}_{\omega} - mJ] + \mathcal{K}_2^{(1)} = -\mathfrak{D}(\xi) - \mathcal{D}_{\omega-\overline{\omega}}X_{A_2} - (m-1)[X_{A_2}, J].$$

By (5.7.35) $\mathfrak{D}(\xi) \in \mathfrak{C}_{1,\mathtt{b}}$ and by the fact that $|\ell - \ell'| \le C$, $|j - j'| \le 2C_S$ (see (5.7.377)) also $\mathcal{D}_{\omega-\overline{\omega}}X_{A_2}, [X_{A_2}, J] \in \mathfrak{C}_{1,\mathtt{b}}$. The bounds (5.7.418) are obtained by (5.7.415) and the estimates for $\mathcal{R}_7$ in (5.7.369). ☐

Step three (order ε^3)

The purpose of this section is to eliminate the terms of size ε^3. In particular, we look for a symplectic change of cooordinates Υ_3 as the time-1 flow map of an Hamiltonian system

$$H_{A_3}(u) := \sum_{j,j'\in S^c} (A_3)_j^{j'}(\varphi)\, u_{j'}\, \overline{u}_j, \tag{5.7.419}$$

with A_3 a self-adjoint operator and thus it has the form

$$\Upsilon_3 := \exp(\varepsilon^3 JA_3) = \mathrm{I}_{H_S^\perp} + \varepsilon^3 JA_3 + \varepsilon^6 R_3, \quad R_3 := \sum_{k\geq 2} \frac{\varepsilon^{3(k-2)}}{k!}(JA_3)^k. \tag{5.7.420}$$

Then

$$\begin{aligned} \mathcal{K}^{(3)} &:= \mathcal{K}^{(2)} \circ \Upsilon_3^{-1} = \mathsf{H}_0 + \varepsilon^2 \mathcal{K}_2^{(2)} + \varepsilon^3 \mathcal{K}_3^{(3)} + o(\varepsilon^3), \\ \mathcal{K}_3^{(3)} &:= \{H_{A_3}, \mathsf{H}_0 + \varepsilon^2 \mathbb{A}\xi\cdot\eta + \frac{\varepsilon^2}{2}\sum_{j\in S^c} \lambda_j(\xi) z_j\, z_{-j}\}_e + \mathcal{K}_3^{(2)}. \end{aligned} \tag{5.7.421}$$

Note that we consider in the normal form also the ε^2-terms. We define the matrix B_3 in the following way (recall (5.7.367))

$$\mathcal{K}_3^{(1)}(u) = \mathcal{K}_3^{(2)}(u) := \sum_{j,j'\in S^c} (\mathsf{B}_3)_j^{j'}(\varphi)\, u_j\, \overline{u}_{j'}. \tag{5.7.422}$$

Note that, by the definition of $\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3$ in (5.7.308) and (5.7.367), if the matrix element $(\mathsf{B}_3)_{j'}^{j}(\ell - \ell') \neq 0$ then $|j-j'| \leq 3C_S$, $|\ell - \ell'| \leq 3$.

We choose A_3 as

$$(A_3)_j^{j'}(\ell) = \begin{cases} -\dfrac{(\mathsf{B}_3)_j^{j'}(\ell)}{\mathrm{i}\delta_{\ell j j'}} & \text{if } \overline{\omega}\cdot\ell + \omega(j) - \omega(j') \neq 0, \\ & j, j' \in S^c,\ |j-j'| \leq 3C_S,\ |\ell| \leq 3, \\ 0 & \text{otherwise} \end{cases} \tag{5.7.423}$$

where

$$\delta_{\ell j j'} := \overline{\omega}\cdot\ell + \varepsilon^2 \mathbb{A}\xi\cdot\ell + \omega(j') - \omega(j) + \varepsilon^2(\omega(j')\lambda_{j'} - \omega(j)\lambda_j) \tag{5.7.424}$$

so that (recall (5.7.407))

$$\mathcal{D}_{\overline{\omega}+\varepsilon^2\mathbb{A}\xi} H_{A_3} + \{H_{A_3}, \mathsf{H}_0 + \frac{\varepsilon^2}{2}\sum_{j\in S^c}\lambda_j(\xi) z_j\, z_{-j}\} + \mathcal{K}_3^{(2)} = 0. \tag{5.7.425}$$

Note that A_3 is self-adjoint since B_3 is self-adjoint and $\overline{\mathrm{i}\delta_{\ell j j'}} = \mathrm{i}\delta_{-\ell j' j}$.
We recall that $\omega \in \mathcal{G}_0$ (see (5.4.5), (5.4.4)), hence by Remark 5.7.29 we have the bound

$$|\overline{\omega}\cdot\ell + \varepsilon^2\mathbb{A}\xi\cdot\ell + \omega(j') - \omega(j) + \varepsilon^2(\omega(j')\lambda_{j'} - \omega(j)\lambda_j)| \geq C\gamma. \tag{5.7.426}$$

Lemma 5.7.42. *For $j, j' \in S^c$, $|j-j'| \leq 3C_S$, $|\ell| \leq 3$ we have that (recall* (5.7.422), (5.7.423)*)*

$$|(A_3)_j^{j'}(\ell)| \leq \frac{C\,\gamma^{-1}}{|j'|^2} \tag{5.7.427}$$

for some constant $C > 0$. Otherwise we have that $(A_3)_j^{j'}(\ell) = 0$.

Proof. Recall the definition of A_3 in (5.7.423). First we estimate the matrix entries of the numerator $\mathtt{B}_3$, see (5.7.422). The following bound on the modulus of the matrix entries of the operators $X_{\mathcal{K}_3} + \mathrm{ad}_{X_{A_1}}[X_{\mathcal{K}_2}]$

$$|(X_{\mathcal{K}_3} + \mathrm{ad}_{X_{A_1}}[X_{\mathcal{K}_2}])_j^{j'}(\ell)| \leq \frac{C}{|j'|^2}$$

is obtained as in the proof of Lemma 5.7.38. Let B be an operator with a Fourier representation as matrix $B_j^{j'}(\ell)$. We omit the index ℓ of the time. We have

$$(\mathrm{ad}^2_{X_{A_1}}[B])_j^{j'} = \sum_{k,k'\in S^c} (A_1)_j^k (A_1)_k^{k'} B_{k'}^{j'} - (A_1)_j^k B_k^{k'} A_{k'}^{j'} + B_j^{k'}(A_1)_{k'}^k (A_1)_k^{j'} - (A_1)_j^{k'} B_{k'}^k (A_1)_k^{j'}. \quad (5.7.428)$$

Then by (5.7.423) $|j-k|, |j-k'|, |k'-k|, |j'-k| \leq 3C_S$, hence the sum above is finite and the number of summands is independent of j and j'.
By (5.7.372) we have that

$$\mathrm{ad}^3_{X_{A_1}}[\mathcal{D}_\omega - mJ] = -\mathrm{ad}^2_{X_{A_1}}[\varepsilon X_{\mathcal{K}_1} + \mathcal{D}_{\omega-\overline{\omega}} X_{A_1}].$$

Hence to estimate the remaining terms of $X_{\mathcal{K}_3^{(2)}}$ we use (5.7.428), the fact that $X_{\mathcal{K}_1}$ is a pseudo differential operator of order -1 (see Remark 5.7.37) and the same arguments of Lemma 5.7.38. Thus the entry (j, j', ℓ) of the numerator is bounded by a pure constant divided by $|j'|^2$. By (5.7.426) we conclude. □

Thus we have the following result.

Lemma 5.7.43. *The linear vector field $X_{A_3} := JA_3$ belongs to the class $\mathfrak{C}_{1,\mathtt{b}}$, in particular it satisfies the following*

$$\mathbb{B}^\gamma_{\varepsilon^3 X_{A_3}}(s, \mathtt{b}) \leq C(s, \mathtt{b})\varepsilon^3\gamma^{-1} = C(s, \mathtt{b})\varepsilon^{1-a}. \quad (5.7.429)$$

Note that X_{A_3} does not depend on $i(\omega)$.

Proof. The bound for the tame constant of the operator $\langle D_x\rangle^{1/2} X_{A_3} \langle D_x\rangle^{1/2}$ is obtained as in Lemma 5.7.31 and by using the bound (5.7.427).
The matrix elements of $\partial^{\mathtt{b}}_{\varphi_m} X_{A_3}$, $[X_{A_3}, \partial_x]$, $[\partial^{\mathtt{b}}_{\varphi_m} X_{A_3}, \partial_x]$ are respectively $\langle \ell_m - \ell'_m\rangle^{\mathtt{b}} \omega(j)(A_3)_j^{j'}(\ell - \ell')$, $(j-j')\omega(j)(A_3)_j^{j'}(\ell-\ell')$, $\langle \ell_m - \ell'_m\rangle^{\mathtt{b}}(j-j')\omega(j)(A_3)_j^{j'}(\ell-\ell')$. Note that by the definition of A_3 in (5.7.423)

$$\langle \ell_m - \ell'_m\rangle^{\mathtt{b}}, |j-j'| \leq C$$

for some constant C depending on the set S. Thus arguing as done for the bound of the tame constant of $\langle D_x\rangle^{1/2} X_{A_3} \langle D_x\rangle^{1/2}$ and by using (5.7.427) one can easily prove that $\partial^{\mathtt{b}}_{\varphi_m} X_{A_3}$, $[X_{A_3}, \partial_x]$, $[\partial^{\mathtt{b}}_{\varphi_m} X_{A_3}, \partial_x]$ are -1-Lip-tame operators with constant $\mathbb{B}^\gamma(s, \mathtt{b})$ given in (5.7.429). □

Corollary 5.7.44. *The transformation $\Upsilon_3 \colon H^s(\mathbb{T}^{\nu+1}) \to H^s(\mathbb{T}^{\nu+1})$ defined in (5.7.420) is invertible and satisfies, for any $u = u(\omega) \in H^s$ Lipschitz in $\omega \in \mathcal{O}^{2\gamma}_\infty$,*

$$\|(\Upsilon_3^{\pm 1} - \mathrm{I})h\|_s^{\gamma, \mathcal{O}^{2\gamma}_\infty} \leq_s \varepsilon^{1-a} C(s_0, \mathtt{b}) \|u\|_s^{\gamma, \mathcal{O}^{2\gamma}_\infty} + \varepsilon^{1-a} C(s, \mathtt{b}) \|u\|_{s_0}^{\gamma, \mathcal{O}^{2\gamma}_\infty}. \quad (5.7.430)$$

Proof. The proof is analogous of the proof of Corollary 5.7.32. In this case we use Lemma 5.7.43 instead of Lemma 5.7.31. □

Lemma 5.7.45. *The transformed operator is (recall* (5.7.365)*)*

$$\mathcal{L}_9 := \Upsilon_3 \mathcal{L}_8 \Upsilon_3^{-1} = \Pi_S^{\perp}\Big(\mathcal{D}_\omega - mJ - \varepsilon^2 \mathfrak{D}(\xi) + \mathcal{R}_9\Big) \tag{5.7.431}$$

where $\mathfrak{D}(\xi)$ *is the diagonal operator of order* -1 *defined in* (5.7.408) *and*

$$\mathcal{R}_9 := \mathcal{R}_8 + \varepsilon^3 \mathrm{ad}_{X_{A_3}}[-\varepsilon^3 X_{\mathcal{K}_3^{(2)}} + \mathcal{R}_8] + \sum_{k\geq 2}\frac{1}{k!}\mathrm{ad}^k_{X_{A_3}}[\mathcal{L}_8]. \tag{5.7.432}$$

Moreover the following holds. The operator $\mathcal{R}_9 \in \mathfrak{C}_{1,\mathtt{b}}$ *with*

$$\begin{aligned} \mathbb{B}^\gamma_{\mathcal{R}_9}(s,\mathtt{b}) &\leq_s \varepsilon^{4-3a} + \varepsilon\gamma^{-1}\|\mathfrak{I}_\delta\|^{\gamma,\mathcal{O}_0}_{s+\tilde\sigma}, \\ \mathbb{B}_{\Delta_{12}\mathcal{R}_9}(s,\mathtt{b}) &\leq_s \varepsilon\gamma^{-1}(\|i_1-i_2\|_{s+\tilde\sigma} + \|\mathfrak{I}_\delta\|_{s+\tilde\sigma}\|i_1-i_2\|_{s_0+\tilde\sigma}), \end{aligned} \tag{5.7.433}$$

for some $\tilde\sigma$ *possibly larger than the one in Lemma 5.7.41.*

Proof. The proof follows the same arguments used for proving Lemma 5.7.33. We note that by (5.7.425) we have (recall (5.3.19) and (5.7.299))

$$\mathrm{ad}_{X_{A_3}}[\mathcal{D}_\omega - mJ - \varepsilon^2\mathfrak{D}(\xi)] = \varepsilon^3 X_{\mathcal{K}_3^{(2)}} + D_{\omega-\overline{\omega}-\varepsilon^2\mathbb{A}\xi}A_3 - (m-1-\varepsilon^2 c(\xi))[X_{A_3}, J] \in \mathfrak{C}_{1,\mathtt{b}},$$

since $|\ell-\ell'|\leq C$ and $|j-j'|\leq C$ (see (5.7.423)). Hence the bounds (5.7.433) follows by (5.7.430) and the estimates for $\mathcal{R}_8$ (see (5.7.418)). □

Proof of Theorem 5.7.3. We choose σ as $\tilde\sigma$ in (5.7.433). Define the map (recall (5.7.343), (5.7.349), (5.7.374), (5.7.420))

$$\Upsilon := \Upsilon_3 \circ \Upsilon_2 \circ \Upsilon_1 \circ \Phi.$$

By (5.7.281), (5.7.300), (5.7.314), (5.7.323), (5.7.363), (5.7.415), (5.7.430), (5.7.329), (5.7.326), Lemma 5.7.18 we have (5.7.17).
The result follows by setting $\mathcal{L} := \mathcal{L}_9$ (see (5.7.431)), $m := m$, $\mathcal{R} := \mathcal{R}_9$ and by noting that (5.7.433) implies (5.7.16), (5.7.331) implies (5.7.11) and (5.7.15). □

5.8 KAM reducibility scheme

We introduce the following parameters

$$\tau = 2\nu + 6, \quad \mathtt{b}_0 := 6\tau + 6, \quad \mathtt{b} = \mathtt{b}_0 + s_0. \tag{5.8.1}$$

The aim of this section is to prove the following theorem.

Theorem 5.8.1. ***(Reducibility)*** *Let* $\gamma_* := \gamma^{3/2}$ *(see* (5.4.6)*). Assume that* $\omega \mapsto i_\delta(\omega)$ *is a Lipschitz function defined on some subset* $\mathcal{O}_0 \subseteq \Omega_\varepsilon$ *(recall* (5.4.2)*), satisfying* (5.6.7) *with* $\mu' \geq \sigma$ *where* $\sigma := \sigma(\nu)$ *is given in Proposition 5.7.3 with* $\mathtt{b}$ *fixed in* (5.8.1)*. Then there exists* $\delta_0 \in (0,1)$*,* $N_0 > 0$*,* $C_0 > 0$*, such that, if*

$$N_0^{C_0}\varepsilon^{4-3a}\gamma_*^{-1} = N_0^{C_0}\varepsilon^{1-(9/2)a} \leq \delta_0, \quad \gamma := \varepsilon^{2+a}, \quad a \ll 1, \tag{5.8.2}$$

then

(i) ***(Eigenvalues)***. *For all $\omega \in \Omega_\varepsilon$ there exists a sequence*

$$d_j^\infty(\omega) := d_j^\infty(\omega, i_\delta(\omega)) := m(\omega)\,\omega(j) + \varepsilon^2\kappa_j(\omega) + r_j^\infty(\omega), \quad j \in S^c, \tag{5.8.3}$$

with m and κ_j in (5.7.331) and (5.7.408) respectively. Furthermore, for all $j \in S^c$

$$\sup_j \langle j\rangle |r_j^{(\infty)}|^{\gamma_*} < C\varepsilon^{4-3a}, \qquad r_j^\infty = -r_{-j}^\infty \tag{5.8.4}$$

for some $C > 0$. All the eigenvalues $\mathrm{i}d_j^\infty$ are purely imaginary.

(ii) ***(Conjugacy)***. *For all ω in the set*

$$\Omega_\infty^{2\gamma_*} := \Omega_\infty^{2\gamma_*}(i_\delta) := \left\{\omega \in \mathcal{O}_\infty^{2\gamma} : |\omega\cdot\ell + d_j^\infty(\omega) - d_k^\infty(\omega)| \geq \frac{2\gamma_*}{\langle\ell\rangle^\tau},\ \forall \ell \in \mathbb{Z}^\nu,\ \forall j,k \in S^c\right\} \tag{5.8.5}$$

there is a real, bounded, invertible, linear operator $\Phi_\infty(\omega)\colon H^s_{S^\perp}(\mathbb{T}^{\nu+1}) \to H^s_{S^\perp}(\mathbb{T}^{\nu+1})$, with bounded inverse $\Phi_\infty^{-1}(\omega)$, that conjugates $\mathcal{L}$ in (5.7.13) to constant coefficients, namely

$$\begin{aligned}\mathcal{L}_\infty(\omega) &:= \Phi_\infty(\omega)\circ\mathcal{L}\circ\Phi_\infty^{-1}(\omega) = \omega\cdot\partial_\varphi + \mathcal{D}_\infty(\omega),\\ \mathcal{D}_\infty(\omega) &:= \mathrm{diag}_{j\in S^c}\{\mathrm{i}d_j^\infty(\omega)\}.\end{aligned} \tag{5.8.6}$$

The transformations $\Phi_\infty, \Phi_\infty^{-1}$ are tame and they satisfy for $s_0 \leq s \leq S_{max}$

$$\|(\Phi_\infty^{\pm1} - \mathrm{I})h\|_s^{\gamma_*,\Omega_\infty^{2\gamma_*}} \leq_s \left(\varepsilon^{4-3a}\gamma_*^{-1} + \varepsilon\gamma^{-1}\gamma_*^{-1}\|\mathfrak{I}_\delta\|_{s+\sigma}^{\gamma,\mathcal{O}_0}\right)\|h\|_{s_0}^{\gamma_*,\Omega_\infty^{2\gamma_*}} + \varepsilon^{4-3a}\gamma_*^{-1}\|h\|_s^{\gamma_*,\Omega_\infty^{2\gamma_*}}. \tag{5.8.7}$$

Moreover $\Phi_\infty, \Phi_\infty^{-1}$ are symplectic, and $\mathcal{L}_\infty$ is a Hamiltonian operator.

In order to prove this theorem we need to work in the class of Lip--1-majorant tame operators. We have the following Lemma

Lemma 5.8.2. *The operator $\mathcal{L}$ in (5.7.13) is of the form $\mathcal{D}_\omega - \mathbf{M}_0$ with*

$$\mathbf{M}_0 = \mathcal{D}_0 + \mathcal{P}_0\,, \quad \mathcal{D}_0 = \mathrm{diag}(\mathrm{i}\,d_j^{(0)})_{j\in S^c}\,, \quad d_j^{(0)} = m\left(\frac{j(4+j^2)}{1+j^2}\right) + \varepsilon^2\kappa_j. \tag{5.8.8}$$

Here the functions $d_j^{(0)}$ are well defined and Lipschitz in the set Ω_ε, $|m-1|^{\gamma,\Omega_\varepsilon} \leq \varepsilon^2$, while $\mathcal{P}_0$ is defined and Lipschitz in ω belonging to the set $\mathcal{O}_\infty^{2\gamma}$ (see (5.7.226)). We have that

$$\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}_0}(s_0, \mathtt{b}_0) \leq \varepsilon^{4-3a}\,, \tag{5.8.9}$$

Moreover m, $\mathcal{P}_0$ and the set $\mathcal{O}_\infty^{2\gamma}$ depend on $i = i(\omega)$ and satisfy the bounds

$$\begin{aligned}&|\Delta_{12} m| \leq \varepsilon\|i_1 - i_2\|_{s_0+\sigma}\\ &\|\langle D_x\rangle^{1/2}\underline{\Delta_{12}\mathcal{P}_0}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})},\ \|\langle D_x\rangle^{1/2}\underline{\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}_0}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})} \leq \varepsilon\gamma^{-1}\|i_1 - i_2\|_{s_0+\sigma}\,,\end{aligned} \tag{5.8.10}$$

for all $\omega \in \mathcal{O}_\infty^{2\gamma}(i_1) \cap \mathcal{O}_\infty^{2\gamma}(i_2)$. Here σ is given in Theorem (5.7.3).

Proof. Lemma 2.3.13 trivially implies that any operator $A \in \mathfrak{C}_{1,\mathtt{b}}$ with $\mathtt{b} := s_0 + \mathtt{b}_0$ (recall the Definition 5.7.2 and the fact that $\gamma_* < \gamma$), satisfies

$$\mathfrak{M}^{\sharp,\gamma_*}_A(-1,s), \mathfrak{M}^{\sharp,\gamma_*}_A(-1,s,\mathtt{b}_0) \leq_s \mathbb{B}^\gamma_A(s,\mathtt{b}). \tag{5.8.11}$$

The same holds for $\|\langle D_x\rangle^{1/2}\underline{\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}_0}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})}$. Thus the bound (5.8.9) and the second estimate in (5.8.10) follow by (5.6.7) and (5.7.16), provided ε is small enough. The first bound in (5.8.10) follows by (5.7.331). □

5.8.1 A reduction algorithm for linear Hamiltonian vector fields

We say that a bounded linear operator $\mathbf{B} = \mathbf{B}(\varphi)$ is Hamiltonian if $\mathbf{B}(\varphi)z$ is a linear Hamiltonian vector field w.r.t. the symplectic form J. This means that the corresponding Hamiltonian $\frac{1}{2}(z, J^{-1}\mathbf{B}(\varphi)z)$ is a real quadratic function provided that $z_j = \bar{z}_{-j}$ and $\varphi \in \mathbb{T}^\nu$.

In matrix elements this means that

$$(J^{-1}\mathbf{B}(\varphi))_j^{j'} = (J^{-1}\mathbf{B}(\varphi))_{j'}^{j}, \quad \overline{(J^{-1}\mathbf{B})_j^{j'}(\ell)} = (J^{-1}\mathbf{B})_{-j}^{-j'}(-\ell)$$

or more explicitely:

$$\mathbf{B}_j^{j'}(\varphi) = -\frac{\omega(j)}{\omega(j')}\mathbf{B}_{-j'}^{-j}(\varphi), \quad \overline{\mathbf{B}_j^{j'}(\ell)} = \mathbf{B}_{-j}^{-j'}(-\ell). \tag{5.8.12}$$

This representation is convenient in the present setting because it keeps track of the Hamiltonian structure and

$$\mathcal{B} = \frac{1}{2}(z, J^{-1}\mathbf{B}(\varphi)z), \quad \mathcal{G} = \frac{1}{2}(z, J^{-1}\mathbf{G}(\varphi)z) \ \Rightarrow \{\mathcal{B}, \mathcal{G}\} = \frac{1}{2}(z, J^{-1}[\mathbf{B}, \mathbf{G}]z).$$

We investigate the reducibility of a Hamiltonian operator of the form (recall (5.7.408))

$$\mathbf{M}_0 = \mathcal{D}_0 + \mathcal{P}_0, \quad \mathcal{D}_0 = \mathrm{diag}(\mathrm{i}\, d_j^{(0)}), \quad d_j^{(0)} = m\left(\frac{j(4+j^2)}{1+j^2}\right) + \varepsilon^2\kappa_j. \tag{5.8.13}$$

Here the functions $d_j^{(0)}$ are well defined and Lipschitz in the set Ω_ε, $|m-1|^{\gamma,\Omega_\varepsilon} \le \varepsilon^2$, while $\mathcal{P}_0$ is defined and Lipschitz in ω belonging to the set $\mathcal{O}_\infty^{2\gamma}$. We fix

$$\mathtt{a} := 6\tau + 4, \qquad \tau_1 := 2\tau + 2, \tag{5.8.14}$$

we require that $\mathcal{P}_0, \langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}_0$ are Lip -1- modulo tame, with constants denoted by $\mathfrak{M}_{\mathcal{P}_0}^{\sharp,\gamma_*}(s)$ and $\mathfrak{M}_{\mathcal{P}_0}^{\sharp,\gamma_*}(s, \mathtt{b}_0)$ respectively, in the set $\mathcal{O}_\infty^{2\gamma}$. Moreover m and $\mathcal{P}_0$ and the set $\mathcal{O}_\infty^{2\gamma}$ depend on $i = i(\omega)$ and satisfy the bounds

$$\begin{aligned} &|\Delta_{12} m| \le \mathtt{B}\|i_1 - i_2\|_{s_0+\sigma} \\ &\|\langle D_x\rangle^{1/2}\underline{\Delta_{12}\mathcal{P}_0}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})},\ \|\langle D_x\rangle^{1/2}\underline{\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}_0}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})} \le \mathbb{E}_0\|i_1 - i_2\|_{s_0+\sigma}, \end{aligned} \tag{5.8.15}$$

for all $\omega \in \mathcal{O}_\infty^{2\gamma}(i_1) \cap \mathcal{O}_\infty^{2\gamma}(i_2)$ with $\mathtt{B}, \mathfrak{M}_{\mathcal{P}_0}^{\sharp,\gamma_*}(s_0), \mathfrak{M}_{\mathcal{P}_0}^{\sharp,\gamma_*}(s_0, \mathtt{b}_0) \le \mathbb{E}_0$. We recall that $\|\cdot\|_{\mathcal{L}(H^{s_0})}$ is the operatorial norm. We associate to the operator (5.8.13) the Hamiltonian

$$\mathcal{H}_0 := \omega\cdot\eta + \frac{1}{2}(z, J^{-1}\mathbf{M}_0 z).$$

Proposition 5.8.3. (Iterative reduction) *Let $\sigma > 0$ be the loss of derivatives in Theorem 5.7.3. Consider an operator of the form* (5.8.13) *. For all $S_{max} > s_0$, there is $N_0 := N_0(S_{max}, \mathtt{b}_0) \in \mathbb{R}_+$ such that, if (recall* (5.8.14)*)*

$$N_0^{\tau_1}\mathfrak{M}_0^{\sharp,\gamma_*}(s_0, \mathtt{b}_0)\gamma_*^{-1} \le 1, \quad \mathfrak{M}_0^{\sharp,\gamma_*}(s, \mathtt{b}) := \mathfrak{M}_{\mathcal{P}_0}^{\sharp,\gamma_*}(s, \mathtt{b}) \tag{5.8.16}$$

then, for all $k \ge 0$:

$(\mathbf{S1})_{\mathbf{k}}$ *there exists a sequence of Hamiltonian operators*

$$\mathbf{M}_k = \mathcal{D}_k + \mathcal{P}_k \,, \qquad \mathcal{D}_k := \mathrm{diag}_{j\in S^c}(\mathrm{i}\, d_j^{(k)}) \,, \tag{5.8.17}$$

with $d_j^{(k)}$ defined for $\omega \in \Omega_\varepsilon$ and

$$d_j^{(k)}(\omega) := d_j^{(0)} + r_j^{(k)}(\omega) \,, \quad r_j^{(0)} := 0 \,, \; r_j^{(k)} \in \mathbb{R} \,, \; r_j^{(k)} = -r_{-j}^{(k)} . \tag{5.8.18}$$

The operators $\mathcal{P}_k$ are defined for $k \geq 1$ in a set

$$\Omega_k^{\gamma_*} := \Big\{ \omega \in \Omega_{k-1}^{\gamma_*} \,:\, |\omega\cdot\ell + d_j^{(k-1)} - d_{j'}^{(k-1)}| \geq \frac{\gamma_*}{\langle \ell \rangle^\tau}, \; \forall |\ell| \leq N_{k-1}, \; \forall j, j' \in S^c \Big\} \tag{5.8.19}$$

where $\Omega_0^{\gamma_*} := \mathcal{O}_\infty^{2\gamma}$ and $N_k := N_0^{(3/2)^k}$. Moreover $\mathcal{P}_k$ and $\langle \partial_\varphi \rangle^{\mathtt{b}_0} \mathcal{P}_k$ are -1-modulo-tame with modulo-tame constants respectively

$$\mathfrak{M}_k^{\sharp,\gamma_*}(s) := \mathfrak{M}_{\mathcal{P}_k}^{\sharp,\gamma_*}(s) \,, \quad \mathfrak{M}_k^{\sharp,\gamma_*}(s,\mathtt{b}_0) := \mathfrak{M}_{\mathcal{P}_k}^{\sharp,\gamma_*}(s,\mathtt{b}_0), \quad k \geq 0 \tag{5.8.20}$$

for all $s \in [s_0, S_{max}]$. Setting $N_{-1} = 1$, we have

$$\mathfrak{M}_k^{\sharp,\gamma_*}(s) \leq \mathfrak{M}_0^{\sharp,\gamma_*}(s,\mathtt{b}_0) N_{k-1}^{-\mathtt{a}} \,, \quad \mathfrak{M}_k^{\sharp,\gamma_*}(s,\mathtt{b}_0) \leq \mathfrak{M}_0^{\sharp,\gamma_*}(s,\mathtt{b}_0) N_{k-1} \,. \tag{5.8.21}$$

While for all $k \geq 1$

$$\langle j \rangle |d_j^{(k)} - d_j^{(k-1)}| \leq \mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b}_0) N_{k-2}^{-\mathtt{a}} . \tag{5.8.22}$$

$(\mathbf{S2})_{\mathbf{k}}$ *For $k \geq 1$, there exists a linear symplectic change of variables $\mathcal{Q}_{k-1}$, defined in $\Omega_k^{\gamma_*}$ and such that*

$$\mathbf{M}_k := \mathcal{Q}_{k-1} \omega\cdot\partial_\varphi \mathcal{Q}_{k-1}^{-1} + \mathcal{Q}_{k-1} \mathbf{M}_{k-1} \mathcal{Q}_{k-1}^{-1} . \tag{5.8.23}$$

The operators

$$\Psi_{k-1} := \mathcal{Q}_{k-1} - \mathrm{I} \tag{5.8.24}$$

and $\langle \partial_\varphi \rangle^{\mathtt{b}_0} \Psi_{k-1}$, are -1-modulo-tame with modulo-tame constants satisfying, for all $s \in [s_0, S_{max}]$,

$$\mathfrak{M}_{\Psi_{k-1}}^{\sharp,\gamma_*}(s) \leq \gamma_*^{-1} N_{k-1}^{\tau_1} N_{k-2}^{-\mathtt{a}} \mathfrak{M}_0^{\sharp,\gamma_*}(s,\mathtt{b}_0) \,, \; \mathfrak{M}_{\Psi_{k-1}}^{\sharp,\gamma_*}(s,\mathtt{b}_0) \leq \gamma_*^{-1} N_{k-1}^{\tau_1} N_{k-2} \mathfrak{M}_0^{\sharp,\gamma_*}(s,\mathtt{b}_0) \,. \tag{5.8.25}$$

$(\mathbf{S3})_{\mathbf{k}}$ *Let $i_1(\omega)$, $i_2(\omega)$ such that $\mathcal{P}_0(i_1)$, $\mathcal{P}_0(i_2)$ satisfy (5.8.15). Then for all $\omega \in \Omega_k^{\gamma_1}(i_1) \cap \Omega_k^{\gamma_2}(i_2)$ with $\gamma_1, \gamma_2 \in [\gamma_*/2, 2\gamma_*]$ we have*

$$\| \langle D_x \rangle^{1/2} \underline{\Delta_{12} \mathcal{P}_k} \langle D_x \rangle^{1/2} \|_{\mathcal{L}(H^{s_0})} \leq \mathbb{E}_0 N_{k-1}^{-\mathtt{a}} \| i_1 - i_2 \|_{s_0+\sigma} , \tag{5.8.26}$$

$$\| \langle D_x \rangle^{1/2} \underline{\langle \partial_\varphi \rangle^{\mathtt{b}_0} \Delta_{12} \mathcal{P}_k} \langle D_x \rangle^{1/2} \|_{\mathcal{L}(H^{s_0})} \leq \mathbb{E}_0 N_{k-1} \| i_1 - i_2 \|_{s_0+\sigma} \,. \tag{5.8.27}$$

Moreover for all $k = 1, \dots, n$, for all $j \in S^c$,

$$\langle j \rangle \big| \Delta_{12} r_j^{(k)} - \Delta_{12} r_j^{(k-1)} \big| \leq \| \langle D_x \rangle^{1/2} \underline{\Delta_{12} \mathcal{P}_k} \langle D_x \rangle^{1/2} \|_{\mathcal{L}(H^{s_0})} \,, \tag{5.8.28}$$

$$\langle j \rangle \; |\Delta_{12} r_j^{(k)}| \leq \mathbb{E}_0 \| i_1 - i_2 \|_{s_0+\sigma} \,. \tag{5.8.29}$$

$(\mathbf{S4})_{\mathbf{k}}$ *Let* i_1, i_2 *be like in* $(\mathbf{S3})_{\mathbf{k}}$ *and* $0 < \rho < \gamma_*/2$. *Then*

$$\mathbb{E}_0 N_{k-1}^{\tau+1} \|i_1 - i_2\|_{s_0+\sigma} \le \rho \quad \Longrightarrow \quad \Omega_k^{\gamma_*}(i_1) \subseteq \Omega_k^{\gamma_*-\rho}(i_2) \,. \tag{5.8.30}$$

The Proposition is proved by applying repeatedly the following **KAM reduction procedure** : Fix any $N \gg 1$ and consider any operator of the form

$$\mathbf{M} = \mathcal{D}(\omega) + \mathcal{P}(\varphi, \omega) \,, \qquad \mathcal{D}(\omega) = \mathrm{diag}(\mathrm{i}\, d_j(\omega))_{j\in\mathbb{Z}} \,, \quad d_j = d_j^{(0)} + r_j .$$

Here the $r_j \in \mathbb{R}$ are well defined and Lipschitz for $\omega \in \Omega_\varepsilon$ and such that

$$r_j = -r_{-j} \,, \quad \sup_j \langle j \rangle |r_j|^{\gamma_*, \Omega_\varepsilon} < 2\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}_0}(s_0, \mathtt{b}_0). \tag{5.8.31}$$

Assume that in a set $\mathcal{O} \equiv \mathcal{O}(i) \subseteq \mathcal{O}^{2\gamma}_\infty(i) \subseteq \Omega_\varepsilon$ the operators $\mathcal{P}, \langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}$ are Hamiltonian, real and -1-modulo tame with

$$\gamma_*^{-1} N^{2\tau+2} \mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0) < 1 \,. \tag{5.8.32}$$

Assume finally that $d_j = d_j(i)$, $\mathcal{P}(i), \langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}(i)$ are Lipschitz w.r.t. i namely for all ω belonging to $\mathcal{O}(i_1) \cap \mathcal{O}(i_2)$

$$\begin{aligned} &\sup_j \langle j \rangle |\Delta_{12} r_j| < 2\mathbb{E}_0 \|i_1 - i_2\|_{s_0+\sigma} \\ &\| \langle D_x \rangle^{1/2} \underline{\Delta_{12} \langle\partial_\varphi\rangle^a \mathcal{P}} \langle D_x \rangle^{1/2} \|_{\mathcal{L}(H^{s_0})} \le \mathbb{B}(a) \|i_1 - i_2\|_{s_0+\sigma} \,, \quad a = 0, \mathtt{b}_0 \end{aligned} \tag{5.8.33}$$

for some constants $\mathbb{B}(0) =: \mathbb{B}$ and $\mathbb{B}(\mathtt{b}_0)$ (recall that $\mathbb{E}_0$ is defined in (5.8.15)). Let

$$\mathcal{C} \equiv \mathcal{C}_{\mathcal{D}}^{(\gamma_*,\tau,N,\mathcal{O})} := \{\omega \in \mathcal{O} \,:\, |\omega \cdot \ell + d_j - d_{j'}| > \frac{\gamma_*}{\langle \ell \rangle^\tau} \,, \quad \forall (\ell, j, j') \neq (0, j, j), \ |\ell| \le N, \, j, j' \in S^c\}. \tag{5.8.34}$$

For $\omega \in \mathcal{C}$ let $\mathcal{A}(\varphi)$ be defined as follows

$$\mathcal{A}_j^{j'}(\ell) = \begin{cases} \dfrac{\mathcal{P}_j^{j'}(\ell)}{\mathrm{i}(\omega\cdot\ell + d_j - d_{j'})} & |\ell| \le N, \\ 0 & \text{otherwise.} \end{cases} \tag{5.8.35}$$

Lemma 5.8.4 (KAM step)**.** *The following holds:*

- $\mathcal{A}$ *in* (5.8.35) *is a Hamiltonian,* -1*-modulo tame matrix with the bounds*

  $$\begin{aligned} &\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s,a) = \gamma_*^{-1} N^{2\tau+1} \mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s,a) \,, \\ &\| \langle D_x \rangle^{1/2} \underline{\Delta_{12} \langle\partial_\varphi\rangle^a \mathcal{A}} \langle D_x \rangle^{1/2} \|_{\mathcal{L}(H^{s_0})} \le C\, \gamma_*^{-1} N^{2\tau+1} (\mathbb{B}(a) + \mathbb{E}_0 \gamma_*^{-1} \mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0, a)) \|i_1 - i_2\|_{s_0+\sigma}) \,, \end{aligned}$$

  for $a = 0, \mathtt{b}_0$ *and for all* $\omega \in \mathcal{C}(i_1) \cap \mathcal{C}(i_2)$.

- *The operator* $\mathcal{Q} = e^{\mathcal{A}} := \sum_{k\ge0} \frac{\mathcal{A}^k}{k!}$ *is well defined and invertible, moreover* $\Psi = \mathcal{Q} - \mathrm{I}$ *is a* -1*-modulo tame operator with the bounds*

  $$\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{Q}-\mathrm{I}}(s,a) \le 2\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s,a) \le 2\gamma_*^{-1} N^{2\tau+1} \mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s,a) \,, \tag{5.8.36}$$

  $$\| \langle D_x \rangle^{1/2} \underline{\Delta_{12} \langle\partial_\varphi\rangle^a \mathcal{Q}} \langle D_x \rangle^{1/2} \|_{\mathcal{L}(H^{s_0})} \le 2\gamma_*^{-1} N^{2\tau+1} (\mathbb{B}(a) + \mathbb{E}_0 \gamma_*^{-1} \mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0, a)) \|i_1 - i_2\|_{s_0+\sigma}) \,,$$

  for $a = 0, \mathtt{b}_0$. *Finally* $z \to \mathcal{Q}z$ *is a symplectic change of variables generated by the time one flow of the Hamiltonian* $\mathcal{S} = \frac{1}{2}(z, J^{-1}\mathcal{A}z)$.

- *Set, for* $\omega \in \mathcal{C}$ *(see* (5.8.34)*),*

$$\mathcal{Q}(\omega \cdot \partial_\varphi \mathcal{Q}^{-1}) + \mathcal{Q}\left(\mathcal{D}(\omega) + \mathcal{P}(\varphi,\omega)\right)\mathcal{Q}^{-1} := \mathbf{M}_+ = \mathcal{D}^+(\omega) + \mathcal{P}^+(\varphi,\omega) \tag{5.8.37}$$

where $\mathcal{D}^+(\omega) = \mathrm{diag}(\mathrm{i}\, d_j^+)$ *is Hamiltonian, diagonal, independent of* φ *and defined for all* $\omega \in \Omega_\varepsilon$ *with*

$$\begin{aligned} & d_j^+ = d_j^{(0)} + r_j^+ , \quad r_j^+ = -r_{-j}^+ , \quad \sup_j \langle j\rangle |r_j - r_j^+|^{\gamma,\Omega_\varepsilon} \le \mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0) , \\ & \sup_j \langle j\rangle |\Delta_{12}(r_j - r_j^+)| \le \mathbb{B}\|i_1 - i_2\|_{s_0+\sigma} , \quad \forall \omega \in \mathcal{C}(i_1)\cap\mathcal{C}(i_2). \end{aligned} \tag{5.8.38}$$

For $\omega \in \mathcal{C}$ *we have the bounds*

$$\mathfrak{M}_{\mathcal{P}^+}^{\sharp,\gamma_*}(s) \le N^{-\mathtt{b}_0}\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s,\mathtt{b}_0) + C(s) N^{2\tau+1}\gamma_*^{-1}\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s)\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0) . \tag{5.8.39}$$

$$\begin{aligned} \mathfrak{M}_{\mathcal{P}^+}^{\sharp,\gamma_*}(s,\mathtt{b}_0) \le {} & \mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s,\mathtt{b}_0) \\ & + N^{2\tau+1}\gamma_*^{-1} C(s,\mathtt{b}_0)\Big(\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s,\mathtt{b}_0)\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0) + \mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s)\Big) . \end{aligned} \tag{5.8.40}$$

Moreover for all $\omega \in \mathcal{C}(i_1)\cap\mathcal{C}(i_2)$

$$\begin{aligned} \|\underline{\Delta_{12}\mathcal{P}^+}\|_{\mathcal{L}(H^{s_0})} \le {} & N^{-\mathtt{b}_0}\mathbb{B}(\mathtt{b}_0)\|i_1 - i_2\|_{s+\sigma(\mathtt{b}_0)} \\ & + C(s_0) N^{2\tau+1}\gamma_*^{-1}\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0)\Big(\mathbb{B} + \gamma_*^{-1}\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0)\mathbb{E}_0\Big)\|i_1 - i_2\|_{s+\sigma} \end{aligned} \tag{5.8.41}$$

$$\begin{aligned} \|\underline{\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}^+}\|_{\mathcal{L}(H^{s_0})} \le {} & \mathbb{B}(\mathtt{b}_0)\|i_1 - i_2\|_{s+\sigma} \\ & + N^{2\tau+1}\gamma_*^{-1} C(s_0,\mathtt{b}_0)\Big(\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\mathbb{B} \\ & + \mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0)\Big(\mathbb{B}(\mathtt{b}_0) + \gamma_*^{-1}\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\mathbb{E}_0\Big) \\ & + \gamma_*^{-1}N^{2\tau+1}\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0)\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\Big(\mathbb{B} + \gamma_*^{-1}\mathfrak{M}_{\mathcal{P}}^{\sharp,\gamma_*}(s_0)\mathbb{E}_0\Big)\Big)\|i_1 - i_2\|_{s+\sigma}. \end{aligned} \tag{5.8.42}$$

- *The action of* $\mathcal{Q}$ *on the Hamiltonian* $\mathcal{H}$ *is given by (see* (5.8.37)*)*

$$\mathcal{H}_+ := e^{\{\mathcal{S},\cdot\}}\mathcal{H} = \omega\cdot\eta + \frac{1}{2}(w, J^{-1}\mathbf{M}^+ w).$$

Proof. The first statement is obvious from the definitions, indeed by (2.3.19)

$$\langle\partial_\varphi\rangle^a\mathcal{A} \preceq \gamma_*^{-1}N^\tau\langle\partial_\varphi\rangle^a\mathcal{P} , \quad \text{for } a = 0, \mathtt{b}_0,$$

while

$$\langle\partial_\varphi\rangle^a\Delta_{\omega,\omega'}\mathcal{A} \preceq \gamma_*^{-1}N^\tau\langle\partial_\varphi\rangle^a\Delta_{\omega,\omega'}\mathcal{P} + \gamma_*^{-2}N^{2\tau+1}\langle\partial_\varphi\rangle^a\mathcal{P} , \quad \text{for } a = 0, \mathtt{b}_0$$

since

$$\Delta_{\omega,\omega'}\mathcal{A}_j^{j'}(\ell) = \frac{\Delta_{\omega,\omega'}\mathcal{P}_j^{j'}(\ell)}{\mathrm{i}\big(\omega\cdot\ell + d_j - d_{j'}\big)} - \mathrm{i}\frac{\mathcal{P}_j^{j'}(\ell)\left([(\omega-\omega')\cdot\ell/(\omega-\omega')] + \Delta_{\omega,\omega'}(d_j - d_{j'})\right)}{\big(\omega\cdot\ell + d_j - d_{j'}\big)^2}$$

and (5.8.31), (5.8.32) hold. So we may apply Lemma 2.3.14 *(i)*. The bounds on Δ_{12} come from applying the Leibniz rule and by (5.8.33)

$$|\Delta_{12}\mathcal{A}_j^{j'}(\ell)| \le \frac{|\Delta_{12}\mathcal{P}_j^{j'}(\ell)|}{|\omega\cdot\ell + d_j - d_{j'}|} + \frac{|\mathcal{P}_j^{j'}(\ell)||\Delta_{12}d_j - \Delta_{12}d_{j'}|}{(\omega\cdot\ell + d_j - d_{j'})^2}. \tag{5.8.43}$$

We remark that in the second summand (recall that $\mathtt{B} \leq \mathbb{E}_0$)

$$\frac{|\Delta_{12} d_j - \Delta_{12} d_{j'}|}{|\omega \cdot \ell + d_j - d_{j'}|} \leq |\Delta_{12} m| \frac{|\omega(j) - \omega(j')|}{|\omega \cdot \ell + d_j - d_{j'}|} + \frac{|\Delta_{12} r_j| + |\Delta_{12} r_{j'}|}{|\omega \cdot \ell + d_j - d_{j'}|}$$
$$\stackrel{(5.8.33),(5.8.15)}{\leq} C\, \gamma_*^{-1} (\mathtt{B} N^{\tau+1} + N^{\tau} \mathbb{E}_0) \|i_1 - i_2\|_{s_0+\sigma} \leq \gamma_*^{-1} N^{\tau+1} \mathbb{E}_0 \|i_1 - i_2\|_{s_0+\sigma}.$$

The estimate on the first summand follows from the estimates on $\Delta_{12} m$ and the fact that if $|\omega(j) - \omega(j')| > C|\ell|$ with $C > 1$ then $|\omega \cdot \ell + d_j - d_{j'}| > \tilde{C}|\omega(j) - \omega(j')|$ with $\tilde{C} > 0$; the estimate on the second summand comes from (5.8.31), (5.8.32). In conclusion we get (recall (5.8.33) for the definition of $\mathbb{B}(a)$)

$$\|\langle D_x \rangle^{1/2} \underline{\Delta_{12} \langle \partial_\varphi \rangle^a \mathcal{A}} \langle D_x \rangle^{1/2} \|_{\mathcal{L}(H^{s_0})} \leq C(\gamma_*^{-1} N^{\tau} \mathbb{B}(a) + \gamma_*^{-2} N^{2\tau+1} \mathbb{E}_0 \mathfrak{M}^{\sharp, \gamma_*}_{\mathcal{P}}(s_0, a)) \|i_1 - i_2\|_{s_0+\sigma}$$

for all $\omega \in \mathcal{C}(i_1) \cap \mathcal{C}(i_2)$. The fact that $\mathcal{A}$ is Hamiltonian follows from (5.8.12) and from the fact that d_j is odd in j (recall (5.7.409) and (5.8.13)).

The second statement comes from the definition of $\mathcal{Q}$, Lemma 2.3.14 *(iv-v)* and the smallness condition (5.8.32). By definition

$$\mathcal{D}^+ + \mathcal{P}^+ = \mathcal{D} + \mathcal{P} - \omega \cdot \partial_\varphi \mathcal{A} + [\mathcal{A}, \mathcal{D} + \mathcal{P}] + \sum_{k \geq 2} \frac{\mathrm{ad}(\mathcal{A})^k}{k!} (\mathcal{D} + \mathcal{P}) - \sum_{k \geq 2} \frac{\mathrm{ad}(\mathcal{A})^{k-1}}{k!} (\omega \cdot \partial_\varphi \mathcal{A}). \quad (5.8.44)$$

Again by definition, $\mathcal{A}$ solves the equation

$$\omega \cdot \partial_\varphi \mathcal{A} + [\mathcal{D}, \mathcal{A}] = \Pi_N \mathcal{P} - [\mathcal{P}], \qquad [\mathcal{P}]_j^{j'}(\ell) = \delta(\ell, 0)\, \delta(j, j')\, P_j^{j'}(\ell).$$

Substituting in (5.8.44) we get

$$\mathcal{D}^+ + \mathcal{P}^+ = \mathcal{D} + [\mathcal{P}] + \Pi_N^\perp \mathcal{P} + \sum_{k \geq 1} \frac{\mathrm{ad}(\mathcal{A})^k}{k!} (\mathcal{P}) - \sum_{k \geq 2} \frac{\mathrm{ad}(\mathcal{A})^{k-1}}{k!} (\Pi_N \mathcal{P} - [\mathcal{P}]\ ) \quad (5.8.45)$$

By the reality condition (5.8.12) we get $\overline{\mathcal{P}_j^j(0)} = \mathcal{P}_{-j}^{-j}(0) = -\mathcal{P}_j^j(0)$, which shows that $\mathcal{P}_j^j(0)$ is purely imaginary and odd in j. By Kirtzbraun Theorem we extend $\mathcal{P}_j^j(0)$ to the whole Ω_ε preserving the $|\cdot|^{\gamma_*}$ norm. We set

$$d_j^+ = d_j - \mathrm{i}(\mathcal{P}_j^j(0))^{\mathrm{Ext}},$$

where $(\cdot)^{\mathrm{Ext}}$ denotes the extension of the eigenvalue at Ω_ε, so that the bound (5.8.38) follows, by Lemma 2.3.14 *(i)*, from the bounds on $\mathcal{P}$ and $\Delta_{12}\mathcal{P}$ (see (5.8.33)). Now for $\omega \in \mathcal{C}$

$$\mathcal{P}^+ = \Pi_N^\perp \mathcal{P} + \sum_{k \geq 1} \frac{\mathrm{ad}(\mathcal{A})^k}{k!} (\mathcal{P}) - \sum_{k \geq 2} \frac{\mathrm{ad}(\mathcal{A})^{k-1}}{k!} (\Pi_N \mathcal{P} - [\mathcal{P}]\ ). \quad (5.8.46)$$

By Lemma 2.3.14-*(iv)* we have

$$\mathfrak{M}^{\sharp, \gamma_*}_{(\mathrm{ad}\mathcal{A})^k \mathcal{P}}(s) \leq C(s)^k \Big((\mathfrak{M}^{\sharp, \gamma_*}_{\mathcal{A}}(s_0))^k \mathfrak{M}^{\sharp, \gamma_*}_{\mathcal{P}}(s) + k (\mathfrak{M}^{\sharp, \gamma_*}_{A}(s_0))^{k-1} \mathfrak{M}^{\sharp, \gamma_*}_{\mathcal{A}}(s) \mathfrak{M}^{\sharp, \gamma_*}_{\mathcal{P}}(s_0) \Big) \quad (5.8.47)$$

which implies (5.8.39), by using also 2.3.14*(iii)*. Finally

$$\begin{aligned}&\mathfrak{M}^{\sharp,\gamma_*}_{(\mathrm{ad}\mathcal{A})^k\mathcal{P}}(s,\mathtt{b}_0)\leq C(s,\mathtt{b}_0)^k\Big((\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^k\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s,\mathtt{b}_0)\\&+k(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^{k-1}\Big(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0)+\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s)\Big)\\&+k(k-1)(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^{k-2}\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s)\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0)\Big)\end{aligned}\tag{5.8.48}$$

which implies (5.8.40). In order to obtain the bounds (5.8.41) and (5.8.42) on Δ_{12}, we just apply Leibniz rule repeatedly in (5.8.46) and then procede as before. More precisely we have for all $\omega\in\mathcal{C}(i_1)\cap\mathcal{C}(i_2)$

$$\Delta_{12}(\mathrm{ad}(\mathcal{A})^k\mathcal{P})=\mathrm{ad}(\mathcal{A})^k\Delta_{12}\mathcal{P}+\sum_{k_1+k_2=k-1}\mathrm{ad}(\mathcal{A})^{k_1}\mathrm{ad}(\Delta_{12}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_2}\mathcal{P}.\ ^{1}$$

Now we note that $\|\langle D_x\rangle^{1/2}\underline{\mathcal{A}}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})}\leq\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0)$ and that for any matrices A,B we have

$$\|\langle D_x\rangle^{1/2}\underline{\mathrm{ad}(A)B}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})}\leq C(s_0)\|\langle D_x\rangle^{1/2}\underline{A}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})}\|\langle D_x\rangle^{1/2}\underline{B}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})}.$$

This implies that for all $\omega\in\mathcal{C}(i_1)\cap\mathcal{C}(i_2)$ (recall (5.8.33) for the definition of $\mathtt{B}$)

$$\begin{aligned}&\|\langle D_x\rangle^{1/2}\underline{\Delta_{12}(\mathrm{ad}(\mathcal{A})^k\mathcal{P})}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})}\leq(C(s_0)\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^k\mathbb{B}\\&+kC(s_0)^k(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^{k-1}\gamma_*^{-1}\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0)(N^\tau\mathbb{B}+\gamma_*^{-1}N^{2\tau+1}\mathbb{E}_0\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0))\|i_1-i_2\|_{s_0+\sigma}.\end{aligned}\tag{5.8.49}$$

Now by definition

$$\Delta_{12}\mathcal{P}^+=\Pi_N^\perp\Delta_{12}\mathcal{P}+\sum_{k\geq1}\Delta_{12}(\frac{\mathrm{ad}(\mathcal{A})^k}{k!}\mathcal{P})-\sum_{k\geq2}\Delta_{12}\big(\frac{\mathrm{ad}(\mathcal{A})^{k-1}}{k!}(\Pi_N\mathcal{P}-[\mathcal{P}])\big)\,,\tag{5.8.50}$$

so we use Lemma 2.3.14 *(iii)* in oder to bound the first summand and (5.8.49) in order to bound the remaining ones. In the same way

$$\begin{aligned}\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}(\mathrm{ad}(\mathcal{A})^k\mathcal{P})&=\mathrm{ad}(\mathcal{A})^k\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}+\sum_{k_1+k_2=k-1}\mathrm{ad}(\mathcal{A})^{k_1}\mathrm{ad}(\Delta_{12}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_2}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}\\&+\sum_{k_1+k_2=k-1}\mathrm{ad}(\mathcal{A})^{k_1}\mathrm{ad}(\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_2}\Delta_{12}\mathcal{P}\\&+\sum_{k_1+k_2=k-1}\mathrm{ad}(\mathcal{A})^{k_1}\mathrm{ad}(\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_2}\mathcal{P}\\&+\sum_{k_1+k_2+k_3=k-2}\mathrm{ad}(\mathcal{A})^{k_1}\mathrm{ad}(\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_2}\mathrm{ad}(\Delta_{12}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_3}\mathcal{P}\\&+\sum_{k_1+k_2+k_3=k-2}\mathrm{ad}(\mathcal{A})^{k_1}\mathrm{ad}(\Delta_{12}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_2}\mathrm{ad}(\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{A})\mathrm{ad}(\mathcal{A})^{k_3}\,,\end{aligned}$$

[1] Recall the usual convention that $a(\Delta_{12}b)c\equiv a(i_1)(\Delta_{12}b)c(i_2)$.

where the last two terms appear only if $k \geq 2$. We proceed as for (5.8.49) and obtain the bound

$$\begin{aligned}
&\|\langle D_x\rangle^{1/2}\underline{\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}(\mathrm{ad}(\mathcal{A})^k\mathcal{P})}\langle D_x\rangle^{1/2}\|_{\mathcal{L}(H^{s_0})} \leq (C(s_0)\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^k\mathbb{B}(\mathtt{b}_0) &(5.8.51)\\
&+ kC(s_0)^k(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^{k-1}\gamma_*^{-1}\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0,\mathtt{b}_0)(N^\tau\mathbb{B} + \gamma_*^{-1}N^{2\tau+1}\mathbb{E}_0\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0))\\
&+ kC(s_0)^k(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^{k-1}\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0,\mathtt{b}_0)\mathbb{B}\\
&+ kC(s_0)^k(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^{k-1}\gamma_*^{-1}\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0)(N^\tau\mathbb{B}(\mathtt{b}_0) + \gamma_*^{-1}N^{2\tau+1}\mathbb{E}_0\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0,\mathtt{b}_0))\\
&+ 2k(k-1)C(s_0)^k(\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0))^{k-2}\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{A}}(s_0,\mathtt{b}_0)\gamma_*^{-1}\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0) &(5.8.52)\\
&(N^\tau\mathbb{B} + \gamma_*^{-1}N^{2\tau+1}\mathbb{E}_0\mathfrak{M}^{\sharp,\gamma_*}_{\mathcal{P}}(s_0))\|i_1 - i_2\|_{s_0+\sigma}
\end{aligned}$$

□

Now we can prove Proposition 5.8.3 by using Lemma 5.8.4.

Proof of Proposition 5.8.3. Items $(\mathbf{S1})_\mathbf{0} - (\mathbf{S4})_\mathbf{0}$ follow trivially from the definitions. So we proceed by induction, assuming that $(\mathbf{S1})_\mathbf{k} - (\mathbf{S4})_\mathbf{k}$ hold.
We start by proving $(\mathbf{S2})_\mathbf{k+1}$. We wish to apply the KAM step with $\mathcal{P}_k = \mathcal{P}$, $d_j^{(k)} = d_j$, $N_k = N$, etc. By the first of (5.8.21), (5.8.16) and (5.8.14), we have that (5.8.32) holds, similarly (5.8.31) holds since, by (5.8.22)

$$\langle j\rangle|d_j^{(k)}| \leq \sum_{h=1}^{k}\langle j\rangle|d_j^{(h)} - d_j^{(h-1)}| \leq 2\mathfrak{M}^{\sharp,\gamma_*}_0(s_0,\mathtt{b}_0).$$

We now choose as change of variables $\mathcal{Q}_k$ the one denoted by $\mathcal{Q}$ in the KAM step. The bounds (5.8.25) follow from (5.8.36) since

$$2\gamma_*^{-1}N_k^{2\tau+1}\mathfrak{M}^{\sharp,\gamma_*}_k(s) \leq 2\gamma_*^{-1}N_k^{\tau_1}\mathfrak{M}^{\sharp,\gamma_*}_0(s,\mathtt{b}_0)N_{k-1}^{-\mathtt{a}}$$
$$2\gamma_*^{-1}N_k^{2\tau+1}\mathfrak{M}^{\sharp,\gamma_*}_k(s,\mathtt{b}_0) \leq 2\gamma_*^{-1}N_k^{\tau_1}\mathfrak{M}^{\sharp,\gamma_*}_0(s,\mathtt{b}_0)N_{k-1}$$

by the definition of τ_1 (see (5.8.14)). Formula (5.8.23) follows by (5.8.37) with $\mathbf{M}_{k+1} = \mathbf{M}_+$.
We now pass to proving $(\mathbf{S1})_\mathbf{k+1}$. As said above $\mathcal{P}_{k+1} = \mathcal{P}^+$ of the KAM step, $d_j^{(k+1)} = d_j^+$, etc...
In order to prove (5.8.22), we use the bound (5.8.38), we get

$$\sup_j\langle j\rangle|d_j^{(k+1)} - d_j^{(k)}|^{\gamma,\Omega_\varepsilon} = \sup_j\langle j\rangle|r_j^{(k+1)} - r_j^{(k)}|^{\gamma,\Omega_\varepsilon} \leq \mathfrak{M}^{\sharp,\gamma_*}_k(s_0) \leq \mathfrak{M}^{\sharp,\gamma_*}_0(s_0,\mathtt{b}_0)N_{k-1}^{-\mathtt{a}}.$$

In order to prove (5.8.21) we use the bounds (5.8.39), (5.8.40). Indeed

$$\begin{aligned}
\mathfrak{M}^{\sharp,\gamma_*}_{k+1}(s) &\leq N_k^{-\mathtt{b}_0}\mathfrak{M}^{\sharp,\gamma_*}_k(s,\mathtt{b}_0) + C(s)N_k^{2\tau+1}\gamma_*^{-1}\mathfrak{M}^{\sharp,\gamma_*}_k(s)\mathfrak{M}^{\sharp,\gamma_*}_k(s_0)\\
&\leq N_k^{-\mathtt{b}_0}\mathfrak{M}^{\sharp,\gamma_*}_0(s,\mathtt{b}_0)N_{k-1} + C(s)N_k^{2\tau+1}N_{k-1}^{-2\mathtt{a}}\gamma_*^{-1}\mathfrak{M}^{\sharp,\gamma_*}_0(s,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_0(s_0,\mathtt{b}_0) \leq N_k^{-\mathtt{a}}\mathfrak{M}^{\sharp,\gamma_*}_0(s,\mathtt{b}_0)
\end{aligned}$$

by using (5.8.16), the fact that $\mathtt{b}_0 - 2/3 > \mathtt{a}$ and $\mathtt{a}/3 > 2\tau + 1$ (recall (5.8.14) and (5.8.1)) and provided that we take $N_0 = N_0(S_{max})$ large, in order to absorb the constants. Regarding the second bound we have

$$\begin{aligned}
\mathfrak{M}^{\sharp,\gamma_*}_{k+1}(s,\mathtt{b}_0) &\leq \mathfrak{M}^{\sharp,\gamma_*}_k(s,\mathtt{b}_0)\\
&\quad + N_k^{2\tau+1}\gamma_*^{-1}C(s,\mathtt{b}_0)\Big(\mathfrak{M}^{\sharp,\gamma_*}_k(s,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_k(s_0) + \mathfrak{M}^{\sharp,\gamma_*}_k(s_0,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_k(s)\Big)\\
&\leq N_{k-1}\mathfrak{M}^{\sharp,\gamma_*}_0(s,\mathtt{b}_0) + N_k^{2\tau+1}N_{k-1}^{1-\mathtt{a}}\gamma_*^{-1}C(s,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_0(s,\mathtt{b}_0)\mathfrak{M}^{\sharp,\gamma_*}_0(s_0,\mathtt{b}_0)
\end{aligned}$$

again (5.8.21) follows by using (5.8.16), the fact that $2\mathtt{a}/3 > 2\tau + 1$ (recall (5.8.14) and (5.8.1)) and provided that we take $N_0 = N_0(S_{max}, \mathtt{b}_0)$ large, in order to absorb the constants.
Item $(\mathbf{S3})_{\mathbf{k+1}}$, follows just as item $(\mathbf{S1})_{\mathbf{k+1}}$ from the corresponding bounds (5.8.41) and (5.8.42). Indeed (here $\mathbb{B}_k, \mathbb{B}_k(\mathtt{b}_0)$ are just the constants $\mathbb{B}$, $\mathbb{B}(\mathtt{b}_0)$ in formula (5.8.33) for $\mathcal{P}_k$)

$$
\begin{aligned}
&\|\underline{\Delta_{12}\mathcal{P}_{k+1}}\|_{\mathcal{L}(H^{s_0})} \leq \Big(N_k^{-\mathtt{b}_0}\mathbb{B}_k(\mathtt{b}_0) + C(s_0)N_k^{2\tau+1}\gamma_*^{-1}\mathfrak{M}_k^{\sharp,\gamma_*}(s_0)\Big(\mathbb{B}_k + \gamma_*^{-1}\mathfrak{M}_k^{\sharp,\gamma_*}(s_0)\mathbb{E}_0\Big)\Big)\|i_1 - i_2\|_{s+\sigma} \\
&\leq \Big(N_k^{-\mathtt{b}_0}\mathbb{E}_0 N_{k-1} + C(s_0)N_k^{2\tau+1}\gamma_*^{-1}N_{k-1}^{-2\mathtt{a}}\mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\mathbb{E}_0\Big(1 + \gamma_*^{-1}\mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\Big)\Big)\|i_1 - i_2\|_{s+\sigma} \\
&\leq \mathbb{E}_0(N_k^{-\mathtt{b}_0+2/3} + 2C(s_0)N_k^{2\tau+1-4/3\mathtt{a}}N_0^{-\tau_1}) \leq \mathbb{E}_0 N_k^{-\mathtt{a}}.
\end{aligned}
$$

Similarly

$$
\begin{aligned}
\|\underline{\Delta_{12}\langle\partial_\varphi\rangle^{\mathtt{b}_0}\mathcal{P}_{k+1}}\|_{\mathcal{L}(H^{s_0})} \leq\ & \mathbb{B}_k(\mathtt{b}_0)\|i_1 - i_2\|_{s+\sigma} \\
&+ N_k^{2\tau+1}\gamma_*^{-1}C(s_0,\mathtt{b}_0)\Big(\mathfrak{M}_k^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\mathbb{B}_k \\
&+ \mathfrak{M}_k^{\sharp,\gamma_*}(s_0)\Big(\mathbb{B}_k(\mathtt{b}_0) + \gamma_*^{-1}\mathfrak{M}_k^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\mathbb{E}_0\Big) \\
&+ \gamma_*^{-1}N_k^{2\tau+1}\mathfrak{M}_k^{\sharp,\gamma_*}(s_0)\mathfrak{M}_k^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\Big(\mathbb{B}_k + \gamma_*^{-1}\mathfrak{M}_k^{\sharp,\gamma_*}(s_0)\mathbb{E}_0\Big)\Big)\|i_1 - i_2\|_{s+\sigma} \\
&\leq \Big(N_{k-1} + N_k^{2\tau+1+2/3-2/3\mathtt{a}}\gamma_*^{-1}C(s_0,\mathtt{b}_0)\mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b}_0) \\
&+\gamma_*^{-2}N_k^{4\tau+2+2/3-4/3\mathtt{a}}(\mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b}_0))^2\Big)\mathbb{E}_0\|i_1 - i_2\|_{s+\sigma},
\end{aligned}
$$

and the result follows as in the previous cases. The estimate (5.8.28) follows from the second formula in (5.8.38) while (5.8.29) follows by (5.8.28) and the bounds (5.8.26).
In order to prove $(\mathbf{S4})_{\mathbf{k+1}}$ we note that for $h \leq k$, $|\ell| \leq N_h$ and $\omega \in \Omega_{h+1}^{\gamma_*}$

$$
\begin{aligned}
|\omega\cdot\ell + d_j^{(h)}(i_2) - d_{j'}^{(h)}(i_2)| &\geq |\omega\cdot\ell + d_j^{(h)}(i_1) - d_{j'}^{(h)}(i_1)| - |\Delta_{12}m(\omega(j) - \omega(j')) + \Delta_{12}(r_j^{(h)} - r_{j'}^{(h)})| \\
&\overset{(5.8.29),(5.8.15)}{\geq} \frac{\gamma_*}{\langle\ell\rangle^\tau} - (\mathtt{B}|\omega(j) - \omega(j')| + \mathbb{E}_0)\|i_1 - i_2\|_{s_0+\sigma}.
\end{aligned}
$$

Now if $|\omega(j) - \omega(j')| > \tilde{C}\langle\ell\rangle$ for an opportune constant $\tilde{C} > 0$ then for $i = i_1, i_2$

$$
|\omega\cdot\ell + d_j^{(h)}(i) - d_{j'}^{(h)}(i)| \geq \frac{1}{2}|\omega(j) - \omega(j')|,
$$

hence $\Omega_h^{\gamma_*} = \emptyset$. Otherwise, if N_0 is large enough to absorb the constants, we get

$$
|\omega\cdot\ell + d_j^{(h)}(i_2) - d_{j'}^{(h)}(i_2)| \geq \frac{\gamma_*}{\langle\ell\rangle^\tau} - (\mathbb{E}_0|\omega(j) - \omega(j')| + \mathbb{E}_0)\|i_1 - i_2\|_{s_0+\sigma} \geq \frac{\gamma_* - \rho}{\langle\ell\rangle^\tau}.
$$

□

Eventually, we are in the position to prove Theorem 5.8.1.

Proof of Theorem 5.8.1. We want to apply Proposition 5.8.3 to the operator $\mathcal{L}$ in (5.7.13) in Proposition 5.7.3. It is convenient to remark that $\mathcal{L}$ gives the dynamics of a quadratic time-dependent Hamiltonian. Passing to the extended phase space, $\mathcal{L}$ corresponds to the Hamiltonian

$$
\mathcal{H} := \mathcal{H}(z,\eta) = \omega\cdot\eta + \frac{1}{2}(z, J^{-1}\mathbf{M}z), \qquad \mathbf{M} = \mathcal{D}_0 + \mathcal{P}_0
$$

where $\mathcal{D}_0, \mathcal{P}_0$ are defined in (5.8.8) and satisfy (5.8.10). The smallness assumption (5.8.16) follows by formula (5.8.9) in Lemma 5.8.2 and by the smallness condition on ε in (5.8.2) provided that N_0 in formula (5.8.2) is chosen as in Proposition 5.8.3. We can conclude that Proposition 5.8.3 applies to the operator $\mathcal{L}$ in (5.7.13).

By (5.8.22) we have that the sequence $(d_j^k)_{k\in\mathbb{N}}$ is Cauchy, hence the limit $d_j^\infty = d_j^{(0)} + r_j^\infty$ exists and, also by (5.8.18), r_j^∞ satisfies (5.8.4).
Now we claim that (recall (5.8.5) and (5.8.19))

$$\Omega_\infty^{2\gamma_*} \subseteq \bigcap_{k\geq 0} \Omega_k^{\gamma_*}. \tag{5.8.53}$$

Indeed we have by (5.8.22), for $|\ell| \leq N_k$

$$\begin{aligned}|\omega\cdot\ell + d_j^k - d_{j'}^k| &\geq |\omega\cdot\ell + d_j^\infty - d_{j'}^\infty| - |r_j^k - r_j^\infty| - |r_{j'}^k - r_{j'}^\infty| \\ &\geq \frac{2\gamma_*}{\langle\ell\rangle^\tau} - \frac{\mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)}{N_{k-2}^{\mathtt{a}}} \geq \frac{\gamma_*}{\langle l\rangle^\tau}\end{aligned}$$

since $\mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b}_0) \leq \gamma_* N_0^{-\tau_1}$ and $\langle\ell\rangle^\tau \leq N_k^\tau \leq N_{k-2}^{\mathtt{a}}$ due to (5.8.14).

We conclude that $\Omega_\infty^{2\gamma_*} \subseteq \Omega_{k+1}^{\gamma_*}$. Thus the sequence $(\Psi_k)_{k\in\mathbb{N}}$ (recall (5.8.24)) is well defined on $\Omega_\infty^{2\gamma_*}$.
We define

$$\Phi_k = \mathcal{Q}_0 \circ \cdots \circ \mathcal{Q}_k.$$

We claim that there exists $\Phi_\infty := \lim_{k\to\infty}\Phi_k$. First we note that by using (5.8.25)

$$\mathfrak{M}_{\Phi_k}^{\sharp,\gamma_*}(s) \leq \sum_{j=0}^k \left(\mathfrak{M}_{\mathcal{Q}_j}^{\sharp,\gamma_*}(s)\prod_{i\neq j}\mathfrak{M}_{\mathcal{Q}_i}^{\sharp,\gamma_*}(s_0)\right) \leq 2\sum_{j=0}^k \mathfrak{M}_{\mathcal{Q}_j}^{\sharp,\gamma_*}(s) \leq C\big(1 + \max_{j=0,\dots,k}\mathfrak{M}_{\Psi_j}^{\sharp,\gamma_*}(s)\big), \qquad \forall k. \tag{5.8.54}$$

By Lemmata 2.3.14 and 2.3.13 we have

$$\begin{aligned}\mathfrak{M}_{\Phi_k-\Phi_{k-1}}^{\gamma_*}(s,\mathtt{b}) &\leq_s \mathfrak{M}_{\Phi_k-\Phi_{k-1}}^{\sharp,\gamma_*}(s,\mathtt{b}_0) \leq_s \mathfrak{M}_{\Psi_k}^{\sharp,\gamma_*}(s,\mathtt{b}_0) + \mathfrak{M}_{\Psi_k}^{\sharp,\gamma_*}(s_0,\mathtt{b}_0)\max_{j=0,\dots,k}\mathfrak{M}_{\Psi_j}^{\sharp,\gamma_*}(s,\mathtt{b}_0) \\ &\quad + \mathfrak{M}_{\Psi_k}^{\sharp,\gamma_*}(s,\mathtt{b}_0)\max_{j=0,\dots,k}\mathfrak{M}_{\Psi_j}^{\gamma_*,\sharp}(s_0,\mathtt{b}_0) \\ &\overset{(5.8.25)}{\leq_s} C(s_0,\mathtt{b})N_k^{\tau_1}N_{k-1}^{-\mathtt{a}}\mathfrak{M}_0^{\sharp,\gamma_*}(s,\mathtt{b}_0)\gamma_*^{-1}\end{aligned}$$

Thus by

$$\|(\Phi_{k+m}-\Phi_k)h\|_s^{\gamma_*,\Omega_\infty^{2\gamma_*}} \leq \sum_{j=k}^{k+m}\|(\Phi_j-\Phi_{j-1})h\|_s^{\gamma_*,\Omega_\infty^{2\gamma_*}}$$

and by (5.8.11) we have that (recall (5.8.16) and (5.8.14))

$$\mathfrak{M}_{\Phi_{k+m}-\Phi_k}^{\gamma_*}(s,\mathtt{b}) \leq C(s_0,\mathtt{b})\mathbb{B}_{\mathcal{P}_0}^{\gamma}(s,\mathtt{b})N_k^{\tau_1}N_{k-1}^{-\mathtt{a}}\gamma_*^{-1} \overset{(5.7.16)}{\leq} C(s_0,\mathtt{b})N_k^{-2(\tau+(1/3))},$$

hence $(\Phi_k)_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{L}(H^s)$ and for Φ_∞ the estimate (5.8.7) holds. The operators Φ_k are close to the identity, hence the same is true for Φ_∞ and by Neumann series it is invertible. It is easy to prove that for Φ_∞^{-1} the estimate (5.8.7) holds. □

5.8.2 Inversion of $\mathcal{L}_\omega$

Let us define (recall (5.4.6))

$$\mathcal{F}_\infty^{2\gamma}(i_\delta) := \{\omega \in \mathcal{O}_0 : |\omega\cdot\ell - d_j^\infty(i_n)| \geq \frac{2\gamma}{\langle\ell\rangle^\tau}, \quad \forall\ell\in\mathbb{Z}^\nu, \forall j\in S^c\} \tag{5.8.55}$$

We have the following result.

Theorem 5.8.5. *Assume the hypotesis of Theorem 5.8.1,* (5.6.7) *with* $\mu' \geq s_0 + \mathtt{b} + \sigma + 2\tau + 1$, *where* σ *is given in Theorem 5.7.3, and* (5.5.6) *with* $\tilde\mu \geq \mu' + \mathfrak{d}$, *where* $\mathfrak{d}$ *is given in Lemma 5.5.3. Then for all* $\omega \in \Omega_\infty := \Omega_\infty^{2\gamma_*}(i_\delta)\cap\mathcal{F}_\infty^{2\gamma}(i_\delta)$ *(see* (5.8.5)*), for any function* $g \in H_{S^\perp}^{s+2\tau+1}(\mathbb{T}^{\nu+1})$ *the equation* $\mathcal{L}_\omega h = g$ *has a solution* $h = \mathcal{L}_\omega^{-1} g \in H_{S^\perp}^s(\mathbb{T}^{\nu+1})$, *satisfying*

$$\begin{aligned}\|\mathcal{L}_\omega^{-1}g\|_s^{\gamma,\Omega_\infty} &\leq_s \gamma^{-1}(\|g\|_{s+2\tau+1}^{\gamma,\Omega_\infty} + \varepsilon\gamma^{-1}\gamma_*^{-1}\|\mathfrak{I}_\delta\|_{s+\mu'}^{\gamma,\Omega_\infty}\|g\|_{s_0}^{\gamma,\Omega_\infty})\\ &\leq_s \gamma^{-1}(\|g\|_{s+2\tau+1}^{\gamma,\Omega_\infty} + \varepsilon\gamma^{-1}\gamma_*^{-1}\{\|\mathfrak{I}_0\|_{s+\mu'+\mathfrak{d}}^{\gamma,\Omega_\infty} + \gamma^{-1}\|\mathfrak{I}_0\|_{s_0+\mathfrak{d}}^{\gamma,\Omega_\infty}\|Z\|_{s+\mu'+\mathfrak{d}}^{\gamma,\Omega_\infty}\}\|g\|_{s_0}^{\gamma,\Omega_\infty}).\end{aligned} \tag{5.8.56}$$

Proof. We conjugated the operator $\mathcal{L}_\omega$ in (5.6.31) to a diagonal operator $\mathcal{L}_\infty = \chi\mathcal{L}_\omega\chi^{-1}$, see (5.8.6), with (recall (5.8.7) and Theorem 5.7.3)

$$\chi := \Phi_\infty \circ \Upsilon. \tag{5.8.57}$$

Moreover, by (5.7.17) and (5.8.7) and Lemma 5.5.3 we had the following estimates

$$\|\chi^{\pm1}h\|_s^{\gamma,\Omega_\infty} \leq_s \|h\|_s^{\gamma,\Omega_\infty} + \varepsilon\gamma^{-1}\gamma_*^{-1}\|\mathfrak{I}_\delta\|_{s+s_0+\mathtt{b}+\sigma}^{\gamma,\mathcal{O}_0}\|h\|_{s_0}^{\gamma,\Omega_\infty}. \tag{5.8.58}$$

We have

$$\mathcal{L}_\infty^{-1}g = \sum_{\ell\neq0,j\neq0}\frac{g_{\ell j}}{\mathrm{i}\big(\omega\cdot\ell - d_j^\infty(\omega)\big)}\, e^{\mathrm{i}(\ell\cdot\varphi+jx)} \tag{5.8.59}$$

and then

$$\|\mathcal{L}_\infty^{-1}g\|_s^{\gamma,\Omega_\infty} \leq \gamma^{-1}\|g\|_{s+2\tau+1}^{\gamma,\Omega_\infty}. \tag{5.8.60}$$

Thus, by (5.8.58) and (5.8.60) we get the estimates (5.8.56). □

If the assumption (5.5.6) is satisfied with $\tilde\mu = s_0 + \mathtt{b} + \sigma + 2\tau + 1 + \mathfrak{d}$, where σ is the loss of derivatives in (5.7.16), $\mathtt{b}$ is given in Theorem 5.8.1 and $\mathfrak{d}$ in Lemma 5.5.3, then Theorem 5.8.5 implies the inversion assumption (5.5.28) with $\sigma' = \tilde\mu$.

5.9 The Nash-Moser nonlinear iteration

In this section we prove Theorem 4.3.2. It will be a consequence of the Nash-Moser theorem 5.9.2.

Consider the finite-dimensional subspaces

$$E_n := \{\mathfrak{I}(\varphi) = (\Theta, y, z)(\varphi) : \Theta = \Pi_n\Theta, y = \Pi_n y, z = \Pi_n z\}$$

where $N_n := N_0^{\chi^n}$ are introduced in (4.6.129), and Π_n are the projectors (which, with a small abuse of notation, we denote with the same symbol)

$$\Pi_n \Theta(\varphi) := \sum_{|\ell|<N_n} \Theta_\ell \, e^{\mathrm{i}\ell\cdot\varphi}, \ \Pi_n y(\varphi) := \sum_{|\ell|<N_n} y_\ell \, e^{\mathrm{i}\ell\cdot\varphi}, \text{ where } \Theta(\varphi) = \sum_{\ell\in\mathbb{Z}^\nu} \Theta_\ell \, e^{\mathrm{i}\ell\cdot\varphi}, \ y(\varphi) = \sum_{\ell\in\mathbb{Z}^\nu} y_\ell \, e^{\mathrm{i}\ell\cdot\varphi},$$

$$\Pi_n z(\varphi, x) := \sum_{|(\ell,j)|<N_n} z_{\ell j} \, e^{\mathrm{i}(\ell\cdot\varphi + jx)}, \text{ where } \quad z(\varphi, x) = \sum_{\ell\in\mathbb{Z}^\nu, j\in S^c} z_{\ell j} \, e^{\mathrm{i}(\ell\cdot\varphi+jx)}. \tag{5.9.1}$$

We define $\Pi_n^\perp = \mathrm{I} - \Pi_n$. The classical smoothing properties hold, namely, for all $\alpha, s \geq 0$,

$$\|\Pi_n \mathfrak{I}\|_{s+\alpha}^{\gamma,\mathcal{O}} \leq N_n^\alpha \|\mathfrak{I}_\delta\|_s^{\gamma,\mathcal{O}}, \quad \forall \mathfrak{I}(\omega) \in H^s, \quad \|\Pi_n^\perp \mathfrak{I}\|_s^{\gamma,\mathcal{O}} \leq N_n^{-\alpha} \|\mathfrak{I}\|_{s+\alpha}^{\gamma,\mathcal{O}}, \quad \forall \mathfrak{I}(\omega) \in H^{s+\alpha}. \tag{5.9.2}$$

We define the following constants

$$\begin{aligned} &\mu_1 := 3\mu + 3, && \alpha := 3\mu_1 + 1, && \alpha_1 := (\alpha - 3\mu)/2, \\ &k := 3(\mu_1 + \rho^{-1}) + 1, && \beta_1 := 6\mu_1 + 3\rho^{-1} + 3, && \frac{1}{2}\left(\frac{1-(9/2)a}{C_1(1+a)}\right) < \rho < \frac{1-(9/2)a}{C_1(1+a)}. \end{aligned} \tag{5.9.3}$$

where $\mu := \mu(\tau, \nu) > 0$ is the "loss of regularity" given by the Theorem 5.5.36 and C_1 is fixed below.

Remark 5.9.1. We remark that $\mu \gg \tilde{\mu}$ given in Theorem 5.8.5 and $\tilde{\mu} \geq \sigma$ where σ is given in Theorem 5.7.3.

Theorem 5.9.2. ***(Nash-Moser)*** *Assume that $f \in C^\infty$. Let $\tau := 2\nu + 6$. Then there exist $C_1 > \max\{\mu_1 + \alpha, C_0\}$ (where $C_0 := C_0(\tau, \nu)$ is the one in Theorem 5.8.1), $\delta_0 := \delta_0(\tau, \nu) > 0$ such that, if*

$$N_0^{C_1} \varepsilon^{b_*+1} \gamma^{-1} \gamma_*^{-1} < \delta_0, \quad \gamma := \varepsilon^{2+a} = \varepsilon^{2b}, \quad \gamma_* := \gamma^{3/2}, \quad N_0 := (\varepsilon\gamma^{-1})^\rho, \quad b_* = 9 - 2b, \tag{5.9.4}$$

then, for all $n \geq 0$:

$(\mathcal{P}1)_n$ *there exists a function $(\mathfrak{I}_n, \zeta_n) \colon \mathcal{G}_n \subseteq \Omega_\varepsilon \to E_{n-1} \times \mathbb{R}^\nu, \omega \mapsto (\mathfrak{I}_n(\omega), \zeta_n(\omega)), (\mathfrak{I}_0, \zeta_0) := 0, E_{-1} := \{0\}$, where the set $\mathcal{G}_0$ is defined in (5.4.5) and the sets $\mathcal{G}_n$ for $n \in \mathbb{N}$ are defined inductively by:*

$$\begin{aligned} \Lambda_{n+1}^{(0)} &:= \left\{\omega \in \mathcal{G}_n : |\omega\cdot\ell + m(i_n)j| \geq \frac{2\,\gamma_n}{\langle\ell\rangle^\tau}, \ \forall j \in S^c, \ell \in \mathbb{Z}^\nu\right\}, \\ \Lambda_{n+1}^{(1)} &:= \left\{\omega \in \mathcal{G}_n : |\omega\cdot\ell + d_j^\infty(i_n)| \geq \frac{2\,\gamma_n}{\langle\ell\rangle^\tau}, \ \forall j \in S^c, \ell \in \mathbb{Z}^\nu\right\}, \\ \Lambda_{n+1}^{(2)} &:= \left\{\omega \in \mathcal{G}_n : |\omega\cdot\ell + d_j^\infty(i_n) - d_k^\infty(i_n)| \geq \frac{2\,\gamma_n^*}{\langle\ell\rangle^\tau}, \ \forall j,k \in S^c, \ell \in \mathbb{Z}^\nu\right\}, \\ \mathcal{G}_{n+1} &:= \bigcap_{i=0}^{2} \Lambda_{n+1}^{(i)}, \end{aligned} \tag{5.9.5}$$

where $\gamma_n := \gamma(1 + 2^{-n})$, $\gamma_n^* := \gamma_*(1 + 2^{-n})$ and $d_j^\infty(\omega) := d_j^\infty(\omega, i_n(\omega))$ are defined in (5.8.3) (and $\mu_0^\infty(\omega) = 0$). Moreover $|\zeta_n|^{\gamma,\mathcal{G}_n} \leq C\|\mathcal{F}(U_n)\|_{s_0}^{\gamma,\mathcal{G}_n}$ and

$$\|\mathfrak{I}_n\|_{s_0+\mu}^{\gamma,\mathcal{G}_n} \leq C_* \varepsilon^{b_*} \gamma^{-1}, \quad \|\mathcal{F}(U_n)\|_{s_0+\mu+3}^{\gamma,\mathcal{G}_n} \leq C_* \varepsilon^{b_*}, \tag{5.9.6}$$

where $U_n := (i_n, \zeta_n)$ *with* $i_n(\varphi) = (\varphi, 0, 0) + \mathfrak{I}_n(\varphi)$. *The differences* $\hat{\mathfrak{I}}_n := \mathfrak{I}_n - \mathfrak{I}_{n-1}$ *(where we set* $\tilde{\mathfrak{I}}_0 := 0$*) is defined on* $\mathcal{G}_n$, *and satisfy*

$$\|\hat{\mathfrak{I}}_1\|_{s_0+\mu}^{\gamma,\mathcal{G}_1} \le C_* \varepsilon^{b_*} \gamma^{-1}, \quad \|\hat{\mathfrak{I}}_n\|_{s_0+\mu}^{\gamma,\mathcal{G}_n} \le C_* \varepsilon^{b_*} \gamma^{-1} N_{n-1}^{-\alpha}, \quad \forall n \ge 2. \tag{5.9.7}$$

$(\mathcal{P}2)_n$ $\|\mathcal{F}(U_n)\|_{s_0}^{\gamma,\mathcal{G}_n} \le C_* \varepsilon^{b_*} N_{n-1}^{-\alpha}$ *where we set* $N_{-1} := 1$.

$(\mathcal{P}3)_n$ *(High Norms).* $\|\mathfrak{I}_n\|_{s_0+\beta_1}^{\gamma,\mathcal{G}_n} \le C_* \varepsilon^{b_*} \gamma^{-1} N_{n-1}^{k}$ *and* $\|\mathcal{F}(U_n)\|_{s_0+\beta_1}^{\gamma,\mathcal{G}_n} \le C_* \varepsilon^{b*} N_{n-1}^{k}$.

$(\mathcal{P}4)_n$ *(Measure). The measure of the "Cantor-like" sets* $\mathcal{G}_n$ *satisfies*

$$|\Omega_\varepsilon \setminus \mathcal{G}_0| \le C_* \varepsilon^{2(\nu-1)} \gamma, \quad |\mathcal{G}_n \setminus \mathcal{G}_{n+1}| \le C_* \varepsilon^{2(\nu-1)} \gamma N_{n-1}^{-1}. \tag{5.9.8}$$

Proof. To simplify the notations we omit the index $\gamma, \mathcal{G}_n$ on the norm $\|\cdot\|_s$.
Proof of $(\mathcal{P}_1)_0, (\mathcal{P}_2)_0, (\mathcal{P}_3)_0$. Recalling (5.4.8), we have, by the second estimate in (5.4.19),

$$\|\mathcal{F}(U_0)\|_s = \|\mathcal{F}((\varphi, 0, 0), 0)\|_s = \|X_P(i_0)\|_s \le_s \varepsilon^{9-2b}.$$

Hence the smallness conditions in $(\mathcal{P}_1)_0, (\mathcal{P}_2)_0, (\mathcal{P}_3)_0$ hold taking $C_* := C_*(s_0 + \beta_1)$ large enough.

Assume that $(\mathcal{P}_1)_n, (\mathcal{P}_2)_n, (\mathcal{P}_3)_n$ *hold for some* $n \ge 0$, *and prove* $(\mathcal{P}_1)_{n+1}, (\mathcal{P}_2)_{n+1}, (\mathcal{P}_3)_{n+1}$. By (5.9.3) and (5.9.4)

$$N_0^{C_1} \varepsilon^{b_*+1} \gamma^{-1} \gamma_*^{-1} = N_0^{C_1} \varepsilon^{4-3a} \gamma_*^{-1} = \varepsilon^{1-(9/2)a - \rho C_1(1+a)} < \delta_0$$

for ε small enough. If we take $C_1 \ge C_0$ then (5.8.2) holds. Moreover (5.9.6) imply (5.5.6), and so (5.6.7), and Theorem 5.8.5 applies. Hence the operator $\mathcal{L}_\omega := \mathcal{L}_\omega(\omega, i_n(\omega))$ in (5.6.31) is defined on $\mathcal{O}_0 = \mathcal{G}_n$ and is invertible for all $\omega \in \mathcal{G}_{n+1}$ since $\mathcal{G}_{n+1} \subseteq \Omega_\infty^{2\gamma_n^*}(i_n) \cap \mathcal{F}_\infty^{2\gamma_n}(i_n)$ and the last estimate in (5.8.56) holds. This means that the assumption (5.5.28) of Theorem 5.5.9 is verified with $\Omega_\infty = \mathcal{G}_{n+1}$. By Theorem 5.5.9 there exists an approximate inverse $\mathbf{T}_n(\omega) := \mathbf{T}_0(\omega, i_n(\omega))$ of the linearized operator $L_n(\omega) := d_{i,\zeta} \mathcal{F}(\omega, i_n(\omega))$, satisfying (5.5.37). By (5.9.4), (5.9.6)

$$\|\mathbf{T}_n g\|_s \le_s \gamma^{-1} (\|g\|_{s+\mu} + \varepsilon \gamma^{-1} \gamma_*^{-1} \{\|\mathfrak{I}_n\|_{s+\mu} + \gamma^{-1} \|\mathfrak{I}_n\|_{s_0+\mu} \|\mathcal{F}(U_n)\|_{s+\mu}\} \|g\|_{s_0+\mu}) \tag{5.9.9}$$

$$\|\mathbf{T}_n g\|_{s_0} \le_{s_0} \gamma^{-1} \|g\|_{s_0+\mu} \tag{5.9.10}$$

and, by (5.5.38), using also (5.9.4), (5.9.6), (5.9.2),

$$\begin{aligned} \|(L_n \circ \mathbf{T}_n - \mathrm{I})g\|_s \le_s & \varepsilon^{2b-1} \gamma^{-2} (\|\mathcal{F}(U_n)\|_{s_0+\mu} \|g\|_{s+\mu} + \|\mathcal{F}(U_n)\|_{s+\mu} \|g\|_{s_0+\mu} \\ & + \varepsilon \gamma^{-1} \|\mathfrak{I}_n\|_{s+\mu} \|\mathcal{F}(U_n)\|_{s_0+\mu} \|g\|_{s_0+\mu}) \end{aligned} \tag{5.9.11}$$

$$\begin{aligned} \|(L_n \circ \mathbf{T}_n - \mathrm{I})g\|_{s_0} & \le_{s_0} \varepsilon^{2b-1} \gamma^{-2} \|\mathcal{F}(U_n)\|_{s_0+\mu} \|g\|_{s_0+\mu} \\ & \le_{s_0} \varepsilon^{2b-1} \gamma^{-2} (\|\Pi_n \mathcal{F}(U_n)\|_{s_0+\mu} + \|\Pi_n^\perp \mathcal{F}(U_n)\|_{s_0+\mu}) \|g\|_{s_0+\mu} \\ & \le_{s_0} \varepsilon^{2b-1} \gamma^{-2} N_n^{\mu} (\|\mathcal{F}(U_n)\|_{s_0} + N_n^{-\beta_1} \|\mathcal{F}(U_n)\|_{s_0+\beta_1}) \|g\|_{s_0+\mu}. \end{aligned} \tag{5.9.12}$$

The index β_1 in (5.9.3) is an ultraviolet cut, and it has to be define in order to obtain the convergence of the iteration scheme.
Now, for all $\omega \in \mathcal{G}_{n+1}$, we can define, for $n \ge 0$,

$$U_{n+1} := U_n + H_{n+1}, \quad H_{n+1} := (\hat{\mathfrak{I}}_{n+1}, \hat{\zeta}_{n+1}) := -\tilde{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) \in E_n \times \mathbb{R}^\nu, \tag{5.9.13}$$

where $\tilde{\Pi}_n(\mathfrak{I},\zeta) := (\Pi_n \mathfrak{I},\zeta)$ with Π_n defined in (5.9.1). Since $L_n := d_{i,\zeta}\mathcal{F}(i_n)$, we write

$$\mathcal{F}(U_{n+1}) = \mathcal{F}(U_n) + L_n H_{n+1} + Q_n,$$

where

$$Q_n := Q(U_n, H_{n+1}), \quad Q(U_n, H) := \mathcal{F}(U_n + H) - \mathcal{F}(U_n) - L_n H, \quad H \in E_n \times \mathbb{R}^\nu. \tag{5.9.14}$$

Then, by the definition of H_{n+1} in (5.9.13), using $[L_n, \Pi_n]$ and writing $\tilde{\Pi}_n^\perp(\mathfrak{I},\zeta) := (\Pi_n^\perp \mathfrak{I}, 0)$ we have

$$\begin{aligned}\mathcal{F}(U_{n+1}) &= \mathcal{F}(U_n) - L_n \tilde{\Pi}_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n = \mathcal{F}(U_n) - L_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + L_n \tilde{\Pi}_n^\perp \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n \\ &= \mathcal{F}(U_n) - \Pi_n L_n \mathbf{T}_n \Pi_n \mathcal{F}(U_n) + (L_n \tilde{\Pi}_n^\perp - \Pi_n^\perp L_n)\mathbf{T}_n \Pi_n \mathcal{F}(U_n) + Q_n \\ &= \Pi_n^\perp \mathcal{F}(U_n) + R_n + Q_n + Q_n'\end{aligned} \tag{5.9.15}$$

where

$$R_n := (L_n \tilde{\Pi}_n^\perp - \Pi_n^\perp L_n)\mathbf{T}_n \Pi_n \mathcal{F}(U_n), \quad Q_n' := -\Pi_n (L_n \mathbf{T}_n - \mathrm{I})\Pi_n \mathcal{F}(U_n). \tag{5.9.16}$$

Lemma 5.9.3. *Define*

$$w_n := \varepsilon\gamma^{-2}\|\mathcal{F}(U_n)\|_{s_0}, \qquad B_n := \varepsilon\gamma^{-1}\|\mathfrak{I}_n\|_{s_0+\beta_1} + \varepsilon\gamma^{-2}\|\mathcal{F}(U_n)\|_{s_0+\beta_1}. \tag{5.9.17}$$

Then there exists $K := K(s_0,\beta_1) > 0$ such that, for all $n \geq 0$, setting $\mu_1 := 3\mu + 3$

$$w_{n+1} \leq K N_n^{\mu_1+\rho^{-1}-\beta_1} B_n + K N_n^{\mu_1} w_n^2, \qquad B_{n+1} \leq K N_n^{\mu_1+\rho^{-1}} B_n. \tag{5.9.18}$$

The proof of Lemma 5.9.3 is similar to the one of Lemma 9.2 in [KdVAut].

Proof of $(\mathcal{P}3)_{n+1}$. By (5.9.18) and $(\mathcal{P}3)_n$

$$B_{n+1} \leq K N_n^{\mu_1+\rho^{-1}} B_n \leq 2C_* K \varepsilon^{b_*+1}\gamma^{-2} N_n^{\mu_1+\rho^{-1}} N_{n-1}^k \leq C_* \varepsilon^{b_*+1}\gamma^{-2} N_n^k, \tag{5.9.19}$$

provided $2K N_n^{\mu_1+\rho^{-1}-k} N_{n-1}^k \leq 1, \forall n \geq 0$. Choosing k as in (5.9.3) and N_0 large enough, i.e. for ε small enough. By (5.9.17) and the bound (5.9.19) $(\mathcal{P}3)_{n+1}$ holds.
Proof of $(\mathcal{P}2)_{n+1}$. Using (5.9.17), (5.9.18) and $(\mathcal{P}2)_n, (\mathcal{P}3)_n$, we get

$$w_{n+1} \leq K N_n^{\mu_1+\rho^{-1}-\beta_1} B_n + K N_n^{\mu_1} w_n^2 \leq K N_n^{\mu_1+\rho^{-1}-\beta_1} 2C_* \varepsilon^{b_*+1}\gamma^{-2} N_{n-1}^k + K N_n^{\mu_1}(C_* \varepsilon^{b_*+1}\gamma^{-2} N_{n-1}^{-\alpha})^2$$

and $w_{n+1} \leq C_* \varepsilon^{b_*+1}\gamma^{-2} N_n^{-\alpha}$ provided that

$$4K N_n^{\mu_1+\rho^{-1}-\beta_1+\alpha} N_{n-1}^k \leq 1, \quad 2K C_* \varepsilon^{b_*+1}\gamma^{-2} N_n^{\mu_1+\alpha} N_{n-1}^{-2\alpha} \leq 1, \ \forall n \geq 0. \tag{5.9.20}$$

The inequalities in (5.9.20) hold by (5.9.4), taking α as in (5.9.3), $C_1 > \mu_1 + \alpha$ and δ_0 in (5.9.4) small enough. By (5.9.17), the inequality $w_{n+1} \leq C_* \varepsilon^{b_*+1}\gamma^{-2} N_n^{-\alpha}$ implies $(\mathcal{P}2)_{n+1}$.

Proof of $(\mathcal{P}_1)_{n+1}$. The bound (5.9.7) for $\hat{\mathfrak{I}}_1$ follows by (5.9.13), (5.9.9) (for $s = s_0 + \mu$) and $\|\mathcal{F}(U_0)\|_{s_0+2\mu} = \|\mathcal{F}((\varphi,0,0),0)\|_{s_0+2\mu} \leq_{s_0+2\mu} \varepsilon^{b_*}$. The bound (5.9.7) for $\hat{\mathfrak{I}}_{n+1}$ follows by (5.9.1), $(\mathcal{P}_2)_n$ and (5.9.3). It remains to prove that (5.9.6) holds at the step $n+1$. We have

$$\|\mathfrak{I}_{n+1}\|_{s_0+\mu} \leq \sum_{k=1}^{n+1} \|\hat{\mathfrak{I}}_k\|_{s_0+\mu} \leq C_* \varepsilon^{b_*} \gamma^{-1} \sum_{k \geq 1} N_{k-1}^{-\alpha_1} \leq C_* \varepsilon^{b_*} \gamma^{-1} \tag{5.9.21}$$

taking α_1 as in (5.9.3) and N_0 large enough, i.e. ε small enough. Moreover, using (5.9.1), $(\mathcal{P}_2)_{n+1}, (\mathcal{P}_3)_{n+1}$, (5.9.3) we get

$$\begin{aligned} \|\mathcal{F}(U_{n+1})\|_{s_0+\mu+1} &\leq N_n^{\mu+1} \|\mathcal{F}(U_{n+1})\|_{s_0} + N_n^{\mu+1-\beta_1} \|\mathcal{F}(U_{n+1})\|_{s_0+\beta_1} \\ &\leq C_* \varepsilon^{b_*} N_n^{\mu+1-\alpha} + C_* \varepsilon^{b_*} N_n^{\mu+1-\beta_1+k} \leq C_* \varepsilon^{b_*}, \end{aligned}$$

which is the second inequality in (5.9.6) at the step $n+1$. The bound $|\zeta_{n+1}|^\gamma \leq C \|\mathcal{F}(U_{n+1})\|_{s_0}^\gamma$ is a consequence of Lemma (5.5.1).

5.9.1 Measure estimates

In this section we prove $(\mathcal{P}_4)_n$ for all $n \geq 0$. First we estimate the measure of $\mathcal{G}_0$.

Lemma 5.9.4. *We have that* $|\Omega_\varepsilon \setminus \mathcal{G}_0| \leq C \varepsilon^{2(\nu-1)} \gamma$ *(see* (5.4.5)*).*

Proof. It is well known that $|\Omega_\varepsilon \setminus \mathcal{G}_0^{(0)}| \leq C \varepsilon^{2(\nu-1)} \gamma$. Thus we focus on the estimate for the measure of $\mathcal{G}_0^{(1)}$. We have that

$$\Omega_\varepsilon \setminus \mathcal{G}_0^{(1)} = \bigcup_{\substack{\ell \in \mathbb{Z}^\nu, |\ell| \leq 3, \\ j,k \in S^c, |j-k| \leq 3C_S}} T_{\ell jk}$$

where

$$T_{\ell jk} := \{\omega \in \Omega_\varepsilon : |\overline{\omega} \cdot \ell + \varepsilon^2 \mathbb{A}\xi \cdot \ell + \omega(j) - \omega(k) + \varepsilon^2(\omega(j)\lambda_j - \omega(k)\lambda_k)| \leq C\gamma\} \tag{5.9.22}$$

for $j,k \in S^c$ such that $\sum_{i=1}^\nu \overline{\jmath}_i \ell_i + j = k$.

Let us first study the ε-independent part of our small divisor i.e.

$$\overline{\omega} \cdot \ell + \omega(j) - \omega(k) = 3 \sum_{i=1}^3 \frac{j_i}{1 + j_i^2} + 3 \frac{j}{1+j^2} - 3 \frac{(\sum_{i=1}^3 j_i + j)}{1 + (\sum_{i=1}^3 j_i + j)^2} \tag{5.9.23}$$

for $j_1, j_2, j_3 \in S$. By our genericity assumption (H1) in (1.2.14) $\sum_{i=1}^3 \frac{j_i}{1+j_i^2} \neq 0$ and hence it is bounded by some positive number $K(S)$. We deduce that for j large enough, i.e. $|j| \geq c(S) \max \overline{\jmath}_i$, (5.9.23) is bounded from below by $K(S)/2$. We note that

$$|\omega(j) - \omega(k)| \leq |\overline{\omega} \cdot \ell + \omega(j) - \omega(k)| + |\overline{\omega} \cdot \ell|$$

then, since $|\ell| \leq 3$,

$$\begin{aligned} &|\bar{\omega} \cdot \ell + \varepsilon^2 \mathbb{A}\xi \cdot \ell + \omega(j)(1 + \varepsilon^2 \lambda_j) - \omega(k)(1 + \varepsilon^2 \lambda_k)| \\ &\geq |\overline{\omega} \cdot \ell + \omega(j) - \omega(k)| - C\varepsilon^2 |\ell| - C' \varepsilon^2 (|\overline{\omega} \cdot \ell + \omega(j) - \omega(k)| + |\overline{\omega} \cdot \ell|) \geq K(S)/4 \end{aligned}$$

for ε small enough. This implies that $T_{\ell jk} = \emptyset$.
We are left to deal with the case $|j| \leq c(S) \max \bar{\jmath}_i$. We write (5.9.23) as $P(\bar{\jmath}_i, j)/Q(\bar{\jmath}_i, j)$ where P, Q are polynomials with integer coefficients. We remark that $1 < Q < C(S)$ due to the condition $|j| \leq c(S) \leq c(S) \max \bar{\jmath}_i$. If $P \neq 0$ then $|P| \geq 1$ and again (5.9.23) is larger than some $K(S)$. We conclude that $T_{\ell jk} = \emptyset$ by reasoning as done in the case j large. Now we study the case in which $P = 0$. Fixed $j_1, j_2, j_3 \in S$ in (5.9.23), then P has degree four in j and so the condition $P = 0$ fixes at most four choices of j. We have for $P = 0$

$$\begin{aligned} &\bar{\omega} \cdot \ell + \varepsilon^2 \mathbb{A}\xi \cdot \ell + \omega(j)(1 + \varepsilon^2 \lambda_j) - \omega(k)(1 + \varepsilon^2 \lambda_k) \\ &= \varepsilon^2 (\mathbb{A}\xi \cdot \ell + (\omega(j') - \omega(j))\vec{v} \cdot \xi + (w_j - w_k) \cdot \xi) \\ &= \varepsilon^2 (\mathbb{A}\xi \cdot \ell + (\bar{\omega} \cdot \ell)\vec{v} \cdot \xi + (w_j - w_k) \cdot \xi), \end{aligned} \tag{5.9.24}$$

where $\vec{v} := (2/3)(1 + \bar{\jmath}_k^2)_{k=1}^{\nu}$. These are a finite number of linear functions of ξ. We compute the derivative in ξ which is

$$(\mathbb{A}^T + \vec{v}\bar{\omega}^T)\ell + (w_j - w_k). \tag{5.9.25}$$

The condition (H3) in (1.2.16) implies that the quantity (5.9.25) is bounded from below by a constant depending on S. This lower bound and Fubini Theorem imply that $|T_{\ell jk}| \leq C_S \varepsilon^{2(\nu-1)} \gamma$ for some C_S depending on S. By the discussion above we have

$$|\Omega_\varepsilon \setminus \mathcal{G}_0^{(1)}| \leq \sum_{|\ell| \leq 3, |j|, |k| \leq K(S)} |T_{\ell jk}| \leq C_S \varepsilon^{2(\nu-1)} \gamma,$$

where $K(S) > 0$ and $C_S > 0$ are constant depending on the set S. This implies the thesis.

□

Fixed $n \geq 1$, we have

$$\mathcal{G}_n \setminus \mathcal{G}_{n+1} = \bigcup_{\ell \in \mathbb{Z}^\nu, j, k \in S^c} \Big(R_{\ell jk}(i_n) \cup Q_{\ell j}(i_n) \cup P_{\ell j}(i_n) \Big) \tag{5.9.26}$$

where

$$\begin{aligned} R_{\ell jk}(i_n) &:= \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + d_j^\infty(i_n) - d_k^\infty(i_n)| < 2\,\gamma_n^* \,\langle \ell \rangle^{-\tau}\}, \\ Q_{\ell j}(i_n) &:= \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + m(i_n) j| < 2\gamma_n \langle \ell \rangle^{-\tau}\} \\ P_{\ell j}(i_n) &:= \{\omega \in \mathcal{G}_n : |\omega \cdot \ell + d_j^\infty(i_n)| < 2\gamma_n \langle \ell \rangle^{-\tau}\} \end{aligned} \tag{5.9.27}$$

Since, by (5.4.5), $R_{\ell jk}(i_n) = \emptyset$ for $j = k$, in the sequel we assume that $j \neq k$.

Lemma 5.9.5. *Let $n \geq 0$. If $R_{\ell jk}(i_n) \neq \emptyset$, then $|\ell| \geq C_1 |\omega(j) - \omega(k)| \geq \frac{C_1}{2} |j - k|$ for some constant $C_1 > 0$ dependent of the tangential set and independent of $\ell, j, k, n, i_n, \omega$.*
If $Q_{\ell j} \neq \emptyset$ then $|\ell| \geq C_2 |j|$ for some constant C_2 dependent of the tangential set and independent of $\ell, j, k, n, i_n, \omega$.
If $P_{\ell j} \neq \emptyset$ then $|\ell| \geq C_3 |j|$ for some constant C_3 dependent of the tangential set and independent of $\ell, j, k, n, i_n, \omega$.

Proof. We claim that $8|\overline{\omega} \cdot \ell| \geq |\omega(j) - \omega(k)|$. Then this would imply

$$|\ell| \geq C_1 |\omega(j) - \omega(k)|, \qquad C_1 := \frac{1}{8|\overline{\omega}|}. \tag{5.9.28}$$

If $R_{\ell jk}(i_n) \neq \emptyset$, then there exist ω such that

$$|d_j^\infty(\omega, i_n(\omega)) - d_k^\infty(\omega, i_n(\omega))| < \frac{2\gamma_n^*}{\langle \ell \rangle^\tau} + 2|\overline{\omega} \cdot \ell|. \tag{5.9.29}$$

Moreover, recall (5.8.4),

$$\begin{aligned} |d_j^\infty(\omega, i_n(\omega)) - d_k^\infty(\omega, i_n(\omega))| &\geq |m||\omega(j) - \omega(k)| - \varepsilon^2|\kappa_j| - \varepsilon^2|\kappa_k| - |r_j^\infty| - |r_k^\infty| \\ &\geq \frac{1}{2}|\omega(j) - \omega(k)| - C\varepsilon^2\left(\frac{1}{|j|} + \frac{1}{|k|}\right) \\ &\geq \frac{1}{3}|\omega(j) - \omega(k)|. \end{aligned} \tag{5.9.30}$$

Thus, for ε small enough

$$2|\overline{\omega}||\ell| \geq 2|\overline{\omega} \cdot \ell| \geq \left(\frac{1}{3} - \frac{2\gamma_n^*}{\langle \ell \rangle^\tau |\omega(j) - \omega(k)|}\right)|\omega(j) - \omega(k)| \geq \frac{1}{4}|\omega(j) - \omega(k)|$$

and this proves the first claim on $R_{\ell jk}$.
If $|mj| \geq 2|\omega \cdot \ell|$ then by (5.4.5)

$$|\omega \cdot \ell + mj| \geq |m||j| - |\omega \cdot \ell| \geq 2|\omega \cdot \ell| - |\omega \cdot \ell| = |\omega \cdot \ell| \geq \frac{\gamma}{\langle \ell \rangle^\tau}.$$

Hence if $Q_{\ell j} \neq \emptyset$ we have

$$|j| \leq \frac{2|\omega \cdot \ell|}{|m|} \leq \frac{1}{C_2}|\ell|, \qquad C_2 := \frac{|m|}{4|\overline{\omega}|}.$$

Following the same arguments and by using that $|d_j| \geq C|j|$ for some constant $C > 0$ we get the last statement. □

Lemma 5.9.6. *For $n \geq 1, |\ell| \leq N_{n-1}$, one has the inclusion $R_{\ell jk}(i_n) \subseteq R_{\ell jk}(i_{n-1})$, $Q_{\ell j}(i_n) \subseteq Q_{\ell j}(i_{n-1})$ and $P_{\ell j}(i_n) \subseteq P_{\ell j}(i_{n-1})$.*

Proof. We first note that, by Lemma 5.9.5, if $|\omega(j) - \omega(k)| > C_1^{-1}|\ell|$ then $R_{\ell jk}(i_n) = R_{\ell jk}(i_{n-1}) = \emptyset$, so that our claim is trivial. Otherwise, if

$$|\omega(j) - \omega(k)| \leq C_1^{-1}|\ell| \leq C_1^{-1}N_{n-1}$$

we claim that for all $j, k \in \mathbb{Z}$ we have (recall (5.8.14))

$$|(d_j^\infty - d_k^\infty)(i_n) - (d_j^\infty - d_k^\infty)(i_{n-1})| \leq \varepsilon^{4-3a}N_n^{-\mathtt{a}} \qquad \forall \omega \in \mathcal{G}_n, \tag{5.9.31}$$

where $d_j^\infty = d_j^\infty(\omega, i_n(\omega))$. We first prove that (5.9.31) implies that $R_{\ell jk}(i_n) \subseteq R_{\ell jk}(i_{n-1})$. For all $j \neq k$, $|\ell| \leq N_{n-1}$, $\omega \in \mathcal{G}_n$ by (5.9.31)

$$\begin{aligned} &|\omega \cdot \ell + d_j^\infty(i_n) - d_k^\infty(i_n)| \geq |\omega \cdot \ell + d_j^\infty(i_{n-1}) - d_k^\infty(i_{n-1})| - |(d_j^\infty - d_k^\infty)(i_n) - (d_j^\infty - d_k^\infty)(i_{n-1})| \\ &\geq 2\gamma_{n-1}^*\langle \ell \rangle^{-\tau} - \varepsilon^{4-3a}N_n^{-\mathtt{a}} \geq 2\gamma_n^*\langle \ell \rangle^{-\tau} \end{aligned} \tag{5.9.32}$$

since $\varepsilon^{4-3a}\gamma_*^{-1}N_n^{\tau-(2/3)\mathtt{a}}2^{n+1} \leq 1$.
Proof of (5.9.31). By (5.8.3) (recall that $\omega(j) := j(4+j^2)/(1+j^2)$)

$$(d_j^\infty - d_k^\infty)(i_n) - (d_j^\infty - d_k^\infty)(i_{n-1}) = (m(i_n) - m(i_{n-1}))(\omega(j) - \omega(k)) \\ + (r_j^\infty(i_n) - r_j^\infty(i_{n-1})) + (r_k^\infty(i_n) - r_k^\infty(i_{n-1})), \tag{5.9.33}$$

where $m = m(\omega, i_n(\omega))$ and similarly for κ_j and r_j^∞. We first apply Proposition 5.8.3-$(\mathbf{S4})_{\mathtt{k}}$ with $k = n+1$, $\gamma_* = \gamma_{n-1}^*$, $\gamma_* - \rho = \gamma_n^*$ and $i_1 \rightsquigarrow i_{n-1}, i_2 \rightsquigarrow i_n$, in order to conclude that

$$\Omega_{n+1}^{\gamma_{n-1}^*}(i_{n-1}) \subseteq \Omega_{n+1}^{\gamma_n^*}(i_n). \tag{5.9.34}$$

The smallness condition in (5.8.30) is satisfied (recall (5.8.15), $\mathbb{E}_0 = \varepsilon\gamma^{-1}$), indeed by (5.9.7) and Remark 5.9.1

$$\varepsilon\gamma^{-1}N_n^{\tau+1}\|i_n - i_{n-1}\|_{s_0+\sigma} \leq C_*\varepsilon^{b_*+1}\gamma^{-2}N_n^{\tau+1}N_{n-1}^{-\alpha}, \tag{5.9.35}$$

$N_n^{\tau-(2/3)\alpha}$ is a decreasing sequence (see (5.9.3)) and by (5.9.4)

$$\varepsilon\gamma^{-1}N_n^{\tau+1}\|i_n - i_{n-1}\|_{s_0+\sigma} \leq \gamma_{n-1}^* - \gamma_n^* = \gamma_* 2^{-n}. \tag{5.9.36}$$

Then by definitions (5.9.5) (see also the proof of Theorem 5.8.1)

$$\mathcal{G}_n \subseteq \mathcal{G}_{n-1} \cap \Omega_\infty^{2\gamma_{n-1}^*}(i_{n-1}) \subseteq \bigcap_{k\geq 0}\Omega_k^{\gamma_{n-1}^*}(i_{n-1}) \subseteq \Omega_{n+1}^{\gamma_{n-1}^*}(i_{n-1}) \overset{(5.9.34)}{\subseteq} \Omega_{n+1}^{\gamma_n^*}(i_n). \tag{5.9.37}$$

For all $\omega \in \mathcal{G}_n \subseteq \Omega_{n+1}^{\gamma_{n-1}^*}(i_{n-1}) \cap \Omega_{n+1}^{\gamma_n^*}(i_n)$, we deduce by Proposition 5.8.3-$(S3)_{\mathtt{k}}$ with $\mathtt{k} = n+1$

$$\langle j\rangle|r_j^{n+1}(i_n) - r_j^{n+1}(i_{n-1})| \overset{(5.8.29)}{\leq} \varepsilon\gamma^{-1}\|i_n - i_{n-1}\|_{s_0+\sigma} \overset{(5.9.7)}{\leq_{S_{max}}} C_*\varepsilon^{b_*+1}\gamma^{-2}N_{n-1}^{-\alpha} \\ \leq_{S_{max}} \varepsilon^{4-3a}N_n^{-(2/3)\alpha}. \tag{5.9.38}$$

We have, by (5.8.22), for any $n \in \mathbb{N}$

$$\langle j\rangle|r_j^\infty(i_n) - r_j^{n+1}(i_n)| \leq \langle j\rangle \sum_{k\geq n+1}|r_j^{k+1}(i_n) - r_j^k(i_n)| \leq \mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b})\sum_{k\geq n}N_k^{-\mathtt{a}} \\ \overset{(5.8.9)}{\leq} \varepsilon^{4-3a}N_n^{-\mathtt{a}}. \tag{5.9.39}$$

Therefore $\forall\omega \in \mathcal{G}_n$, $\forall j \in \mathbb{Z}$ we have

$$\langle j\rangle|r_j^\infty(i_n) - r_j^\infty(i_{n-1})| \leq \langle j\rangle\Big(|r_j^{n+1}(i_n) - r_j^{n+1}(i_{n-1})| + |r_j^\infty(i_n) - r_j^{n+1}(i_n)| \\ + |r_j^\infty(i_{n-1}) - r_j^{n+1}(i_{n-1})|\Big) \overset{(5.9.38),(5.9.39)}{\leq_{S_{max}}} \varepsilon^{4-3a}N_n^{-(2/3)\alpha} + \mathfrak{M}_0^{\sharp,\gamma_*}(s_0,\mathtt{b})N_n^{-\mathtt{a}} \leq_{S_{max}} \varepsilon^{4-3a}N_n^{-\mathtt{a}}.$$

Now we prove that $Q_{\ell j}(i_{n-1}) \subseteq Q_{\ell j}(i_n)$. We have

$$|m(i_n) - m(i_{n-1})||j| \overset{(5.7.331)}{\leq} C\varepsilon^3\|i_n - i_{n-1}\|_{s_0+2}|j| \overset{(5.9.7)}{\leq} C\varepsilon^{b_*+3}\gamma^{-1}N_{n-1}^{-\alpha}|j| \\ \leq C\varepsilon^{b_*+3}\gamma^{-1}N_{n-1}^{-\alpha}|\ell| \tag{5.9.40}$$

and then

$$\begin{aligned} |\omega\cdot\ell+m(i_n)j| &\ge |\omega\cdot\ell+m(i_{n-1})j|-|m(i_n)-m(i_{n-1})||j| \\ &\ge 2\gamma_{n-1}\langle\ell\rangle^{-\tau}-\varepsilon^{b_*+3}\gamma^{-1}N_{n-1}^{-\alpha+1}\ge 2\gamma_n\langle\ell\rangle^{-\tau} \end{aligned} \tag{5.9.41}$$

since $|\ell|\le N_{n-1}$.

Now we prove that $P_{\ell j}(i_{n-1})\subseteq P_{\ell j}(i_n)$.

We claim that for all $j,k\in\mathbb{Z}$ we have (recall (5.8.14))

$$|d_j^\infty(i_n)-d_j^\infty(i_{n-1})|\le\varepsilon^{4-3a}N_n^{-\mathtt{a}} \qquad \forall\omega\in\mathcal{G}_n. \tag{5.9.42}$$

We first prove that (5.9.31) implies that $P_{\ell jk}(i_n)\subseteq P_{\ell jk}(i_{n-1})$. For all $j\neq k$, $|\ell|\le N_{n-1}$, $\omega\in\mathcal{G}_n$ by (5.9.31)

$$\begin{aligned} |\omega\cdot\ell+d_j^\infty(i_n)| &\ge |\omega\cdot\ell+d_j^\infty(i_{n-1})|-|d_j^\infty(i_n)-d_j^\infty(i_{n-1})| \\ &\ge 2\gamma_{n-1}\langle\ell\rangle^{-\tau}-\varepsilon^{4-3a}N_n^{-\mathtt{a}}\ge 2\gamma_n\langle\ell\rangle^{-\tau} \end{aligned} \tag{5.9.43}$$

since $\varepsilon^{4-3a}\gamma^{-1}N_n^{\tau-(2/3)\mathtt{a}}2^{n+1}\le 1$. The proof of (5.9.42) follows by the same arguments used for the proof of (5.9.31).

□

By definition, $R_{\ell jk}(i_n),Q_{\ell j}(i_n),P_{\ell j}(i_n)\subseteq\mathcal{G}_n$ (see (5.9.27)). By Lemma 5.9.6, for $n\ge 1$ and $|l|\le N_{n-1}$ we also have $R_{\ell jk}(i_n)\subseteq R_{\ell jk}(i_{n-1})$, $Q_{\ell j}(i_n)\subseteq Q_{\ell j}(i_{n-1})$ and $P_{\ell j}(i_n)\subseteq P_{\ell j}(i_{n-1})$. On the other hand, $R_{\ell jk}(i_n)\cap\mathcal{G}_n=\emptyset$, $Q_{\ell j}(i_n)\cap\mathcal{G}_n=\emptyset$ and $P_{\ell j}(i_n)\cap\mathcal{G}_n=\emptyset$ (see (5.9.5)). As a consequence, $R_{\ell jk}(i_n),Q_{\ell j}(i_n),P_{\ell j}(i_n)=\emptyset$ for all $|\ell|\le N_{n-1}$, and

$$\mathcal{G}_n\setminus\mathcal{G}_{n+1}\subseteq\bigcup_{\substack{j,k\in S^c\\ |\ell|>N_{n-1}}}\Big(R_{\ell jk}(i_n)\cup Q_{\ell j}(i_n)\cup P_{\ell j}(i_n)\Big),\quad\forall n\ge 1. \tag{5.9.44}$$

By (5.9.27), we have to bound the measure of the sublevels of the function $\omega\mapsto\phi(\omega)$ defined by

$$\begin{aligned} &\phi_R(\omega):=\mathrm{i}\omega\cdot\ell+d_j^\infty(\omega)-d_k^\infty(\omega)=\mathrm{i}\omega\cdot\ell+\mathrm{i}m(\omega)(\omega(j)-\omega(k))+\mathrm{i}\varepsilon^2(\kappa_j-\kappa_k)+(r_j^\infty-r_k^\infty)(\omega),\\ &\phi_Q(\omega):=\mathrm{i}\omega\cdot\ell+m(\omega)j. \end{aligned} \tag{5.9.45}$$

Note that ϕ also depends on ℓ,j,k,i_n. We recall that

$$\begin{aligned} &m=1+\varepsilon^2c(\xi)+\mathtt{r}_m(\omega), \qquad c(\xi)=\vec v\cdot\xi,\quad \vec v\in\mathbb{R}^\nu\quad \vec v:=(2/3)(1+\bar\jmath_k^2)_{k=1}^\nu\\ &\kappa_j(\xi)=w_j\cdot\xi,\qquad w_j\in\mathbb{R}^\nu,\quad w_j:=-(2/3)\omega(j)\mathtt{G}_j(\bar\jmath_1,\dots,\bar\jmath_\nu)\\ &\mathtt{G}_j(\bar\jmath_1,\dots,\bar\jmath_\nu):=\left(\frac{(1+\bar\jmath_k^2)(7+5\bar\jmath_k^2+\bar\jmath_k^4+3j^2)}{(3+j^2)^2+(6+j^2)\bar\jmath_k^2+\bar\jmath_k^4}\right)_{k=1}^\nu\in\mathbb{R}^\nu \end{aligned} \tag{5.9.46}$$

where

$$\mathtt{r}_m:=\varepsilon^4d(\xi)+O(\varepsilon^5),\quad |\nabla_\omega\mathtt{r}_m|\le C(S)\varepsilon^2+O(\varepsilon^5\gamma^{-1}) \tag{5.9.47}$$

$$|\mathtt{r}_m|^\gamma\le C\varepsilon^4,\qquad |\mathtt{G}_j|\le\mathtt{C}|j|^{-2} \tag{5.9.48}$$

with $\mathtt{C}=\mathtt{C}(S)$. It will be useful to consider $\phi(\omega)$ in (5.9.45) as a small perturbation of an affine function in ω. We write it as

$$\begin{aligned} &\phi_R(\omega):=a_{jk}+b_{\ell jk}\cdot\omega+q_{jk}(\omega),\qquad \ell\in\mathbb{Z}^\nu,\ j,k\in S^c,\\ &\phi_Q(\omega):=f_j+g_{\ell j}\cdot\omega+h_j(\omega),\qquad \ell\in\mathbb{Z}^\nu,\ j,k\in S^c, \end{aligned} \tag{5.9.49}$$

where, by (5.3.19), and calling

$$\mathtt{f}(\xi) := -\varepsilon^{-2}\mathbb{A}^{-1}(\omega - \overline{\omega} - \varepsilon^2\mathbb{A}\xi), \quad |\mathtt{f}(\xi)|^{sup} \le C\varepsilon^4, \quad |\mathtt{f}(\xi)|^{lip} \le C\varepsilon^2 \tag{5.9.50}$$

$$a_{jk} := \mathrm{i}\Big((\omega(j) - \omega(k))[1 - c(\mathbb{A}^{-1}\overline{\omega})] + (w_k - w_j)\cdot \mathbb{A}^{-1}\overline{\omega}\Big), \qquad f_j := \mathrm{i}j\Big(1 - c(\mathbb{A}^{-1}\overline{\omega})\Big), \tag{5.9.51}$$

$$b_{ljk} := \mathrm{i}\Big(\ell + \mathbb{A}^{-T}v(\omega(j) - \omega(k)) + \mathbb{A}^{-T}(w_j - w_k)\Big), \qquad g_{lj} := \mathrm{i}\Big(\ell + j\,\mathbb{A}^{-T}\vec{v}\Big), \tag{5.9.52}$$

$$q_{jk}(\omega) := (\mathtt{r}_m(\omega) + \vec{v}\cdot \mathtt{f}(\xi))(\omega(j) - \omega(k)) + \varepsilon^2(w_j - w_k)\cdot \mathtt{f}(\xi) + r_j^\infty(\omega) - r_k^\infty(\omega), \tag{5.9.53}$$

$$h_j := \mathtt{r}_m(\omega)j + (\vec{v}\cdot \mathtt{f}(\xi))j \tag{5.9.54}$$

and by (5.9.48), (5.9.54),

$$\begin{aligned} &|q_{jk}(\omega)|^{sup} \le C\varepsilon^4|j-k| + \varepsilon^{4-3a}, \\ &|q_{jk}(\omega)|^{lip} \le |\mathtt{r}_m(\omega)|^{lip}|\omega(j) - \omega(k)| + |r_j^\infty - r_k^\infty|^{lip} \le C(S)\varepsilon^2|j-k| + \varepsilon^{1-4a}, \\ &|h_j(\omega)|^{sup} \le C\varepsilon^3|j|, \qquad |h_j(\omega)|^{lip} \le C\varepsilon^3\gamma^{-1}|j|. \end{aligned} \tag{5.9.55}$$

By Lemma 5.9.5 it is sufficient to study the measure of the resonant sets $R_{\ell jk}(i_n)$ defined in (5.9.27) for $(\ell, j, k) \neq (0, j, j)$. In particular we will prove the following Lemma.

Lemma 5.9.7. *Let us define for $\eta \le C(S)\sqrt{\varepsilon}$ and $\sigma \in \mathbb{N} > 0$*

$$\begin{aligned} R_{\ell jk}(\eta, \sigma) &:= \{\omega : |\omega\cdot\ell + d_j^\infty - d_k^\infty| \le \frac{2\eta}{\langle \ell\rangle^\sigma}\}, \\ Q_{\ell j}(\eta, \sigma) &:= \{\omega : |\omega\cdot\ell + mj| \le \frac{2\eta}{\langle \ell\rangle^\sigma}\}, \\ P_{\ell j}(\eta, \sigma) &:= \{\omega : |\omega\cdot\ell + d_j^\infty| \le \frac{2\eta}{\langle \ell\rangle^\sigma}\}. \end{aligned} \tag{5.9.56}$$

Then for a generic choice of the tangential sites we have that $|R_{\ell jk}(\eta, \sigma)| \le \varepsilon^{2(\nu-1)}\eta\langle\ell\rangle^{-\sigma}$. The same holds for $Q_{\ell j}(\eta, \sigma)$ and $P_{\ell j}(\eta, s)$.

We give the proof of Lemma 5.9.7 for the set $R_{\ell jk}$, which is the most difficult case.

Lemma 5.9.8. *Let $\gamma_0 = \varepsilon^{1/4}$ and $\tau_0 \in \mathbb{N}$. If $|\ell| \le \varepsilon^{-1/(4\tau_0)}$ and*

$$|\overline{\omega}\cdot\ell + \omega(j) - \omega(k)| \ge \gamma_0\langle\ell\rangle^{-\tau_0} \tag{5.9.57}$$

then $R_{\ell jk}(\eta, \sigma) = \emptyset$.

Proof. We have by Lemma 5.9.5 (recall (5.9.45))

$$\begin{aligned} |\omega\cdot\ell + d_j^\infty - d_k^\infty| &\ge \gamma_0\langle\ell\rangle^{-\tau_0} - |m-1||\omega(j) - \omega(k)| - C(S)\,\varepsilon^2\left(\frac{|j-k|}{\min\{|j|,|k|\}}\right) - (r_j^\infty - r_k^\infty) \\ &\ge \gamma_0\langle\ell\rangle^{-\tau_0} - C(S)\varepsilon^{2-1/(4\tau_0)} \ge \sqrt{\varepsilon} - C(S)\varepsilon^{3/2} \ge \frac{\sqrt{\varepsilon}}{2}. \end{aligned} \tag{5.9.58}$$

□

Lemma 5.9.9. *Let $\gamma_0 = \varepsilon^{1/4}$ and $\tau_0 \in \mathbb{N} > 0$. If $|\ell| \le \varepsilon^{-1/(4\tau_0)}$ and*

$$|\overline{\omega} \cdot \ell + \omega(j) - \omega(k)| \le \gamma_0 \langle \ell \rangle^{-\tau_0} \tag{5.9.59}$$

then $|R_{\ell jk}(\eta, \sigma)| \le \varepsilon^{2(\nu-1)} \eta \langle \ell \rangle^{-\sigma}$.

Proof. Let us call $p := \overline{\omega} \cdot \ell + \omega(j) - \omega(k)$ and note that $p = C\varepsilon^{1/4}$. By definition (see (5.9.52), (5.9.54))

$$\nabla_\omega \Phi_R(\omega) = \mathrm{i} b_{\ell jk} + \nabla_\omega q_{jk} = \mathrm{i}\Big(\ell - \mathbb{A}^{-T} \vec{v}(\overline{\omega} \cdot \ell) - \mathbb{A}^{-T}(w_j - w_k) + O(|\ell| \varepsilon^2) + O(\varepsilon^{1/4})\Big)$$

since

$$|r_j^\infty|^{lip} = O(\varepsilon^{1-}), \qquad |\nabla_\omega \mathtt{r}_m(\omega)|^{sup} |\omega(j) - \omega(k)| \le C(S)\varepsilon, \quad |\mathbb{A}^{-1}(\mathtt{G}_j - \mathtt{G}_k))\, p|^{sup} \le C\varepsilon^{1/4}. \tag{5.9.60}$$

By the assumption (1.2.16) we have

$$\Big(\mathrm{I} - \mathbb{A}^{-T} \vec{v}(\overline{\omega})^T\Big)\ell \ne \mathbb{A}^{-T}(w_j - w_k), \qquad j, k \in S^c. \tag{5.9.61}$$

Thus $|b_{\ell jk}| \ge c$ for some c independent of ε.

□

Lemma 5.9.10. *If $|\ell| \ge \varepsilon^{-1/(4\tau_0)}$ then $|R_{\ell jk}(\eta, \sigma)| \le \varepsilon^{2(\nu-1)} \eta \langle \ell \rangle^{-\sigma}$.*

Proof. We consider two cases: $|j-k| \ge \mathtt{C}_5$ and $|j-k| < \mathtt{C}_5$ for some constant $\mathtt{C}_5 > 0$ to be determined.

In the first case we have (recall (5.9.48))

$$|a_{jk}| \ge |j-k| \Big(|1 - \mathbb{A}^{-1} \overline{\omega} \cdot \vec{v}| - \frac{\mathbb{A}^{-T}(w_j - w_k)}{|j-k|} | \Big) \ge \delta |j-k| \tag{5.9.62}$$

with $\delta > 0$ for $\mathtt{C}_5$ large enough. In particular the constant δ does not depend on S by Lemma B.1.1.

By (5.9.27), (5.9.55)

$$|b_{\ell jk} \cdot \omega| \ge |a_{jk}| - |\phi_R(\omega)| - |q_{jk}(\omega)| \ge (\delta - \frac{\eta}{\langle \ell \rangle^\sigma \mathtt{C}_5} - C\varepsilon^3)|j-k| \ge \frac{\delta}{2}|j-k|,$$

for ε small enough and $\sigma \ge 1$.

If $b := b_{\ell jk}$ we have $|b \cdot \omega| \le 2|b||\overline{\omega}|$, because $|\omega| \le 2|\overline{\omega}|$. Hence $|b| \ge \delta_1 |j-k|$ where $\delta_1 := \delta/(2|\overline{\omega}|)$. Split $\omega = s\hat{b} + v$ where $\hat{b} := b/|b|$ and $v \cdot b = 0$. Let $\Psi_R(s) := \phi(s\hat{b} + v)$. For ε small enough, by (5.9.55), we get

$$\begin{aligned} |\Psi_R(s_1) - \Psi_R(s_2)| &\ge (|b| - |q_{jk}|^{lip})|s_1 - s_2| \ge \big(\delta_1 - \varepsilon^3 \gamma^{-1}\big)\, |j-k|\, |s_1 - s_2| \\ &\ge \frac{\delta_1}{2} |j-k|\, |s_1 - s_2|. \end{aligned} \tag{5.9.63}$$

As a consequence, the set $\Delta_{\ell jk}(i_n) := \{s : s\hat{\ell} + v \in R_{ljk}(i_n)\}$ has Lebesgue measure

$$|\Delta_{\ell jk}(i_n)| \le \frac{2}{\delta_1 |j-k|} \frac{4\eta}{\langle \ell \rangle^\tau} \le \frac{C\eta}{\langle \ell \rangle^\tau}$$

for some $C>0$. The Lemma follows by Fubini's theorem.

In the second case we estimate directly the derivative of $\Phi_R(\omega)$, namely

$$|b_{\ell jk}| \geq |\ell| - \mathtt{C}_5|\mathbb{A}^{-T}\vec{v}| - |\mathbb{A}^{-T}(w_j - w_k)| \geq \varepsilon^{-1/(4\tau_0)} - C(S) > 0 \tag{5.9.64}$$

for ε small enough. We conclude as in (5.9.63). □

The proof of Lemma 5.9.7 for the sets $R_{\ell jk}$ follows by Lemmata 5.9.8, 5.9.9, 5.9.10. The proof for the sets Q_{lj} and $P_{\ell j}$ follows using the same arguments used for $R_{\ell jk}$.

Lemma 5.9.11. *Let* $|j|,|k| \geq \mathtt{C}_6\langle\ell\rangle^{\tau_1}\gamma^{-(1/2)}$ *with* $\tau \geq \tau_1 \geq \tau_2$ *then*

$$R_{\ell jk}(\gamma^{3/2},\tau) \subseteq Q_{\ell,j-k}(\gamma,\tau_2). \tag{5.9.65}$$

Proof. We have that

$$\begin{aligned}|\omega\cdot\ell + d_j^\infty - d_k^\infty| &\geq |\omega\cdot\ell + m(j-k)| - |m||\omega(j) - j + k - \omega(k)| - \varepsilon^2|w_j - w_k| - |r_j^\infty| - |r_k^\infty| \\ &\geq \frac{2\gamma}{\langle\ell\rangle^{\tau_2}} - 2|j-k|\frac{C}{|j||k|} - \frac{\tilde{C}\varepsilon^2}{\min\{|j|,|k|\}} \\ &\geq \frac{2\gamma}{\langle\ell\rangle^{\tau_2}} - \frac{C\gamma}{\mathtt{C}_6\langle\ell\rangle^{2\tau_1-1}} - \frac{\tilde{C}\varepsilon^2\sqrt{\gamma}}{\mathtt{C}_6\langle\ell\rangle^{\tau_1}} \geq \frac{\gamma}{\langle\ell\rangle^{\tau_2}}\Big(2 - \frac{C}{2\mathtt{C}_6\langle\ell\rangle^{2\tau_1-\tau_2-1}} - \frac{\tilde{C}\varepsilon^2}{2\sqrt{\gamma}\mathtt{C}_6\langle\ell\rangle^{\tau_1-\tau_2}}\Big) \\ &\geq \frac{\gamma}{\langle\ell\rangle^{\tau_2}} \geq \frac{\gamma^{3/2}}{\langle\ell\rangle^{\tau}}\end{aligned} \tag{5.9.66}$$

for $\mathtt{C}_6$ big enough and since $\varepsilon^2(\sqrt{\gamma})^{-1} \ll 1$. □

We are in position to prove (5.9.8). Let $\tau > \max\{\tau_1+\nu+2,\tau_2\}$ and $\tau_2 - 1 \geq (\nu+1)$. We have by Lemma 5.9.27

$$\Big|\bigcup_{\ell\in\mathbb{Z}^\nu, j,k\in S^c} R_{\ell jk}(i_n)\Big| \leq \sum_{|\ell|>N_{n-1},|j|,|k|\geq\mathtt{C}_6\langle\ell\rangle^{\tau_1}\gamma^{-(1/2)}} |R_{\ell jk}(i_n)| + \sum_{|\ell|>N_{n-1},|j|,|k|\leq\mathtt{C}_6\langle\ell\rangle^{\tau_1}\gamma^{-(1/2)}} |R_{\ell jk}(i_n)|.$$

On one hand we have that, using Lemma 5.9.11,

$$\begin{aligned}\sum_{|\ell|>N_{n-1},|j|,|k|\geq\mathtt{C}_6\langle\ell\rangle^{\tau_1}\gamma^{-(1/2)}} |R_{\ell jk}(i_n)| &\leq C\sum_{j-k=h,|h|\leq C|\ell|}\varepsilon^{2(\nu-1)}\gamma\langle\ell\rangle^{-\tau_2} \leq C\varepsilon^{2(\nu-1)}\gamma\sum_{|\ell|\geq N_{n-1}}\langle\ell\rangle^{-(\tau_2-1)} \\ &\leq C\varepsilon^{2(\nu-1)}\gamma N_{n-1}^{-1},\end{aligned}$$

for some $C>0$. On the other hand

$$\begin{aligned}\sum_{|\ell|>N_{n-1},|j|,|k|\leq\mathtt{C}_6\langle\ell\rangle^{\tau_1}\gamma^{-(1/2)},|j-k|\leq C|\ell|} |R_{\ell jk}(i_n)| &\leq C\gamma^{(3/2)}\varepsilon^{2(\nu-1)}\sum_{|\ell|\geq N_{n-1}}\frac{|\ell|\langle\ell\rangle^{\tau_1}}{\sqrt{\gamma}\langle\ell\rangle^{\tau}} \\ &\leq C\gamma\varepsilon^{2(\nu-1)}\sum_{|\ell|\geq N_{n-1}}\langle\ell\rangle^{-(\tau-\tau_1-1)} \\ &\leq C\gamma\varepsilon^{2(\nu-1)}N_{n-1}^{-1}.\end{aligned}$$

We conclude by fixing $\tau_2 = \tau_1 = \nu+2$. Recalling (5.8.1) we check that actually $\tau > \tau_1+\nu+2$. The discussion above implies estimates (5.9.8).

□

APPENDIX A

Technical Lemmata

In this Section we present standard tame and Lipschitz estimates for composition of functions and changes of variables which are used in the Thesis. We refer to the Appendix of [7] and [5] (and references therein) for more details. In particular in [7] the results below are proved for the Lipschitz norm (2.1.8).

Let us denote $W^{s,\infty} := W^{s,\infty}(\mathbb{T}^d, \mathbb{C})$ and $L^\infty := L^\infty(\mathbb{T}^d, \mathbb{C})$ with $d \geq 1$. The norms of these spaces are respectively indicated with $|\cdot|_{s,\infty}$, $|\cdot|_{L^\infty} := |\cdot|_{0,\infty}$ and are defined by

$$|u|_{L^\infty} := \sup_{x\in\mathbb{T}^d} |u(x)|, \quad |u|_{s,\infty} := \sum_{s_1\leq s} |D^{s_1}u|_{L^\infty}, \quad |D^{s_1}u|_{L^\infty} := \sup_{|\vec{s}_1|=s_1} |\partial_x^{\vec{s}_1} u|_{L^\infty}, \tag{A.0.1}$$

here D^s is the s-th Fréchet derivative with respect to x, hence D^s is a symmetric multi-linear operator.

Let us denote with $H^s := H^s(\mathbb{T}^d, \mathbb{C})$ the space of Sobolev functions on $\mathbb{T}^d$ defined in (2.1.2). We shall actually use the equivalent norm

$$\|u\|_s := \|u\|_{H^s(\mathbb{T}^d)} := \|u\|_{L^2(\mathbb{T}^d)} + \|D^s u\|_{L^2(\mathbb{T}^d)}, \quad \|D^s u\|_{L^2(\mathbb{T}^d)} := \sup_{|\vec{s}|=s} \|\partial_x^{\vec{s}} u\|_{L^2(\mathbb{T}^d)}. \tag{A.0.2}$$

Remark A.0.1. In the following Lemmata the estimates which hold for the Lipschitz norm $\|\cdot\|_s^{Lip(\gamma)}$, defined in (2.1.8), hold also for the slightly different Lipschitz norm $\|\cdot\|_s^{\gamma,\mathcal{O}}$ defined in (2.1.9). One can repeat the proofs of these Lemmata for the bounds of the variation $(u(\omega) - u(\omega'))/(\omega - \omega')$ (see for instance the Appendix of [7]) and then pass to the equivalent (to $\|\cdot\|_s^{\gamma,\mathcal{O}}$) norm $\max\{\|u\|_s^{sup}, \gamma\|u\|_{s-1}^{lip}\}$ to prove the same bounds for $\|\cdot\|_s^{\gamma,\mathcal{O}}$.
We remark that the main difference in estimating with the norm $\|\cdot\|_s^{\gamma,\mathcal{O}}$ is for the bound (A.0.8) in Lemma A.0.5.

Lemma A.0.2. *Let $s_0 > d/2 + 1$. Then the following holds.*

(*i*) **Embedding**. $|u|_{L^\infty} \leq \|u\|_{s_0}$ *for all* $u \in H^{s_0}$.

(*ii*) **Algebra**. $\|uv\|_{s_0} \leq C(s_0)\|u\|_{s_0}\|v\|_{s_0}$ *for all* $u, v \in H^{s_0}$.

(*iii*) **Interpolation**. *For* $0 \leq s_1 \leq s \leq s_2$, $s = \lambda s_1 + (1-\lambda)s_2$, $\lambda \in [0,1]$,

$$\|u\|_s \leq \|u\|_{s_1}^{\lambda} \|u\|_{s_2}^{1-\lambda}, \quad \forall u \in H^{s_2}.$$

Let $a_0, b_0 \geq 0$ *and* $p, q > 0$. *For all* $u \in H^{a_0+p+q}$, $v \in H^{b_0+p+q}$

$$\|u\|_{a_0+p}\|v\|_{b_0+q} \leq \|u\|_{a_0+p+q}\|v\|_{b_0} + \|u\|_{a_0}\|v\|_{b_0+p+q}.$$

Similarly

$$|u|_{s,\infty} \le C(s_1,s_2)|u|_{s_1,\infty}^{\lambda}|v|_{s_2,\infty}^{1-\lambda} \quad \forall u \in W^{s_2,\infty}$$

and for all $u \in W^{a_0+p+q}$, $v \in W^{b_0+p+q}$

$$|u|_{a_0+p,\infty}|v|_{b_0+q,\infty} \le C(a_0,b_0,p,q)|u|_{a_0+p+q,\infty}|v|_{b_0,\infty} + |u|_{a_0,\infty}|v|_{b_0+p+q,\infty}.$$

(iv) **Asymmetric tame product**. *For* $s \ge s_0$

$$\|u\,v\|_s \le C(s_0)\|u\|_s\|v\|_{s_0} + C(s)\|u\|_{s_0}\|v\|_s, \quad \forall u,v \in H^s.$$

(v) **Asymmetric tame product in** $W^{s,\infty}$. *For* $s \in \mathbb{N}$

$$|uv|_{s,\infty} \le \frac{3}{2}|u|_{L^\infty}|v|_{s,\infty} + C(s)|u|_{s,\infty}|v|_{L^\infty}, \quad \forall u,v \in W^{s,\infty}.$$

If $u = u(\omega)$ *and* $v = v(\omega)$ *depend in a Lipschitz way on a parameter* $\mathcal{O} \subset \mathbb{R}^\nu$, *all the previous statements hold by replacing* $|\cdot|_{s,\infty}$, $\|\cdot\|_s$ *with the Lipschitz norms* $|\cdot|_{s,\infty}^{Lip(\gamma)}$, $\|\cdot\|_s^{Lip(\gamma)}$.

The following Lemma is a classical result on tame estimates for the composition of functions. We refer to [81] and Lemma $B.3$-(iii) in the Appendix of [5] for the proof.

Fix $p \in \mathbb{N}$. A function $f\colon \mathbb{T}^d \times B_1 \to \mathbb{C}$, where $B_1 := \{y \in \mathbb{R}^m : |y| < 1\}$, induces the composition operator

$$\tilde{f}(u)(x) := f(x, u(x), Du(x), D^2u(x), \dots, D^p u(x)) \tag{A.0.3}$$

where $D^k u(x)$, $k = 1, \dots, p$ denotes the partial derivatives $\partial_x^\alpha u$ of order $|\alpha| = k$.

Lemma A.0.3. (Composition of functions) *Assume* $f \in C^r(\mathbb{T}^d \times B_1)$. *Then for all* $u \in H^{r+p}$ *such that* $|u|_{p,\infty} < 1$, *the composition operator* (A.0.3) *is well defined and*

$$\|\tilde{f}(u)\|_r \le C\|f\|_{C^r}(\|u\|_{r+p} + 1),$$

where the constant C *depends on* r, d, p. *If* $f \in C^{r+2}$ *then for all* $|u|_{p,\infty}, |h|_{p,\infty} < 1/2$,

$$\|\tilde{f}(u+h) - \sum_{i=0}^{k} \frac{\tilde{f}^{(i)}(u)}{i!}[h^i]\|_r \le C\,\|f\|_{C^{r+2}}\|h\|_{L^\infty}^k(\|h\|_{r+p} + |h|_{L^\infty}\|u\|_{r+p}). \tag{A.0.4}$$

The same holds by replacing $\|\cdot\|_r$ *with the norm* $|\cdot|_{r,\infty}$.

Lemma A.0.4. *Lemma* 6.3 *in [7] Let* $d \in \mathbb{N}$, $d/2 < s_0 \le s$, $p \ge 0$, $\gamma > 0$. *Let* F *be a* C^1*-map satisfying the tame estimates: for all* $\|u\|_{s_0+p} \le 1, h \in H^{s+p}$,

$$\|F(u)\|_s \le C(s)(1 + \|u\|_{s+p}),$$
$$\|\partial_u F(u)[h]\|_s \le C(s)(\|h\|_{s+p} + \|u\|_{s+p}\|h\|_{s_0+p}).$$

For $\mathcal{O} \subset \mathbb{R}^\nu$, *let* $u(\omega)$ *be a Lipschitz family of functions parametrized by* $\omega \in \mathcal{O}$ *with* $\|u\|_{s_0+p}^{Lip(\gamma)} \le 1$. *Then*

$$\|F(u)\|_s^{Lip(\gamma)} \le C(s)(1 + \|u\|_{s+p}^{Lip(\gamma)}).$$

The following Lemma is classical, see for instance Appendix G of [69].

Lemma A.0.5. (Change of variable) *Let $p : \mathbb{R}^d \to \mathbb{R}^d$ be a $2\pi-$periodic in each variable function in $W^{s,\infty}$, $s \geq 1$, with $|p|_{1,\infty} \leq 1/2$. Let $f(x) = x + p(x)$. Then:*
(i) f is invertible, its inverse is $f^{-1}(y) = g(y) = y + q(y)$ where q is $2\pi-$periodic in each variable, $q \in W^{s,\infty}(\mathbb{T}^d; \mathbb{R}^d)$ and $|q|_{s,\infty} \leq C|p|_{s,\infty}$. More precisely,

$$|q|_{L^\infty} = |p|_{L^\infty}, \ |Dq|_{L^\infty} \leq 2|Dp|_{L^\infty}, \ |Dq|_{s-1,\infty} \leq C|Dp|_{s-1,\infty}, \tag{A.0.5}$$

where the constant C depends on d, s.

Moreover, assume that $p = p_\lambda$ depends in a Lipschitz way by a parameter $\lambda \in \mathcal{O} \subset \mathbb{R}^\nu$, and suppose, as above, that $|D_x p_\lambda|_{L^\infty} \leq 1/2$ for all λ. Then $q = q_\lambda$ is also Lipschitz in λ, and

$$|q|_{s,\infty}^{Lip(\gamma)} \leq C \left(|p|_{s,\infty}^{Lip(\gamma)} + \{\sup_{\lambda \in \mathcal{O}} |p_\lambda|_{s+1,\infty}\} |p|_{L^\infty}^{Lip(\gamma)} \right) \leq C|p|_{s+1,\infty}^{Lip(\gamma)}, \tag{A.0.6}$$

the constant C depends on d, s (it is independent on γ).

(ii) If $u \in H^s(\mathbb{T}^d; \mathbb{C})$, then $u \circ f(x) = u(x + p(x)) \in H^s$, and, with the same C as in (i) one has

$$\|u \circ f\|_s \leq \|u\|_s + C(\|u\|_s |p|_{1,\infty} + |Dp|_{s-1,\infty} \|u\|_1), \tag{A.0.7a}$$

$$\|u \circ f - u\|_s \leq C(|p|_{L^\infty} \|u\|_{s+1} + |p|_{s,\infty} \|u\|_2), \tag{A.0.7b}$$

$$\|u \circ f\|_s^{Lip(\gamma)} \leq C(\|u\|_{s+1}^{Lip(\gamma)} + |p|_{s,\infty}^{Lip(\gamma)} \|u\|_2^{Lip(\gamma)}). \tag{A.0.7c}$$

Moreover

$$\|u \circ f\|_s^{\gamma,\mathcal{O}} \leq \|u\|_s^{\gamma,\mathcal{O}} + C(\|u\|_s^{\gamma,\mathcal{O}} \|p\|_{s_0+1}^{\gamma,\mathcal{O}} + \|p\|_{s+1}^{\gamma,\mathcal{O}} \|u\|_{s_0}^{\gamma,\mathcal{O}}), \tag{A.0.8}$$

the constant C depends on d, s (it is independent on γ).

(iii) Part (ii) holds also with the Sobolev and Lipschitz norms replaced by $|\cdot|_{s,\infty}$ and $|\cdot|_{s,\infty}^{Lip(\gamma)}$.

Proof. All the items are proved in [5] and the Appendix of [7], except for the bounds (A.0.7a) and (A.0.8). These estimates are easily proved by following the proof of Lemma $B.4$-(ii) in [5] and by treating in a different way some terms arising from the Faa di Bruno's formula. More precisely, we consider the expression

$$D^s(u \circ f) = \sum_{k=1}^{s} \sum_{j_1+\dots+j_k=s,\, j_i \geq 1} C_k \ (D^k u)[D^{j_1} f, \dots, D^{j_k} f]$$

and we note that if $j_i = 1$ then $D^{j_i} f = \mathrm{I} + Dp$. We split the sum above in the following way

$$\begin{aligned} D^s(u \circ f) = D^s u + \sum_{r=1}^{s} C_r \ (D^s u)[\underbrace{Dp, \dots, Dp}_{r \text{ times}}, \underbrace{\mathrm{I}, \dots, \mathrm{I}}_{s-r \text{ times}}] \\ + \sum_{k=1}^{s-1} \sum_{\substack{j_1+\dots+j_k=s, \\ \prod_i j_i > 1}} C_k \ (D^k u)[D^{j_1} f, \dots, D^{j_k} f]. \end{aligned} \tag{A.0.9}$$

Thus from the first term of the right hand side of (A.0.9) we get the first term of the right hand side of (A.0.7a). For the remaining terms one can follow exactly the same proof of Lemma $B.4$-(ii) in [5].
The bound (A.0.8) follows by (A.0.7a) for the estimate of $\sup_\lambda \|u\circ f\|_s$. For the Lipschitz variation of λ the proof follows the arguments used for (A.0.7c), see the Appendix in [7]. We note that $u(x+p_\lambda(x)) - u(x+p_{\lambda'}(x))$ is estimated with the x-derivative of u, but the norm $\|\cdot\|^{\gamma,\mathcal{O}}$ requires the estimate of the Lipschitz variation of λ only for the norm $\|\cdot\|_{s-1}$. This is why in (A.0.7c) there is a loss of one derivative which does not appear in (A.0.8). □

We now give a lemma on symbols defined on $\mathbb{T}^d$. Recalling Definition 2.2.2 and (2.2.10) we define

$$|Aw|_{m,s,\alpha} := \sup_{\xi\in\mathbb{R}^d} \max_{0\le|\vec{\alpha}_1|\le\alpha} \|\partial_\xi^{\vec{\alpha}_1} Aw\|_s \langle\xi\rangle^{-m+|\vec{\alpha}_1|} \tag{A.0.10}$$

where

$$\partial_y^{\vec{\alpha}} := \prod_{i=1}^d \partial_{y_i}^{\alpha^{(i)}}, \quad \vec{\alpha} := (\alpha^{(1)},\dots,\alpha^{(d)}).$$

Lemma A.0.6. *Let $\mathcal{O}$ be a subset of $\mathbb{R}^\nu$. Let $p = p_\lambda$ as in the previous lemma, let A be the linear operator defined for all $w = w_\lambda(x,\xi) \in S^m(\mathbb{T}^d)$, $\lambda\in\mathcal{O}$, as*

$$Aw = w(f(x), g(x)\xi), \quad f(x) := x + p(x), \quad g(x) = (\mathrm{I} + Dp)^{-1}, \quad x\in\mathbb{T}^d, \xi\in\mathbb{R}^d \tag{A.0.11}$$

such that $\|p\|^{\gamma,\mathcal{O}}_{2s_0+2} < 1$. Then A is bounded, namely $Aw\in S^m$, with

$$|Aw|^{\gamma,\mathcal{O}}_{m,s,\alpha} \le_{s,m,\alpha} |w|^{\gamma,\mathcal{O}}_{m,s,\alpha} + \sum_{\substack{s_1+s_2+s_3=s,\\ s_1<s, s_1,s_2,s_3\ge 0,\\ s_1+s_2\ge 1}} |w|^{\gamma,\mathcal{O}}_{m,s_1,\alpha+s_2} \|p\|^{\gamma,\mathcal{O}}_{s_3+s_0+2}. \tag{A.0.12}$$

Proof. We adopt the notation $|\cdot|_{W^{s,\infty}}$ instead of $|\cdot|_{s,\infty}$ (see (A.0.1)) in order to avoid confusion with the norm of the symbols. We also denote with D^s_ξ the s-th Fréchet derivative with respect to ξ.
We study

$$D^\alpha_\xi D^s w(f, g\xi) = \sum_{k=1}^s \sum_{\substack{r=0,\\ \sum(j_i+n_i)=s}}^k C_{krjn}\, (D_\xi^{k-r+\alpha} D^r w)[D^{j_1} f,\dots, D^{j_r} f, D^{n_1} g\,\xi,\dots, D^{n_{k-r}} g\,\xi, \underbrace{g,\dots,g}_{\alpha \text{ times}}] \tag{A.0.13}$$

where $j := (j_1,\dots,j_r)$, $n := (n_1,\dots,n_{k-r})$. In the following formulas we shall denote $\underbrace{g,\dots,g}_{\alpha\text{ times}}$ by g^α. For $k=1$ and $r=0$ we get from the expression (A.0.13) (and estimating $|g|_{L^\infty}\le 2$)

$$\|(D_\xi^{1+\alpha} w)[D^s g\,\xi, g^\alpha]\|_{L^2(\mathbb{T}^d)} \le_\alpha |w|_{m,0,\alpha+1} |D^2 p|_{W^{s-1,\infty}} \tag{A.0.14}$$

and for $r=1$

$$\|(D^\alpha_\xi Dw)[D^s f, g^\alpha]\|_{L^2(\mathbb{T}^d)} \le_\alpha |w|_{m,1,\alpha} |D^2 p|_{W^{s-2,\infty}}. \tag{A.0.15}$$

For $k = s$ we have that $j_i = n_i = 1$ for all i and we get from (A.0.13)

$$\|\sum_{r=0}^{s}(D_\xi^{s-r+\alpha}D^r w)[\underbrace{Df,\ldots,Df}_{r \text{ times}},\underbrace{Dg\,\xi,\ldots,Dg\,\xi}_{s-r \text{ times}},g^\alpha]\|_{L^2(\mathbb{T}^d)} \le \sum_{r=0}^{s}|w|_{m,r,\alpha+(s-r)}|f|^r_{W^{1,\infty}}|D^2p|^{s-r}_{L^\infty}$$
$$\le_s \sum_{\substack{s_1+s_2=s,\\ s_1,s_2\ge 0}}|w|_{m,s_1,\alpha+s_2}|D^2p|^{s_2}_{L^\infty} \le |w|_{m,s,\alpha} + \sum_{\substack{s_1+s_2=s,\\ s_1,s_2\ge 0, s_1<s}}|w|_{m,s_1,\alpha+s_2}|D^2p|_{L^\infty}. \tag{A.0.16}$$

It remains to estimate

$$\sum_{k=2}^{s-1}\sum_{\substack{r=0,\\ \sum(j_i+n_i)=s}}^{k} C_{krjn}\,(D_\xi^{k-r+\alpha}D^r w)[D^{j_1}f,\ldots,D^{j_r}f,D^{n_1}g\,\xi,\ldots,D^{n_{k-r}}g\,\xi,g^\alpha]. \tag{A.0.17}$$

We call $\ell \ge 1$ the number of indices j_i that are ≥ 2 and we rename these ones σ_i. Then $\sum_i(\sigma_i+n_i) = s-(k-\ell) = s-k+\ell$. The L^2-norm of (A.0.17) can be estimated by

$$\sum_{k=2}^{s-1}\sum_{r=0}^{k}\sum_{\ell}|w|_{m,r,\alpha+(k-r)}|Df|^{k-\ell}_{L^\infty}|D^{\sigma_1}f|_{L^\infty}\ldots|D^{\sigma_\ell}f|_{L^\infty}|D^{n_1}g|_{L^\infty}\ldots|D^{n_{k-r}}g|_{L^\infty}$$
$$\le_s \sum_{k=2}^{s-1}\sum_{r=0}^{k}\sum_{\ell}|w|_{m,r,\alpha+(k-r)}|D^{\sigma_1-2}D^2p|_{L^\infty}\ldots|D^{\sigma_\ell-2}D^2p|_{L^\infty}|D^{n_1-1}D^2p|_{L^\infty}\ldots|D^{n_{k-r}-1}D^2p|_{L^\infty}$$
$$\le_s \sum_{k=2}^{s-1}\sum_{r=0}^{k}\sum_{\ell}|w|_{m,r,\alpha+(k-r)}|D^2p|^{k+\ell-r-1}_{L^\infty}|D^2p|_{W^{s-2k-\ell+r,\infty}}$$
$$\le_s \sum_{k=2}^{s-1}\sum_{r=0}^{k}|w|_{m,r,\alpha+(k-r)}|D^2p|_{W^{s-k-1,\infty}} \le \sum_{\substack{s_1+s_2+s_3=s-1,\\ s>s_1,s_2,s_3\ge 0}}|w|_{s_1,\alpha+s_2}|D^2p|_{W^{s_3,\infty}}. \tag{A.0.18}$$

Then by (A.0.14), (A.0.15), (A.0.16), (A.0.18) we have (A.0.12) for $|Aw|_{m,s,\alpha}$. For the Lipschitz variation we observe that

$$\Delta_{\lambda,\lambda'}(w(\lambda,f(\lambda),g(\lambda)\xi)) = A(\Delta_{\lambda,\lambda'}w) + A\,Dw[\Delta_{\lambda,\lambda'}f] + A\,D_\xi w[\Delta_{\lambda,\lambda'}g\,\xi]. \tag{A.0.19}$$

One follows exactly the strategy above but considering $s-1$ derivatives instead of s (recall (2.2.11)). This is important since in formula (A.0.19) we have one extra derivative either in x or ξ. □

A.1 Properties of torus diffeomorphisms

We give some properties of $\mathcal{A}^\tau$ defined in (5.7.59). First of all we recall that $\mathcal{A}^\tau$ is the flow of

$$\begin{cases}\partial_\tau\mathcal{A}^\tau = \mathtt{X}\mathcal{A}^\tau,\\ \mathcal{A}^0 = \mathrm{I}.\end{cases}, \qquad \mathtt{X} := \partial_x\circ b, \quad b := \frac{\beta}{1+\tau\beta_x}. \tag{A.1.1}$$

Lemma A.1.1. *Assume that $\beta := \beta(\omega, i(\omega)) \in H^s(\mathbb{T}^{\nu+1})$ (see (5.7.88)) for some $s \ge s_0$, is lipschitz in $\omega \in \mathcal{O} \subseteq \Omega_\varepsilon$ and Lipschitz in the variable i. If $\|\beta\|^{\gamma,\mathcal{O}}_{s_0+\mu} \le 1$, for some $\mu \gg 1$, then, for any*

$s \geq s_0$ and $u \in H^s$ with $u = u(\omega)$ depending in a Lipschitz way on $\omega \in \mathcal{O}$, one has

$$\sup_{\tau\in[0,1]} \|\mathcal{A}^\tau u\|_s^{\gamma,\mathcal{O}} \leq_s \Big(\|u\|_s^{\gamma,\mathcal{O}} + \|\beta\|_{s+s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s_0}^{\gamma,\mathcal{O}}\Big) \tag{A.1.2}$$

$$\sup_{\tau\in[0,1]} \|(\mathcal{A}^\tau - \mathrm{I})u\|_s^{\gamma,\mathcal{O}} \leq_s \Big(\|\beta\|_{s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s+1}^{\gamma,\mathcal{O}} + \|\beta\|_{s+s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s_0}^{\gamma,\mathcal{O}}\Big) \tag{A.1.3}$$

$$\sup_{\tau\in[0,1]} \|(\mathcal{A}^\tau)^* u\|_s^{\gamma,\mathcal{O}} \leq_s \Big(\|u\|_s^{\gamma,\mathcal{O}} + \|\beta\|_{s+s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s_0}^{\gamma,\mathcal{O}}\Big) \tag{A.1.4}$$

$$\sup_{\tau\in[0,1]} \|((\mathcal{A}^\tau)^* - \mathrm{I})u\|_s^{\gamma,\mathcal{O}} \leq_s \Big(\|b\|_{s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s+1}^{\gamma,\mathcal{O}} + \|b\|_{s+s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s_0}^{\gamma,\mathcal{O}}\Big). \tag{A.1.5}$$

The inverse map $(\mathcal{A}^\tau)^{-1}$ satisfies the same estimates.

Proof. The bounds (A.1.2)-(A.1.5) in norm $\|\cdot\|_s$ follows by an explicit computation using the formula (5.7.59) and applying Lemma (A.0.5) in Appendix A. If $\beta = \beta(\omega)$ is a function of the parameters $\omega \in \mathcal{O}$, hence we need to study the term

$$\sup_{\omega_1\neq\omega_2} \frac{\|(\mathcal{A}^\tau(\omega_1) - \mathcal{A}^\tau(\omega_2))u\|_{s-1}}{|\omega_1 - \omega_2|} \tag{A.1.6}$$

in order to estimate the Lip-norm introduced in (2.1.9). We reason as follows. By (5.7.59) we have for $\omega_1, \omega_2 \in \mathcal{O}$

$$\begin{aligned}(\mathcal{A}^\tau(\omega_1) - \mathcal{A}^\tau(\omega_2))u &= (1 + \tau\beta_x(\omega_1))\big[u(\omega_1, x + \beta(\omega_1)) - u(\omega_1, x + \beta(\omega_2))\big]\\ &\quad + (1 + \tau\beta_x(\omega_1))\big[u(\omega_1, x + \beta(\omega_2)) - u(\omega_2, x + \beta(\omega_2))\big]\\ &\quad + \tau u(\omega_1, x + \beta(\omega_2))(\beta_x(\omega_1) - \beta_x(\omega_2)).\end{aligned} \tag{A.1.7}$$

Using the (A.0.7b) and interpolation arguments we get

$$\begin{aligned}\|u(\omega_1, x + \beta(\omega_1)) - u(\omega_1, x + \beta(\omega_2))\|_{s-1} &\leq_s \|\beta(\omega_1) - \beta(\omega_2)\|_{s_0}\|u\|_s\\ &\quad + \|\beta(\omega_1) - \beta(\omega_2)\|_{s+1}\|u\|_{s_0}\\ &\leq_s \Big(\|\beta\|_{s+s_0+1}^{\gamma,\mathcal{O}}\|u\|_{s_0}^{\gamma,\mathcal{O}} + \|\beta\|_{s_0}^{\gamma,\mathcal{O}}\|u\|_s^{\gamma,\mathcal{O}}\Big)|\omega_1 - \omega_2|.\end{aligned}$$

The term we have estimated above is the most critical one among the summand in (A.1.7). The other estimates follow by the fact that $u(\omega, \varphi, x)$ and $\beta(\omega, \varphi, x)$ are Lipschitz functions of $\omega \in \mathcal{O}$. One can reason in the same way to get the estimates on the inverse map $(\mathcal{A}^\tau)^{-1}$ (see (5.7.59)). □

Lemma A.1.2. *Let $\mathtt{b} \in \mathbb{N}$. For any $|\alpha| \leq \mathtt{b}$, $m_1, m_2 \in \mathbb{R}$ such that $m_1 + m_2 = |\alpha|$, for any $s \geq s_0$ there exists a constant $\mu = \mu(|\alpha|, m_1, m_2)$ and $\delta = \delta(m_1, s)$ such that if*

$$\|\beta\|_{2s_0+|m_1|+2} \leq \delta, \quad \|\beta\|_{s_0+\mu}^{\gamma,\mathcal{O}} \leq 1, \tag{A.1.8}$$

then one has

$$\sup_{\tau\in[0,1]} \|\langle D_x\rangle^{-m_1}\partial_\varphi^\alpha \mathcal{A}^\tau(\varphi)\langle D_x\rangle^{-m_2} u\|_s^{\gamma,\mathcal{O}} \leq_{s,\mathtt{b},m_1,m_2} \|u\|_s + \|\beta\|_{s+\mu}^{\gamma,\mathcal{O}}\|u\|_{s_0}. \tag{A.1.9}$$

The inverse map $(\mathcal{A}^\tau)^{-1}$ satisfies the same estimate.

Proof. We prove the bound (A.1.9) for the $\|\cdot\|_s$ norm since one can obtain the bound in the Lipschitz norm $\|\cdot\|_s^{\gamma,\mathcal{O}}$ using the same arguments (recall also the reasoning used in (A.1.7)). We take $h \in C^\infty$, so that $\partial_\varphi^\alpha \mathcal{A}^\tau(\varphi)h \in C^\infty$ for any $|\alpha| \leq \mathtt{b}$ and we prove the bound (A.1.9) in this case. The thesis will follows by density.
We argue by induction on α. Given $\alpha \in \mathbb{N}^\nu$ we write $\alpha' \prec \alpha$ if $\alpha'_n \leq \alpha_n$ for any $n = 1, \dots, \nu$ and $\alpha' \neq \alpha$.
Let us check (A.1.9) for $\alpha = 0$. Let us define $\Psi^\tau := \langle D_x \rangle^m \mathcal{A}^\tau(\varphi) \langle D_x \rangle^{-m}$ with $m = -m_1 = m_2$. One has that $\Psi^0 := \mathrm{I}$ (where I is the identity operator). One can check that Ψ^τ solves the problem (recall (A.1.1))

$$\partial_\tau \Psi^\tau = \mathtt{X}\Psi^\tau + G^\tau \Psi^\tau, \tag{A.1.10}$$

where $G^\tau := [\langle D_x \rangle^m, \mathtt{X}]\langle D_x \rangle^{-m}$. Therefore by Duhamel principle one has

$$\Psi^\tau = \mathcal{A}^\tau + \mathcal{A}^\tau \int_0^\tau (\mathcal{A}^\sigma)^{-1} G^\sigma \Psi^\sigma d\sigma.$$

By Lemma 2.2.5 and (2.2.23) one has that

$$|G^\tau|_{0,s,0} \leq_s \|\beta\|_{s+m+3}, \quad s \geq s_0,$$

hence by estimate (A.1.2), Lemma 2.3.8 we have

$$\begin{aligned} \sup_{\tau \in [0,1]} \|\Psi^\tau h\|_s \leq_s \|h\|_s + \|\beta\|_{s+s_0+1}\|h\|_{s_0} + \|\beta\|_{s_0+m+3} \sup_{\tau \in [0,1]} \|\Psi^\tau h\|_s \\ + (\|\beta\|_{s+m+3} + \|\beta\|_{s+s_0+1}) \sup_{\tau \in [0,1]} \|\Psi^\tau h\|_{s_0}. \end{aligned} \tag{A.1.11}$$

For δ in (A.1.8) small enough, then the (A.1.11) for $s = s_0$ implies that $\sup_{\tau \in [0,1]} \|\Psi^\tau h\|_{s_0} \leq_{s_0} \|h\|_{s_0}$. Using this bound in (A.1.11) one gets the (A.1.9).
Now assume that the bound (A.1.9) holds for any $\alpha' \prec \alpha$ with $|\alpha| \leq \mathtt{b}$ and $m_1, m_2 \in \mathbb{R}$ with $m_1 + m_2 = |\alpha'|$. We now prove the estimate (A.1.9) for the operator $\langle D_x \rangle^{-m_1} \partial_\varphi^\alpha \mathcal{A}^\tau(\varphi) \langle D_x \rangle^{-m_2}$ for $m_1 + m_2 = |\alpha|$. Differentiating the (A.1.1) and using the Duhamel formula we get that

$$\begin{aligned} \partial_\varphi^\alpha \mathcal{A}^\tau(\varphi) &= \mathcal{A}^\tau(\varphi) \int_0^\tau (\mathcal{A}^\sigma(\varphi))^{-1} F_\alpha^\sigma d\sigma, \\ F_\alpha^\sigma &:= \sum_{\alpha_1 + \alpha_2 = \alpha} C(\alpha_1, \alpha_2) \partial_x (\partial_\varphi^{\alpha_1} b) \partial_\varphi^{\alpha_2} \mathcal{A}^\sigma(\varphi). \end{aligned} \tag{A.1.12}$$

For any $m_1 + m_2 = |\alpha|$ and any $\tau, s \in [0,1]$ we write

$$\begin{aligned} &\langle D_x \rangle^{-m_1} \partial_x (\partial_\varphi^{\alpha_1} b) \partial_\varphi^{\alpha_2} \mathcal{A}^\sigma(\varphi) \langle D_x \rangle^{-m_2} \\ &= \langle D_x \rangle^{-m_1} \partial_x (\partial_\varphi^{\alpha_1} b) \langle D_x \rangle^{-m_2+|\alpha_2|} \langle D_x \rangle^{m_2-|\alpha_2|} \partial_\varphi^{\alpha_2} \mathcal{A}^\sigma(\varphi) \langle D_x \rangle^{-m_2}. \end{aligned} \tag{A.1.13}$$

Hence in order to estimate the operator $\langle D_x \rangle^{-m_1} (\mathcal{A}^\sigma(\varphi))^{-1} F_\alpha^\sigma \langle D_x \rangle^{-m_2}$ we need to estimate, uniformly in $\tau, s \in [0,1]$ the term

$$\Big(\langle D_x \rangle^{-m_1} \mathcal{A}^\tau (\mathcal{A}^\sigma)^{-1} \langle D_x \rangle^{m_1}\Big)\Big(\langle D_x \rangle^{-m_1} \partial_x (\partial_\varphi^{\alpha_1} b) \langle D_x \rangle^{-m_2+|\alpha_2|}\Big)\Big(\langle D_x \rangle^{m_2-|\alpha_2|} \partial_\varphi^{\alpha_2} \mathcal{A}^\sigma(\varphi) \langle D_x \rangle^{-m_2}\Big). \tag{A.1.14}$$

For $s \geq s_0$, by the inductive hypothesis one has

$$\|\langle D_x\rangle^{-m_1}\mathcal{A}^\tau(\mathcal{A}^\sigma)^{-1}\langle D_x\rangle^{m_1}h\|_s \leq_{s,m_1} \|h\|_s + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mu}\|h\|_{s_0}, \tag{A.1.15}$$

$$\|\langle D_x\rangle^{m_2-|\alpha_2|}\partial_\varphi^{\alpha_2}\mathcal{A}^\sigma(\varphi)\langle D_x\rangle^{-m_2}h\|_s \leq_{s,\mathtt{b},m_2} \|h\|_s + \|\beta\|^{\gamma,\mathcal{O}}_{s+\mu}\|h\|_{s_0}. \tag{A.1.16}$$

provided that $\alpha_1 \neq 0$. We estimate the second factor in (A.1.14). We first note that

$$-m_1 - m_2 + 1 + |\alpha_2| = 1 + |\alpha_2| - |\alpha| \leq 0.$$

This implies that $\langle D_x\rangle^{-m_1}\partial_x(\partial_\varphi^{\alpha_1}b)\langle D_x\rangle^{-m_2+|\alpha_2|}$ belongs to OPS^0, and in particular, using Lemma 2.2.5 and (2.2.12), we obtain

$$|\langle D_x\rangle^{-m_1}\partial_x(\partial_\varphi^{\alpha_1}b)\langle D_x\rangle^{-m_2+|\alpha_2|}|_{0,s,0} \leq_{\mathtt{b},m_1,m_2} \|a\|^{\gamma,\mathcal{O}}_{s+|m_1|+|\alpha_2|}. \tag{A.1.17}$$

To obtain the bound (A.1.9) it is enough to use bounds (A.1.15), (A.1.16),(A.1.17), Lemma 2.3.8 and recall the smallness assumption (A.1.8). □

Lemma A.1.3. *Let $\mathtt{b} \in \mathbb{N}$. For any $|\alpha| \leq \mathtt{b}$, $m_1, m_2 \in \mathbb{R}$ such that $m_1 + m_2 = |\alpha| + 1$, for any $s \geq s_0$ there exists a constant $\mu = \mu(|\alpha|, m_1, m_2)$ and $\delta = \delta(s, m_1) > 0$ such that if*

$$\|\beta\|_{s_0+\mu} \leq \delta, \tag{A.1.18}$$

then one has

$$\begin{aligned}&\sup_{\tau\in[0,1]}\|\langle D_x\rangle^{-m_1}\partial_\varphi^\alpha\Delta_{12}\mathcal{A}^\tau(\varphi)\langle D_x\rangle^{-m_2}u\|_s \\ &\qquad\leq_{s,\mathtt{b},m_1,m_2} \|u\|_s\|\Delta_{12}\beta\|_{s_0+\mu} + \|u\|_{s_0+1}(\|\Delta_{12}\beta\|_{s+\mu} + \|\Delta_{12}\beta\|_{s_0+\mu}\|\beta\|_{s+\mu}).\end{aligned} \tag{A.1.19}$$

The operators $\Delta_{12}(\mathcal{A})^*$, $\Delta_{12}(\mathcal{A})^{-1}$ satisfy the same estimate.

Proof. The Lemma can be proved arguing as in the proof of Lemma A.1.2. □

APPENDIX B

PROOFS

Here we collect the proofs of some lemmata.

Proof of Lemma 5.7.18. Let us define $\Upsilon^\tau := (\Psi^\tau)^{-1} \circ \Phi^\tau$. Then by substituting in (5.7.143) we have

$$\begin{aligned}(\partial_\tau \Psi^\tau)(\Upsilon^\tau u) + \Psi^\tau(\partial_\tau \Upsilon^\tau u) &= (J \circ b)\Pi_S^\perp[\Psi^\tau(\Upsilon^\tau u)] - \Pi_S[(J \circ b)\Pi_S^\perp[\Psi^\tau(\Upsilon^\tau u)] \\ &= (J \circ b)\Psi^\tau(\Upsilon^\tau u) - \Big((J \circ b)\Pi_S[\Psi^\tau(\Upsilon^\tau u)] + \Pi_S[(J \circ b)\Pi_S^\perp[\Psi^\tau(\Upsilon^\tau u)]\Big).\end{aligned} \tag{B.0.1}$$

Thus by (5.7.89) we have

$$\begin{cases}\partial_\tau \Upsilon^\tau u = -\big(\Psi_*^\tau \mathtt{Z}\big)\Upsilon^\tau u, \\ \Upsilon^0 u = u\end{cases} \tag{B.0.2}$$

with

$$\mathtt{Z}u := (J \circ b)\Pi_S[u] + \Pi_S[(J \circ b)\Pi_S^\perp[u] = \sum_{j \in S} \big(g_j(\tau), u\big)_{L^2(\mathbb{T}_x)} \chi_j(\tau) + \sum_{j \in S} \big(\tilde{g}_j(\tau), u\big)_{L^2(\mathbb{T}_x)} \tilde{\chi}_j(\tau),$$

where

$$g_j = \tilde{\chi}_j := e^{\mathrm{i}jx}, \quad \chi_j = J(b(\tau)\, e^{\mathrm{i}jx}), \quad \tilde{g}_j := \omega(j)\Pi_S^\perp[b(\tau)\, e^{\mathrm{i}jx}].$$

Equation (B.0.2) is well posed on H^s.

We show that $\Upsilon^\tau - \mathrm{I}$ is of the form (5.6.5). By Taylor expansion at $\tau = 0$ we get

$$\Upsilon^\tau u - u = -\tau\big(\Psi_*^\tau \mathtt{Z}(\Upsilon^\tau u)\big)_{|\tau=0} + \int_0^\tau (1-t)\big(\partial_{tt}\Upsilon^t(u)\big)\, dt. \tag{B.0.3}$$

Note that

$$\Psi_*^\tau \mathtt{Z}(\Upsilon^\tau u) = \sum_{j \in S} \big((\Phi^\tau)^* g_j(\tau), u\big)_{L^2(\mathbb{T}_x)} (\Psi^\tau)^{-1}\chi_j(\tau) + \sum_{j \in S} \big((\Phi^\tau)^* \tilde{g}_j(\tau), u\big)_{L^2(\mathbb{T}_x)} (\Psi^\tau)^{-1}\tilde{\chi}_j(\tau)$$

has already the form (5.6.5) and

$$\Big(\Psi_*^\tau \mathtt{Z}(\Upsilon^\tau u)\Big)_{|\tau=0} = \mathtt{Z}u.$$

We denoted by $(\Psi^\tau)^*$ the flow of the adjoint PDE

$$\partial_\tau(\Psi^\tau)^* u = -bJ((\Psi^\tau)^* u), \tag{B.0.4}$$

by $(\Phi^\tau)^*$ the flow of

$$\partial_\tau (\Phi^\tau)^* u = -\Pi_S^\perp \, b \, J \, \Pi_S^\perp [(\Phi^\tau)^* u]. \tag{B.0.5}$$

These maps satisfy estimates like (5.7.139),(5.7.140). We have

$$\int_0^\tau (1-t)\partial_{tt}\Upsilon^t(u)\,ds = \sum_{j\in S}\int_0^\tau (1-t)\big(\mathtt{g}_j(t),u\big)_{L^2(\mathbb{T}_x)}\,\mathtt{f}_j(t)\,dt + \sum_{j\in S}\int_0^\tau (1-t)\big(\tilde{\mathtt{g}}_j(t),u\big)_{L^2(\mathbb{T}_x)}\,\tilde{\mathtt{f}}_j(t)\,dt$$

where

$$\mathtt{g}_j := \Big(-(\Psi_*^s \mathtt{Z})^*(\Phi^s)^* - (\Upsilon^s)^* b\, J((\Psi^s)^*)\Big)(g_j), \quad \mathtt{f}_j := -(J\circ b)((\Psi^s)^{-1})\chi_j + (\Psi^s)^{-1}J(b'(s)e^{\mathrm{i}jx}),$$

$$\tilde{\mathtt{g}}_j := \Big(-(\Psi_*^s \mathtt{Z})^*(\Phi^s)^* - (\Upsilon^s)^* b\, J((\Psi^s)^*)\Big)(\tilde g_j) + (\Phi^s)^*\omega(j)\Pi_S^\perp[b'(s)e^{\mathrm{i}jx}], \quad \tilde{\mathtt{f}}_j := -(J\circ b)((\Psi^s)^{-1})\tilde\chi_j.$$

Thus by (B.0.3), for $\tau = 1$ we get

$$\Upsilon^1 u - u = \mathcal{R}u, \qquad \mathcal{R}u := \mathcal{R}_1 u + \mathcal{R}_2 u \tag{B.0.6}$$

$$\mathcal{R}_1 u := -\mathtt{Z}u, \qquad \mathcal{R}_2 u := \int_0^1 (1-t)\partial_{tt}\Upsilon^t(u)\,dt \tag{B.0.7}$$

where $\mathcal{R}_1$ has the finite dimensional form (5.6.19) and $\mathcal{R}_2$ has the form (5.6.5). Hence by Lemma 5.7.8 we have

$$\mathbb{M}^\gamma_{\mathcal{R}_1}(s,\mathtt{b}) \le_s \sum_{|j|\le C}\sup_{\tau\in[0,1]} (\|\mathtt{f}_j(\tau)\|_s^{\gamma,\mathcal{O}_0}\|\mathtt{g}_j(\tau)\|_{s_0}^{\gamma,\mathcal{O}_0} + \|\mathtt{f}_j(\tau)\|_{s_0}^{\gamma,\mathcal{O}_0}\|\mathtt{g}_j(\tau)\|_s^{\gamma,\mathcal{O}_0}). \tag{B.0.8}$$

$$\mathbb{M}^\gamma_{\mathcal{R}_2}(s,\mathtt{b}) \le_s \sum_{|j|\le C}\sup_{\tau\in[0,1]} (\|\tilde{\mathtt{f}}_j(\tau)\|_s^{\gamma,\mathcal{O}_0}\|\tilde{\mathtt{g}}_j(\tau)\|_{s_0}^{\gamma,\mathcal{O}_0} + \|\tilde{\mathtt{f}}_j(\tau)\|_{s_0}^{\gamma,\mathcal{O}_0}\|\tilde{\mathtt{g}}_j(\tau)\|_s^{\gamma,\mathcal{O}_0}). \tag{B.0.9}$$

By using the estimates in Proposition 5.7.16 we have

$$\begin{aligned} &\|\mathtt{f}_j\|_s^{\gamma,\mathcal{O}_0}, \|\tilde{\mathtt{g}}_j\|_s^{\gamma,\mathcal{O}_0} \le_s \|b\|_{s+s_0+2}^{\gamma,\mathcal{O}_0} + \|\partial_\tau b(\tau)\|_{s+1}^{\gamma,\mathcal{O}_0}, \\ &\|\tilde{\mathtt{f}}_j\|_s^{\gamma,\mathcal{O}_0}, \|\mathtt{g}_j\|_s^{\gamma,\mathcal{O}_0} \le_s \|b\|_{s+s_0+2}^{\gamma,\mathcal{O}_0}. \end{aligned} \tag{B.0.10}$$

In the same way, the bounds for the variation on the i-variable (5.7.145) follows by the estimates on the derivatives of the coefficients $\mathtt{g}_j, \tilde{\mathtt{g}}_j, \chi_j, \tilde\chi_j$ whose depend on the variation Δ_{12} of the flows Φ^τ, Ψ^τ and their adjoints, for instance recall (5.7.142). We have proved that $\Upsilon^1 u = u + \mathcal{R}u$ and hence $\Phi^1 u = \Psi^1\circ(\mathrm{I}+\mathcal{R})\mathrm{u}$. By (B.0.8), (B.0.9) and (B.0.10) we obtain (5.7.144). □

Proof of Lemma 5.7.24. $(\mathcal{P}_{1,2})_0$ are trivial.
Now suppose that $(\mathcal{P}_{1,2})_n$ hold and we prove that $(\mathcal{P}_{1,2})_{n+1}$ also hold.
We have to prove that the $(n+1)$-th diffeomorphism of the torus is well defined from H^s to itself for all $s \ge s_0$.
We show that (5.7.230) holds with $K \rightsquigarrow K_n$ and $a \rightsquigarrow a_n$.
We first recall that $\|a\|_s^{\gamma,\mathcal{O}} \le \gamma^{-1}\|a\|_s^{\gamma,\mathcal{O}}$. We set $\lambda = 1/(s_1 - s_0)$. By classical interpolation arguments (see Appendix A) we have that

$$\gamma^{-1}\|a_n\|_{s_0+1}^{\gamma,\mathcal{O}_n} \le (\gamma^{-1}\|a_n\|_{s_0}^{\gamma,\mathcal{O}_n})^{1-\lambda}(\gamma^{-1}\|a_n\|_{s_1}^{\gamma,\mathcal{O}_n})^\lambda \overset{(5.7.256)}{\le} C(s_1)\,\delta_0(s_1)K_0^{\mu(1-\lambda)}K_n^{-\mu(1-\lambda)}. \tag{B.0.11}$$

Therefore we have

$$C(s_1)\,\delta_n(s_0+1)K_n^{2\tau+4} \le C(s_1)\,\delta_0(s_1)\,K_n^{2\tau+4-(1-\lambda)\mu}K_0^{(1-\lambda)\mu}.$$

Since $\lambda \ll 1/2$ then $\mu(1-\lambda) > 2\tau+4$, then $K_n^{2\tau+4-(1-\lambda)\mu}$ is a decreasing sequence and by (5.7.250) (since $\varrho \ge 3$)

$$C(s_1)\,\delta_0(s_1)\,K_n^{2\tau+4-(1-\lambda)\mu}K_0^{(1-\lambda)\mu} \le C(s_1)\,\delta_0(s_1)K_0^{2\tau+4} < 1 \quad \Rightarrow \quad C(s_1)\,\delta_n(s_0+1)K_n^{2\tau+4} \le 1. \tag{B.0.12}$$

Now we apply the KAM step and have $a_+ \rightsquigarrow a_{n+1}$.

$$\begin{aligned}\delta_{n+1}(s_0) &\le \gamma^{-1}C_{s_0}(K_n^{s_0-s_1}\|a_n\|_{s_1}^{\gamma,\mathcal{O}_n} + \gamma^{-1}K_n^{2\tau+2}(\|a_n\|_{s_0}^{\gamma,\mathcal{O}_n})^2)(1+\gamma^{-1}K_n^{2\tau+2}\|a_n\|_{s_0}^{\gamma,\mathcal{O}_n})\\ &\le C_{s_0}(K_n^{s_0-s_1}\delta_n(s_1) + K_n^{2\tau+2}(\delta_n(s_0)^2)(1+K_n^{2\tau+2}\delta_n(s_0)\ ).\end{aligned} \tag{B.0.13}$$

We first note that $K_n^{2\tau+2-\mu}\delta_0(s_1)K_0^{\mu} < 1$. Indeed since $\mu > 2\tau+2$, this is a decreasing sequence and by (5.7.250) and the definition of ϱ (see (5.7.249))

$$C_{s_0}K_n^{2\tau+2-\mu}\delta_0(s_1)K_0^{\mu} \le C_{s_0}K_0^{2\tau+2}\delta_0(s_1) < 1.$$

Hence

$$\delta_{n+1}(s_0) \le C_{s_0}\big(K_n^{s_0-s_1}\delta_n(s_1) + K_n^{2\tau+2}\delta_n(s_0)^2\big)$$

and then we have to prove that

$$\begin{cases} C_{s_0}\,\delta_n(s_1)K_n^{-(s_1-s_0)} < \frac{1}{2}\delta_0(s_1)\,K_0^{\mu}\,K_{n+1}^{-\mu}\\ C_{s_0}\,\delta_n(s_0)^2K_n^{2\tau+2} < \frac{1}{2}\delta_0(s_1)\,K_0^{\mu}\,K_{n+1}^{-\mu}.\end{cases} \tag{B.0.14}$$

Thus, by the inductive hypotesis, we have to prove

$$2C_{s_0}K_n^{-(s_1-s_0)+1+\chi\mu}K_0^{-\mu} < 1\,, \quad 2C_{s_0}\delta_0(s_1)K_0^{\mu}K_n^{2\tau+2-(2-\chi)\mu} < 1. \tag{B.0.15}$$

Since

$$s_1 - s_0 > 1+\chi\mu\,, \quad \mu > \frac{2\tau+2}{2-\chi}$$

then the sequences in (B.0.15) are decreasing and we just need

$$2C_{s_0}K_0^{-(s_1-s_0)+1+\mu(\chi-1)} < 1\,, \quad 2C_{s_0}\delta_0(s_1)K_0^{2\tau+2+\mu(\chi-1)} < 1$$

which follows by taking K_0 sufficiently large ($K_0 > 2C_{s_0}$) and by (5.7.250), since

$$\varrho \ge 3+\mu(\chi-1).$$

Regarding the estimates in high norm, by (5.7.245) we have for all $s \ge s_0$

$$\|a_{n+1}\|_s^{\gamma,\mathcal{O}_{n+1}} \le \|a_n\|_s^{\gamma,\mathcal{O}_n} + C(s)\gamma^{-1}K_n^{2\tau+2}\|a_n\|_{s_0}^{\gamma,\mathcal{O}_n}\|a_n\|_s^{\gamma,\mathcal{O}_n}.$$

Note that there exists $n_0(s) \in \mathbb{N}$ sufficiently large such that for any $n \ge n_0(s)$

$$C(s)K_n^{2\tau+2-\mu}K_0^{\mu}\delta_0(s_1) \le 2^{-n-3}. \tag{B.0.16}$$

Then we have by (B.0.16)

$$\delta_{n+1}(s) \le \delta_n(s)(1 + C_s K_n^{2\tau+2-\mu} K_0^{\mu} \delta_0(s_1)) \le C_1(s)\delta_0(s)(1 + \sum_{j=1}^{n} 2^{-j})(1 + 2^{-n-3})$$
$$\le C_1(s)\delta_0(s)(1 + \sum_{j=1}^{n+1} 2^{-j})$$

for $s \ge s_0$. For $s \ge s_0$, setting $\lambda = 1/(s - s_0 + 1) \le 1$, we have

$$\|a_{n+1}\|_s^{\gamma,\mathcal{O}_{n+1}} \le (\|a_{n+1}\|_{s_0}^{\gamma,\mathcal{O}_{n+1}})^{\lambda} (\|a_{n+1}\|_{s+1}^{\gamma,\mathcal{O}_{n+1}})^{1-\lambda}, \tag{B.0.17}$$

from which we may deduce that

$$\delta_{n+1}(s) \le K_{n+1}^{-\lambda\mu}(K_0^{\mu}\delta_0(s_1))^{\lambda}(\delta_{n+1}(s+1))^{1-\lambda} \le 2\, C_1(s+1)\, K_{n+1}^{-\lambda\mu} K_0^{\lambda\mu} \max(\delta_0(s_1), \delta_0(s+1)) \tag{B.0.18}$$

now by (5.7.258) and (B.0.18)

$$\|\alpha_{n+1}\|_s^{\gamma,\mathcal{O}_0} \le \delta_{n+1}(s + 2\tau + 1) \le C_2(s)\, K_{n+1}^{-\lambda\mu} K_0^{\lambda\mu} \max(\delta_0(s_1), \delta_0(s + 2\tau + 2)), \quad s \ge s_0. \tag{B.0.19}$$

It remains to prove (5.7.257) for $n + 1$.

By the inductive hypothesis, (5.7.239) and (5.7.255) we have

$$\varepsilon_{n+1} \le \varepsilon_0[\delta_n(1 + 2\delta_n)] \overset{(B.0.11)}{\le} 2^{-(n+1)}\varepsilon_0. \tag{B.0.20}$$

The (B.0.20) implies the last bound in (5.7.255) in $(\mathcal{P}_1)_{n+1}$.

Now we prove $(\mathcal{P}_2)_{n+1}$. By construction

$$\beta_{n+1}(x) = \alpha_{n+1}(x) + \beta_n(x + \alpha_{n+1}(x))$$

thus, by (A.0.8)

$$\|\beta_{n+1}\|_s^{\gamma,\mathcal{O}_0} \le \|\alpha_{n+1}\|_s^{\gamma,\mathcal{O}_0} + \|\beta_n\|_s^{\gamma,\mathcal{O}_0}(1 + C_s\|\alpha_{n+1}\|_{s_0+1}^{\gamma,\mathcal{O}_0}) + C_s\|\beta_n\|_{s_0}^{\gamma,\mathcal{O}_0}\|\alpha_{n+1}\|_{s+1}^{\gamma,\mathcal{O}_0}.$$

By the inductive hypotesis (see (5.7.260))

$$\|\beta_n\|_s^{\gamma,\mathcal{O}_0} \le C_3(s)\max(\delta_0(s_1), \delta_0(s + 2\tau + 3)) \sum_{j=0}^{n} 2^{-j}, \qquad s \ge s_0 \tag{B.0.21}$$

and

$$\|\beta_n\|_{s_0}^{\gamma,\mathcal{O}_0} \le C(s_1)\delta_0(s_1) \sum_{j=0}^{n} 2^{-j}. \tag{B.0.22}$$

By (B.0.19), (B.0.21) and (B.0.22) we have

$$\|\beta_{n+1}\|_s^{\gamma,\mathcal{O}_0} \le C_2(s)\, K_{n+1}^{-\lambda\mu} K_0^{\lambda\mu}\, \mathtt{M}(s+1)\, (1 + C_s\|\beta_n\|_{s_0}^{\gamma,\mathcal{O}_0}) +$$
$$C_3(s)\, \mathtt{M}(s+1) \sum_{j=0}^{n} 2^{-j}(1 + C_s K_{n+1}^{2\tau+2-\mu} K_0^{\mu}\delta_0(s_1)) \le$$
$$C_3(s)\, \max(\delta_0(s_1), \delta_0(s + 2\tau + 3)) \sum_{j=0}^{n+1} 2^{-j}$$

provided that $K_n^{-\lambda\mu} K_0^{\lambda\mu} \le C_s\, 2^{-(n+1)}$ which holds for $n \ge n_0(s)$ for some $n_0(s)$ sufficiently large. By (B.0.19), (B.0.21) and (B.0.22) we have

$$\begin{aligned} &\|\beta_{n+1} - \beta_n\|_s^{\gamma,\mathcal{O}_0} \le \|\alpha_{n+1}\|_s^{\gamma,\mathcal{O}_0} + C(s)(\|\beta_n\|_{s+1}^{\gamma,\mathcal{O}_0}\|\alpha_{n+1}\|_{s_0}^{\gamma,\mathcal{O}_0} + \|\beta_n\|_{s_0+1}^{\gamma,\mathcal{O}_0}\|\alpha_{n+1}\|_s^{\gamma,\mathcal{O}_0}) \le \\ &\le C_4(s)\, \mathtt{M}(s+2)\, 2^{-n} \end{aligned}$$

for $C_4(s) \ge C_2(s) + C_s\, C_2(s)\, \delta_0(s_1)\, C_3(s+1) + 2C_s^2\, C_2(s)$. □

B.1 Generic conditions

In this Section we prove the genericity of the conditions assumed in the Chapter 5.

Lemma B.1.1. *There exists a constant* $\mathtt{c}_1 > 0$ *independent of the set* S^+ *(recall* (1.2.8)*) such that the assumption* (1.2.15) *is generic in* $\mathcal{V}(\mathtt{c}_1)$.

Proof. Recalling (5.3.9) and (5.3.19) we define a matrix $\mathbb{K}$ so that $\mathbb{A} = (2/9)\,\mathrm{diag}\Big(\omega(\bar{\jmath}_i)(1+\bar{\jmath}_i^2)\Big)\mathbb{K}$. In this way the entries of $\mathbb{K}$ are bounded by some constant independent of the $\bar{\jmath}_i$. Let us assume that $\bar{\jmath}_1$ is the largest of the $\bar{\jmath}_i$ (which we recall are positive) and write

$$\bar{\jmath}_1 \rightsquigarrow 1/x, \quad \bar{\jmath}_i \rightsquigarrow p_i/x\,, \quad 0 < p_i \le 1\,, \tag{B.1.1}$$

so that $P(x,p) = \det(\mathbb{K})$ is a rational function of its variables. It is easily seen that $\mathbb{K}$ computed at $p_i = 1$ for all i, coincides with the matrix (recall that $U_{ij} = 1\ \forall i, j = 1, \dots, \nu$)

$$M(x) := \mathrm{I} + 2\, g(x)\,(U - \mathrm{I}), \qquad g(x) := \frac{1}{3x^2+1}$$

so that its determinant is

$$\det M = \frac{(2x^2 + (2\nu - 1))}{(3x^2+1)}$$

we note that this function is > 1 at $x = 0$. We conclude that there exists $x_0 < 1$ and a constant $\mathtt{c}$ so that for all

$$x < x_0, |p_i - 1| \le \mathtt{c} \quad \text{one has} \quad \det \mathbb{K} \ge 1/2.$$

This implies that, for suitable choice of $\mathtt{c}_1$, the bound (1.2.15) holds for any choice of tangential sites in the set given (1.2.12). In (1.2.11) the determinant of $\mathbb{A}$ is an analytic, non identically zero, function of $\bar{\jmath}_i$. Hence for a generic choice of the tangential site it is different from zero. In particular in the ball of radius $\mathtt{c}_1^{-1}$ (recall that the variables $\bar{\jmath}_k$ are integers) the minimum of $\det \mathbb{A}$ depends only on the constant $\mathtt{c}_1$. This implies (1.2.15) for a suitable constant $\mathtt{c}_2$. Let us check (1.2.15). We first note that

$$\det(\mathrm{I} - \mathbb{A}^{-1}\vec{v}\,\overline{\omega}^T) = 1 - \mathbb{A}^{-1}\vec{v}\cdot\overline{\omega}.$$

The function $\vec{v}\cdot\mathbb{A}^{-1}\overline{\omega}$ is an algebraic function of the variables $\bar{\jmath}_i$. We must show that it is not identically zero. Consider the change of variables (B.1.1) and set $\lambda = 1/x$. By an explicit computation one can note that the matrix $\mathbb{A}$ in (5.3.9) at $(\bar{\jmath}_i)_{i=1}^{\nu} = \lambda\vec{1}$ (recall that $\vec{1} := (1, \dots, 1)$) is given by

$$\begin{aligned} &\mathbb{A} := d(\lambda)\big[\mathrm{I} + e(\lambda)U\big], \\ &\mathbb{A}^{-1} := \frac{1}{d(\lambda)}\big[\mathrm{I} - f(\lambda)U\big], \\ &d(\lambda) := \frac{2(4+\lambda^2)(3+2\lambda^2+\lambda^4)}{9\lambda(3+\lambda^2)}, \qquad e(\lambda) := \frac{2\lambda^2}{3+2\lambda^2+\lambda^4}, \quad f(\lambda) := \frac{e(\lambda)}{1+\nu e(\lambda)}. \end{aligned} \tag{B.1.2}$$

By (B.1.2)

$$\mathbb{A}^{-1}\vec{v}\cdot\varpi = \frac{2}{3}\lambda(4+\lambda^2)(\vec{1}\cdot\mathbb{A}^{-1}\vec{1}) = \frac{2}{3}\lambda(4+\lambda^2)\frac{9\nu\lambda(3+\lambda^2)}{2(4+\lambda^2)(3+2(1+\nu)\lambda^2+\lambda^4)}. \tag{B.1.3}$$

We note that, for $x=0$, one has

$$|\det(\mathrm{I}-\mathbb{A}^{-1}\vec{v}\varpi^T)| \geq 1,$$

hence there exists $x_0<1$ and a constant $\mathtt{c}$ so that for all

$$x<x_0, |p_i-1|\leq \mathtt{c} \qquad \text{one has} \qquad \det(\mathrm{I}-\mathbb{A}^{-1}\vec{v}\varpi^T)\geq 1/2.$$

One concludes as done for (1.2.15).

□

Lemma B.1.2. *There exists a constant* $\mathtt{c}_2>0$ *independent of the set* S *such that the assumption* (1.2.16) *is generic in* $\mathcal{V}(\mathtt{c}_2)$*.*

Proof. Similarly to the previous Lemma, we define $\mathbb{H}$ so that (see (5.3.9)) $\mathbb{A}=\Omega\,\mathbb{H}\,\mathrm{diag}(\vec{v}_i)$ and $w_j=\mathrm{diag}(\vec{v}_i)b_j$ with $|b_j|\leq C$ with C independent of the $\jmath_i$ and j. Similarly to the previous Lemma, the entries of $\mathbb{H}$ are uniformly bounded. Then

$$\mathbb{A}^{-T}(w_j-w_k)=\Omega^{-1}\mathbb{H}^{-T}(b_j-b_k). \tag{B.1.4}$$

Following B.1.1 we can conclude that there exists some constant $C>0$, independent of $\jmath_i$, such that

$$|\mathbb{H}^{-T}(b_j-b_k)|\leq C \tag{B.1.5}$$

In the same way we can conclude that

$$|(\mathrm{I}-\mathbb{A}^{-T}\vec{v}(\varpi)^T)y|\leq C|y| \qquad \forall y\in\mathbb{R}^\nu. \tag{B.1.6}$$

Moreover (recall (5.3.9) for the definition of Ω)

$$\mathrm{I}-\mathbb{A}^{-T}\vec{v}(\varpi)^T=\Omega^{-1}(\mathrm{I}-\mathbb{H}^{-T}U)\Omega, \tag{B.1.7}$$

then in the set (1.2.12) (with a possible smaller constant $\mathtt{c}_1$) the entries of this matrix are uniformly bounded by some constant independent of the $\jmath_i$. By Lemma B.1.1 the determinant is bounded from below, so that there exists R independent of S such that if $\max(\jmath_i)>R$ then

$$\left|\left(\mathrm{I}-\mathbb{A}^{-T}\vec{v}(\varpi)^T\right)^{-1}\Omega^{-1}\mathbb{H}^{-T}(b_j-b_k)\right|<1. \tag{B.1.8}$$

In the case $\max(\jmath_i)\leq R$ consider $\max\{|j|,|k|\}\geq M$ with $M>0$ independent of the set S. Then if M is sufficiently large then (B.1.8) holds.
In the case $\max(\jmath_i)\leq R$ and $\max\{|j|,|k|\}\leq M$ we reason as follows.
Note that

$$\left|\left(\mathrm{I}-\mathbb{A}^{-T}\vec{v}(\varpi)^T\right)^{-1}\Omega^{-1}\mathbb{H}^{-T}(b_j-b_k)\right|\leq K$$

where $K > 0$ independent of S, hence for $|\ell| > K$ then (1.2.16) holds. Otherwise we show that the following rational function of the tangential sites are non identically zero

$$P(\ell, j, k, \overline{\jmath}_i) = \Big(\mathrm{I} - \mathbb{A}^{-T}\vec{v}(\overline{\omega})^T\Big)\ell - \mathbb{A}^{-T}(w_j - w_k).$$

We remark that $\overline{\jmath_i} \mapsto P(\ell, j, k, \overline{\jmath_i})$ are a finite number of the functions since $|\ell| \le K$ and $|j|, |k| \le M$. By evaluating the w_j at $(\overline{\jmath}_k)_{k=1}^{\nu} = \lambda\vec{1}$ we get

$$w_j = w[j, \lambda] := -\frac{2}{3}\frac{j(4+j^2)}{(1+j^2)}\frac{(1+\lambda^2)(7+5\lambda^2+\lambda^4+3j^2)}{(3+\lambda^2-\lambda j+j^2)(3+\lambda^2+\lambda j+j^2)}.$$

Hence condition (1.2.16) reads

$$\ell \neq \frac{1}{d(\lambda)}\big[\mathrm{I} - r(\lambda)U\big](w[j,\lambda] - w[k,\lambda])\vec{1} = \frac{1-\nu r(\lambda)}{d(\lambda)}(w[j,\lambda] - w[k,\lambda])\vec{1}. \tag{B.1.9}$$

We show that there exists $\lambda \in \mathbb{R}$ such that the function

$$s[j,k,\lambda] := \frac{1-\nu r(\lambda)}{d(\lambda)}(w[j,\lambda] - w[k,\lambda]) \tag{B.1.10}$$

does not belong to $\mathbb{Z}$ for any $\nu \in \mathbb{N}$ and any $|j|, |k| \le M$. We have that

$$s[j,k,1] = \frac{12}{5(3+10\nu)}\left(\frac{k(4+k^2)(13+3k^2)}{16+23k^2+8k^4+k^6} - \frac{j(4+j^2)(13+3j^2)}{16+23j^2+8j^4+j^6}\right), \tag{B.1.11}$$

and we note that

$$|s[j,k,1]| \le \frac{12}{5(3+10\nu)} \cdot 4 < 1,$$

for any $\nu \in \mathbb{N}$. This implies that for any fixed j, k, ℓ the functions $P(\ell, j, k, \overline{\jmath_i})$ are not identically zero. Hence the thesis follows. $\square$

APPENDIX C

Integrable structure of the Degasperis-Procesi equation

We write the complete expression of the constants of motion (5.1.1) computed thanks to the procedure of Section 4 in [48].

$$K_1 := \frac{1}{2}\int_{\mathbb{T}} (J^{-1}u_x)\, u\, dx, \tag{C.0.1}$$

$$K_2 := \int_{\mathbb{T}} (-1+u-u_{xx})^{\frac{1}{3}}\, dx, \tag{C.0.2}$$

$$K_3 := -\frac{1}{2}\int_{\mathbb{T}} (-1+u-u_{xx})^{-\frac{7}{3}}\,(u_x-u_{xxx})^2 + 9(-1+u-u_{xx})^{\frac{1}{3}}\, dx, \tag{C.0.3}$$

$$\begin{aligned} K_4 := \int_{\mathbb{T}} m^{-\frac{2}{3}}\Big(&- \frac{\partial_x^6 m}{27m^2} + \frac{2m_{xxxx}}{9m^2} + \frac{17m_{xxx}^2}{27m^3} - \frac{1880m_{xx}^3}{729m^4} - \frac{31m_{xx}^2}{27m^3} \\ &- \frac{7m_{xx}}{27m^2} + \frac{379351m_x^6}{19683m^7} + \frac{3283m_x^4}{729m^5} \\ &+ \frac{46m_x^2}{81m^3} + \frac{11\partial_x^5 m m_x}{27m^3} + \frac{77m_{xxxx}m_{xx}}{81m^3} \\ &- \frac{682m_{xxxx}m_x^2}{243m^4} + \frac{3394m_{xxx}m_x^3}{243m^5} \\ &- \frac{5m_{xxx}m_x}{3m^3} - \frac{108386m_x^4 m_{xx}}{2187m^6} + \frac{22240m_x^2m_{xx}^2}{729m^5} \\ &+ \frac{1688m_x^2 m_{xx}}{243m^4} - \frac{290m_{xxx}m_x m_{xx}}{27m^4} - \frac{1}{27m}\Big)\, dx. \end{aligned} \tag{C.0.4}$$

In Section 5.2 we use the infinite dimensional version of the following lemma about the normal form of commuting Hamiltonians.

Lemma C.0.1. *If two finite dimensional Hamiltonians*

$$H := H^{(2)} + H^{(\geq 3)}, \qquad K := K^{(2)} + K^{(\geq 3)} \tag{C.0.5}$$

are such that $\{H, K\} = 0$, *then, for any* $n \in \mathbb{N}\setminus\{0\}$, *there exist a Birkhoff transformation* Φ_{n+2} *which puts* H *and* K *in normal form up to order* n

$$H\circ\Phi_{n+2} = H^{(2)} + Z_n + H^{(\geq n+3)}, \qquad K\circ\Phi_{n+2} = K^{(2)} + W_n + K^{(\geq n+3)} \tag{C.0.6}$$

where Z_n *and* W_n *are polynomial of maximal degree at most* $n+2$, *commuting both with* $H^{(2)}$ *and* $K^{(2)}$.

Proof. By assumption

$$\{H,K\} = \{H^{(2)},K^{(2)}\} + \{H^{(3)},K^{(2)}\} + \{H^{(2)},K^{(3)}\} + \text{ h.o.t} \equiv 0\,, \tag{C.0.7}$$

then this identity holds for each degree, namely

$$\{H^{(2)},K^{(2)}\} = 0\,, \quad \{H^{(3)},K^{(2)}\} + \{H^{(2)},K^{(3)}\} = 0\,, \dots \tag{C.0.8}$$

We decompose $\mathcal{P}^{(k)}$, the space of polynomials of finite degree k, as

$$\mathcal{P}^{(k)} = \text{ Ker}(H^{(2)}) \oplus \text{Rg}(H^{(2)})\,, \quad \mathcal{P}^{(k)} = \text{ Ker}(K^{(2)}) \oplus \text{Rg}(K^{(2)})$$

where $\text{Ker}(H^{(2)})$ is the kernel of the adjoint action $\{H^{(2)},\cdot\}$, same for $\text{Ker}(K^{(2)})$.
First we prove the following claim: given an integer $k \geq 2$, if

$$\{H^{(k)},K^{(2)}\} + \{H^{(2)},K^{(k)}\} = 0, \tag{C.0.9}$$

then

$$\Pi_{\text{Rg}(H^{(2)})}H^{(k)} = \Pi_{\text{Rg}(K^{(2)})}\Pi_{\text{Rg}(H^{(2)})}H^{(k)}, \quad \Pi_{\text{Rg}(K^{(2)})}K^{(k)} = \Pi_{\text{Rg}(H^{(2)})}\Pi_{\text{Rg}(K^{(2)})}K^{(k)}. \tag{C.0.10}$$

In particular, we prove that

$$\Pi_{\text{Rg}(K^{(2)})}\Pi_{\text{Ker}(H^{(2)})}H^{(k)} = 0\,, \quad \Pi_{\text{Rg}(H^{(2)})}\Pi_{\text{Ker}(K^{(2)})}K^{(k)} = 0. \tag{C.0.11}$$

The main point is that, given $f \in \text{ Ker}_{H^{(2)}}$, one has $\{K^{(2)},f\} \in \text{ Ker}_{H^{(2)}}$, namely the adjoint actions of $H^{(2)}$ and $K^{(2)}$ commute. Indeed $\{H^{(2)},K^{(2)}\} = 0$ and the result follows by the Jacobi identity. Thus $\{\Pi_{\text{Ker}(H^{(2)})}H^{(k)},K^{(2)}\} \in \text{Ker}(H^{(2)})$ and by (C.0.9) we have

$$\{\Pi_{\text{Ker}(H^{(2)})}H^{(k)},K^{(2)}\} = -\{\Pi_{\text{Rg}(H^{(2)})}H^{(k)},K^{(2)}\} - \{H^{(2)},K^{(k)}\} \in \text{Rg}(H^{(2)}).$$

Therefore $\{\Pi_{\text{Ker}(H^{(2)})}H^{(k)},K^{(2)}\} = 0$ and (C.0.11) follows. By exchanging $K^{(k)}, H^{(k)}$ and repeating the same arguments we deduce our claim.
We now proceed the proof by induction on the number n of the Birkhoff normal form steps.
The base of the induction is trivial, because $\Phi^{(0)}$ is the identity map. Thus suppose that we have performed $n \geq 0$ steps. By substituting H and K in (C.0.7) with $H \circ \Phi_n$ and $K \circ \Phi_n$, and looking at each degree separately, we get

$$\begin{aligned}
&\{H^{(2)},K^{(2)}\} = 0\,,\\
&\{Z_n,K^{(2)}\} + \{H^{(2)},W_n\} + \Pi^{(\leq n+2)}\{Z_n,W_n\} = 0\,,\\
&\Pi^{(n+3)}\{Z_n,W_n\} + \{H^{(n+3)},K^{(2)}\} + \{H^{(2)},K^{(n+3)}\} = 0\,,\dots
\end{aligned} \tag{C.0.12}$$

By the inductive hypothesis Z_n, W_n belong to $\text{Ker}(H^{(2)})\cap\text{Ker}(K^{(2)})$ so $\{Z_n,K^{(2)}\} = \{H^{(2)},W_n\} = 0$ and by (C.0.12) $\Pi^{(\leq n+2)}\{Z_n,W_n\} = 0$.
Again by the inductive hypothesis (and the Jacobi identity) we know that $\{Z_n,W_n\} \in \text{Ker}(H^{(2)}) \cap \text{Ker}(K^{(2)})$. Then, by (C.0.12), we have $\Pi^{(n+3)}\{Z_n,W_n\} = 0$, because

$$\Pi_{\text{Ker}(H^{(2)})}\{H^{(2)},K^{(n+3)}\} = 0, \quad \Pi_{\text{Ker}(K^{(2)})}\{H^{(n+3)},K^{(2)}\} = 0. \tag{C.0.13}$$

Hence we conclude that $\Pi^{(\leq n+3)}\{Z_n, W_n\} = 0$. By this fact, the third equation in (C.0.12) becomes

$$\{H^{(n+3)}, K^{(2)}\} + \{H^{(2)}, K^{(n+3)}\} = 0 \tag{C.0.14}$$

and, by the claim above, we have

$$\Pi_{\mathrm{Rg}(H^{(2)})} H^{(n+3)} = \Pi_{\mathrm{Rg}(H^{(2)})} \Pi_{\mathrm{Rg}(K^{(2)})} H^{(n+3)}$$

and

$$\Pi_{\mathrm{Rg}(K^{(2)})} K^{(n+3)} = \Pi_{\mathrm{Rg}(H^{(2)})} \Pi_{\mathrm{Rg}(K^{(2)})} K^{(n+3)}.$$

In order to obtain the normal form at step $n+1$ of (C.0.5) we apply a change of variables $\Phi^{(n)}$ generated by the Hamiltonian $F^{(n)}$ and we define $\Phi_{n+1} := \Phi^{(n)}\Phi_n$. The function $F^{(n)}$ is determined by solving the homological equation for H

$$\{H^{(2)}, F^{(n)}\} = -\Pi_{\mathrm{Rg}(H^{(2)})} H^{(n+3)} \overset{(C.0.13)}{=} -\Pi_{\mathrm{Rg}(K^{(2)})} \Pi_{\mathrm{Rg}(H^{(2)})} H^{(n+3)}\,. \tag{C.0.15}$$

We now prove that $F^{(n)}$ solves also the homological equation for K. Indeed by (C.0.15)

$$\{K^{(2)}, F^{(n)}\} = -(\mathrm{ad} H^{(2)})^{-1} \{K^{(2)}, \Pi_{\mathrm{Rg}(K^{(2)})} \Pi_{\mathrm{Rg}(H^{(2)})} H^{(n+3)}\}$$

and by (C.0.12), (C.0.11) we get

$$\{K^{(2)}, \Pi_{\mathrm{Rg}(K^{(2)})} \Pi_{\mathrm{Rg}(H^{(2)})} H^{(n+3)}\} = \{H^{(2)}, \Pi_{\mathrm{Rg}(K^{(2)})} \Pi_{\mathrm{Rg}(H^{(2)})} K^{(n+3)}\}. \tag{C.0.16}$$

□

Bibliography

[1] Alazard T., Baldi P., *Gravity capillary standing water waves*, Arch. Ration. Mech. Anal. 217, 741-830 (2015).

[2] Arnold V.I., *Proof of Kolmogorov's Theorem on the Preservation of Quasi-Periodic Motions under Small Perturbation of the Hamiltonian*, Uspekhi Mat. N. USSR, Vol. 18, Ser 5 (113), pp. 13-40 (1963).

[3] Arnold V.I., *Small denominators and problems of stability of motion in classical celestial mechanics*, Russian Math. Surveys 18, 85-193 (1963).

[4] Baldi P., *Periodic solutions of forced Kirchoff equations*, Ann. Scuola Norm. Sup. Pisa, Cl. Sci. (5) 8, 117-141 (2009).

[5] Baldi P., *Periodic solutions of fully nonlinear autonomous equations of Benjamin-Ono type*, Ann. Inst. H. Poincaré (C) Anal. Non Linéaire 30, 33-77 (2013).

[6] Baldi P., Berti M., Haus E., Montalto R., *Time quasi-periodic gravity water waves in finite depth*, preprint, arXiv:1708.01517 (2017).

[7] Baldi P., Berti M., Montalto R., *KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation*, Math Annalen 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 (2015).

[9] Baldi P., Berti M., Montalto R., *KAM for autonomous quasi-linear perturbations of mKdV*, Bollettino dell'Unione Matematica Italiana, Volume 9, Issue 2, 143-188 (2016).

[10] Bambusi D., Berti M., Magistrelli E., *Degenerate KAM theory for partial differential equations*, Journal of Differential Equations, Volume 250, Issue 8, Pages 3379-3397 (2011).

[11] Berti M., *Nonlinear oscillations of Hamiltonian PDEs*, Progress in Nonlin. Diff. Eq. Appl. 74, Birkhäuser Boston (2008).

[12] Berti M., *KAM for PDEs*, Bollettino dell'Unione Matematica Italiana, Volume 9, Issue 2, pp 115-142 (2016).

[13] Berti M., Biasco L., *Branching of Cantor Manifolds of Elliptic Tori and Applications to PDEs*, Communications in Mathematical Physics, Volume 305, Issue 3, pp 741-796 (2011).

[14] Berti M., Biasco L., Procesi M., *KAM theory for the Hamiltonian DNLW*, Ann. Sci. Ec. Norm. Sup. (4), VOL. 46, fascicule 2, 301-373 (2013).

[15] Berti M., Biasco L., Procesi M., *KAM theory for the reversible derivative wave equation*, Arch. Rational Mech. Anal., 212 , 905-955 (2014).

[16] Berti M., Bolle P., *Cantor families of periodic solutions for completely resonant nonlinear wave equations*, Duke Math. J. 134, 359-419 (2006).

[17] Berti M., Bolle P., *Cantor families of periodic solutions for wave equations via a variational principle*, Advances in Mathematics, 217, 1671-1727 (2008).

[18] Berti M., Bolle P., *Cantor families of periodic solutions of wave equations with C^k nonlinearities*, NoDEA 15, 247-276 (2008).

[19] Berti M., Bolle P., *Sobolev quasi periodic solutions of multidimensional wave equations with a multiplicative potential*, Nonlinearity 25, 2579-2613 (2012).

[20] Berti M., Bolle P., *Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential*, Eur. Jour. Math. 15, 229-286 (2013).

[21] Berti M., Bolle P., *A Nash-Moser approach to KAM theory*, Fields Institute Communications, volume 75 "Hamiltonian PDEs and Applications", 255-284 (2014).

[22] Berti M., Bolle P., *Quasi-periodic solutions for autonomous NLW on $\mathbb{T}^d$ with a multiplicative potential*, in preparation.

[23] Berti M., Bolle P., Procesi M., *An abstract Nash-Moser theorem with parameters and applications to PDEs*, Annales de l'Institute H. Poincaré, Analyse nonlinéaire, 27, 377 - 399 (2010).

[24] Berti M., Corsi L., Procesi M., *An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous spaces* , M. Commun. Math. Phys. 334, no. 3, 1413-1454 (2015).

[25] Berti M., Delort J. M., *Almost global existence of solutions for capillarity-gravity water waves equations with periodic spatial boundary conditions*, preprint, arXiv:1702.04674 (2017).

[26] Berti M., Kappeler T., Montalto R., *Large KAM tori for perturbations of the dNLS equation*, preprint, arXiv:1603.09252 (2016).

[27] Berti M., Montalto R., *Quasi-periodic standing wave solutions of gravity-capillary water waves*, preprint arXiv:1602.02411, to appear in Memoirs of the AMS, MEMO 891 (2016).

[28] Berti M., Montalto R., *Quasi-periodic water waves, Journal of Fixed Point Theory and Applications*, Volume 19, Issue 1, pp 129-156 (2017).

[29] Berti M., Procesi M., *Nonlinear wave and Schrödinger equations on compact Lie groups and homogeneous spaces*, Duke Math. J. Volume 159, Number 3, 479-538 (2011).

[30] Bourgain J., *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices, 11 (1994).

[31] Bourgain J., *Gibbs measures and quasi-periodic solutions for nonlinear Hamiltonian partial different equations*, 23-43, Gelfand Math. Sem., Birkhäuser Boston, Boston, MA (1996).

[32] Bourgain J., *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Annals of Math. 148, 363-439 (1998).

[33] Bourgain J., *Periodic solutions of nonlinear wave equations*, Harmonic analysis and partial differential equations, Chicago Lectures in Math., Univ. Chicago Press, 69-97 (1999).

[34] Bourgain J., *Green's function estimates for lattice Schrödinger operators and applications*, Annals of Mathematics Studies 158, Princeton University Press, Princeton (2005).

[35] Bourgain J., *On invariant tori of full dimension for the 1d periodic NLS*, J. Funct. Anal. 229 (2005).

[36] Chierchia L., Perfetti P., *Second order Hamiltonian equations on* $\mathbb{T}^\infty$ *and almost-periodic solutions*, Journal of Differential Equations, n. 21, pp 1, 172-201 (1995).

[37] Chierchia L., You J., *KAM Tori for 1D Nonlinear Wave Equations with Periodic Boundary Conditions*, Commun. Math. Phys. 211, 497-525 (2000).

[38] Constantin A., Lannes D., *The Hydrodynamical Relevance of the Camassa-Holm and Degasperis-Procesi Equations*, Arch. Rational Mech. Anal. 192 (2009).

[39] Corsi L., Feola R., Procesi M., *Finite dimensional invariant KAM tori for tame vector fields*, preprint, arXiv:1611.01641 (2017).

[40] Craig W., *Birkhoff normal forms for water waves. Mathematical problems in the theory of water waves*, (Luminy, 1995), 57-74, Contemp. Math., 200, Amer. Math. Soc., Providence, RI (1996).

[41] Craig W., *Problèmes de petits diviseurs dans les équations aux dérivées partielles*, Panoramas et Synthèses 9, Société Mathématique de France, Paris (2000).

[42] Craig. W., *Transformation theory of Hamiltonian PDEs and the problem of water waves*, proceedings of the Advanced Study Institute on Hamiltonian Dynamical Systems and Applications, NATO Science for Peace and Security Series B: Springer - Verlag, 67-83 (2008).

[43] Craig W., Nicholls D., *Travelling two and three dimensional capillary gravity water waves*, SIAM J. Math. Anal. 32, 323-359 (2000).

[44] Craig W., Sulem C., *Numerical simulation of gravity waves*, J. Comput. Phys. 108 no. 1, 73-83 (1993).

[45] Craig W., Wayne C., *Newton's method and periodic solutions of nonlinear wave equation*, Comm. Pure Appl. Math. 46, 1409-1498 (1993).

[46] Craig W., Wayne C., *Periodic solutions of nonlinear Schrödinger equations and Nash Moser method*, in: J. Semanis (Ed.), Hamiltonian Mechanics, Torun, 1993, NATO Adv. Sci. Inst. Ser. B Phys., vol 331, Plenum, 103 - 122 (1994).

[47] Craig W., Worfolk P., *An integrable normal form for water waves in infinite depth*, Phys. D 84, no. 3-4, 513-531 (1995).

[48] Degasperis A., Holm D. D., Hone A. N. W., *A new integrable Equation with Peakon Solutions*, Theoretical and Mathematical Physics, Volume 133, Issue 2, pp 1463-1474 (2002).

[49] Degasperis A., Procesi M., *Asymptotic Integrability*, in Symmetry and Perturbation Theory (A. Degasperis and G. Gaeta, eds.), World Scientific Publishing, 23-37 (1999).

[50] Dullin H., Gottwald G., Holm D.D., *An Integrable Shallow Water Equation with Linear and Nonlinear Dispersion*, Physical Review Letters 87 (19), 194501 (2001).

[51] Eliasson L.H., *Perturbations of stable invariant tori for Hamiltonian systems*, Annali della scuola normale superiore di Pisa, Cl. Sci. (4) 15, no 1, 115-147 (1988).

[52] Eliasson L.H., Grebert B., Kuksin S., *KAM for the nonlinear Beam equation*, Geometric and Functional Analysis, Volume 26, Issue 6, pp 1588-1715 (2016).

[53] Eliasson, L.H., Kuksin S.B., *On reducibility of Schrödinger equations with quasiperiodic in time potentials*, Comm. Math. Phys. 286, 125-135 (2009).

[54] Eliasson L.H., Kuksin S., *KAM for non-linear Schrödinger equation*, Annals of Math., 172, 371-435 (2010).

[55] Feola R., *KAM for quasi-linear forced hamiltonian NLS*, preprint, arXiv:1602.01341 (2017).

[56] Feola R., Procesi M., *Quasi-periodic solutions for fully nonlinear forced reversible Schrödinger equations*, Journal of Differential Equations, Volume 259, Issue 7, 3389-3447, (2015).

[57] Geng J., You J., *A KAM theorem for the one dimensional Schrödinger equation with periodic boundary conditions*, J. Diff. Equ. 209, 1 - 56 (2005).

[58] Geng J., You J., *A KAM Theorem for Hamiltonian Partial Differential Equations in Higher Dimensional Spaces*, Communications in Mathematical Physics, Volume 262, Issue 2, pp 343-372 (2006).

[59] Geng J., Xu X., You J., *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math. 226, 5361-5402 (2011).

[60] Gentile G., Mastropietro V., Procesi M., *Periodic solutions for completely resonant nonlinear wave equations with Dirichlet boundary conditions*, Comm. Math. Phys. 256 no. 2, 437-490 (2005).

[61] Giuliani F., *Quasi-periodic solutions for quasi-linear generalized KdV equations*, Journal of Differential Equations, Volume 262, Issue 10, Pages 5052-5132 (2017).

[62] Hörmander L., *The boundary problems of physical geodesy*, Arch. Ration. Mech. Anal. 62, no. 1, 1-52 (1976).

[63] Hörmander L., *The analysis of linear partial differential operators III. Pseudo-differential operartors*, Springer-Verlag, Berlin (1990).

[64] Huang G., Kuksin K., *KdV equation under periodic boundary conditions and its perturbations*, Nonlinearity, Volume 27 (9) (2014).

[65] Iooss G., Plotnikov P.I., *Existence of multimodal standing gravity waves*, J.Math.Fluid Mech.7 , 349-364 (2005).

[66] Iooss G., Plotnikov P.I., *Multimodal standing gravity waves: a completely resonant system*, J. Math. Fluid Mech. 7, 110-126 (2005).

[67] Iooss G., Plotnikov P.I., *Small divisor problem in the theory of three-dimensional water gravity waves*, Mem. Amer. Math. Soc. 200 no. 940 (2009).

[68] Iooss G., Plotnikov P.I., *Asymmetrical three-dimensional travelling gravity waves*, Arch. Ration. Mech. Anal. 200 no. 3, 789-880 (2011).

[69] Iooss G., Plotnikov P.I., Toland J.F., *Standing waves on an infinitely deep perfect fluid under gravity*, Arch. Rational Mech. Anal. 177 no. 3, 367-478 (2005).

[70] Kappeler T., Pöschel J., *KAM and KdV*, Springer (2003).

[71] Kolmogorov A. N., *On the persistence of conditionally periodic motions under a small perturbation of the Hamiltonian function*, Dokl. Akad. Nauk. SSSR 98, 527-530 (1954).

[72] Kuksin S., *Hamiltonian perturbations of infinite-dimensional linear systems with imaginary spectrum*, Funktsional. Anal. i Prilozhen. 21 no. 3, 22-37, 95 (1987).

[73] Kuksin S., *A KAM theorem for equations of the Korteweg-de Vries type*, Rev. Math. Phys. 10, no. 3, 1-64 (1998).

[74] Kuksin S., *Analysis of Hamiltonian PDEs*, Oxford Lecture Series in Mathematics and its Applications, 19. Oxford University Press (2000).

[75] Kuksin S,. Pöschel J., *Invariant Cantor Manifolds of Quasi-Periodic Oscillations for a Nonlinear Schrödinger Equation*, Ann. of Math. 143, 149-179 (1996).

[76] Liu J., Yuan X., *A KAM theorem for Hamiltonian Partial Differential Equations with Unbounded Perturbations*, Comm. Math. Phys., 307 (3), 629-673 (2011).

[77] Métivier G., *Para-differential Calculus and Applications to the Cauchy Problem for Nonlinear Systems*, Pubblicazioni Scuola Normale Pisa, 5 (2008).

[78] Montalto R., *A reducibility result for a class of linear wave equations with unbounded perturbations on* $\mathbb{T}^d$, arXiv:1702.06880 (2017).

[79] Montalto R., *Quasi-periodic solutions of forced Kirchoff equation*, Nonlinear Differ. Equ. Appl. 24: 9 (2017).

[80] Moser J., *A new technique for the construction of solutions of nonlinear differential equations*, Proc. Nat. Acad. Sci. 47, 1824-1831 (1961).

[81] Moser J., *A rapidly convergent iteration method and non-linear partial differential equations (I)*, Ann. Sc. Norm. Sup. Pisa, Cl. Sci. (3) 20 no. 2, 265-315 (1966).

[82] Moser J., *A rapidly convergent iteration method and non-linear partial differential equations (II)*, Ann. Sc. Norm. Sup. Pisa, Cl. Sci. (3) 20 no. 2, 499-535 (1966).

[83] Moser J., *Convergent series expansion for quasi-periodic motions*, Math. Ann. 169, 136-176 (1967).

[84] Plotnikov P.I., Toland, J.F., *Nash-Moser theory for standing water waves*, Arch. Ration. Mech. Anal. 159, 1-83 (2001).

[85] Pöschel J., *On elliptic lower dimensional tori in Hamiltonian systems*, Math Z. 202, 559-608 (1989).

[86] Pöschel J., *A KAM-Theorem for some nonlinear partial differential equations*, Ann. Sc. Norm. Sup. Pisa, Cl. Sci. (4) 23, 119-148 (1996).

[87] Pöschel J., *On the construction of almost periodic solutions for a nonlinear Schrödinger equation*, Ergodic Theory Dynam. Systems 22, no. 5, 1537-1549 (2002).

[88] Pöschel J., *Quasi-periodic solutions for a nonlinear wave equation*, Comment. Math. Helv. 71, no. 2, 269-296 (1996).

[89] Procesi C., Procesi M., *A normal form for the Schrödinger equation with analytic non-linearities*, Communications in Mathematical Physics 312, no. 2, 501-557 (2012).

[90] Procesi C., Procesi M., *A KAM algorithm for the non-linear Schrödinger equation*, Advances in Math. 272 , 399-470 (2015).

[91] Procesi M., Xu X., *Quasi-Töplitz Functions in KAM Theorem*, SIAM J.Math. Anal. 45 no. 4, 2148-2181 (2013).

[92] Saranen J., Vainikko G., *Periodic Integral and Pseudodifferential Equations with Numerical Approximation*, Springer Monographs in Mathematics (2002).

[93] Taylor M. E., *Pseudodifferential Operators and Nonlinear PDEs*, Progress in Mathematics, Birkhäuser (1991).

[94] Wang W. M., *Supercritical nonlinear Schrödinger equations* I*: quasi-periodic solutions*, Duke Math. J. 165, no. 6, 1129-1192 (2016).

[95] Wayne E., *Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory*, Comm. Math. Phys. 127, 479-528 (1990).

[96] Zakharov V. E., *Stability of periodic waves of finite amplitude on the surface of a deep fluid*, Journal of Applied Mechanics and Technical Physics 9 no. 2, 190-194 (1968).

[97] Zehnder E., *Generalized implicit function theorems with applications to some small divisors problems* $I - II$, Comm. Pure Appl. Math., 28, 91-140, (1975); 29, 49-113, (1976).

[98] Zhang J., Gao M., Yuan X., *KAM tori for reversible partial differential equations*, Nonlinearity 24 , 1189-1228 (2011).